Anti-unification and the Theory of Semirings*

David M. Cerna †

The Czech Academy of Science,
Institute of Computer Science (CAS ICS)
Pod Vodárenskou věží 271/2 182 07, Prague, Czechia

Johannes Kepler Univerisity (JKU),
Research Institute for Symbolic Computation (RISC)
Altenbergerstrasse 69, 4040 Linz, Austria

September 3, 2020

Abstract

It was recently shown that anti-unification over an equational theory consisting of only unit equations (more than one) is nullary. Such pure theories are artificial and are of little effect on practical aspects of anti-unification. In this work, we extend these nullarity results to the theory of semirings, a heavily studied theory with many practical applications. Furthermore, our argument holds over semirings with commutative multiplication and/or idempotent addition. We also cover a few open questions discussed in previous work.

1 Introduction

Unification is a process by which two symbolic expressions may be identified through variable replacement. *Anti-unification* (generalization), on the other hand, is a process that derives from a set of symbolic expressions a new symbolic expression possessing certain commonalities shared between its members. As the naming suggests, Anti-unification is the duel problem of unification.

Furthermore, both have been considered in a variety of settings, in particular, over first-order equational theories, what is commonly referred to as *E-unification* (*E-generalization*). Modulo a first-order equational theory, the unification/anti-unification problem is said to be nullary (type zero), if there exist problem instances for which a minimal complete set of solutions does not exist. Concerning unification, the first known, though artificial, first-order equational theory with this property was found by Fages and Huet [17]. Shortly after their publication, Baader and Schmidt-Schauss independently showed that unification over the theory of idempotent semigroups is nullary [4, 30].

Though, anti-unification was first introduced in the 1970's by Plotkin and Reynolds [26, 27] and has been studied in various forms and over numerous equational theories [1, 12, 13, 15, 25, 22, 9, 20], not much is known about nullary anti-unification. The first example, to the best of our knowledge, was presented in [14] where an artificial equational theory with this property is provided (the theory consisted of unit equations only). Extension by *permutative equations* (A concept introduced in [31]) was discussed but not formally proven.

One of the many applications of anti-unification is inductive theorem proving based on tree grammars [3]. In this work, the Herbrand sequents of a sequence of cut-free proofs is transformed into a minimal tree grammar (in terms of productions) using the methods described in [16]. The construction of these minimal tree grammars uses syntactic anti-unification to minimize the number of productions.

Other applications of anti-unification are recursion scheme detection in functional programs [8], inductive synthesis of recursive functions [29], learning fixes from software code repositories [28, 7], preventing bugs and misconfiguration [23], as well as uses within the fields of natural language processing

*Supported by the Linz Institute of Technology (LIT) Math_{LP} project (LIT- 2019-7-YOU-213) funded by the state of upper Austria.

†E-mail: dcerna@{cs.cas.cz, risc.jku.at}

and linguistics [18, 2]. Thus, gaining a better understanding of its behavior over important first-order equational theories is essential. The prominence of algebraic structures, whose equational theory includes unit equations, in programming language theory, suggests that understanding of anti-unification in the presence of such axioms can benefit future progress in this area, especially over more complex equational theories such as semirings. Concerning inductive theorem proving methods, i.e. [3], the type of anti-unification used directly influences the theory over which invariants can be discovered. Concerning Constraint Logic Programming (CLP) [11], our result implies that generalizing the constraints occurring within a semiring-based CLP may not result in a least general generalization (lgg), i.e. there may not exist a "best" constraint capturing the common structure of the compared constraints. One may desire to generalize the constraints in CLPs for reasons outlined above as well as software clone analysis and pattern detection.

Also, note that unification over monoidal theories [24, 5] is often solved within a corresponding semiring, i.e. one can construct equations over the semiring which correspond to instances of the unification problem over the monoidal theory. The main result of our paper provides insight into the structures used to solve unification over these theories. Furthermore, generalization was also considered in [5] using a similar technique. Note that the term signature of the equational theories considered in [5] contained a single operator while in this work we consider theories with two operators.

Overall, while it is not clear how these results can be put directly into practice, it is nonetheless important to have insight into the various properties of such an important equational theory.

In this paper, we extend the results of [14] as follows:

(1) Anti-unification over an equational theory with two idempotent function symbols, each of which has an associated unit equation, is nullary. A corollary of this result is that the general algorithm presented in [14] is incomplete, i.e. There exist generalizations that are incomparable to the generalizations found by the algorithm. In other words, the generalizations found by the algorithm do not form a complete set.

(2) Anti-unification over the theory of semirings [19, 11, 21] is nullary, and thus we provide the first natural example of a first-order theory over which anti-unification is nullary.

Furthermore, a trivial extension of our argument shows nullarity of anti-unifica-tion over semirings with commutative multiplication and/or idempotent addition. Though, as we discuss below, our argument does not extend to semirings with idempotent multiplication. This is a side effect of *Nilpotency.*

2 Preliminaries

We assume familiarity with the basic notions of unification theory, see [6]. Variables are denoted by x, y, z, terms by s, t, r, generalizations by g, substitutions by σ, θ, μ, and natural numbers by n, m, k, q, l. Terms are constructed inductively from a signature Σ containing function symbols of fixed arity and a countable set of variables $\mathcal{V}$. We denote the set of all terms inductively constructable over Σ by $\mathcal{T}(\Sigma, \mathcal{V})$. Constants are function symbols of arity 0. We denote the set of variables appearing in a term t by $var(t)$. The *depth of a term* t is defined inductively as follow:

- If t is a constant or a variable then $depth(t) = 1$.
- If $t = f(t_1, \ldots, t_n)$ and $n \geq 1$, then $depth(t) = (\max_{1 \leq i \leq n}\{depth(t_i)\}) + 1$.

The *head of a term* t is defined as follows:

- If t is a variable then $head(t) = t$.
- If $t = f(t_1, \ldots, t_n)$ for $f \in \Sigma$ and $n \geq 0$, then $head(t) = f$.

For a substitution σ, $dom(\sigma)$ denotes the set of variables which σ does not map to themselves. Let $S \subseteq dom(\sigma)$, by $\sigma|_S$ we denote the substitution mapping every variable to itself but those occurring in S which are mapped to the same term as σ, i.e for $x \in S$, $x\sigma = x\sigma|_S$. Let $\mathcal{E}$ be a set of equations and $=_\mathcal{E}$ be the equality of terms induced by $\mathcal{E}$.

A term $g \in \mathcal{T}(\Sigma, \mathcal{V})$ is referred to as an $\mathcal{E}$-generalization of terms $s, t \in \mathcal{T}(\Sigma, \mathcal{V})$ if there exist substitutions σ_1 and σ_2 such that $g\sigma_1 =_\mathcal{E} s$ and $g\sigma_2 =_\mathcal{E} t$. We refer to such g as a solution to the *anti-unification problem (AUP)* $s \triangleq t$.

The set of all $\mathcal{E}$-generalizations of $s, t \in \mathcal{T}(\Sigma, \mathcal{V})$ is denoted by $csg_{\mathcal{E}}(s,t)$. A quasi-ordering $\leq_{\mathcal{E}}$ may be defined on $\mathcal{T}(\Sigma, \mathcal{V})$ as follows: $g \leq_{\mathcal{E}} g'$ if and only if there exists σ such that $g\sigma =_{\mathcal{E}} g'$. By $g <_{\mathcal{E}} g'$, we imply that $g' \not\leq_{\mathcal{E}} g$.

A minimal complete set of $\mathcal{E}$-generalizations of s and t, denoted $mcsg_{\mathcal{E}}(s,t)$ has the following properties:

1. $mcsg_{\mathcal{E}}(s,t) \subseteq csg_{\mathcal{E}}(s,t)$.
2. For all $g' \in csg_{\mathcal{E}}(s,t)$ there exists a $g \in mcsg_{\mathcal{E}}(s,t)$ such that $g' \leq_{\mathcal{E}} g$.
3. If $g, g' \in mcsg_{\mathcal{E}}(s,t)$ and $g \leq_{\mathcal{E}} g'$, then $g = g'$.

A subset of $csg_E(s,t)$ is called *complete* if the first two conditions hold. Equational theories are classified by the size and existence of the $mcsg_{\mathcal{E}}(s,t)$:

- *Unitary*: The $mcsg_{\mathcal{E}}(s,t)$ exists for all terms s, t and is always singleton.
- *Finitary*: The $mcsg_{\mathcal{E}}(s,t)$ exists, is finite for all terms s, t, and there exist terms s', t' for which $1 < |mcsg_{\mathcal{E}}(s',t')| < \infty$.
- *Infinitary*: The $mcsg_{\mathcal{E}}(s,t)$ exists for all terms s, t and there exists at least one pair of terms for which it is infinite.
- *Nullary*: There exist terms s, t such that $mcsg_{\mathcal{E}}(s,t)$ does not exist.

For example, syntactic anti-unification (AU) is unitary [26, 27], associative AU is finitary [1, 10], idempotent AU is infinitary [13], and AU with multiple unit equations is nullary [14].

The main result of this work concerns anti-unification over semirings, $\mathsf{SR} =$

$$\left\{\begin{array}{cl}
x + \mathbf{0} = x \ , \ \mathbf{0} + x = x & \text{Unital } + \ (\mathsf{U}^{+}_{\mathbf{0}}) \\
x + (y + z) = (x + y) + z & \text{Associativity } + \ (\mathsf{A}^{+}) \\
x + y = y + x & \text{Commutativity } + \ (\mathsf{C}^{+}) \\
x \cdot \mathbf{1} = x \ , \ \mathbf{1} \cdot x = x & \text{Unital } \cdot \ (\mathsf{U}^{\cdot}_{\mathbf{1}}) \\
x \cdot (y \cdot z) = (x \cdot y) \cdot z & \text{Associativity } \cdot \ \ (\mathsf{A}^{\cdot}) \\
x + y = y + x & \text{Commutativity } \cdot \ \ (\mathsf{C}^{\cdot}) \\
x \cdot \mathbf{0} = \mathbf{0} \ , \ \mathbf{0} \cdot x = \mathbf{0} & \text{Nilpotency } \cdot \ (\mathsf{N}^{\cdot}_{0}) \\
x \cdot (y + z) = (x \cdot y) + (x \cdot z) & \text{Distributivity } \cdot \text{ over } + \text{ left } \ (\mathsf{D}^{(\cdot,+)}_{l}) \\
(y + z) \cdot x = (y \cdot x) + (z \cdot x) & \text{Distributivity } \cdot \text{ over } + \text{ right } \ (\mathsf{D}^{(\cdot,+)}_{r})
\end{array}\right\}.$$

Though we include commutativity of multiplication in the above definition, in Section 3.2, we assume non-commutative multiplication to align the results with the conventional definition of semirings. Note that adding commutative actually simplifies the argumentation presented in Section 3.2. We say that a term t is in *semiring normal form* (*SR-normal form*) if t does not contain any subterms of the form $\mathbf{0} + s$, $s + \mathbf{0}$, $\mathbf{0} \cdot s$, $s \cdot \mathbf{0}$, $s \cdot \mathbf{1}$, $\mathbf{1} \cdot s$, $s \cdot (r + l)$, or $(r + l) \cdot s$. Additionally, we will need to consider the equational theories $\mathsf{SR}^{\mathsf{ND-}}$ (without nilpotency and distributivity), $\mathsf{SR}^{\mathsf{D-}}$ (without distributivity), and $\mathsf{SR}^{\mathsf{N-}}$ (without nilpotency). Normal forms for these theories are defined analogously to SR-normal form.

2.1 Relationship to Previous Work

In [14], it was shown that anti-unification over the equational theory $\mathsf{U}_2 = \mathsf{U}^{+}_{\mathbf{0}} \cup \mathsf{U}^{\cdot}_{\mathbf{1}}$ is nullary. As discussed in [14], we conjectured that the argument presented there holds when U_2 is extended by any combination of the following equations: A^{+}, C^{+}, $\mathsf{A}^{\cdot}$, and $\mathsf{C}^{\cdot}$. The results presented in Section 3.2 together with [14] prove this conjecture.

However, the argument fails for the theory $\mathsf{UI}_2 = \mathsf{U}_2 \cup \mathsf{I}^{+} \cup \mathsf{I}^{\cdot}$ where

$$\begin{array}{rl}
x + x = x & \text{Idempontency } + \ (\mathsf{I}^{+}) \\
x \cdot x = x & \text{Idempontency } \cdot \ \ (\mathsf{I}^{\cdot})
\end{array}$$

The problem being that the nullarity argument presented in [14] relies on the seed term $x + (x \cdot y)$ (or $x \cdot (x + y)$) which collapses to a variable over the theory UI_2, i.e. $x + (x \cdot y) \leq_{\mathsf{UI}_2} x$. Thus, the anti-unification type of UI_2 (as well as ACUI_2) was left as an open question. In Section 3.1, we show that a different more complex seed term allows the nullarity argument to go through. This seed term also proves that the general algorithm presented in [14] is not complete (the generalizations found by the algorithm do not form a complete set), another open question of the previous work.

The Theory $\mathsf{U_2N} = \mathsf{U_2} \cup \mathsf{N_0^{\cdot}}$, i.e. extending $\mathsf{U_2}$ by Nilpotency $\cdot$, is equally problematic. Note that while $x + (x \cdot y)$ collapses to a variable over the theory $\mathsf{U_2N}$, i.e. $x + (x \cdot y) \leq_{\mathsf{U_2N}} x$, the term $x \cdot (x + y)$ does not. Instead $x \cdot (x + y) =_{\mathsf{U_2N}} x \cdot x$. As we shall show in Section 3.2 a related seed term $\prod x$ is enough to prove nullarity of anti-unification over semirings. The approach taken in [14] is not enough to show that this is indeed the case. A corollary of the argument presented in Section 3.2 is that the simpler theory $\mathsf{UN} = \mathsf{U_1^{\cdot}} \cup \mathsf{N_0^{\cdot}}$ consisting of Unital $\cdot$ and Nilpotency $\cdot$ is also nullary.

As a final comment concerning the relationship between the previous results and this work, let us consider the theory $\mathsf{UI_2N} = \mathsf{UI_2} \cup \mathsf{N_0^{\cdot}}$. While choosing the right seed term allowed us to prove nullarity of $\mathsf{UI_2}$ and strengthening our argument allows us to prove nullarity of $\mathsf{U_2N}$, the theory $\mathsf{UI_2N}$ breaks our method, specifically Theorem 1 (Theorem 7 of [14]) which relies on an increase in the number of variable occurrences. Thus, finding the anti-unification type of $\mathsf{UI_2N}$ will require a different approach. We conjecture that it is of type infinitary and an approach similar to [13] is required. We conjecture similar for fully idempotent semirings.

As a final remark, we would like to point out that adding additional equations to the theory does not imply its anti-unification type will remain the same or increase, in some cases it may decrease [5]. Note that the generalizations used to prove nullarity of anti-unification over semirings are significantly simpler than what was needed when proving nullarity of the theory $\mathsf{U_2}$. This simplification most likely explains the failure of our argument for $\mathsf{UI_2N}$, i.e. the generalizations are even simpler in this case and perhaps may explain why the semiring extension of this theory is favored in practice.

3 Results

In this section we present the main contributions discussed in Section 2.1. Before we continue, an interesting observation to point out is that these results rely on variables having multiple occurrences within the generalizations. If we limit ourselves to *linear* generalizations, each variable may only occur once, our argument no longer works. It was shown in [14] that linear anti-unification over $\mathsf{U_2}$ is type finitary. Extending this to semirings is elementary. What remains open is whether one may extend this linear anti-unification result to semirings with idempotent $+$ or to $\mathsf{UI_2N}$. It is known that anti-unification over I^+ is type infinitary [13], but the linear case was never investigated.

3.1 Anti-unification over $\mathsf{UI_2}$ is Nullary

As mentioned in Section 2.1, the proof presented in [14] fails for the theory $\mathsf{UI_2}$ because the seed terms used in this earlier work collapses to a variable, i.e. $x + (x \cdot y) \leq_{\mathsf{UI_2}} z$ as well as $x \cdot (x + y) \leq_{\mathsf{UI_2}} z$. As it turned out, there exists seed terms which do not collapse, namely $((x + y) \cdot \mathbf{0}) + x$ and $((x \cdot y) + \mathbf{1}) \cdot x$. Note the following:

$$((x + y) \cdot \mathbf{0}) + x\{x \mapsto \mathbf{0}\ ,\ y \mapsto \mathbf{1}\} =_{\mathsf{UI_2}} ((\mathbf{0} + \mathbf{1}) \cdot \mathbf{0}) + \mathbf{0} =_{\mathsf{UI_2}} \mathbf{0}$$

$$((x + y) \cdot \mathbf{0}) + x\{x \mapsto \mathbf{1}\ ,\ y \mapsto \mathbf{0}\} =_{\mathsf{UI_2}} ((\mathbf{1} + \mathbf{0}) \cdot \mathbf{0}) + \mathbf{1} =_{\mathsf{UI_2}} \mathbf{1}$$

If one considers possible decompositions using the inference rules discussed in [14] these generalizations are clearly not reachable from the initial anti-unification problem $\mathbf{0} \triangleq \mathbf{1}$. The inference rules do not allow one to place a $\mathbf{0}$ in the scope of $\cdot$ or a $\mathbf{1}$ in the scope of a $+$. This implies that the general algorithm presented in Section 6 of [14] is incomplete. It is not entirely clear if a complete algorithm can be constructed in any other sense than what is presented in [12]. On a final note, our nullarity result can easily be extended to the theory where both $+$ and $\cdot$ are also associative-commutative (using the results presented in the following section), that is the theory $\mathsf{ACUI_2}$, thus proving that this extended theory is also nullary.

3.2 Anti-unification over Semirings is Nullary

We show that an $mcsg_{\mathsf{SR}}(\mathbf{0}, \mathbf{1})$ does not exist by first showing only a single variable is needed for generalization (Definition 2, Corollaries 1 & 2, Theorem 1) and then, using these restricted generalizations, showing that any complete set of generalizations will necessarily contain $\leq_{\mathsf{SR}}$-comparable generalizations (Lemma 3 and Theorem 2).

Definition 1 *Let g be a generalization in $\mathcal{E}$-normal form of $t \triangleq s$ and σ_1, σ_2 substitutions such that $g\sigma_1 =_{\mathcal{E}} t$, $g\sigma_2 =_{\mathcal{E}} s$, and for every $\{x \mapsto r\} \in \sigma_i$, for $i \in \{1, 2\}$, r is in $\mathcal{E}$-normal form. We refer to the triple (g, σ_1, σ_2) as a $(t, s, \mathcal{E})$-*generalization*.*

Observe that in many cases the substitutions associated with a $(t, s, \mathcal{E})$-generaliza-tions are not unique. For example, nilpotency guarantees infinitely many possible associated substitutions. An important part of the argument presented in this section is that while the substitutions associated with a $(\mathbf{0}, \mathbf{1}, \mathsf{SR})$-generalization need not be unique we can assume the substitutions take a particular form.

Definition 2 *Let (g, σ_1, σ_2) be a $(t, s, \mathcal{E})$-generalization and S be a set of terms in $\mathcal{E}$-normal form. We say that (g, σ_1, σ_2) is $(S, \mathcal{E})$-*reduced *if the following hold:*

1. *For every $x \in var(g)$, $x\sigma_1 \neq_{\mathcal{E}} x\sigma_2$.*
2. *For all $x, y \in var(g)$ either $x = y$, or for some $\theta \in \{\sigma_1, \sigma_2\}$, $x\theta \neq_{\mathcal{E}} y\theta$.*
3. *For all $x \in var(g)$ and $i \in \{1, 2\}$, $x\sigma_i \in S$.*

*If only (3) holds, we refer to (g, σ_1, σ_2) as $(S, \mathcal{E})$-*reducible.

The following lemmas provides an essential property of $(\mathbf{0}, \mathbf{1}, \mathsf{SR})$-generaliza-tions which will allow us to transform them into $(\{\mathbf{0}, \mathbf{1}\}, \mathsf{SR})$-reduced generalizations. We first prove these properties for the weakest considered theory and describe how to lift it to SR.

Lemma 1 *Let t be a term and σ a substitution such that for all $x \in var(t)$, $x\sigma$ is in $\mathsf{SR}^{\mathsf{ND}\text{-}}$-normal form . If $t\sigma =_{\mathsf{SR}^{\mathsf{ND}\text{-}}} t'$ for $t' \in \{\mathbf{0}, \mathbf{1}\}$, then for all $x \in var(t)$ $x\sigma \in \{\mathbf{0}, \mathbf{1}\}$.*

Proof We prove this lemma by induction over the term depth of t. If $depth(t) = 1$ then either $t \in \{\mathbf{0}, \mathbf{1}\}$ or t is a variable. If $t \in \{\mathbf{0}, \mathbf{1}\}$, then obviously $t\sigma \in \{\mathbf{0}, \mathbf{1}\}$. If t is a variable, then by our assumptions on σ, $t\sigma \in \{\mathbf{0}, \mathbf{1}\}$, thus completing the basecase.

As an induction hypothesis we assume that the statement holds for $depth(t) \leq m$ and show it holds for $depth(t) = m + 1$. Note that $head(t) \in \{+, \cdot\}$ and by associativity of these operators t can be written as a sum $(\sum)$ or product $(\prod)$:

a) If $head(t) = +$ then $t =_{\mathsf{SR}^{\mathsf{ND}\text{-}}} \sum_{i=1}^{k} t_i$. This implies that $t\sigma =_{\mathsf{SR}^{\mathsf{ND}\text{-}}} t'$ for $t' \in \{\mathbf{0}, \mathbf{1}\}$ iff for at most one $1 \leq i \leq k$, $t_i\sigma =_{\mathsf{SR}^{\mathsf{ND}\text{-}}} \mathbf{1}$ and for all $1 \leq j \leq k$ such that $i \neq j$, $t_j\sigma =_{\mathsf{SR}^{\mathsf{ND}\text{-}}} \mathbf{0}$. Furthermore, for all $1 \leq i \leq k$, $depth(t_i) < depth(t)$. From these observations it follows that the induction hypothesis is applicable to all t_i for $1 \leq i \leq k$.

b) If $head(t) = \cdot$ then $t =_{\mathsf{SR}^{\mathsf{ND}\text{-}}} \prod_{i=1}^{k} t_i$. This implies that $t\sigma =_{\mathsf{SR}^{\mathsf{ND}\text{-}}} t'$ for $t' \in \{\mathbf{0}, \mathbf{1}\}$ iff for at most one t_i, $t_i\sigma =_{\mathsf{SR}^{\mathsf{ND}\text{-}}} \mathbf{0}$ and for all $1 \leq j \leq k$ such that $i \neq j$, $t_j\sigma =_{\mathsf{SR}^{\mathsf{ND}\text{-}}} \mathbf{1}$. Furthermore, for all $1 \leq i \leq k$, $depth(t_i) < depth(t)$. From these observations it follows that the induction hypothesis is applicable to all t_i for $1 \leq i \leq k$ and thus we have completed the proof.

□

Using Lemma 1 we can now make the first steps towards the construction of $(\{\mathbf{0}, \mathbf{1}\}, \mathsf{SR})$-reduced generalizations.

Lemma 2 *Let (g, σ_1, σ_2) be a $(\mathbf{0}, \mathbf{1}, \mathsf{SR}^{\mathsf{ND}\text{-}})$-generalization. Then for all $x \in var(g)$ and $i \in \{1, 2\}$, $x\sigma_i \in \{\mathbf{0}, \mathbf{1}\}$.*

Proof We can prove this lemma by induction over the term depth of g. As a basecase consider $depth(g) = 1$, i.e. g is a variable. W.l.o.g, let us assume $g = x$. Now let $\sigma_1 = \{x \mapsto t_1\}$ and $\sigma_2 = \{x \mapsto t_2\}$. By the definition of a $(\mathbf{0}, \mathbf{1}, \mathsf{SR}^{\mathsf{ND}\text{-}})$-generalization, both t_1 and t_2 must be in $\mathsf{SR}^{\mathsf{ND}\text{-}}$-normal form. Thus, $g\sigma_1 =_{\mathsf{SR}^{\mathsf{ND}\text{-}}} \mathbf{0}$ iff $t_1 = \mathbf{0}$. Similar holds for $g\sigma_2$.

Let us consider for our induction hypothesis that the lemma holds for $depth(g) = n$ and show that it holds for $depth(g) = n + 1$. As in Lemma 1 we take advantage of the sum $(\sum)$ and product $(\prod)$ normal forms:

a) If $head(g) = +$ then $g =_{\mathsf{SR}^{\mathsf{ND}\text{-}}} \sum_{i=1}^{k} t_i$. This implies that $g\sigma_1 =_{\mathsf{SR}^{\mathsf{ND}\text{-}}} \mathbf{0}$ and $g\sigma_2 =_{\mathsf{SR}^{\mathsf{ND}\text{-}}} \mathbf{1}$ if and only if

   - $t_i\sigma_1 =_{\mathsf{SR}^{\mathsf{ND}\text{-}}} \mathbf{0}$ for all t_i
   - there exists a unique $1 \leq j \leq k$ such that $t_j\sigma_2 =_{\mathsf{SR}^{\mathsf{ND}\text{-}}} \mathbf{1}$ and for $1 \leq l \leq k$, such that $j \neq l$, $t_l\sigma_2 = \mathbf{0}$.

Note that $(t_j, \sigma_1, \sigma_2)$ is a $(\mathbf{0}, \mathbf{1}, \mathsf{SR}^{\mathsf{ND-}})$-generalization of $depth(t_j) \leq n$, and thus by the induction hypothesis all variables occurring in t_j are mapped to either $\mathbf{0}$ or $\mathbf{1}$. If $var(g) = var(t_j)$ then we are done. If $var(g) \setminus var(t_j) \neq \emptyset$ then further consideration is necessary.

Let $\sigma_1' = \sigma_1|_{var(t_j)}$ and $\sigma_2' = \sigma_2|_{var(t_j)}$. Note that $g\sigma_1' =_{\mathsf{SR}^{\mathsf{ND-}}} \mathbf{0} + s_1 =_{\mathsf{SR}^{\mathsf{ND-}}} s_1$ and $g\sigma_2' =_{\mathsf{SR}^{\mathsf{ND-}}} \mathbf{1} + s_2$ for some terms s_1 and s_2. Furthermore, $s_1\sigma_1 =_{\mathsf{SR}^{\mathsf{ND-}}} s_2\sigma_2 =_{\mathsf{SR}^{\mathsf{ND-}}} \mathbf{0}$. From Lemma 1 it follows that for all variables x occurring in s_1 and s_2 that $x\sigma_1, x\sigma_2 \in \{\mathbf{0}, \mathbf{1}\}$.

b) If $head(g) = \cdot$ then $g =_{\mathsf{SR}^{\mathsf{ND-}}} \prod_{i=1}^{k} t_i$. This implies that $g\sigma_1 =_{\mathsf{SR}^{\mathsf{ND-}}} \mathbf{0}$ and $g\sigma_2 =_{\mathsf{SR}^{\mathsf{ND-}}} \mathbf{1}$ if and only if

- there exists a unique $1 \leq j \leq k$ such that $t_j\sigma_1 =_{\mathsf{SR}^{\mathsf{ND-}}} \mathbf{0}$ and $1 \leq l \leq k$, such that $l \neq k$, $t_l\sigma_1 = \mathbf{1}$
- $t_i\sigma_2 =_{\mathsf{SR}^{\mathsf{ND-}}} \mathbf{1}$ for all t_i.

Note that $(t_j, \sigma_1, \sigma_2)$ is a $(\mathbf{0}, \mathbf{1}, \mathsf{SR}^{\mathsf{ND-}})$-generalization of $depth(t_j) \leq n$, and thus by the induction hypothesis all variables occurring in t_j are mapped to either $\mathbf{0}$ or $\mathbf{1}$. If $var(g) = var(t_j)$ then we are done. If $var(g) \setminus var(t_j) \neq \emptyset$ then further consideration is necessary.

Let $\sigma_1' = \sigma_1|_{var(t_j)}$ and $\sigma_2' = \sigma_2|_{var(t_j)}$. Because $\cdot$ is associative and not associative-commutative we have to consider three different cases concerning $g\sigma_1'$, namely, $g\sigma_1' =_{\mathsf{SR}^{\mathsf{ND-}}} \mathbf{0} \cdot s_1$, $g\sigma_1' =_{\mathsf{SR}^{\mathsf{ND-}}} s_1 \cdot \mathbf{0}$, and $g\sigma_1' =_{\mathsf{SR}^{\mathsf{ND-}}} s_1 \cdot \mathbf{0} \cdot s_1'$. These cases are handled in a similar fashion and thus we only handle one explicitly. W.l.o.g, let $g\sigma_1' =_{\mathsf{SR}^{\mathsf{ND-}}} \mathbf{0} \cdot s_1$ and $g\sigma_2' =_{\mathsf{SR}^{\mathsf{ND-}}} \mathbf{1} \cdot s_2 =_{\mathsf{SR}^{\mathsf{ND-}}} s_2$ for some terms s_1 and s_2. Furthermore, $s_1\sigma_1 =_{\mathsf{SR}^{\mathsf{ND-}}} s_2\sigma_2 =_{\mathsf{SR}^{\mathsf{ND-}}} \mathbf{1}$. From Lemma 1 it follows that for all variables x occurring in s_1 and s_2 that $x\sigma_1, x\sigma_2 \in \{\mathbf{0}, \mathbf{1}\}$.

□

In the following corollary we add the nilpotency equation and extend the argument of Lemma 2.

Corollary 1 *Let (g, σ_1, σ_2) be a $(\mathbf{0}, \mathbf{1}, \mathsf{SR}^{\mathsf{D-}})$-generalization and $S \subseteq var(g)$ be the set of $x \in var(g)$ for which there exists $i \in \{1, 2\}$, such that $x\sigma_i \notin \{\mathbf{0}, \mathbf{1}\}$. Then there exists a $(\mathbf{0}, \mathbf{1}, \mathsf{SR}^{\mathsf{D-}})$-generalization $(g, \sigma_1^*, \sigma_2^*)$ such that σ_1^*, and σ_2^* coincide everywhere with σ_1 and σ_2 except on S, where for $x \in S$ and $i \in \{1, 2\}$, $x\sigma_i^* \in \{\mathbf{0}, \mathbf{1}\}$.*

Proof This may be proven using a similar argument as Lemma 2. As one would expect the only significant change concerns part b when we consider a product. Note that $g\sigma_1' =_{\mathsf{SR}^{\mathsf{D-}}} \mathbf{0} \cdot s_1 =_{\mathsf{SR}^{\mathsf{D-}}} \mathbf{0}$ implying that σ_1 can map the variables of s_1 to anything. Thus, we construct a substitution σ_1^* which coincides with σ_1 everywhere except on $var(s_1)$. For $var(s_1)$ we define σ_1^* as follows: for $x \in var(s_1)$, $x\sigma_1^* = \mathbf{0}$. Similar changes would have to be made to lift Lemma 1 from $\mathsf{SR}^{\mathsf{ND-}}$ to $\mathsf{SR}^{\mathsf{D-}}$. □

In the next corollary, we extend this to SR by pointing out that distributivity does not add any additional annihilating power.

Corollary 2 *Let (g, σ_1, σ_2) be a $(\mathbf{0}, \mathbf{1}, \mathsf{SR})$-generalization and $S \subseteq var(g)$ be the set of $x \in var(g)$ for which there exists $i \in \{1, 2\}$, such that $x\sigma_i \notin \{\mathbf{0}, \mathbf{1}\}$. Then there exists a $(\mathbf{0}, \mathbf{1}, \mathsf{SR})$-generalization $(g, \sigma_1^*, \sigma_2^*)$ such that σ_1^*, and σ_2^* coincide everywhere with σ_1 and σ_2 except on S, where for $x \in S$ and $i \in \{1, 2\}$, $x\sigma_i^* \in \{\mathbf{0}, \mathbf{1}\}$.*

Proof We may prove this statement by once again following an argument similar to Lemma 2, however, we may assume that g is in a sum-prod ($\sum\prod$) form. Observe that $g =_{\mathsf{SR}} \sum_{i=1}^{m} \prod_{j=1}^{k_i} t_{(i,j)}$ where $t_{(i,j)}$ is either a constant or variable. Thus, $g\sigma_1 =_{\mathsf{SR}} \mathbf{0}$ and $g\sigma_2 =_{\mathsf{SR}} \mathbf{1}$ if and only if

- for all $1 \leq i \leq m$, $\left(\prod_{j=1}^{k_i} t_{(i,j)}\right) \sigma_1 =_{\mathsf{SR}} \mathbf{0}$.
- there exists a unique $1 \leq l_1 \leq m$ such that $\left(\prod_{j=1}^{k_{l_1}} t_{(l_1,j)}\right) \sigma_2 =_{\mathsf{SR}} \mathbf{1}$ and for all $1 \leq l_2 \leq m$, such that $l_2 \neq l_1$, $\left(\prod_{j=1}^{k_{l_2}} t_{(l_2,j)}\right) \sigma_2 = \mathbf{0}$

Note, $(\prod_{j=1}^{k_{l_1}} t_{(l_1,j)}, \sigma_1, \sigma_2)$ is a $(\mathbf{0}, \mathbf{1}, \mathsf{SR})$-generalization and $depth(\prod_{j=1}^{k_{l_1}} t_{(l_1,j)}) < depth(g)$, and thus by the induction hypothesis all variables occurring in $\prod_{j=1}^{k_{l_1}} t_{(l_1,j)}$ are mapped to either $\mathbf{0}$ or $\mathbf{1}$. To finish this argument we need to consider the variables which do not occur in $\prod_{j=1}^{k_{l_1}} t_{(l_1,j)}$ but do occur in g.

Essentially, we follow the argument outlined in Corollary 1 for constructing the appropriate substitutions. □

The above statements imply that we only need to consider $(\{\mathbf{0},\mathbf{1}\},\mathsf{SR})$-reducible $(\mathbf{0},\mathbf{1},\mathsf{SR})$-generalizations. We now show that non-trivial generalizations of this form exist and that $(\{\mathbf{0},\mathbf{1}\},\mathsf{SR})$-reducible generalizations can be transformed into $(\{\mathbf{0},\mathbf{1}\},\mathsf{SR})$-reduced generalizations.

Lemma 3 *There exists a $(\{\mathbf{0},\mathbf{1}\},SR)$-reduced $(\mathbf{0},\mathbf{1},SR)$-generalization $(g,\sigma_1,$ $\sigma_2)$ such that $x <_{SR} g$.*

Proof For example, $(\prod_{i=1}^{n} x,\sigma_1,\sigma_2)$ or $(x\cdot(x+y),\sigma_1,\sigma_2)$, where $n>1$. □

Observation: For $n \geq 1$, $\prod_{i=1}^{n} x <_{\mathsf{SR}} \prod_{i=1}^{2n} x$, $x\cdot(x+y) <_{\mathsf{SR}} (x\cdot x)$, and $x \leq_{\mathsf{SR}} x + (x\cdot y)$.

Theorem 1 *For every $(\{\mathbf{0},\mathbf{1}\},SR)$-reduced $(\mathbf{0},\mathbf{1},SR)$-generalization (g,σ_1,σ_2) there exists a $(\{\mathbf{0},\mathbf{1}\},SR)$-reduced $(\mathbf{0},\mathbf{1},SR)$-generalization (g',σ_1',σ_2') such that $g \leq_{SR} g'$, and $|var(g')| = 1$.*

Proof Using Corollary 2 and the proof of Theorem 2 in [14], we can deduce that a $(g'',\sigma_1'',\sigma_2'')$ may be constructed such that $g \leq_{\mathsf{SR}} g''$ and $1 \leq |var(g'')| \leq 2$. This is possible because we have shown that the substitutions σ_1'' and σ_2'' need only map variables to $\mathbf{0}$ and $\mathbf{1}$. There are only two ways this may be done, for a variable x either $x\sigma_1'' = \mathbf{0}$ and $x\sigma_2'' = \mathbf{1}$ or vice versa. Thus through substitution we can reduce the number of variables to at most two.

By Theorem 3 in [14], there exists a substitution θ such that $g'' = g\theta$. Note that g'' (if in SR-normal form) must be of the following form

$$\left(\sum_{i=1}^{n}\left(\left(\prod_{j=1}^{m_i} y\right)\left(\prod_{j=1}^{k_i} x\right)\right)\right) + \sum_{i=1}^{q}\mathbf{1}$$

This term is a generalization of $\mathbf{0} \triangleq \mathbf{1}$, when $n>1$, if and only if $q=0$ and if there exists a unique i such that $m_i = 0$ and $k_i > 0$ (or $k_i = 0$ and $m_i > 0$). If $n=1$, then we only require either $k_1 > 0$ and/or $m_1 > 0$ for it to be a generalization. This implies (in the former case) $g \leq_{\mathsf{SR}} g\theta\{y \mapsto \mathbf{0}\} = g'$ and $|var(g')| = 1$ and (in the latter case) w.l.o.g $g \leq_{\mathsf{SR}} g\theta\{y \mapsto \mathbf{1}\} = g'$. □

Observation: We only need to consider $(\mathbf{0},\mathbf{1},\mathsf{SR})$-generalizations of the form $(\prod_{j=1}^{n} x, \{x \mapsto \mathbf{0}\}, \{x \mapsto \mathbf{1}\})$ for $n \geq 1$. The following theorem easily follows:

Theorem 2 *Every complete subset of $csg_E(s,t)$ contains g and g' such that $g <_{SR} g'$.*

Proof Let $\mathcal{C}$ by a complete subset of $csg_E(s,t)$ and let $g \in \mathcal{C}$. By Corollaries 1 & 2 we can construct a $(\{\mathbf{0},\mathbf{1}\},\mathsf{SR})$-reducible $(\mathbf{0},\mathbf{1},\mathsf{SR})$-generalization (g,σ_1,σ_2). By Theorem 1, we can derive $(g'',\sigma_1'',\sigma_2'')$ from (g,σ_1,σ_2) such that $|var(g'')| = 1$. Let $g' = g''\{x \mapsto g''\}$. Then by the first bservation $g <_{\mathsf{SR}} g'$. By completeness of $\mathcal{C}$, there exists μ such that $g'\mu \in \mathcal{C}$. □

Corollary 3 *Anti-unification over the theory of semirings is nullary.*

4 Conclusion

Extension of our results to semirings with commutative multiplication and/or idempotent addition is elementary. However, our argument relies on non-idempotent multiplication, and thus nullarity of anti-unification over fully idempotent semirings remains open as discussed earlier concerning $\mathsf{UI_2N}$. Anti-unification over idempotent equational theories has been investigated [13], though theories including unit equations were not considered. We leave the analysis of such theories to future work, though we conjecture that $\mathsf{UI_2N}$, as well as fully idempotent semirings, are anti-unification type infinitary.

Acknowledgements

I would like to thank Temur Kutsia for introducing me to this area and for the discussions which lead to the above results.

References

[1] María Alpuente, Santiago Escobar, Javier Espert, and José Meseguer. A modular order-sorted equational generalization algorithm. *Inf. Comput.*, 235:98–136, 2014.

[2] Nino Amiridze and Temur Kutsia. Anti-unification and natural language processing. EasyChair Preprint no. 203, EasyChair, 2018.

[3] Sebastian Eberhard andStefan Hetzl. Inductive theorem proving based on tree grammars. *Ann. Pure Appl. Logic*, 166(6):665–700, 2015.

[4] Franz Baader. The theory of idempotent semigroups is of unification type zero. *Journal of Automated Reasoning*, 2:283–289, 1986.

[5] Franz Baader. Unification, weak unification, upper bound, lower bound, and generalization problems. In Ronald V. Book, editor, *Rewriting Techniques and Applications, 4th International Conference, RTA-91, Como, Italy, April 10-12, 1991, Proceedings*, volume 488 of *Lecture Notes in Computer Science*, pages 86–97. Springer, 1991.

[6] Franz Baader and Wayne Snyder. Unification theory. In Alan Robinson and Andrei Voronkov, editors, *Handbook of Automated Reasoning*, pages 445 – 533. North-Holland, Amsterdam, 2001.

[7] Johannes Bader, Andrew Scott, Michael Pradel, and Satish Chandra. Getafix: learning to fix bugs automatically. *Proc. ACM Program. Lang.*, 3(OOPSLA):159:1–159:27, 2019.

[8] Adam D. Barwell, Christopher Brown, and Kevin Hammond. Finding parallel functional pearls: Automatic parallel recursion scheme detection in Haskell functions via anti-unification. *Future Generation Comp. Syst.*, 79:669–686, 2018.

[9] Alexander Baumgartner, Temur Kutsia, Jordi Levy, and Mateu Villaret. Higher-order pattern anti-unification in linear time. *J. Autom. Reasoning*, 58(2):293–310, 2017.

[10] Alexander Baumgartner, Temur Kutsia, Jordi Levy, and Mateu Villaret. Term-graph anti-unification. In *3rd International Conference on Formal Structures for Computation and Deduction, FSCD 2018, July 9-12, 2018, Oxford, UK*, pages 9:1–9:17, 2018.

[11] Stefano Bistarelli, Ugo Montanari, and Francesca Rossi. Semiring-based constraint logic programming: Syntax and semantics. *ACM Trans. Program. Lang. Syst.*, 23(1):1âĂŞ29, January 2001.

[12] Jochen Burghardt. E-generalization using grammars. *Artificial intelligence*, 165(1):1–35, 2005.

[13] David Cerna and Temur Kutsia. Idempotent anti-unification. *ACM Trans. Comput. Logic*, 21(2), November 2019.

[14] David Cerna and Temur Kutsia. Unital anti-unification: type and algorithms. In *5th International Conference on Formal Structures for Computation and Deduction, FSCD 2020, june 29- july 6, 2020*, 2020. Preprint: https://www3.risc.jku.at/publications/download/risc_6080/RiscReport.pdf.

[15] David M. Cerna and Temur Kutsia. A generic framework for higher-order generalizations. In *4th International Conference on Formal Structures for Computation and Deduction, FSCD 2019, June 24-30, 2019, Dortmund, Germany*, volume 131 of *LIPIcs*, pages 10:1–10:19, 2019.

[16] Sebastian Eberhard, Gabriel Ebner, and Stefan Hetzl. Algorithmic compression of finite tree languages by rigid acyclic grammars. *ACM Trans. Comput. Log.*, 18(4):26:1–26:20, 2017.

[17] FranÃğois Fages and GÃľrard Huet. Complete sets of unifiers and matchers in equational theories. *Theoretical Computer Science*, 43:189 – 200, 1986.

[18] Boris Galitsky. *Developing Enterprise Chatbots - Learning Linguistic Structures.* Springer, 2019.

[19] Jonathan S. Golan. *Semirings and their Applications.* Springer Netherlands, 1 edition, 1999.

[20] Boris Konev and Temur Kutsia. Anti-unification of concepts in description logic EL. In *Principles of Knowledge Representation and Reasoning: Proceedings of the Fifteenth International Conference, KR 2016, Cape Town, South Africa, April 25-29, 2016.*, pages 227–236, 2016.

[21] Werner Kuich and Arto Salomaa, editors. *Semirings, Automata, Languages*. Springer-Verlag, Berlin, Heidelberg, 1985.

[22] Jianguo Lu, John Mylopoulos, Masateru Harao, and Masami Hagiya. Higher order generalization and its application in program verification. *Annals of Mathematics and Artificial Intelligence*, 28(1-4):107–126, 2000.

[23] Sonu Mehta, Ranjita Bhagwan, Rahul Kumar, Chetan Bansal, Chandra Maddila, B. Ashok, Sumit Asthana, Christian Bird, and Aditya Kumar. Rex: Preventing bugs and misconfiguration in large services using correlated change analysis. In *17th USENIX Symposium on Networked Systems Design and Implementation (NSDI 20)*, pages 435–448, February 2020.

[24] Werner Nutt. Unification in monoidal theories. In Mark E. Stickel, editor, *10th International Conference on Automated Deduction*, pages 618–632, Berlin, Heidelberg, 1990. Springer Berlin Heidelberg.

[25] Frank Pfenning. Unification and anti-unification in the calculus of constructions. In *LICS*, volume 91, pages 74–85, 1991.

[26] Gordon D. Plotkin. A note on inductive generalization. *Machine Intell.*, 5(1):153–163, 1970.

[27] John C. Reynolds. Transformational systems and the algebraic structure of atomic formulas. *Machine Intel.*, 5(1):135–151, 1970.

[28] Reudismam Rolim, Gustavo Soares, Rohit Gheyi, and Loris D'Antoni. Learning quick fixes from code repositories. *CoRR*, abs/1803.03806, 2018.

[29] Ute Schmid. *Inductive Synthesis of Functional Programs, Universal Planning, Folding of Finite Programs, and Schema Abstraction by Analogical Reasoning*, volume 2654 of *Lecture Notes in Computer Science*. Springer, 2003.

[30] Manfred Schmidt-Schauss. Unification under associativity and idempotence is of type nullary. *Journal of Automated Reasoning*, 2:277–281, 1986.

[31] Jörg H. Siekmann. Unification theory. *J. Symb. Comput.*, 7(3/4):207–274, 1989.