On Anti-Unification over Absorption, Associative, and Commutative Theories

Andrés Felipe González Barragán[1], David M. Cerna[2], Mauricio Ayala-Rincón[1], and Temur Kutsia[3]

[1] University of Brasilia, Brasilia D.F., Brazil
andres.felipe@aluno.unb.br, ayala@unb.br
[2] Czech Academy of Sciences, Prague, Czech Republic
dcerna@cs.cas.cz
[3] RISC, Johannes Kepler University, Linz, Austria
kutsia@risc.jku.at

Abstract

Algebras as monoids, semi-groups, and Abelian semi-groups, including absorption operators with their relative absorption constants, are also equipped with commutative (C) and associative (A) properties such as the product operator with the constant zero: $x * 0 \approx 0 * x \approx 0$. We present a sound algorithm and some examples of the anti-unification problem for absorption ($\mathfrak{a}$) theories, including A or C operators.

1 Introduction

Anti-unification (AU) or generalization is a crucial method of reasoning. The problem of AU consists of finding commonalities between two expressions. algorithms aiming to solve this problem find a set of terms that minimally express all possible similarities between input expressions. The problem was introduced by Plotkin and Reynolds, addressing the (syntactic) first-order languages [7, 8]. AU has been studied in several equational theories, such as theories with associative (A) and commutative (C) operators [1], unital [5], and absorption ($\mathfrak{a}$) theories [3]. Moreover, one of the relatively unexplored areas is the investigation of combinations between these theories, as highlighted in related works in [2] and [4]. This abstract discusses the combinations of absorption theories with associative or commutative operators. For a survey on anti-unification, see [6]. In a recent paper [3], the authors presented a sound and complete algorithm that solves anti-unification modulo absorption theories, theories with operators that satisfy the axioms $\{f(\varepsilon_f, x) \approx \varepsilon_f, f(x, \varepsilon_f) \approx \varepsilon_f\}$. This work aims to present recent advancements, introducing two distinct extensions of the anti-unification problem modulo absorption. We consider absorption symbols together with associative and commutative symbols, treated separately in the same set of axioms. The inclusion of this kind of symbols raises new generalizations that were not considered before either in $\mathfrak{a}$- or C- or A-theories, then we assemble the existing algorithms in [3, 1] and introduce new rules to handle these generalizations. It is important to highlight that here in this new approach, the role of $\star$ in the expansions of the absorption constants within commutative or associative properties could lead us to new generalizations that need to be captured for the algorithm. This algorithm is terminating, sound, and capable of capturing generalizations for this kind of combination.

1.1 Preliminaries

Let $\mathcal{V}$ be a countable set of variables and $\mathcal{F}$ a set of function symbols with a fixed arity. Additionally, we assume $\mathcal{F}$ to contain a special constant $\star$, referred to as the *wild card*. The set of terms derived from $\mathcal{F}$ and $\mathcal{V}$ is denoted by $\mathcal{T}(\mathcal{F},\mathcal{V})$, whose members are constructed using the grammar $t ::= x \mid f(t_1,\dots,t_n)$, where $x \in \mathcal{V}$ and $f \in \mathcal{F}$ with arity $n \geq 0$. When $n = 0$, f is called a *constant*. The set of variables occurring in t is denoted by $var(t)$. The *size* of a term is defined inductively as: $size(x) = 1$, and $size(f(t_1,\dots,t_n)) = 1 + \sum_{i=1}^{n} size(t_i)$. Let σ be a substitution, $dom(\sigma)$, and $rvar(\sigma)$ denote the domain and the set of variables occurring in terms of the range of σ, respectively. The *head* of a term t is defined as $head(x) = x$ and $head(f(t_1,\dots,t_n)) = f$, for $n \geq 0$.

The focus of this work is anti-unification modulo equational theories E that may include commutative symbols, for short C-symbols, with axioms for commutativity, $\{f(x,y) = f(y,x)\}$, associative symbols (A-symbols), with axioms for Associativity, $\{f(f(x,y),z) = f(x,f(y,z))\}$, and absorption symbols, for short $\mathfrak{a}$-symbols, with absorption axioms, $\{f(x,\varepsilon_f) \approx \varepsilon_f,\ f(\varepsilon_f,x) \approx \varepsilon_f\}$. Symbols f and ε_f are called *related* $\mathfrak{a}$*-symbols*. An $(E)(E')$-theory includes E-symbols and E'-symbols and an EE'-theory includes symbols holding E- and E'-axioms simultaneously.

Definition 1 (E-generalization, $\preceq_E$)**.** *The generalization relation of the theory induced by E holds for terms $r, s \in \mathcal{T}(\mathcal{F},\mathcal{V})$, written $r \preceq_E s$, if there exists a substitution σ such that $r\sigma \approx_E s$. An E-generalization r of s and t is a term r such that $r \preceq_E s$ and $r \preceq_E t$.*

Example 1. *Consider $\mathfrak{a}C = \{f(\varepsilon_f,x) \approx \varepsilon_f, f(x,y) \approx f(y,x)\}$. Then $f(x,a)$ is an $\mathfrak{a}C$-generalization of ε_f and $f(a,a)$: $f(x,a)\{x \mapsto \varepsilon_f\} \approx_{\mathfrak{a}C} \varepsilon_f$ and $f(x,a)\{x \mapsto a\} \approx_{\mathfrak{a}C} f(a,a)$.*

Definition 2 (Minimal complete set of E-generalizations)**.** *The* minimal complete set of E-generalizations *of the terms s and t, denoted as $mcsg_E(s,t)$, is a subset of $\mathcal{G}_E(s,t)$, the set of all E-generalizations of s and t, satisfying: (i) for each $r \in \mathcal{G}_E(s,t)$ there exists $r' \in mcsg_E(s,t)$ such that $r \preceq_E r'$; (ii) if $r, r' \in mcsg_E(s,t)$ and $r \preceq_E r'$, then $r = r'$.*

Example 2. *Continuing Example 1, notice that $mcsg_{\mathfrak{a}C}(\varepsilon_f, f(a,a)) = \{f(x,a), f(x,x)\}$, and $mcsg_{\mathfrak{a}}(\varepsilon_f, f(a,a)) = \{f(x,a), f(a,x), f(x,x)\}$.*

Definition 3 (Anti-unification type)**.** *The anti-unification type of an equational theory E is said to be* unitary *if $mcsg_E(s,t)$ is a singleton for all terms s and t; it is* finitary *if it is not unitary but $mcsg_E(s,t)$ is always finite; it is* infinitary *if it is neither unitary nor finitary but $mcsg_E(s,t)$ always exists; otherwise, it is said to be* nullary*.*

Syntactic AU is *unitary* [7, 8], AU over (A) and (C) theories is *finitary* [1], AU over ($\mathfrak{a}$) theories is *infinitary* [4], and AU with a disjoint combination of unital equations is *nullary* [5].

An *anti-unification triple* (AUT), $s \triangleq_x t$, consists of a *label* $x \in \mathcal{V}$, and two terms s and t. Given a set A of AUTs, $labels(A) = \{x \mid s \triangleq_x t \in A\}$ and $size(A) = \sum_{s \triangleq_x t \in A} \big(size(s) + size(t)\big)$. A set of AUTs is *valid* if its labels are pairwise disjoint. A *wild* AUT is of the form either $\star \triangleq_x s$ or $s \triangleq_x \star$. A non-wild AUT $s \triangleq_x t$ is *solved* over an absorption theory $\mathfrak{a}$ if $head(s)$ and $head(t)$ are different and they are not related $\mathfrak{a}$-symbols.

The label x in an AUT $s \triangleq_x t$, as a variable, is a most general generalization of the terms s and t, and it is used to associate the generalizations of s and t. The wild card plays an important role when anti-unification problems are decomposed, and related $\mathfrak{a}$-symbols appear in the head of AUTs; they will represent any possible term expanding and $\mathfrak{a}$-constant symbol needs to be expanded ($\varepsilon_f \approx_{\mathfrak{a}} f(\varepsilon_f,\star)$ or $\varepsilon_f \approx_{\mathfrak{a}} f(\star,\varepsilon_f)$), see [3].

2 Anti-Unification in Absorption Theories with Commutative or Associative Properties

Several algebras having absorption property like semi-groups, Abelian semi-groups, and monoids may include the associative or/and commutative property. Interesting examples of these algebras are the integers with multiplication with zero as absorption constant; the integers with the greatest common divisor *gcd* with one as the absorption constant; the $n \times n$-matrices over reals with the product and the zero matrix; the powerset of a given set with the intersection $\cap$ with $\varnothing$ as absorption constant; Boolean algebras with two binary operations, where each operation is associative, commutative, and has zero element. This section shows how generalizations for $\mathfrak{a}$ theories with A, C symbols differ from generalizations of pure $\mathfrak{a}$ theories presented in [3].

Example 3. *The set* $mcsg_{\mathfrak{a}}(\varepsilon_f, f(a,a)) = \{f(x,a), f(a,x), f(x,x)\}$*, which is different from the set* $mcsg_{\mathfrak{a}C}(\varepsilon_f, f(a,a))$ *(see Example 2), is computed by the algorithm in [3]. Also, for the more elaborated example,* $mcsg_{\mathfrak{a}C}(\varepsilon_f, f(f(a,a), f(a,a)))$ *does not include* $\mathfrak{a}$ *minimal generalizations as* $f(f(a,a), f(u,a))$ *and* $f(f(a,u), f(a,u))$ *in* $mcsg_{\mathfrak{a}}(\varepsilon_f, f(f(a,a), f(a,a)))$.

An algorithm to compute generalizations in ($\mathfrak{a}$)($\mathfrak{a}$C)(C)-theories should include rules to treat C symbols as in [1], and also adaptations of the expansion and merge rules introduced in [3] for $\mathfrak{a}$ theories, to deal with $\mathfrak{a}$C symbols.

Example 4. *The set* $mcsg_{\mathfrak{a}A}(g(\varepsilon_f, a), g(f(f(a,a), f(a,a)), f(a, f(a,a))))$ *includes the* $\mathfrak{a}$*A-generalization* $g(f(x,y),y)$*, where* g *is a syntactic symbol and* f *is an* $\mathfrak{a}$*A-symbol. Notice that this is not an* $\mathfrak{a}$*-generalization.*

In the case of ($\mathfrak{a}$)($\mathfrak{a}$A)(A)-theories, standard flattened notation is used, and for an A-symbol, f, the flattened term $f(t)$ equals t. An algorithm to compute the generalizations requires designing specialized rules, adapted from [3], to deal with $\mathfrak{a}$A symbols.

3 Algorithm for Absorption with Commutative or Associative Theories

Tables 1, 2, and 3 present inference rules for theories with $\mathfrak{a}$-, $\mathfrak{a}$C-, $\mathfrak{a}$A-, A-, and C-symbols. The algorithm AUnif consists of applying these rules exhaustively, returning a set of objects from which generalizations of the input AUTs may be derived. The inference rules work on *configurations*, defined below.

Definition 4 (Configuration)**.** *A configuration is a quadruple of the form* $\langle A; S; D; \theta\rangle$*, where,* A *is a valid set of AUTs (*active set*);* S *is a valid set of solved AUTs (*store*);* D *is a valid set of* wild *AUTs (*delayed set*);* θ *is a* substitution *such that* $rvar(\theta) = labels(A) \cup labels(S) \cup labels(D)$ *(*anti-unifier*); and with the property that* $labels(A), labels(S), labels(D)$*, and* $dom(\theta)$ *are pairwise disjoint.*

All terms occurring in a configuration are in their $\mathfrak{a}$-normal forms. For $\mathfrak{a}$A- and A-symbols, all terms in a configuration are considered in the flattened form.

Table 1 contains rules: Decompose ($\overset{Dec}{\Longrightarrow}$), Solve ($\overset{Sol}{\Longrightarrow}$), Expansions for Left Absorption, ($\overset{ExpL1}{\Longrightarrow}$ and $\overset{ExpL2}{\Longrightarrow}$), Expansions for Right Absorption ($\overset{ExpR1}{\Longrightarrow}$ and $\overset{ExpR2}{\Longrightarrow}$), and Expansion Absorption in Both sides ($\overset{ExpB1}{\Longrightarrow}$ and $\overset{ExpB2}{\Longrightarrow}$), representing the common rules. Table 2 has the extra rules

Commutative ($\overset{Com}{\Longrightarrow}$), and Table 3 shows the additional rules Associativity Left and Right ($\overset{AL}{\Longrightarrow}$) and ($\overset{AR}{\Longrightarrow}$), and Absorption-Associative Left and Right 1,2 ($\overset{\mathfrak{a}\text{AL1}}{\Longrightarrow}$), ($\overset{\mathfrak{a}\text{AL2}}{\Longrightarrow}$), ($\overset{\mathfrak{a}\text{AR1}}{\Longrightarrow}$), and ($\overset{\mathfrak{a}\text{AR2}}{\Longrightarrow}$).

By $\mathcal{C} \Longrightarrow^* \mathcal{C}'$ we denote a finite sequence of inference rule applications starting at $\mathcal{C}$ and ending with $\mathcal{C}'$. In both cases we say $\mathcal{C}'$ is *derived* from $\mathcal{C}$. An initial configuration is a configuration of the form $\langle A; \varnothing; \varnothing; \iota\rangle$, where $\iota = \{\mathsf{f}_A(x) \mapsto x \mid x \in \mathit{labels}(A)\}$ with $\mathsf{f}_A : \mathcal{V} \to (\mathcal{V} \setminus \mathit{labels}(A))$ being a bijection over variables. A configuration $\mathcal{C}$ is referred to as *final* if no inference rule applies to $\mathcal{C}$. We denote the set of final configurations finitely derived from an initial configuration $\mathcal{C}$ by AUnif($\mathcal{C}$).

Example 5. *Notice that* AUnif *computes the generalization $f(x,a)$ for the problem in Example 3 using the rules (ExpL1) and (Sol); and the generalization $g(f(x,z),y)$ for the problem in Example 4 using the rules (Dec),($\mathfrak{a}AL1$) for $k = 1$, and (Sol).*

Table 1: Inference rules common to all theories.

$$(\overset{Dec}{\Longrightarrow}) \quad \frac{\langle\{f(s_1,\dots,s_n) \triangleq_x f(t_1,\dots,t_n)\} \uplus A; S; D; \theta\rangle}{\langle\{s_1 \triangleq_{y_1} t_1,\dots,s_n \triangleq_{y_n} t_n\} \cup A; S; D; \theta\{x \mapsto f(y_1,\dots,y_n)\}\rangle}$$

where f is any symbol, $n \geq 0$, and $y_1,\dots,y_n$ are fresh variables.

$$(\overset{Sol}{\Longrightarrow}) \quad \frac{\langle\{s \triangleq_x t\} \uplus A; S; D; \theta\rangle}{\langle A; \{s \triangleq_x t\} \cup S; D; \theta\rangle}$$

where $\mathit{head}(s) \neq \mathit{head}(t)$ and they are not related $\mathfrak{a}$-symbols.

$$(\overset{Mer}{\Longrightarrow}) \quad \frac{\langle\varnothing; \{s_1 \triangleq_x t_1, s_2 \triangleq_y t_2\} \cup S; T; \theta\rangle}{\langle\varnothing; \{s_2 \triangleq_y t_2\} \cup S; T; \theta\{x \mapsto y\}\rangle}$$

where $s_1 \approx_E s_2$, $t_1 \approx_E t_2$, $x \neq y$, and E is an equational theory.

In the following rules, f is an $\mathfrak{a}$-, $\mathfrak{a}$C-, or $\mathfrak{a}$A-symbol, and y_1, y_2 are fresh variables:

$$(\overset{ExpL1}{\Longrightarrow}) \quad \frac{\langle\{\varepsilon_f \triangleq_x f(t_1,t_2)\} \uplus A; S; D; \theta\rangle}{\langle\{\varepsilon_f \triangleq_{y_1} t_1\} \cup A; S; \{\star \triangleq_{y_2} t_2\} \cup D; \theta\{x \mapsto f(y_1,y_2)\}\rangle}$$

$$(\overset{ExpL2}{\Longrightarrow}) \quad \frac{\langle\{\varepsilon_f \triangleq_x f(t_1,t_2)\} \uplus A; S; D; \theta\rangle}{\langle\{\varepsilon_f \triangleq_{y_2} t_2\} \cup A; S; \{\star \triangleq_{y_1} t_1\} \cup D; \theta\{x \mapsto f(y_1,y_2)\}\rangle}$$

$$(\overset{ExpR1}{\Longrightarrow}) \quad \frac{\langle\{f(s_1,s_2) \triangleq_x \varepsilon_f\} \uplus A; S; D; \theta\rangle}{\langle\{s_1 \triangleq_{y_1} \varepsilon_f\} \cup A; S; \{s_2 \triangleq_{y_2} \star\} \cup D; \theta\{x \mapsto f(y_1,y_2)\}\rangle}$$

$$(\overset{ExpR2}{\Longrightarrow}) \quad \frac{\langle\{f(s_1,s_2) \triangleq_x \varepsilon_f\} \uplus A; S; D; \theta\rangle}{\langle\{s_2 \triangleq_{y_2} \varepsilon_f\} \cup A; S; \{s_1 \triangleq_{y_1} \star\} \cup D; \theta\{x \mapsto f(y_1,y_2)\}\rangle}$$

$$(\overset{ExpB1}{\Longrightarrow}) \quad \frac{\langle\{\varepsilon_f \triangleq_x \varepsilon_f\} \uplus A; S; D; \theta\rangle}{\langle A; S; \{\varepsilon_f \triangleq_{y_1} \star,\ \star \triangleq_{y_2} \varepsilon_f\} \cup D; \theta\{x \mapsto f(y_1,y_2)\}\rangle}$$

$$(\overset{ExpB2}{\Longrightarrow}) \quad \frac{\langle\{\varepsilon_f \triangleq_x \varepsilon_f\} \uplus A; S; D; \theta\rangle}{\langle A; S; \{\star \triangleq_{y_1} \varepsilon_f,\ \varepsilon_f \triangleq_{y_2} \star\} \cup D; \theta\{x \mapsto f(y_1,y_2))\}\rangle}$$

Lemma 1 (Configuration Preservation)**.** *Let $\mathcal{C}$ be a configuration and $\mathcal{C} \Longrightarrow^* \mathcal{C}'$. Then $\mathcal{C}'$ is a configuration.*

Proof. According to the rules in Tables 1,2 and 3 we can have the following two cases:

Table 2: Inference rule for αC- and C-symbols.

$$(\overset{Com}{\Longrightarrow}) \quad \frac{\langle \{f(s_1,s_2) \triangleq_x f(t_1,t_2)\} \cup A; S; T; \theta\rangle}{\langle \{s_1 \triangleq_{y_1} t_2, s_2 \triangleq_{y_2} t_1\} \cup A; S; T; \theta\{x \mapsto f(y_1,y_2)\}\rangle}$$

for f an αC- or C-symbol and y_1, y_2 fresh variables.

Table 3: Inference rule for αA and A symbols.

In the next two rules, g is either an αA-symbol or an A-symbol:

$$(\overset{AL}{\Longrightarrow}) \quad \frac{\langle \{g(s_1,\ldots,s_n) \triangleq_x g(t_1,\ldots,t_m)\} \cup A; S; T; \theta\rangle}{\langle \{s_1 \triangleq_{y_1} g(t_1,\ldots,t_k), g(s_2,\ldots,s_n) \triangleq_{y_2} g(t_{k+1},\ldots,t_m)\} \cup A; S; T; \theta\{x \mapsto g(y_1,y_2)\}\rangle}$$

for $1 \le k \le m-1$ and y_1, y_2 are fresh variables.

$$(\overset{AR}{\Longrightarrow}) \quad \frac{\langle\{\{g(s_1,\ldots,s_n) \triangleq_x g(t_1,\ldots,t_m)\} \cup A; S; T; \theta\rangle}{\langle \{g(s_1,\ldots,s_k) \triangleq_{y_1} t_1, g(s_{k+1},\ldots,s_n) \triangleq_{y_2} g(t_2,\ldots,t_m)\} \cup A; S; T; \theta\{x \mapsto g(y_1,y_2)\}\rangle}$$

for $1 \le k \le n-1$ and y_1, y_2 are fresh variables.

Next five rules apply to αA-symbols, and $1 \le k \le n-1$, and y_1, y_2 are fresh variables:

$$(\overset{\alpha\mathbf{AL1}}{\Longrightarrow}) \quad \frac{\langle \{\varepsilon_f \triangleq_x f(t_1,\ldots,t_n)\} \cup A; S; T; \theta\rangle}{\langle \{\varepsilon_f \triangleq_{y_1} f(t_1,\ldots,t_k)\} \cup A; S; \{\varepsilon_f \triangleq_{y_2} f(t_{k+1},\ldots,t_n)\} \cup T; \theta\{x \mapsto f(y_1,y_2)\}\rangle}$$

$$(\overset{\alpha\mathbf{AL2}}{\Longrightarrow}) \quad \frac{\langle \{\varepsilon_f \triangleq_x f(t_1,\ldots,t_n)\} \cup A; S; T; \theta\rangle}{\langle \{\varepsilon_f \triangleq_{y_2} f(t_{k+1},\ldots,t_n)\} \cup A; S; \{\varepsilon_f \triangleq_{y_1} f(t_1,\ldots,t_k)\} \cup T; \theta\{x \mapsto f(y_1,y_2)\}\rangle}$$

$$(\overset{\alpha\mathbf{AR1}}{\Longrightarrow}) \quad \frac{\langle \{f(s_1,\ldots,s_n) \triangleq_x \varepsilon_f\} \cup A; S; T; \theta\rangle}{\langle \{f(t_1,\ldots,t_k) \triangleq_{y_1} \varepsilon_f\} \cup A; S; \{f(t_{k+1},\ldots,t_n) \triangleq_{y_2} \varepsilon_f\} \cup T; \theta\{x \mapsto f(y_1,y_2)\}\rangle}$$

$$(\overset{\alpha\mathbf{AR2}}{\Longrightarrow}) \quad \frac{\langle \{f(s_1,\ldots,s_n) \triangleq_x \varepsilon_f\} \cup A; S; T; \theta\rangle}{\langle \{f(t_{k+1},\ldots,t_n) \triangleq_{y_2} \varepsilon_f\} \cup A; S; \{f(t_1,\ldots,t_k) \triangleq_{y_1} \varepsilon_f\} \cup T; \theta\{x \mapsto f(y_1,y_2)\}\rangle}$$

- A rule removes an AUT $s \triangleq_x t$ from the active set of $\mathcal{C}$. Then either $s \triangleq_x t$ occurs in the store of $\mathcal{C}'$, or the anti-unifier component of $\mathcal{C}'$ is the composition of the anti-unifier component of $\mathcal{C}$ with $\{x \mapsto r\}$, where $var(r)$ are fresh variables labeling newly added AUTs in the active and delayed sets of $\mathcal{C}'$.
- A rule removes an AUT $s \triangleq_x t$ from the store of $\mathcal{C}$. Then the store of $\mathcal{C}'$ is a subset of the store of $\mathcal{C}$ and the anti-unifier component of $\mathcal{C}'$ is the composition of the anti-unifier component of $\mathcal{C}$ with $\{x \mapsto y\}$, where y is a label of an AUT in the store of $\mathcal{C}$ such that $x \neq y$.

In both cases, the properties of a configuration are preserved. □

Theorem 1 (Termination)**.** *Let $\mathcal{C}$ be a configuration. Then* AUNIF$(\mathcal{C})$ *is finitely computable.*

Proof. The termination of AUNIF is proved using a lexicographical measure over configurations. The measure for $\mathcal{C} = \langle A; S; T; \theta\rangle$ is given by $(size(A), size(S))$. All rules except (Mer) decrease the first component, and (Mer) maintains the first but decreases the second component. □

Termination (Theorem 1) guarantees that always is possible to obtain a final configuration. Configurations of the form $\langle \varnothing; S; D; \theta\rangle$ where S has no duplicated AUTs, except for the label, are final configurations since no rule can be applied. Since all rules, except for (Mer), decrease the size of the active set, it becomes empty, and the rule (Mer) will eliminate all such possible duplication in the store. Configuration preservation (Lemma 1) and termination allow proving the soundness of AUnif for the combination of the $\mathfrak{a}$-, $\mathfrak{a}$C-, $\mathfrak{a}$A-, C-, and A-theories.

Theorem 2 (Soundness). *Let $\langle \emptyset; S_n; D_n; \theta_n \rangle \in \text{AUnif}(\langle A_0; S_0; D_0; \theta_0 \rangle)$, and E be any combination of the theories $\mathfrak{a}$, $\mathfrak{a}C$, $\mathfrak{a}A$, C, and A. Then, for all $s \triangleq_x t \in A_0 \cup S_0$, $x\theta_n \in \mathcal{G}_E(s,t)$.*

Proof. The proof is by induction on the length of derivations, analyzing each rule application. □

4 Work in progress

Work in progress addresses adaptation of AUnif to allow combinations in which (AC)- and ($\mathfrak{a}$AC)-symbols are also allowed. Of course, it also aims to prove completeness. For theories with C-, $\mathfrak{a}$-, and $\mathfrak{a}$C-symbols, currently under study, the proof requires induction on the occurrence of variables in the possible generalizations interrelated with structural analysis of the AUTs under the action of AUnif. The analysis is much more elaborate than the applied on the proof of completeness for $\mathfrak{a}$-theories in [3]. Additionally, succeeding in the completeness proof will imply that the anti-unification problem is infinitary.

Acknowledgments. The research is supported by the Brazilian Coordination for the Improvement of Higher Education Personnel (CAPES), under the PrInt program. It is partially supported by the Austrian Science Fund (FWF), under project P 35530, the Georgian Rustaveli National Science Foundation Project FR-21-16725, and by the Czech Science Foundation Grant No. 22-06414L, and the Cost Action CA20111 EuroProofNet.

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] María Alpuente, Santiago Escobar, Javier Espert, and José Meseguer. Order-sorted equational generalization algorithm revisited. *Ann. Math. Artif. Intell.*, 90(5):499–522, 2022.

[3] Mauricio Ayala-Rincón, David M. Cerna, Andrés Felipe González Barragán, and Temur Kutsia. Equational anti-unification over absorption theories. In *Automated Reasoning - 12th International Joint Conference, IJCAR 2024, Nancy, France, Proceedings*, volume 14740 of *Lecture Notes in Computer Science*. Springer, 2024.

[4] David M. Cerna. Anti-unification and the theory of semirings. *Theor. Comput. Sci.*, 848:133–139, 2020.

[5] David M. Cerna and Temur Kutsia. Unital anti-unification: Type and algorithms. In *5th Int. Conference on Formal Structures for Computation and Deduction, FSCD*, volume 167 of *LIPIcs*, pages 26:1–26:20, 2020.

[6] David M. Cerna and Temur Kutsia. Anti-unification and generalization: A survey. In *Proceedings of the 32nd Int. Joint Conference on Artificial Intelligence, IJCAI*, pages 6563–6573, 2023.

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

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