Bounded ACh Unification

Guilherme Borges Brandão

Advisor: Daniele Nantes Sobrinho
Department of Mathematics
Universidade de Brasília
Funded by a CNPq Masters scholarship

February 7th, 2024

Table of Contents

Table of Contents

Unification is a method to find solutions for a set of equations.

Example

Consider the equation $f(X, Y) \stackrel{?}{=} f(a, b)$, where X, Y are variables and a, b are constants. If f is an uninterpreted function symbol, the solution for this equation is $\{X \mapsto a, Y \mapsto b\}$.

Unification Modulo ACh

We're interested in the problem of ACh Unification, i.e. unification when we have an homomorphism h and the function symbol $+$ is associative and commutative.

Example

The set of equations

$$\Gamma = \{h(h(X_1) + X_2)) =^? h(Y_1 + h(Y_2))\}$$

is an ACh unification problem.

A Step Back

- Unfortunately, it has been proved that this problem is undecidable [Nar96].
- But, recently, Eeralla and Lynch [EL20] proposed an approximation of the ACh Unification problem by introducing a bound on the number of h symbols occurring in a term.
- Our goal is to verify the proof of correctness of the ACh Unification algorithm presented in [EL20].

Table of Contents

Table of Contents

Flattened form

For convenience, we assume that our unification problem is always in flattened form, that is, every equation has one the following forms:

- $X \stackrel{?}{=} Y$ or
- $X \stackrel{?}{=} X_1 + \ldots + X_n$ or
- $X \stackrel{?}{=} f(X_1, \ldots, X_n)$ or
- $X \stackrel{?}{=} h(Y)$

where $X, Y, X_1, \ldots, X_n$ are variables.

h-depth of a variable

- **Given:** A unification problem Γ, $X \in var(\Gamma)$ and h a unary function symbol.
- **Define:** the *h-depth of X* as the maximum number of h symbols along a path to X.
- **Define:** the h-depth set of Γ as

$$\Delta := \{(X, h_d(X, \Gamma)) | X \in var(\Gamma)\}$$

- **Define:** $MaxVal(\Delta) := \max\{c \mid (X, c) \in \Delta\}$

Problem Format

- For our problem we use a set triple $\Gamma||\Delta||\sigma$ where
  - Γ is our ACh-unification problem
  - Δ is an h-depth set of Γ, where all the elements initially have the form $(V, 0)$, with $V \in Var(\Gamma)$
  - σ is a substitution.
- Given a bound $\kappa \in \mathbb{N}$:
  $\Gamma||\Delta||\sigma$ is said to be in *solved form* if $\Gamma = \emptyset$ and $MaxVal(\Delta) \leq \kappa$.

Table of Contents

Flattening Rules

The first step is to put the equations in Γ in flattened form. The rules are:

Flatten Both Sides

$$\frac{\{t_1 \stackrel{?}{=} t_2\} \cup \Gamma || \Delta || \sigma}{\{V \stackrel{?}{=} t_1, V \stackrel{?}{=} t_1\} \cup \Gamma || \{(V, 0)\} \cup \Delta || \sigma} \text{if } t_1, t_2 \notin \mathcal{V}$$

Flatten Right (Left is similar)

$$\frac{\{t \stackrel{?}{=} t_1 + t_2\} \cup \Gamma || \Delta || \sigma}{\{t \stackrel{?}{=} t_1 + V, V \stackrel{?}{=} t_2\} \cup \Gamma || \{(V, 0)\} \cup \Delta || \sigma} \text{if } t_2 \notin \mathcal{V}$$

Flatten under h

$$\frac{\{t \stackrel{?}{=} h(t_1)\} \cup \Gamma || \Delta || \sigma}{\{t \stackrel{?}{=} h(V), V \stackrel{?}{=} t_1\} \cup \Gamma || \{(V, 1)\} \cup \Delta || \sigma} \text{if } t_1 \notin \mathcal{V}$$

Update Δ

$$\textsc{Update } h \quad \frac{\{X \stackrel{?}{=} h(Y)\} \cup \Gamma || \{(X, c_1), (Y, c_2)\} \cup \Delta || \sigma}{\{X \stackrel{?}{=} h(Y)\} \cup \Gamma || \{(X, c_1), (Y, c_1 + 1)\} \cup \Delta || \sigma} \text{if } c_2 < c_1 + 1$$

$$\textsc{Update Right + (Left is Similar)} \quad \frac{\{X \stackrel{?}{=} Y_1 + Y_2\} \cup \Gamma || \{(X, c_1), (Y_1, c_2), (Y_2, c_3)\} \cup \Delta || \sigma}{\{X \stackrel{?}{=} Y_1 + Y_2\} \cup \Gamma || \{(X, c_1), (Y_1, c_2), (Y_2, c_1)\} \cup \Delta || \sigma} \text{if } c_3 < c_1$$

Bound Check

- Given κ as a bound for our problem:

$$\textsc{Bound Check}\quad \frac{\Gamma||\Delta||\sigma}{\perp} \quad \text{if } MaxVal(\Delta) > \kappa.$$

Splitting Rule

This rule takes homomorphism theory into consideration

Splitting (Split)

$$\frac{\{X \stackrel{?}{=} h(Y), X \stackrel{?}{=} X_1 + ... + X_n\} \cup \Gamma || \Delta || \sigma}{\{X \stackrel{?}{=} h(Y), Y \stackrel{?}{=} V_1 + ... + V_n, X_1 \stackrel{?}{=} h(V_1), ..., X_n \stackrel{?}{=} h(V_n)\} \cup \Gamma || \Delta' || \sigma}$$

where $X \neq Y$, $X \neq X_i$ for any $i = 1, \ldots, n$ and

$$\Delta' := \{(V_1, 1), ..., (V_n, 1)\} \cup \Delta$$

AC Unification Rule

Remarks:

- This rule invokes any AC unification algorithm to unify the AC part of the problem.
- We will use the standard algorithm by Fages and Stickel [Fag87] and the work by Gabriel Silva et. al. [AFSS22]

AC Unification Rule

AC UNIFICATION (AC)

$$\frac{\Psi \cup \Gamma || \Delta || \sigma}{GetEqs(\theta_1) \cup \Gamma \vee \ldots \vee GetEqs(\theta_n) \cup \Gamma || \Delta || \sigma}$$

$$\text{where } Unify(\Psi) = \{\theta_1, \ldots, \theta_n\}$$

- Here Ψ represents all the equations with the $+$ on the right hand side
- $Unify$ is a function that returns one of the complete set of unifiers returned by the AC unification algorithm.
- $GetEqs$ is a function that takes a substitution and returns one's equational form.

Table of Contents

Algorithm

We're currently studying the Algorithm which gives an order to the application of the rules

Algorithm 1: Unify$_{\mathsf{AC_h}}$

Input: An equation set Γ, an empty h-depth set Δ, an empty set σ and a bound $\kappa \in \mathbb{N}$

Output: A complete set of κ-bounded ACh unifiers $\{\sigma_1, \ldots, \sigma_n\}$ or $\perp$ indicating that the problem has no solution.

Begin

1. Apply Algorithm 2 (Flattening) on Γ
2. **Repeat**

   (Apply (VE1) exhaustively after each of the following rule applications)

   (a) Apply (TRIV) exhaustively to eliminate equations of the form $t \stackrel{?}{=} t$;

   (b) Apply the (OC), i.e., **If** any variable on the left side occurs on the right **then** return $\perp$;

   (c) Apply the (BC), i.e., **If** $\mathcal{M}ax\mathcal{V}al(\Delta) > \kappa$ **then** return $\perp$;

   (d) **If** at least one of the h-depth update rules ((Uh), (UL) or (UR)) is applicable **then** applly the rule and go to (c) **else** go to next step;

   (e) Apply (OR) exhaustively;

   (f) **If** (SPLIT) is applicable **then** apply the rule and go to (a);

   (g) Apply (CLASH), i.e., **If** the top symbols of the left and right sides of an equation do not match **then** return $\perp$;

   (h) **If** (DEC) is applicable **then** apply the rule and go to (a);

   (i) **If** there is at least one variable X occuring left side in at least two equations of the form $X \stackrel{?}{=} Y_1 + \ldots Y_n$ and $X \stackrel{?}{=} Z_1 + \ldots + Z_n$, **then** apply the (AC) rule and go to (d) **else** go to next step;

   (j) Apply (VE2) exhaustively and return the output;

End

Algorithm 2: Flattening

Input: An equation set Γ

Output: An equation set Γ' where all of the equations are in *flattened form*.

while *any of the flattening rules can be applied* **do**
- Apply (FBS)
- Apply (FL)
- Apply (FR)
- Apply (FU)
- Apply (FLFUN)

Table of Contents

Summary

	Results	Status
Termination	Lemma 2, Theorem 4	✗ fix the measure/proof new definition
Soundness	Theorem 9	✓ proofs verified
Completeness	Theorem 11 Theorem 12	✓ completed omitted proofs ✗ imprecision. Open.

Table of Contents

Termination of $\mathsf{Unify}_{\mathsf{AC_h}}$- Measure

Let $\Gamma \parallel \Delta \parallel \sigma$ be a triple. Consider the following measure for $\Gamma \parallel \Delta \parallel \sigma$:

$$\mathbb{M}_{\mathfrak{I}_{ACh}}(\Gamma, \Delta, \sigma) := (\kappa - a, n_X, |\mathcal{S}ym(\Gamma)|, m, p, |\Gamma|, \overline{h_d}(\Delta)), \text{ where}$$

- κ the given bound.
- a is be the number of applications of the (AC) rule.
- m: number of equations on the form $t \stackrel{?}{=} X \in \Gamma$, with $t \notin \mathcal{V}$.
- p: number of isolated variables in Γ $(X \stackrel{?}{=} t,\ X \notin \mathcal{V}ar(t))$.

Imprecision found

Lemma 2. *Let $\Gamma||\Delta||\sigma$ be a set triple, and κ be a natural number (bound) given as an input to the algorithm. Then, the maximum number of times the AC unification applied is κ.*

Proof. The only time the AC unification is invoked on the problem is when there is at least one non-solved variable in the problem. A variable x occurs at least in two equations as $x \stackrel{?}{=} y_1 + \cdots +$

org/10.1017/S0960129520000183 Published online by Cambridge University Press

y_n and $x \stackrel{?}{=} z_1 + \cdots + z_m$. On each application of the AC unification, the lowest depth of non-solved variables get solved, and there is no other rule that makes these variables non-solved again. Hence, the maximum number of times the AC unification could be applied is the κ. $\square$

Here the authors claim that there are no rules that makes a solved variable non-solved without defining what a solved variable is.

Imprecision found

- We figured that the concept of solved/non variable is related to the AC part of the problem.
- We came up with a definition for AC-solved variable;
- With that, we managed to prove the affirmation and that the measure decreases;

Table of Contents

Soundness

Definition (Satisfiability)

Let $\Gamma \parallel \Delta \parallel \sigma$ be a set triple and $\kappa \in \mathbb{N}$ be a given bound.
A substitution θ satisfies $\Gamma \parallel \Delta \parallel \sigma$ iff

1. $\theta \vDash \Gamma$, ($s \stackrel{?}{=} t \in \Gamma$ implies $s\theta = t\theta$)
2. $\theta \vDash \sigma$, ($X \mapsto t \in \sigma$ implies $X\theta = t\theta$)
3. $\mathcal{M}ax\mathcal{V}al(\Delta) \leq \kappa$

(Notation: $\theta \vDash \Gamma \parallel \Delta \parallel \sigma$)

Soundness

Then, by induction on the number of steps, we obtain:

Theorem 1

Let $\Gamma \parallel \Delta \parallel \sigma$ *and* $\Gamma' \parallel \Delta' \parallel \sigma' = \bigvee_{i \in I}(\Gamma_i \parallel \Delta_i \parallel \sigma_i)$ *be two ACh unification problems such that*

$$\Gamma \parallel \Delta \parallel \sigma \Rightarrow^*_{\mathfrak{I}_{ACh}} \Gamma' \parallel \Delta' \parallel \sigma'.$$

If θ *is a substitution such that* $\theta \vDash \Gamma_i \parallel \Delta_i \parallel \sigma_i$*, then* $\theta \vDash \Gamma \parallel \Delta \parallel \sigma$*.*

Soundness

Corollary 1 (Soundness)

Let Γ *be a set of equations. Suppose that*

$$\Gamma \parallel \Delta \parallel \sigma \Rightarrow^{*}_{\mathfrak{I}_{ACh}} \bigvee_{i \in I} (\Gamma_i \parallel \Delta_i \parallel \sigma_i),$$

where for each i*, there are no applicable rules to* $\Gamma_i \parallel \Delta_i \parallel \sigma_i$ *and let* $\mathcal{S} = \{\sigma_i \mid \Gamma_i = \emptyset\}$.
Then, any element of $\mathcal{S}$ *is an ACh-unifier of* Γ.

Proof: For all $\sigma_i \in \mathcal{S}$, we have that $\sigma_i \vDash \Gamma_i \parallel \Delta_i \parallel \sigma_i$. Hence, by Theorem 1, $\sigma_i \vDash \Gamma \parallel \Delta \parallel \sigma$. Therefore, σ_i is an ACh-unifier of Γ. ■

Table of Contents

Completeness

Considering all the possible forms of Γ (in flattened form), we can prove

Lemma 1

Let $\Gamma \parallel \Delta \parallel \sigma$ be a set triple which is not in solved form, and θ be a substitution such that $\theta \vDash \Gamma \parallel \Delta \parallel \sigma$. Then, there exists an inference

$$\Gamma \parallel \Delta \parallel \sigma \Rightarrow_{\mathfrak{I}_{ACh}} \bigvee_{i \in I} (\Gamma_i \parallel \Delta_i \parallel \sigma_i),$$

an i and θ_0 such that $\mathcal{D}om(\theta_0) \subset \mathcal{V}ar(\Gamma_i) \setminus \mathcal{V}ar(\Gamma)$ and $\theta\theta_0 \vDash \Gamma_i \parallel \Delta_i \parallel \sigma_i$.

Completeness

Since $\mathsf{Unify}_{\mathsf{AC_h}}$ terminates, we obtain the following result by construction:

Theorem 2

Let $\Gamma \parallel \Delta \parallel \sigma$ *be a triple which is not in solved form, and* θ *be a substitution such that* $\theta \vDash \Gamma \parallel \Delta \parallel \sigma$. *Then, there exists a sequence of inferences*

$$\Gamma \parallel \Delta \parallel \sigma \Rightarrow^{+}_{\mathfrak{I}_{ACh}} \bigvee_{i \in I} (\Gamma_i \parallel \Delta_i \parallel \sigma_i),$$

and an i *and* θ_0 *such that* $\theta\theta_0 \vDash \Gamma_i \parallel \Delta_i \parallel \sigma_i$.

Completeness

Corollary 2 (Completeness)

Let $\Gamma \parallel \Delta \parallel \sigma$ *be a triple. Suppose that*

$$\Gamma \parallel \Delta \parallel \sigma \Rightarrow^{*}_{\mathfrak{I}_{ACh}} \bigvee_{i \in I} (\Gamma_i \parallel \Delta_i \parallel \sigma_i),$$

where, for each i*, there are no rules left to be applied. Let* $S = \{\sigma_i \mid \Gamma_i = \emptyset\}$*. Then, for each ACh Unifier* θ *of* Γ*, there exists a* $\sigma_j \in S$*, such that* $\sigma_j \lesssim^{\mathcal{V}ar(\Gamma)}_{ACh} \theta$

Completeness

Proof: Let θ be an ACh Unifier of Γ, then, by Theorem 2 we have that there exist inferences such that

$$\Gamma \parallel \Delta \parallel \sigma \Rightarrow^{*}_{\mathfrak{I}_{ACh}} \bigvee_{i \in I} (\Gamma_i \parallel \Delta_i \parallel \sigma_i),$$

and there exists $j \in I$, and θ_0 such that $\theta\theta_0 \vDash \Gamma_j \parallel \Delta_j \parallel \sigma_j$ and $\Gamma_j = \emptyset$. For any $X \in \mathcal{D}om(\sigma_j)$, consider that $X \mapsto t_X \in \sigma_j$, that is, $X\sigma_j = t_X (*)$. Since $\theta\theta_0 \vDash \sigma_j$, by definition, we have

$$X\theta\theta_0 =_{ACh} t_X\theta\theta_0.$$

By $(*)$, we obtain

$$X\theta\theta_0 =_{ACh} X\sigma_j\theta\theta_0.$$

But we cannot guarantee that it is possible to remove θ_0!!!

Next steps

For future work would like to

- Investigate the imprecision found in the last proof. (Open for suggestions)
- Refine the definition of AC-solved variables.

References

Mauricio Ayala-Rincón, Maribel Fernández, Gabriel Ferreira Silva, and Daniele Nantes Sobrinho.
A certified algorithm for ac-unification.
In Amy P. Felty, editor, *7th FSCD 2022, August 2-5, 2022, Haifa, Israel*, volume 228 of *LIPIcs*, pages 8:1–8:21. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022.

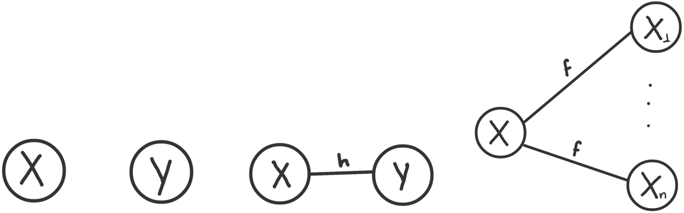
Ajay Kumar Eeralla and Christopher Lynch.
Bounded ach unification.
Math. Struct. Comput. Sci., 30(6):664–682, 2020.

François Fages.
Associative-commutative unification.
J. Symb. Comput., 3(3):257–275, 1987.

Paliath Narendran.
Solving linear equations over polynomial semirings.
In *Proceedings, 11th Annual IEEE Symposium on Logic in Computer Science, New Brunswick, New Jersey, USA, July 27-30, 1996*, pages 466–472. IEEE Computer Society, 1996.

Graph $\mathbb{G}(\Gamma)$

- For convenience, we can assume that our problem is in *flattened form*, that is, every equation is in the form $X \stackrel{?}{=} Y, X \stackrel{?}{=} h(Y), X \stackrel{?}{=} f(Y_1, ..., Y_n)$ or $X \stackrel{?}{=} Y_1 + ... + Y_n$.
- So, given an unification problem Γ in flattened form, we may define the Graph of Γ where
  - each node represents a variable;
  - each edge represents a function symbol;

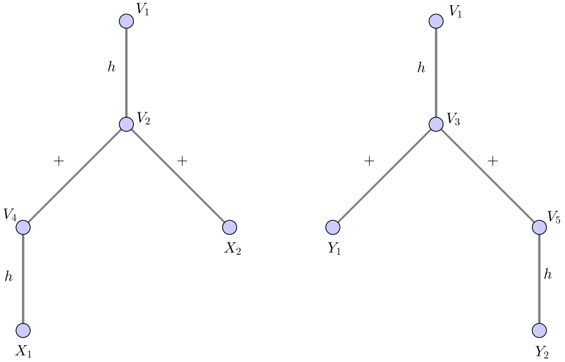


(a) $X =^? Y$ (b) $X =^? h(Y)$ (c) $X =^? f(X_1, \ldots, X_n)$

Graph $\mathbb{G}(\Gamma)$

Consider $\Gamma = \{h(h(X_1) + X_2)) =^? h(Y_1 + h(Y_2))\}$

- In flattened form we have $\Gamma = \{V_1 \stackrel{?}{=} h(V_2), V_1 \stackrel{?}{=} h(V_3), V_2 \stackrel{?}{=} V_4 + X_2, V_3 \stackrel{?}{=} Y_1 + V_5, V_4 \stackrel{?}{=} h(X_1), V_5 \stackrel{?}{=} h(Y_2)\}$;
- So, $\mathbb{G}(\Gamma)$ should be like the figure below



Variable Elimination Rules

Variable Elimination 1 (VE1)

$$\frac{\{X \stackrel{?}{=} Y\} \cup \Gamma || \Delta || \sigma}{\Gamma\{X \mapsto Y\} || \Delta || \sigma\{X \mapsto Y\} \cup \{X \mapsto Y\}} \text{if } X \neq Y$$

Variable Elimination 2 (VE2)

$$\frac{\{X \stackrel{?}{=} t\} \cup \Gamma || \Delta || \sigma}{\Gamma\{X \mapsto t\} || \Delta || \sigma\{X \mapsto Y\} \cup \{X \mapsto t\}} \text{if } X \notin Var(t)$$