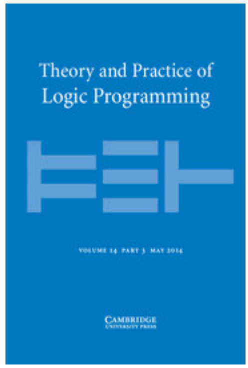


Optimal multibinding unification for sharing and linearity analysis

GIANLUCA AMATO and FRANCESCA SCOZZARI

Theory and Practice of Logic Programming / Volume 14 / Issue 03 / May 2014, pp 379 - 400
DOI: 10.1017/S1471068413000070, Published online: 09 August 2013

TLP **14** (3): 379–400, 2014. © Cambridge University Press 2013
doi:10.1017/S1471068413000070 First published online 9 August 2013

Optimal multibinding unification for sharing and linearity analysis

GIANLUCA AMATO and FRANCESCA SCOZZARI
Dipartimento di Economia, Università di Chieti-Pescara, Pescara, Italy
(*e-mail:* gamato, fscozzari@unich.it)

submitted 28 September 2012; revised 15 March 2013; accepted 31 May 2013

Abstract

In the analysis of logic programs, abstract domains for detecting sharing properties are widely used. Recently, the new domain $\mathtt{ShLin}^{\omega}$ has been introduced to generalize both sharing and linearity information. This domain is endowed with an optimal abstract operator for single-binding unification. The authors claim that the repeated application of this operator is also optimal for multibinding unification. This is the proof of such a claim.

KEYWORDS: static analysis, abstract interpretation, sharing, linearity, unification

1 Introduction

In the abstract interpretation-based static analysis of logic programs, many abstract domains for encoding sharing information have been proposed. For instance, in the original and most studied Sharing domain of Jacobs and Langen (1992), the substitution $\theta = \{x/s(u,v), y/g(u,u,u), z/v\}$ is abstracted into $\{uxy, vxz\}$, where the sharing group uxy means that $\theta(u)$, $\theta(x)$, and $\theta(y)$ share a common variable, namely u.

Since Sharing is not very precise, it is often combined with other domains handling freeness, linearity, groundness or structural information (see Bagnara *et al.* 2005 for a comparative evaluation). In particular, adding some kind of linearity information seems to be very profitable, both for the gain in precision and speed which can be obtained, and for the fact that it can be easily and elegantly embedded inside the sharing groups (see King 1994). For example, if we know that x, y and z do not share, nothing can be said after the unification with $\{z/t(x,y)\}$. However, if we also know that z is linear, then we may conclude that x and y do not share after the unification.

Recently, the new abstract domain $\mathtt{ShLin}^{\omega}$ has been proposed by Amato and Scozzari (2010) as a generalization of Sharing. It is able to encode the *amount* of non-linearity in a substitution, by keeping track of the exact number of occurrences of the same variable in a term. The above substitution θ is abstracted into $\{uxy^3, vxz\}$ by $\mathtt{ShLin}^{\omega}$, with the additional information that the variable u occurs three times in $\theta(y)$. The authors provide a constructive characterization of the optimal abstract unification operator for single-binding substitutions (i.e., substitutions $\{x/t\}$ with a single variable x). Such operator is used to derive optimal (single-binding) abstract unification operators for both the domains $\mathtt{Sharing} \times \mathtt{Lin}$ (Hans and Winkler 1992; Muthukumar and Hermenegildo 1992) and $\mathtt{ShLin}^2$ (King 1994). These were the first optimality results for domains combining aliasing and linearity information.

In the same paper, the authors claim that computing abstract unification over $\mathtt{ShLin}^\omega$ one binding at a time yields the best abstract unification for multibinding substitutions. In this paper, we prove this claim.

For this purpose, we introduce a parallel abstract unification operator, which computes the abstract unification over $\mathtt{ShLin}^\omega$ by considering all the bindings at the same time. We prove that (1) the parallel unification operator and the standard (sequential) one do coincide over $\mathtt{ShLin}^\omega$ and that (2) the parallel unification operator is optimal.

2 Preliminaries

Given a set A, we use $\wp(A)$ for the powerset of A, $\wp_f(A)$ for the set of finite subsets of A and $|A|$ for the cardinality of A. $\mathbb{N}$ is the set of natural numbers with zero.

2.1 Multisets

A *multiset* is a set where repetitions are allowed. We denote by $\{\!\!\{v_1, \dots, v_m\}\!\!\}$ a multiset, where $v_1, \dots, v_m$ is a sequence with (possible) repetitions. We denote by $\{\!\!\{\}\!\!\}$ the empty multiset. We will often use the polynomial notation $v_1^{i_1} \dots v_n^{i_n}$, where $v_1, \dots, v_n$ is a sequence without repetitions, to denote a multiset A whose element v_j appears i_j times. The set $\{v_j \mid i_j > 0\}$ is called the *support* of A and is denoted by $\llbracket A \rrbracket$. We also use the functional notation $A : \{v_1, \dots, v_n\} \to \mathbb{N}$, where $A(v_j) = i_j$.

In this paper, we only consider multisets whose support is *finite*. We denote with $\wp_m(X)$ the set of all the multisets whose support is *any finite subset* of X. For example, both a^2c^4 and $a^1b^2c^3$ are elements of $\wp_m(\{a, b, c\})$.

The new fundamental operation for multisets is the *sum*, defined as

$$A \uplus B = \lambda v \in \llbracket A \rrbracket \cup \llbracket B \rrbracket . A(v) + B(v). \tag{1}$$

For instance, the sum of a^2c^4 and $a^1b^2c^3$ is $a^3b^2c^7$. Given a multiset A and $X \subseteq \llbracket A \rrbracket$, the *restriction* of A over X, denoted by $A|_X$, is the only multiset B such that $\llbracket B \rrbracket = X$ and $B(v) = A(v)$ for each $v \in X$.

2.2 Multigraphs

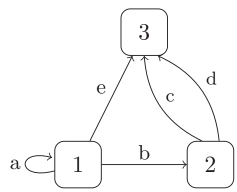
We call (directed) *multigraph* a graph where multiple distinguished edges are allowed between nodes. We use the definition of multigraph which is customary in category theory (Mac Lane 1971).

Definition 2.1 (*Multigraph*)
A *multigraph* G is a tuple $\langle N_G, E_G, \mathrm{src}_G, \mathrm{tgt}_G \rangle$, where $N_G \neq \emptyset$ and E_G are the sets of *nodes* and *edges* respectively, $\mathrm{src}_G : E_G \to N_G$ is the *source* function which maps each edge to its starting node, and $\mathrm{tgt}_G : E_G \to N_G$ is the *target* function which maps each edge to its ending node.

We write $e : n_1 \to n_2 \in G$ to denote an edge $e \in E_G$ such that $\mathrm{src}_G(e) = n_1$ and $\mathrm{tgt}_G(e) = n_2$. We call *in-degree* (respectively *out-degree*) of a node n the cardinality of the set $\{e \in E_G \mid \mathrm{tgt}_G(e) = n\}$ (respectively $\{e \in E_G \mid \mathrm{src}_G(e) = n\}$).

Given a multigraph G, a *path* $\pi : n_1 \to n_k$ is a *non-empty* sequence of nodes $n_1 \dots n_k$ such that, for each $i \in \{1, \dots, k-1\}$, there is either an edge $n_i \to n_{i+1} \in G$ or an edge $n_{i+1} \to n_i \in G$. Nodes n_1 and n_k are the *endpoints* of π, and we say that π *connects* n_1 and n_k. A multigraph is *connected* when all pairs of nodes are connected by at least one path.

Example 2.2
Consider the multigraph G such that $N_G = \{1, 2, 3\}$, $E_G = \{a, b, c, d, e\}$, $\mathrm{src}_G = \{a \mapsto 1, b \mapsto 1, c \mapsto 2, d \mapsto 2, e \mapsto 1\}$ and $\mathrm{tgt}_G = \{a \mapsto 1, b \mapsto 2, c \mapsto 3, d \mapsto 3, e \mapsto 3\}$. It may be depicted as follows:



Note that the edges c and d have the same starting and ending nodes, but different names. According to our definition, the graph is connected.

2.3 Abstract interpretation

Given two sets C and A of concrete and abstract objects respectively, an *abstract interpretation* (Cousot and Cousot 1992) is given by an approximation relation $\rhd \subseteq A \times C$. When $a \rhd c$ holds, this means that a is a correct abstraction of c. We work in a framework where (1) A is a complete lattice, (2) $a \rhd c$ and $a \leqslant a'$ imply $a' \rhd c$, (3) each c has a least correct abstraction in A given by $\alpha(c)$.

Given a function $f : C \to C$, we say that $\tilde{f} : A \to A$ is a correct abstraction of f, and we write $\tilde{f} \rhd f$, when

$$a \rhd c \Rightarrow \tilde{f}(a) \rhd f(c).$$

We say that $\tilde{f} : A \to A$ is the *optimal* abstraction of f when it is correct and, for each $f' : A \to A$,

$$f' \rhd f \Rightarrow \tilde{f} \leqslant f'$$

with the standard pointwise ordering.

2.4 Terms and substitutions

In the following, we fix a first-order signature and a denumerable set of variables $\mathcal{V}$. Given a term or other syntactic object o, we denote by $vars(o)$ the set of variables occurring in o and by $occ(v, o)$ the number of occurrences of v in o. When it does not cause ambiguities, we abuse the notation and prefer to use o itself in the place of $vars(o)$. For example, if t is a term and $x \in \mathcal{V}$, then $x \in t$ should be read as $x \in vars(t)$.

We denote by ϵ the empty substitution, by $\{x_1/t_1, \dots, x_p/t_p\}$ a substitution θ with $\theta(x_i) = t_i \neq x_i$, and by $\mathrm{dom}(\theta) = \{x \in \mathcal{V} \mid \theta(x) \neq x\}$ and $\mathrm{rng}(\theta) = \cup_{x \in \mathrm{dom}(\theta)} vars(\theta(x))$ the domain and range of θ respectively. Let $vars(\theta)$ be the set $\mathrm{dom}(\theta) \cup \mathrm{rng}(\theta)$, and given $U \in \wp_f(\mathcal{V})$, let $\theta|_U$ be the projection of θ over U, i.e., the unique substitution such that $\theta|_U(x) = \theta(x)$ if $x \in U$ and $\theta|_U(x) = x$ otherwise. Given θ_1 and θ_2 two substitutions with disjoint domains, we denote by $\theta_1 \cup \theta_2$ the substitution θ such that $\mathrm{dom}(\theta) = \mathrm{dom}(\theta_1) \cup \mathrm{dom}(\theta_2)$ and $\theta(x) = \theta_i(x)$ if $x \in \mathrm{dom}(\theta_i)$, for each $i \in \{1, 2\}$. The application of a substitution θ to a term t is written as $t\theta$ or $\theta(t)$. Given two substitutions θ and δ, their composition, denoted by $\theta \circ \delta$, is given by $(\theta \circ \delta)(x) = \delta(\theta(x))$. A substitution θ is idempotent when $\theta \circ \theta = \theta$ or, equivalently, when $\mathrm{dom}(\theta) \cap \mathrm{rng}(\theta) = \emptyset$. A substitution ρ is called renaming if it is a bijection from $\mathcal{V}$ to $\mathcal{V}$ (this is equivalent to saying that there exists a substitution ρ^{-1} such that $\rho \circ \rho^{-1} = \rho^{-1} \circ \rho = \epsilon$). The sets of idempotent substitutions and renamings are denoted by *ISubst* and *Ren*, respectively. Given a set of equations E, we write $\theta = \mathrm{mgu}(E)$ to denote that θ is a most general unifier of E. Conversely, $\mathrm{Eq}(\theta) = \{x = \theta(x) \mid x \in \mathrm{dom}(\theta)\}$.

A *position* is a sequence of positive natural numbers. Given a term t and a position ξ, we define $t(\xi)$ inductively as follows:

$$t(\epsilon) = t \qquad \text{(where } \epsilon \text{ denotes the empty sequence)}$$

$$t(i \cdot \xi') = \begin{cases} t_i(\xi') & \text{if } t \text{ is } s(t_1, \dots, t_p) \text{ and } i \leqslant p; \\ \text{undefined} & \text{otherwise.} \end{cases} \tag{2}$$

For any variable x, an *occurrence* of x in t is a position ξ such that $t(\xi) = x$.

In the rest of the paper, we use U, V, W to denote finite sets of variables; u, v, w, x, y, z for variables; r, s for term symbols; t for terms; $\beta, \eta, \theta, \delta$ for substitutions; ρ for renamings.

2.5 Existential substitutions

The denotational semantics of logic programs is not generally interested in substitutions, but in appropriate equivalence classes which abstract away from the particular renaming of clauses used during selective linear definite-clause (SLD) derivations. Among the many choices available in the literature (e.g. Jacobs and Langen 1992; Marriott *et al.* 1994; Levi and Spoto 2003), we adopt the domain of existential substitutions (Amato and Scozzari 2009).

Given $\theta_1, \theta_2 \in ISubst$ and $U \in \wp_f(\mathcal{V})$, consider the equivalence relation $\sim_U$ given by

$$\theta_1 \sim_U \theta_2 \iff \exists \rho \in Ren. \forall v \in U.\ \theta_1(v) = \rho(\theta_2(v)), \tag{3}$$

and let $ISubst_{\sim_U}$ be the quotient set of $ISubst$ w.r.t. $\sim_U$. The domain $ISubst_\sim$ of *existential substitutions* is defined as the union of all the $ISubst_{\sim_U}$ for $U \in \wp_f(\mathcal{V})$, namely

$$ISubst_\sim = \bigcup_{U \in \wp_f(\mathcal{V})} ISubst_{\sim_U}. \tag{4}$$

In the following, we write $[\theta]_U$ for the equivalence class of θ w.r.t. $\sim_U$. To ease the notation, we often omit braces from the sets of variables of interest when they are given extensionally. So we write $[\theta]_{x,y}$ instead of $[\theta]_{\{x,y\}}$.

Given $U \in \wp_f(\mathcal{V})$, $[\delta]_U \in ISubst_\sim$ and $\theta \in ISubst$, the most general unifier of θ and $[\delta]_U$ may be obtained from the mgu of θ and a suitably chosen representative for δ, where variables not of interest are renamed apart. In formulas:

$$\mathrm{mgu}([\delta]_U, \theta) = [\mathrm{mgu}(\delta', \theta)]_{U \cup vars(\theta)}, \tag{5}$$

where $\delta \sim_U \delta' \in ISubst$ and $vars(\delta') \cap vars(\theta) \subseteq U$.

2.6 *The domain* $\mathtt{ShLin}^\omega$

The domain $\mathtt{ShLin}^\omega$ (Amato and Scozzari 2010) generalizes $\mathtt{Sharing}$ by recording multiplicity of variables in sharing groups. We will call a multiset of variables (an element of $\wp_m(\mathcal{V})$) an *ω-sharing group*. Given a substitution θ and a variable $v \in \mathcal{V}$, we denote by $\theta^{-1}(v)$ the ω-sharing group $\lambda w \in \mathcal{V}.occ(v, \theta(w))$, which maps each variable w to the number of occurrences of v in $\theta(w)$.

Given a set of variables U and a set of ω-sharing groups $S \subseteq \wp_m(U)$, we say that $[S]_U$ *correctly approximates* a substitution $[\theta]_W$ if $U = W$ and for each $v \in \mathcal{V}$, $\theta^{-1}(v)|_U \in S$. We write $[S]_U \rhd [\theta]_W$ to mean that $[S]_U$ correctly approximates $[\theta]_W$. Therefore, $[S]_U \rhd [\theta]_U$ when S contains all the ω-sharing groups in θ, restricted to the variables in U.

Definition 2.3 ($\mathtt{ShLin}^\omega$)
The domain $\mathtt{ShLin}^\omega$ is defined as

$$\mathtt{ShLin}^\omega = \{[S]_U \mid U \in \wp_f(\mathcal{V}), S \subseteq \wp_m(U), S \neq \emptyset \Rightarrow \{\!\!\{\}\!\!\} \in S\}, \tag{6}$$

and ordered by $[S_1]_{U_1} \leqslant_\omega [S_2]_{U_2}$ iff $U_1 = U_2$ and $S_1 \subseteq S_2$.

In order to ease the notation, we write $[\{\{\!\!\{\}\!\!\}, B_1, \ldots, B_n\}]_U$ as $[B_1, \ldots, B_n]_U$ by omitting the braces and the empty multiset. Moreover, if $X \in \mathtt{ShLin}^\omega$, we write $B \in X$ in place of $X = [S]_U \wedge B \in S$. The best correct abstraction of a substitution $[\theta]_U$ is

$$\alpha_\omega([\theta]_U) = [\{\theta^{-1}(v)|_U \mid v \in \mathcal{V}\}]_U. \tag{7}$$

Example 2.4
Given $\theta = \{x/s(y, u, y), z/s(u, u), v/u\}$ and $U = \{w, x, y, z\}$, we have $\theta^{-1}(u) = uvxz^2$, $\theta^{-1}(y) = x^2y$, $\theta^{-1}(z) = \theta^{-1}(v) = \theta^{-1}(x) = \{\!\!\{\}\!\!\}$ and $\theta^{-1}(v) = v$ for all the other variables (included w). Projecting over U we obtain $\alpha_\omega([\theta]_U) = [xz^2, x^2y, w]_U$.

Definition 2.5 (*Multiplicity of ω-sharing groups*)
The *multiplicity* of an ω-sharing group B in a term t is defined as:

$$\chi(B,t) = \sum_{v \in \llbracket B \rrbracket} B(v) \cdot occ(v,t). \tag{8}$$

For instance, $\chi(w^2x^3yz^4, r(x,y,s(x,y,z),v)) = 2 \cdot 0 + 3 \cdot 2 + 1 \cdot 2 + 4 \cdot 1 = 12$.

3 Parallel abstract unification

We want to find the optimal abstract operator in $\mathtt{ShLin}^{\omega}$ corresponding to unification. Amato and Scozzari (2010) define the operator mgu_{ω}, which is optimal for single-binding substitutions. The cornerstone of their abstract unification is the concept of *sharing graph* which plays the same role of alternating paths (Søndergaard 1986; King 2000) for pair sharing analysis. The authors claim that, by applying mgu_{ω} one binding at a time, we get an optimal operator for multibinding substitutions.

Here, in order to prove this claim, we proceed along these steps:

1. We define a new operator $\mathrm{mgu}_{\mathrm{p}}$ which computes the abstract unification with a multibinding substitution in one step. This is based on a generalization of the concept of a sharing graph with multiple layers. For this reason, we speak of *parallel sharing graph* and *parallel abstract unification*;
2. We prove that parallel abstract unification ($\mathrm{mgu}_{\mathrm{p}}$) is actually the same as the *sequential abstract unification* (mgu_{ω});
3. We prove that parallel abstract unification is optimal with respect to concrete unification.

If $[S]_U \rhd [\delta]_U$ and we unify $[\delta]_U$ with θ, some of the ω-sharing groups in S may be glued together to obtain a larger resultant group. It happens that the gluing may be represented by special families of labeled multigraphs which we call *parallel sharing graphs.*

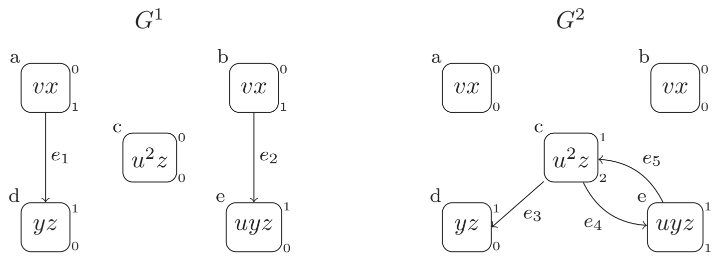
Definition 3.1 (*Parallel sharing graph*)
A *parallel sharing graph* for a set of ω-sharing groups S and the idempotent substitution $\theta = \{x_1/t_1, \ldots, x_p/t_p\}$ is a family $\mathscr{G} = \{G^i\}_{i\in[1,p]}$ of multigraphs over the same set of nodes $N_{\mathscr{G}}$, equipped with a labeling function $l_{\mathscr{G}} : N_{\mathscr{G}} \to S$, such that

- for each node $n \in N_{\mathscr{G}}$ and each $i \in [1,p]$, the out-degree of n in G^i is equal to $\chi(l_{\mathscr{G}}(n), x_i)$ and the in-degree of n in G^i is equal to $\chi(l_{\mathscr{G}}(n), t_i)$;
- the sets of edges E_{G^i} are all pairwise disjoint;
- $\underline{\mathscr{G}}$ (the *flattening* of $\mathscr{G}$) is connected.

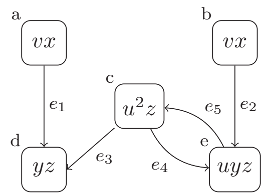
In the last condition, $\underline{\mathscr{G}}$ is defined as the multigraph $\langle N_{\mathscr{G}}, E, \mathrm{src}_{\mathscr{G}}, \mathrm{tgt}_{\mathscr{G}} \rangle$ where $E = \cup_{i\in[1,p]} E_{G^i}$ and $\mathrm{src}_{\mathscr{G}} : E \to N_{\mathscr{G}}$ maps $x \in E_{G^i}$ to $\mathrm{src}_{G^i}(e)$ ($\mathrm{tgt}_{\mathscr{G}}$ is defined analogously). Each of the G^i's which make up $\mathscr{G}$ is called a *layer* of the sharing graph.

Since in this paper we only use parallel sharing graphs, in the following we will call them just *sharing graphs.*

Example 3.2
Let $S = \{u^2z, uyz, vx, yz\}$ and $\theta = \{x/y, u/r(z)\}$. Consider the sharing graph $\mathscr{G} = \{G^1, G^2\}$ over the set of nodes $N_{\mathscr{G}} = \{a, b, c, d, e\}$ labeled by $l_{\mathscr{G}} = \{a \mapsto vx, b \mapsto vx, c \mapsto u^2z, d \mapsto yz, e \mapsto uyz\}$:



The left layer (G^1) is for the binding x/y, while the right one (G^2) is for the binding $u/r(z)$. Each node is annotated with its name, label, in- and out-degree. Its flattening is the following connected multigraph:



Let us motivate the three conditions of Definition 3.1. If $[S]_U \rhd [\delta]_U$ and we compute $\mathrm{mgu}([\delta]_U, \theta)$, each G^i represents a possible way the sharing groups in δ may be joined together as a result of binding x_i/t_i, that is unifying $\delta(x_i)$ and $\delta(t_i)$. We may restrict our attention to the case when, as a result of the unification, variables are only bound to other variables, not to composed terms. In other words, we assume that, for each position ξ, the term $\delta(x_i)(\xi)$ is a variable iff $\delta(t_i)(\xi)$ is a variable. Each node in the sharing graph represents a variable w_n such that $\delta^{-1}(w_n)|_U$ is the node label. Each edge $e : n_1 \to n_2$ in G^i represents a position ξ such that $\delta(x_i)(\xi) = w_{n_1}$ and $\delta(t_i)(\xi) = w_{n_2}$. The result is that the variables w_{n_1} and w_{n_2} are aliased; hence, the ω-sharing groups $l_{\mathscr{G}}(n_1)$ and $l_{\mathscr{G}}(n_2)$ are joined together.

According to this correspondence, the number of edges departing from n should be equal to the number of occurrences of w_n in $\delta(x_i)$, that is $\chi(l_{\mathscr{G}}(n), x_i)$. Analogously for the in-degree of nodes. This justifies the first condition in the definition.

The second condition ensures that, in the flattening, no edges share the same identifier and therefore $\mathrm{src}_{\mathscr{G}}$ and $\mathrm{tgt}_{\mathscr{G}}$ are well defined. Remember that, since an edge is just an element of a set with associated source and target nodes, this does not preclude the possibility to have different edges with the same source and target nodes.

Finally, the third condition is needed since we want each sharing graph to represent a single non-empty sharing group. If the flattening were not connected, some pairs of variables would not be aliased, and the result of the unification of θ with δ would contain more than one non-empty sharing group.

Example 3.3
Consider the sharing graph in Example 3.2. Let us associate with each node n the variable w_n and consider the substitution

$$\delta = \{u/r(s(w_c, w_c, w_e)), v/s(w_a, w_b), x/s(w_a, w_b), y/s(w_d, w_e), z/s(w_d, w_e, w_c)\}.$$

This substitution is built according to the variables that appear in the nodes. For instance, the first binding $u/r(s(w_c, w_c, w_e))$ suggests that the variable u appears in the nodes c (twice) and e.

We now want to unify δ with θ. The first binding x/y in θ unifies $\delta(x) = s(w_a, w_b)$ with $\delta(y) = s(w_d, w_e)$. This causes variables (w_a, w_d) and (w_b, w_e) to be aliased, exactly as described by the arrows e_1 and e_2 in the left graph. The second binding $u/r(z)$ unifies $\delta(u) = r(s(w_c, w_c, w_e))$ with $\delta(r(z)) = r(s(w_d, w_e, w_c))$, which yields the aliasing of the pairs (w_c, w_d), (w_c, w_e) and (w_e, w_c), as described by the arrows e_3, e_4 and e_5. By transitivity, all pairs of variables are aliased.

Definition 3.4 (*Resultant ω-sharing group*)
The *resultant ω-sharing group* of the sharing graph $\mathscr{G}$ is

$$res(\mathscr{G}) = \biguplus_{s \in N_{\mathscr{G}}} l_{\mathscr{G}}(s). \tag{9}$$

Example 3.5
Consider again the sharing graph in Example 3.2. The resultant sharing group is $u^3v^2x^2y^2z^3$. This is exactly the only non-empty sharing group in $\alpha_\omega([\eta]_U)$ where $U = vars(S)$ and

$$\begin{aligned}\eta = \mathrm{mgu}(\delta, \theta) = \{&u/r(s(w_a, w_a, w_a)), v/s(w_a, w_a), x/s(w_a, w_a),\\ &\times y/s(w_a, w_a), z/s(w_a, w_a, w_a), w_b/w_a, w_c/w_a, w_d/w_a, w_e/w_a\}.\end{aligned}$$

Definition 3.6 (*Parallel abstract mgu*)
Given a set of ω-sharing groups S and an idempotent substitution θ, the *abstract parallel unification* of S and θ is given by

$$\mathrm{mgu_p}(S, \theta) = \{res(\mathscr{G}) \mid \mathscr{G} \text{ is a sharing graph for } S \text{ and } \theta\}. \tag{10}$$

This is lifted to the domain $\texttt{ShLin}^\omega$:

$$\mathrm{mgu_p}([S]_U, \theta) = [\mathrm{mgu_p}(S \cup \{\{\!\{v\}\!\} \mid v \in vars(\theta) \setminus U\}, \theta)]_{U \cup vars(\theta)}. \tag{11}$$

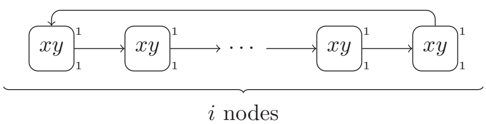
It is worth noting that, given any set of ω-sharing groups S and substitution θ, there exist many different sharing graphs for S and θ. Each sharing graph yields a resultant sharing group which must be included in the result of the abstract unification operator. Of course, different sharing graphs may give the same resultant sharing group. The abstract unification operator is defined by collecting all the resultant sharing groups.

Example 3.7
We show another sharing graph for the same S and θ of Example 3.2. We omit from the picture the names of edges and nodes, since they are not relevant here:

The resultant sharing group is u^2vxyz.

It is worth noting that the domain $\mathtt{ShLin}^\omega$ is not amenable to a direct implementation. Actually, it may be the case that even the mgu of a finite set of ω-sharing groups with a single-binding substitution generates an infinite set of ω-sharing groups (see Example 3.8 later). However, it is an invaluable theoretical device to study the abstract operators for its abstractions, such as $\mathtt{Sharing} \times \mathtt{Lin}$ and $\mathtt{ShLin}^2$.

Example 3.8
It holds that $\mathrm{mgu_p}(\{xy\}, \{x/y\}) = \{\{\!\!\}\} \cup \{x^i y^i \mid i \geqslant 1\}$. Actually, for each $i \geqslant 1$, the following is a single-layer sharing graph:



3.1 *Coincidence of parallel and sequential abstract unification*

For concrete substitutions, unification may be performed one binding at a time. On an abstract domain, computing one binding at a time generally incurs in a loss of precision. However, there are well-known domains when this does not happen, such as $\mathtt{Def}$ (Armstrong *et al.* 1994) and $\mathtt{Sharing}$. We will show that computing one binding at a time does not cause loss of precision on the abstract domain $\mathtt{ShLin}^\omega$.

Definition 3.9 (*Abstract sequential unification*)
Given a set of ω-sharing groups S and an idempotent substitution θ, the *abstract sequential unification* of S and θ, denoted by $\mathrm{mgu}_\omega(S, \theta)$, is given by:

$$\begin{aligned} \mathrm{mgu}_\omega(S, \epsilon) &= S \\ \mathrm{mgu}_\omega(S, \{x/t\} \cup \theta)) &= \mathrm{mgu}_\omega(\mathrm{mgu_p}(S, \{x/t\}), \theta) \end{aligned} \tag{12}$$

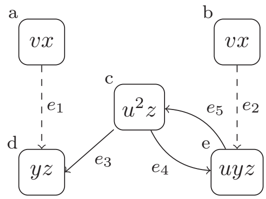
The definition may be lifted to the domain $\mathtt{ShLin}^\omega$ as for $\mathrm{mgu_p}$. It is immediate to check that $\mathrm{mgu_p}$ and mgu_ω are equivalent for single-binding substitutions. We will prove that this holds for any substitution.

In Amato and Scozzari (2010), the abstract sequential unification mgu_ω has been introduced starting from the definition of a sharing graph for single-binding unification. This is essentially a sharing graph with a single layer. Hence, it is immediate to check that the definition of mgu_ω given above is the same as the definition of mgu_ω given by Amato and Scozzari (2010).

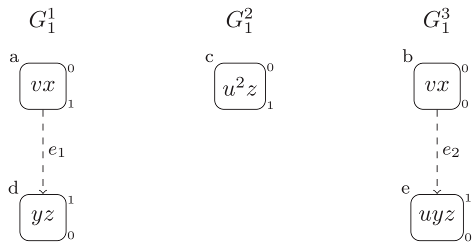
Before introducing the formal proof of coincidence between sequential and parallel abstract unification, we try to convey the intuitive idea behind it with an example.

Example 3.10

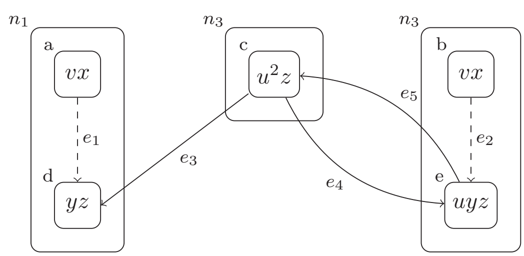
Consider again the sharing graph $\mathcal{G}$ given in Example 3.2 for $S = \{u^2z, uyz, xv, yz\}$ and $\theta = \{x/y, u/r(z)\}$. For the sake of conciseness, we can draw $\mathcal{G}$ with a single picture, omitting the in- and out-degree annotations on the nodes, and with the edges in different styles, according to the layers they come from:



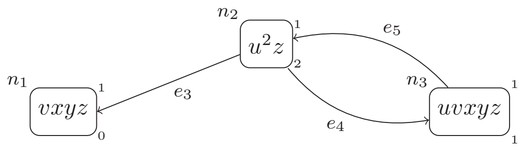
As we said before, the resultant sharing group of $\mathcal{G}$ is $u^3v^2x^2y^2z^3$. The same sharing group may be obtained by first computing $S' = \mathrm{mgu}_{\mathrm{p}}(S, \{x/y\})$ and later $\mathrm{mgu}_{\mathrm{p}}(S', \{u/r(z)\})$. Consider the three connected components in the multigraph G^1, corresponding to the dashed arrows:



Each of them alone may be viewed as a sharing graph with a single layer for the substitution $\{x/y\}$. Therefore, $vxyz$, u^2z and $uvxyz$ are elements of S'. Now, in the original sharing graph, we collapse these connected components:



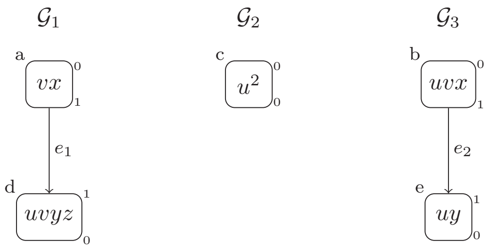
and we get



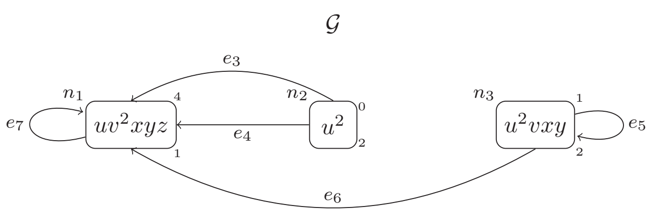
which is a sharing graph for S' and $\{u/r(z)\}$. Note that, in this new sharing graph, the nodes correspond to the connected components of G^1 and the edges are the same as in the original G^2, but with different source and target. The edge e_3 from c to d is now an edge from n_2 to n_1, since d is in the first connected component and c in the second one. We obtain, as expected, that $u^3v^2x^2y^2z^3 \in \mathrm{mgu}_\omega(S', \{u/r(z)\})$.

Example 3.11

We now show an example of the converse, i.e., how to move from sequential to parallel unification. Assume $S = \{vx, u^2, uvx, uvyz, uy\}$ and $\theta = \{x/y, u/s(z, s(z, v))\}$. The following are single-layer sharing graphs for S and $\{x/y\}$:

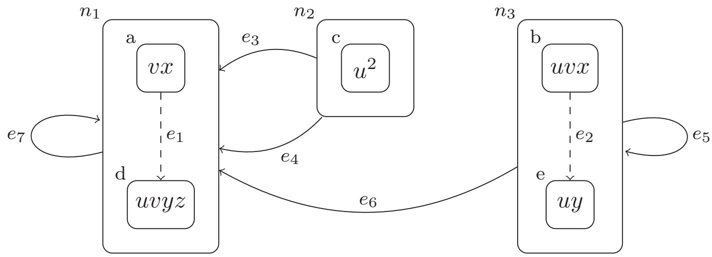


Note that we have chosen disjoint sets of nodes $N_{\mathscr{G}_1} = \{a, d\}$, $N_{\mathscr{G}_2} = \{c\}$ and $N_{\mathscr{G}_3} = \{b, e\}$, and disjoint sets of edges $E_{\mathscr{G}_1} = \{e_1\}$, $E_{\mathscr{G}_2} = \{\}$ and $E_{\mathscr{G}_3} = \{e_2\}$. By definition, the corresponding resultant sharing groups, i.e., uv^2xyz, u^2 and u^2vxy are elements of $S' = \mathrm{mgu}_\omega(S, \{x/y\})$. Now consider the following sharing graph $\mathscr{G}$ for S' and the binding $u/s(z, s(z, v))$:

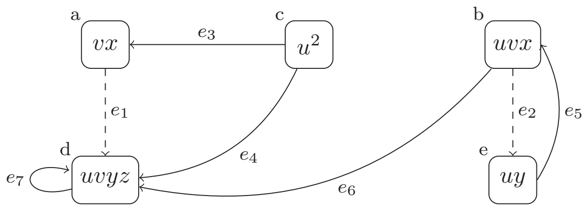


(13)

We need to build a sharing graph for S and $\{x/y, u/s(z, s(z, v))\}$ from these pieces. The idea is to replace, in the graph $\mathscr{G}$, the nodes n_1, n_2 and n_3 with the graphs $\mathscr{G}_1$, $\mathscr{G}_2$ and $\mathscr{G}_3$, respectively:



For each edge in $\mathscr{G}$ we need to specify its target and source as a node in $\{a, b, c, d, e\}$, since giving only the connected component is not enough. For example, the target of e_1 should be either a or d. We may choose the targets freely, subject to the conditions on the in-/out- degree of nodes. Since $\chi(vx, s(z, s(z, v))) = 1$ and $\chi(uvyz, s(z, s(z, v))) = 3$, among e_3, e_4, e_6 and e_7, three edges should be targeted at $uvyz$ and one should be targeted at vx. Among the many others, this is a possible sharing graph, where the different layers are depicted through different line styles:



Note that the self-loop on the node n_3 has become an edge from uy to uvx.

The ideas presented in the previous examples are formalized in the following result.

Lemma 3.12
Given a set of ω-sharing groups S and an idempotent substitution θ, we have that $\mathrm{mgu_p}(S, \{x_1/t_1\} \cup \theta) = \mathrm{mgu_p}(\mathrm{mgu_p}(S, \{x_1/t_1\}), \theta)$.

Proof
If $\theta = \epsilon$, the result easily follows since $\mathrm{mgu_p}(S, \epsilon) = S$. In the case $\theta \neq \epsilon$, we separately prove the two sides of the equality.

First part: $\subseteq$ inclusion. Let $B \in \mathrm{mgu_p}(S, \{x_1/t_1\} \cup \theta)$. We want to prove that $B \in \mathrm{mgu_p}(\mathrm{mgu_p}(S, \{x_1/t_1\}), \theta)$. To this aim, we will provide a sharing graph $\mathscr{G}'$ for $\mathrm{mgu_p}(S, \{x_1/t_1\})$ and θ such that $res(\mathscr{G}') = B$.

Let $\theta = \{x_2/t_2, \ldots, x_p/t_p\}$. By definition, there exists a sharing graph $\mathscr{G} = \{G^i\}_{i \in [1,p]}$ such that $B = res(\mathscr{G})$. We decompose G^1 into its connected components $G^1_1, \ldots, G^1_k$. Note that each G^1_j, labeled with the obvious restriction of $l_{\mathscr{G}}$, is a sharing graph for S and $\{x_1/t_1\}$, therefore $res(G^1_j) \in \mathrm{mgu_p}(S, \{x_1/t_1\})$.

We now show a sharing graph $\mathscr{G}'$ for $\mathrm{mgu_p}(S, \{x_1/t_1\})$ and θ and prove that $res(\mathscr{G}') = B$. For any $i \in [2, p]$, let G_i be the multigraph obtained from G^i by collapsing each of the connected components $G^1_1, \ldots, G^1_k$ to a single node. Formally:

- $N_{G_i} = \{1, \ldots, k\}$;

- $E_{G_i} = E_{G^i}$;
- $\mathrm{src}_{G_i}(e) = j$ iff $\mathrm{src}_{G^i}(e) \in G^1_j$;
- symmetrically for tgt_{G_i}.

We want to prove that $\mathscr{G}' = \{G_i\}_{i\in[2,p]}$, endowed with the labeling function $l_{\mathscr{G}'}(j) = res(G^1_j)$, is a sharing graph for $\mathrm{mgu_p}(S, \{x_1/t_1\})$ and θ. By definition of sharing graph, we need to check that, first, the conditions on the out-degree and the in-degree hold for each node; second, the sets of edges are pairwise disjoint; third, the flattening is connected.

First condition. We now show that the conditions on the out-degree and the in-degree of the nodes hold. Given any node $j \in [1, k]$ we have that the out-degree of j in G_i is

$$
\begin{aligned}
&|\{e \in E_{G_i} \mid \mathrm{src}_{G_i}(e) = j\}| = |\{e \in E_{G^i} \mid \mathrm{src}_{G^i}(e) \in G^1_j\}| \\
&\quad = \sum_{n\in N_{G^1_j}} |\{e \in E_{G^i} \mid \mathrm{src}_{G^i}(e) = n\}| = \sum_{n\in N_{G^1_j}} \chi(l_{\mathscr{G}}(n), x_i) \\
&\quad = \sum_{n\in N_{G^1_j}} \sum_{v\in\mathscr{V}} l_{\mathscr{G}}(n)(v) \cdot occ(v, x_i) = \sum_{v\in\mathscr{V}} \sum_{n\in N_{G^1_j}} l_{\mathscr{G}}(n)(v) \cdot occ(v, x_i) \\
&\quad = \sum_{v\in\mathscr{V}} res(G^1_j)(v) \cdot occ(v, x_i) = \sum_{v\in\mathscr{V}} l_{\mathscr{G}'}(j)(v) \cdot occ(v, x_i) \\
&\quad = \chi(l_{\mathscr{G}'}(j), x_i).
\end{aligned}
$$

Symmetrically, we have that the in-degree of j in G_i is $\chi(l_{\mathscr{G}'}(j), t_i)$.

Second condition. It is immediate to check that the sets of edges E_{G_i} are pairwise disjoint.

Third condition. We prove that $\underline{\mathscr{G}'}$ is connected. Assume that we want to find a path from i to j. Since $\underline{\mathscr{G}}$ is connected, there is a path π from some $n_1 \in N_{G^1_i}$ to some $n_2 \in N_{G^1_j}$. A path from i to j may be obtained in two steps:

1. by replacing each node n in π with $\bar{n}$, where $\bar{n}$ is the unique $m \in [1, k]$ such that $m \in N_{G^1_m}$;
2. by replacing each subsequence $\bar{n}\bar{n}$ with a single node $\bar{n}$. Such a situation may arise when π contains the subsequence nm with $n \to m \in G^1_q$ for some q. The corresponding edge $q \to q$ may not exist in $\mathscr{G}'$, but being a self-loop it may be deleted.

Finally, we need to show that $res(\mathscr{G}') = B$. It is easy to check that $res(\mathscr{G}') = \biguplus_{i\in[1,k]} l_{\mathscr{G}'}(i) = \biguplus_{i\in[1,k]} res(G^1_i) = \biguplus_{i\in[1,k]} \biguplus_{n\in N_{G^1_i}} l_{\mathscr{G}}(n) = \biguplus_{n\in N_{\mathscr{G}}} l_{\mathscr{G}}(n) = res(\mathscr{G})$.

Second part: $\supseteq$ inclusion. Let $S' = \mathrm{mgu_p}(S, \{x_1/t_1\})$ and $B \in \mathrm{mgu_p}(S', \theta)$, where $\theta = \{x_2/t_2, \ldots, x_p/t_p\}$. We show that there exists a sharing graph $\mathscr{G}'$ for S and $\{x_1/t_1, \ldots, x_p/t_p\}$ such that $res(\mathscr{G}') = B$.

By definition, there is a sharing graph $\mathscr{G} = \{G^i\}_{i\in[2,p]}$ for S' and θ such that $res(\mathscr{G}) = B$. Since $S' = \mathrm{mgu_p}(S, \{x_1/t_1\})$, for each node $k \in N_{\mathscr{G}}$ we have a sharing graph G_k such that $res(G_k) = l_{\mathscr{G}}(k)$. Without loss of generality, we may choose these graphs in such a way that the sets N_{G_k} are pairwise disjoint and disjoint from $N_{\mathscr{G}}$.

For each multigraph G^i, with $i \in [2, p]$, we build a new multigraph $\bar{G}^i$ obtained by replacing each node k in G^i with the set of nodes of the generating graph G_k. Then, we pack the $\bar{G}^i$'s and G_k's into a sharing graph $\mathscr{G}$. Formally, $\mathscr{G} = \{\bar{G}^i\}_{i\in[1,p]}$ such that:

- $N_{\mathscr{G}} = \bigcup_{k \in N_{\mathscr{G}}} N_{G_k}$;
- $\bar{G}^1$ is the union of the graphs G_k;
- for $i \in [2, p]$, $E_{\bar{G}^i} = E_{G^i}$;
- for $i \in [2, p]$, $\mathrm{src}_{\bar{G}^i}$ is chosen freely, subject to the following conditions:
  - — if $\mathrm{src}_{G^i}(e) = k$ then $\mathrm{src}_{\bar{G}^i}(e)$ is a node in N_{G_k};
  - — the out-degree of each node n in $\bar{G}^i$ is $\chi(l(n), x_i)$.

  This is always possible since $l_{\mathscr{G}}(k) = res(G_k) = \biguplus_{n \in N_{G_k}} l_{G_k}(n)$ and therefore $\chi(l_{\mathscr{G}}(k), x_i) = \sum_{n \in N_{G_k}} \chi(l_{G_k}(n), x_i)$. Symmetrically for $\mathrm{tgt}_{\bar{G}^i}$.
- the labeling function $l : N_{\mathscr{G}} \to \wp_m(\mathscr{V})$ is the disjoint union of all l_{G_k}. Namely, $l(n) = l_{G_k}(n)$ iff $n \in N_{G_k}$.

We now want to prove that $\mathscr{G}'$ is a sharing graph for $\{x_1/t_1, \ldots, x_p/t_p\}$ and S. The only thing we need to prove is that $\underline{\mathscr{G}'}$ is connected (the other conditions hold by construction).

Assume that there is an edge $i \to j$ in G^k, and consider nodes $n_i \in N_{G_i}$ and $n_j \in N_{G_j}$. We prove that there is a path in $\underline{\mathscr{G}'}$ from n_i to n_j. Actually, there is in $\bar{G}^k$ at least an edge $m_i \to m_j$ from a node $m_i \in N_{G_i}$ to $m_j \in N_{G_j}$. Since G_i and G_j are connected, there are in $\bar{G}^1$ two paths $\pi : n_i \to m_i$ and $\pi' : m_j \to n_j$. Therefore, $\pi\pi'$ is a path in $\underline{\mathscr{G}'}$ from n_i to n_j.

Now, given two generic nodes n_i, n_j, where $n_i \in N_{G_i}$ and $n_j \in N_{G_j}$, we know there is a path π in $\underline{\mathscr{G}}$ from i to j. Applying the result of the previous paragraph to each edge in π, we immediately get that n_i and n_j are connected.

Finally, it is easy to check that $res(\mathscr{G}') = B$ and this concludes the proof of the theorem. □

By exploiting the previous lemma, it is now a trivial task to show that parallel and sequential unification compute the same result.

Theorem 3.13
The abstract operators mgu_ω and mgu_p coincide.

Proof
The proof is by induction on the number of bindings in θ. Clearly, $\mathrm{mgu}_\omega(S, \epsilon) = S = \mathrm{mgu}_\mathrm{p}(S, \epsilon)$. Assume that $\mathrm{mgu}_\omega(S, \theta) = \mathrm{mgu}_\mathrm{p}(S, \theta)$ for each S. It follows that

$$\begin{array}{ll} \mathrm{mgu}_\omega(S, \{x/t\} \cup \theta) & \\ = \mathrm{mgu}_\omega(\mathrm{mgu}_\mathrm{p}(S, \{x/t\}), \theta) & [\text{by definition of } \mathrm{mgu}_\omega] \\ = \mathrm{mgu}_\mathrm{p}(\mathrm{mgu}_\mathrm{p}(S, \{x/t\}), \theta) & [\text{by induction hypothesis}] \\ = \mathrm{mgu}_\mathrm{p}(S, \{x/t\} \cup \theta) & [\text{by Lemma 3.12}] \end{array}$$

and this proves the theorem. □

3.2 Optimality of abstract unification

An immediate consequence of Theorem 3.13 is that $\mathrm{mgu_p}$ is correct, since it coincides with mgu_ω which has been proved correct in Amato and Scozzari (2010). We now want to prove that it is optimal. First, we prove optimality in the special case of $\mathrm{mgu_p}([S]_U, \theta)$ with $vars(\theta) \subseteq U$. Next, we extend this result to the general case.

In the Example 3.3, we have already shown how to build a substitution δ which mimics the effect of a sharing graph. We now give another example, introducing the terminology to be used in the proof of optimality to come.

Example 3.14

We refer to Example 3.11. Let $U = \{u, v, x, y, z\}$ be the set of variables of interest. We show how to build a substitution δ such that $[S]_U \rhd [\delta]_U$ and $u^3v^2x^2y^2z^3 \in \alpha_\omega(\mathrm{mgu}([\delta]_U, \theta))$.

For each node $n \in \{a, b, c, d, e\}$ of the sharing graph in 13, we consider a different fresh variable w_n. For any variable $v \in U \setminus \mathrm{dom}(\theta) = \{v, y, z\}$, we define $\delta(v)$ as the following term of arity $\sum_n l_{\mathscr{G}}(n)(v)$:

$$\delta(v) = r(\underbrace{w_a, \ldots, w_a}_{l_{\mathscr{G}}(a)(v) \text{ times}}, \underbrace{w_b, \ldots, w_b}_{l_{\mathscr{G}}(b)(v) \text{ times}}, \ldots, \underbrace{w_e, \ldots, w_e}_{l_{\mathscr{G}}(e)(v) \text{ times}}).$$

Since $l_{\mathscr{G}}(n)$ is the label of the node n and $l_{\mathscr{G}}(n)(v)$ is the multiplicity of v in such a label, we have:

$$\delta(v) = r(w_a, w_b, w_d) \qquad \delta(y) = r(w_d, w_e) \qquad \delta(z) = r(w_d).$$

For the variables in $\mathrm{dom}(\theta) = \{u, x\}$ we define δ in a different way. Consider the first layer of the sharing graph, corresponding to the binding x/y, and denote by f^1 an injective map from occurrences of variables w_n in $\delta(\theta(x))$ to edges targeted at n. In this case, we have $\delta(\theta(x)) = r(w_d, w_e)$ and $f^1 = \{\xi_d \mapsto e_1, \xi_e \mapsto e_2\}$, where $\xi_d = 1$ and $\xi_e = 2$ are the positions of w_d and w_e in $\delta(\theta(x))$.

Analogously, we define f^2 for the binding $u/s(z, s(z, v))$. In this case, $\delta(\theta(u)) = s(r(w_d), s(r(w_d), r(w_a, w_b, w_d))$ and a possible f^2 is $\{1 \cdot 1 \mapsto e_7, 2 \cdot 1 \cdot 1 \mapsto e_4, 2 \cdot 2 \cdot 1 \mapsto e_3, 2 \cdot 2 \cdot 2 \mapsto e_5, 2 \cdot 2 \cdot 3 \mapsto e_6\}$. In this case, other values for f^2 are possible: we could exchange the assignments for $1 \cdot 1$, $2 \cdot 1 \cdot 1$ and $2 \cdot 2 \cdot 3$ freely.

We now define $\delta(x) = f^1(\delta(\theta(x))) = r(w_a, w_b)$. Here, we denote with $f^1(t)$ the result of replacing, in the term t, the variable in position ξ with the variable associated with the source of $f^1(\xi)$. Analogously, we define $\delta(u) = f^2(\delta(\theta(u))) = s(r(w_d), s(r(w_c), r(w_c, w_e, w_b))$.

Note that the terms $\delta(x)$ and $\delta(u)$ are obtained by replacing in $\theta(x)$ and $\theta(u)$ each occurrence ξ of variable v with a variant of $\delta(v)$. We call $\delta^1_1(y)$ the term which replaces y in position 1 of $\theta(x)$, i.e., $r(w_d, w_e)$. Analogously, we define $\delta^2_1(y) = r(w_d)$, $\delta^2_{2\cdot 1}(y) = r(w_c)$, and $\delta^2_{2\cdot 2}(v) = r(w_c, w_e, w_b)$ for the replacements in $\theta(u)$ of y in position 1, y in position $2 \cdot 1$ and v in position $2 \cdot 2$, respectively. This terminology will be used in the proof.

We have that $[S]_U \rhd [\delta]_U$ and $\mathrm{mgu}(\delta, \theta) = \mathrm{mgu}(\{x = y, u = s(z, s(z, v)), u = s(r(w_d), s(r(w_c), r(w_c, w_e, w_b)), v = r(w_a, w_b, w_d), x = r(w_a, w_b), y = r(w_d, w_e), z = r(w_d)\})$

is

$$\begin{aligned}\mathrm{mgu}(\delta,\theta) &= \theta \circ \mathrm{mgu}(\{v = r(w_a, w_b, w_d), y = r(w_d, w_e), z = r(w_d)\} \cup \\ &\quad \times \{y = r(w_a, w_b), s(z, s(z, v)) = s(r(w_d), s(r(w_c), r(w_c, w_e, w_b)))\}) \\ &= \theta \circ \mathrm{mgu}(\{v = r(w_a, w_b, w_d), y = r(w_d, w_e), z = r(w_d)\} \cup \\ &\quad \times \{y = r(w_a, w_b), z = r(w_d), z = r(w_c), v = r(w_c, w_e, w_b)\} \\ &= \theta \circ \delta|_{U \setminus \mathrm{dom}(\theta)} \circ \mathrm{mgu}(\{r(w_d, w_e) = r(w_a, w_b), r(w_d) = r(w_d), \\ &\quad \times r(w_d) = r(w_c), r(w_a, w_b, w_d) = r(w_c, w_e, w_b)\}) \\ &= \theta \circ \delta|_{U \setminus \mathrm{dom}(\theta)} \circ \mathrm{mgu}(\{w_d = w_a, w_e = w_b, w_d = w_d, \\ &\quad \times w_d = w_c, w_a = w_c, w_b = w_e, w_d = w_b\}).\end{aligned}$$

In the last formula, we have an equation $w_n = w_m$ for each edge $n \to m$ in the sharing graph. Since the graph is connected, we pick a variable, say it is w_d, and we solve the set of equations with respect to that variable, obtaining:

$$\begin{aligned}\mathrm{mgu}(\delta,\theta) &= \theta \circ \delta|_{U \setminus \mathrm{dom}(\theta)} \circ \{w_a/w_d, w_b/w_d, w_c/w_d, w_e/w_d\} \\ &= \{u/s(r(w_d), s(r(w_d), r(w_d, w_d, w_d)), v/r(w_d, w_d, w_d), \\ &\quad \times x/r(w_d, w_d), y/r(w_d, w_d), z/r(w_d)\}.\end{aligned}$$

We get $\alpha_\omega(\mathrm{mgu}([\delta]_U, \theta)) = [u^5 v^3 x^2 r^2 z]_U$, where $u^5 v^3 x^2 r^2 z = \mathrm{mgu}(\delta,\theta)^{-1}(w_d)|_U$.

The above example shows how to find a substitution whose fresh variables are aliased according to the arrows in a sharing graph. The same idea is exploited in the next theorem for proving the optimality of the abstract unification operator $\mathrm{mgu}_\mathrm{p}([S]_U, \theta)$.

Theorem 3.15
The parallel unification $\mathrm{mgu}_\mathrm{p}([S]_U, \theta)$ is optimal with respect to mgu, under the assumption that $vars(\theta) \subseteq U$, that is:

$$\forall B \in \mathrm{mgu}_\mathrm{p}([S]_U, \theta)\ \exists [\delta]_U \in \mathit{ISubst}_\sim.\ [S]_U \rhd [\delta]_U \text{ and } B \in \alpha_\omega(\mathrm{mgu}([\delta]_U, \theta)).$$

Proof
Let $\theta = \{x_1/t_1, \ldots, x_p/t_p\}$ and $B \in \mathrm{mgu}_\mathrm{p}([S]_U, \theta)$. By definition of mgu_p, there exists a sharing graph $\mathscr{G} = \{G^i\}_{i \in [1,p]}$ such that $B = res(\mathscr{G})$. Let $N_\mathscr{G} = \{n_1, \ldots, n_k\}$. We want to define a substitution δ such that $[S]_U \rhd [\delta]_U$ and $B \in \alpha_\omega(\mathrm{mgu}([\delta]_U, \theta))$. If $B = \{\!\!\{\}\!\!\}$ this is trivial, just take $\delta = \epsilon$; hence, we assume that $B \neq \{\!\!\{\}\!\!\}$. The structure of the proof is as follows: first, we define a substitution δ which unifies with θ; second, we show that δ is approximated by $[S]_U$, namely, $[S]_U \rhd [\delta]_U$; third, we show that $B \in \alpha_\omega(\mathrm{mgu}([\delta]_U, \theta))$.

First part. We define a substitution δ which unifies with θ. For each node $n \in N_\mathscr{G}$ we consider a fresh variable w_n and we denote by W the set of all these new variables.

For any $y \in U \setminus \mathrm{dom}(\theta)$ we define as $\delta(y)$ the term of arity $\sum_{n \in N_\mathscr{G}} l_\mathscr{G}(n)(y)$ given by:

$$\delta(y) = r(\underbrace{w_{n_1}, \ldots, w_{n_1}}_{l_\mathscr{G}(n_1)(y)\text{ times}}, \underbrace{w_{n_2}, \ldots, w_{n_2}}_{l_\mathscr{G}(n_2)(y)\text{ times}}, \ldots, \underbrace{w_{n_k}, \ldots, w_{n_k}}_{l_\mathscr{G}(n_k)(y)\text{ times}}).$$

For any $x_i \in \mathrm{dom}(\theta)$, consider an injective function f^i which maps each occurrence of a variable w_n in $\delta(t_i)$ to an edge in E_{G^i} targeted at n. Note that the map exists since the number of occurrences of w_n in $\delta(t_i)$ is exactly $\sum_{y \in t_i} occ(y, t_i) \cdot occ(w_n, \delta(y)) = \sum_{y \in t_i} occ(y, t_i) \cdot l_{\mathscr{G}}(n)(y) = \chi(l_{\mathscr{G}}(n), t_i)$, which is the in-degree of n. Then, we define $\delta(x_i)$ as $f^i(\delta(t_i))$, where $f^i(\delta(t_i))$ is the result of replacing, in $\delta(t_i)$, the variable in position ξ with the variable associated with the source of $f^i(\xi)$.

The image of f^i is the set of all the edges in G^i. Given an edge $e : n_1 \to n_2$, the sharing group associated with n_2 should contain at least a variable $y \in t_i$; hence, w_{n_2} will occur in $\delta(y)$ and e will be $f^i(\xi)$ for some occurrence ξ of w_{n_2} in $\delta(t_i)$.

Second part. Now we show that $[S]_U \rhd [\delta]_U$. We need to consider all the variables $v \in \mathscr{V}$ and check that $\delta^{-1}(v)|_U \in S$. We distinguish several cases:

- Let us choose as v the variable w_n for some $n \in N_{\mathscr{G}}$. By construction, for each $y \in U \setminus \mathrm{dom}(\theta)$, we have that $occ(w_n, \delta(y)) = l_{\mathscr{G}}(n)(y)$. Since $\mathscr{G}$ is a sharing graph, for any $x_i \in \mathrm{dom}(\theta)$ there are $l_{\mathscr{G}}(n)(x_i)$ edges in E^i departing from n. They are all in the image of f^i, hence $occ(w_n, \delta(x_i)) = l_{\mathscr{G}}(n)(x_i)$. We obtain the required result, which is $\delta^{-1}(w_n)|_U = l_{\mathscr{G}}(n) \in S$.
- If we choose a variable $v \in U$, then $v \in \mathrm{dom}(\delta)$ and $\delta^{-1}(v) = \{\!\!\{\}\!\!\} \in S$;
- Finally, if $v \notin U \cup W$, then $\delta^{-1}(v) = \{\!\!\{v\}\!\!\}$ and $\delta^{-1}(v)|_U = \{\!\!\{\}\!\!\} \in S$.

Third part. We now show that $B \in \alpha_\omega(\mathrm{mgu}([\delta]_U, \theta))$. Note that $\delta(x_i)$ is obtained by replacing, in t_i, each occurrence ξ of a variable $y \in \mathrm{rng}(\theta)$ with a variant of $\delta(y)$. We denote this variant by $\delta^i_\xi(y) = f^i(\delta(t_i))(\xi)$. By the definition of mgu over $ISubst_\sim$, we have that $\mathrm{mgu}([\delta]_U, \theta) = [\mathrm{mgu}(\delta, \theta)]_U$. We obtain:

$$
\begin{aligned}
\eta &= \mathrm{mgu}(\delta, \theta) \\
&= \theta \circ \mathrm{mgu}\big(\mathrm{Eq}(\delta|_{U \setminus \mathrm{dom}(\theta)}) \cup \{\theta(x_i) = \delta(x_i)) \mid i \in [1,p]\}\big) \\
&= \theta \circ \mathrm{mgu}\big(\mathrm{Eq}(\delta|_{U \setminus \mathrm{dom}(\theta)}) \cup \{t_i = f^i(\delta(t_i)) \mid i \in [1,p]\}\big) \\
&= \theta \circ \mathrm{mgu}(\mathrm{Eq}(\delta|_{U \setminus \mathrm{dom}(\theta)}) \cup \{y = \delta^i_\xi(y) \mid i \in [1,p], t_i(\xi) = y\}\big) \\
&= \theta \circ \delta_{U \setminus \mathrm{dom}(\theta)} \circ \mathrm{mgu}\big(\{\delta(y) = \delta^i_\xi(y) \mid i \in [1,p], t_i(\xi) = y\}\big).
\end{aligned}
\tag{14}
$$

The set of equations $F = \{\delta(y) = \delta^i_\xi(y) \mid i \in [1,p], t_i(\xi) = y\}$ has a solution, given by aliasing some variables. We show that, for any edge $e : n \to m \in E_{G^i}$, it follows from F that $w_n = w_m$. Since the image of f^i is the set of all the edges in G^i, there is an occurrence ξ of w_m in $\delta(t_i)$ such that $f^i(\xi) = e$. Occurrence ξ may be written as $\xi' \cdot \xi''$, where ξ' is an occurrence of a variable $y \in t_i$ and ξ'' is an occurrence of w_m in $\delta(y)$. Therefore, $\delta(y) = \delta^i_{\xi'}(y) \in F$, and from this follows that $w_m = \delta(y)(\xi'')$ is unified with $w_n = \delta^i_{\xi'}(y)(\xi'')$.

Since this holds for any edge in E_{G^i} and for any $i \in [1,p]$, it follows that for any edge $n \to m \in E_{\underline{\mathscr{G}}}$ the equation $w_m = w_n$ is entailed by F. We know that $\underline{\mathscr{G}}$ is connected; hence, for any $n, m \in N_{\mathscr{G}}$, the set of equations in F implies $w_n = w_m$. We choose a particular node $\bar{n} \in N_{\mathscr{G}}$ and, for what we said before, we have

$\mathrm{mgu}(F) = \{w_n/w_{\bar{n}} \mid n \in N_{\mathscr{G}}, n \neq \bar{n}\}$. We show that $\eta^{-1}(w_{\bar{n}})|_U = B$.

$$\begin{aligned}
\eta^{-1}(w_{\bar{n}})|_U &= \theta^{-1}(\delta^{-1}|_{U \setminus \mathrm{dom}(\theta)}(\{\!\{w_{n_1}, \ldots, w_{n_k}\}\!\}))|_U \\
&= \theta^{-1}(\{\!\{w_{n_1}, \ldots, w_{n_k}\}\!\} \uplus \lambda y \in U \setminus \mathrm{dom}(\theta). \sum_{n \in N_{\mathscr{G}}} l_{\mathscr{G}}(n)(y))|_U \\
&= \lambda y \in U \setminus \mathrm{dom}(\theta). \sum_{n \in N_{\mathscr{G}}} l_{\mathscr{G}}(n)(y) \ \uplus \\
&\quad \times \lambda x \in \mathrm{dom}(\theta). \sum_{y \in \mathscr{V}} occ(y, \theta(x)) \cdot \sum_{n \in N_{\mathscr{G}}} l_{\mathscr{G}}(n)(y) \\
&= \lambda y \in U \setminus \mathrm{dom}(\theta). \sum_{n \in N_{\mathscr{G}}} l_{\mathscr{G}}(n)(y) \uplus \lambda x \in \mathrm{dom}(\theta). \sum_{n \in N_{\mathscr{G}}} \chi(l_{\mathscr{G}}(n), \theta(x)).
\end{aligned}$$

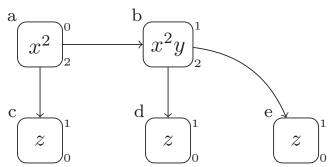
Since $\mathscr{G}$ is a sharing graph, the total in-degree for G^i, i.e., $\sum_{n \in N_{\mathscr{G}}} \chi(l_{\mathscr{G}}(n), t_i)$, is equal to the total out-degree $\sum_{n \in N_{\mathscr{G}}} \chi(l_{\mathscr{G}}(n), x_i)$. Hence

$$\begin{aligned}
\eta^{-1}(w_{\bar{n}})|_U &= \lambda y \in U \setminus \mathrm{dom}(\theta). \sum_{n \in N_{\mathscr{G}}} l_{\mathscr{G}}(n)(y) \uplus \lambda x \in \mathrm{dom}(\theta). \sum_{n \in N_{\mathscr{G}}} \chi(l_{\mathscr{G}}(n), x) \\
&= \lambda x \in U. \sum_{n \in N_{\mathscr{G}}} l_{\mathscr{G}}(n)(x) \\
&= res(\mathscr{G}) = B.
\end{aligned}$$

This concludes the proof. $\square$

The previous proof requires $vars(\theta) \subseteq U$. However, the same construction also works when this condition does not hold.

Example 3.16
Let $U = \{x, y\}, S = \{x^2, x^2y\}, \theta = \{x/s(y, z)\}$ and assume that we want to compute $\mathrm{mgu}_\omega([S]_U, \theta)$. By extending the domain of variables of interest to $V = \{x, y, z\}$, we obtain $[S']_V = [x^2, x^2y, z]_{x,y,z}$. One of the sharing graphs for θ and $[S']_V$ is



Following the proof of the previous theorem, we obtain the substitution

$$\delta' = \{x/s(r(w_a), r(w_a, w_b, w_b)), y/r(w_b), z/r(w_c, w_d, w_e)\},$$

where $[S']_V \rhd [\delta']_V$ and $x^4yz^3 \in \alpha_\omega(\mathrm{mgu}_\omega([\delta']_V, \theta))$. However, what we are looking for is a substitution δ such that $[S]_U \rhd [\delta]_U$ and $x^4yz^3 \in \alpha_\omega(\mathrm{mgu}_\omega([\delta]_U, \theta))$. Nonetheless, we may choose $\delta = \delta'$ (or, if we prefer, $\delta = \delta'|_{\{x,y\}}$) to get the required substitution.

This is not a fortuitous coincidence. We will prove that it happens consistently and, therefore, $\mathrm{mgu}_\mathrm{p}([S]_U, \theta)$ is optimal even when $vars(\theta) \not\subseteq U$.

Note that it is not an obvious result. The operation $\mathrm{mgu}_\mathrm{p}([S]_U, \theta)$ is designed by first extending the set of variables of interest of the abstract object in order to include all the variables in $vars(\theta) \setminus U$ and then performing the real operation. This construction does not always yield optimal operators. For example, Amato and Scozzari (2009) show that this is not the case for `Sharing`.

Theorem 3.17 (*Optimality of* $\mathrm{mgu_p}$)

The abstract parallel unification $\mathrm{mgu_p}([S]_U, \theta)$ is optimal, that is:

$$\forall B \in \mathrm{mgu_p}([S]_U, \theta)\ \exists [\delta]_U \in \mathit{ISubst}_\sim.\ [S]_U \rhd [\delta]_U \text{ and } B \in \alpha_\omega(\mathrm{mgu}([\delta]_U, \theta)).$$

Proof

Given $[S]_U \in \mathtt{ShLin}^\omega$ and $\theta \in \mathit{ISubst}$, proving optimality amounts to show that, for each $B \in \mathrm{mgu_p}([S]_U, \theta)$, there is $[\delta]_U \in \mathit{ISubst}_\sim$ such that $[S]_U \rhd [\delta]_U$ and $B \in \alpha_\omega(\mathrm{mgu}([\delta]_U, \theta))$. By definition $B \in \mathrm{mgu_p}([S]_U, \theta)$ iff $B \in \mathrm{mgu_p}(S', \theta)$ for $S' = S \cup \{\{\!\{v\}\!\} \mid v \in vars(\theta) \setminus U\}$. In the rest of the proof, assume $\theta = \{x_1/t_1, \ldots, x_p/t_p\}$, $V = U \cup vars(\theta)$ and $B \in \mathrm{mgu_p}(S', \theta)$.

Using the previous theorem, we find δ such that $B \in \alpha_\omega(\mathrm{mgu}([\delta]_V, \theta))$ and $[S']_V \rhd [\delta]_V$. We want to prove that $[S]_U \rhd [\delta]_U$ and $B \in \alpha_\omega(\mathrm{mgu}([\delta]_U, \theta))$.

We first prove $[S]_U \rhd [\delta]_U$. Given any $v \in \mathcal{V}$, since $[S']_V \rhd [\delta]_V$, we have $\delta^{-1}(v) \cap V \in S'$. There are two cases: either $\delta^{-1}(v) \cap V \in S$ or $\delta^{-1}(v) \cap V = \{\!\{w\}\!\}$ for some $w \in V$. In the first case, $[\![\delta^{-1}(v) \cap V]\!] \subseteq U$, hence $\delta^{-1}(v) \cap U \in S$. In the latter, $\delta^{-1}(v) \cap U = \{\!\{\}\!\} \in S$. Therefore, $[S]_U \rhd [\delta]_U$.

In order to prove that $B \in \alpha_\omega(\mathrm{mgu}([\delta]_U, \theta))$, we need to study the relationship between $\mathrm{mgu}([\delta]_U, \theta) = [\mathrm{mgu}(\delta|_U, \theta)]_V$ and $\mathrm{mgu}([\delta]_V, \theta) = [\mathrm{mgu}(\delta, \theta)]_V$. We split δ into $\delta|_{U \cup \mathrm{rng}(\theta)}$ and $\delta|_{\mathrm{dom}(\theta) \setminus U}$. With the same considerations which led to (14), we have:

$$\begin{aligned}
\mathrm{mgu}\big(\delta, \theta\big) &= \mathrm{mgu}(\mathrm{Eq}(\theta) \cup \mathrm{Eq}(\delta|_{U \cup \mathrm{rng}(\theta)}) \cup \mathrm{Eq}(\delta|_{\mathrm{dom}(\theta) \setminus U})\big) \\
&= \mathrm{mgu}\big(\mathrm{Eq}(\theta) \cup \mathrm{Eq}(\delta|_{U \cup \mathrm{rng}(\theta)}) \cup \{x_i = f^i(\delta(t_i)) \mid x_i \in \mathrm{dom}(\theta) \setminus U\}\big) \\
&= \mathrm{mgu}\big(\mathrm{Eq}(\theta) \cup \mathrm{Eq}(\delta|_{U \cup \mathrm{rng}(\theta)}) \cup \{\delta(t_i) = f^i(\delta(t_i)) \mid x_i \in \mathrm{dom}(\theta) \setminus U\}\big).
\end{aligned}$$

If $x_i \notin U$, then x_i appears in S' only in the multiset $\{\!\{x_i\}\!\}$. Hence, $\delta(x_i) = f^i(\delta(t_i))$ is linear and independent from the other variables, i.e., no variables in $f^i(\delta(t_i))$ appear in either θ or other bindings in δ. As a result, $\mathrm{mgu}(\delta, \theta)$ may be rewritten as

$$\beta \circ \mathrm{mgu}\big(\mathrm{Eq}(\theta) \cup \mathrm{Eq}(\delta|_{U \cup \mathrm{rng}(\theta)})\big),$$

where β is a substitution such that $\mathrm{dom}(\beta) = \mathrm{rng}(\delta|_{\mathrm{dom}(\theta) \setminus U})$.

We now split $U \cup \mathrm{rng}(\theta)$ into U, U_1 and U_2 where $U_1 = (\mathrm{rng}(\theta) \setminus U) \cap vars(\theta(U))$ and $U_2 = (\mathrm{rng}(\theta) \setminus U) \setminus vars(\theta(U))$. If $y \in U_1$ there exists $i_y \in [1, p]$ and a position ξ_y such that $x_{i_y} \in U$, $\theta(x_{i_y})(\xi_i) = t_{i_y}(\xi_y) = y$ and $\delta(x_{i_y})(\xi_y) = \delta^{i_y}_{\xi_y}(y)$. Since $y \notin U$, then $\delta(y)$ is linear and independent from θ and the other bindings in δ. Therefore

$$\begin{aligned}
&\mathrm{mgu}\big(\mathrm{Eq}(\theta) \cup \mathrm{Eq}(\delta|_{U\cup U_2}) \cup \mathrm{Eq}(\delta|_{U_1})\big)\\
&\quad = \mathrm{mgu}\big(\mathrm{Eq}(\theta) \cup \mathrm{Eq}(\delta|_{U\cup U_2}) \cup \{\delta^{i_y}_{\xi_y}(y) = \delta(y) \mid y \in U_1\}\big)\\
&\qquad \times \text{[by equations } x_{i_y} = t_{i_y} \text{ in } \mathrm{Eq}(\theta) \text{ and } x_{i_y} = f^{i_y}(\delta(t_{i_y})) \text{ in } \mathrm{Eq}(\delta|_U),\\
&\qquad \times \text{ restricted to position } \xi_y]\\
&\quad = \beta' \circ \mathrm{mgu}\big(\mathrm{Eq}(\theta) \cup \mathrm{Eq}(\delta|_{U\cup U_2})\big)\\
&\qquad \times \text{[linearity and independence of } \delta(y)]
\end{aligned}$$

where $\beta' = \mathrm{mgu}(\{\delta^{i_y}_{\xi_y}(y) = \delta(y) \mid y \in U_1\})$ and $\mathrm{dom}(\beta') = \mathrm{rng}(\delta|_{U_1})$. We may proceed as follows:

$$\begin{aligned}
&\mathrm{mgu}(\mathrm{Eq}(\theta) \cup \mathrm{Eq}(\delta|_U) \cup \mathrm{Eq}(\delta|_{U_2}))\\
&\quad = \mathrm{mgu}(\mathrm{Eq}(\theta|_U) \cup \mathrm{Eq}(\theta|_{\mathrm{dom}(\theta)\setminus U}) \cup \mathrm{Eq}(\delta|_U) \cup \mathrm{Eq}(\delta|_{U_2}))\\
&\quad = \theta|_{\mathrm{dom}(\theta)\setminus U} \circ \mathrm{mgu}(\mathrm{Eq}(\theta|_U) \cup \mathrm{Eq}(\delta|_U) \cup \mathrm{Eq}(\delta|_{U_2}))\\
&\qquad \times \text{[since } \mathrm{dom}(\theta) \setminus U \text{ is disjoint from the variables in } \theta|_U,\ \delta|_U,\ \delta_{U_2}]\\
&\quad = \theta|_{\mathrm{dom}(\theta)\setminus U} \circ \mathrm{mgu}(\theta|_U, \delta|_U) \circ \delta|_{U_2}\\
&\qquad \times \text{[since } \mathit{vars}(\delta|_{U_2}) \text{ is disjoint from } \mathit{vars}(\theta|_U) \cup \mathit{vars}(\delta|_U)].
\end{aligned}$$

Note that $\theta|_{\mathrm{dom}(\theta)\setminus U} \circ \mathrm{mgu}(\theta|_U, \delta|_U)$ is $\mathrm{mgu}(\delta|_U, \theta)$, and will be denoted in the following by η. Therefore, we have

$$\mathrm{mgu}(\delta, \theta) = \beta \circ \beta' \circ \eta \circ \delta|_{U_2}.$$

We check that $B \in \alpha_\omega([\eta]_V)$. Let v be the variable such that $(\beta \circ \beta' \circ \eta \circ \delta|_{U_2})^{-1}(v) \cap V = B$. We want to find $\bar{v}$ such that $\eta^{-1}(\bar{v}) \cap V = B$. First of all, since β and β' have no variables in common with V, then $(\beta \circ \beta' \circ \eta \circ \delta|_{U_2})^{-1}(v) \cap V = (\eta \circ \delta|_{U_2})^{-1}(v) \cap V$. If $v \notin \mathit{vars}(\delta|_{U_2})$, then $(\eta \circ \delta|_{U_2})^{-1} = \eta^{-1}(v)$ and we get the required result with $\bar{v} = v$. If $v \in \mathrm{rng}(\delta|_{U_2})$ we know that v only occurs once in $\delta|_{U_2}$ and never in η. Then, $(\eta \circ \delta|_{U_2})^{-1}(v) = \eta^{-1}(y) \uplus \eta^{-1}(v) = \eta^{-1}(y) \uplus \{\!\{v\}\!\}$ for the unique $y \in U_2$ such that $v \in \mathit{vars}(\delta(y))$. Therefore, since $v \notin V$, we may choose $\bar{v} = y$. Finally, if $v \in \mathrm{dom}(\delta|_{U_2})$ then $(\eta \circ \delta|_{U_2})^{-1}(v) = \{\!\{\}\!\}$ and we may take $\bar{v}$ to be any variable not in $V \cup \mathit{vars}(\eta)$. □

4 Related work

Proving optimality results for abstract unification operators on domains involving sharing information is a difficult task. The well-known domain `Sharing` (without any linearity or freeness information) is the unique domain whose abstract unification operator has been proved optimal in the general case of multibinding substitutions (Cortesi and Filé (1999), later extended by Amato and Scozzari (2009) to substitutions with variables out of the set of interest).

In the simpler case of single-binding unification, the only optimality results for domains combining sharing and linearity appeared in Amato and Scozzari (2010).

As far as we know, this is the first optimality result for domains involving linearity information for multibinding substitutions.

Although $\mathtt{ShLin}^{\omega}$ is not amenable to a direct implementation, as future work we plan to design suitable abstractions using numerical domains. The idea is to consider ω-sharing groups with symbolic multiplicities constrained by linear inequalities, such as $x^{\alpha}y^{\beta}$ with $\alpha = \beta + 2$. We plan to implement in our analyzers Random (Amato *et al.* 2010b; Amato and Scozzari 2012b) and Jandom (Amato *et al.* 2013) an abstract domain based on (template) parallelotopes (Amato *et al.* 2009, 2010a, 2012; Amato and Scozzari 2012a), exploiting the recent localized (Amato and Scozzari 2013) iteration strategies.

References

Amato, G., Di Nardo Di Maio, S. and Scozzari, F. 2013. Numerical static analysis with Soot. In *Proc. of the ACM SIGPLAN International Workshop on State of the Art in Java Program Analysis*. SOAP '13. ACM, New York, NY, USA.

Amato, G., Lipton, J. and McGrail, R. 2009. On the algebraic structure of declarative programming languages. *Theoretical Computer Science 410,* 46, 4626–4671.

Amato, G., Parton, M. and Scozzari, F. 2010a. Deriving numerical abstract domains via principal component analysis. In *Proc. of the SAS 2010*, R. Cousot and M. Martel, Eds. Lecture Notes in Computer Science, vol. 6337. Springer, Berlin Heidelberg, 134–150.

Amato, G., Parton, M. and Scozzari, F. 2010b. A tool which mines partial execution traces to improve static analysis. In *Proc. of the RV 2010*, H. Barringer, *et al.*, Eds. Lecture Notes in Computer Science, vol. 6418. Springer, Berlin, Heidelberg, 475–479.

Amato, G., Parton, M. and Scozzari, F. 2012. Discovering invariants via simple component analysis. *Journal of Symbolic Computation 47,* 12, 1533–1560.

Amato, G. and Scozzari, F. 2009. Optimality in goal-dependent analysis of sharing. *Theory and Practice of Logic Programming 9,* 5(September), 617–689.

Amato, G. and Scozzari, F. 2010. On the interaction between sharing and linearity. *Theory and Practice of Logic Programming 10,* 1(January), 49–112.

Amato, G. and Scozzari, F. 2012a. The abstract domain of parallelotopes. In *Proc. of the Fourth International Workshop on Numerical and Symbolic Abstract Domains, NSAD 2012*, J. Midtgaard and M. Might, Eds. Electronic Notes in Theoretical Computer Science, vol. 287. Elsevier, Amsterdam, The Netherlands, 17–28.

Amato, G. and Scozzari, F. 2012b. Random: R-based analyzer for numerical domains. In *Proc. of the LPAR-18, 2012*, N. Bjrner and A. Voronkov, Eds. Lecture Notes in Computer Science, vol. 7180. Springer, Berlin, Heidelberg, 375–382.

Amato, G. and Scozzari, F. 2013. Localizing widening and narrowing. In *Proc. of the SAS 2013*, F. Logozzo and M. Fähndrich, Eds. Lecture Notes in Computer Science, vol. 7935. Springer, Berlin, Heidelberg, 25–42.

Armstrong, T., Marriott, K., Schachte, P. and Søndergaard, H. 1994. Boolean functions for dependency analysis: Algebraic properties and efficient representation. In *Proc. SAS 1994*, B. Le Charlier, Ed. Lecture Notes in Computer Science, vol. 864. Springer, Berlin, Heidelberg, 266–280.

Bagnara, R., Zaffanella, E. and Hill, P. M. 2005. Enhanced sharing analysis techniques: A comprehensive evaluation. *Theory and Practice of Logic Programming 5,* 1–2(January), 1–43.

Cortesi, A. and Filé, G. 1999. Sharing is optimal. *The Journal of Logic Programming 38,* 3(March), 371–386.

Cousot, P. and Cousot, R. 1992. Abstract interpretation frameworks. *Journal of Logic and Computation 2,* 4 (August), 511–549.

Hans, W. and Winkler, S. 1992. *Aliasing and Groundness Analysis of Logic Programs through Abstract Interpretation and Its Safety* [online]. Technical Report 92–27, Technical University of Aachen (RWTH Aachen). URL: http://sunsite.informatik.rwth-aachen.de/Publications/AIB. Accessed 14 March 14 2013.

Jacobs, D. and Langen, A. 1992. Static analysis of logic programs for independent AND parallelism. *The Journal of Logic Programming 13,* 2–3(July), 291–314.

King, A. 1994. A synergistic analysis for sharing and groundness which traces linearity. In *Proc. of the ESOP 1994*, D. Sannella, Ed. Lecture Notes in Computer Science, vol. 788. Springer, Berlin, Heidelberg, 363–378.

King, A. 2000. Pair-sharing over rational trees. *The Journal of Logic Programming 46,* 1–2(November–December), 139–155.

Levi, G. and Spoto, F. 2003. Pair-independence and freeness analysis through linear refinement. *Information and Computation 182,* 1(April), 14–52.

Mac Lane, S. 1971. *Categories for the Working Mathematician.* Graduate Texts in Mathematics, vol. 5. Springer, Berlin, Heidelberg.

Marriott, K., Søndergaard, H. and Jones, N. D. 1994. Denotational abstract interpretation of logic programs. *ACM Transactions on Programming Languages and Systems 16,* 3(May), 607–648.

Muthukumar, K. and Hermenegildo, M. V. 1992. Compile-time derivation of variable dependency using abstract interpretation. *The Journal of Logic Programming 13,* 2–3(July), 315–347.

Søndergaard, H. 1986. An application of abstract interpretation of logic programs: Occur check reduction. In *Proc. of the ESOP 1986*, B. Robinet and R. Wilhelm, Eds. Lecture Notes in Computer Science, vol. 213. Springer, Berlin, Heidelberg, 327–338.