A Correct, Precise and Efficient Integration of Set-Sharing, Freeness and Linearity for the Analysis of Finite and Rational Tree Languages*

Patricia M. Hill[†] Enea Zaffanella[‡] Roberto Bagnara[§]

December 21, 2001

Abstract

It is well-known that freeness and linearity information positively interact with aliasing information, therefore allowing both the precision and the efficiency of the sharing analysis of logic programs to be improved. In this paper we present a novel combination of set-sharing with freeness and linearity information, which is characterized by an improved abstract unification operator. We provide a new abstraction function and prove the correctness of the analysis for both the finite tree and the rational tree cases. Moreover, we show that the same notion of redundant information as identified in [3] also applies to this abstract domain combination: this allows for the implementation of an abstract unification operator running in polynomial time and achieving the same precision on all the considered observable properties.

1 Introduction

Even though the set-sharing domain is, in a sense, remarkably precise, more precision is attainable by combining it with other domains. In particular, freeness and linearity information have received much attention by the literature on sharing analysis (recall that a variable is said to be free if it is not bound to a non-variable term; it is linear if it is not bound to a term containing multiple occurrences of another variable).

As argued informally by Søndergaard [34], the mutual interaction between linearity and aliasing information can improve the accuracy of a sharing analysis. This observation has been formally applied in [11] to the specification of the abstract mgu operator for the domain ASub. In his PhD thesis [30], Langen proposed a similar integration with linearity, but for the set-sharing domain. He also shown how the aliasing information allows to compute freeness with

*The present work has been partly funded by MURST project "Certificazione automatica di programmi mediante interpretazione astratta" and by the University of Parma's FIL scientific research project (ex 60%) "Matematica pura e matematica applicata". The work of P. M. Hill and E. Zaffanella has been partly funded by EPSRC under grant M05645.

[†]School of Computing, University of Leeds, UK, hill@comp.leeds.ac.uk.

[‡]Department of Mathematics, University of Parma,, Italy, zaffanella@cs.unipr.it.

[§]Department of Mathematics, University of Parma, Italy, bagnara@cs.unipr.it.

a good degree of accuracy (however, freeness information was not exploited to improve aliasing). King [27] also shown how a more refined tracking of linearity allows for further precision improvements.

The synergy attainable from a bi-directional interaction between aliasing and freeness information was initially pointed out by Muthukumar and Hermenegildo [32, 33]. Since then, several authors considered the integration of set-sharing with freeness, sometimes also including additional explicit structural information [9, 10, 29, 18].

Building on the results obtained in [34], [11] and [32], but independently from [30], Hans and Winkler [20] proposed a combined integration of freeness and linearity information with set-sharing. Similar combinations have been proposed in [5, 6, 7]. From a more pragmatic point of view, Codish et al. [12, 13] integrate the information captured by the domains of [34] and [32] by performing the analysis with both the domains at the same time, exchanging information between the two components at each step.

Most of the above proposals differ in the carrier of the underlying abstract domain. Even when considering the simplest domain combinations, where no explicit structural information is considered, there is no general consensus on the specification of the abstract unification procedure. From a theoretical point of view, once the abstract domain has been related to the concrete one by means of a Galois connection, it is always possible to specify the best correct approximation of each operator of the concrete semantics. However, empirical observations suggest that sub-optimal operators are likely to result in better complexity/precision trade-offs [4]. As a consequence, it is almost impossible to identify "the right combination" of variable aliasing with freeness and linearity information, at least when practical issues, such as the complexity of the abstract unification procedure, are taken into account.

Given this state of affairs, we will now consider a domain combination whose carrier is essentially the same as specified by Langen [30] and Hans and Winkler [20] (the same domain combination was also considered by Bruynooghe et al. [6, 7], but with the addition of compoundness and explicit structural information). The novelty of our proposal lies in the specification of an improved abstract unification procedure, better exploiting the interaction between sharing and linearity. As a matter of fact, we provide an example showing that all previous approaches to the combination of set-sharing with freeness and linearity are not uniformly more precise than the analysis based on the ASub domain [11, 28, 34].

By extending the results of [22] to this combination, we provide a new abstraction function that can be applied to any logic language computing on domains of syntactic structures, with or without the occurs-check; by using this abstraction function, we also prove the correctness of the new abstract unification procedure. Moreover, we show that the same notion of redundant information as identified in [2, 3, 37] also applies to this abstract domain combination. As a consequence, it is possible to implement an algorithm for abstract unification running in polynomial time and still obtain the same precision on all the considered observables, which are groundness, independence, freeness and linearity.

2 Preliminaries

For a set S, $\wp(S)$ is the powerset of S. The cardinality of S is denoted by $\# S$ and the emptyset is denoted by $\varnothing$. The notation $\wp_{\mathrm{f}}(S)$ stands for the set of all the *finite* subsets of S, while the notation $S \subseteq_{\mathrm{f}} T$ stands for $S \in \wp_{\mathrm{f}}(T)$.

2.1 Terms and Trees

Let *Sig* denote a possibly infinite set of function symbols, ranked over the set of natural numbers. Let *Vars* denote a denumerable set of variables, disjoint from *Sig*. Then *Terms* denotes the free algebra of all (possibly infinite) terms in the signature *Sig* having variables in *Vars*. Thus a term can be seen as an ordered labeled tree, possibly having some infinite paths and possibly containing variables: every inner node is labeled with a function symbol in *Sig* with a rank matching the number of the node's immediate descendants, whereas every leaf is labeled by either a variable in *Vars* or a function symbol in *Sig* having rank 0 (a constant). It is assumed that *Sig* contains at least two distinct function symbols, one having rank 0 (so that there exist finite terms having no variables) and one having rank greater than 0 (so that there exist infinite terms).

If $t \in \mathit{Terms}$ then $\mathrm{vars}(t)$ and $\mathrm{mvars}(t)$ denote the set and the multiset of variables occurring in t, respectively. We will also write $\mathrm{vars}(o)$ to denote the set of variables occurring in an arbitrary syntactic object o. To prove a few of the results of this thesis, it is useful to assume a total ordering, denoted with '$\leq$', on *Vars*.

Suppose $s, t \in \mathit{Terms}$: s and t are *independent* if $\mathrm{vars}(s) \cap \mathrm{vars}(t) = \varnothing$; if $y \in \mathrm{vars}(t)$ and $\neg\bigl(y \Subset \mathrm{mvars}(t)\bigr)$ we say that variable y *occurs linearly in* t, more briefly written using the predication $\mathrm{occ\_lin}(y, t)$; t is said to be *ground* if $\mathrm{vars}(t) = \varnothing$; t is *free* if $t \in \mathit{Vars}$; t is *linear* if, for all $y \in \mathrm{vars}(t)$, we have $\mathrm{occ\_lin}(y, t)$; finally, t is a *finite term* (or *Herbrand term*) if it contains a finite number of occurrences of function symbols. The sets of all ground, linear and finite terms are denoted by *GTerms*, *LTerms* and *HTerms*, respectively.

The function $\mathrm{size} \colon \mathit{HTerms} \to \mathbb{N}$, for each $t \in \mathit{HTerms}$, is defined by

$$\mathrm{size}(t) \stackrel{\mathrm{def}}{=} \begin{cases} 1, & \text{if } t \in \mathit{Vars}; \\ 1 + \sum_{i=1}^{n} \mathrm{size}(t_i), & \text{if } t = f(t_1, \ldots, t_n). \end{cases}$$

A path $p \in \bigl(\mathbb{N} \setminus \{0\}\bigr)^{\star}$ is any finite sequence of (non-zero) natural numbers. Given a path p and a (possibly infinite) term $t \in \mathit{Terms}$, we denote by $t[p]$ the subterm of t found by following path p. Formally,

$$t[p] = \begin{cases} t & \text{if } p = \epsilon; \\ t_i[q] & \text{if } p = i \,.\, q \wedge (1 \leq i \leq n) \wedge t = f(t_1, \ldots, t_n). \end{cases}$$

Note that $t[p]$ is only defined for those paths p actually corresponding to subterms of t.

2.2 Substitutions

A *substitution* is a total function $\sigma \colon \mathit{Vars} \to \mathit{HTerms}$ that is the identity almost everywhere; in other words, the *domain* of σ,

$$\mathrm{dom}(\sigma) \stackrel{\mathrm{def}}{=} \bigl\{\, x \in \mathit{Vars} \bigm| \sigma(x) \neq x \,\bigr\},$$

is finite. Given a substitution $\sigma\colon \mathit{Vars} \to \mathit{HTerms}$, we overload the symbol 'σ' so as to denote also the function $\sigma\colon \mathit{HTerms} \to \mathit{HTerms}$ defined as follows, for each term $t \in \mathit{HTerms}$:

$$\sigma(t) \stackrel{\text{def}}{=} \begin{cases} t, & \text{if } t \text{ is a constant symbol;} \\ \sigma(t), & \text{if } t \in \mathit{Vars}; \\ f\bigl(\sigma(t_1), \ldots, \sigma(t_n)\bigr), & \text{if } t = f(t_1, \ldots, t_n). \end{cases}$$

If $t \in \mathit{HTerms}$, we write $t\sigma$ to denote $\sigma(t)$. Note that, for each substitution σ and each finite term $t \in \mathit{HTerms}$, it holds $\operatorname{size}(t) \leq \operatorname{size}(t\sigma)$.

If $x \in \mathit{Vars}$ and $t \in \mathit{HTerms} \setminus \{x\}$, then $x \mapsto t$ is called a *binding*. The set of all bindings is denoted by *Bind*. Substitutions are denoted by the set of their bindings, thus a substitution σ is identified with the (finite) set

$$\bigl\{\, x \mapsto \sigma(x) \bigm| x \in \operatorname{dom}(\sigma) \,\bigr\}.$$

We denote by $\operatorname{vars}(\sigma)$ the set of variables occurring in the bindings of σ. We also define the sets $\operatorname{param}(\sigma)$ and $\operatorname{range}(\sigma)$ (the *parameter variables* and the *range variables* of σ, respectively) as

$$\operatorname{param}(\sigma) \stackrel{\text{def}}{=} \operatorname{vars}(\sigma) \setminus \operatorname{dom}(\sigma),$$

$$\operatorname{range}(\sigma) \stackrel{\text{def}}{=} \bigl\{\, y \in \operatorname{vars}(t) \bigm| (x \mapsto t) \in \sigma \,\bigr\}.$$

A substitution is said to be *circular* if, for $n > 1$, it has the form

$$\{x_1 \mapsto x_2, \ldots, x_{n-1} \mapsto x_n, x_n \mapsto x_1\},$$

where x_1, ..., x_n are distinct variables. A substitution is in *rational solved form* if it has no circular subset. The set of all substitutions in rational solved form is denoted by *RSubst*. A substitution σ is *idempotent* if, for all $t \in \mathit{Terms}$, we have $t\sigma\sigma = t\sigma$. Equivalently, σ is idempotent if and only if $\operatorname{dom}(\sigma) \cap \operatorname{range}(\sigma) = \varnothing$. The set of all idempotent substitutions is denoted by *ISubst* and $\mathit{ISubst} \subset \mathit{RSubst}$.

Example 1 *The following hold:*

$$\begin{aligned} \bigl\{x \mapsto y, y \mapsto a\bigr\} &\in \mathit{RSubst} \setminus \mathit{ISubst}, \\ \bigl\{x \mapsto a, y \mapsto a\bigr\} &\in \mathit{ISubst}, \\ \bigl\{x \mapsto y, y \mapsto g(y)\bigr\} &\in \mathit{RSubst} \setminus \mathit{ISubst}, \\ \bigl\{x \mapsto y, y \mapsto g(x)\bigr\} &\in \mathit{RSubst} \setminus \mathit{ISubst}, \\ \bigl\{x \mapsto y, y \mapsto x\bigr\} &\notin \mathit{RSubst}, \\ \bigl\{x \mapsto y, y \mapsto x, z \mapsto a\bigr\} &\notin \mathit{RSubst}. \end{aligned}$$

We have assumed that there is a total ordering '$\leq$' for *Vars*. We say that $\sigma \in \mathit{RSubst}$ is *ordered* (with respect to this ordering) if, for each binding $(x \mapsto y) \in \sigma$ such that $y \in \operatorname{param}(\sigma)$, we have $y < x$.

We will sometimes write $t[x/s]$ to denote $t\{x \mapsto s\}$.

The composition of substitutions is defined in the usual way. Thus $\tau \circ \sigma$ is the substitution such that, for all terms $t \in \mathit{HTerms}$,

$$(\tau \circ \sigma)(t) = \tau\bigl(\sigma(t)\bigr)$$

and has the formulation

$$\tau \circ \sigma = \left\{ x \mapsto x\sigma\tau \;\middle|\; x \in \mathrm{dom}(\sigma), x \neq x\sigma\tau \right\} \\ \cup \left\{ x \mapsto x\tau \;\middle|\; x \in \mathrm{dom}(\tau) \setminus \mathrm{dom}(\sigma) \right\}. \quad (1)$$

As usual, σ^0 denotes the identity function (i.e., the empty substitution) and, when $i > 0$, σ^i denotes the substitution $(\sigma \circ \sigma^{i-1})$.

For each $\sigma \in \mathit{RSubst}$, $s \in \mathit{HTerms}$, the sequence of finite terms

$$\sigma^0(s), \sigma^1(s), \sigma^2(s), \ldots$$

converges to a (possibly infinite) term, denoted $\sigma^\infty(s)$ [23, 28]. Therefore, the function $\mathrm{rt}\colon \mathit{HTerms} \times \mathit{RSubst} \to \mathit{Terms}$ such that

$$\mathrm{rt}(s, \sigma) \stackrel{\mathrm{def}}{=} \sigma^\infty(s)$$

is well defined. Note that, in general, this function is not a substitution: while having a finite domain, its "bindings" $x \mapsto \mathrm{rt}(x, \sigma)$ can map a domain variable x into a term $\mathrm{rt}(x, \sigma) \in \mathit{Terms} \setminus \mathit{HTerms}$. However, as the name of the function suggests, the term $\mathrm{rt}(x, \sigma)$ is granted to be *rational*, meaning that it can only have a finite number of distinct subterms. Rational terms, even though infinite in the sense that they admit paths of infinite length, can be finitely represented.

We have the following useful result regarding rt and substitutions that are equivalent with respect to any given syntactic equality theory.

Proposition 2 *Let $\sigma, \tau \in \mathit{RSubst}$ be satisfiable in the syntactic equality theory T and suppose that $T \vdash \forall(\sigma \leftrightarrow \tau)$. Then*

$$\mathrm{rt}(y, \sigma) \in \mathit{Vars} \iff \mathrm{rt}(y, \tau) \in \mathit{Vars}, \quad (2)$$

$$\mathrm{rt}(y, \sigma) \in \mathit{GTerms} \iff \mathrm{rt}(y, \tau) \in \mathit{GTerms}, \quad (3)$$

$$\mathrm{rt}(y, \sigma) \in \mathit{LTerms} \iff \mathrm{rt}(y, \tau) \in \mathit{LTerms}. \quad (4)$$

2.3 Equality Theories

An *equation* is of the form $s = t$ where $s, t \in \mathit{HTerms}$. Eqs denotes the set of all equations. A substitution σ may be regarded as a finite set of equations, that is, as the set $\left\{ x = t \;\middle|\; (x \mapsto t) \in \sigma \right\}$. We say that a set of equations e is in *rational solved form* if $\left\{ s \mapsto t \;\middle|\; (s = t) \in e \right\} \in \mathit{RSubst}$. In the rest of the paper, we will often write a substitution $\sigma \in \mathit{RSubst}$ to denote a set of equations in rational solved form (and vice versa).

As is common in research work involving equality, we overload the symbol '=' and use it to denote both equality and to represent syntactic identity. The context makes it clear what is intended.

Let $\{r, s, t, s_1, \ldots, s_n, t_1, \ldots, t_n\} \subseteq \mathit{HTerms}$. We assume that any equality theory T over Terms includes the *congruence axioms* denoted by the following schemata:

$$s = s, \quad (5)$$

$$s = t \leftrightarrow t = s, \quad (6)$$

$$r = s \land s = t \to r = t, \quad (7)$$

$$s_1 = t_1 \wedge \cdots \wedge s_n = t_n \to f(s_1, \ldots, s_n) = f(t_1, \ldots, t_n). \tag{8}$$

In logic programming and most implementations of Prolog it is usual to assume an equality theory based on syntactic identity. This consists of the congruence axioms together with the *identity axioms* denoted by the following schemata, where f and g are distinct function symbols or $n \neq m$:

$$f(s_1, \ldots, s_n) = f(t_1, \ldots, t_n) \to s_1 = t_1 \wedge \cdots \wedge s_n = t_n, \tag{9}$$

$$\neg\big(f(s_1, \ldots, s_n) = g(t_1, \ldots, t_m)\big). \tag{10}$$

The axioms characterized by schemata (9) and (10) ensure the equality theory depends only on the syntax. The equality theory for a non-syntactic domain replaces these axioms by ones that depend instead on the semantics of the domain and, in particular, on the interpretation given to functor symbols.

The equality theory of Clark [8], denoted $\mathcal{FT}$, on which pure logic programming is based, usually called the *Herbrand* equality theory, is given by the congruence axioms, the identity axioms, and the axiom schema

$$\forall z \in \mathit{Vars} : \forall t \in (\mathit{HTerms} \setminus \mathit{Vars}) : z \in \mathrm{vars}(t) \to \neg(z = t). \tag{11}$$

Axioms characterized by the schema (11) are called the *occurs-check axioms* and are an essential part of the standard unification procedure in SLD-resolution.

An alternative approach used in some implementations of logic programming systems, such as Prolog II, SICStus and Oz, does not require the occurs-check axioms. This approach is based on the theory of rational trees [14, 15], denoted $\mathcal{RT}$. It assumes the congruence axioms and the identity axioms together with a *uniqueness axiom* for each substitution in rational solved form. Informally speaking these state that, after assigning a ground rational tree to each parameter variable, the substitution uniquely defines a ground rational tree for each of its domain variables. Note that being in rational solved form is a very weak property. Indeed, unification algorithms returning a set of equations in rational solved form are allowed to be much more "lazy" than one would usually expect (e.g., see the first substitution in Example 1). We refer the interested reader to [25, 26, 31] for details on the subject.

In the sequel we will use the expression "equality theory" to denote any consistent, decidable theory T satisfying the congruence axioms. We will also use the expression "syntactic equality theory" to denote any equality theory T also satisfying the identity axioms. When the equality theory T is clear from the context, it is convenient to adopt the notations $\sigma \implies \tau$ and $\sigma \iff \tau$, where σ, τ are sets of equations, to denote $T \vdash \forall(\sigma \to \tau)$ and $T \vdash \forall(\sigma \leftrightarrow \tau)$, respectively.

Given an equality theory T, and a set of equations in rational solved form σ, we say that σ is *satisfiable* in T if $T \vdash \forall \mathit{Vars} \setminus \mathrm{dom}(\sigma) : \exists \mathrm{dom}(\sigma) \mathrel{.} \sigma$. If T is a syntactic equality theory that also includes the occurs-check axioms, and σ is satisfiable in T, then we say that σ is *Herbrand*.

Given a satisfiable set of equations $e \in \wp_{\mathrm{f}}(\mathit{Eqs})$ in an equality theory T, then a substitution $\sigma \in \mathit{RSubst}$ is called a *solution for e in T* if σ is satisfiable in T and $T \vdash \forall(\sigma \to e)$. If $\mathrm{vars}(\sigma) \subseteq \mathrm{vars}(e)$, then σ is said to be a *relevant* solution for e. In addition, σ is a *most general solution for e in T* if $T \vdash \forall(\sigma \leftrightarrow e)$. In this thesis, a most general solution is always a relevant solution of e. When the theory T is clear from the context, the set of all the relevant most general solutions for e in T is denoted by $\mathrm{mgs}(e)$.

Observe that, given an equality theory T, a set of equations in rational solved form may not be satisfiable in T. For example, $\exists x \,.\, \bigl\{x = f(x)\bigr\}$ is false in the equality theory $\mathcal{FT}$.

2.4 Galois Connections and uco's

Given two complete lattices $(C, \leq_C)$ and $(A, \leq_A)$, a *Galois connection* is a pair of monotonic functions $\alpha \colon C \to A$ and $\gamma \colon A \to C$ such that

$$\forall c \in C : c \leq_C \gamma\bigl(\alpha(c)\bigr), \qquad \forall a \in A : \alpha\bigl(\gamma(a)\bigr) \leq_A a.$$

The functions α and γ are said to be the abstraction and concretization functions, respectively.

An *upper closure operator* (uco) $\rho \colon C \to C$ on the complete lattice $(C, \leq_C)$ is a monotonic, idempotent and extensive[1] self-map. The set of all uco's on C, denoted by $\mathrm{uco}(C)$, is itself a complete lattice. Given a Galois connection, the function $\rho \stackrel{\mathrm{def}}{=} \gamma \circ \alpha$ is an element of $\mathrm{uco}(C)$. The presentation of abstract interpretation in terms of Galois connections can be rephrased by using uco's. In particular, the partial order $\sqsubseteq$ defined on $\mathrm{uco}(C)$ formalizes the intuition of an abstract domain being more precise than another one; moreover, given two elements $\rho_1, \rho_2 \in \mathrm{uco}(C)$, their reduced product, denoted $\rho_1 \sqcap \rho_2$, is their glb on $\mathrm{uco}(C)$.

2.5 The Set-Sharing Domain

The set-sharing domain of Jacobs and Langen [24], encodes both aliasing and groundness information. Let $\mathit{VI} \subseteq_{\mathrm{f}} \mathit{Vars}$ be a fixed and finite set of variables of interest. An element of the set-sharing domain (a *sharing set*) is a set of subsets of VI (the *sharing groups*). Note that the empty set is not a sharing group.

Definition 3 (The *set-sharing* lattice.) *Let $\mathit{SG} \stackrel{\mathrm{def}}{=} \wp(\mathit{VI}) \setminus \{\varnothing\}$ be the set of* sharing groups. *The set-sharing lattice is defined as $\mathit{SH} \stackrel{\mathrm{def}}{=} \wp(\mathit{SG})$, ordered by subset inclusion.*

The following operators on SH are needed for the specification of the abstract semantics.

Definition 4 (Auxiliary operators on SH.) *For each $sh, sh_1, sh_2 \in \mathit{SH}$ and each $V \subseteq \mathit{VI}$, we define the following functions:*
the star-union *function $(\cdot)^\star \colon \mathit{SH} \to \mathit{SH}$, is defined as*

$$sh^\star \stackrel{\mathrm{def}}{=} \bigl\{\, S \in \mathit{SG} \bigm| \exists n \geq 1 \,.\, \exists S_1, \ldots, S_n \in sh \,.\, S = S_1 \cup \cdots \cup S_n \,\bigr\};$$

the extraction of the relevant component of sh with respect to V *is encoded by* $\mathrm{rel} \colon \wp(\mathit{VI}) \times \mathit{SH} \to \mathit{SH}$ *defined as*

$$\mathrm{rel}(V, sh) \stackrel{\mathrm{def}}{=} \{\, S \in sh \mid S \cap V \neq \varnothing \,\};$$

[1] Namely, $c \leq_C \rho(c)$ for each $c \in C$.

the irrelevant component of sh *with respect to* V *is thus defined as*

$$\overline{\mathrm{rel}}(V, sh) \stackrel{\mathrm{def}}{=} sh \setminus \mathrm{rel}(V, sh);$$

the binary union *function* $\mathrm{bin}\colon SH \times SH \to SH$ *is defined as*

$$\mathrm{bin}(sh_1, sh_2) \stackrel{\mathrm{def}}{=} \{\, S_1 \cup S_2 \mid S_1 \in sh_1, S_2 \in sh_2 \,\};$$

the self-bin-union *operation on* SH *is defined as*

$$sh^2 \stackrel{\mathrm{def}}{=} \mathrm{bin}(sh, sh);$$

the abstract existential quantification *function* $\mathrm{aexists}\colon SH \times \wp(VI) \to SH$ *is defined as*

$$\mathrm{aexists}(sh, V) \stackrel{\mathrm{def}}{=} \bigl\{\, S \setminus V \bigm| S \in sh, S \setminus V \neq \varnothing \,\bigr\} \cup \bigl\{\, \{x\} \bigm| x \in V \,\bigr\}.$$

In [2, 3] it was shown that the domain SH contains many elements that are redundant for the computation of the actual *observable* properties of the analysis, that is groundness and pair-sharing. The following formalization of these observables is a rewording of the definitions provided in [36, 37].

Definition 5 (The observables of SH.) *The* groundness *and* pair-sharing *observables (on SH)* $\rho_{\scriptscriptstyle Con}, \rho_{\scriptscriptstyle PS} \in \mathrm{uco}(SH)$ *are defined, for each* $sh \in SH$*, by*

$$\rho_{\scriptscriptstyle Con}(sh) \stackrel{\mathrm{def}}{=} \bigl\{\, S \in SG \bigm| S \subseteq \mathrm{vars}(sh) \,\bigr\},$$
$$\rho_{\scriptscriptstyle PS}(sh) \stackrel{\mathrm{def}}{=} \bigl\{\, S \in SG \bigm| (P \subseteq S \wedge \#P = 2) \implies (\exists T \in sh \,.\, P \subseteq T) \,\bigr\}.$$

Definition 6 (The *pair-sharing dependency* lattice PSD.) *The operator* $\rho_{\scriptscriptstyle PSD} \in \mathrm{uco}(SH)$ *is defined, for each* $sh \in SH$*, by*

$$\rho_{\scriptscriptstyle PSD}(sh) \stackrel{\mathrm{def}}{=} \Bigl\{\, S \in SG \Bigm| \forall y \in S : S = \bigcup \bigl\{\, U \in sh \bigm| \{y\} \subseteq U \subseteq S \,\bigr\} \,\Bigr\}.$$

The pair-sharing dependency *lattice is* $PSD \stackrel{\mathrm{def}}{=} \rho_{\scriptscriptstyle PSD}(SH)$.

3 The Domain SFL

The abstract domain SFL is made up of three components, providing different kinds of sharing information regarding the set of variables of interest VI: the first component is the set-sharing domain SH of Jacobs and Langen [24]; the other two components provide freeness and linearity information, each represented by simply recording those variables of interest that are known to enjoy the corresponding property.

Definition 7 (The domain SFL.) *Let* $F \stackrel{\mathrm{def}}{=} \wp(VI)$ *and* $L \stackrel{\mathrm{def}}{=} \wp(VI)$ *be partially ordered by reverse subset inclusion. The abstract domain SFL is defined as*

$$SFL \stackrel{\mathrm{def}}{=} \bigl\{\, \langle sh, f, l \rangle \bigm| sh \in SH, f \in F, l \in L \,\bigr\}$$

and is ordered by $\leq_S$, the component-wise extension of the orderings defined on the sub-domains. With this ordering, SFL is a complete lattice whose least upper bound operation is denoted by alub_S. *The bottom element* $(\varnothing, \mathit{VI}, \mathit{VI})$ *will be denoted by* $\bot_S$.

The domain *SFL* contains many redundancies: that is, different abstract elements represent the same set of concrete computation states. For instance, any element $d = \langle sh, f, l\rangle \in \mathit{SFL}$ where $f \nsubseteq \mathrm{vars}(sh)$, such as $\bot_S$, represent the semantics of those program fragments that have no successful computations: this is because any free variable necessarily shares (at least, with itself). Similarly, the element d has the same meaning as the element $\langle sh, f, l'\rangle$, where $l' = \bigl(\mathit{VI} \setminus \mathrm{vars}(sh)\bigr) \cup f \cup l$: in this case, the reason is that any variable that is either ground or free is also necessarily linear.

All these redundancies can be removed by taking, as abstract domain, the image of the concrete domain under the abstraction function. Apart from the simple cases shown above, it is somehow difficult to *explicitly* characterize such a set. For instance, as observed in [18], the element

$$\bigl\langle \{xy, yz, xz\}, \{x, y, z\}, \{x, y, z\}\bigr\rangle \in \mathit{SFL}$$

like $\bot_S$ does not correspond to the abstraction of any concrete computation state. It is worth stressing that these "spurious" elements do not compromise the correctness of the analysis and, although they can affect the precision of the analysis, they rarely occur in practice [4, 35].

3.1 The Abstraction Function

When the concrete domain is based on the theory of finite trees, idempotent substitutions provide a finitely computable *strong normal form* for domain elements, meaning that different substitutions describe different sets of finite trees.[2] In contrast, when working on a concrete domain based on the theory of rational trees, substitutions in rational solved form, while being finitely computable, no longer satisfy this property: there can be an infinite set of substitutions in rational solved form all describing the same set of rational trees (i.e., the same element in the "intended" semantics). For instance, the substitutions

$$\sigma_n = \{x \mapsto \overbrace{f(\cdots f(}^{n} x)\cdots)\}$$

for $n = 1, 2, \ldots$, all map the variable x into the same rational tree (which is usually denoted by f^ω).

Ideally, a strong normal form for the set of rational trees described by a substitution $\sigma \in \mathit{RSubst}$ can be obtained by computing the limit σ^∞. The problem is that we may end up with $\sigma^\infty \notin \mathit{RSubst}$, as σ^∞ can map domain variables to infinite rational terms.

This poses a non-trivial problem when trying to define "good" abstraction functions, since it would be really desirable for this function to map any two equivalent concrete elements to the same abstract element. As shown in [22], the classical abstraction function for set-sharing analysis [16, 24], which was

[2] As usual, this is modulo the possible renaming of variables.

defined for idempotent substitutions only, does not enjoy this property when applied, as it is, to arbitrary substitutions in rational solved form. A possibility is to look for a more general abstraction function that allows to obtain the desired property. For example, in [21, 22] the sharing group operator 'sg' of [24] is replaced by an occurrence operator, 'occ', defined by means of a fixpoint computation. However, to simplify the presentation, here we define 'occ' directly by exploiting the fact that the number of iterations needed to reach the fixpoint is bounded by the number of bindings in the substitution.

Definition 8 (Occurrence operator.) *For each $\sigma \in \mathit{RSubst}$ and $v \in \mathit{Vars}$, the* occurrence operator $\mathrm{occ}\colon \mathit{RSubst} \times \mathit{Vars} \to \wp_{\mathrm{f}}(\mathit{Vars})$ *is defined as*

$$\mathrm{occ}(\sigma, v) \stackrel{\mathrm{def}}{=} \bigl\{\, y \in \mathit{Vars} \bigm| n = \#\sigma, v \in \mathrm{vars}(y\sigma^n) \setminus \mathrm{dom}(\sigma) \,\bigr\}.$$

For each $\sigma \in \mathit{RSubst}$, $\mathrm{ssets}\colon \mathit{RSubst} \to \mathit{SH}$ *is defined as*

$$\mathrm{ssets}(\sigma) \stackrel{\mathrm{def}}{=} \bigl\{\, \mathrm{occ}(\sigma, v) \cap \mathit{VI} \bigm| v \in \mathit{Vars} \,\bigr\} \setminus \{\varnothing\}.$$

The operator 'ssets' is introduced for notational convenience only; its additive extension corresponds to the abstraction function mapping concrete elements into elements of the set-sharing domain SH.

Example 9 *Let*

$$\begin{aligned}
\sigma &= \bigl\{x_1 \mapsto f(x_2), x_2 \mapsto g(x_3, x_4), x_3 \mapsto x_1\bigr\},\\
\tau &= \bigl\{x_1 \mapsto f(g(x_3, x_4)), x_2 \mapsto g(x_3, x_4), x_3 \mapsto f(g(x_3, x_4))\bigr\}.
\end{aligned}$$

Then $\mathrm{dom}(\sigma) = \mathrm{dom}(\tau) = \{x_1, x_2, x_3\}$ *so that* $\mathrm{occ}(\sigma, x_i) = \mathrm{occ}(\tau, x_i) = \varnothing$*, for* $i = 1, 2, 3$ *and* $\mathrm{occ}(\sigma, x_4) = \mathrm{occ}(\tau, x_4) = \{x_1, x_2, x_3, x_4\}$.

In a similar way, it is possible to define suitable operators for freeness, groundness and linearity. As all ground trees are linear, a knowledge of the definite groundness information in substitutions can be useful for proving properties concerning the linearity abstraction. Groundness is already encoded in the previously defined abstraction for set-sharing; nonetheless, for both a simplified notation and a clearer intuitive reading, we now explicitly define the set of variables that are associated to ground trees by a substitution in *RSubst*.

Definition 10 (Groundness operator.) *For each $\sigma \in \mathit{RSubst}$, the* groundness operator $\mathrm{gvars}\colon \mathit{RSubst} \to \wp_{\mathrm{f}}(\mathit{Vars})$ *is defined as*

$$\mathrm{gvars}(\sigma) \stackrel{\mathrm{def}}{=} \bigl\{\, y \in \mathrm{dom}(\sigma) \bigm| \forall v \in \mathrm{param}(\sigma) : y \notin \mathrm{occ}(\sigma, v) \,\bigr\}.$$

Example 11 *Consider $\sigma \in \mathit{RSubst}$, where*

$$\sigma = \bigl\{x_1 \mapsto x_2, x_2 \mapsto f(a), x_3 \mapsto x_4, x_4 \mapsto f(x_2, x_4)\bigr\}.$$

Then $\mathrm{gvars}(\sigma) = \{x_1, x_2, x_3, x_4\}$. *Observe that* $x_1 \in \mathrm{gvars}(\sigma)$ *although* $x_1\sigma \in \mathit{Vars}$*. Also,* $x_3 \in \mathrm{gvars}(\sigma)$ *although* $\mathrm{vars}(x_3\sigma^i) = \{x_2, x_4\} \neq \varnothing$ *for all* $i \geq 2$.

As for possible sharing, the definite freeness information can be extracted from a substitution in rational solved form by observing the result of a bounded number of applications of the substitution.

Definition 12 (Freeness operator.) *For each $\sigma \in RSubst$, the* freeness operator fvars: $RSubst \to \wp(Vars)$ *is defined as*

$$\mathrm{fvars}(\sigma) \stackrel{\mathrm{def}}{=} \{\, y \in Vars \mid n = \#\sigma, y\sigma^n \in Vars \,\}.$$

As $\sigma \in RSubst$ has no circular subset, $y \in \mathrm{fvars}(\sigma)$ implies $y\sigma^n \in Vars \setminus \mathrm{dom}(\sigma)$.

Example 13 *Consider $\sigma \in RSubst$, where*

$$\sigma = \bigl\{x_1 \mapsto x_2, x_2 \mapsto f(x_3), x_3 \mapsto x_4, x_4 \mapsto x_5\bigr\}.$$

Then, $\mathrm{fvars}(\sigma) = \{x_3, x_4, x_5\}$*. Thus, $x_1 \notin \mathrm{fvars}(\sigma)$ although $x_1\sigma \in Vars$. Also, $x_3 \in \mathrm{fvars}(\sigma)$ although $x_3\sigma \in \mathrm{dom}(\sigma)$.*

As in previous cases, the definite linearity information can be extracted by observing the result of a bounded number of applications of the considered substitution.

Definition 14 (Linearity operator.) *For each $\sigma \in RSubst$, the* linearity operator lvars: $RSubst \to \wp(Vars)$ *is defined as*

$$\mathrm{lvars}(\sigma) \stackrel{\mathrm{def}}{=} \bigl\{\, y \in Vars \bigm| n = \#\sigma, \forall z \in \mathrm{vars}(y\sigma^n) \setminus \mathrm{dom}(\sigma) : \mathrm{occ\_lin}(z, y\sigma^{2n}) \,\bigr\}.$$

In the next example we consider the extraction of linearity from two substitutions. The substitution σ shows that, in contrast with respect to set-sharing and freeness, for linearity we may need to compute up to $2n$ applications, where $n = \#\sigma$; the substitution τ shows that, when observing the term $y\tau^{2n}$, multiple occurrences of domain variables have to be disregarded.

Example 15 *Let $VI = \{x_1, x_2, x_3, x_4\}$ and consider $\sigma \in RSubst$, where*

$$\sigma = \bigl\{x_1 \mapsto x_2, x_2 \mapsto x_3, x_3 \mapsto f(x_1, x_4)\bigr\}.$$

Then, $\mathrm{lvars}(\sigma) = \{x_4\}$*. Observe that $x_1 \notin \mathrm{lvars}(\sigma)$ since $x_4 \notin \mathrm{dom}(\sigma)$, $x_4 \in x_1\sigma^3 = f(x_1, x_4)$ and $x_1\sigma^6 = f\bigl(f(x_1, x_4), x_4\bigr)$, so that $\mathrm{occ\_lin}(x_4, x_1\sigma^6)$ does not hold. Note also that $\mathrm{occ\_lin}(x_4, x_1\sigma^i)$ holds for $i = 3$, 4, 5.*

Let now $\tau \in RSubst$, where

$$\tau = \bigl\{x_1 \mapsto f(x_2, x_2), x_2 \mapsto f(x_2)\bigr\}.$$

Then $\mathrm{lvars}(\tau) = VI$*. Note that we have $x_1 \in \mathrm{lvars}(\tau)$, although, for all $i > 0$, $x_2 \in \mathrm{dom}(\tau)$ occurs more than once in the term $x_1\tau^i$.*

The occurrence, groundness, freeness and linearity operators precisely capture the intended properties over the domain of rational trees.

Proposition 16 *If $\sigma \in RSubst$ and $y, v \in Vars$ then*

$$y \in \mathrm{occ}(\sigma, v) \iff v \in \mathrm{vars}\bigl(\mathrm{rt}(y, \sigma)\bigr), \tag{12}$$

$$y \in \mathrm{gvars}(\sigma) \iff \mathrm{rt}(y, \sigma) \in GTerms, \tag{13}$$

$$y \in \mathrm{fvars}(\sigma) \iff \mathrm{rt}(y, \sigma) \in Vars, \tag{14}$$

$$y \in \mathrm{lvars}(\sigma) \iff \mathrm{rt}(y, \sigma) \in LTerms. \tag{15}$$

Moreover, the properties of sharing, groundness, freeness and linearity are invariant with respect to substitutions that are equivalent in the given syntactic equality theory.

Proposition 17 *Let $\sigma, \tau \in RSubst$ be satisfiable in the syntactic equality theory T and suppose that $T \vdash \forall(\sigma \leftrightarrow \tau)$. Then*

$$\mathrm{ssets}(\sigma) = \mathrm{ssets}(\tau), \tag{16}$$
$$\mathrm{gvars}(\sigma) = \mathrm{gvars}(\tau), \tag{17}$$
$$\mathrm{fvars}(\sigma) = \mathrm{fvars}(\tau), \tag{18}$$
$$\mathrm{lvars}(\sigma) = \mathrm{lvars}(\tau). \tag{19}$$

We are now in position to define the abstraction function mapping rational trees to elements of the domain *SFL*.

Definition 18 (The abstraction function for *SFL*.) *For each substitution $\sigma \in RSubst$, the function $\alpha_S \colon RSubst \to SFL$ is defined by*

$$\alpha_S(\sigma) \stackrel{\mathrm{def}}{=} \bigl\langle\, \mathrm{ssets}(\sigma),\, \mathrm{fvars}(\sigma) \cap \mathit{VI},\, \mathrm{lvars}(\sigma) \cap \mathit{VI} \,\bigr\rangle,$$

The concrete domain $\wp(RSubst)$ is related to SFL by means of the abstraction function *$\alpha_S \colon \wp(RSubst) \to SFL$ such that, for each $\Sigma \in \wp(RSubst)$,*

$$\alpha_S(\Sigma) \stackrel{\mathrm{def}}{=} \mathrm{alub}_S\bigl\{\, \alpha_S(\sigma) \bigm| \sigma \in \Sigma \,\bigr\}.$$

Since the abstraction function α_S is additive, the concretization function is given by the adjoint [17]

$$\gamma_S\bigl(\langle sh, f, l\rangle\bigr) \stackrel{\mathrm{def}}{=} \bigl\{\, \sigma \in RSubst \bigm| \mathrm{ssets}(\sigma) \subseteq sh, \mathrm{fvars}(\sigma) \supseteq f, \mathrm{lvars}(\sigma) \supseteq l \,\bigr\}.$$

With the definitions given in this section, one of our objectives is fulfilled: substitutions in *RSubst* that are equivalent have the same abstraction. The following is a simple consequence of Definition 18 and Proposition 17.

Corollary 19 *Let $\sigma, \tau \in RSubst$ be satisfiable in the syntactic equality theory T and suppose $T \vdash \forall(\sigma \leftrightarrow \tau)$. Then $\alpha_S(\sigma) = \alpha_S(\tau)$.*

3.2 The Abstract Operators

The specification of the abstract unification operator on the domain *SFL* is rather complex, since it is based on a very detailed case analysis. To achieve some modularity, that will be also useful when proving its correctness, in the next definition we introduce several auxiliary abstract operators.

Definition 20 (Auxiliary operators.) *Let $s, t \in HTerms$ be finite terms such that $\mathrm{vars}(s) \cup \mathrm{vars}(t) \subseteq \mathit{VI}$. For each $d = \langle sh, f, l\rangle \in SFL$ we define the following predicates:*
s and t are independent in *d if and only if $\mathrm{ind}_d \colon HTerms^2 \to Bool$ holds for (s, t), where*

$$\mathrm{ind}_d(s, t) \stackrel{\mathrm{def}}{=} \Bigl(\mathrm{rel}\bigl(\mathrm{vars}(s), sh\bigr) \cap \mathrm{rel}\bigl(\mathrm{vars}(t), sh\bigr) = \varnothing\Bigr);$$

t is ground in d *if and only if* $\mathrm{ground}_d \colon \mathit{HTerms} \to \mathit{Bool}$ *holds for* t*, where*

$$\mathrm{ground}_d(t) \stackrel{\mathrm{def}}{=} \bigl(\mathrm{vars}(t) \subseteq \mathit{VI} \setminus \mathrm{vars}(sh)\bigr);$$

$y \in \mathrm{vars}(t)$ occurs linearly (in t) in d *if and only if* $\mathrm{occ\_lin}_d \colon \mathit{VI} \times \mathit{HTerms} \to \mathit{Bool}$ *holds for* (y, t)*, where*

$$\begin{aligned}\mathrm{occ\_lin}_d(y,t) \stackrel{\mathrm{def}}{=} \mathrm{ground}_d(y) \lor \Bigl(&\mathrm{occ\_lin}(y,t) \land (y \in l)\\ &\land \forall z \in \mathrm{vars}(t) : \bigl(y \neq z \implies \mathrm{ind}_d(y,z)\bigr)\Bigr);\end{aligned}$$

t is free in d *if and only if* $\mathrm{free}_d \colon \mathit{HTerms} \to \mathit{Bool}$ *holds for* t*, where*

$$\mathrm{free}_d(t) \stackrel{\mathrm{def}}{=} \exists y \in \mathit{VI} \mathrel{.} (y = t) \land (y \in f);$$

t is linear in d *if and only if* $\mathrm{lin}_d \colon \mathit{HTerms} \to \mathit{Bool}$ *holds for* t*, where*

$$\mathrm{lin}_d(t) \stackrel{\mathrm{def}}{=} \forall y \in \mathrm{vars}(t) : \mathrm{occ\_lin}_d(y,t).$$

The function $\mathrm{share\_with}_d \colon \mathit{HTerms} \to \wp(\mathit{VI})$ *yields the set of variables of interest that may share with the given term. For each* $t \in \mathit{HTerms}$*,*

$$\mathrm{share\_with}_d(t) \stackrel{\mathrm{def}}{=} \mathrm{vars}\Bigl(\mathrm{rel}\bigl(\mathrm{vars}(t), sh\bigr)\Bigr).$$

The function $\mathrm{cyclic}_x^t \colon \mathit{SH} \to \mathit{SH}$ *strengthens the sharing set* sh *by forcing the coupling of* x *with* t*. For each* $sh \in \mathit{SH}$ *and each* $(x \mapsto t) \in \mathit{Bind}$*,*

$$\mathrm{cyclic}_x^t(sh) \stackrel{\mathrm{def}}{=} \overline{\mathrm{rel}}\bigl(\{x\} \cup \mathrm{vars}(t), sh\bigr) \cup \mathrm{rel}\bigl(\mathrm{vars}(t) \setminus \{x\}, sh\bigr).$$

As a first correctness result, we have that the auxiliary operators correctly approximate the corresponding concrete properties.

Theorem 21 *Let* $d \in \mathit{SFL}$*,* $\sigma \in \gamma_S(d)$*,* $y \in \mathit{VI}$ *and* $s, t \in \mathit{HTerms}$ *be such that* $\mathrm{vars}(s) \cup \mathrm{vars}(t) \subseteq \mathit{VI}$*. Then*

$$\mathrm{ind}_d(s,t) \implies \mathrm{vars}\bigl(\mathrm{rt}(s,\sigma)\bigr) \cap \mathrm{vars}\bigl(\mathrm{rt}(t,\sigma)\bigr) = \varnothing; \tag{20}$$

$$\mathrm{ind}_d(y,t) \iff y \notin \mathrm{share\_with}_d(t); \tag{21}$$

$$\mathrm{free}_d(t) \implies \mathrm{rt}(t,\sigma) \in \mathit{Vars}; \tag{22}$$

$$\mathrm{ground}_d(t) \implies \mathrm{rt}(t,\sigma) \in \mathit{GTerms}; \tag{23}$$

$$\mathrm{lin}_d(t) \implies \mathrm{rt}(t,\sigma) \in \mathit{LTerms}. \tag{24}$$

We now introduce the abstract mgu operator, specifying how a single binding affects each component of the domain *SFL* in the context of a syntactic equality theory T.

Definition 22 **(amgu_S.)** *The function* $\mathrm{amgu}_S \colon \mathit{SFL} \times \mathit{Bind} \to \mathit{SFL}$ *captures the effects of a binding on an element of SFL. Let* $d = \langle sh, f, l \rangle \in \mathit{SFL}$ *and* $(x \mapsto t) \in \mathit{Bind}$*, where* $\{x\} \cup \mathrm{vars}(t) \subseteq \mathit{VI}$*. Let also*

$$sh' \stackrel{\mathrm{def}}{=} \mathrm{cyclic}_x^t(sh_- \cup sh''),$$

where

$$
\begin{aligned}
sh_{-} &\stackrel{\mathrm{def}}{=} \overline{\mathrm{rel}}\bigl(\{x\} \cup \mathrm{vars}(t), sh\bigr),\\
sh_x &\stackrel{\mathrm{def}}{=} \mathrm{rel}\bigl(\{x\}, sh\bigr),\\
sh_t &\stackrel{\mathrm{def}}{=} \mathrm{rel}\bigl(\mathrm{vars}(t), sh\bigr),\\
sh_{xt} &\stackrel{\mathrm{def}}{=} sh_x \cap sh_t,
\end{aligned}
$$

$$
sh'' \stackrel{\mathrm{def}}{=} \begin{cases}
\mathrm{bin}(sh_x, sh_t), & \text{if } \mathrm{free}_d(x) \vee \mathrm{free}_d(t);\\
\mathrm{bin}\bigl(sh_x \cup \mathrm{bin}(sh_x, sh_{xt}^{\star}), & \\
\qquad\qquad sh_t \cup \mathrm{bin}(sh_t, sh_{xt}^{\star})\bigr), & \text{if } \mathrm{lin}_d(x) \wedge \mathrm{lin}_d(t);\\
\mathrm{bin}(sh_x^{\star}, sh_t), & \text{if } \mathrm{lin}_d(x);\\
\mathrm{bin}(sh_x, sh_t^{\star}), & \text{if } \mathrm{lin}_d(t);\\
\mathrm{bin}(sh_x^{\star}, sh_t^{\star}), & \text{otherwise.}
\end{cases}
$$

Letting $S_x \stackrel{\mathrm{def}}{=} \mathrm{share\_with}_d(x)$ *and* $S_t \stackrel{\mathrm{def}}{=} \mathrm{share\_with}_d(t)$*, we also define*

$$
f' \stackrel{\mathrm{def}}{=} \begin{cases}
f, & \text{if } \mathrm{free}_d(x) \wedge \mathrm{free}_d(t);\\
f \setminus S_x, & \text{if } \mathrm{free}_d(x);\\
f \setminus S_t, & \text{if } \mathrm{free}_d(t);\\
f \setminus (S_x \cup S_t), & \text{otherwise;}
\end{cases}
$$

$$
l' \stackrel{\mathrm{def}}{=} \bigl(VI \setminus \mathrm{vars}(sh')\bigr) \cup f' \cup l'',
$$

where

$$
l'' \stackrel{\mathrm{def}}{=} \begin{cases}
l \setminus (S_x \cap S_t), & \text{if } \mathrm{lin}_d(x) \wedge \mathrm{lin}_d(t);\\
l \setminus S_x, & \text{if } \mathrm{lin}_d(x);\\
l \setminus S_t, & \text{if } \mathrm{lin}_d(t);\\
l \setminus (S_x \cup S_t), & \text{otherwise.}
\end{cases}
$$

Then

$$
\mathrm{amgu}_S\bigl(d, x \mapsto t\bigr) \stackrel{\mathrm{def}}{=} \begin{cases}
\bot_S, & \text{if } d = \bot_S \vee \bigl(T = \mathcal{FT} \wedge x \in \mathrm{vars}(t)\bigr);\\
\langle sh', f', l' \rangle & \text{otherwise.}
\end{cases}
$$

The next result states that the abstract mgu operator is a correct approximation of the concrete one.

Theorem 23 *Let* $d \in SFL$ *and* $(x \mapsto t) \in Bind$*, where* $\{x\} \cup \mathrm{vars}(t) \subseteq VI$*. Then, for all* $\sigma \in \gamma_S(d)$ *and* $\tau \in \mathrm{mgs}\bigl(\sigma \cup \{x = t\}\bigr)$ *in the syntactic equality theory* T*, we have*

$$
\tau \in \gamma_S\bigl(\mathrm{amgu}_S(d, x \mapsto t)\bigr).
$$

We now highlight the similarities and differences of the operator amgu_S with respect to the corresponding ones defined in the "classical" proposals for an integration of set-sharing with freeness and linearity, such as [6, 20, 30]. Note that, when comparing our domain with the proposal in [6], we deliberately ignore all those enhancements that depend on properties that cannot be represented in *SFL* (i.e., compoundness and explicit structural information).

- In the computation of the set-sharing component, the main difference can be observed in the second, third and fourth cases of the definition of sh'': here we omit one of the star-unions even when the terms x and t possibly share. In contrast, in [6, 20, 30] the corresponding star-union is avoided only when $\mathrm{ind}_d(x, t)$ holds. Note that when $\mathrm{ind}_d(x, t)$ holds in the second case of sh'', then we have $sh_{xt} = \varnothing$; thus, the whole computation for this case reduces to $sh'' = \mathrm{bin}(sh_x, sh_t)$, as was the case in the previous proposals.

- Another improvement on the set-sharing component can be observed in the definition of sh': the cyclic_x^t operator allows the set-sharing description to be further enhanced when dealing with *definitely cyclic bindings*, i.e., when $x \in \mathrm{vars}(t)$. This is the rewording of a similar enhancement proposed in [1] for the domain *Pos* in the context of groundness analysis. Its net effect is to recover some groundness and sharing dependencies that would have been unnecessarily lost when using the standard operators. When $x \notin \mathrm{vars}(t)$, we have $\mathrm{cyclic}_x^t(sh_- \cup sh'') = sh_- \cup sh''$.

- The computation of the freeness component f' is the same as specified in [6, 20], and is more precise than the one defined in [30].

- The computation of the linearity component l' is the same as specified in [6], and is more precise than those defined in [20, 30].

In the following examples we show that the improvements in the abstract computation of the sharing component allow, in particular cases, to derive better information than that obtainable by using the classical abstract unification operators.

Example 24 *Let $VI = \{x, x_1, x_2, y, y_1, y_2, z\}$ and $\sigma \in RSubst$ such that*

$$\sigma \stackrel{\mathrm{def}}{=} \bigl\{x \mapsto f(x_1, x_2, z), y \mapsto f(y_1, z, y_2)\bigr\}.$$

By Definition 18, we have $d \stackrel{\mathrm{def}}{=} \alpha_S\bigl(\{\sigma\}\bigr) = \langle sh, f, l\rangle$, where

$$\begin{aligned} sh &= \{xx_1, xx_2, xyz, yy_1, yy_2\}, \\ f &= VI \setminus \{x, y\}, \\ l &= VI. \end{aligned}$$

Consider the binding $(x \mapsto y) \in Bind$. In the concrete, we compute (a substitution equivalent to) $\tau \in \mathrm{mgs}\bigl(\sigma \cup \{x = y\}\bigr)$, where

$$\tau = \bigl\{x \mapsto f(y_1, y_2, y_2), y \mapsto f(y_1, y_2, y_2), x_1 \mapsto y_1, x_2 \mapsto y_2, z \mapsto y_2\bigr\}.$$

Note that $\alpha_S\bigl(\{\tau\}\bigr) = \langle sh_\tau, f_\tau, l_\tau\rangle$, where $sh_\tau = \{xx_1yy_1, xx_2yy_2z\}$, so that the pairs of variables $P_x = \{x_1, x_2\}$ and $P_y = \{y_1, y_2\}$ keep their independence.

When abstractly evaluating the binding, both $\mathrm{lin}_d(x)$ and $\mathrm{lin}_d(y)$ hold so that we apply the second case of the definition of sh''. By using the notation of Definition 22, we have

$$sh_x = \{xx_1, xx_2, xyz\}, \qquad sh_- = \varnothing,$$

$$sh_t = \{yy_1, yy_2, xyz\}, \qquad\qquad sh_{xt} = \{xyz\}.$$

Since we compute the star-closure of sh_{xt} only, we obtain the set-sharing component

$$sh' = \{xx_1yy_1, xx_1yy_2, xx_1yz, xx_2yy_1, xx_2yy_2, xx_2yz, xyy_1z, xyy_2z, xyz\}.$$

Thus, we precisely capture the fact that pairs P_x and P_y keep their independence.

In contrast, since $\mathrm{ind}_d(x, y)$ does not hold, all of the classical definitions of abstract unification would have required the star-closure of both sh_x and sh_t, resulting in an abstract element including, among others, the sharing group $S = \{x, x_1, x_2, y, y_1, y_2\}$. Since $P_x \cup P_y \subset S$, this independence information would have been unnecessarily lost.

Similar examples can be devised for the third and fourth cases of the definition of sh'', where only one side of the binding is known to be linear. Example 24 has another interesting, unexpected consequence. By repeating the above abstract computation on the domain ASub (e.g., using the abstract semantics operators specified in [28]), we discover that even this simpler domain precisely captures the independence of pairs P_x and P_y. Therefore, the example provides a formal proof that all the classical approaches based on set-sharing are not *uniformly* more precise than the pair-sharing domain ASub. Such a property is enjoyed by our combination *SFL* with the improved abstract unification operator. The next example shows the precision improvements arising from the use of the 'cyclic_x^t' operator.

Example 25 *Let $VI = \{x, x_1, x_2, y\}$ and $\sigma \stackrel{\mathrm{def}}{=} \bigl\{x \mapsto f(x_1, x_2)\bigr\}$. By Definition 18, we have $d \stackrel{\mathrm{def}}{=} \alpha_S\bigl(\{\sigma\}\bigr) = \langle sh, f, l\rangle$, where*

$$\begin{aligned} sh &= \{xx_1, xx_2, y\}, \\ f &= VI \setminus \{x\}, \\ l &= VI. \end{aligned}$$

Let $t = f(x, y)$ and consider the cyclic binding $(x \mapsto t) \in Bind$. In the concrete, we compute (a substitution equivalent to) $\tau \in \mathrm{mgs}\bigl(\sigma \cup \{x = t\}\bigr)$, where

$$\tau = \bigl\{x \mapsto f(x_1, x_2), x_1 \mapsto f(x_1, x_2), y \mapsto x_2, \bigr\}.$$

Note that if we further instantiate τ by grounding y, then variables x, x_1 and x_2 would become ground too. Formally, $\alpha_S\bigl(\{\tau\}\bigr) = \langle sh_\tau, f_\tau, l_\tau\rangle$, where $sh_\tau = \{xyx_1x_2\}$. Thus, as observed above, y covers x, x_1 and x_2.

When abstractly evaluating the binding, we compute

$$\begin{aligned} sh_x &= \{xx_1, xx_2\}, \\ sh_t &= \{xx_1, xx_2, y\}, \\ sh_{xt} &= sh_x, \\ sh_- \cup sh'' &= \{xx_1, xx_1x_2, xx_1x_2y, xx_1y, xx_2, xx_2y\}, \\ sh' &= \mathrm{cyclic}_x^t(sh_- \cup sh'') \\ &= \{xx_1x_2y, xx_1y, xx_2y\}. \end{aligned}$$

Note that, in the element $sh_- \cup sh''$ (which is the abstract element that would have been computed when not exploiting the 'cyclic_x^t' operator) variable y covers none of variables x, x_1 and x_2. Thus, by applying the 'cyclic_x^t' operator, this covering information is restored.

The full abstract unification operator aunify_S, capturing the effect of a sequence of bindings on an abstract element, can now be specified by a straightforward inductive definition using the operator amgu_S.

Definition 26 **(aunify_S.)** *The operator $\mathrm{aunify}_S \colon \mathit{SFL} \times \mathit{Bind}^\star \to \mathit{SFL}$ is defined, for each $d \in \mathit{SFL}$ and each sequence of bindings $bs \in \mathit{Bind}^\star$, by*

$$\mathrm{aunify}_S(d, bs) \stackrel{\mathrm{def}}{=} \begin{cases} d, & \text{if } bs = \epsilon; \\ \mathrm{aunify}_S\bigl(\mathrm{amgu}_S(d, x \mapsto t), bs'\bigr), & \text{if } bs = (x \mapsto t) \mathbin{.} bs'. \end{cases}$$

Note that the second argument of aunify_S is a *sequence* of bindings (i.e., it is not a substitution, which is a *set* of bindings), because amgu_S is neither commutative nor idempotent, so that the multiplicity and the actual order of application of the bindings can influence the overall result of the abstract computation. The correctness of the aunify_S operator is simply inherited from the correctness of the underlying amgu_S operator. In particular, any reordering of the bindings in the sequence bs still results in a correct implementation of aunify_S.

The 'merge-over-all-path' operator on the domain *SFL* is provided by alub_S and is correct by definition. Finally, we define the abstract existential quantification operator for the domain *SFL*.

Definition 27 **($\mathrm{aexists}_S$.)** *The function $\mathrm{aexists}_S \colon \mathit{SFL} \times \wp_\mathrm{f}(\mathit{VI}) \to \mathit{SFL}$ provides the* abstract existential quantification *of an element with respect to a subset of the variables of interest. For each $d \stackrel{\mathrm{def}}{=} \langle sh, f, l \rangle \in \mathit{SFL}$ and $V \subseteq \mathit{VI}$,*

$$\mathrm{aexists}_S\bigl(\langle sh, f, l\rangle, V\bigr) \stackrel{\mathrm{def}}{=} \bigl\langle \mathrm{aexists}(sh, V), f \cup V, l \cup V \bigr\rangle.$$

Note that the correctness of the $\mathrm{aexists}_S$ operator does not pose any problems.

4 SFL_2: Eliminating Redundancies

As done in [3] for the plain set-sharing domain, even when considering the richer domain *SFL* it is natural to question whether it contains redundancies with respect to the computation of the observable properties of the analysis.

It is worth stressing that the results presented in [3] and [37] cannot be simply inherited by the new domain. The concept of "redundancy" depends on both the starting domain and the given observables: in the *SFL* domain both of these have changed. First of all, as can be seen by looking at the definition of amgu_S, freeness and linearity positively interact in the computation of sharing information: *a priori* it is an open issue whether or not the "redundant" sharing groups can play a role in such an interaction. Secondly, since freeness and linearity information can be themselves usefully exploited in a number of applications of static analysis (e.g., in the optimized implementation of concrete unification or in occurs-check reduction), these properties have to be included in the observables.

This stated, we will now show that the domain *SFL* can be simplified by applying the same notion of redundancy as identified in [3]. Namely, in the definition of *SFL* it is possible to replace the set-sharing component *SH* by *PSD* without affecting the precision on groundness, pair-sharing, freeness and linearity. In order to prove such a claim, we now formalize the new observable properties.

Definition 28 (The observables of *SFL*.) *The (overloaded)* groundness *and* pair-sharing *observables* $\rho_{Con}, \rho_{PS} \in \mathrm{uco}(SFL)$ *are defined, for each* $\langle sh, f, l\rangle \in SFL$*, by*

$$\rho_{Con}\bigl(\langle sh, f, l\rangle\bigr) \stackrel{\mathrm{def}}{=} \bigl\langle \rho_{Con}(sh), \varnothing, \varnothing\bigr\rangle,$$
$$\rho_{PS}\bigl(\langle sh, f, l\rangle\bigr) \stackrel{\mathrm{def}}{=} \bigl\langle \rho_{PS}(sh), \varnothing, \varnothing\bigr\rangle;$$

the freeness *and* linearity *observables* $\rho_F, \rho_L \in \mathrm{uco}(SFL)$ *are defined, for each* $\langle sh, f, l\rangle \in SFL$*, by*

$$\rho_F\bigl(\langle sh, f, l\rangle\bigr) \stackrel{\mathrm{def}}{=} \langle SG, f, \varnothing\rangle,$$
$$\rho_L\bigl(\langle sh, f, l\rangle\bigr) \stackrel{\mathrm{def}}{=} \langle SG, \varnothing, l\rangle.$$

The overloading of ρ_{PSD} working on the domain *SFL* leaves the freeness and linearity components untouched.

Definition 29 (Non-redundant *SFL*.) *The operator* $\rho_{PSD} \in \mathrm{uco}(SFL)$ *is defined, for each* $\langle sh, f, l\rangle \in SFL$*, by*

$$\rho_{PSD}\bigl(\langle sh, f, l\rangle\bigr) \stackrel{\mathrm{def}}{=} \bigl\langle \rho_{PSD}(sh), f, l\bigr\rangle.$$

This operator induces the lattice $SFL_2 \stackrel{\mathrm{def}}{=} \rho_{PSD}(SFL)$*.*

As proved in [37], we have that $\rho_{PSD} \sqsubseteq (\rho_{Con} \sqcap \rho_{PS})$; by the above definitions, it is also straightforward to observe that $\rho_{PSD} \sqsubseteq (\rho_F \sqcap \rho_L)$; thus, ρ_{PSD} is more precise than the reduced product $(\rho_{Con} \sqcap \rho_{PS} \sqcap \rho_F \sqcap \rho_L)$. Informally, this means that the domain SFL_2 is able to *represent* all of our observable properties without precision losses.

The next theorem shows that ρ_{PSD} is a congruence with respect to the aunify_S, alub_S and $\mathrm{aexists}_S$ operators. This means that the domain SFL_2 is able to *propagate* the information on the observables as precisely as *SFL*, therefore providing a completeness result.

Theorem 30 *Let* $d_1, d_2 \in SFL$ *be such that* $\rho_{PSD}(d_1) = \rho_{PSD}(d_2)$*. Then, for each sequence of bindings* $bs \in Bind^\star$*, for each* $d' \in SFL$ *and* $V \in \wp(VI)$*,*

$$\rho_{PSD}\bigl(\mathrm{aunify}_S(d_1, bs)\bigr) = \rho_{PSD}\bigl(\mathrm{aunify}_S(d_2, bs)\bigr),$$
$$\rho_{PSD}\bigl(\mathrm{alub}_S(d', d_1)\bigr) = \rho_{PSD}\bigl(\mathrm{alub}_S(d', d_2)\bigr),$$
$$\rho_{PSD}\bigl(\mathrm{aexists}_S(d_1, V)\bigr) = \rho_{PSD}\bigl(\mathrm{aexists}_S(d_2, V)\bigr).$$

Finally, by providing the minimality result, we show that the domain SFL_2 is indeed the generalized quotient of *SFL* with respect to the reduced product $(\rho_{Con} \sqcap \rho_{PS} \sqcap \rho_F \sqcap \rho_L)$.

Theorem 31 *For each* $i \in \{1,2\}$*, let* $d_i = \langle sh_i, f_i, l_i \rangle \in \mathit{SFL}$ *be such that* $\rho_{\scriptscriptstyle PSD}(d_1) \neq \rho_{\scriptscriptstyle PSD}(d_2)$*. Then there exist a sequence of bindings* $bs \in \mathit{Bind}^\star$ *and* $\rho \in \{\rho_{\scriptscriptstyle Con}, \rho_{\scriptscriptstyle PS}, \rho_{\scriptscriptstyle F}, \rho_{\scriptscriptstyle L}\}$ *such that*

$$\rho\bigl(\mathrm{aunify}_{\scriptscriptstyle S}(d_1, bs)\bigr) \neq \rho\bigl(\mathrm{aunify}_{\scriptscriptstyle S}(d_2, bs)\bigr).$$

As far as the implementation is concerned, the results proved in [3] for the domain *PSD* can also be applied to SFL_2. In particular, in the definition of $\mathrm{amgu}_{\scriptscriptstyle S}$ every occurrence of the star-union operator can be safely replaced by the self-bin-union operator. As a consequence, it is possible to provide an implementation where the time complexity of the $\mathrm{amgu}_{\scriptscriptstyle S}$ operator is bounded by a polynomial in the number of sharing groups of the set-sharing component.

The following result provides another optimization that can be applied when both terms x and t are definitely linear, but none of them is definitely free (i.e., when we compute sh'' by the second case stated in Definition 22).

Theorem 32 *Let* $sh \in \mathit{SH}$ *and* $(x \mapsto t) \in \mathit{Bind}$*, where* $\{x\} \cup \mathrm{vars}(t) \subseteq \mathit{VI}$*. Let* $W = \mathrm{vars}(t) \setminus \{x\}$ *and* $sh_- \stackrel{\mathrm{def}}{=} \overline{\mathrm{rel}}\bigl(\{x\} \cup \mathrm{vars}(t), sh\bigr)$*,* $sh_x \stackrel{\mathrm{def}}{=} \mathrm{rel}\bigl(\{x\}, sh\bigr)$*,* $sh_t \stackrel{\mathrm{def}}{=} \mathrm{rel}\bigl(\mathrm{vars}(t), sh\bigr)$*,* $sh_W \stackrel{\mathrm{def}}{=} \mathrm{rel}(W, sh)$*,* $sh_{xt} \stackrel{\mathrm{def}}{=} sh_x \cap sh_t$ *and*

$$sh^\diamond \stackrel{\mathrm{def}}{=} sh_- \cup \mathrm{bin}\bigl(sh_x \cup \mathrm{bin}(sh_x, sh_{xt}^\star), sh_t \cup \mathrm{bin}(sh_t, sh_{xt}^\star)\bigr).$$

Then it holds

$$\rho_{\scriptscriptstyle PSD}\bigl(\mathrm{cyclic}_x^t(sh^\diamond)\bigr) = \begin{cases} \rho_{\scriptscriptstyle PSD}\bigl(sh_- \cup \mathrm{bin}(sh_x, sh_t)\bigr), & \text{if } x \notin \mathrm{vars}(t); \\ \rho_{\scriptscriptstyle PSD}\bigl(sh_- \cup \mathrm{bin}(sh_x^2, sh_W)\bigr), & \text{otherwise.} \end{cases}$$

Therefore, even when terms x and t possibly share (i.e., when $sh_{xt} \neq \varnothing$), by using SFL_2 we can avoid the expensive computation of at least one of the two inner binary unions in the expression for $sh^\diamond$.

5 Conclusion

In this paper we have introduced the abstract domain *SFL*, combining the set-sharing domain *SH* with freeness and linearity information. While the carrier of *SFL* can be considered standard, we have provided the specification of a new abstract unification operator, showing examples where this operator achieves more precision than the classical proposals. The main contributions of this paper are the following:

- we have defined a precise abstraction function, mapping arbitrary substitutions in rational solved form into their *most precise* approximation on *SFL*;
- using this abstraction function, we have provided the mandatory proof of *correctness* for the new abstract unification operator, *for both finite-tree and rational-tree languages*;
- we have shown that, in the definition of *SFL*, we can replace the set-sharing domain *SH* by its non-redundant version *PSD*. As a consequence, it is possible to implement an algorithm for abstract unification running in *polynomial time* and still obtain the same precision on all the considered observables, that is groundness, independence, freeness and linearity.

References

[1] R. Bagnara. *Data-Flow Analysis for Constraint Logic-Based Languages.* PhD thesis, Dipartimento di Informatica, Università di Pisa, Pisa, Italy, 1997. Printed as Report TD-1/97.

[2] R. Bagnara, P. M. Hill, and E. Zaffanella. Set-sharing is redundant for pair-sharing. In P. Van Hentenryck, editor, *Static Analysis: Proceedings of the 4th International Symposium*, volume 1302 of *Lecture Notes in Computer Science*, pages 53–67, Paris, France, 1997. Springer-Verlag, Berlin.

[3] R. Bagnara, P. M. Hill, and E. Zaffanella. Set-sharing is redundant for pair-sharing. *Theoretical Computer Science*, 2002. To appear.

[4] R. Bagnara, E. Zaffanella, and P. M. Hill. Enhanced sharing analysis techniques: A comprehensive evaluation. In M. Gabbrielli and F. Pfenning, editors, *Proceedings of the 2nd International ACM SIGPLAN Conference on Principles and Practice of Declarative Programming*, pages 103–114, Montreal, Canada, 2000. Association for Computing Machinery.

[5] M. Bruynooghe and M. Codish. Freeness, sharing, linearity and correctness — All at once. In P. Cousot, M. Falaschi, G. Filé, and A. Rauzy, editors, *Static Analysis, Proceedings of the Third International Workshop*, volume 724 of *Lecture Notes in Computer Science*, pages 153–164, Padova, Italy, 1993. Springer-Verlag, Berlin. An extended version is available as Technical Report CW 179, Department of Computer Science, K.U. Leuven, September 1993.

[6] M. Bruynooghe, M. Codish, and A. Mulkers. Abstract unification for a composite domain deriving sharing and freeness properties of program variables. In F. S. de Boer and M. Gabbrielli, editors, *Verification and Analysis of Logic Languages, Proceedings of the W2 Post-Conference Workshop, International Conference on Logic Programming*, pages 213–230, Santa Margherita Ligure, Italy, 1994.

[7] M. Bruynooghe, M. Codish, and A. Mulkers. A composite domain for freeness, sharing, and compoundness analysis of logic programs. Technical Report CW 196, Department of Computer Science, K.U. Leuven, Belgium, July 1994.

[8] K. L. Clark. Negation as failure. In H. Gallaire and J. Minker, editors, *Logic and Databases*, pages 293–322, Toulouse, France, 1978. Plenum Press.

[9] M. Codish, D. Dams, G. Filé, and M. Bruynooghe. Freeness analysis for logic programs — and correctness? In D. S. Warren, editor, *Logic Programming: Proceedings of the Tenth International Conference on Logic Programming*, MIT Press Series in Logic Programming, pages 116–131, Budapest, Hungary, 1993. The MIT Press. An extended version is available as Technical Report CW 161, Department of Computer Science, K.U. Leuven, December 1992.

[10] M. Codish, D. Dams, G. Filé, and M. Bruynooghe. On the design of a correct freeness analysis for logic programs. *Journal of Logic Programming*, 28(3):181–206, 1996.

[11] M. Codish, D. Dams, and E. Yardeni. Derivation and safety of an abstract unification algorithm for groundness and aliasing analysis. In Furukawa [19], pages 79–93.

[12] M. Codish, A. Mulkers, M. Bruynooghe, M. Garcìa de la Banda, and M. Hermenegildo. Improving abstract interpretations by combining domains. In *Proceedings of the ACM SIGPLAN Symposium on Partial Evaluation and Semantics-Based Program Manipulation*, pages 194–205, Copenhagen, Denmark, 1993. ACM Press. Also available as Technical Report CW 162, Department of Computer Science, K.U. Leuven, December 1992.

[13] M. Codish, A. Mulkers, M. Bruynooghe, M. Garcìa de la Banda, and M. Hermenegildo. Improving abstract interpretations by combining domains. *ACM Transactions on Programming Languages and Systems*, 17(1):28–44, January 1995.

[14] A. Colmerauer. Prolog and infinite trees. In K. L. Clark and S. Å. Tärnlund, editors, *Logic Programming, APIC Studies in Data Processing*, volume 16, pages 231–251. Academic Press, New York, 1982.

[15] A. Colmerauer. Equations and inequations on finite and infinite trees. In *Proceedings of the International Conference on Fifth Generation Computer Systems (FGCS'84)*, pages 85–99, Tokyo, Japan, 1984. ICOT.

[16] A. Cortesi and G. Filé. Sharing is optimal. *Journal of Logic Programming*, 38(3):371–386, 1999.

[17] P. Cousot and R. Cousot. Abstract interpretation: A unified lattice model for static analysis of programs by construction or approximation of fixpoints. In *Proceedings of the Fourth Annual ACM Symposium on Principles of Programming Languages*, pages 238–252, 1977.

[18] G. Filé. Share $\times$ Free: Simple and correct. Technical Report 15, Dipartimento di Matematica, Università di Padova, December 1994.

[19] K. Furukawa, editor. *Logic Programming: Proceedings of the Eighth International Conference on Logic Programming*, MIT Press Series in Logic Programming, Paris, France, 1991. The MIT Press.

[20] W. Hans and S. Winkler. Aliasing and groundness analysis of logic programs through abstract interpretation and its safety. Technical Report 92–27, Technical University of Aachen (RWTH Aachen), 1992.

[21] P. M. Hill, R. Bagnara, and E. Zaffanella. The correctness of set-sharing. In G. Levi, editor, *Static Analysis: Proceedings of the 5th International Symposium*, volume 1503 of *Lecture Notes in Computer Science*, pages 99–114, Pisa, Italy, 1998. Springer-Verlag, Berlin.

[22] P. M. Hill, R. Bagnara, and E. Zaffanella. Soundness, idempotence and commutativity of set-sharing. *Theory and Practice of Logic Programming*, 2(2):155–201, 2002. To appear. Available at `http://arXiv.org/abs/cs.PL/0102030`.

[23] B. Intrigila and M. Venturini Zilli. A remark on infinite matching vs infinite unification. *Journal of Symbolic Computation*, 21(3):2289–2292, 1996.

[24] D. Jacobs and A. Langen. Accurate and efficient approximation of variable aliasing in logic programs. In E. L. Lusk and R. A. Overbeek, editors, *Logic Programming: Proceedings of the North American Conference*, MIT Press Series in Logic Programming, pages 154–165, Cleveland, Ohio, USA, 1989. The MIT Press.

[25] J. Jaffar, J-L. Lassez, and M. J. Maher. Prolog-II as an instance of the logic programming scheme. In M. Wirsing, editor, *Formal Descriptions of Programming Concepts III*, pages 275–299. North-Holland, 1987.

[26] T. Keisu. *Tree Constraints*. PhD thesis, The Royal Institute of Technology, Stockholm, Sweden, May 1994. Also available in the SICS Dissertation Series: SICS/D–16–SE.

[27] A. King. A synergistic analysis for sharing and groundness which traces linearity. In D. Sannella, editor, *Proceedings of the Fifth European Symposium on Programming*, volume 788 of *Lecture Notes in Computer Science*, pages 363–378, Edinburgh, UK, 1994. Springer-Verlag, Berlin.

[28] A. King. Pair-sharing over rational trees. *Journal of Logic Programming*, 46(1–2):139–155, 2000.

[29] A. King and P. Soper. Depth-k sharing and freeness. In P. Van Hentenryck, editor, *Logic Programming: Proceedings of the Eleventh International Conference on Logic Programming*, MIT Press Series in Logic Programming, pages 553–568, Santa Margherita Ligure, Italy, 1994. The MIT Press.

[30] A. Langen. *Advanced Techniques for Approximating Variable Aliasing in Logic Programs*. PhD thesis, Computer Science Department, University of Southern California, 1990. Printed as Report TR 91-05.

[31] M. J. Maher. Complete axiomatizations of the algebras of finite, rational and infinite trees. In *Proceedings, Third Annual Symposium on Logic in Computer Science*, pages 348–357, Edinburgh, Scotland, 1988. IEEE Computer Society.

[32] K. Muthukumar and M. Hermenegildo. Combined determination of sharing and freeness of program variables through abstract interpretation. In Furukawa [19], pages 49–63. An extended version appeared in [33].

[33] K. Muthukumar and M. Hermenegildo. Compile-time derivation of variable dependency using abstract interpretation. *Journal of Logic Programming*, 13(2&3):315–347, 1992.

[34] H. Søndergaard. An application of abstract interpretation of logic programs: Occur check reduction. In B. Robinet and R. Wilhelm, editors, *Proceedings of the 1986 European Symposium on Programming*, volume 213 of *Lecture Notes in Computer Science*, pages 327–338. Springer-Verlag, Berlin, 1986.

[35] E. Zaffanella. *Correctness, Precision and Efficiency in the Sharing Analysis of Real Logic Languages*. PhD thesis, School of Computing, University of Leeds, Leeds, U.K., 2001. Available at `http://www.cs.unipr.it/~zaffanella/`.

[36] E. Zaffanella, P. M. Hill, and R. Bagnara. Decomposing non-redundant sharing by complementation. In A. Cortesi and G. Filé, editors, *Static Analysis: Proceedings of the 6th International Symposium*, volume 1694 of *Lecture Notes in Computer Science*, pages 69–84, Venice, Italy, 1999. Springer-Verlag, Berlin.

[37] E. Zaffanella, P. M. Hill, and R. Bagnara. Decomposing non-redundant sharing by complementation. *Theory and Practice of Logic Programming*, 2(2):233–261, 2002. To appear. Available at `http://arXiv.org/abs/cs.PL/0101025`.

A Proofs of the Results of Section 2

The next three lemmas are needed for the proof of Proposition 2. The first two are proven in [22].

Lemma 33 *Let T be an equality theory, $\sigma \in \mathit{RSubst}$ and $t \in \mathit{HTerms}$. Then*

$$T \vdash \forall\bigl(\sigma \rightarrow (t = t\sigma)\bigr).$$

Lemma 34 *Let $\sigma \in \mathit{RSubst}$ be satisfiable in the equality theory T and consider $x \mapsto t$ such that $x \notin \mathrm{dom}(\sigma)$ and $t \in \mathit{GTerms} \cap \mathit{HTerms}$. Then, $\sigma' \stackrel{\mathrm{def}}{=} \sigma \cup \{x \mapsto t\} \in \mathit{RSubst}$ and σ' is satisfiable in T.*

Lemma 35 *Let $\sigma \in \mathit{RSubst}$ be satisfiable in the syntactic equality theory T. Suppose $s, t \in \mathit{HTerms}$ are such that $T \vdash \forall\bigl(\sigma \rightarrow (s = t)\bigr)$. Then $\mathrm{rt}(s, \sigma) = \mathrm{rt}(t, \sigma)$.*

Proof. We suppose, toward a contradiction, that $\mathrm{rt}(s, \sigma) \neq \mathrm{rt}(t, \sigma)$. Then, there must exist a finite path p such that:

a. $x = \mathrm{rt}(s, \sigma)[p] \in \mathit{Vars} \setminus \mathrm{dom}(\sigma)$, $y = \mathrm{rt}(t, \sigma)[p] \in \mathit{Vars} \setminus \mathrm{dom}(\sigma)$ and $x \neq y$; or

b. $x = \mathrm{rt}(s, \sigma)[p] \in \mathit{Vars} \setminus \mathrm{dom}(\sigma)$ and $r = \mathrm{rt}(t, \sigma)[p] \notin \mathit{Vars}$ or, symmetrically, we have $r = \mathrm{rt}(s, \sigma)[p] \notin \mathit{Vars}$ and $x = \mathrm{rt}(t, \sigma)[p] \in \mathit{Vars} \setminus \mathrm{dom}(\sigma)$; or

c. $r_1 = \mathrm{rt}(s, \sigma)[p] \notin \mathit{Vars}$, $r_2 = \mathrm{rt}(t, \sigma)[p] \notin \mathit{Vars}$ and r_1 and r_2 have different principal functors.

Then, by definition of 'rt', there must exists an index $i \in \mathbb{N}$ such that one of these holds:

1. $x = s\sigma^i[p] \in \mathit{Vars} \setminus \mathrm{dom}(\sigma)$, $y = t\sigma^i[p] \in \mathit{Vars} \setminus \mathrm{dom}(\sigma)$ and $x \neq y$; or
2. $x = s\sigma^i[p] \in \mathit{Vars} \setminus \mathrm{dom}(\sigma)$ and $r = t\sigma^i[p] \notin \mathit{Vars}$ or, symmetrically, we have $r = s\sigma^i[p] \notin \mathit{Vars}$ and $x = t\sigma^i[p] \in \mathit{Vars} \setminus \mathrm{dom}(\sigma)$; or

3. $r_1 = s\sigma^i[p] \notin \mathit{Vars}$ and $r_2 = t\sigma^i[p] \notin \mathit{Vars}$ have different principal functors.

By Lemma 33, we have $T \vdash \forall\bigl(\sigma \rightarrow (s\sigma^i = t\sigma^i)\bigr)$; from this, by the identity axioms, we obtain that

$$T \vdash \forall\Bigl(\sigma \rightarrow \bigl(s\sigma^i[p] = t\sigma^i[p]\bigr)\Bigr). \tag{25}$$

We now prove that each case leads to a contradiction.

Consider case 1. Let $r_1, r_2 \in \mathit{GTerms} \cap \mathit{HTerms}$ be two ground and finite terms having different principal functors, so that $T \vdash \forall(r_1 \neq r_2)$. By Lemma 34, we have that $\sigma' = \sigma \cup \{x \mapsto r_1, y \mapsto r_2\} \in \mathit{RSubst}$ is satisfiable; moreover, $T \vdash \forall(\sigma' \rightarrow \sigma)$, $T \vdash \forall\bigl(\sigma' \rightarrow (x = r_1)\bigr)$ and $T \vdash \forall\bigl(\sigma' \rightarrow (y = r_2)\bigr)$. This is a contradiction, since, by (25), we have $T \vdash \forall\bigl(\sigma \rightarrow (x = y)\bigr)$.

Consider case 2. Without loss of generality, consider the first subcase, where $x = s\sigma^i$ and $r = t\sigma^i[p] \notin \mathit{Vars}$. Let $r' \in \mathit{GTerms} \cap \mathit{HTerms}$ be such that r and r' have different principal functors, so that $T \vdash \forall(r \neq r')$. By Lemma 34, $\sigma' = \sigma \cup \{x \mapsto r'\} \in \mathit{RSubst}$ is satisfiable; we also have $T \vdash \forall(\sigma' \rightarrow \sigma)$ and $T \vdash \forall\bigl(\sigma' \rightarrow (x = r')\bigr)$. This is a contradiction, since, by (25), $T \vdash \forall\bigl(\sigma \rightarrow (x = r)\bigr)$.

Finally, consider case 3. In this case $T \vdash \forall(r_1 \neq r_2)$. This immediately leads to a contradiction, since, by (25), $T \vdash \forall\bigl(\sigma \rightarrow (r_1 = r_2)\bigr)$.

Proof of Proposition 2 on page 5 For each stated equivalence, we will prove only one implication since the other one will follow by symmetry.

Consider (2). Reasoning by contraposition, suppose $\mathrm{rt}(y, \sigma) \notin \mathit{Vars}$. Then there exists an index $i \geq 0$ such that $y\sigma^i \notin \mathit{Vars}$. Since $T \vdash \forall(\tau \rightarrow \sigma)$, by Lemma 33 we have $T \vdash \forall\bigl(\tau \rightarrow (y = y\sigma^i)\bigr)$. By Lemma 35, we obtain $\mathrm{rt}(y, \tau) = \mathrm{rt}(y\sigma^i, \tau)$, so that $\mathrm{rt}(y, \tau) \notin \mathit{Vars}$.

Consider (3). We suppose, toward a contradiction, that $\mathrm{rt}(y, \sigma) \in \mathit{GTerms}$ but $\mathrm{rt}(y, \tau) \notin \mathit{GTerms}$. Then, there must exist a finite path p such that:

a. $r = \mathrm{rt}(y, \sigma)[p] \in \mathit{GTerms}$ and $x = \mathrm{rt}(y, \tau)[p] \in \mathit{Vars} \setminus \mathrm{dom}(\tau)$; or

b. $r_1 = \mathrm{rt}(y, \sigma)[p] \notin \mathit{Vars}$, $r_2 = \mathrm{rt}(y, \tau)[p] \notin \mathit{Vars}$ and r_1 and r_2 have different principal functors.

Then, by definition of 'rt', there must exists an index $i \in \mathbb{N}$ such that one of these holds:

1. $r = y\sigma^i[p] \notin \mathit{Vars}$ and $x = y\tau^i[p] \in \mathit{Vars} \setminus \mathrm{dom}(\tau)$; or

2. $r_1 = y\sigma^i[p] \notin \mathit{Vars}$ and $r_2 = y\tau^i[p] \notin \mathit{Vars}$ have different principal functors.

By Lemma 33, we have $T \vdash \forall\bigl(\sigma \rightarrow (y\sigma^i = y\tau^i)\bigr)$; from this, by the identity axioms, we obtain that

$$T \vdash \forall\Bigl(\sigma \rightarrow \bigl(y\sigma^i[p] = y\tau^i[p]\bigr)\Bigr). \tag{26}$$

We now prove that both cases lead to a contradiction.

Consider case 1. Let $r' \in \mathit{GTerms} \cap \mathit{HTerms}$ be such that r and r' have different principal functors, so that $T \vdash \forall(r \neq r')$. By Lemma 34, $\tau' = \tau \cup \{x \mapsto$

$r'\} \in \mathit{RSubst}$ is satisfiable; we also have $T \vdash \forall(\tau' \to \tau)$ and $T \vdash \forall\bigl(\tau' \to (x = r')\bigr)$. This is a contradiction, since, by (26), $T \vdash \forall\bigl(\tau \to (x = r)\bigr)$.

Finally, consider case 2. In this case $T \vdash \forall(r_1 \neq r_2)$. This immediately leads to a contradiction, since, by (26), $T \vdash \forall\bigl(\sigma \to (r_1 = r_2)\bigr)$.

Consider (4). Reasoning by contraposition, suppose that $\mathrm{rt}(y, \tau) \notin \mathit{LTerms}$, so that there exists $v \in \mathrm{vars}\bigl(\mathrm{rt}(y, \tau)\bigr)$ such that $\mathrm{occ\_lin}\bigl(v, \mathrm{rt}(y, \tau)\bigr)$ does not hold. By definition of 'rt', there exists an index $i \geq 0$ such that $v \in \mathrm{vars}(y\tau^i)$ and $\mathrm{occ\_lin}(v, y\tau^i)$ does not hold. Thus, as $v \notin \mathrm{dom}(\tau)$, so that $\mathrm{rt}(v, \tau) = v \in \mathit{Vars}$. By (2), $\mathrm{rt}(v, \sigma) = w \in \mathit{Vars} \setminus \mathrm{dom}(\sigma)$. Therefore there exists $j \geq 0$ such that $w = v\sigma^j$. Hence, we obtain that $w \in \mathrm{vars}\bigl(\mathrm{rt}(y\tau^i\sigma^j, \sigma)\bigr)$ and $\mathrm{occ\_lin}\bigl(w, \mathrm{rt}(y\tau^i\sigma^j, \sigma)\bigr)$ does not hold, so that $\mathrm{rt}(y\tau^i\sigma^j, \sigma) \notin \mathit{LTerms}$. Since $T \vdash \forall(\sigma \to \tau)$, by Lemma 33 we have $T \vdash \forall\bigl(\sigma \to (y = y\tau^i\sigma^j)\bigr)$. By Lemma 35, we obtain $\mathrm{rt}(y, \sigma) = \mathrm{rt}(y\tau^i\sigma^j, \sigma)$, so that $\mathrm{rt}(y, \sigma) \notin \mathit{LTerms}$.

B Proofs of the Results of Subsection 3.1.

The definition of idempotence requires that repeated applications of a substitution do not change the syntactic structure of a term. However, several abstractions of terms, such as the ones commonly used for sharing analysis, are only interested in the variables and not in the structure that contains them. Thus, an obvious way to relax the definition of idempotence to allow for a non-Herbrand substitution is to ignore the structure and just require that its repeated application leaves the set of variables in a term invariant.

Definition 36 (Variable-idempotence.) *A substitution $\sigma \in \mathit{RSubst}$ is* variable-idempotent *if and only if for all $t \in \mathit{HTerms}$ we have*

$$\mathrm{vars}(t\sigma\sigma) = \mathrm{vars}(t\sigma).$$

The set of variable-idempotent substitutions is denoted VSubst.

Note that any idempotent substitution is also variable-idempotent, so that $\mathit{ISubst} \subset \mathit{VSubst} \subset \mathit{RSubst}$. The above definition of variable-idempotence, which is the same originally provided in [21], is a bit stronger than the one adopted in [22] (*weak* variable-idempotence), where we were disregarding those variables that are in the domain of the substitution. Since any variable-idempotent substitution is also weak variable-idempotent, almost all the results proven in [22] still apply.

The next result provides an alternative characterization of variable-idempotence.

Lemma 37 *Let $\sigma \in \mathit{RSubst}$. Then*

$$\sigma \in \mathit{VSubst} \iff \forall(x \mapsto r) \in \sigma : \mathrm{vars}(r\sigma) = \mathrm{vars}(r).$$

Proof. Suppose first that $\sigma \in \mathit{VSubst}$ and let $(x \mapsto r) \in \sigma$. Then

$$\mathrm{vars}(x\sigma\sigma) = \mathrm{vars}(x\sigma)$$

and hence, $\mathrm{vars}(r\sigma) = \mathrm{vars}(r)$.

Next, suppose that for all $(x \mapsto r) \in \sigma$, $\mathrm{vars}(r\sigma) = \mathrm{vars}(r)$ and consider $t \in \mathit{HTerms}$. We will show that $\mathrm{vars}(t\sigma\sigma) = \mathrm{vars}(t\sigma)$ by induction on the size of t. If t is a constant or $t \in \mathit{Vars} \setminus \mathrm{dom}(\sigma)$, then the result follows from the fact that $t\sigma = t$. If $t \in \mathrm{dom}(\sigma)$, then there exists $(y \mapsto s) \in \sigma$ such that $t = y$, so that $t\sigma = s$. Thus, we have

$$\mathrm{vars}(t\sigma\sigma) = \mathrm{vars}(s\sigma) = \mathrm{vars}(s) = \mathrm{vars}(t\sigma).$$

Finally, if $t = f(t_1, \ldots, t_n)$, then by the inductive hypothesis $\mathrm{vars}(t_i\sigma\sigma) = \mathrm{vars}(t_i\sigma)$ for $i = 1, \ldots, n$. Therefore we have

$$\mathrm{vars}(t\sigma\sigma) = \bigcup_{i=1}^{n} \mathrm{vars}(t_i\sigma\sigma) = \bigcup_{i=1}^{n} \mathrm{vars}(t_i\sigma) = \mathrm{vars}(t\sigma).$$

Thus, by Definition 36, as $\sigma \in \mathit{RSubst}$, $\sigma \in \mathit{VSubst}$.

The next result provides a sufficient condition for a variable-idempotent substitution so that all of its subsets are variable-idempotent too.

Lemma 38 *Let $\sigma \in \mathit{VSubst}$ be such that for all $y \in \mathrm{range}(\sigma)$, $y \in \mathrm{vars}(y\sigma)$. Then, for all $\sigma' \subseteq \sigma$, $\sigma' \in \mathit{VSubst}$.*

Proof. Let $(x \mapsto t) \in \sigma' \subseteq \sigma$. We will prove that $\mathrm{vars}(t\sigma') = \mathrm{vars}(t)$, so that the thesis will follow from Lemma 37.

To prove the first implication, let $y \in \mathrm{vars}(t\sigma')$, so that $y \in \mathrm{range}(\sigma)$. If it also holds $y \in \mathrm{dom}(\sigma)$, then by the hypothesis $y \in \mathrm{vars}(y\sigma)$, so that $y \in \mathrm{vars}(t\sigma)$. Otherwise, if $y \notin \mathrm{dom}(\sigma)$, then again $y \in \mathrm{vars}(t\sigma)$. Thus, in both cases, since $\sigma \in \mathit{VSubst}$, by Lemma 37 we obtain $y \in \mathrm{vars}(t)$.

To prove the other implication, let $y \in \mathrm{vars}(t)$, so that $y \in \mathrm{range}(\sigma)$. If $y \notin \mathrm{dom}(\sigma')$ then $y \in \mathrm{vars}(t\sigma')$. Otherwise, if $y \in \mathrm{dom}(\sigma')$, then we have $y \in \mathrm{dom}(\sigma) \cap \mathrm{range}(\sigma)$. Thus, by hypothesis, $y \in \mathrm{vars}(y\sigma)$. Since $y\sigma = y\sigma'$, we have $y \in \mathrm{vars}(y\sigma')$, so that $y \in \mathrm{vars}(t\sigma')$.

The following result concerns the composition of variable idempotent substitutions.

Lemma 39 *Let $\sigma, \tau \in \mathit{VSubst}$, where $\mathrm{dom}(\sigma) \cap \mathrm{vars}(\tau) = \varnothing$. Then $\tau \circ \sigma$ has the following properties.*

1. *$T \vdash \forall\big((\tau \circ \sigma) \leftrightarrow (\tau \cup \sigma)\big)$, for any equality theory T;*
2. *$\mathrm{dom}(\tau \circ \sigma) = \mathrm{dom}(\sigma) \cup \mathrm{dom}(\tau)$;*
3. *$\tau \circ \sigma \in \mathit{VSubst}$.*

Proof. We have that $(\tau \cup \sigma) \in \mathit{RSubst}$ because, by hypothesis, $\sigma, \tau \in \mathit{RSubst}$ and $\mathrm{dom}(\sigma) \cap \mathrm{vars}(\tau) = \varnothing$. It follows from (1) that $\tau \circ \sigma$ can be obtained from $(\tau \cup \sigma)$ by a sequence of $\mathcal{S}$-steps so that, by Theorem 47, we have properties 1 and 2.

To prove property 3, we will show that, for all terms $t \in \mathit{HTerms}$,

$$\mathrm{vars}(t\sigma\tau) = \mathrm{vars}(t\sigma\tau\sigma\tau).$$

- We start by proving the inclusion $\mathrm{vars}(t\sigma\tau) \subseteq \mathrm{vars}(t\sigma\tau\sigma\tau)$. Thus, let $z \in \mathrm{vars}(t\sigma\tau)$.

  First note that, if $z \notin \mathrm{dom}(\sigma) \cup \mathrm{dom}(\tau)$, then the result is trivial.

  Suppose $z \in \mathrm{dom}(\sigma)$. By hypothesis, $z \notin \mathrm{vars}(\tau)$ so that $z \in \mathrm{vars}(t\sigma)$. Since σ is variable-idempotent, $z \in \mathrm{vars}(t\sigma\sigma)$, so that there exists $v \in \mathrm{vars}(t\sigma) \cap \mathrm{dom}(\sigma)$ such that $z \in \mathrm{vars}(v\sigma)$. Thus $v \notin \mathrm{vars}(\tau)$, so that $v \in \mathrm{vars}(t\sigma\tau)$. Therefore $z \in \mathrm{vars}(t\sigma\tau\sigma)$ and, since $z \notin \mathrm{vars}(\tau)$, we can conclude $z \in \mathrm{vars}(t\sigma\tau\sigma\tau)$.

  Otherwise, let $z \in \mathrm{dom}(\tau)$, so that $z \notin \mathrm{dom}(\sigma)$. There exists $v \in \mathrm{vars}(t\sigma) \cap \mathrm{dom}(\tau)$ such that $z \in \mathrm{vars}(v\tau)$. Since τ is variable-idempotent, $z \in \mathrm{vars}(v\tau\tau)$ so that there exists $w \in \mathrm{vars}(v\tau) \cap \mathrm{dom}(\tau)$ such that $z \in \mathrm{vars}(w\tau)$. Since $w \notin \mathrm{dom}(\sigma)$ then $w \in \mathrm{vars}(t\sigma\tau\sigma)$. Therefore we can conclude $z \in \mathrm{vars}(t\sigma\tau\sigma\tau)$.

- To prove the other inclusion, let $z \in \mathrm{vars}(t\sigma\tau\sigma\tau)$, so that there exists $v \in \mathrm{vars}(t\sigma\tau\sigma)$ such that $z \in \mathrm{vars}(v\tau)$. Similarly, there exists $w \in \mathrm{vars}(t\sigma\tau)$ such that $v \in \mathrm{vars}(w\sigma)$.

  Suppose $v \neq w$. Then $w \in \mathrm{dom}(\sigma)$, so that by hypothesis $w \notin \mathrm{vars}(\tau)$. As a consequence, $w \in \mathrm{vars}(t\sigma)$, $v \in \mathrm{vars}(t\sigma\sigma)$ and $z \in \mathrm{vars}(t\sigma\sigma\tau)$. Thus, as $\sigma \in \mathit{VSubst}$, we obtain $z \in \mathrm{vars}(t\sigma\tau)$.

  Otherwise, if $v = w$, there exists $x \in \mathrm{vars}(t\sigma)$ such that $z \in \mathrm{vars}(x\tau\tau)$. Thus, $z \in \mathrm{vars}(t\sigma\tau\tau)$ and, since $\tau \in \mathit{VSubst}$, $z \in \mathrm{vars}(t\sigma\tau)$.

The proof of the following result is the same as the proof of [22, Theorem 2] (where weaker properties were stated).

Proposition 40 *Suppose T is an equality theory and $\sigma \in \mathit{RSubst}$. Then there exists $\sigma' \in \mathit{VSubst}$ such that* $\mathrm{dom}(\sigma) = \mathrm{dom}(\sigma')$, $\mathrm{vars}(\sigma) = \mathrm{vars}(\sigma')$ *and* $T \vdash \forall(\sigma \leftrightarrow \sigma')$*; also, for all* $y \in \mathrm{dom}(\sigma') \in \mathrm{range}(\sigma')$, $y \in \mathrm{vars}(y\sigma')$.

The following result is proven in [22, Lemma 6].

Lemma 41 *Let $\tau, \sigma \in \mathit{VSubst}$ be satisfiable in the syntactic equality theory T and suppose $T \vdash \forall(\tau \rightarrow \sigma)$. In addition, let $s, t \in \mathit{HTerms}$ be such that* $T \vdash \forall\bigl(\tau \rightarrow (s = t)\bigr)$ *and* $v \in \mathrm{vars}(s) \setminus \mathrm{dom}(\tau)$. *Then there exists a variable* $z \in \mathrm{vars}(t\sigma) \setminus \mathrm{dom}(\sigma)$ *such that* $v \in \mathrm{vars}(z\tau)$.

When $\sigma \in \mathit{VSubst}$, the following simplified characterizations for the operators occ, fvars, gvars and lvars can be used.

Proposition 42 *For each $\sigma \in \mathit{VSubst}$ and $v \in \mathit{Vars}$, we have*

$$\mathrm{occ}(\sigma, v) = \bigl\{\, y \in \mathit{Vars} \bigm| v \in \mathrm{vars}(y\sigma) \setminus \mathrm{dom}(\sigma) \,\bigr\}, \tag{27}$$

$$\mathrm{gvars}(\sigma) = \bigl\{\, y \in \mathit{Vars} \bigm| \mathrm{vars}(y\sigma) \subseteq \mathrm{dom}(\sigma) \,\bigr\}, \tag{28}$$

$$\mathrm{fvars}(\sigma) = \bigl\{\, y \in \mathit{Vars} \bigm| y\sigma \in \mathit{Vars} \setminus \mathrm{dom}(\sigma) \,\bigr\}, \tag{29}$$

$$\mathrm{lvars}(\sigma) = \left\{\, y \in \mathit{Vars} \;\middle|\; \begin{array}{l} \forall z \in \mathrm{vars}(y\sigma) \setminus \mathrm{dom}(\sigma) : \mathrm{occ\_lin}(z, y\sigma), \\ \forall z \in \mathrm{vars}(y\sigma) \cap \mathrm{dom}(\sigma) : z \in \mathrm{gvars}(\sigma) \end{array} \right\}. \tag{30}$$

Proof. Let $\#\sigma = n$. If $n = 0$, then, by Definitions 10, 12 and 14, $\mathrm{gvars}(\sigma) = \varnothing$ and $\mathrm{fvars}(\sigma) = \mathrm{lvars}(\sigma) = \mathit{Vars}$. Thus equations (28), (29) and (30) hold. We assume now that $n > 0$ and prove each equation separately.

Consider equation (28). By Definition 10, $y \in \mathrm{gvars}(\sigma)$ if and only if, for all $v \in \mathit{Vars}$, we have $y \notin \mathrm{occ}(\sigma, v)$. By Proposition 42, this holds if and only if there does not exist $v \in \mathit{Vars}$ such that $v \in \mathrm{vars}(y\sigma) \setminus \mathrm{dom}(\sigma)$, i.e., if and only if $\mathrm{vars}(y\sigma) \subseteq \mathrm{dom}(\sigma)$.

Consider equation (29). By Definition 12, $y \in \mathrm{fvars}(\sigma)$ if and only if $y\sigma^n \in \mathit{Vars}$. First note that, if $y\sigma^n \notin \mathit{Vars}$, then $y\sigma \notin \mathit{Vars} \setminus \mathrm{dom}(\sigma)$. Conversely, if $y\sigma^n \in \mathit{Vars}$, then $y\sigma \in \mathit{Vars}$ and, as $\sigma \in \mathit{VSubst}$, $y\sigma = y\sigma^n$. Thus, if $n = 1$, $y\sigma \notin \mathrm{dom}(\sigma)$ and, if $n > 1$, then $y\sigma = y\sigma^2$ and, again $y\sigma \notin \mathrm{dom}(\sigma)$.

Consider equation (30). First, suppose that, for some $z \in \mathrm{vars}(y\sigma)$, either $z \notin \mathrm{dom}(\sigma)$ and $\mathrm{occ\_lin}(z, y\sigma)$ does not hold or $z \notin \mathrm{gvars}(\sigma)$. We show that, in both cases, $y \notin \mathrm{lvars}(\sigma)$. In the first case it follows that $z \in \mathrm{vars}(y\sigma^n)$ and $\mathrm{occ\_lin}(z, y\sigma^{2n})$ does not hold. For the second case, by (28), there exists $w \in \mathrm{vars}(z\sigma) \setminus \mathrm{dom}(\sigma)$. As $\sigma \in \mathit{VSubst}$, $w \in \mathrm{vars}(y\sigma)$. As $n \geq 1$, $w \neq z$ and $z \in \mathrm{vars}(y\sigma)$, so that $\mathrm{occ\_lin}(w, y\sigma^2)$ does not hold. Thus, as $w \notin \mathrm{dom}(\sigma)$ and $n \geq 1$, $\mathrm{occ\_lin}(w, y\sigma^{2n})$ does not hold. Therefore, by Definition 14, $y \notin \mathrm{lvars}(\sigma)$.

Secondly, suppose that $y \notin \mathrm{lvars}(\sigma)$. Then, by Definition 14, there exists $z \in \mathrm{vars}(y\sigma^n) \setminus \mathrm{dom}(\sigma)$ and $\mathrm{occ\_lin}(z, y\sigma^{2n})$ does not hold. Thus, as $\sigma \in \mathit{VSubst}$, $z \in \mathrm{vars}(y\sigma)$. Also, if $\mathrm{occ\_lin}(z, y\sigma)$ holds, there must exist $v \in \mathrm{vars}(y\sigma) \cap \mathrm{dom}(\sigma)$ and $z \in \mathrm{vars}(v\sigma^{2n-1})$. Thus, as $\sigma \in \mathit{VSubst}$, $z \in \mathrm{vars}(v\sigma)$ and $\mathrm{vars}(v\sigma) \setminus \mathrm{dom}(\sigma) \neq \varnothing$ and hence, by (28), $z \notin \mathrm{gvars}(\sigma)$.

The following proposition shows that, for a substitution $\sigma \in \mathit{VSubst}$, the occurrence, groundness, freeness and linearity operators precisely capture the intended properties.

Proposition 43 *Let $\sigma \in \mathit{VSubst}$, $y \in \mathit{VI}$ and $v \in \mathit{Vars}$. Then:*

$$y \in \mathrm{occ}(\sigma, v) \iff v \in \mathrm{vars}\bigl(\mathrm{rt}(y, \sigma)\bigr), \tag{31}$$

$$y \in \mathrm{gvars}(\sigma) \iff \mathrm{rt}(y, \sigma) \in \mathit{GTerms}, \tag{32}$$

$$y \in \mathrm{fvars}(\sigma) \iff \mathrm{rt}(y, \sigma) \in \mathit{Vars}, \tag{33}$$

$$y \in \mathrm{lvars}(\sigma) \iff \mathrm{rt}(y, \sigma) \in \mathit{LTerms}. \tag{34}$$

Proof. We first prove item (31). By Proposition 42, $y \in \mathrm{occ}(\sigma, v)$ if and only if $v \in \mathrm{vars}(y\sigma) \setminus \mathrm{dom}(\sigma)$.

To prove the first implication ($\Rightarrow$), let $v \in \mathrm{vars}(y\sigma) \setminus \mathrm{dom}(\sigma)$. Then, for all $i > 0$, we have $v \in \mathrm{vars}(y\sigma^i) \setminus \mathrm{dom}(\sigma)$, so that $v \in \mathrm{vars}\bigl(\mathrm{rt}(y, \sigma)\bigr)$.

To prove the other implication ($\Leftarrow$), assume that $v \in \mathrm{vars}\bigl(\mathrm{rt}(y, \sigma)\bigr)$. We prove by contradiction that $v \in \mathrm{vars}(y\sigma) \setminus \mathrm{dom}(\sigma)$. In fact, assume that $v \notin \mathrm{vars}(y\sigma) \setminus \mathrm{dom}(\sigma)$. Then, since $\sigma \in \mathit{VSubst}$, by Definition 36 we obtain $v \notin \mathrm{vars}(y\sigma\sigma) \setminus \mathrm{dom}(\sigma)$ so that, for all $i > 0$, $v \notin \mathrm{vars}(y\sigma^i) \setminus \mathrm{dom}(\sigma)$. Thus, by definition of 'rt', $v \notin \mathrm{vars}\bigl(\mathrm{rt}(y, \sigma)\bigr)$.

We now prove item (32). By Definition 10, we have $y \in \mathrm{gvars}(\sigma)$ if and only if $y \notin \mathrm{occ}(\sigma, v)$, for all $v \in \mathit{Vars}$. By Proposition 16, this is equivalent to $v \notin \mathrm{vars}\bigl(\mathrm{rt}(y, \sigma)\bigr)$, for all $v \in \mathit{Vars}$. Thus, $\mathrm{vars}\bigl(\mathrm{rt}(y, \sigma)\bigr) = \varnothing$ and $\mathrm{rt}(y, \sigma) \in \mathit{GTerms}$.

We now prove item (33). By Proposition 42, $y \in \mathrm{fvars}(\sigma)$ if and only if $y\sigma \in \mathit{Vars} \setminus \mathrm{dom}(\sigma)$.

To prove the first implication ($\Rightarrow$), let $y\sigma \in \mathit{Vars} \setminus \mathrm{dom}(\sigma)$. Then, $\mathrm{rt}(y,\sigma) \in \mathit{Vars} \setminus \mathrm{dom}(\sigma)$ and, more generally, $\mathrm{rt}(y,\sigma) \in \mathit{Vars}$.

To prove the other implication ($\Leftarrow$), assume that $\mathrm{rt}(y,\sigma) \in \mathit{Vars}$. We prove by contradiction that $y\sigma \in \mathit{Vars} \setminus \mathrm{dom}(\sigma)$. In fact, assume that $y\sigma \notin \mathit{Vars} \setminus \mathrm{dom}(\sigma)$. We have two cases:

1. if $y\sigma \notin \mathit{Vars}$ then, by definition, $\mathrm{rt}(y,\sigma) \notin \mathit{Vars}$.

2. otherwise, let $y\sigma \in \mathrm{dom}(\sigma)$. Thus, we have $y \neq y\sigma \neq y\sigma\sigma$, so that $(y \mapsto y\sigma) \in \sigma$ and $(y\sigma \mapsto y\sigma\sigma) \in \sigma$. Since $\sigma \in \mathit{VSubst}$, we also have $\{y\sigma\} = \mathrm{vars}(y\sigma) = \mathrm{vars}(y\sigma\sigma)$. Therefore, $y\sigma\sigma \notin \mathit{Vars}$, so that there exists an $n > 0$ such that $y\sigma\sigma = f(t_1, \dots, t_n)$, and $\mathrm{size}(y\sigma\sigma) > 1$. Since $\mathrm{rt}(y,\sigma) \in \mathit{Vars}$, this is a contradiction because we also have $\mathrm{size}(y\sigma^2) \leq \mathrm{size}\bigl(\mathrm{rt}(y,\sigma)\bigr) = 1$.

Finally, we prove item (34). In order to prove the first implication ($\Rightarrow$), assume $y \in \mathrm{lvars}(\sigma)$ so that, by Proposition 42, we have

$$\forall z \in \mathrm{vars}(y\sigma) \setminus \mathrm{dom}(\sigma) : \mathrm{occ\_lin}(z, y\sigma), \tag{35}$$

$$\forall z \in \mathrm{vars}(y\sigma) \cap \mathrm{dom}(\sigma) : \mathrm{vars}(z\sigma) \subseteq \mathrm{dom}(\sigma). \tag{36}$$

We need to show that $\mathrm{rt}(y,\sigma) \in \mathit{LTerms}$ and we proceed by contradiction, negating the conclusion. Thus assume there exists $v \in \mathrm{vars}\bigl(\mathrm{rt}(y,\sigma)\bigr)$ such that $\mathrm{occ\_lin}\bigl(v, \mathrm{rt}(y,\sigma)\bigr)$ does not hold. Note that $v \notin \mathrm{dom}(\sigma)$; also, since $\sigma \in \mathit{VSubst}$, $\mathrm{vars}(y\sigma) = \mathrm{vars}(y\sigma^i)$, for all $i > 0$, so that $v \in \mathrm{vars}(y\sigma)$. If $\mathrm{occ\_lin}(v, y\sigma)$ does not hold, then we obtain the negation of equation (35), hence a contradiction. So, assume that $\mathrm{occ\_lin}(v, y\sigma)$ hold. As a consequence, there exists an index $j > 1$ such that $\mathrm{occ\_lin}(v, y\sigma^{j-1})$ holds and $\mathrm{occ\_lin}(v, y\sigma^j)$ does not hold. Thus, there exists $w \in \mathrm{vars}(y\sigma^{j-1}) \cap \mathrm{dom}(\sigma)$ such that $v \in \mathrm{vars}(w\sigma) \setminus \mathrm{dom}(\sigma)$. Since $\sigma \in \mathit{VSubst}$, $w \in \mathrm{vars}(y\sigma^{j-1})$ if and only if $w \in \mathrm{vars}(y\sigma)$. Hence $j = 2$, $w \in \mathrm{vars}(y\sigma) \cap \mathrm{dom}(\sigma)$ and $\mathrm{vars}(w\sigma) \not\subseteq \mathrm{dom}(\sigma)$. Hence we have contradicted equation (36).

To prove the other implication ($\Leftarrow$), assume $\mathrm{rt}(y,\sigma) \in \mathit{LTerms}$, so that, by definition, we have $\mathrm{occ\_lin}\bigl(z, \mathrm{rt}(y,\sigma)\bigr)$, for all $z \in \mathrm{vars}\bigl(\mathrm{rt}(y,\sigma)\bigr)$. We need to show that equations (35) and (36) hold. We proceed by contradiction, negating the conclusion. There are two cases.

1. Assume that equation (35) does not hold, i.e., there exists a variable $z \in \mathrm{vars}(y\sigma) \setminus \mathrm{dom}(\sigma)$ such that $\mathrm{occ\_lin}(z, y\sigma)$ does not hold. Then, for all $i > 0$, we have that $z \in \mathrm{vars}(y\sigma^i)$, but $\mathrm{occ\_lin}(z, y\sigma^i)$ does not hold. Hence, $\mathrm{occ\_lin}\bigl(z, \mathrm{rt}(y,\sigma)\bigr)$ does not hold and $\mathrm{rt}(y,\sigma) \notin \mathit{LTerms}$, therefore obtaining the contradiction.

2. Assume now equation (36) does not hold. Namely, there exists a variable $z \in \mathrm{vars}(y\sigma) \cap \mathrm{dom}(\sigma)$ such that $\mathrm{vars}(z\sigma) \not\subseteq \mathrm{dom}(\sigma)$. Thus, let $v \in \mathrm{vars}(z\sigma) \setminus \mathrm{dom}(\sigma)$. Since $\sigma \in \mathit{VSubst}$ and $v \in \mathrm{vars}(y\sigma\sigma)$, then $v \in \mathrm{vars}(y\sigma)$. Then, since $z \in \mathrm{vars}(y\sigma) \cap \mathrm{dom}(\sigma)$, $\mathrm{occ\_lin}(v, y\sigma\sigma)$ does not hold. Also, since $v \notin \mathrm{dom}(\sigma)$, for all $i \geq 2$, $\mathrm{occ\_lin}(v, y\sigma^i)$ does not hold. By definition, $\mathrm{occ\_lin}\bigl(v, \mathrm{rt}(y,\sigma)\bigr)$ does not hold and $\mathrm{rt}(y,\sigma) \notin \mathit{LTerms}$, obtaining the contradiction.

Proof of Proposition 16 on page 11. By Proposition 40, there exists $\tau \in \mathit{VSubst}$ such that $\sigma \iff \tau$, $\mathrm{dom}(\sigma) = \mathrm{dom}(\tau)$ and $T \vdash \forall(\sigma \leftrightarrow \tau)$. By Proposition 17, we have $\mathrm{gvars}(\sigma) = \mathrm{gvars}(\tau)$, $\mathrm{fvars}(\sigma) = \mathrm{fvars}(\tau)$ and $\mathrm{lvars}(\sigma) = \mathrm{lvars}(\tau)$. From all of the above, by Proposition 43, we obtain

$$
\begin{aligned}
y \in \mathrm{fvars}(\sigma) &\iff \mathrm{rt}(y, \tau) \in \mathit{Vars}, \\
y \in \mathrm{lvars}(\sigma) &\iff \mathrm{rt}(y, \tau) \in \mathit{LTerms}.
\end{aligned}
$$

Thus, the equivalences (13), (14) and (15) follow from Proposition 2.

In order to prove one of the statements of Proposition 17, in particular, to show that the linearity operators precisely capture the intended properties even for arbitrary substitutions in *RSubst*, we prove two preliminary results.

Lemma 44 *Let T be a syntactic equality theory, $\sigma, \tau \in \mathit{RSubst}$ and $s, t \in \mathit{Terms}$ where $T \vdash \forall(\sigma \leftrightarrow \tau)$, $T \vdash \forall(\sigma \to s = t)$, $\mathrm{dom}(\sigma) = \mathrm{dom}(\tau)$ and $\#\tau = n$. Then, if $z \in \mathrm{vars}(s) \setminus \mathrm{dom}(\sigma)$, $z \in \mathrm{vars}(t\tau^n)$.*

Proof. Suppose that $z \in \mathrm{vars}(s) \setminus \mathrm{dom}(\sigma)$, then we show that, by induction on the depth of s, that $z \in \mathrm{vars}(t\tau^n)$.

Suppose first that $s = z$. Then $\mathrm{rt}(s, \sigma) = z$. By Lemma 35, $\mathrm{rt}(s, \sigma) = \mathrm{rt}(t, \sigma)$ and hence $\mathrm{rt}(t, \sigma) = z$. Thus, for all $i \in \mathbb{N}$, $t\sigma^i \in \mathit{Vars}$ and for some $j \in \mathbb{N}$, $t\sigma^j = z$. Since $\tau \in \mathit{RSubst}$, σ has no circular subsets, $t\sigma^n = z$.

Secondly, suppose that $s = f(s_1, \ldots, s_m)$ and $z \in \mathrm{vars}(s_j)$, for some $j \in \{1, \ldots, m\}$. We first show that $t\tau^n \notin \mathit{Vars}$. Suppose, by contraposition that $t\tau^n \in \mathit{Vars}$, then as τ contains no circular subsets, $t\tau^n \in \mathit{Vars} \setminus \mathrm{dom}(\tau)$. Then, as $\mathrm{dom}(\sigma) = \mathrm{dom}(\tau)$, $\mathrm{rt}(t\tau^n, \sigma) \in \mathit{Vars}$. By Lemma 33, we have $T \vdash \forall\bigl(\sigma \to (s = t\tau^n)\bigr)$. Thus, by Lemma 35, $\mathrm{rt}(s, \sigma) = \mathrm{rt}(t\tau^n, \sigma)$ so that $\mathrm{rt}(s, \sigma) \in \mathit{Vars}$ contradicting the assumption that $s \notin \mathit{Vars}$. Thus we must have $t\tau^n = f(t_1, \ldots, t_m)$, for some $t_1, \ldots, t_m \in \mathit{Terms}$ where, for each $i = 1, \ldots, n$, $T \vdash \forall\bigl(\sigma \to (s_i = t_i)\bigr)$. Hence, by the inductive hypothesis, $z \in \mathrm{vars}(t_j\tau^n)$. It follows that $z \in \mathrm{vars}(t\tau^{2n})$. This means that $\mathrm{occ}(z, \tau) \cap \mathrm{vars}(t) \neq \varnothing$ and, therefore $\mathrm{occ}_n(z, \tau) \cap \mathrm{vars}(t) \neq \varnothing$. Thus $z \in \mathrm{vars}(t\tau^{2n})$.

Lemma 45 *Let T be a syntactic equality theory, $\sigma, \tau \in \mathit{RSubst}$, $s, t \in \mathit{Terms}$ and $z \in \mathit{Vars} \setminus \mathrm{dom}(\sigma)$ where $T \vdash \forall(\sigma \leftrightarrow \tau)$, $T \vdash \forall(\sigma \to s = t)$, $\mathrm{dom}(\sigma) = \mathrm{dom}(\tau)$ and $\#\tau = n$. Then, if $z \in \mathrm{vars}(s) \cap \mathrm{vars}(t)$ and $\mathrm{occ\_lin}(z, s\sigma^n)$ does not hold, then $\mathrm{occ\_lin}(z, t\tau^n)$ does not hold*

Proof. Suppose that $z \in \mathrm{vars}(s) \cap \mathrm{vars}(t)$ and $\mathrm{occ\_lin}(s, \sigma^n)$ does not hold. We show that, by induction on the depth of s, that $\mathrm{occ\_lin}(t, \tau^n)$ does not hold.

Suppose first that $s = z$. Then $\mathrm{occ\_lin}(z, s\sigma^n)$ does not hold and the result follows.

Secondly, suppose that $s = f(s_1, \ldots, s_m)$ and $z \in \mathrm{vars}(s_j)$, for some $j \in \{1, \ldots, m\}$. We first show that $t \notin \mathit{Vars}$. Suppose, by contraposition that $t \in \mathit{Vars}$, then as $z \in \mathrm{vars}(t)$, $t = z$. Then, as $\mathrm{dom}(\sigma) = \mathrm{dom}(\tau)$, $\mathrm{rt}(t, \sigma) = z$. Thus, by Lemma 35, $\mathrm{rt}(s, \sigma) = \mathrm{rt}(t, \sigma) = z$ so that $\mathrm{rt}(s, \sigma) \in \mathit{Vars}$ contradicting the assumption that $s \notin \mathit{Vars}$. Thus we must have $t = f(t_1, \ldots, t_m)$, for some $t_1, \ldots, t_m \in \mathit{Terms}$ where, for each $i = 1, \ldots, n$, $T \vdash \forall\bigl(\sigma \to (s_i = t_i)\bigr)$. By Lemma 44, $z \in \mathrm{vars}(t_j)$ If $\mathrm{occ\_lin}(z, s_j\sigma^n$ does not hold, then, by the inductive hypothesis, $\mathrm{occ\_lin}(z, t_j\tau^n$ does not hold and hence, $\mathrm{occ\_lin}(z, t\tau^n$ does not hold.

Otherwise, there must exist $k \in \{1, \dots, m\}$ where $k \neq j$ such that $z \in \mathrm{vars}(s_k)$. By Lemma 44, $z \in \mathrm{vars}(t_k)$ As $k \neq j$, $\mathrm{occ\_lin}(t\tau^n)$.

The next three results are needed in the proof of Proposition 17 below. The first Lemma is a consequence of the proof of [22, Lemma 15]. The other two results correspond to [22, Theorems 1 and 2].

Lemma 46 *Let $\sigma, \sigma' \in RSubst$ be such that $\sigma \overset{\mathcal{S}}{\longmapsto}{}^{*} \sigma'$. Then, for all $v \in Vars$, we have* $\mathrm{occ}(\sigma, v) = \mathrm{occ}(\sigma', v)$.

Theorem 47 *Let T be an equality theory, $\sigma \in RSubst$ and $\sigma \overset{\mathcal{S}}{\longmapsto}{}^{*} \sigma'$. Then $\sigma' \in RSubst$,* $\mathrm{vars}(\sigma) = \mathrm{vars}(\sigma')$, $\mathrm{dom}(\sigma) = \mathrm{dom}(\sigma')$ *and* $T \vdash \forall(\sigma \leftrightarrow \sigma')$.

Theorem 48 *Let $\sigma \in RSubst$. Then there exists $\sigma' \in VSubst$ such that $\sigma \overset{\mathcal{S}}{\longmapsto}{}^{*} \sigma'$ and $y \in \mathrm{dom}(\sigma') \cap \mathrm{range}(\sigma')$ implies $y \in \mathrm{vars}(y\sigma')$.*

Proof of Proposition 17 on page 12. Let $n = \#\sigma$. Then, as $\mathrm{dom}(\sigma) = \mathrm{dom}(\tau)$, $n = \#\tau$.

Consider first (16). By Theorem 48, there exists $\tau \in VSubst$ such that $\sigma \overset{\mathcal{S}}{\longmapsto}{}^{*} \tau$. By Theorem 47, $\mathrm{dom}(\sigma) = \mathrm{dom}(\tau)$ and $T \vdash \forall(\sigma \leftrightarrow \tau)$. By Lemma 46, $\mathrm{occ}(\sigma, v) = \mathrm{occ}(\tau, v)$. Moreover, by Proposition 43, we have $y \in \mathrm{occ}(\tau, v)$ if and only if $v \in \mathrm{vars}\bigl(\mathrm{rt}(y, \tau)\bigr)$. Thus, we have $y \in \mathrm{occ}(\sigma, v)$ if and only if $v \in \mathrm{vars}\bigl(\mathrm{rt}(y, \tau)\bigr)$ and, to complete the proof, it is sufficient to show that

$$v \in \mathrm{vars}\bigl(\mathrm{rt}(y, \sigma)\bigr) \quad \iff \quad v \in \mathrm{vars}\bigl(\mathrm{rt}(y, \tau)\bigr).$$

We only prove the first implication, since the other one follows by symmetry.

Suppose $v \in \mathrm{vars}\bigl(\mathrm{rt}(y, \sigma)\bigr)$. Then there exists an index $i \geq 0$ such that $v \in \mathrm{vars}(y\sigma^i)$. Note that $v \notin \mathrm{dom}(\sigma) = \mathrm{dom}(\tau)$, so that $v \in \mathrm{vars}\bigl(\mathrm{rt}(y\sigma^i, \tau)\bigr)$. Since $T \vdash \forall(\tau \rightarrow \sigma)$, by Lemma 33 we have $T \vdash \forall\bigl(\tau \rightarrow (y = y\sigma^i)\bigr)$. By Lemma 35, $\mathrm{rt}(y, \tau) = \mathrm{rt}(y\sigma^i, \tau)$. Thus, $v \in \mathrm{vars}\bigl(\mathrm{rt}(y, \tau)\bigr)$.

Consider (17). By Definition 10, this is a simple consequence of (16) proved above.

Consider (18). Suppose that $y \in \mathrm{fvars}(\sigma)$. Then, by Definition 12, $y\sigma^n \in Vars \setminus \mathrm{dom}(\sigma)$ and hence, $y\sigma^n \in Vars \setminus \mathrm{dom}(\tau)$. By Lemma 33, we have $T \vdash \forall\bigl(\tau \rightarrow (y = y\sigma^n)\bigr)$. Thus, by Lemma 35, $\mathrm{rt}(y, \tau) = \mathrm{rt}(y\sigma^n, \tau) = y\sigma^n$. Thus $\mathrm{rt}(y, \tau) \in Vars \setminus \mathrm{dom}(\tau)$. Thus $y\tau^i \in Vars$ for all $i \in \mathbb{N}$. As $\tau \in RSubst$, there are no circular subsets of τ so that $y\tau^n \notin \mathrm{dom}(\tau)$ so that, by Definition 12, $y \in \mathrm{fvars}(\tau)$.

Consider (18). Suppose $y \notin \mathrm{lvars}(\sigma)$, Then there exists $z \in \mathrm{vars}(y\sigma^n)$ such that $\mathrm{occ\_lin}(z, y\sigma^{2n})$ does not hold. By Lemma 44, $z \in \mathrm{vars}(y\tau^n)$. As $T \vdash \forall\bigl(\tau \rightarrow (y\sigma^n = y\tau^n)\bigr)$, by Lemma 45, $\mathrm{occ\_lin}(z, y\tau^{2n})$ does not hold.

C Proofs of the Results of Subsection 3.2

In this section we prove that the abstract unification operator defined on the domain *SFL* is a correct approximation of the concrete unification procedure. As already discussed, this result applies to any syntactic equality theory, including both the finite-tree theory $\mathcal{FT}$ and the rational-tree theory $\mathcal{RT}$. In contrast, all the published proofs of correctness for a combination of set-sharing with freeness and linearity information assume the occurs-check is performed. As a consequence, they do not apply to the vast majority of implemented logic languages.

First we recall the definition of the abstract mgu operator on SH and the corresponding correctness result, as provided in [22].

Definition 49 (amgu.) *The function* amgu: $SH \times Bind \to SH$ *captures the effects of a binding on an* SH *element. Suppose* $x \in Vars$, $r \in HTerms$, *and* $sh \in SH$. *Let* $v_x \stackrel{\mathrm{def}}{=} \{x\}$, $v_r \stackrel{\mathrm{def}}{=} \mathrm{vars}(r)$ *and* $v_{xr} \stackrel{\mathrm{def}}{=} v_x \cup v_r$. *Then*

$$\mathrm{amgu}(sh, x \mapsto r) \stackrel{\mathrm{def}}{=} \overline{\mathrm{rel}}(v_{xr}, sh) \cup \mathrm{bin}\bigl(\mathrm{rel}(v_x, sh)^\star, \mathrm{rel}(v_r, sh)^\star\bigr).$$

Theorem 50 *Let* $sh \in SH$ *and* $(x \mapsto t) \in Bind$, *where* $\{x\} \cup \mathrm{vars}(t) \subseteq VI$. *Let* $\sigma \in RSubst$ *be such that* $\mathrm{ssets}(\sigma) \subseteq sh$ *and suppose that* $\mu \in \mathrm{mgs}\bigl(\{x = t\} \cup \sigma\bigr)$ *in the syntactic equality theory* T. *Then*

$$\mathrm{ssets}(\mu) \subseteq \mathrm{amgu}(sh, x \mapsto t).$$

We now prove Theorem 21, stating that the auxiliary operators introduced in Definition 20 correctly approximate the intended properties.

Proof of Theorem 21 on page 13. Let $d = \langle sh, f, l \rangle$, $V_s = \mathrm{vars}(s)$ and $V_t = \mathrm{vars}(t)$. By Definition 18, we have $\forall v \in Vars : \mathrm{occ}(\sigma, v) \cap VI \in sh \cup \{\varnothing\}$, $f \subseteq \mathrm{fvars}(\sigma)$, and $l \subseteq \mathrm{lvars}(\sigma)$.

Consider the implication (20). By Definition 20, the hypothesis and Proposition 16, we have

$$\begin{aligned}
\mathrm{ind}_d(s,t) &\iff \mathrm{rel}(V_s, sh) \cap \mathrm{rel}(V_t, sh) = \varnothing \\
&\iff \forall S \in sh, w_1 \in V_s, w_2 \in V_t : \{w_1, w_2\} \nsubseteq S \\
&\implies \forall v \in Vars, w_1 \in V_s, w_2 \in V_t : \{w_1, w_2\} \nsubseteq \mathrm{occ}(\sigma, v) \\
&\iff \forall w_1 \in V_s, w_2 \in V_t : \mathrm{vars}\bigl(\mathrm{rt}(w_1, \sigma)\bigr) \cap \mathrm{vars}\bigl(\mathrm{rt}(w_1, \sigma)\bigr) = \varnothing \\
&\iff \mathrm{vars}\bigl(\mathrm{rt}(s, \sigma)\bigr) \cap \mathrm{vars}\bigl(\mathrm{rt}(t, \sigma)\bigr) = \varnothing.
\end{aligned}$$

Consider the equivalence (21). By Definition 20, we have

$$\begin{aligned}
\mathrm{ind}_d(y,t) &\iff \mathrm{rel}\bigl(\{y\}, sh\bigr) \cap \mathrm{rel}(V_t, sh) = \varnothing \\
&\iff \forall S \in \mathrm{rel}(V_t, sh) : y \notin S \\
&\iff y \notin \mathrm{share_with}_d(t).
\end{aligned}$$

Consider now the implication (22). By Definition 20, the hypothesis and Proposition 16, we have

$$\begin{aligned}
\mathrm{free}_d(t) &\iff t \in f \\
&\implies t \in \mathrm{fvars}(\sigma) \\
&\iff \mathrm{rt}(t, \sigma) \in Vars.
\end{aligned}$$

Consider now the implication (23). By Definition 20, the hypothesis and Proposition 16, we have

$$\mathrm{ground}_d(t) \iff \mathrm{vars}(t) \subseteq VI \setminus \mathrm{vars}(sh)$$

$$
\begin{aligned}
&\implies \forall w \in V_t : w \in \mathrm{gvars}(\sigma) \\
&\iff \forall w \in V_t : \mathrm{vars}\bigl(\mathrm{rt}(w, \sigma)\bigr) = \varnothing \\
&\iff \mathrm{rt}(t, \sigma) \in \mathit{GTerms}.
\end{aligned}
$$

Finally consider the implication (24). By Definition 20, the hypothesis, the above results and Proposition 16, we have

$$
\begin{aligned}
\mathrm{lin}_d(t) &\iff \forall y, z \in \mathrm{vars}(t) : \mathrm{ground}_d(y) \\
&\qquad \lor \Bigl((y \in l) \land \mathrm{occ\_lin}(y, t) \land \bigl(y \neq z \implies \mathrm{ind}_d(y, z)\bigr) \Bigr) \\
&\implies \forall y, z \in \mathrm{vars}(t) : \mathrm{ground}_d(y) \\
&\qquad \lor \Bigl(\bigl(y \in \mathrm{lvars}(\sigma)\bigr) \land \mathrm{occ\_lin}(y, t) \land \bigl(y \neq z \implies \mathrm{ind}_d(y, z)\bigr) \Bigr) \\
&\implies \forall y, z \in \mathrm{vars}(t) : \mathrm{rt}(y, \sigma) \in \mathit{GTerms} \\
&\qquad \lor \Bigl(\bigl(\mathrm{rt}(y, \sigma) \in \mathit{LTerms}\bigr) \land \mathrm{occ\_lin}(y, t) \\
&\qquad\qquad \land \bigl(y \neq z \implies \mathrm{vars}(\mathrm{rt}(y, \sigma)) \cap \mathrm{vars}(\mathrm{rt}(z, \sigma)) = \varnothing\bigr) \Bigr) \\
&\iff \mathrm{rt}(t, \sigma) \in \mathit{LTerms}.
\end{aligned}
$$

The following simple lemma will be systematically used in the following correctness proofs.

Lemma 51 *Assume T is an equality theory and $\sigma \in \mathit{RSubst}$. Then, for each $s, t \in \mathit{HTerms}$,*

$$
\mathrm{mgs}\bigl(\sigma \cup \{s = t\}\bigr) = \mathrm{mgs}\bigl(\sigma \cup \{s = t\sigma\}\bigr).
$$

Proof. First, note, using the congruence axioms (6) and (7), that, for any terms $p, q, r \in \mathit{HTerms}$,

$$
T \vdash \forall (p = q \land q = r) \leftrightarrow \forall (p = r \land q = r).
$$

Secondly note that, using Lemma 33, for any substitution $\tau \in \mathit{RSubst}$ and term $r \in \mathit{HTerms}$, $T \vdash \forall\bigl(\tau \to (r = r\tau)\bigr)$, so that

$$
T \vdash \forall\bigl(\tau \leftrightarrow \tau \cup \{r = r\tau\}\bigr).
$$

Using these results,

$$
\begin{aligned}
&T \vdash \forall\bigl(\sigma \cup \{s = t\} \leftrightarrow \sigma \cup \{s = t, t = t\sigma\}\bigr), \\
&T \vdash \forall\bigl(\sigma \cup \{s = t\} \leftrightarrow \sigma \cup \{s = t\sigma, t = t\sigma\}\bigr), \\
&T \vdash \forall\bigl(\sigma \cup \{s = t\} \leftrightarrow \sigma \cup \{s = t\sigma\}\bigr).
\end{aligned}
$$

The thesis follows by the definition of mgs.

The following Lemma, which will be used several times in the following proofs without an explicit reference to it, states the well-known result that groundness is closed by entailment.

Lemma 52 *Let $\sigma, \tau \in \mathit{RSubst}$ be satisfiable in the syntactic equality theory T and such that $T \vdash \forall(\tau \to \sigma)$. Then $\mathrm{gvars}(\sigma) \subseteq \mathrm{gvars}(\tau)$.*

Proof. We prove the result by showing that $x \notin \text{gvars}(\tau)$ implies $x \notin \text{gvars}(\sigma)$.

By Proposition 40, we can assume there exist $\sigma', \tau' \in \mathit{VSubst}$ such that $T \vdash \forall(\sigma \leftrightarrow \sigma')$ and $T \vdash \forall(\tau \leftrightarrow \tau')$, so that $T \vdash \forall(\tau' \to \sigma')$. Also, by Proposition 17, we have $\text{gvars}(\sigma) = \text{gvars}(\sigma')$ and $\text{gvars}(\tau) = \text{gvars}(\tau')$. Therefore, it is sufficient to prove that $x \notin \text{gvars}(\tau')$ implies $x \notin \text{gvars}(\sigma')$.

Assume $x \notin \text{gvars}(\tau')$. By Definition 10, there exists $v \in \mathit{Vars}$ such that $x \in \text{occ}(\tau', v)$. By Proposition 42, $v \in \text{vars}(x\tau') \setminus \text{dom}(\tau')$. Also, by Lemma 33, $T \vdash \forall(\tau' \to x\tau' = x)$. Therefore, by Lemma 41 (taking $s = x\tau'$ and $t = x$) there exists $z \in \text{vars}(x\sigma') \setminus \text{dom}(\sigma')$ such that $v \in \text{vars}(z\tau')$. By Definition 8, we have $x \in \text{occ}(\sigma', z)$ so that, by Definition 10, $x \notin \text{gvars}(\sigma')$.

Another useful result is the following.

Lemma 53 *Let $e \subseteq \mathit{Eqs}$ be satisfiable in the syntactic equality theory T. If $\sigma, \tau \in \text{mgs}(e)$, then $\text{ssets}(\sigma) = \text{ssets}(\tau)$, $\text{fvars}(\sigma) = \text{fvars}(\tau)$, $\text{gvars}(\sigma) = \text{gvars}(\tau)$ and $\text{lvars}(\sigma) = \text{lvars}(\tau)$.*

Proof. By definition of mgs, we have $\sigma, \tau \in \mathit{RSubst}$ and $\sigma \iff e \iff \tau$. Thus, all the above equivalences follow from Proposition 17.

We now introduce a bit of terminology that will be helpful in order to simplify the notation in the following proofs.

Given $V \subseteq \mathit{Vars}$, we say that $t \in \mathit{HTerms}$ is V-linear if $\text{occ\_lin}(v, t)$ holds for all variables $v \in \text{vars}(t) \cap V$. Note that if a term is V-linear, then it is also W-linear, for all $W \subseteq V$. This terminology also applies to n-tuples of terms, by simply regarding the n-tuple construction as a term functor of arity n. Moreover, if $\bar{s}, \bar{t} \in \mathit{HTerms}^n$ are such that $\text{mgs}(\bar{s} = \bar{t}) \neq \varnothing$, then we write $\text{gvars}(\bar{s} = \bar{t})$ to denote the set $\text{gvars}(\mu)$, where $\mu \in \text{mgs}(\bar{s} = \bar{t})$. Note that, by Lemma 53, this notation is not ambiguous.

Lemma 54 *Let $\bar{s}, \bar{t} \in \mathit{HTerms}^n$ be such that $\text{mgs}(\bar{s} = \bar{t}) \neq \varnothing$. Let also $G \stackrel{\text{def}}{=} \text{gvars}(\bar{s} = \bar{t})$. If $\bar{s}$ is $(\mathit{Vars} \setminus G)$-linear, then there exists $\mu \in \text{mgs}(\bar{s} = \bar{t})$ such that, for each $z \in \mathit{Vars} \setminus \text{vars}(\bar{s})$,*

1. *$z\mu$ is $(\mathit{Vars} \setminus G)$-linear;*
2. *$\text{vars}(z\mu) \cap \text{dom}(\mu) \subseteq G$;*
3. *$\forall z' \in \mathit{Vars} \setminus \text{vars}(\bar{s}) : z \neq z' \implies \text{vars}(z\mu) \cap \text{vars}(z'\mu) \subseteq G$.*

Proof. We assume that the congruence and identity axioms hold.

Let $\bar{s} = (s_1, \ldots, s_n)$, and $\bar{t} = (t_1, \ldots, t_n)$, $W = \text{vars}(\bar{s}) \cup \text{vars}(\bar{t})$ and $V = \mathit{Vars} \setminus G$. We assume that $\bar{s}$ is V-linear and prove that the result holds by induction on the number of variables in W.

Suppose first that, for some $i = 1, \ldots, n$, we have $s_i = f(r_1, \ldots, r_m)$ and $t_i = f(u_1, \ldots, u_m)$, where $m \geq 0$. Let

$$\bar{s}_i \stackrel{\text{def}}{=} (s_1, \ldots, s_{i-1}, r_1, \ldots, r_m, s_{i+1}, \ldots, s_n),$$
$$\bar{t}_i \stackrel{\text{def}}{=} (t_1, \ldots, t_{i-1}, u_1, \ldots, u_m, t_{i+1}, \ldots, t_n).$$

Then $\text{mvars}(\bar{s}_i) = \text{mvars}(\bar{s})$ and $\text{mvars}(\bar{t}_i) = \text{mvars}(\bar{t})$ so that, since $\bar{s}$ is V-linear, $\bar{s}_i$ is V-linear. Moreover, by the congruence axiom (8), we have $\text{mgs}(\bar{s}_i =$

$\bar{t}_i) = \mathrm{mgs}(\bar{s} = \bar{t})$. (Note that in the case that s_i and t_i are identical constants, the equation $s_i = t_i$ is just removed.) Thus, as $\bar{s}$ and $\bar{t}$ are finite sequences of finite terms, we can assume that $s_i \in \mathit{Vars}$ or $t_i \in \mathit{Vars}$, for all $i = 1, \ldots, n$.

Secondly, suppose that for some $i = 1, \ldots, n$, $s_i = t_i$. By the previous paragraph, we can assume that $s_i \in \mathit{Vars}$. Let

$$\bar{s}_i \stackrel{\mathrm{def}}{=} (s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_n),$$
$$\bar{t}_i \stackrel{\mathrm{def}}{=} (t_1, \ldots, t_{i-1}, t_{i+1}, \ldots, t_n).$$

Then $\mathrm{mvars}(\bar{s}_i) \cup \{s_i\} = \mathrm{mvars}(\bar{s})$ and $\mathrm{mvars}(\bar{t}_i) \cup \{s_i\} = \mathrm{mvars}(\bar{t})$ so that, as $\bar{s}$ is V-linear, $\bar{s}_i$ is V-linear. Furthermore, by the congruence axiom (5), $\mathrm{mgs}(\bar{s}_i = \bar{t}_i) = \mathrm{mgs}(\bar{s} = \bar{t})$. Thus, as $\bar{s}$ and $\bar{t}$ are sequences of finite length n, we can assume that $s_i \neq t_i$, for all $i = 1, \ldots, n$.

Therefore, for the rest of the proof, we will assume that $s_i \neq t_i$ and $s_i \in \mathit{Vars}$ or $t_i \in \mathit{Vars}$, for all $i = 1, \ldots, n$.

The base case is when $W = \varnothing$, so that we have $\mathrm{vars}(\mu) = \varnothing$ for all $\mu \in \mathrm{mgs}(\bar{s} = \bar{t})$. Thus all three properties hold trivially.

To prove the inductive step, we assume that $W \neq \varnothing$, so that $n > 0$. Note that, in the case that $\mathrm{vars}(\bar{t}) \subseteq \mathrm{vars}(\bar{s})$, then all three properties hold trivially. This is because for all $\mu \in \mathrm{mgs}(\bar{s} = \bar{t})$ and for all $z \in \mathit{Vars} \setminus \mathrm{vars}(\bar{s})$, we have $z \notin \mathrm{vars}(\mu)$. Similarly, the three properties hold trivially whenever $\mathrm{vars}(\bar{t}) \subseteq G$. This is because, if $z \in \mathrm{dom}(\mu)$, then $\mathrm{vars}(z\mu) \subseteq G$. We therefore assume for the rest of the proof that for some $i = 1, \ldots, n$, $\mathrm{vars}(t_i) \setminus \bigl(\mathrm{vars}(\bar{s}) \cup G\bigr) \neq \varnothing$. As the order of equations is irrelevant, without loss of generality we assume that this property holds when $i = 1$, so that $\mathrm{vars}(t_1) \setminus \bigl(\mathrm{vars}(\bar{s}) \cup G\bigr) \neq \varnothing$. This can be re-written as

$$\mathrm{vars}(t_1) \cap \bigl(V \setminus \mathrm{vars}(\bar{s})\bigr) \neq \varnothing. \tag{37}$$

Note that this implies that $t_1 \neq s_1$. By Proposition 16, another consequence of the above assumption is that, for all $\mu \in \mathrm{mgs}(\bar{s} = \bar{t})$, we have $\mathrm{rt}(t_1, \mu) \notin \mathit{GTerms}$. Since $\mu \implies \{s_1 = t_1\}$, by Lemma 35, we obtain $\mathrm{rt}(s_1, \mu) \notin \mathit{GTerms}$, so that again by Proposition 16, $\mathrm{vars}(s_1) \setminus G \neq \varnothing$. This in turn can be re-written as

$$\mathrm{vars}(s_1) \cap V \neq \varnothing. \tag{38}$$

By exploiting (37) and (38), we can identify three different cases:

a. for all $i = 1, \ldots, n$, $V \cap \mathrm{vars}(s_i) \cap \mathrm{vars}(t_i) \neq \varnothing$;

b. $s_1 \in V \setminus \mathrm{vars}(t_1)$;

c. $t_1 \in V \setminus \mathrm{vars}(\bar{s})$ and $s_1 \notin \mathit{Vars}$;

Case a. For all $i = 1, \ldots, n$, $V \cap \mathrm{vars}(s_i) \cap \mathrm{vars}(t_i) \neq \varnothing$.

For each $i = 1, \ldots, n$, we are assuming that $s_i \in V$ or $t_i \in V$. Therefore, for each $i = 1, \ldots, n$, $s_i \in \mathrm{vars}(t_i)$ or $t_i \in \mathrm{vars}(s_i)$ so that, without loss of generality, we can assume, for some k where $0 \leq k \leq n$, $s_i \in V$ if $1 \leq i \leq k$ and $t_i \in V$ if $k + 1 \leq i \leq n$.

Let $\mu \subseteq \mathit{Eqs}$ be defined as

$$\mu \stackrel{\mathrm{def}}{=} \{s_1 = t_1, \ldots, s_k = t_k\} \cup \{t_{k+1} = s_{k+1}, \ldots, t_n = s_n\}.$$

We show that $\mu \in \mathrm{mgs}(\bar{s} = \bar{t})$. First we must show that $\mu \in \mathit{RSubst}$. As $\bar{s}$ is V-linear, $(s_1, \ldots, s_k)$ is linear; $(t_{k+1}, \ldots, t_n)$ is also linear, because $\bar{s}$ is V-linear and $t_i \in V \cap \mathrm{vars}(s_i)$ if $k+1 \leq i \leq n$; moreover, for the same reasons, $\{s_1, \ldots, s_k\} \cap \{t_{k+1}, \ldots, t_n\} = \varnothing$. As we are assuming that, for all $i = 1, \ldots, n$, $s_i \neq t_i$ and $V \cap \mathrm{vars}(s_i) \cap \mathrm{vars}(t_i) \neq \varnothing$, it follows that $t_i \notin \mathit{Vars}$ when $1 \leq i \leq k$ and $s_i \notin \mathit{Vars}$ when $k+1 \leq i \leq n$, so that each equation in μ is a binding and μ has no circular subsets. Thus $\mu \in \mathit{RSubst}$ and hence, by the congruence axiom (6), $\mu \in \mathrm{mgs}(\bar{s} = \bar{t})$.

As $\{t_{k+1}, \ldots, t_n\} \subseteq \mathrm{vars}\bigl((s_{k+1}, \ldots, s_n)\bigr)$, we have $\mathrm{dom}(\mu) \setminus \mathrm{vars}(\bar{s}) = \varnothing$ so that the required result holds trivially.

Case b. Suppose $s_1 \in V \setminus \mathrm{vars}(t_1)$ Let

$$\begin{aligned} \bar{s}_1 &\stackrel{\mathrm{def}}{=} (s_2, \ldots, s_n), \\ \bar{t}_1 &\stackrel{\mathrm{def}}{=} \bigl(t_2[s_1/t_1], \ldots, t_n[s_1/t_1]\bigr). \end{aligned}$$

As $\bar{s}$ is V-linear, $\bar{s}_1$ is V-linear and $s_1 \notin \mathrm{vars}(\bar{s}_1)$. Also, all occurrences of s_1 in $\bar{t}$ are replaced in $\bar{t}_1$ by t_1 so that, as $s_1 \notin \mathrm{vars}(t_1)$ (by the assumption for this case), $s_1 \notin \mathrm{vars}(\bar{t}_1)$. Thus,

$$s_1 \notin W_1 \stackrel{\mathrm{def}}{=} \mathrm{vars}(\bar{s}_1) \cup \mathrm{vars}(\bar{t}_1), \tag{39}$$

so that $W_1 \subset W$. Let $G_1 \stackrel{\mathrm{def}}{=} \mathrm{gvars}(\bar{s}_1 = \bar{t}_1)$ and $V_1 \stackrel{\mathrm{def}}{=} \mathit{Vars} \setminus G_1$. Note that $G_1 \subseteq G$ and, by the assumption for this case, $s_1 \in V$, so that $s_1 \notin G$. As a consequence, $G_1 = G$, $V_1 = V$ and $\bar{s}_1$ is V_1-linear, so that the inductive hypothesis applies to $\bar{s}_1$ and $\bar{t}_1$. Thus, there exists $\mu_1 \in \mathrm{mgs}(\bar{s}_1 = \bar{t}_1)$ such that, for each $z \in \mathit{Vars} \setminus \mathrm{vars}(\bar{s_1})$, the three inductive properties hold.

Let $\mu \subseteq \mathit{Eqs}$ be defined as

$$\mu \stackrel{\mathrm{def}}{=} \{s_1 = t_1\mu_1\} \cup \mu_1.$$

We show that $\mu \in \mathrm{mgs}(\bar{s} = \bar{t})$. By (39), we have $s_1 \notin \mathrm{vars}(\mu_1)$ so that $s_1 \notin \mathrm{dom}(\mu_1)$. Also, since $\mu_1 \in \mathit{RSubst}$, μ has no identities or circular subsets. Thus we have $\mu \in \mathit{RSubst}$. By Lemma 51, $\mu \in \mathrm{mgs}(\bar{s} = \bar{t})$.

Suppose that $z \in \mathit{Vars} \setminus \mathrm{vars}(\bar{s})$. Then, as $\mathrm{vars}(\bar{s}) = \mathrm{vars}(\bar{s}_1) \cup \{s_1\}$, $z \in \mathit{Vars} \setminus \mathrm{vars}(\bar{s_1})$. Thus, the inductive properties 1, 2 and 3 using μ_1 and $\bar{s}_1$ can be applied to $z\mu_1$. Knowing this, we now show that the same properties using μ and $\bar{s}$ can be applied to $z\mu$. Since $\mathrm{dom}(\mu) = \mathrm{dom}(\mu_1) \cup \{s_1\}$ and $z \neq s_1$, we have $z\mu_1 = z\mu$ and $s_1 \notin \mathrm{vars}(z\mu)$. Each property is proved separately.

1. By the inductive property 1, we have $z\mu_1$ is V_1-linear. As $z\mu = z\mu_1$ and $V = V_1$, $z\mu$ is V-linear.

2. By inductive property 2, we have $\mathrm{vars}(z\mu_1) \cap \mathrm{dom}(\mu_1) \subseteq G_1$. Since $z\mu = z\mu_1$, $G = G_1$, $\mathrm{dom}(\mu) = \mathrm{dom}(\mu_1) \cup \{s_1\}$ and $s_1 \notin \mathrm{vars}(z\mu)$, we obtain $\mathrm{vars}(z\mu) \cap \mathrm{dom}(\mu) \subseteq G$.

3. Let $z' \in \mathit{Vars} \setminus \mathrm{vars}(\bar{s})$ be such that $z \neq z'$. Since $z' \notin \mathrm{vars}(\bar{s})$, we have $z' \notin \mathrm{vars}(\bar{s}_1)$ and $z'\mu = z'\mu_1$. By applying inductive property 3, $\mathrm{vars}(z\mu_1) \cap \mathrm{vars}(z'\mu_1) \subseteq G_1$. As $z\mu = z\mu_1$, $z'\mu = z'\mu_1$ and $G = G_1$, we obtain $\mathrm{vars}(z\mu) \cap \mathrm{vars}(z'\mu) \subseteq G$.

Case c. Assume that $t_1 \in V \setminus \mathrm{vars}(\bar{s})$ and $s_1 \notin \mathit{Vars}$. Let

$$\bar{s}_1 \stackrel{\mathrm{def}}{=} (s_2, \ldots, s_n),$$
$$\bar{t}_1 \stackrel{\mathrm{def}}{=} \bigl(t_2[t_1/s_1], \ldots, t_n[t_1/s_1]\bigr).$$

As $\bar{s}$ is V-linear, $\bar{s}_1$ is V-linear. Also, since by the assumption for this case $t_1 \notin \mathrm{vars}(\bar{s})$, we have $t_1 \notin \mathrm{vars}(\bar{s}_1)$. Moreover, all occurrences of t_1 in $\bar{t}$ are replaced in $\bar{t}_1$ by s_1 so that $t_1 \notin \mathrm{vars}(\bar{t}_1)$. Thus

$$t_1 \notin W_1 \stackrel{\mathrm{def}}{=} \mathrm{vars}(\bar{s}_1) \cup \mathrm{vars}(\bar{t}_1), \tag{40}$$

so that $W_1 \subset W$. Let $G_1 \stackrel{\mathrm{def}}{=} \mathrm{gvars}(\bar{s}_1 = \bar{t}_1)$ and $V_1 \stackrel{\mathrm{def}}{=} \mathit{Vars} \setminus G_1$. Note that $G_1 \subseteq G$ and, by the assumption for this case, $t_1 \in V$, so that $t_1 \notin G$. As a consequence, $G_1 = G$, $V_1 = V$ and $\bar{s}_1$ is V_1-linear, so that the inductive hypothesis applies to $\bar{s}_1$ and $\bar{t}_1$. Thus, there exists $\mu_1 \in \mathrm{mgs}(\bar{s}_1 = \bar{t}_1)$ such that, for each $z \in \mathit{Vars} \setminus \mathrm{vars}(\bar{s_1})$, the three inductive properties hold.

Let $\mu \subseteq \mathit{Eqs}$ be defined as

$$\mu \stackrel{\mathrm{def}}{=} \{t_1 = s_1\mu_1\} \cup \mu_1. \tag{41}$$

Note that, by (40), $t_1 \notin \mathrm{vars}(\mu_1)$ and, in particular, $t_1 \notin \mathrm{dom}(\mu_1)$. Thus, since $\mu_1 \in \mathit{RSubst}$, μ has no identities or circular subsets so that $\mu \in \mathit{RSubst}$. By Lemma 51, $\mu \in \mathrm{mgs}(\bar{s} = \bar{t})$.

Suppose that $z \in \mathit{Vars} \setminus \mathrm{vars}(\bar{s})$. Then either $z \neq t_1$, so that $z\mu = z\mu_1$, or $z = t_1$, so that $z\mu = s_1\mu_1$. We show in each case that $z\mu$ satisfies the three required properties.

1. Suppose $z \neq t_1$. By inductive property 1, $z\mu_1$ is V_1-linear. As $z\mu = z\mu_1$ and $V = V_1$, $z\mu$ is V-linear.

   Otherwise, let $z = t_1$, so that $z\mu = s_1\mu_1$. Consider an arbitrary variable $u \in \mathrm{vars}(s_1)$. Then $u \in \mathit{Vars} \setminus \mathrm{vars}(\bar{s}_1)$ and the inductive properties using μ_1 and $\bar{s}_1$ can be applied to $u\mu_1$. Therefore, by property 1, $u\mu_1$ is V_1-linear. Moreover, by property 3, we have

   $$\forall u' \in \mathit{Vars} \setminus \mathrm{vars}(\bar{s}_1) : u \neq u' \implies \mathrm{vars}(u\mu_1) \cap \mathrm{vars}(u'\mu_1) \subseteq G_1.$$

   In particular, this holds for all $u' \in \mathrm{vars}(s_1)$ such that $u \neq u'$. As a consequence, $z\mu = s_1\mu_1$ is V_1-linear. As $V = V_1$, $z\mu$ is V-linear.

2. Suppose $z \neq t_1$. By property 2, we have $\mathrm{vars}(z\mu_1) \cap \mathrm{dom}(\mu_1) \subseteq G_1$. Since $z\mu = z\mu_1$, $G = G_1$, $\mathrm{dom}(\mu) = \mathrm{dom}(\mu_1) \cup \{t_1\}$ and $t_1 \notin \mathrm{vars}(z\mu)$, we obtain $\mathrm{vars}(z\mu) \cap \mathrm{dom}(\mu) \subseteq G$.

   Otherwise, let $z = t_1$ so that $z\mu = s_1\mu_1$. Consider $u \in \mathrm{vars}(s_1)$. Then $u \in \mathit{Vars} \setminus \mathrm{vars}(\bar{s}_1)$, so that the inductive properties using μ_1 and $\bar{s}_1$ can be applied to $u\mu_1$. By property 2, $\mathrm{vars}(u\mu_1) \cap \mathrm{dom}(\mu_1) \subseteq G_1$. As this holds for all $u \in \mathrm{vars}(s_1)$, we have $\mathrm{vars}(s_1\mu_1) \cap \mathrm{dom}(\mu_1) \subseteq G_1$. As $z\mu = s_1\mu_1$, $G = G_1$, $\mathrm{dom}(\mu) = \mathrm{dom}(\mu_1) \cup \{t_1\}$ and $t_1 \notin \mathrm{vars}(z\mu)$, we obtain $\mathrm{vars}(z\mu) \cap \mathrm{dom}(\mu) \subseteq G$.

3. Suppose $z \neq t_1$ and let $z' \in \mathit{Vars} \setminus \mathrm{vars}(\bar{s})$ be such that $z \neq z'$. Then, by inductive property 3, we have $\mathrm{vars}(z\mu_1) \cap \mathrm{vars}(z'\mu_1) \subseteq G_1$. Since $z\mu = z\mu_1$ and $G = G_1$, if also $z' \neq t_1$ (so that $z'\mu = z'\mu_1$) we obtain $\mathrm{vars}(z\mu) \cap \mathrm{vars}(z'\mu) \subseteq G$. Otherwise, let $z' = t_1$ (so that $z'\mu = s_1\mu_1$). We will show that

$$\forall u \in \mathrm{vars}(s_1) : \mathrm{vars}(z\mu_1) \cap \mathrm{vars}(u\mu_1) \subseteq G_1. \tag{42}$$

In fact, in the case that $u \in G_1$ then $\mathrm{vars}(u\mu_1) \subseteq G_1$. On the other hand, if $u \in \mathrm{vars}(s_1) \setminus G_1$, then we have $u \neq z$. As $\bar{s}$ is V-linear, $u \in \mathit{Vars} \setminus \mathrm{vars}(\bar{s}_1)$ so that the property holds by inductive property 3 (taking $z' = u$). As (42) holds, we have $\mathrm{vars}(z\mu_1) \cap \mathrm{vars}(s_1\mu_1) \subseteq G_1$. Thus, by observing that $z\mu = z\mu_1$, $z'\mu = s_1\mu_1$ and $G = G_1$, we can conclude $\mathrm{vars}(z\mu) \cap \mathrm{vars}(z'\mu) \subseteq G$.

Otherwise, let $z = t_1$ so that $z\mu = s_1\mu_1$. Let $z' \in \mathit{Vars} \setminus \mathrm{vars}(\bar{s})$ be such that $z \neq z'$ (note that this implies $z' \neq t_1$, so that $z'\mu = z'\mu_1$). We will prove that, for all $u \in \mathrm{vars}(s_1)$,

$$\mathrm{vars}(u\mu_1) \cap \mathrm{vars}(z'\mu_1) \subseteq G_1. \tag{43}$$

In fact, if $u \in G_1$ then $\mathrm{vars}(u\mu_1) \subseteq G_1$. Suppose now $u \in \mathrm{vars}(s_1) \setminus G_1$. As $\bar{s}$ is V-linear, $u \in \mathit{Vars} \setminus \mathrm{vars}(\bar{s}_1)$ so that the inductive property 3 can be applied to $u\mu_1$. Thus, for all $u' \in \mathit{Vars} \setminus \mathrm{vars}(\bar{s}_1)$, if $u \neq u'$ we have $\mathrm{vars}(u\mu_1) \cap \mathrm{vars}(u'\mu_1) \subseteq G_1$. In particular, since $u \in \mathrm{vars}(s_1)$ and $z' \notin \mathrm{vars}(\bar{s})$, we have $u \neq z'$ so that, by taking $u' = z'$, we obtain (43). As the choice of $u \in \mathrm{vars}(s_1)$ is arbitrary, we have $\mathrm{vars}(s_1\mu_1) \cap \mathrm{vars}(z'\mu_1) \subseteq G_1$. By observing that $z\mu = s_1\mu_1$, $z'\mu = z'\mu_1$ and $G = G_1$ we obtain $\mathrm{vars}(z\mu) \cap \mathrm{vars}(z'\mu) \subseteq G$.

Corollary 55 *There exists $\mu' \in \mathit{VSubst}$ that (under the same hypotheses) satisfies all the properties stated for $\mu \in \mathit{RSubst}$ in Lemma 54.*

Proof. We start by proving that the properties stated for $\mu \in \mathit{RSubst}$ in Lemma 54 are invariant under the application of an $\mathcal{S}$-step.

Suppose that $\mu \in \mathit{RSubst}$ satisfies the properties stated in Lemma 54 and $\mu \overset{\mathcal{S}}{\longmapsto} \mu'$. First note that, by Theorem 47, we have $\mu' \in \mathrm{mgs}(\bar{s} = \bar{t})$ and $\mathrm{dom}(\mu') = \mathrm{dom}(\mu)$. Also, by Lemma 53, $\mathrm{gvars}(\mu') = \mathrm{gvars}(\mu) = G$. By definition of $\mathcal{S}$-step, there exist $\{x \mapsto t, y \mapsto s\} \subseteq \mu$ such that $x \neq y$ and

$$\mu' \stackrel{\mathrm{def}}{=} \bigl(\mu \setminus \{y \mapsto s\}\bigr) \cup \{y \mapsto s[x/t]\}.$$

Let $z \in \mathit{Vars} \setminus \mathrm{vars}(\bar{s})$ and consider the term $z\mu'$. If $z \neq y$ or $x \notin \mathrm{vars}(s)$ then we have $z\mu' = z\mu$ and there is nothing to prove. Therefore, assume $z = y$, so that $z\mu = s$, and $x \in \mathrm{vars}(z\mu)$, so that $z\mu' = z\mu[x/t]$. Note that $x \in \mathrm{vars}(z\mu) \cap \mathrm{dom}(\mu)$ so that, by property 2, we have $x \in G$. As a consequence, $\mathrm{vars}(t) \subseteq G$ and we obtain

$$\mathrm{vars}(z\mu) \setminus G = \mathrm{vars}(z\mu') \setminus G.$$

From this, it is easy to conclude that properties 1, 2 and 3 hold for μ'.

By a simple induction, the above result generalizes to any finite sequence of $\mathcal{S}$-steps. Then, by Theorem 48 it follows than we can construct such a $\mu' \in \mathit{VSubst}$.

In the following three sections, we prove the correctness of the abstract unification operator on each component of the *SFL* domain. A further section will join all of these results to establish the whole correctness of amgu_S.

C.1 The Correctness for Set-Sharing

Proposition 56 *Let $d = \langle sh, f, l\rangle \in \mathit{SFL}$, $\sigma \in \gamma_S(d) \cap \mathit{VSubst}$ and $(x \mapsto t) \in \mathit{Bind}$, where $\{x\} \cup \mathrm{vars}(t) \subseteq \mathit{VI}$ and $y \in \mathrm{dom}(\sigma) \cap \mathrm{range}(\sigma)$ implies $y \in \mathrm{vars}(y\sigma)$. Suppose that $\{r, r'\} = \{x, t\}$ and $\mathrm{free}_d(r)$ holds. For all $\tau \in \mathrm{mgs}\bigl(\sigma \cup \{x = t\}\bigr)$ in the syntactic equality theory T, letting*

$$
\begin{aligned}
sh_- &= \overline{\mathrm{rel}}\bigl(\{x\} \cup \mathrm{vars}(t), sh\bigr),\\
sh_r &= \mathrm{rel}\bigl(\mathrm{vars}(r), sh\bigr),\\
sh_{r'} &= \mathrm{rel}\bigl(\mathrm{vars}(r'), sh\bigr),
\end{aligned}
$$

we have

$$sh_- \cup \mathrm{bin}(sh_r, sh_{r'}) \supseteq \mathrm{ssets}(\tau). \tag{44}$$

Proof. We assume that the congruence and identity axioms hold. Note that if $\sigma \cup \{x = t\}$ is not satisfiable, then the result is trivial. We therefore assume, for the rest of the proof, that $\sigma \cup \{x = t\}$ is satisfiable in T. It follows from Lemma 53 that we just have to show that there exists $\tau \in \mathrm{mgs}\bigl(\sigma \cup \{x = t\}\bigr)$ such that (44) holds.

Since $\mathrm{free}_d(r)$ holds, by Theorem 21, $\mathrm{rt}(r, \sigma) \in \mathit{Vars}$ and hence, by Proposition 42,

$$r\sigma \in \mathit{Vars} \setminus \mathrm{dom}(\sigma). \tag{45}$$

Let

$$
\begin{aligned}
R_- &= \overline{\mathrm{rel}}\bigl(\{x\} \cup \mathrm{vars}(t), \mathrm{ssets}(\sigma)\bigr),\\
R_r &= \mathrm{rel}\bigl(\mathrm{vars}(r), \mathrm{ssets}(\sigma)\bigr),\\
R_{r'} &= \mathrm{rel}\bigl(\mathrm{vars}(r'), \mathrm{ssets}(\sigma)\bigr).
\end{aligned}
$$

Since $\sigma \in \gamma_S(d)$, we have $sh \supseteq \mathrm{ssets}(\sigma)$ so that, using the monotonicity of $\overline{\mathrm{rel}}$, rel and bin, we obtain

$$sh_- \cup \mathrm{bin}(sh_r, sh_{r'}) \supseteq R_- \cup \mathrm{bin}(R_r, R_{r'}).$$

Thus, in order to prove (44) it is sufficient to show that

$$R_- \cup \mathrm{bin}(R_r, R_{r'}) \supseteq \mathrm{ssets}(\tau). \tag{46}$$

Note that, by Definition 8 and (45), we obtain

$$R_r = \bigl\{\mathrm{occ}(\sigma, r\sigma)\bigr\}. \tag{47}$$

Suppose first $x\sigma = t\sigma$. Then we have $\sigma \in \mathrm{mgs}\bigl(\sigma \cup \{x\sigma = t\sigma\}\bigr)$, so that by Lemma 51, $\sigma \in \mathrm{mgs}\bigl(\sigma \cup \{x = t\}\bigr)$. Thus, take $\tau \stackrel{\mathrm{def}}{=} \sigma$. Moreover, by (47), we also have $R_{r'} = \bigl\{\mathrm{occ}(\sigma, r\sigma)\bigr\} = R_r$, so that $R_r = \mathrm{bin}(R_r, R_{r'})$. As a consequence,

$$\begin{aligned} R_- \cup \mathrm{bin}(R_r, R_{r'}) &= \bigl(\mathrm{ssets}(\sigma) \setminus R_r\bigr) \cup R_r \\ &= \mathrm{ssets}(\sigma) \\ &= \mathrm{ssets}(\tau). \end{aligned}$$

Otherwise, let $x\sigma \neq t\sigma$ and let $\nu, \mu, \tau \in \mathit{RSubst}$ be defined as

$$\begin{aligned} \nu &\stackrel{\mathrm{def}}{=} \bigl\{\, (y \mapsto s) \in \sigma \bigm| y \notin \mathrm{vars}(x\sigma) \cup \mathrm{vars}(t\sigma) \,\bigr\}, \\ \mu &\stackrel{\mathrm{def}}{=} \{r\sigma \mapsto r'\sigma\}, \\ \tau &\stackrel{\mathrm{def}}{=} \nu \circ \mu. \end{aligned}$$

As $\nu \subseteq \sigma \in \mathit{VSubst}$, by Lemma 38 we have $\nu \in \mathit{VSubst}$; also, $\mu \in \mathit{VSubst}$ because it has a single binding; moreover, by construction, $\mathrm{dom}(\nu) \cap \mathrm{vars}(\mu) = \varnothing$; thus we can apply Lemma 39 to obtain $\tau \in \mathit{VSubst}$. By applying Lemma 51, we also have $\tau \in \mathrm{mgs}\bigl(\sigma \cup \{x = t\}\bigr)$.

Suppose $S \in \mathrm{ssets}(\tau)$. By Definition 49 and Theorem 50, we have

$$S \in R_- \cup \mathrm{bin}(R_r^\star, R_{r'}^\star).$$

If $S \in R_-$, then (46) holds trivially. Therefore suppose $S \in \mathrm{bin}(R_r^\star, R_{r'}^\star)$, so that there exist $S_r \in R_r^\star$ and $S_{r'} \in R_{r'}^\star$ such that $S = S_r \cup S_{r'}$. Note that, by (47), $R_r^\star = R_r$. Thus, to prove (46) holds, it is sufficient to show that $S_{r'} \in R_{r'}$.

As $S \in \mathrm{ssets}(\tau)$, by Proposition 42, there exists $v \in \mathit{Vars} \setminus \mathrm{dom}(\tau)$ such that $S = \{\, y \in \mathit{VI} \mid v \in \mathrm{vars}(y\tau) \,\}$. As $v \notin \mathrm{dom}(\tau)$, $v \neq r\sigma$.

Let $y \in S$. We show that $y \in \mathrm{occ}(\sigma, r\sigma) \cup \mathrm{occ}(\sigma, v)$. Using Lemma 33, we have $T \vdash \forall(\tau \to y\tau = y)$. Therefore, since $T \vdash \forall(\tau \to \sigma)$, by Lemma 41 (replacing $s = t$ by $y\tau = y$), there exists $z \in \mathrm{vars}(y\sigma) \setminus \mathrm{dom}(\sigma)$ such that $v \in \mathrm{vars}(z\tau)$. If $z = r\sigma$, then $y \in S_r$. If $z \neq r\sigma$, then $z \notin \mathrm{dom}(\tau)$ and $z\tau = z$. Therefore $v = z$ and $y \in \mathrm{occ}(\sigma, v)$.

Proposition 57 *Let $d = \langle sh, f, l\rangle \in \mathit{SFL}$, $\sigma \in \gamma_S(d) \cap \mathit{VSubst}$ and $(x \mapsto t) \in \mathit{Bind}$, where $\{x\} \cup \mathrm{vars}(t) \subseteq \mathit{VI}$ and $y \in \mathrm{dom}(\sigma) \cap \mathrm{range}(\sigma)$ implies $y \in \mathrm{vars}(y\sigma)$. Suppose that $\mathrm{lin}_d(x)$ and $\mathrm{lin}_d(t)$ hold. For all $\tau \in \mathrm{mgs}\bigl(\sigma \cup \{x = t\}\bigr)$ in the syntactic equality theory T, letting*

$$\begin{aligned} sh_- &= \overline{\mathrm{rel}}\bigl(\{x\} \cup \mathrm{vars}(t), sh\bigr), \\ sh_x &= \mathrm{rel}\bigl(\{x\}, sh\bigr), \\ sh_t &= \mathrm{rel}\bigl(\mathrm{vars}(t), sh\bigr), \\ sh_{xt} &= sh_x \cap sh_t, \end{aligned}$$

we have

$$sh_- \cup \mathrm{bin}\bigl(sh_x \cup \mathrm{bin}(sh_x, sh_{xt}^\star), sh_t \cup \mathrm{bin}(sh_t, sh_{xt}^\star)\bigr) \supseteq \mathrm{ssets}(\tau). \tag{48}$$

Proof. We assume that the congruence and identity axioms hold. Note that if $\sigma \cup \{x = t\}$ is not satisfiable, then the result is trivial. We therefore assume, for the rest of the proof, that $\sigma \cup \{x = t\}$ is satisfiable in T. It follows from Lemma 53 that we just have to show that there exists $\tau \in \mathrm{mgs}\bigl(\sigma \cup \{x = t\}\bigr)$ such that (48) holds.

For every $r \in \{x, t\}$, by hypothesis, $\mathrm{lin}_d(r)$ holds. Then, by Theorem 21, $\mathrm{rt}(r, \sigma) \in \mathit{LTerms}$ and hence, by Proposition 16, $\mathrm{vars}(r) \subseteq \mathrm{lvars}(\sigma)$. Thus, by Proposition 42,

$$\begin{aligned} \forall v \in \mathrm{vars}(r\sigma) \setminus \mathrm{dom}(\sigma) &: \mathrm{occ\_lin}(v, r\sigma), \\ \mathrm{vars}(r\sigma) \cap \mathrm{dom}(\sigma) &\subseteq \mathrm{gvars}(\sigma). \end{aligned} \tag{49}$$

Therefore, by defining $V_\sigma \stackrel{\mathrm{def}}{=} \mathit{Vars} \setminus \mathrm{gvars}(\sigma)$, we obtain that both terms $x\sigma$ and $t\sigma$ are V_σ-linear. Let

$$\begin{aligned} \{u_1, \ldots, u_k\} &\stackrel{\mathrm{def}}{=} \mathrm{dom}(\sigma) \cap \bigl(\mathrm{vars}(x\sigma) \cup \mathrm{vars}(t\sigma)\bigr), \\ \bar{s} &\stackrel{\mathrm{def}}{=} (u_1, \ldots, u_k, x\sigma), \\ \bar{t} &\stackrel{\mathrm{def}}{=} (u_1\sigma, \ldots, u_k\sigma, t\sigma). \end{aligned}$$

Since $x\sigma$ is V_σ-linear, it follows from (49) (letting $r = x$) that $\bar{s}$ is V_σ-linear. It also follows from (49) (applied twice, once with $r = x$ and once with $r = t$) that, for each $i = 1, \ldots, k$ we have $u_i \in \mathrm{gvars}(\sigma)$, so that $\mathrm{vars}(u_i\sigma) \subseteq \mathrm{gvars}(\sigma)$. Therefore, since $t\sigma$ is V_σ-linear, $\bar{t}$ is also V_σ-linear. By Lemma 51 and the congruence axioms, $\sigma \cup \{x = t\} \implies \bar{s} = \bar{t}$. Thus, as $\sigma \cup \{x = t\}$ is satisfiable, there exists $\mu \in \mathrm{mgs}(\bar{s} = \bar{t})$. Let $V_\mu = \mathit{Vars} \setminus \mathrm{gvars}(\mu)$; since $\mathrm{gvars}(\sigma) \subseteq \mathrm{gvars}(\mu)$, then $V_\mu \subseteq V_\sigma$ and $\bar{s}$, $\bar{t}$ are also V_μ-linear. Therefore, we can apply Lemma 54 and Corollary 55 so that, by case (3), there exists $\mu \in \mathrm{mgs}(\bar{s} = \bar{t}) \cap \mathit{VSubst}$ such that, for all $w, w' \in \mathit{Vars}$ where $w \neq w'$ and either $\{w, w'\} \cap \mathrm{vars}(\bar{s}) = \varnothing$ or $\{w, w'\} \cap \mathrm{vars}(\bar{t}) = \varnothing$,

$$\mathrm{vars}(w\mu) \cap \mathrm{vars}(w'\mu) \subseteq \mathrm{gvars}(\mu). \tag{50}$$

Note that, since $\sigma \in \mathit{VSubst}$, we have $\mathrm{vars}(u_i\sigma) \subseteq \mathrm{vars}(x\sigma) \cup \mathrm{vars}(t\sigma)$ for each $i = 1, \ldots, k$. Therefore

$$\mathrm{vars}(\mu) \subseteq \mathrm{vars}(x\sigma) \cup \mathrm{vars}(t\sigma). \tag{51}$$

Let $\nu, \tau \in \mathit{RSubst}$ be defined as

$$\begin{aligned} \nu &\stackrel{\mathrm{def}}{=} \bigl\{\, (y \mapsto s) \in \sigma \bigm| y \notin \mathrm{vars}(x\sigma) \cup \mathrm{vars}(t\sigma) \,\bigr\}, \\ \tau &\stackrel{\mathrm{def}}{=} \nu \circ \mu. \end{aligned}$$

As $\nu \subseteq \sigma \in \mathit{VSubst}$, by Lemma 38 we have $\nu \in \mathit{VSubst}$; moreover, by (51), $\mathrm{dom}(\nu) \cap \mathrm{vars}(\mu) = \varnothing$; thus we can apply Lemma 39 to obtain $\tau \in \mathit{VSubst}$. By applying Lemma 51, we also have $\tau \in \mathrm{mgs}\bigl(\sigma \cup \{x = t\}\bigr)$.

Let

$$\begin{aligned} R_- &= \overline{\mathrm{rel}}\bigl(\{x\} \cup \mathrm{vars}(t), \mathrm{ssets}(\sigma)\bigr), \\ R_x &= \mathrm{rel}\bigl(\{x\}, \mathrm{ssets}(\sigma)\bigr), \\ R_t &= \mathrm{rel}\bigl(\mathrm{vars}(t), \mathrm{ssets}(\sigma)\bigr), \\ R_{xt} &= R_x \cap R_t. \end{aligned}$$

Since $\sigma \in \gamma_S(d)$, we have $sh \supseteq \mathrm{ssets}(\sigma)$ so that, using the monotonicity of $\overline{\mathrm{rel}}$, rel, $(\cdot)^\star$ and bin, we obtain

$$sh_- \cup \mathrm{bin}\bigl(sh_x \cup \mathrm{bin}(sh_x, sh_{xt}^\star), sh_t \cup \mathrm{bin}(sh_t, sh_{xt}^\star)\bigr) \\ \supseteq R_- \cup \mathrm{bin}\bigl(R_x \cup \mathrm{bin}(R_x, R_{xt}^\star), R_t \cup \mathrm{bin}(R_t, R_{xt}^\star)\bigr).$$

It follows that, in order to prove (48), it is sufficient to show

$$R_- \cup \mathrm{bin}\bigl(R_x \cup \mathrm{bin}(R_x, R_{xt}^\star), R_t \cup \mathrm{bin}(R_t, R_{xt}^\star)\bigr) \supseteq \mathrm{ssets}(\tau). \tag{52}$$

Let S be an arbitrary sharing set in $\mathrm{ssets}(\tau)$. By Definition 49 and Theorem 50, we have

$$S \in R_- \cup \mathrm{bin}(R_x^\star, R_t^\star).$$

If $S \in R_-$, then (52) holds trivially. Therefore suppose $S \in \mathrm{bin}(R_x^\star, R_t^\star)$, so that there exist $S_x \in R_x^\star$ and $S_t \in R_t^\star$ such that $S = S_x \cup S_t$. We prove that (52) holds by showing that $S_x \in R_x \cup \mathrm{bin}(R_x, R_{xt}^\star)$ and $S_t \in R_t \cup \mathrm{bin}(R_t, R_{xt}^\star)$.

We first show that $S_t \in R_t \cup \mathrm{bin}(R_t, R_{xt}^\star)$. As $S_t \in R_t^\star$, $S_t = S_1 \cup S_2$ where $S_1 \in (R_t \setminus R_{xt})^\star \cup \{\varnothing\}$ and $S_2 \in R_{xt}^\star \cup \{\varnothing\}$. Note that as $S_t \neq \varnothing$, we cannot have $S_1 = S_2 = \varnothing$. Suppose first that $S_1 = \varnothing$ so that $S_t = S_2 \neq \varnothing$. Then $S_t \in R_{xt}^\star$. However, since $R_{xt} \subseteq R_t$, $R_{xt}^\star \subseteq \mathrm{bin}(R_t, R_{xt}^\star)$. Thus $S_t \in \mathrm{bin}(R_t, R_{xt}^\star)$. Suppose next that $S_1 \neq \varnothing$. As $R_t \setminus R_{xt} = R_t \setminus R_x$, we have $S_1 = \bigcup\bigl\{\, \mathrm{occ}(\sigma, w) \bigm| w \in S_1 \setminus \mathrm{vars}(x\sigma) \,\bigr\}$. However, as $\mathrm{occ}(\sigma, w) = \varnothing$ for all $w \in \mathrm{dom}(\sigma)$ and $\mathrm{vars}(\bar{s}) \setminus \mathrm{vars}(x\sigma) \subseteq \mathrm{dom}(\sigma)$,

$$S_1 = \bigcup\Bigl\{\, \mathrm{occ}(\sigma, w) \Bigm| w \in S_1 \setminus \bigl(\mathrm{dom}(\sigma) \cup \mathrm{vars}(\bar{s})\bigr) \,\Bigr\}$$

Let $w_1, w_2 \in S_1 \setminus \bigl(\mathrm{dom}(\sigma) \cup \mathrm{vars}(\bar{s})\bigr)$. Then, as $S_1 \subseteq S$, $S \in \mathrm{ssets}(\tau)$ and $\tau \in \mathit{VSubst}$, by Proposition 42, there exists $v \in \mathit{Vars} \setminus \mathrm{dom}(\tau)$ such that $v \in \mathrm{vars}(w_1\tau) \cap \mathrm{vars}(w_2\tau)$. However, since $w_i \notin \mathrm{dom}(\sigma)$, we have $w_i\tau = w_i\mu$, for $i \in \{1, 2\}$. Thus, noting that $v \notin \mathrm{dom}(\tau)$ implies $v \notin \mathrm{gvars}(\tau)$, we can apply (50) to conclude that $w_1 = w_2$. As the choice of w_1 and w_2 was arbitrary, there exists a unique variable $w \in S_1 \setminus \bigl(\mathrm{dom}(\sigma) \cup \mathrm{vars}(\bar{s})\bigr)$ such that $S_1 = \mathrm{occ}(\sigma, w)$. Thus $S_1 \in R_t$. If $S_2 = \varnothing$ then $S_t = S_1 \in R_t$. If $S_2 \neq \varnothing$, then $S_t = S_1 \cup S_2 \in \mathrm{bin}(R_t, R_{xt}^\star)$. Hence, in both cases, $S_t \in R_t \cup \mathrm{bin}(R_t, R_{xt}^\star)$.

By the same reasoning, (replacing x by t, t by x and $\bar{s}$ by $\bar{t}$ in the previous paragraph) we obtain $S_x \in R_x \cup \mathrm{bin}(R_x, R_{xt}^\star)$, thus completing the proof.

Corollary 58 *Let $d = \langle sh, f, l \rangle \in \mathit{SFL}$, $\sigma \in \gamma_S(d) \cap \mathit{VSubst}$ and $(x \mapsto t) \in \mathit{Bind}$, where $\{x\} \cup \mathrm{vars}(t) \subseteq \mathit{VI}$ and $y \in \mathrm{dom}(\sigma) \cap \mathrm{range}(\sigma)$ implies $y \in \mathrm{vars}(y\sigma)$. Suppose that $\mathrm{lin}_d(x)$, $\mathrm{lin}_d(t)$ and $\mathrm{ind}_d(x, t)$ hold. For all $\tau \in \mathrm{mgs}\bigl(\sigma \cup \{x = t\}\bigr)$ in the syntactic equality theory T, letting*

$$\begin{aligned} sh_- &= \overline{\mathrm{rel}}\bigl(\{x\} \cup \mathrm{vars}(t), sh\bigr), \\ sh_x &= \mathrm{rel}\bigl(\{x\}, sh\bigr), \\ sh_t &= \mathrm{rel}\bigl(\mathrm{vars}(t), sh\bigr) \end{aligned}$$

we have

$$sh_- \cup \mathrm{bin}(sh_x, sh_t) \supseteq \mathrm{ssets}(\tau).$$

Proof. Since $\mathrm{ind}_d(x, t)$ holds, by Definition 20 we have $sh_x \cap sh_t = \varnothing$. The result then follows from Proposition 57.

Proposition 59 *Let $d = \langle sh, f, l \rangle \in SFL$, $\sigma \in \gamma_S(d) \cap VSubst$ and $(x \mapsto t) \in Bind$, where $\{x\} \cup \mathrm{vars}(t) \subseteq VI$ and $y \in \mathrm{dom}(\sigma) \cap \mathrm{range}(\sigma)$ implies $y \in \mathrm{vars}(y\sigma)$. Suppose that $\{r, r'\} = \{x, t\}$ and $\mathrm{lin}_d(r)$ holds. For all $\tau \in \mathrm{mgs}\bigl(\sigma \cup \{x = t\}\bigr)$ in the syntactic equality theory T, letting*

$$
\begin{aligned}
sh_- &= \overline{\mathrm{rel}}\bigl(\{x\} \cup \mathrm{vars}(t), sh\bigr),\\
sh_r &= \mathrm{rel}\bigl(\mathrm{vars}(r), sh\bigr),\\
sh_{r'} &= \mathrm{rel}\bigl(\mathrm{vars}(r'), sh\bigr),
\end{aligned}
$$

we have

$$
sh_- \cup \mathrm{bin}(sh_r^\star, sh_{r'}) \supseteq \mathrm{ssets}(\tau). \tag{53}
$$

Proof. We assume that the congruence and identity axioms hold. Note that if $\sigma \cup \{x = t\}$ is not satisfiable, then the result is trivial. We therefore assume, for the rest of the proof, that $\sigma \cup \{x = t\}$ is satisfiable in T. It follows from Lemma 53 that we just have to show that there exists $\tau \in \mathrm{mgs}\bigl(\sigma \cup \{x = t\}\bigr)$ such that (53) holds.

Since $\mathrm{lin}_d(r)$ holds, by Theorem 21, $\mathrm{rt}(r, \sigma) \in LTerms$ and hence, by Proposition 16, $\mathrm{vars}(r) \subseteq \mathrm{lvars}(\sigma)$. Thus, by Proposition 42,

$$
\begin{aligned}
\forall v \in \mathrm{vars}(r\sigma) \setminus \mathrm{dom}(\sigma) &: \mathrm{occ\_lin}(v, r\sigma),\\
\mathrm{vars}(r\sigma) \cap \mathrm{dom}(\sigma) &\subseteq \mathrm{gvars}(\sigma).
\end{aligned} \tag{54}
$$

Therefore, by defining $V_\sigma \stackrel{\mathrm{def}}{=} Vars \setminus \mathrm{gvars}(\sigma)$, we obtain that the term $r\sigma$ is V_σ-linear. Let

$$
\begin{aligned}
\{u_1, \ldots, u_k\} &\stackrel{\mathrm{def}}{=} \mathrm{dom}(\sigma) \cap \bigl(\mathrm{vars}(x\sigma) \cup \mathrm{vars}(t\sigma)\bigr),\\
\bar{s} &\stackrel{\mathrm{def}}{=} (u_1, \ldots, u_k, r\sigma),\\
\bar{t} &\stackrel{\mathrm{def}}{=} (u_1\sigma, \ldots, u_k\sigma, r'\sigma).
\end{aligned}
$$

Since $r\sigma$ is V_σ-linear it follows from (54) that $\bar{s}$ is V_σ-linear. By Lemma 51 and the congruence axioms, $\sigma \cup \{x = t\} \implies \bar{s} = \bar{t}$. Thus, as $\sigma \cup \{x = t\}$ is satisfiable, there exists $\mu \in \mathrm{mgs}(\bar{s} = \bar{t})$. Let $V_\mu = Vars \setminus \mathrm{gvars}(\mu)$; since $\mathrm{gvars}(\sigma) \subseteq \mathrm{gvars}(\mu)$, then $V_\mu \subseteq V_\sigma$ and $\bar{s}$ is also V_μ-linear. Therefore, we can apply Lemma 54 and Corollary 55 so that, by case (3), there exists $\mu \in \mathrm{mgs}(\bar{s} = \bar{t}) \cap VSubst$ such that, for all $w, w' \in Vars \setminus \mathrm{vars}(\bar{s})$ where $w \neq w'$,

$$
\mathrm{vars}(w\mu) \cap \mathrm{vars}(w'\mu) \subseteq \mathrm{gvars}(\mu). \tag{55}
$$

Note that, since $\sigma \in VSubst$, we have $\mathrm{vars}(u_i\sigma) \subseteq \mathrm{vars}(x\sigma) \cup \mathrm{vars}(t\sigma)$ for each $i = 1, \ldots, k$. Therefore

$$
\mathrm{vars}(\mu) \subseteq \mathrm{vars}(x\sigma) \cup \mathrm{vars}(t\sigma). \tag{56}
$$

Let $\nu, \tau \in RSubst$ be defined as

$$
\nu \stackrel{\mathrm{def}}{=} \bigl\{\, (y \mapsto s) \in \sigma \bigm| y \notin \mathrm{vars}(x\sigma) \cup \mathrm{vars}(t\sigma) \,\bigr\},
$$

$$\tau \stackrel{\text{def}}{=} \nu \circ \mu.$$

As $\nu \subseteq \sigma \in \mathit{VSubst}$, by Lemma 38 we have $\nu \in \mathit{VSubst}$; moreover, by (56), $\mathrm{dom}(\nu) \cap \mathrm{vars}(\mu) = \varnothing$; thus we can apply Lemma 39 to obtain $\tau \in \mathit{VSubst}$. By applying Lemma 51, we also have $\tau \in \mathrm{mgs}\bigl(\sigma \cup \{x = t\}\bigr)$.

Let

$$\begin{aligned} R_- &= \overline{\mathrm{rel}}\bigl(\{x\} \cup \mathrm{vars}(t), \mathrm{ssets}(\sigma)\bigr), \\ R_r &= \mathrm{rel}\bigl(\mathrm{vars}(r), \mathrm{ssets}(\sigma)\bigr), \\ R_{r'} &= \mathrm{rel}\bigl(\mathrm{vars}(r'), \mathrm{ssets}(\sigma)\bigr). \end{aligned}$$

Since $\sigma \in \gamma_S(d)$, we have $sh \supseteq \mathrm{ssets}(\sigma)$ so that, using the monotonicity of $\overline{\mathrm{rel}}$, rel, $(\cdot)^\star$ and bin, we obtain

$$sh_- \cup \mathrm{bin}(sh_r^\star, sh_{r'}) \supseteq R_- \cup \mathrm{bin}(R_r^\star, R_{r'}).$$

Thus, in order to prove (53) it is sufficient to show that

$$R_- \cup \mathrm{bin}(R_r^\star, R_{r'}) \supseteq \mathrm{ssets}(\tau). \tag{57}$$

Suppose $S \in \mathrm{ssets}(\tau)$. By Definition 49 and Theorem 50, we have

$$S \in R_- \cup \mathrm{bin}(R_r^\star, R_{r'}^\star).$$

If $S \in R_-$, then (57) holds trivially. Therefore suppose $S \in \mathrm{bin}(R_r^\star, R_{r'}^\star)$, so that there exist $S_r \in R_r^\star$ and $S_{r'} \in R_{r'}^\star$ such that $S = S_r \cup S_{r'}$. First note that, since $S_{r'} \in R_{r'}^\star$, there exists $S' \subseteq S_{r'}$ such that $S' \in R_{r'}$. Moreover, if $S_{r'} \in R_r^\star$, then we have $S \in R_r^\star$, so that $S = S \cup S' \in \mathrm{bin}(R_r^\star, R_{r'})$, proving (57). Thus, we now assume that $S_{r'} \notin R_r^\star$.

As $S_{r'} \in R_{r'}^\star$, we have $S_{r'} = \bigcup\bigl\{\, \mathrm{occ}(\sigma, w) \bigm| w \in S_{r'} \setminus \mathrm{dom}(\sigma) \,\bigr\}$. From this, as $S_{r'} \notin R_r^\star$, $\mathrm{vars}(\bar{s}) \subseteq \mathrm{dom}(\sigma) \cup \mathrm{vars}(r\sigma)$ and $\mathrm{vars}(r\sigma) \subseteq \mathrm{vars}(\bar{s})$, we obtain

$$S_{r'} = \bigcup\Bigl\{\, \mathrm{occ}(\sigma, w) \Bigm| w \in S_{r'} \setminus \bigl(\mathrm{dom}(\sigma) \cup \mathrm{vars}(\bar{s})\bigr) \,\Bigr\}.$$

Let $w_1, w_2 \in S_{r'} \setminus \bigl(\mathrm{dom}(\sigma) \cup \mathrm{vars}(\bar{s})\bigr)$. Note that, as $S_{r'} \subseteq S$, $S \in \mathrm{ssets}(\tau)$ and $\tau \in \mathit{VSubst}$, by Proposition 42 there exists $v \in \mathit{Vars} \setminus \mathrm{dom}(\tau)$ such that $v \in \mathrm{vars}(w_1\tau) \cap \mathrm{vars}(w_2\tau)$. However, since $w_i \notin \mathrm{dom}(\sigma)$, we have $w_i\tau = w_i\mu$, for $i \in \{1, 2\}$. Thus, noting that $v \notin \mathrm{dom}(\tau)$ implies $v \notin \mathrm{gvars}(\tau)$, we can apply (55) to conclude that $w_1 = w_2$. As the choice of w_1 and w_2 was arbitrary, there exists a unique variable $w \in S_{r'} \setminus \bigl(\mathrm{dom}(\sigma) \cup \mathrm{vars}(\bar{s})\bigr)$ such that $S_{r'} = \mathrm{occ}(\sigma, w)$. Thus $S_{r'} \in R_{r'}$ and (57) holds.

Lemma 60 *Let $sh \in \mathit{SH}$ and $V, W \subseteq \mathit{VI}$, where $\mathrm{rel}(V, sh) \subseteq \mathrm{rel}(W, sh)$. Let $(x \mapsto t) \in \mathit{Bind}$ and $sh' \stackrel{\text{def}}{=} \mathrm{amgu}(sh, x \mapsto t)$, where $\{x\} \cup \mathrm{vars}(t) \subseteq \mathit{VI}$. Then, $\mathrm{rel}(V, sh') \subseteq \mathrm{rel}(W, sh')$.*

Proof. Suppose that $S \in \mathrm{rel}(V, sh')$. By Definition 49, we have two cases.

1. Suppose first that $S \in \overline{\mathrm{rel}}\bigl(\{x\} \cup \mathrm{vars}(t), sh\bigr)$. Then, $S \in sh$ and, in particular, $S \in \mathrm{rel}(V, sh)$. Thus, by hypothesis, $S \in \mathrm{rel}(W, sh)$ and we can conclude $S \in \mathrm{rel}(W, sh')$.

2. Otherwise, let $S \in \operatorname{bin}\Bigl(\operatorname{rel}\bigl(\{x\}, sh\bigr)^\star, \operatorname{rel}\bigl(\operatorname{vars}(t), sh\bigr)^\star\Bigr)$. Then, it holds $S = S_0 \cup \cdots \cup S_n$ where $n \in \mathbb{N}$ and $S_i \in sh$, for each $0 \leq i \leq n$. Moreover, since $S \in \operatorname{rel}(V, sh')$, there exists an index $j \in \{0, \ldots, n\}$ such that $S_j \in \operatorname{rel}(V, sh)$. Hence, by the hypothesis, $S_j \in \operatorname{rel}(W, sh)$ and, since $S_j \subseteq S$, we can conclude $S \in \operatorname{rel}(W, sh')$.

Proposition 61 *Let $sh \in SH$, $(x \mapsto t) \in Bind$, where $x \in \operatorname{vars}(t) \subseteq VI$. Let $\tau \in RSubst$ be satisfiable in the syntactic equality theory T and suppose that $T \vdash \forall(\tau \rightarrow x = t)$ and $\operatorname{ssets}(\tau) \subseteq sh$. Then $\operatorname{ssets}(\tau) \subseteq \operatorname{cyclic}_x^t(sh)$.*

Proof. Take $V_x = \{x\}$ and $W_t = \operatorname{vars}(t) \setminus \{x\}$. Let $\tau = \{x_1 \mapsto t_1, \ldots, x_n \mapsto t_n\}$, where $n = \#\tau$. Define

$$\tau_0 \stackrel{\text{def}}{=} \{x \mapsto t\}, \qquad sh_0 \stackrel{\text{def}}{=} \operatorname{ssets}\bigl(\{x \mapsto t\}\bigr),$$

and, for each $i = 1, \ldots, n$,

$$\tau_i \in \operatorname{mgs}\bigl(\{x_1 = t_1, \ldots, x_i = t_i\} \cup \{x = t\}\bigr), \quad sh_i \stackrel{\text{def}}{=} \operatorname{amgu}(sh_{i-1}, x_i \mapsto t_i).$$

We show by induction on $i = 0, \ldots, n$ that $\operatorname{ssets}(\tau_i) \subseteq sh_i$ and

$$\operatorname{rel}(V_x, sh_i) \subseteq \operatorname{rel}(W_t, sh_i). \tag{58}$$

The base case, when $i = 0$, follows directly from Definition 8; note that (58) holds because $x \in \operatorname{dom}(\tau_0)$, so that $\operatorname{occ}(\tau_0, x) = \varnothing$.

Consider the inductive case, when $0 < i \leq n$. By the inductive hypothesis, $\operatorname{ssets}(\tau_{i-1}) \subseteq sh_{i-1}$ so that, by Theorem 50, $\operatorname{ssets}(\tau_i) \subseteq sh_i$. By the inductive hypothesis, (58) holds for sh_{i-1}. Thus, by Lemma 60 (taking $V = V_x$ and $W = W_t$), we obtain that (58) also holds for sh_i.

By taking $i = n$, we obtain $\operatorname{rel}(V_x, sh_n) \subseteq \operatorname{rel}(W_t, sh_n)$. Note that, by hypothesis, we have $\tau \in \operatorname{mgs}\bigl(\tau \cup \{x = t\}\bigr) = \operatorname{mgs}(\tau_n)$, so that $T \vdash \forall(\tau \leftrightarrow \tau_n)$. By Lemma 53, we have $\operatorname{ssets}(\tau) = \operatorname{ssets}(\tau_n)$, so that $\operatorname{ssets}(\tau) \subseteq sh_n$. As a consequence, $\operatorname{ssets}(\tau) \subseteq sh \cap sh_n$. Thus, by Definition 20, we obtain $\operatorname{ssets}(\tau) \subseteq \operatorname{cyclic}_x^t(sh)$.

C.2 The Correctness for Freeness

Proposition 62 *Let $\sigma \in VSubst$ and $(x \mapsto y) \in Bind$, where $\{x, y\} \subseteq VI$. Suppose also that $\{x, y\} \subseteq \operatorname{fvars}(\sigma)$. Then, for all $\tau \in \operatorname{mgs}\bigl(\sigma \cup \{x = y\}\bigr)$ in the syntactic equality theory T, we have*

$$\operatorname{fvars}(\sigma) \subseteq \operatorname{fvars}(\tau). \tag{59}$$

Proof. We assume that the congruence and identity axioms hold. Note that if $\sigma \cup \{x = y\}$ is not satisfiable in T, then the result is trivial. We therefore assume, for the rest of the proof, that $\sigma \cup \{x = y\}$ is satisfiable in T. It follows from Lemma 53 that we just have to show that there exists $\tau \in \operatorname{mgs}\bigl(\sigma \cup \{x = y\}\bigr)$ such that (59) holds.

As $\{x, y\} \subseteq \operatorname{fvars}(\sigma)$ we have, using Proposition 42,

$$\{x\sigma, y\sigma\} \subseteq Vars \setminus \operatorname{dom}(\sigma). \tag{60}$$

Suppose first $x\sigma = y\sigma$. Then we have $\sigma \in \operatorname{mgs}\bigl(\sigma \cup \{x\sigma = y\sigma\}\bigr)$, so that by Lemma 51, $\sigma \in \operatorname{mgs}\bigl(\sigma \cup \{x = y\}\bigr)$. Thus, by taking $\tau \stackrel{\mathrm{def}}{=} \sigma$, we trivially obtain $\operatorname{fvars}(\sigma) = \operatorname{fvars}(\tau)$.

Otherwise, let $x\sigma \neq y\sigma$ and take $\tau \stackrel{\mathrm{def}}{=} \sigma \cup \{x\sigma = y\sigma\}$. Then, since $\sigma \in \mathit{RSubst}$, it follows from (60) that $\tau \in \mathit{Eqs}$ has no identities or circular subsets so that $\tau \in \mathit{RSubst}$. Note that $\operatorname{dom}(\tau) = \operatorname{dom}(\sigma) \cup \{x\sigma\}$. By Lemma 51, $\tau \in \operatorname{mgs}\bigl(\sigma \cup \{x = t\}\bigr)$. Suppose $z \in \operatorname{fvars}(\sigma)$, so that by Proposition 42, $z\sigma \in \mathit{Vars} \setminus \operatorname{dom}(\sigma)$. Then we show that $z \in \operatorname{fvars}(\tau)$. If $z\sigma = x\sigma$ then $z\tau = y\sigma$. On the other hand, if $z\sigma \neq x\sigma$, we have $z\tau = z\sigma$. In both cases, $z\tau \in \mathit{Vars} \setminus \operatorname{dom}(\tau)$, which implies $z \in \operatorname{fvars}(\tau)$.

Lemma 63 *Let $d \in \mathit{SFL}$ and $\sigma \in \gamma_S(d) \cap \mathit{VSubst}$. Let also $y \in \mathit{VI}$ and $t \in \mathit{HTerms}$ be such that* $\operatorname{vars}(t) \subseteq \mathit{VI}$ *and* $y \notin \operatorname{share\_with}_d(t)$. *Then* $\operatorname{vars}(y\sigma) \cap \operatorname{vars}(t\sigma) \subseteq \operatorname{dom}(\sigma)$.

Proof. Let $d = \langle sh, f, l \rangle$ and $V_t = \operatorname{vars}(t)$ so that, by Definition 20, $y \notin \operatorname{vars}\bigl(\operatorname{rel}(V_t, sh)\bigr)$. Thus

$$\forall w \in V_t, S \in sh : \{y, w\} \nsubseteq S.$$

By Definition 18, since $\sigma \in \gamma_S(d)$, this implies

$$\forall v \in \mathit{Vars}, w \in V_t : \{y, w\} \nsubseteq \operatorname{occ}(\sigma, v).$$

Thus, since $\sigma \in \mathit{VSubst}$, by Proposition 42 we obtain

$$\forall w \in V_t : \operatorname{vars}(y\sigma) \cap \operatorname{vars}(w\sigma) \subseteq \operatorname{dom}(\sigma),$$

which is equivalent to the thesis.

Proposition 64 *Let $d \in \mathit{SFL}$, $\sigma \in \gamma_S(d) \cap \mathit{VSubst}$ and $(x \mapsto t) \in \mathit{Bind}$, where $\{x\} \cup \operatorname{vars}(t) \subseteq \mathit{VI}$. Suppose also that $x \in \operatorname{fvars}(\sigma)$. Then, for all $\tau \in \operatorname{mgs}\bigl(\sigma \cup \{x = t\}\bigr)$ in the syntactic equality theory T, we have*

$$\operatorname{fvars}(\sigma) \setminus \operatorname{share\_with}_d(x) \subseteq \operatorname{fvars}(\tau). \tag{61}$$

Proof. We assume that the congruence and identity axioms hold. Note that if $\sigma \cup \{x = t\}$ is not satisfiable, then the result is trivial. We therefore assume, for the rest of the proof, that $\sigma \cup \{x = t\}$ is satisfiable in T. It follows from Lemma 53 that we just have to show that there exists $\tau \in \operatorname{mgs}\bigl(\sigma \cup \{x = t\}\bigr)$ such that (61) holds.

As $x \in \operatorname{fvars}(\sigma)$ we have, using Proposition 42,

$$x\sigma \in \mathit{Vars} \setminus \operatorname{dom}(\sigma). \tag{62}$$

Suppose first $x\sigma = t\sigma$. Then we have $\sigma \in \operatorname{mgs}\bigl(\sigma \cup \{x\sigma = t\sigma\}\bigr)$, so that by Lemma 51, $\sigma \in \operatorname{mgs}\bigl(\sigma \cup \{x = t\}\bigr)$. Thus, by taking $\tau \stackrel{\mathrm{def}}{=} \sigma$, we trivially obtain $\operatorname{fvars}(\sigma) = \operatorname{fvars}(\tau)$, which implies the thesis.

Otherwise, let $x\sigma \neq t\sigma$ and take $\tau \stackrel{\mathrm{def}}{=} \sigma \cup \{x\sigma = t\sigma\}$. Then, since $\sigma \in \mathit{RSubst}$, it follows from (62) that $\tau \in \mathit{Eqs}$ has no identities or circular subsets so that $\tau \in \mathit{RSubst}$. By Lemma 51, $\tau \in \operatorname{mgs}\bigl(\sigma \cup \{x = t\}\bigr)$.

Suppose $y \in \mathrm{fvars}(\sigma) \setminus \mathrm{share\_with}_d(x)$. Then we show that $y \in \mathrm{fvars}(\tau)$. Since $y \in \mathrm{fvars}(\sigma)$, it follows by Proposition 42 that

$$y\sigma \in \mathit{Vars} \setminus \mathrm{dom}(\sigma). \tag{63}$$

Since $y \notin \mathrm{share\_with}_d(x)$ and $\sigma \in \mathit{VSubst}$, it follows by Lemma 63 that we have $\mathrm{vars}(y\sigma) \cap \mathrm{vars}(x\sigma) \subseteq \mathrm{dom}(\sigma)$. From this, by using (62) and (63), we obtain $y\sigma \neq x\sigma$; from this, again by (62), we derive $y \neq x\sigma$. Thus $y\tau = y\sigma$, so that $y \in \mathrm{fvars}(\tau)$.

Proposition 65 *Let $d \in \mathit{SFL}$, $\sigma \in \gamma_S(d) \cap \mathit{VSubst}$ and $(x \mapsto t) \in \mathit{Bind}$, where $\{x\} \cup \mathrm{vars}(t) \subseteq \mathit{VI}$. Then, for all $\tau \in \mathrm{mgs}\big(\sigma \cup \{x = t\}\big)$ in a syntactic equality theory T, we have*

$$\mathrm{fvars}(\sigma) \setminus \big(\mathrm{share\_with}_d(x) \cup \mathrm{share\_with}_d(t)\big) \subseteq \mathrm{fvars}(\tau). \tag{64}$$

Proof. We assume that the congruence and identity axioms hold. Note that if $\sigma \cup \{x = t\}$ is not satisfiable, then the result is trivial. We therefore assume, for the rest of the proof, that $\sigma \cup \{x = t\}$ is satisfiable in T. It follows from Lemma 53 that we just have to show that there exists $\tau \in \mathrm{mgs}\big(\sigma \cup \{x = t\}\big)$ such that (64) holds.

Let

$$\begin{aligned}
\{u_1, \ldots, u_k\} &\stackrel{\mathrm{def}}{=} \mathrm{dom}(\sigma) \cap \big(\mathrm{vars}(x\sigma) \cup \mathrm{vars}(t\sigma)\big),\\
\bar{s} &\stackrel{\mathrm{def}}{=} (u_1, \ldots, u_k, x\sigma),\\
\bar{t} &\stackrel{\mathrm{def}}{=} (u_1\sigma, \ldots, u_k\sigma, t\sigma).
\end{aligned}$$

Note that, since $\sigma \in \mathit{VSubst}$, for each $i = 1, \ldots, k$, we have

$$\mathrm{vars}(u_i\sigma) \subseteq \mathrm{vars}(x\sigma) \cup \mathrm{vars}(t\sigma).$$

Thus, for any $\mu \in \mathrm{mgs}(\bar{s} = \bar{t})$, we have

$$\mathrm{vars}(\mu) \subseteq \mathrm{vars}(x\sigma) \cup \mathrm{vars}(t\sigma). \tag{65}$$

Let

$$\begin{aligned}
\nu &\stackrel{\mathrm{def}}{=} \Big\{\, z = z\sigma\mu \;\Big|\; z \in \mathrm{dom}(\sigma) \setminus \big(\mathrm{vars}(x\sigma) \cup \mathrm{vars}(t\sigma)\big) \,\Big\},\\
\tau &\stackrel{\mathrm{def}}{=} \nu \cup \mu.
\end{aligned}$$

Then, as $\sigma, \mu \in \mathit{RSubst}$, it follows from (65) that $\nu, \tau \in \mathit{Eqs}$ have no identities or circular subsets so that $\nu, \tau \in \mathit{RSubst}$. Thus, using Lemma 51 and the assumption that $\sigma \cup \{x = t\}$ is satisfiable, $\tau \in \mathrm{mgs}\big(\sigma \cup \{x = t\}\big)$.

Suppose $y \in \mathrm{fvars}(\sigma) \setminus \big(\mathrm{share\_with}_d(x) \cup \mathrm{share\_with}_d(t)\big)$. We show that $y \in \mathrm{fvars}(\tau)$. As $y \in \mathrm{fvars}(\sigma)$, by Proposition 42,

$$y\sigma \in \mathit{Vars} \setminus \mathrm{dom}(\sigma). \tag{66}$$

As $y \notin \mathrm{share\_with}_d(x) \cup \mathrm{share\_with}_d(t)$, it follows from Lemma 63 that

$$\mathrm{vars}(y\sigma) \cap \big(\mathrm{vars}(x\sigma) \cup \mathrm{vars}(t\sigma)\big) \subseteq \mathrm{dom}(\sigma).$$

From this, by using (66), we obtain

$$y\sigma \notin \mathrm{vars}(x\sigma) \cup \mathrm{vars}(t\sigma). \tag{67}$$

By using either (67), if $y \notin \mathrm{dom}(\sigma)$, or the fact that $\sigma \in \mathit{VSubst}$, if $y \in \mathrm{dom}(\sigma)$, it follows that

$$y \notin \mathrm{vars}(x\sigma) \cup \mathrm{vars}(t\sigma). \tag{68}$$

Thus, by (68), $y\tau = y\nu$ and, by (67), $y\nu = y\sigma$, so that $y\tau = y\sigma$ and $y \in \mathrm{fvars}(\tau)$.

C.3 The Correctness for Linearity

Lemma 66 *Let $\sigma \in \mathit{VSubst}$ and $y \in \mathrm{dom}(\sigma)$. If $y \in \mathrm{lvars}(\sigma) \setminus \mathrm{gvars}(\sigma)$ then $y \notin \mathrm{vars}(y\sigma)$.*

Proof. Suppose, by contraposition, that $y \in \mathrm{vars}(y\sigma)$. Since $\sigma \in \mathit{VSubst}$ and $y \in \mathrm{lvars}(\sigma)$, by Proposition 42 we have

$$\mathrm{vars}(y\sigma) \cap \mathrm{dom}(\sigma) \subseteq \mathrm{gvars}(\sigma).$$

Therefore $y \in \mathrm{gvars}(\sigma)$, contradicting the assumption.

The following simple consequence of Proposition 16 will be used in the sequel.

Lemma 67 *If $\sigma \in \mathit{RSubst}$ then $\mathrm{fvars}(\sigma) \cup \mathrm{gvars}(\sigma) \subseteq \mathrm{lvars}(\sigma)$.*

Proof. Let $y \in \mathrm{fvars}(\sigma) \cup \mathrm{gvars}(\sigma)$. It follows from Proposition 16 that we have $\mathrm{rt}(y, \sigma) \in \mathit{Vars} \cup \mathit{GTerms}$. However, $\mathit{Vars} \cup \mathit{GTerms} \subset \mathit{LTerms}$ so that, again by Proposition 16, $y \in \mathrm{lvars}(\sigma)$.

Proposition 68 *Let $d \in \mathit{SFL}$, $\sigma \in \gamma_S(d) \cap \mathit{VSubst}$ and $(x \mapsto t) \in \mathit{Bind}$, where $\{x\} \cup \mathrm{vars}(t) \subseteq \mathit{VI}$ and $y \in \mathrm{dom}(\sigma) \cap \mathrm{range}(\sigma)$ implies $y \in \mathrm{vars}(y\sigma)$. Suppose that $\{r, r'\} = \{x, t\}$ and $\mathrm{lin}_d(r)$ holds. For all $\tau \in \mathrm{mgs}\bigl(\sigma \cup \{x = t\}\bigr)$ in the syntactic equality theory T, we have*

$$\mathrm{lvars}(\sigma) \setminus \mathrm{share\_with}_d(r) \subseteq \mathrm{lvars}(\tau). \tag{69}$$

Proof. We assume that the congruence and identity axioms hold. Note that if $\sigma \cup \{x = t\}$ is not satisfiable, then the result is trivial. We therefore assume, for the rest of the proof, that $\sigma \cup \{x = t\}$ is satisfiable in T. It follows from Lemma 53 that we just have to show that there exists $\tau \in \mathrm{mgs}\bigl(\sigma \cup \{x = t\}\bigr)$ such that (69) holds.

Since $\mathrm{lin}_d(r)$ holds, then $\mathrm{vars}(r) \subseteq \mathrm{lvars}(\sigma)$. Thus, by Proposition 42,

$$\forall v \in \mathrm{vars}(r\sigma) \setminus \mathrm{dom}(\sigma) : \mathrm{occ\_lin}(v, r\sigma), \tag{70}$$
$$\mathrm{vars}(r\sigma) \cap \mathrm{dom}(\sigma) \subseteq \mathrm{gvars}(\sigma). \tag{71}$$

Therefore, by defining $V_\sigma \stackrel{\mathrm{def}}{=} \mathit{Vars} \setminus \mathrm{gvars}(\sigma)$, we obtain that the term $r\sigma$ is V_σ-linear. Let

$$\{u_1, \ldots, u_k\} \stackrel{\mathrm{def}}{=} \mathrm{dom}(\sigma) \cap \bigl(\mathrm{vars}(x\sigma) \cup \mathrm{vars}(t\sigma)\bigr),$$
$$\bar{s} \stackrel{\mathrm{def}}{=} (u_1, \ldots, u_k, r\sigma),$$

$$\bar{t} \stackrel{\text{def}}{=} (u_1\sigma, \dots, u_k\sigma, r'\sigma).$$

Since $r\sigma$ is V_σ-linear it follows from (71) that $\bar{s}$ is V_σ-linear. By Lemma 51 and the congruence axioms, $\sigma \cup \{x = t\} \implies \bar{s} = \bar{t}$. Thus, as $\sigma \cup \{x = t\}$ is satisfiable, there exists $\mu \in \mathrm{mgs}(\bar{s} = \bar{t})$. Let $V_\mu = \mathit{Vars} \setminus \mathrm{gvars}(\mu)$; since $\mathrm{gvars}(\sigma) \subseteq \mathrm{gvars}(\mu)$, then $V_\mu \subseteq V_\sigma$ and $\bar{s}$ is also V_μ-linear. Therefore, we can apply Lemma 54 so that there exists $\mu \in \mathrm{mgs}(\bar{s} = \bar{t})$ such that, for all $w \in W \stackrel{\text{def}}{=} \mathit{Vars} \setminus \mathrm{vars}(\bar{s})$,

$$w\mu \text{ is } V_\mu\text{-linear}; \tag{72}$$

$$\mathrm{vars}(w\mu) \cap \mathrm{dom}(\mu) \subseteq \mathrm{gvars}(\mu); \tag{73}$$

$$\forall w' \in W : w \neq w' \implies \mathrm{vars}(w\mu) \cap \mathrm{vars}(w'\mu) \subseteq \mathrm{gvars}(\mu). \tag{74}$$

Note that, since $\sigma \in \mathit{VSubst}$, we have $\mathrm{vars}(u_i\sigma) \subseteq \mathrm{vars}(x\sigma) \cup \mathrm{vars}(t\sigma)$ for each $i = 1, \dots, k$. Therefore

$$\mathrm{vars}(\mu) \subseteq \mathrm{vars}(x\sigma) \cup \mathrm{vars}(t\sigma). \tag{75}$$

Let $\nu, \tau \subseteq \mathit{Eqs}$ be defined as

$$\nu \stackrel{\text{def}}{=} \left\{\, z = z\sigma\mu \;\middle|\; z \in \mathrm{dom}(\sigma) \setminus \bigl(\mathrm{vars}(x\sigma) \cup \mathrm{vars}(t\sigma)\bigr) \,\right\},$$

$$\tau \stackrel{\text{def}}{=} \nu \cup \mu.$$

Then, as $\sigma, \mu \in \mathit{RSubst}$, it follows from (75) that ν and τ have no identities or circular subsets so that $\nu, \tau \in \mathit{RSubst}$. By applying Lemma 51, we obtain $\tau \in \mathrm{mgs}\bigl(\sigma \cup \{x = t\}\bigr)$.

Suppose $y \in \mathrm{lvars}(\sigma) \setminus \mathrm{share\_with}_d(r)$. Then we show that $y \in \mathrm{lvars}(\tau)$.

If $y \in \mathrm{gvars}(\sigma)$ then $y \in \mathrm{gvars}(\tau)$. Thus, by Lemma 67, $y \in \mathrm{lvars}(\tau)$. Therefore, for the rest of the proof, we assume $y \in \mathrm{lvars}(\sigma) \setminus \mathrm{gvars}(\sigma)$. Thus, by Lemma 66, we have $y \in \mathrm{dom}(\sigma)$ implies $y \notin \mathrm{vars}(y\sigma)$ so that, by the hypothesis,

$$y \notin \mathrm{dom}(\sigma) \cap \mathrm{range}(\sigma). \tag{76}$$

As $y \in \mathrm{lvars}(\sigma)$, by Proposition 42,

$$\forall v \in \mathrm{vars}(y\sigma) \setminus \mathrm{dom}(\sigma) : \mathrm{occ\_lin}(v, y\sigma), \tag{77}$$

$$\mathrm{vars}(y\sigma) \cap \mathrm{dom}(\sigma) \subseteq \mathrm{gvars}(\sigma). \tag{78}$$

As $y \notin \mathrm{share\_with}_d(r)$, by Definition 20 and Theorem 21, we have $\mathrm{vars}(y\sigma) \cap \mathrm{vars}(r\sigma) \subseteq \mathrm{gvars}(\sigma)$, so that, by (78), we obtain

$$\mathrm{vars}(y\sigma) \cap \mathrm{vars}(\bar{s}) \subseteq \mathrm{gvars}(\sigma). \tag{79}$$

We now prove that $y \in \mathrm{lvars}(\tau)$, by showing that

$$\forall v \in \mathrm{vars}(y\tau) \setminus \mathrm{dom}(\tau) : \mathrm{occ\_lin}(v, y\tau), \tag{80}$$

$$\mathrm{vars}(y\tau) \cap \mathrm{dom}(\tau) \subseteq \mathrm{gvars}(\tau). \tag{81}$$

Since $\mathrm{dom}(\tau) = \mathrm{dom}(\nu) \cup \mathrm{dom}(\mu)$, we have three cases.

1. Suppose first that $y \notin \mathrm{dom}(\tau)$. Then (80) holds because $\mathrm{occ\_lin}(y, y)$ is always true; similarly, (81) is true, because $\mathrm{vars}(y\tau) = \{y\}$.

2. Next, suppose that $y \in \mathrm{dom}(\nu)$, so that $y\tau = y\nu = y\sigma\mu$.

   To prove (80) we have to prove, for all $w \in \mathrm{vars}(y\sigma)$,

$$\forall v \in \mathrm{vars}(w\mu) \setminus \mathrm{gvars}(\tau) : \mathrm{occ\_lin}(v, w\mu), \tag{82}$$

$$\forall w' \in \mathrm{vars}(y\sigma) \setminus \{w\} : \mathrm{vars}(w\mu) \cap \mathrm{vars}(w'\mu) \subseteq \mathrm{gvars}(\tau). \tag{83}$$

   We consider two subcases. If $w \in \mathrm{vars}(\bar{s})$ then, by (79), $w \in \mathrm{gvars}(\sigma)$. This implies $w \in \mathrm{gvars}(\tau)$ and, as a consequence, $\mathrm{vars}(w\mu) \subseteq \mathrm{gvars}(\tau)$. Thus, both (82) and (83) hold. Otherwise, if $w \notin \mathrm{vars}(\bar{s})$, then we have $w \in W$ and both (72) and (74) hold. Therefore, since $\mathrm{gvars}(\mu) \subseteq \mathrm{gvars}(\tau)$, (82) follows from (72). As for (83), this follows either from (74), when $w' \notin \mathrm{vars}(\bar{s})$, or from (79), when $w' \in \mathrm{vars}(\bar{s})$.

   In order to prove (73), note that

$$\mathrm{vars}(y\sigma\mu) \cap \mathrm{dom}(\tau) = \bigcup \bigl\{\, \mathrm{vars}(w\mu) \cap \mathrm{dom}(\tau) \bigm| w \in \mathrm{vars}(y\sigma) \,\bigr\}.$$

   We will prove that, for all $w \in \mathrm{vars}(y\sigma)$, $\mathrm{vars}(w\mu) \cap \mathrm{dom}(\tau) \subseteq \mathrm{gvars}(\tau)$. Let $w \in \mathrm{vars}(y\sigma)$. Suppose first that $w \in \mathrm{gvars}(\sigma)$. As a consequence, $w \in \mathrm{gvars}(\tau)$, which implies $\mathrm{vars}(w\tau) \subseteq \mathrm{gvars}(\tau)$. In particular, $\mathrm{vars}(w\tau) \cap \mathrm{dom}(\tau) \subseteq \mathrm{gvars}(\tau)$. Otherwise, let $w \notin \mathrm{gvars}(\sigma)$ so that, by (78), $w \notin \mathrm{dom}(\sigma)$. If also $w \notin \mathrm{dom}(\mu)$, then $w = w\mu \notin \mathrm{dom}(\tau)$ and there is nothing to prove. If $w \in \mathrm{dom}(\mu)$, then $\mathrm{vars}(w\mu) \subseteq \mathrm{vars}(\mu)$ so that, by (75) and the definition of ν, $\mathrm{vars}(w\mu) \cap \mathrm{dom}(\nu) = \varnothing$. Moreover, by (79), we have $w \notin \mathrm{vars}(\bar{s})$. Thus (73) applies so that, as $\mathrm{gvars}(\mu) \subseteq \mathrm{gvars}(\tau)$, $\mathrm{vars}(w\mu) \cap \mathrm{dom}(\tau) \subseteq \mathrm{gvars}(\tau)$.

3. Finally, suppose $y \in \mathrm{dom}(\mu)$, so that $y\tau = y\mu$. First, we prove that $y \notin \mathrm{vars}(\bar{s})$. In fact, by definition of $\bar{s}$, if $y \in \mathrm{vars}(\bar{s})$ then $y \in \mathrm{vars}(r\sigma) \cup \mathrm{vars}(r'\sigma)$. Now, if $y \in \mathrm{dom}(\sigma)$, then $y \in \mathrm{range}(\sigma)$, therefore contradicting (76). Otherwise, if $y \notin \mathrm{dom}(\sigma)$, then $y \in \mathrm{vars}(y\sigma)$ and, by (79), $y \in \mathrm{gvars}(\sigma)$, contradicting our previous assumption.

   Thus, we have $y \notin \mathrm{vars}(\bar{s})$, so that $y \in W$ and (72) holds. Note that (80) follows because $\mathrm{gvars}(\mu) \subseteq \mathrm{gvars}(\tau) \subseteq \mathrm{dom}(\tau)$, so that $\mathrm{vars}(y\tau) \setminus \mathrm{dom}(\tau) \subseteq V_\mu$.

   Similarly, (73) holds and we can obtain (81) by observing that $\mathrm{gvars}(\mu) \subseteq \mathrm{gvars}(\tau)$.

Proposition 69 *Let $d \in SFL$, $\sigma \in \gamma_S(d) \cap VSubst$ and $(x \mapsto t) \in Bind$, where $\{x\} \cup \mathrm{vars}(t) \subseteq VI$ and $y \in \mathrm{dom}(\sigma) \cap \mathrm{range}(\sigma)$ implies $y \in \mathrm{vars}(y\sigma)$. For all $\tau \in \mathrm{mgs}\bigl(\sigma \cup \{x = t\}\bigr)$ in the syntactic equality theory T, we have*

$$\mathrm{lvars}(\sigma) \setminus \bigl(\mathrm{share\_with}_d(x) \cup \mathrm{share\_with}_d(t)\bigr) \subseteq \mathrm{lvars}(\tau). \tag{84}$$

Proof. We assume that the congruence and identity axioms hold. Note that if $\sigma \cup \{x = t\}$ is not satisfiable, then the result is trivial. We therefore assume, for the rest of the proof, that $\sigma \cup \{x = t\}$ is satisfiable in T. It follows from Lemma 53 that we just have to show that there exists $\tau \in \mathrm{mgs}\bigl(\sigma \cup \{x = t\}\bigr)$ such that (84) holds.

Let

$$\{u_1, \ldots, u_k\} \stackrel{\mathrm{def}}{=} \mathrm{dom}(\sigma) \cap \bigl(\mathrm{vars}(x\sigma) \cup \mathrm{vars}(t\sigma)\bigr),$$

$$\bar{s} \stackrel{\text{def}}{=} (u_1, \dots, u_k, x\sigma),$$
$$\bar{t} \stackrel{\text{def}}{=} (u_1\sigma, \dots, u_k\sigma, t\sigma).$$

By Lemma 51 and the congruence axioms, $\sigma \cup \{x = t\} \implies \bar{s} = \bar{t}$. Thus, as $\sigma \cup \{x = t\}$ is satisfiable, there exists $\mu \in \mathrm{mgs}(\bar{s} = \bar{t})$. Note that, since $\sigma \in \mathit{VSubst}$, we have $\mathrm{vars}(u_i\sigma) \subseteq \mathrm{vars}(x\sigma) \cup \mathrm{vars}(t\sigma)$ for each $i = 1, \dots, k$. Therefore

$$\mathrm{vars}(\mu) \subseteq \mathrm{vars}(x\sigma) \cup \mathrm{vars}(t\sigma). \tag{85}$$

Let $\nu, \tau \subseteq \mathit{Eqs}$ be defined as

$$\nu \stackrel{\text{def}}{=} \Bigl\{\, z = z\sigma\mu \Bigm| z \in \mathrm{dom}(\sigma) \setminus \bigl(\mathrm{vars}(x\sigma) \cup \mathrm{vars}(t\sigma)\bigr) \,\Bigr\},$$
$$\tau \stackrel{\text{def}}{=} \nu \cup \mu.$$

Then, as $\sigma, \mu \in \mathit{RSubst}$, it follows from (85) that ν and τ have no identities or circular subsets so that $\nu, \tau \in \mathit{RSubst}$. By applying Lemma 51, we obtain $\tau \in \mathrm{mgs}\bigl(\sigma \cup \{x = t\}\bigr)$.

Suppose $y \in \mathrm{lvars}(\sigma) \setminus \bigl(\mathrm{share\_with}_d(x) \cup \mathrm{share\_with}_d(t)\bigr)$. Then we show that $y \in \mathrm{lvars}(\tau)$.

If $y \in \mathrm{gvars}(\sigma)$ then $y \in \mathrm{gvars}(\tau)$. Thus, by Lemma 67, $y \in \mathrm{lvars}(\tau)$. Therefore, for the rest of the proof, we assume $y \in \mathrm{lvars}(\sigma) \setminus \mathrm{gvars}(\sigma)$. Thus, by Lemma 66, we have $y \in \mathrm{dom}(\sigma)$ implies $y \notin \mathrm{vars}(y\sigma)$ so that, by the hypothesis,

$$y \notin \mathrm{dom}(\sigma) \cap \mathrm{range}(\sigma). \tag{86}$$

As $y \in \mathrm{lvars}(\sigma)$, by Proposition 42,

$$\forall v \in \mathrm{vars}(y\sigma) \setminus \mathrm{dom}(\sigma) : \mathrm{occ\_lin}(v, y\sigma), \tag{87}$$
$$\mathrm{vars}(y\sigma) \cap \mathrm{dom}(\sigma) \subseteq \mathrm{gvars}(\sigma). \tag{88}$$

As $y \notin \mathrm{share\_with}_d(x) \cup \mathrm{share\_with}_d(t)$, by Definition 20 and Theorem 21, we have

$$\mathrm{vars}(y\sigma) \cap \bigl(\mathrm{vars}(x\sigma) \cup \mathrm{vars}(t\sigma)\bigr) \subseteq \mathrm{gvars}(\sigma). \tag{89}$$

Moreover, by (85), this implies

$$\mathrm{vars}(y\sigma) \cap \mathrm{vars}(\mu) \subseteq \mathrm{gvars}(\sigma). \tag{90}$$

We now prove that $y \in \mathrm{lvars}(\tau)$, by showing that

$$\forall v \in \mathrm{vars}(y\tau) \setminus \mathrm{dom}(\tau) : \mathrm{occ\_lin}(v, y\tau), \tag{91}$$
$$\mathrm{vars}(y\tau) \cap \mathrm{dom}(\tau) \subseteq \mathrm{gvars}(\tau). \tag{92}$$

Since $\mathrm{dom}(\tau) = \mathrm{dom}(\nu) \cup \mathrm{dom}(\mu)$, we have three cases.

1. Suppose first that $y \notin \mathrm{dom}(\tau)$. Then (91) holds because $\mathrm{occ\_lin}(y, y)$ is always true; similarly, (92) is true, because $\mathrm{vars}(y\tau) = \{y\}$.

2. Next, suppose that $y \in \mathrm{dom}(\nu)$, so that $y\tau = y\nu = y\sigma\mu$.

   To prove (91), consider $v \in \mathrm{vars}(y\sigma\mu) \setminus \mathrm{dom}(\tau)$. If $v \in \mathrm{vars}(\mu)$ then there exists a variable $w \in \mathrm{vars}(y\sigma) \cap \mathrm{vars}(\mu)$ such that $v \in \mathrm{vars}(w\mu)$.

By (90), we have $w \in \mathrm{gvars}(\sigma)$, which implies $w \in \mathrm{gvars}(\tau)$. Thus, $\mathrm{vars}(w\mu) \subseteq \mathrm{gvars}(\tau) \subseteq \mathrm{dom}(\tau)$, therefore contradicting our assumption that $v \notin \mathrm{dom}(\tau)$. Therefore, $v \notin \mathrm{vars}(\mu)$, so that $v \in \mathrm{vars}(y\sigma)$. Since $\mathrm{dom}(\sigma) \subseteq \mathrm{dom}(\tau)$, by (87), we have $\mathrm{occ\_lin}(v, y\sigma)$. Moreover, for all $w \in \mathrm{vars}(y\sigma) \cap \mathrm{dom}(\mu)$, we have $v \notin \mathrm{vars}(w\mu)$, because $v \notin \mathrm{vars}(\mu)$. Thus, we obtain $\mathrm{occ\_lin}(v, y\sigma\mu)$.

In order to prove (92), note that

$$\mathrm{vars}(y\sigma\mu) \cap \mathrm{dom}(\tau) = \bigcup \bigl\{ \mathrm{vars}(w\mu) \cap \mathrm{dom}(\tau) \bigm| w \in \mathrm{vars}(y\sigma) \bigr\}.$$

We will prove that, for all $w \in \mathrm{vars}(y\sigma)$, $\mathrm{vars}(w\mu) \cap \mathrm{dom}(\tau) \subseteq \mathrm{gvars}(\tau)$. Let $w \in \mathrm{vars}(y\sigma)$. If $w \in \mathrm{vars}(\mu)$ then, by (90), $w \in \mathrm{gvars}(\sigma)$, which implies $w \in \mathrm{gvars}(\tau)$. Thus, we obtain $\mathrm{vars}(w\mu) \subseteq \mathrm{gvars}(\tau)$. Otherwise, let $w \notin \mathrm{vars}(\mu)$, so that $w = w\mu$. Thus, $w \notin \mathrm{dom}(\mu)$. If also $w \notin \mathrm{dom}(\sigma)$, then $w \notin \mathrm{dom}(\tau)$ and there is nothing to prove. If $w \in \mathrm{dom}(\sigma)$, by (88), $w \in \mathrm{gvars}(\sigma)$, which implies $w \in \mathrm{gvars}(\tau)$. Thus, $\mathrm{vars}(w\mu) \subseteq \mathrm{gvars}(\tau)$.

3. Finally, suppose $y \in \mathrm{dom}(\mu)$, so that $y\tau = y\mu$. Suppose also that $y \in \mathrm{vars}(x\sigma) \cup \mathrm{vars}(t\sigma)$. Now, if $y \in \mathrm{dom}(\sigma)$, then $y \in \mathrm{range}(\sigma)$, thus contradicting (86). Otherwise, if $y \notin \mathrm{dom}(\sigma)$, then $y \in \mathrm{vars}(y\sigma)$ and, by (89), $y \in \mathrm{gvars}(\sigma)$, contradicting our previous assumption.

   Thus, we have $y \notin \mathrm{vars}(x\sigma) \cup \mathrm{vars}(t\sigma)$, so that, by (85), we obtain $y\mu = y$. As a consequence, (91) holds trivially. As for (92), this follows from (88) if $y \in \mathrm{dom}(\sigma)$, while being trivial if $y \notin \mathrm{dom}(\sigma)$.

C.4 Putting Results Together

By exploiting the correctness results regarding each of the three components of the domain *SFL*, we now prove the correctness of the amgu_S operator. We start by proving a restricted result that only applies to variable-idempotent substitutions.

Lemma 70 *Let $d = \langle sh, f, l \rangle \in SFL$ and $\sigma \in \gamma_S(d) \cap VSubst$ be such that $y \in \mathrm{dom}(\sigma) \cap \mathrm{range}(\sigma)$ implies $y \in \mathrm{vars}(y\sigma)$. Let also $(x \mapsto t) \in Bind$, where $\mathrm{vars}(\sigma) \cup \mathrm{vars}(x \mapsto t) \subseteq VI$, and suppose there exists $\tau \in \mathrm{mgs}\bigl(\sigma \cup \{x = t\}\bigr)$ in the syntactic equality theory T. Then $\tau \in \gamma_S\bigl(\mathrm{amgu}_S(d, x \mapsto t)\bigr)$.*

Proof. Let $d' = \langle sh', f', l' \rangle \stackrel{\mathrm{def}}{=} \mathrm{amgu}_S(d, x \mapsto t)$. Note that, by the existence of σ, τ as specified in the hypotheses, we have both $d \neq \bot_S$ and $d' \neq \bot_S$. Since $\sigma \in \gamma_S(d)$, it follows from Definition 18 that $\mathrm{ssets}(\sigma) \subseteq sh$, $\mathrm{fvars}(\sigma) \supseteq f$ and $\mathrm{lvars}(\sigma) \supseteq l$. Therefore, to prove $\tau \in \gamma_S(d')$, we have to show that

$$\mathrm{ssets}(\tau) \subseteq sh', \tag{93}$$
$$\mathrm{fvars}(\tau) \supseteq f', \tag{94}$$
$$\mathrm{lvars}(\tau) \supseteq l'. \tag{95}$$

We prove each inclusion separately.

(93). By Definition 22, we have $sh' = \mathrm{cyclic}_x^t(sh_- \cup sh'')$. We will show that

$$\mathrm{ssets}(\tau) \subseteq sh_- \cup sh''. \tag{96}$$

From this, the thesis will follow by Proposition 61. To prove (96) we need to consider five cases.

1. $\mathrm{free}_d(x) \lor \mathrm{free}_d(t)$ holds.

   We can apply Proposition 56, taking $r = x$ when $\mathrm{free}_d(x)$ holds and $r = t$ when $\mathrm{free}_d(x)$ does not hold, to conclude that $\mathrm{ssets}(\tau) \subseteq \mathit{sh}'$.

2. $\mathrm{lin}_d(x) \land \mathrm{lin}_d(t)$ holds.

   We can apply Proposition 57, obtaining

$$\mathrm{ssets}(\tau) \subseteq \mathit{sh}_- \cup \mathrm{bin}\bigl(\mathit{sh}_x \cup \mathrm{bin}(\mathit{sh}_x, \mathit{sh}_{xt}^\star), \mathit{sh}_t \cup \mathrm{bin}(\mathit{sh}_t, \mathit{sh}_{xt}^\star)\bigr) \stackrel{\mathrm{def}}{=} \mathit{sh}'.$$

3. $\mathrm{lin}_d(x)$ holds.

   We can apply Proposition 59, taking $r = x$ to conclude that

$$\mathrm{ssets}(\tau) \subseteq \mathit{sh}_- \cup \mathrm{bin}(\mathit{sh}_x^\star, \mathit{sh}_t) \stackrel{\mathrm{def}}{=} \mathit{sh}'.$$

4. $\mathrm{lin}_d(t)$ holds. This case is symmetric to the previous one.

5. When nothing is known about x and t, we can apply Theorem 50 to conclude that

$$\mathrm{ssets}(\tau) \subseteq \mathit{sh}_- \cup \mathrm{bin}(\mathit{sh}_x^\star, \mathit{sh}_t^\star) \stackrel{\mathrm{def}}{=} \mathit{sh}'.$$

(94). In order to show that $\mathrm{fvars}(\tau) \supseteq f'$, according to Definition 22, we consider four cases.

1. $\mathrm{free}_d(x) \land \mathrm{free}_d(t)$ holds.

   By Definition 20, $\{x, t\} \subseteq f \subseteq \mathrm{fvars}(\sigma)$. Therefore we can apply Proposition 62 (where y is replaced by $t \in \mathit{VI}$) to conclude that

$$f' \stackrel{\mathrm{def}}{=} f \subseteq \mathrm{fvars}(\sigma) \subseteq \mathrm{fvars}(\tau).$$

2. $\mathrm{free}_d(x)$ holds.

   By Definition 20, $x \in f \subseteq \mathrm{fvars}(\sigma)$. Therefore we can apply Proposition 64 to conclude that

$$\begin{aligned} f' &\stackrel{\mathrm{def}}{=} f \setminus \mathrm{share\_with}_d(x) \\ &\subseteq \mathrm{fvars}(\sigma) \setminus \mathrm{share\_with}_d(x) \\ &\subseteq \mathrm{fvars}(\tau). \end{aligned}$$

3. $\mathrm{free}_d(t)$ holds.

   This case is symmetric to the previous one.

4. When nothing is known about x and t, we can apply Proposition 65 to conclude that

$$\begin{aligned} f' &\stackrel{\mathrm{def}}{=} f \setminus \bigl(\mathrm{share\_with}_d(x) \cup \mathrm{share\_with}_d(t)\bigr) \\ &\subseteq \mathrm{fvars}(\sigma) \setminus \bigl(\mathrm{share\_with}_d(x) \cup \mathrm{share\_with}_d(t)\bigr) \\ &\subseteq \mathrm{fvars}(\tau). \end{aligned}$$

(95). In order to show that $\mathrm{lvars}(\tau) \supseteq l'$, according to Definition 22, we start by proving $\mathrm{lvars}(\tau) \supseteq l''$. There are four cases that have to be considered.

1. $\mathrm{lin}_d(x) \wedge \mathrm{lin}_d(t)$ holds.

   We can apply Proposition 68 twice, the first time taking $r = x$ and the second time taking $r = t$, to conclude that

$$
\begin{aligned}
l'' &\stackrel{\mathrm{def}}{=} l \setminus \bigl(\mathrm{share\_with}_d(x) \cap \mathrm{share\_with}_d(t)\bigr) \\
&= \bigl(l \setminus \mathrm{share\_with}_d(x)\bigr) \cup \bigl(l \setminus \mathrm{share\_with}_d(t)\bigr) \\
&\subseteq \bigl(\mathrm{lvars}(\sigma) \setminus \mathrm{share\_with}_d(x)\bigr) \cup \bigl(\mathrm{lvars}(\sigma) \setminus \mathrm{share\_with}_d(t)\bigr) \\
&\subseteq \mathrm{lvars}(\tau).
\end{aligned}
$$

2. $\mathrm{lin}_d(x)$ holds.

   We can apply Proposition 68 (where we take $r = x$) to conclude that

$$
\begin{aligned}
l'' &\stackrel{\mathrm{def}}{=} l \setminus \mathrm{share\_with}_d(x) \\
&\subseteq \mathrm{lvars}(\sigma) \setminus \mathrm{share\_with}_d(x) \\
&\subseteq \mathrm{lvars}(\tau).
\end{aligned}
$$

3. $\mathrm{lin}_d(t)$ holds.

   This case is symmetric to the previous one.

4. When nothing is known about x and t, we can apply Proposition 69 to conclude that

$$
\begin{aligned}
l'' &\stackrel{\mathrm{def}}{=} l \setminus \bigl(\mathrm{share\_with}_d(x) \cup \mathrm{share\_with}_d(t)\bigr) \\
&\subseteq \mathrm{lvars}(\sigma) \setminus \bigl(\mathrm{share\_with}_d(x) \cup \mathrm{share\_with}_d(t)\bigr) \\
&\subseteq \mathrm{lvars}(\tau).
\end{aligned}
$$

Therefore, we have $\mathrm{lvars}(\tau) \supseteq l''$. By (93) proved above, $\mathrm{ssets}(\tau) \subseteq sh'$. By Definition 10, $\mathit{VI} \setminus \mathrm{vars}(sh') \subseteq \mathrm{gvars}(\tau)$. Moreover, by (94) proved above, we have $\mathrm{fvars}(\tau) \supseteq f'$. Thus, by applying Lemma 67 we obtain the thesis:

$$
\begin{aligned}
\mathrm{lvars}(\tau) &\supseteq \mathrm{gvars}(\tau) \cup \mathrm{fvars}(\tau) \cup \mathrm{lvars}(\tau) \\
&\supseteq \bigl(\mathit{VI} \setminus \mathrm{vars}(sh')\bigr) \cup f' \cup l'' \\
&\stackrel{\mathrm{def}}{=} l'.
\end{aligned}
$$

Finally, by exploiting the results of Subsection 3.1, we drop the assumption about variable-idempotent substitutions.

Proof of Theorem 23 on page 14. Let $d' = \mathrm{amgu}_S(d, x \mapsto t)$.

If $d = \bot_S$ then we have $d' = \bot_S$ and the result holds trivially, since $\gamma_S(d) = \varnothing$. Similarly, if $T = \mathcal{FT}$ is the theory of finite trees and $x \in \mathrm{vars}(t)$, then $d' = \bot_S$. Again, the result holds trivially, since the equation $\{x = t\}$ is not satisfiable in $\mathcal{FT}$, so that $\mathrm{mgs}\bigl(\sigma \cup \{x = t\}\bigr) = \varnothing$.

Therefore suppose there exists $\sigma \in \gamma_S(d)$ and $\tau \in \mathrm{mgs}\bigl(\sigma \cup \{x = t\}\bigr)$. By Proposition 40, there exists $\sigma' \in \mathit{VSubst}$ such that $T \vdash \forall(\sigma \leftrightarrow \sigma')$ and $y \in \mathrm{dom}(\sigma') \cap \mathrm{range}(\sigma')$ implies $y \in \mathrm{vars}(y\sigma')$. By Corollary 19 and Definition 18, we have $\sigma \in \gamma_S(d)$ if and only if $\sigma' \in \gamma_S(d)$. Therefore, the result follows by application of Lemma 70.

D Proofs of the Results of Section 4

We will use results from [3], that show that the domain *PSD* is the weakest abstraction of *SH* achieving the same precision for the computation of groundness and pair-sharing.

Theorem 71 *Suppose* $sh_1, sh_2 \in SH$*. Then, for each* $(x \mapsto t) \in Bind$*,*

$$\rho(sh_1) = \rho(sh_2) \implies \rho\bigl(\mathrm{amgu}(sh_1, x \mapsto t)\bigr) = \rho\bigl(\mathrm{amgu}(sh_2, x \mapsto t)\bigr).$$

Theorem 72 *Let* $sh_1, sh_2 \in SH$ *be such that* $\rho_{PSD}(sh_1) \neq \rho_{PSD}(sh_2)$*. Then there exist* $\sigma \in RSubst$ *such that*

$$\rho_{PSD}\bigl(\mathrm{aunify}(sh_1, \sigma)\bigr) \neq \rho_{PSD}\bigl(\mathrm{aunify}(sh_2, \sigma)\bigr).$$

The following result in [3] shows that, for groundness and pair-sharing information, the exponential star-union operator can be replaced by a quadratic operation without loss of precision.

Theorem 73 *Let* $sh \in SH$ *and* $(x \mapsto t) \in Bind$*. Let also*

$$sh_- \stackrel{\mathrm{def}}{=} \overline{\mathrm{rel}}\bigl(\{x\} \cup \mathrm{vars}(t), sh\bigr), \quad sh_x \stackrel{\mathrm{def}}{=} \mathrm{rel}\bigl(\{x\}, sh\bigr), \quad sh_t \stackrel{\mathrm{def}}{=} \mathrm{rel}\bigl(\mathrm{vars}(t), sh\bigr).$$

Then

$$\rho\bigl(\mathrm{amgu}(sh, x \mapsto t)\bigr) = \rho\bigl(sh_- \cup \mathrm{bin}(sh_x^2, sh_t^2)\bigr).$$

The next three lemmas show that the precision of the abstract evaluation of the operators specified in Definition 20 is not affected by ρ_{PSD}.

Lemma 74 *For each* $V \subseteq VI$ *and* $sh \in SH$ *it holds*

$$\mathrm{vars}\bigl(\mathrm{rel}(V, sh)\bigr) = \mathrm{vars}\Bigl(\mathrm{rel}\bigl(V, \rho_{PSD}(sh)\bigr)\Bigr).$$

Proof. If $V = \varnothing$, the result is trivial. Thus, assume $V \neq \varnothing$.

The first inclusion ($\subseteq$) follows from the extensivity of ρ_{PSD} and the monotonicity of the operators rel and vars. To prove the other inclusion, let $S \in \mathrm{rel}\bigl(V, \rho_{PSD}(sh)\bigr)$. By Definition 6, we have

$$\forall x \in S : S = \bigcup \bigl\{\, T \in sh \bigm| \{x\} \subseteq T \subseteq S \,\bigr\}.$$

In particular, for all $x \in S \cap V$, it holds

$$\begin{aligned}
S &= \bigcup \bigl\{\, T \in sh \bigm| \{x\} \subseteq T \subseteq S \,\bigr\} \\
&= \bigcup \bigl\{\, T \in \mathrm{rel}(V, sh) \bigm| \{x\} \subseteq T \subseteq S \,\bigr\} \\
&\subseteq \bigcup \mathrm{rel}(V, sh) \\
&= \mathrm{vars}\bigl(\mathrm{rel}(V, sh)\bigr).
\end{aligned}$$

Since the choice of S was arbitrary, we obtain the desired inclusion

$$\mathrm{vars}\bigl(\mathrm{rel}(V, sh)\bigr) \supseteq \mathrm{vars}\Bigl(\mathrm{rel}\bigl(V, \rho_{PSD}(sh)\bigr)\Bigr).$$

Lemma 75 *For each $V, W \subseteq VI$ and $sh \in SH$ it holds*

$$\bigl(\mathrm{rel}(V, sh) \cap \mathrm{rel}(W, sh) = \varnothing\bigr) \iff \Bigl(\mathrm{rel}\bigl(V, \rho_{\scriptscriptstyle PSD}(sh)\bigr) \cap \mathrm{rel}\bigl(W, \rho_{\scriptscriptstyle PSD}(sh)\bigr) = \varnothing\Bigr).$$

Proof. To prove the first implication ($\Rightarrow$), we reason by contraposition and suppose

$$\mathrm{rel}\bigl(V, \rho_{\scriptscriptstyle PSD}(sh)\bigr) \cap \mathrm{rel}\bigl(W, \rho_{\scriptscriptstyle PSD}(sh)\bigr) \neq \varnothing.$$

Thus, there exists $S \in \rho_{\scriptscriptstyle PSD}(sh)$ such that $S \cap V \neq \varnothing$ and $S \cap W \neq \varnothing$. Consider $x \in S \cap V$ and $y \in S \cap W$, so that we have $\{x, y\} \subseteq S$.

By Definition 6, we have

$$\forall v \in S : S = \bigcup \bigl\{\, T \in sh \bigm| \{v\} \subseteq T \subseteq S \,\bigr\}.$$

In particular, by taking $v = x$, there must exist a sharing group $T \in sh$ such that $\{x, y\} \subseteq T$, so that

$$T \in \mathrm{rel}(V, sh) \cap \mathrm{rel}(W, sh) \neq \varnothing.$$

The other implication ($\Leftarrow$) follows by the extensivity of $\rho_{\scriptscriptstyle PSD}$ and the monotonicity of rel.

Lemma 76 *For each $i \in \{1, 2\}$, let $d_i = \langle sh_i, f, l\rangle \in SFL$ and suppose that $\rho_{\scriptscriptstyle PSD}(sh_1) = \rho_{\scriptscriptstyle PSD}(sh_2)$. Then, for all $s, t \in HTerms$ and $y \in VI$,*

$$\mathrm{ind}_{d_1}(s, t) \iff \mathrm{ind}_{d_2}(s, t); \tag{97}$$
$$\mathrm{free}_{d_1}(t) \iff \mathrm{free}_{d_2}(t); \tag{98}$$
$$\mathrm{ground}_{d_1}(t) \iff \mathrm{ground}_{d_2}(t); \tag{99}$$
$$\mathrm{occ\_lin}_{d_1}(y, t) \iff \mathrm{occ\_lin}_{d_2}(y, t); \tag{100}$$
$$\mathrm{lin}_{d_1}(t) \iff \mathrm{lin}_{d_2}(t); \tag{101}$$
$$\mathrm{share\_with}_{d_1}(t) = \mathrm{share\_with}_{d_2}(t). \tag{102}$$

Proof. Consider equivalence (97) and let $V = \mathrm{vars}(s)$, $W = \mathrm{vars}(t)$. By Definition 20, Lemma 75 and the hypothesis, we obtain

$$\begin{aligned}
\mathrm{ind}_{d_1}(s, t) &\iff \mathrm{rel}(V, sh_1) \cap \mathrm{rel}(W, sh_1) \neq \varnothing \\
&\iff \mathrm{rel}\bigl(V, \rho_{\scriptscriptstyle PSD}(sh_1)\bigr) \cap \mathrm{rel}\bigl(W, \rho_{\scriptscriptstyle PSD}(sh_1)\bigr) \neq \varnothing \\
&\iff \mathrm{rel}\bigl(V, \rho_{\scriptscriptstyle PSD}(sh_2)\bigr) \cap \mathrm{rel}\bigl(W, \rho_{\scriptscriptstyle PSD}(sh_2)\bigr) \neq \varnothing \\
&\iff \mathrm{rel}(V, sh_2) \cap \mathrm{rel}(W, sh_2) \neq \varnothing \\
&\iff \mathrm{ind}_{d_2}(s, t).
\end{aligned}$$

The proof of (98) follows easily from Definition 20, since the predicate $\mathrm{free}_{d_i}(t)$ does not depend on the sharing component sh_i of d_i.

Consider now (99). By Definition 20, Lemma 74 and the hypothesis, we obtain

$$\begin{aligned}
\mathrm{ground}_{d_1}(t) &\iff \mathrm{vars}(t) \subseteq VI \setminus \mathrm{vars}(sh_1) \\
&\iff \mathrm{vars}(t) \subseteq VI \setminus \mathrm{vars}\bigl(\mathrm{rel}(VI, sh_1)\bigr) \\
&\iff \mathrm{vars}(t) \subseteq VI \setminus \mathrm{vars}\Bigl(\mathrm{rel}\bigl(VI, \rho_{\scriptscriptstyle PSD}(sh_1)\bigr)\Bigr)
\end{aligned}$$

$$
\begin{aligned}
&\iff \operatorname{vars}(t) \subseteq VI \setminus \operatorname{vars}\Bigl(\operatorname{rel}\bigl(VI, \rho_{\scriptscriptstyle PSD}(sh_2)\bigr)\Bigr) \\
&\iff \operatorname{vars}(t) \subseteq VI \setminus \operatorname{vars}\bigl(\operatorname{rel}(VI, sh_2)\bigr) \\
&\iff \operatorname{vars}(t) \subseteq VI \setminus \operatorname{vars}(sh_2) \\
&\iff \operatorname{ground}_{d_2}(t).
\end{aligned}
$$

The proof of (100) follows easily from Definition 20, by applying the equivalences (97) and (99). Similarly, the proof of (101) follows from Definition 20 and (100). Finally, equation (102) follows from Definition 20 and Lemma 74.

Since both ρ (by [22, Theorem 7]) and $(\cdot)^\star$ are upper closure operators it follows that

$$
sh_1 \subseteq \rho(sh_2) \iff \rho(sh_1) \subseteq \rho(sh_2), \tag{103}
$$

$$
sh_1 \subseteq sh_2^\star \iff sh_1^\star \subseteq sh_2^\star. \tag{104}
$$

The following lemma needed below is stated and proved as Lemma 19 in [22].

Lemma 77 *For each $sh_1, sh_2 \in SH$ and each $V \in \wp_{\mathrm{f}}(\mathit{Vars})$,*

$$
sh_1 \subseteq \rho(sh_2) \implies \operatorname{rel}(V, sh_1)^\star \subseteq \operatorname{rel}(V, sh_2)^\star.
$$

Lemma 78 *Let $sh_1, sh_2 \in SH$ be such that $sh_1 \subseteq \rho_{\scriptscriptstyle PSD}(sh_2)$. For each $V, W \subseteq VI$ and each $i \in \{1, 2\}$, let also*

$$
\begin{aligned}
sh_{-,i} &= \overline{\operatorname{rel}}(V \cup W, sh_i), \\
sh_{x,i} &= \operatorname{rel}(V, sh_i), \\
sh_{t,i} &= \operatorname{rel}(W, sh_i).
\end{aligned}
$$

Then, we have

$$
\operatorname{bin}(sh_{x,1}, sh_{t,1}) \subseteq \rho_{\scriptscriptstyle PSD}\bigl(sh_{-,2} \cup \operatorname{bin}(sh_{x,2}, sh_{t,2})\bigr); \tag{105}
$$

$$
\operatorname{bin}(sh_{x,1}, sh_{t,1}^\star) \subseteq \rho_{\scriptscriptstyle PSD}\bigl(sh_{-,2} \cup \operatorname{bin}(sh_{x,2}, sh_{t,2}^\star)\bigr). \tag{106}
$$

Proof. Let $S \stackrel{\mathrm{def}}{=} S_x \cup S_t \in \operatorname{bin}(sh_{x,1}, sh_{t,1})$ where $S_x \in sh_{x,1}$ and $S_t \in sh_{t,1}$. Consider an arbitrary variable $y \in S$.

Suppose first that $y \in S_x$ and let $w \in W \cap S_t$. Since $S_x, S_t \in sh_1$, by hypothesis $S_x, S_t \in \rho_{\scriptscriptstyle PSD}(sh_2)$. By Definition 6, $S = \bigcup(A \cup B)$, where

$$
\begin{aligned}
A &\stackrel{\mathrm{def}}{=} \bigl\{\, S' \in sh_2 \bigm| \{y\} \subseteq S' \subseteq S_x \,\bigr\}, \\
B &\stackrel{\mathrm{def}}{=} \bigl\{\, S' \in sh_2 \bigm| \{w\} \subseteq S' \subseteq S_t \,\bigr\}.
\end{aligned}
$$

In particular, we can write $A \cup B = sh'_{-} \cup sh'_x \cup sh'_t$, where

$$
\begin{aligned}
sh'_{-} &\stackrel{\mathrm{def}}{=} \overline{\operatorname{rel}}(V \cup W, A), \\
sh'_x &\stackrel{\mathrm{def}}{=} \operatorname{rel}(V, A), \\
sh'_t &\stackrel{\mathrm{def}}{=} \operatorname{rel}(W, A \cup B).
\end{aligned}
$$

Since $V \cap S_x \neq \varnothing$ and $v \in W \cap S_t$, we have $sh'_x \neq \varnothing$ and $sh'_t \neq \varnothing$. Therefore $\bigcup(sh'_x \cup sh'_t) = \bigcup \operatorname{bin}(sh'_x, sh'_t)$, so that

$$S = \bigcup\bigl(sh'_- \cup \operatorname{bin}(sh'_x, sh'_t)\bigr).$$

By construction, we have $\{y\} \subseteq S' \subseteq S$ for all $S' \in A$. Thus, it also holds $\{y\} \subseteq S' \subseteq S$ for all $S' \in sh'_- \cup \operatorname{bin}(sh'_x, sh'_t)$, so that

$$S = \bigcup\bigl\{\, S' \in sh'_- \cup \operatorname{bin}(sh'_x, sh'_t) \bigm| \{y\} \subseteq S' \subseteq S \,\bigr\}.$$

By a symmetric argument, the same conclusion can be obtained when $y \in S_t$. As the choice of y was arbitrary, by Definition 6,

$$S \in \rho_{\scriptscriptstyle PSD}\bigl(sh'_- \cup \operatorname{bin}(sh'_x, sh'_t)\bigr). \tag{107}$$

Note that $sh'_- \subseteq sh_{-,2}$, $sh'_x \subseteq sh_{x,2}$ and $sh'_t \subseteq sh_{t,2}$, so that it holds

$$sh'_- \cup \operatorname{bin}(sh'_x, sh'_t) \subseteq sh_{-,2} \cup \operatorname{bin}(sh_{x,2}, sh_{t,2}).$$

Then, (105) follows from (107) by the monotonicity of $\rho_{\scriptscriptstyle PSD}$.

To prove (106), let $S \stackrel{\text{def}}{=} S_x \cup T_t \in \operatorname{bin}(sh_{x,1}, sh^{\star}_{t,1})$, where $S_x \in sh_{x,1}$ and $T_t \in sh^{\star}_{t,1}$. Consider an arbitrary variable $y \in S$.

Suppose that $y \in S_x$. Since $S_x \in sh_1$, by hypothesis $S_x \in \rho_{\scriptscriptstyle PSD}(sh_2)$, so that by Definition 6 we have $S = \bigcup\bigl(A \cup \{T_t\}\bigr)$, where

$$A \stackrel{\text{def}}{=} \bigl\{\, S' \in sh_2 \bigm| \{y\} \subseteq S' \subseteq S_x \,\bigr\}.$$

In particular, we can write $S = \bigcup(sh'_- \cup sh'_x \cup sh'_t)$, where

$$\begin{aligned} sh'_- &\stackrel{\text{def}}{=} \overline{\operatorname{rel}}(V \cup W, A),\\ sh'_x &\stackrel{\text{def}}{=} \operatorname{rel}(V, A),\\ sh'_t &\stackrel{\text{def}}{=} \operatorname{rel}(W, A) \cup \{T_t\}. \end{aligned}$$

Since $V \cap S_x \neq \varnothing$, we have $sh'_x \neq \varnothing$. Since it also holds $sh'_t \neq \varnothing$, we have $\bigcup\bigl(sh'_x \cup sh'_t\bigr) = \bigcup \operatorname{bin}\bigl(sh'_x, sh'_t\bigr)$, so that

$$S = \bigcup\bigl(sh'_- \cup \operatorname{bin}(sh'_x, sh'_t)\bigr).$$

By construction, we have $\{y\} \subseteq S' \subseteq S$ for all $S' \in A$. Thus, it also holds $\{y\} \subseteq S' \subseteq S$ for all $S' \in sh'_- \cup \operatorname{bin}(sh'_x, sh'_t)$, so that

$$S = \bigcup\bigl\{\, S' \in sh'_- \cup \operatorname{bin}(sh'_x, sh'_t) \bigm| \{y\} \subseteq S' \subseteq S \,\bigr\}. \tag{108}$$

Suppose now that $y \in T_t$ and let $v \in V \cap S_x$. Since $S_x \in sh_1 \subseteq \rho_{\scriptscriptstyle PSD}(sh_2)$, by Definition 6 we have $S = \bigcup\bigl(B \cup \{T_t\}\bigr)$, where

$$B \stackrel{\text{def}}{=} \bigl\{\, S' \in sh_2 \bigm| \{v\} \subseteq S' \subseteq S_x \,\bigr\}.$$

In particular, we can write $S = \bigcup(sh'_- \cup sh'_x \cup sh'_t)$, where

$$sh'_- \stackrel{\text{def}}{=} \varnothing,$$

$$sh'_x \stackrel{\text{def}}{=} \text{rel}(V, B) = B,$$
$$sh'_t \stackrel{\text{def}}{=} \{T_t\}.$$

Since $v \in V \cap S_x$, we have $sh'_x \neq \varnothing$; since it also holds $sh'_t \neq \varnothing$, we obtain $\bigcup(sh'_x \cup sh'_t) = \bigcup\bigl(\text{bin}(sh'_x, sh'_t)\bigr)$. Since $y \in T_t$ and $sh'_- = \varnothing$, we have $\{y\} \subseteq S' \subseteq S$ for all $S' \in sh'_- \cup \text{bin}(sh'_x, sh'_t)$, so that (108) holds even in this case.

As the choice of y was arbitrary, by (108) and Definition 6,

$$S \in \rho_{\scriptscriptstyle PSD}\bigl(sh'_- \cup \text{bin}(sh'_x, sh'_t)\bigr). \tag{109}$$

Clearly, $sh'_- \subseteq sh_{-,2}$ and $sh'_x \subseteq sh_{x,2}$; also, by Lemma 77, $T_t \in sh^{\star}_{t,2}$, so that $sh'_t \subseteq sh^{\star}_{t,2}$. Thus, $sh'_- \cup \text{bin}(sh'_x, sh'_t) \subseteq sh_{-,2} \cup \text{bin}(sh_{x,2}, sh^{\star}_{t,2})$. The thesis (106) follows from (109) by the monotonicity of $\rho_{\scriptscriptstyle PSD}$.

Lemma 79 *Let $sh \in SH$ and $V, W \subseteq VI$. Let also $sh_x \stackrel{\text{def}}{=} \text{rel}(V, sh)$, $sh_t \stackrel{\text{def}}{=} \text{rel}(W, sh)$, $sh_{xt} \stackrel{\text{def}}{=} sh_x \cap sh_t$ and*

$$sh^{\diamond} \stackrel{\text{def}}{=} \text{bin}\bigl(sh_x \cup \text{bin}(sh_x, sh^{\star}_{xt}), sh_t \cup \text{bin}(sh_t, sh^{\star}_{xt})\bigr).$$

Then, $\rho_{\scriptscriptstyle PSD}(sh^{\diamond}) = \rho_{\scriptscriptstyle PSD}\bigl(\text{bin}(sh_x, sh_t)\bigr)$.

Proof. Observe that, since $sh_{xt} \subseteq sh_x$ and $sh_{xt} \subseteq sh_t$,

$$\text{bin}\bigl(\text{bin}(sh_x, sh^{\star}_{xt}), \text{bin}(sh_t, sh^{\star}_{xt})\bigr) = \text{bin}\bigl(\text{bin}(sh_x, sh_t), sh^{\star}_{xt}\bigr).$$

Thus

$$sh^{\diamond} = \text{bin}(sh_x, sh_t) \cup \text{bin}\bigl(\text{bin}(sh_x, sh_t), sh^{\star}_{xt}\bigr). \tag{110}$$

Thus, the inclusion $\rho_{\scriptscriptstyle PSD}(sh^{\diamond}) \supseteq \rho_{\scriptscriptstyle PSD}\bigl(\text{bin}(sh_x, sh_t)\bigr)$ follows by the monotonicity of $\rho_{\scriptscriptstyle PSD}$. We now prove the other inclusion

$$\rho_{\scriptscriptstyle PSD}(sh^{\diamond}) \subseteq \rho_{\scriptscriptstyle PSD}\bigl(\text{bin}(sh_x, sh_t)\bigr). \tag{111}$$

Let $S \in sh^{\diamond}$. Then, by (110), $S = S_x \cup S_t \cup T_{xt}$, where $S_x \in sh_x$, $S_t \in sh_t$, $T_{xt} \in sh^{\star}_{xt} \cup \varnothing$. Thus, for some $k \geq 0$, $T_{xt} = T_1 \cup \cdots \cup T_k$, where $T_i \in sh_{xt}$ for each $i = 1, \ldots, k$.

Consider an arbitrary variable $y \in S$. We will show that

$$S = \bigcup\bigl\{\, S' \in \text{bin}(sh_x, sh_t) \bigm| \{y\} \subseteq S' \subseteq S \,\bigr\}. \tag{112}$$

Suppose first that $y \in S_x$. Then, since

$$S = (S_x \cup S_t) \cup (S_x \cup T_1) \cup \cdots \cup (S_x \cup T_k),$$

it follows that (112) holds. By a similar argument, (112) holds also when $y \in S_t$, by taking

$$S = (S_x \cup S_t) \cup (S_t \cup T_1) \cup \cdots \cup (S_t \cup T_k).$$

On the other hand, suppose now $y \notin S_x \cup S_t$, so that $k > 0$ and there exists $j \in \{1, \ldots, k\}$ such that $y \in T_j$. In this case we have

$$S = (S_x \cup T_j) \cup (T_j \cup S_t) \cup (T_1 \cup T_j) \cup \cdots \cup (T_k \cup T_j),$$

so that, since $T_j \in sh_{xt}$, (112) still holds.

As the choice of $y \in S$ was arbitrary, by (112) and Definition 6 it follows $S \in \rho_{\scriptscriptstyle PSD}\bigl(\text{bin}(sh_x, sh_t)\bigr)$. Thus, by monotonicity and idempotence of $\rho_{\scriptscriptstyle PSD}$, (111) holds.

Lemma 80 *Let $sh \in SH$ and $(x \mapsto t) \in Bind$, where $\{x\} \cup \mathrm{vars}(t) \subseteq VI$. Consider $W = \mathrm{vars}(t) \setminus \{x\}$ and let $sh_x, sh_t, sh_W \in SH$ be defined as*

$$\begin{aligned} sh_x &= \mathrm{rel}\bigl(\{x\}, sh\bigr), \\ sh_t &= \mathrm{rel}\bigl(\mathrm{vars}(t), sh\bigr), \\ sh_W &= \mathrm{rel}(W, sh). \end{aligned}$$

Then

$$\mathrm{bin}(sh_x, sh_W) = \mathrm{cyclic}_x^t\bigl(\mathrm{bin}(sh_x, sh_t)\bigr); \tag{113}$$

$$\mathrm{bin}(sh_x^\star, sh_W) = \mathrm{cyclic}_x^t\bigl(\mathrm{bin}(sh_x^\star, sh_t)\bigr); \tag{114}$$

$$\mathrm{bin}(sh_x^\star, sh_W^\star) = \mathrm{cyclic}_x^t\bigl(\mathrm{bin}(sh_x^\star, sh_t^\star)\bigr). \tag{115}$$

Proof. We start by proving equations (113) and (114) at the same time. Therefore, let $sh'_x \in \{sh_x, sh_x^\star\}$.

To prove the first inclusions ($\subseteq$), we assume $S \in \mathrm{bin}(sh'_x, sh_W)$ and show that $S \in \mathrm{cyclic}_x^t\bigl(\mathrm{bin}(sh'_x, sh_t)\bigr)$. Since $sh_W \subseteq sh_t$, we have $S \in \mathrm{bin}(sh'_x, sh_t)$. Moreover, $S = S_x \cup S_W$, where $x \in S_x$ and $W \cap S_W \neq \varnothing$. Thus $W \cap S \neq \varnothing$ and hence, by Definition 20, $S \in \mathrm{cyclic}_x^t\bigl(\mathrm{bin}(sh'_x, sh_t)\bigr)$.

To prove the opposite inclusions ($\supseteq$), let $S \in \mathrm{cyclic}_x^t\bigl(\mathrm{bin}(sh'_x, sh_t)\bigr)$. Then, $S \in \mathrm{bin}(sh'_x, sh_t)$, so that $S = S_x \cup S_t$, where $S_x \in sh'_x$ and $S_t \in sh_t$. Thus $x \in S$ and, by Definition 20, $S \in \mathrm{rel}(W, sh)$. If $\mathrm{vars}(t) \cap S_t \neq \{x\}$, then $S_t \in sh_W$ and $S \in \mathrm{bin}(sh'_x, sh_W)$, so that the two inclusions hold. Otherwise, let $\mathrm{vars}(t) \cap S_t = \{x\}$, so that $S_t \in sh'_x$ and $S_t \notin sh_W$. Since we know $S \in \mathrm{rel}(W, sh)$, there exists $w \in W \cap S_x$, so that $S_x \in sh_W$. First, consider the case when $sh'_x = sh_x$. Then $S_t \in sh_x$ and we have $S \in \mathrm{bin}(sh_x, sh_W)$, proving the second inclusion for (113). Secondly, consider the case when $sh'_x = sh_x^\star$. Then we can write $S_x = S_W \cup S_x$, where $S_W \in sh_W$. Thus $S = (S_x \cup S_t) \cup S_W \in \mathrm{bin}(sh_x^\star, sh_W)$, proving the second inclusion for (114).

Finally, we prove equation (115).

To prove the first inclusion ($\subseteq$), we assume $S \in \mathrm{bin}(sh_x^\star, sh_W^\star)$ and show that $S \in \mathrm{cyclic}_x^t\bigl(\mathrm{bin}(sh_x^\star, sh_t^\star)\bigr)$. As $sh_W \subseteq sh_t$, we have $S \in \mathrm{bin}(sh_x^\star, sh_t^\star)$. Moreover, $S = S_x \cup S_W$, where $x \in S_x$ and $W \cap S_W \neq \varnothing$. Thus $W \cap S \neq \varnothing$ and hence, by Definition 20, $S \in \mathrm{cyclic}_x^t\bigl(\mathrm{bin}(sh_x^\star, sh_t^\star)\bigr)$.

To prove the opposite inclusion ($\supseteq$), let $S \in \mathrm{cyclic}_x^t\bigl(\mathrm{bin}(sh_x^\star, sh_t^\star)\bigr)$. Then, we have $S \in \mathrm{bin}(sh_x^\star, sh_t^\star)$, so that $S = S_x \cup S_t$, where $S_x \in sh_x^\star$ and $S_t \in sh_t^\star$. Thus $x \in S$ and, by Definition 20, $S \in \mathrm{rel}(W, sh)$. Suppose first that $\mathrm{vars}(t) \cap S_t \neq \{x\}$. Then $S_t = S_W \cup S_{xt}$, where $S_W \in sh_W^\star$ and $S_{xt} \in sh_x^\star \cup \{\varnothing\}$. Thus $S = (S_x \cup S_{xt}) \cup S_W \in \mathrm{bin}(sh_x^\star, sh_W^\star)$. Suppose next that $\mathrm{vars}(t) \cap S_t = \{x\}$, so that $S_t \in sh_x^\star$ $W \cap S_x \neq \varnothing$. Then $S_x = S_W \cup S_x$, where $S_W \in sh_W^\star$. Thus $S = (S_x \cup S_t) \cup S_W \in \mathrm{bin}(sh_x^\star, sh_W^\star)$, completing the proof.

For the next theorem, we will use the following lemma corresponding to Lemmas 18 in [3].

Lemma 81 *For each $sh \in SH$ and $V \in \wp_{\mathrm{f}}(Vars)$, $\overline{\mathrm{rel}}\bigl(V, \rho(sh)\bigr) = \rho\bigl(\overline{\mathrm{rel}}(V, sh)\bigr)$.*

Theorem 82 *Let $d_1, d_2 \in SFL$ be such that $\rho_{PSD}(d_1) = \rho_{PSD}(d_2)$. Then, for all $(x \mapsto t) \in Bind$,*

$$\rho_{PSD}\bigl(\mathrm{amgu}_S(d_1, x \mapsto t)\bigr) = \rho_{PSD}\bigl(\mathrm{amgu}_S(d_2, x \mapsto t)\bigr).$$

Proof. Let $d_1 = \langle sh_1, f, l \rangle$. Then, by definition of $\rho_{\scriptscriptstyle PSD}$ on *SFL*, it holds $d_2 = \langle sh_2, f, l \rangle$, where $\rho_{\scriptscriptstyle PSD}(sh_1) = \rho_{\scriptscriptstyle PSD}(sh_2)$.

For each $i \in \{1, 2\}$, let $\langle sh'_i, f'_i, l'_i \rangle = \mathrm{amgu}_S\bigl(d_i, x \mapsto t\bigr)$. We will prove the following results:

$$\rho_{\scriptscriptstyle PSD}(sh'_1) = \rho_{\scriptscriptstyle PSD}(sh'_2), \tag{116}$$

$$f'_1 = f'_2, \tag{117}$$

$$l'_1 = l'_2. \tag{118}$$

Equation (116). We will prove the result

$$sh_1 \subseteq \rho_{\scriptscriptstyle PSD}(sh_2) \implies sh'_1 \subseteq \rho_{\scriptscriptstyle PSD}(sh'_2). \tag{119}$$

Then, by using (104), we obtain

$$\rho_{\scriptscriptstyle PSD}(sh_1) \subseteq \rho_{\scriptscriptstyle PSD}(sh_2) \implies \rho_{\scriptscriptstyle PSD}(sh'_1) \subseteq \rho_{\scriptscriptstyle PSD}(sh'_2),$$

from which the thesis follows by symmetry.

Let $W = \mathrm{vars}(t) \setminus \{x\}$. For each $i \in \{1, 2\}$, let

$$\begin{aligned}
sh_{-,i} &= \overline{\mathrm{rel}}\bigl(\{x\} \cup \mathrm{vars}(t), sh_i\bigr),\\
sh_{x,i} &= \mathrm{rel}\bigl(\{x\}, sh_i\bigr),\\
sh_{t,i} &= \mathrm{rel}\bigl(\mathrm{vars}(t), sh_i\bigr),\\
sh_{xt,i} &= sh_{x,i} \cap sh_{t,i},\\
sh_{W,i} &= \mathrm{rel}(W, sh_i).
\end{aligned}$$

To prove (119), assume that $sh_1 \subseteq \rho_{\scriptscriptstyle PSD}(sh_2)$. By Definitions 22 and 20, for each $i = 1$, 2 we have

$$sh'_i = \mathrm{cyclic}_x^t(sh_{-,i} \cup sh''_i) = sh_{-,i} \cup \mathrm{cyclic}_x^t(sh''_i).$$

We first show that $sh_{-,1} \subseteq \rho_{\scriptscriptstyle PSD}\bigl(sh_{-,2} \cup \mathrm{cyclic}_x^t(sh''_2)\bigr)$. By the definition of $sh_{-,1}$, the assumption and the monotonicity of $\overline{\mathrm{rel}}$, we have

$$sh_{-,1} \subseteq \overline{\mathrm{rel}}\bigl(\{x\} \cup \mathrm{vars}(t), \rho_{\scriptscriptstyle PSD}(sh_2)\bigr).$$

Thus, by Lemma 81, $sh_{-,1} \subseteq \rho_{\scriptscriptstyle PSD}(sh_{-,2})$, from which the required result follows by monotonicity of $\rho_{\scriptscriptstyle PSD}$.

We next show that $\mathrm{cyclic}_x^t(sh''_1) \subseteq \rho_{\scriptscriptstyle PSD}\bigl(sh_{-,2} \cup \mathrm{cyclic}_x^t(sh''_2)\bigr)$. By applying cases (98) and (101) of Lemma 76, it can be seen that sh''_1 and sh''_2 are each computed by selecting the same alternative branch of Definition 22. We have five cases.

1. In the first case, for each $i = 1$, 2, we have $sh''_i = \mathrm{bin}(sh_{x,i}, sh_{t,i})$. By case (113) of Lemma 80, $\mathrm{cyclic}_x^t(sh''_i) = \mathrm{bin}(sh_{x,i}, sh_{W,i})$, for each $i = 1$, 2. Thus, by case (105) of Lemma 78, where we take $V = \{x\}$,

   $$\mathrm{bin}(sh_{x,1}, sh_{W,1}) \subseteq \rho_{\scriptscriptstyle PSD}\bigl(sh_{-,2} \cup \mathrm{bin}(sh_{x,2}, sh_{W,2})\bigr),$$

   from which the thesis follows.

2. In the second case we have, for each $i = 1, 2$,

$$sh_i'' = \mathrm{bin}\bigl(sh_{x,i} \cup \mathrm{bin}(sh_{x,i}, sh^\star_{xt,i}), sh_{t,i} \cup \mathrm{bin}(sh_{t,i}, sh^\star_{xt,i})\bigr).$$

   There are two cases.

   First suppose that $x \notin \mathrm{vars}(t)$, so that $\mathrm{cyclic}_x^t(sh_i'') = sh_i''$. Then, by Lemma 79, for each $i = 1, 2$, we have $sh_i'' \subseteq \rho_{\scriptscriptstyle PSD}\bigl(\mathrm{bin}(sh_{x,i}, sh_{t,i})\bigr)$. Therefore, by case (105) of Lemma 78 and the monotonicity of $\rho_{\scriptscriptstyle PSD}$, we obtain

$$sh_1'' \subseteq \rho_{\scriptscriptstyle PSD}\bigl(sh_{-,2} \cup \mathrm{bin}(sh_{x,2}, sh_{t,2})\bigr),$$

   so that the thesis holds.

   Secondly, suppose that $x \in \mathrm{vars}(t)$. In this case, for each $i = 1, 2$, we have $sh_{xt,i} = sh_{x,i}$, so that $sh_i'' = \mathrm{bin}(sh^\star_{x,i}, sh_{t,i})$. This case is therefore equivalent to the third case, proven below.

3. In the third case, for each $i = 1, 2$, we have $sh_i'' = \mathrm{bin}(sh^\star_{x,i}, sh_{t,i})$. By case (114) of Lemma 80, $\mathrm{cyclic}_x^t(sh_i'') = \mathrm{bin}(sh^\star_{x,i}, sh_{W,i})$. Thus, by case (106) of Lemma 78, where we take $V = \{x\}$, we obtain

$$sh_1'' \subseteq \rho_{\scriptscriptstyle PSD}\bigl(sh_{-,2} \cup \mathrm{bin}(sh^\star_{x,2}, sh_{W,2})\bigr),$$

   so that the thesis holds.

4. In the fourth case, for each $i = 1, 2$, we have $sh_i'' = \mathrm{bin}(sh_{x,i}, sh^\star_{t,i})$. Moreover, as $\mathrm{lin}_d(t)$ holds and $\mathrm{lin}_d(x)$ does not hold, we can assume that $x \notin \mathrm{vars}(t)$, so that $\mathrm{cyclic}_x^t(sh_i'') = sh_i''$. Thus, by case (106) of Lemma 78, where we exchange the usual roles of V and W, we obtain

$$sh_1'' \subseteq \rho_{\scriptscriptstyle PSD}\bigl(sh_{-,2} \cup \mathrm{bin}(sh_{x,2}, sh^\star_{t,2})\bigr),$$

   so that the thesis holds.

5. In the fifth case we have, for $i = 1, 2$, $sh_i'' = \mathrm{bin}(sh^\star_{x,i}, sh^\star_{t,i})$. By case (115) of Lemma 80, $\mathrm{cyclic}_x^t(sh_i'') = \mathrm{bin}(sh^\star_{x,i}, sh^\star_{W,i})$. The thesis follows from Theorem 71, by replacing the term t by an arbitrary term $t' \in \mathit{HTerms}$ such that $\mathrm{vars}(t') = W$.

Equation (117). Consider the computation of f_i' as specified in Definition 22. By applying case (98) of Lemma 76, it can be seen that f_1' and f_2' are computed by selecting the same alternative branch. The thesis $f_1' = f_2'$ thus follows from case (102) of Lemma 76.

Equation (118). Consider the computation of l_i' as specified in Definition 22: for each $i \in \{1, 2\}$ we have

$$l_i' = \bigl(\mathit{VI} \setminus \mathrm{vars}(sh_i')\bigr) \cup f_i' \cup l_i''.$$

Let $r \in \mathit{HTerms}$ be such that $\mathrm{vars}(r) = \mathit{VI}$. Then, for each $i \in \{1, 2\}$, we have $\mathrm{vars}(sh_i') = \mathrm{share\_with}_{d_i'}(r)$; also, by equation (116), we know that $\rho_{\scriptscriptstyle PSD}(sh_1') = \rho_{\scriptscriptstyle PSD}(sh_2')$; thus, by case (102) of Lemma 76, we obtain $\mathrm{vars}(sh_1') = \mathrm{vars}(sh_2')$. By equation (117), we also know that $f_1' = f_2'$. Therefore, to complete the proof, we only need to prove that $l_1'' = l_2''$. Consider the computation of l_i'' as specified in Definition 22. By case (101) of Lemma 76, it can be seen that, in the computations of l_1'' and l_2'', the same alternative branch is selected. Hence, the thesis is obtained by applying case (102) of Lemma 76.

Theorem 83 *Let $d_1, d_2 \in SFL$ be such that $\rho_{PSD}(d_1) = \rho_{PSD}(d_2)$. Then, for each sequence of bindings $bs \in Bind^\star$,*

$$\rho_{PSD}\bigl(\mathrm{aunify}_S(d_1, bs)\bigr) = \rho_{PSD}\bigl(\mathrm{aunify}_S(d_2, bs)\bigr).$$

Proof. The proof is by induction on the length of bs. The base case, when $|bs| = 0$ and thus $bs = \epsilon$, is obvious from the definition of aunify_S. For the inductive case, when $|bs| = m > 0$, let $bs = (x \mapsto t) \,.\, bs'$. By the hypothesis and Theorem 82, we have

$$\rho_{PSD}\bigl(\mathrm{amgu}_S(d_1, x \mapsto t)\bigr) = \rho_{PSD}\bigl(\mathrm{amgu}_S(d_2, x \mapsto t)\bigr). \tag{120}$$

Moreover, for each $i \in \{1, 2\}$, by definition of aunify_S we have

$$\mathrm{aunify}_S(d_i, bs) = \mathrm{aunify}_S\bigl(\mathrm{amgu}_S(d_i, x \mapsto t), bs'\bigr).$$

Thus, by (120), we can apply the inductive hypothesis and conclude the proof, since $|bs'| = m - 1 < m$.

For the next theorem, we will use the following lemma, corresponding to Lemma 24 in [3].

Lemma 84 *Let $sh_1, sh_2 \in SH$ be such that $\rho_{PSD}(sh_1) = \rho_{PSD}(sh_2)$. Then, for each $V \subseteq VI$,*

$$\rho_{PSD}\bigl(\mathrm{aexists}(sh_1, V)\bigr) = \rho_{PSD}\bigl(\mathrm{aexists}(sh_2, V)\bigr).$$

Theorem 85 *Let $d_1, d_2 \in SH$ be such that $\rho_{PSD}(d_1) = \rho_{PSD}(d_2)$. Then, for each $V \subseteq VI$,*

$$\rho_{PSD}\bigl(\mathrm{aexists}_S(d_1, V)\bigr) = \rho_{PSD}\bigl(\mathrm{aexists}_S(d_2, V)\bigr).$$

Proof. Let $d_i = \langle sh_i, f_i, l_i \rangle$, for each $i = 1$, 2. By applying Definitions 27 and 29, for each $i = 1$, 2, we have

$$\begin{aligned}\rho_{PSD}\bigl(\mathrm{aexists}_S(d_i, V)\bigr) &= \rho_{PSD}\Bigl(\bigl\langle \mathrm{aexists}(sh_i), f_i \cup V, l_i \cup V \bigr\rangle\Bigr) \\ &= \Bigl\langle \rho_{PSD}\bigl(\mathrm{aexists}(sh_i)\bigr), f_i \cup V, l_i \cup V \Bigr\rangle.\end{aligned}$$

By the hypothesis and Definition 29, we also have $\rho_{PSD}(sh_1) = \rho_{PSD}(sh_2)$, $f_1 = f_2$ and $l_1 = l_2$. Thus, to complete the proof, we only need to show that

$$\rho_{PSD}\bigl(\mathrm{aexists}(sh_1)\bigr) = \rho_{PSD}\bigl(\mathrm{aexists}(sh_2)\bigr).$$

This follows from Lemma 84.

Proof of Theorem 30 on page 18. The congruence properties for aunify_S and $\mathrm{aexists}_S$ follow from Theorems 83 and 85, respectively. The congruence property for alub_S holds, as usual, because ρ_{PSD} is an upper closure operator.

Proof of Theorem 31 on page 19. Suppose $\rho_{PSD}(d_1) \neq \rho_{PSD}(d_2)$. By Definition 29, we have three cases:

1. Suppose $\rho_{\scriptscriptstyle PSD}(sh_1) \neq \rho_{\scriptscriptstyle PSD}(sh_2)$. Let t be a ground and finite term and let

$$\sigma \stackrel{\mathrm{def}}{=} \{\, x \mapsto t \mid x \in \mathit{VI} \setminus S \,\}.$$

   Since σ binds all of its domain variables to terms that are ground and finite, then no binary union and/or star-union needs to be computed. As a consequence, the behavior of amgu_S on the sharing component is the same as the behavior of 'amgu'. The proof therefore follows from Theorem 72.

2. Suppose now $f_1 \neq f_2$. In this case, by taking $\rho = \rho_F$ and $bs = \epsilon$, we obtain

$$\begin{aligned} \rho_F\bigl(\mathrm{aunify}_S(d_1, \epsilon)\bigr) &= \rho_F(d_1) \\ &= \langle SG, f_1, \varnothing \rangle \\ &\neq \langle SG, f_2, \varnothing \rangle \\ &= \rho_F(d_2) \\ &= \rho_F\bigl(\mathrm{aunify}_S(d_2, \epsilon)\bigr). \end{aligned}$$

3. Finally, suppose $l_1 \neq l_2$. Similarly to the previous case, by taking $\rho = \rho_L$ and $bs = \epsilon$, we obtain

$$\begin{aligned} \rho_L\bigl(\mathrm{aunify}_S(d_1, \epsilon)\bigr) &= \rho_L(d_1) \\ &= \langle SG, \varnothing, l_1 \rangle \\ &\neq \langle SG, \varnothing, l_2 \rangle \\ &= \rho_L(d_2) \\ &= \rho_L\bigl(\mathrm{aunify}_S(d_2, \epsilon)\bigr). \end{aligned}$$

Proof of Theorem 32 on page 19. Suppose first that $x \notin \mathrm{vars}(t)$. Then it holds

$$\mathrm{cyclic}_x^t(sh_- \cup sh^\diamond) = sh_- \cup sh^\diamond,$$

so that the thesis is a simple corollary of Lemma 79, where $V = \{x\}$ and $W = \mathrm{vars}(t)$.

Suppose now $x \in \mathrm{vars}(t)$. Then we have $sh_x = sh_{xt}$, so that $sh^\diamond = \mathrm{bin}(sh_x^\star, sh_t)$. In this case the thesis is a corollary of Theorem 73.