# An Isabelle/HOL Formalization of Narrowing and Multiset Narrowing for $E$-Unifiability, Reachability and Infeasibility

**Dohan Kim**
Department of Computer Science, University of Innsbruck, Austria

**Abstract**

We present an Isabelle/HOL formalization of narrowing for $E$-unifiability, reachability, and infeasibility. Given a semi-complete rewrite system $\mathcal{R}$ and two terms $s$ and $t$, we show a formalized proof that if narrowing terminates, then it provides a decision procedure for $\mathcal{R}$-unifiability for $s$ and $t$, where $\mathcal{R}$ is viewed as a set of equations. Furthermore, we present multiset narrowing and its formalization for multiset reachability and reachability analysis, providing decision procedures using certain restricted conditions on multiset reachability and reachability problems. Our multiset narrowing also provides a complete method for $E$-unifiability problems consisting of multiple goals if $E$ can be represented by a complete rewrite system.

**2012 ACM Subject Classification** Theory of computation → Logic and verification; Theory of computation → Equational logic and rewriting

**Keywords and phrases** Narrowing, Multiset Narrowing, Unifiability, Reachability

**Digital Object Identifier** 10.4230/LIPIcs.ITP.2024.24

**Funding** Austrian Science Fund (FWF) project I5943.

**Acknowledgements** The author would like to thank René Thiemann for his valuable comments and discussion on narrowing with special encoding and multiset narrowing.

## 1 Introduction

*Narrowing* [13,18,23] generalizes rewriting in the sense that matching is replaced by unification. Narrowing is a widely used technique for solving $E$-unification problems using term rewriting systems, where equational unification (or $E$-unification) is concerned with making terms equivalent w.r.t. an equational theory $E$ [4]. For example, consider $E = \{f(x, 0) \approx x\}$. Then, two terms $f(y, z)$ and 0 are not syntactically unifiable, but they are $E$-unifiable using the substitution $\theta := \{y \mapsto 0, z \mapsto 0\}$ because $f(y, z)\theta = f(0, 0) \approx_E 0$. Given a complete rewrite system $\mathcal{R}$ representing $E$, narrowing is known to be *complete* for $E$-unification in the sense that for every solution of a given $E$-unification problem for $s$ and $t$, a more general solution can be found by narrowing [18]. It is also known that the semi-completeness of $\mathcal{R}$ suffices for the completeness of narrowing w.r.t. $E$-unification [23, 30].

In logic programming [20] and constraint based theorem proving [19, 25], it is often sufficient to decide the solvability of $E$-unification problems, called *$E$-unifiability* [29]. Given a set of equations $E$ and two terms $s$ and $t$, it is generally undecidable whether there exists a substitution $\sigma$ such that $s\sigma \approx_E t\sigma$ holds or not [4]. It is a natural question to ask when this $E$-unifiability problem is decidable. $E$-unifiability using narrowing was considered in [29] using a complete rewrite system $\mathcal{R}$. However, it focuses on the complexity result of narrowing w.r.t. $E$-unifiability, where narrowing is used as a complete semi-decision procedure for $E$-unifiability.

Given a semi-complete rewrite system $\mathcal{R}$ corresponding to $E$, we present a new formalized proof that (ordinary) narrowing may provide a decision procedure for $E$-unifiability if it terminates. Roughly speaking, if the narrowing procedure terminates, then it either reaches the *success* state or not. If it reaches the success state, then we show that it yields an $E$-unifier. Otherwise, we show that there is no $E$-unifier. We provide this correctness proof of narrowing for the $E$-unifiability problem in the proof assistant Isabelle/HOL.

Narrowing was originally studied in the context of equational unification, but later it was also studied in the context of the *reachability problem* [14, 22, 27, 28]. Given a rewrite system $\mathcal{R}$ and two terms $s$ and $t$, the reachability problem is stated as follows: is there a substitution $\sigma$ such that $s\sigma \to_{\mathcal{R}}^* t\sigma$? We say that this reachability problem is *satisfiable* if there is such a substitution $\sigma$. If no such a substitution exists, then this problem is said to be *infeasible* [21].

Narrowing is known to be *weakly complete* [22] for reachability analysis in the sense that it can find all $\mathcal{R}$ normalized solutions if some reasonable executability assumptions on $\mathcal{R}$ are provided. In [28], the authors proposed a semi-decision procedure, called *back-and-force narrowing*, for solving reachability goals, which is guaranteed to find a solution if it exists.

In this paper, we provide a formalized proof of some sufficient conditions of satisfying reachability problems using ordinary narrowing. Also, given a semi-complete TRS $\mathcal{R}$ and two terms $s$ and $t$, where $t$ is a *strongly-irreducible term* [7] (e.g. a constructor term), we show a formalized proof that if narrowing terminates, then it can provide a decision procedure whether the reachability problem from $s$ to $t$ is satisfiable or infeasible.

Ordinary narrowing (without special encoding) has some limitations on $E$-unifiability and reachability analysis. In particular, it is not (directly) applicable to $E$-unifiability and reachability analysis consisting of multiple goals. $E$-unification consisting of multiple goals is considered in [15, 24] using inference rules, but they are not concerned with $E$-unifiability consisting of multiple goals. Meanwhile, reachability analysis consisting of multiple goals is considered in [22, 28], but they are not concerned with $E$-unification/$E$-unifiability.

One may also use narrowing with special encoding for considering multiple narrowing goals. For example, if $u_1$ (resp. $v_1$) and $u_2$ (resp. $v_2$) are $E$-unifiable, then $\bar{f}(u_1, u_2)$ and $\bar{f}(v_1, v_2)$ are also $E$-unifiable, where $\bar{f}$ is a new symbol. This encoding is applicable to narrowing-based $E$-unification/$E$-unifiability consisting of multiple goals (cf. [9, 11, 12]), but has some limitations on reachability and multiset reachability analysis, which will be discussed later in this paper.

We present multiset narrowing based on multiset rewriting in order to generalize narrowing in multiset setting because identical elements (or states) in a multiset can reach different elements (or states). For example, consider the multiset $S = \{f(x, y), f(x, y)\}$, the (renamed) rewrite system $\mathcal{R} = \{f(a, b) \to d, f(a, z_1) \to g(z_1), f(z_2, a) \to d, g(a) \to c\}$, the target multiset $G = \{c, d\}$, and a variant of a reachability problem: is there a substitution $\sigma$ such that $S\sigma$ can reach $G$ by $\mathcal{R}$? If we simply use the rule $f(a, b) \to d$ using the substitution $\{x \mapsto a, y \mapsto b\}$, then $f(x, y)\sigma$ reaches $d$ but it does not reach $c$ using the rewrite steps by $\mathcal{R}$. Using multiset narrowing discussed later in this paper, we can find a substitution $\sigma = \{x \mapsto a, y \mapsto a\}$, which allows $S\sigma$ to reach $G$ using the rewriting steps by $\mathcal{R}$, i.e., multiset narrowing provides a means to solve multiset reachability problems.

Furthermore, both $E$-unifiability and reachability analysis are considered in the unified multiset narrowing framework. Our multiset narrowing works on multisets of ordinary terms for multiset reachability analysis, multisets of equational terms for $E$-unification/$E$-unifiability along with certain restricted cases of reachability analysis, and multisets of pairs of terms for reachability analysis. It is applicable to $E$-unification problems and (ordinary) reachability problems consisting of multiple goals, which is generic in the sense that it simply encapsulates (ordinary) rewriting/narrowing for multiset rewriting/narrowing. In particular, it provides a complete method for $E$-unification and $E$-unifiability consisting of multiple goals, where $E$ is represented by a complete rewrite system.

Meanwhile, Isabelle [26] is a generic proof assistant, i.e., a computer program that allows its users to express concepts in mathematics and computer science and to prove them using a logical calculus. While formalization of term rewriting has been done extensively in Isabelle (e.g., IsaFoR [1]), formalization of narrowing has not been done much yet in proof assistants including Isabelle. Our formalization of narrowing is built on IsaFoR (Isabelle/HOL *Formalization of Rewriting*). The relevant Isabelle theory files inside IsaFoR[1] under the directory thys/Narrowing are as follows:

| | |
|---|---|
| Narrowing.thy | Equational_Narrowing.thy |
| Multiset_Narrowing.thy | Equational_Narrowing_Unification.thy |
| Equational_Narrowing_Reachability.thy | Multiset_Narrowing_Unification.thy |
| Multiset_Narrowing_Reachability.thy | |

In the remainder of this paper, we provide hyperlinks (marked by ☑) to an HTML rendering for our formalized proofs in Isabelle/HOL.

## 2 Preliminaries

The definitions and results in this section can be found in [3, 6, 10, 18, 23]. We consider first-order terms over some signature $\mathcal{F}$ (consisting of function symbols $f, g, h, \ldots$ with fixed arities) and some infinite set of variables $x, y, z, \ldots \in \mathcal{V}$. A *position* within a term is a list of indices where $\varepsilon$ denotes the empty position, also called the *root position*. The set of positions of a term are defined as $\mathcal{P}os(x) = \{\varepsilon\}$ and $\mathcal{P}os(f(t_1, \ldots, t_n)) = \{\varepsilon\} \cup \{ip \mid 1 \leq i \leq n, p \in \mathcal{P}os(t_i)\}$. Given $p \in \mathcal{P}os(t)$, we write $t|_p$ for the subterm of $t$ at position $p$, i.e., $t|_\varepsilon = t$ and $f(t_1, \ldots, t_n)|_{ip} = t_i|_p$. The set of positions $\mathcal{P}os(t)$ of a term $t$ is partitioned into function positions $\mathcal{FP}os(t)$ and variable positions $\mathcal{VP}os(t)$, where $\mathcal{FP}os(t) = \{p \in \mathcal{P}os(t) \mid t|_p \notin \mathcal{V}\}$. For $p \in \mathcal{P}os(t)$, we denote by $t[s]_p$ the term that is obtained from $t$ by replacing the subterm at position $p$ by $s$.

The set of variables occurring in a term $t$ is denoted by $\mathcal{V}(t)$.

A *substitution* $\sigma$ is a mapping from $\mathcal{V}$ to $T(\mathcal{F}, \mathcal{V})$ such that $\{x \in \mathcal{V} \mid x\sigma \neq x\}$ is finite. This set is called the *domain* of $\sigma$, which is denoted by $\mathcal{D}\sigma$, while the set of variables introduced by $\sigma$ is denoted by $\mathcal{I}\sigma$. Substitutions are extended to mappings from $T(\mathcal{F}, \mathcal{V})$ to $T(\mathcal{F}, \mathcal{V})$ in the obvious way. In the remainder of this paper, we also write $s\sigma := \sigma(s)$ for substitutions $\sigma$ and terms $s$, and $(\sigma \circ \theta)(s) := s\theta\sigma$ for substitutions $\theta$, $\sigma$ and terms $s$.

The *restriction* $\sigma \upharpoonright_{\mathcal{V}}$ of a substitution $\sigma$ to $\mathcal{V}$ is defined as follows:

$$\sigma \upharpoonright_{\mathcal{V}} x = \left\{ \begin{array}{ll} x\sigma & \text{if } x \in \mathcal{V} \\ x & \text{otherwise} \end{array} \right\}$$

A *variable renaming* is a bijective substitution from $\mathcal{V}$ to $\mathcal{V}$. We write $\sigma = \tau[\mathcal{V}]$ if $\sigma \upharpoonright_{\mathcal{V}} = \tau \upharpoonright_{\mathcal{V}}$ and $\sigma \leqslant \tau[\mathcal{V}]$ if there is a substitution $\theta$ such that $\theta \circ \sigma = \tau[\mathcal{V}]$.

An *equation* is a pair $(s, t)$ of terms, written $s \approx t$. We denote by $\approx_E$ the least congruence on $T(\mathcal{F}, \mathcal{V})$ that is closed under substitutions and contains a set of equations $E$. If $s \approx_E t$ for two terms $s$ and $t$, then $s$ and $t$ are *E-equivalent*.

A substitution $\sigma$ is a *unifier* of two terms $s$ and $t$ if $s\sigma = t\sigma$. It is a *most general unifier* (or *mgu* for short) if for every unifier $\theta$ of $s$ and $t$, there exists a substitution $\lambda$ such that $\theta = \lambda \circ \sigma$. Two terms $s$ and $t$ are *E-unifiable* if there exists a substitution $\sigma$ such that $s\sigma \approx_E t\sigma$.

[1] http://cl-informatik.uibk.ac.at/isafor/#downloads
http://cl-informatik.uibk.ac.at/experiments/ITP2024/ceta_with_narrowing.zip for this paper.

A TRS $\mathcal{R}$ is a set of ordered pairs of terms, called *rules*, where a rule is usually written $\ell \to r$. For each rule $\ell \to r$, we assume that the set of variables occurring in $\ell$ includes the set of variables occurring in $r$, i.e., $\mathcal{V}(\ell) \supseteq \mathcal{V}(r)$. The induced rewrite relation is written as $\to_{\mathcal{R}}$ and can be defined either via positions or via contexts: $s \to_{\mathcal{R}} t$ if there is some $\ell \to r \in \mathcal{R}$ and substitution $\sigma$ such that $s|_p = \ell\sigma$ and $t = s[r\sigma]_p$ for some $p \in \mathcal{P}os(s)$ (or equivalently $s = C[\ell\sigma]$ and $t = C[r\sigma]$ for some context $C$).

A substitution $\sigma$ is *normalized* (w.r.t. a TRS $\mathcal{R}$) if $x\sigma$ is a normal form for every $x \in \mathcal{D}\sigma$. A substitution $\sigma$ is *normalizable* (w.r.t. a TRS $\mathcal{R}$) if $x\sigma$ has a normal form for every $x \in \mathcal{D}\sigma$.

A TRS $\mathcal{R}$ is *confluent* if ${}_{\mathcal{R}}^{*}\!\leftarrow \cdot \to_{\mathcal{R}}^{*} \subseteq \to_{\mathcal{R}}^{*} \cdot {}_{\mathcal{R}}^{*}\!\leftarrow$. A TRS $\mathcal{R}$ is *strongly normalizing* (SN) if there is no infinite reduction sequence $t_1 \to_{\mathcal{R}} t_2 \to_{\mathcal{R}} t_3 \to_{\mathcal{R}} \cdots$. A TRS $\mathcal{R}$ is *weakly normalizing* (WN) if every term has a normal form. A TRS $\mathcal{R}$ is *complete* if it is confluent and strongly normalizing. A TRS $\mathcal{R}$ is *semi-complete* if it is confluent and weakly normalizing.

A term $t$ is *strongly irreducible* (w.r.t. $\mathcal{R}$) if $t\sigma$ is a normal form (w.r.t. $\mathcal{R}$) for all normalized substitutions $\sigma$.

A *multiset* is a collection of elements in which elements can occur more than once. More formally, a multiset is a function from an element set $S$ to the natural numbers, giving the multiplicity of each element. This paper is only concerned with finite multisets.

## 3 Narrowing

▶ **Definition 1.** *A term $t$ is* narrowable *into a term $t'$ if there exist a position $p \in \mathcal{F}Pos(t)$, a variant*[2] *$\ell \to r$ of a rewrite rule in $\mathcal{R}$, and a substitution $\sigma$ such that*

- *$\sigma$ is a most general unifier of $t|_p$ and $\ell$,*
- *$t' = t[r]_p\sigma$.*

*Then, we write $t \rightsquigarrow_{[p,\ell\to r,\sigma]} t'$ or simply $t \rightsquigarrow_{\sigma,\mathcal{R}} t'$ (or more simply $\rightsquigarrow$). The relation $\rightsquigarrow$ is called* narrowing*. Also, we write $t \rightsquigarrow_{\sigma,\mathcal{R}}^{*} t'$ if there exists a narrowing derivation $t = t_1 \rightsquigarrow_{\sigma_1,\mathcal{R}} t_2 \rightsquigarrow_{\sigma_2,\mathcal{R}} \cdots \rightsquigarrow_{\sigma_{n-1},\mathcal{R}} t_n = t'$ such that $\sigma = \sigma_{n-1} \circ \cdots \circ \sigma_2 \circ \sigma_1$. If $n = 1$, then $\sigma = \varepsilon$.*

▶ **Lemma 2** (Lifting Lemma). *Let $\mathcal{R}$ be a TRS. Suppose that we have terms $s$ and $t$, a normalized substitution $\theta$ and a set of variables $V$ such that $\mathcal{V}(s) \cup \mathcal{D}\theta \subseteq V$ and $t = s\theta$. If $t \to_{\mathcal{R}}^{*} t'$, then there exist a term $s'$ and substitutions $\theta'$, $\sigma$ such that*

- *$s \rightsquigarrow_{\sigma,\mathcal{R}}^{*} s'$,*
- *$s'\theta' = t'$,*
- *$\theta' \circ \sigma = \theta[V]$,*
- *$\theta'$ is normalized.* ☑

Now, we may add a fresh binary function symbol $\approx^?$ and a fresh constant $\top$ to the set of function symbols and assume that $\mathcal{R}$ contains the rewrite rule $x \approx^? x \to \top$

▶ **Definition 3.** Equational terms *are the terms of the following form $s \approx^? t$, where $s$ and $t$ do not contain any occurrences of $\approx^?$ and $\top$.*

We may use the lifting lemma for equational terms because equational terms are simply some specific type of terms. We often denote equational terms using uppercase letters, such as $S$, $T$, $U$, etc, while ordinary terms are denoted by lowercase letters, such as $s$, $t$, $u$, etc. We assume that if $S$ is an equational term, then $S\sigma$ is also an equational term for any substitution $\sigma$. In other words, any substitution does not allow to introduce the special symbols $\approx^?$ and $\top$ in its range.

[2] See Definition 3.1 in [23] for details.

In our Isabelle/HOL formalization, the definition of narrowing (see Definition 1) is done using inductive_set in Isabelle. Here, $s$ narrows into $t$ iff $(s, t, \delta) \in$ narrowing_step.[3]

**inductive_set** *narrowing_step where*
"$(t = (replace\_at\ s\ p\ (snd\ rl)) \cdot \delta \wedge \omega \bullet rl \in \mathcal{R} \wedge (vars\_term\ s \cap vars\_rule\ rl = \{\}) \wedge p \in fun\_poss\ s \wedge mgu\ (s|_p)\ (fst\ rl) = Some\ \delta) \Rightarrow (s,\ t,\ \delta) \in narrowing\_step$"

Above, the renaming $\omega$ is applied to the rule $rl$, expressed by $\omega \bullet rl$, so that no variable shares between $s$ and $rl$. This corresponds to a variant of a rewrite rule $l \to r$ in Definition 1, where $l \to r$ is denoted here by $rl$. For renaming, we use the earlier formalization of *permutation for renaming* [17] in IsaFoR. Now, we formalize whether a narrowing derivation $s \rightsquigarrow_\sigma^* t$ holds or not, which cannot simply use the reflexive and transitive closure of the relation derived from narrowing_step because $\sigma$ should be combined and computed for the narrowing steps from $s$ to $t$.

**definition** *narrowing_derivation where*
"$narrowing\_derivation\ s\ s'\ \sigma \longleftrightarrow (\exists n. (\exists f\ \tau.\ f\ 0 = s \wedge f\ n = s' \wedge (\forall i < n.((f\ i), (f\ (Suc\ i)), (\tau\ i)) \in narrowing\_step) \wedge (if\ n = 0\ then\ \sigma = Var\ else\ \sigma = compose\ (map\ (\lambda i.(\tau\ i))[0..< n]))))$"

Above, $s \rightsquigarrow_\sigma^* t$ is true, denoted by $(s, t, \sigma) \in$ narrowing_derivation, if there are functions $f$ and $\tau$ forming the chains of narrowing steps and their corresponding narrowing substitutions, where the end points of the chain formed by $f$ are $s$ and $t$, respectively, and $\sigma$ is the composition of all substitutions of the chain formed by the function $\tau$. (Here, if the length of the chain is 0, then $\sigma$ is simply the identity substitution (i.e., $\sigma = Var$).)

Next, we need to formalize equational terms in Definition 3 in order to formalize the results in Sections 4 and 5. Formalization of equational terms needs some special treatment because of the new symbols $\approx^?$ and $\top$. Also, $s$ and $t$ in an equational term $s \approx^? t$ should not contain any occurrences of $\approx^?$ and $\top$. We introduce two function symbols using locale additional_function_symbols. Here, the binary function symbol $\doteq$ corresponds to $\approx^?$ in Definition 3. In the following, a term $t$ is a wf_equational_term if $t$ is either the constant $\top$ (i.e., Fun $\top$ [ ]) or it is an equational term of the form $u \approx^? v$, where the binary symbol $\approx^?$ and the constant $\top$ do not occur in any of $u$ and $v$.

**locale** *additional_function_symbols* = **fixes** *DOTEQ* :: "$'f$" ("$\doteq$") **and** *TOP* :: "$'f$" ("$\top$")
**begin**

**definition** *wf_equational_term where*
"$wf\_equational\_term\ t \longleftrightarrow ((t = Fun\ \top\ []) \vee (\exists u\, v.\ t = Fun\ \doteq\ [u :: ('f,'v)\ term,\ v :: ('f,'v)\ term] \wedge (\doteq, 2) \notin funas\_term\ u \wedge (\doteq, 2) \notin funas\_term\ v) \wedge (\top, 0) \notin funas\_term\ u \wedge (\top, 0) \notin funas\_term\ v))$"
...
**end**

Above, the *term* is represented by the datatype in IsaFoR:

**datatype** $(\alpha, \beta)\ term = Var\ \beta \mid Fun\ \alpha\ ((\alpha, \beta)\ term\ list)$

where $\alpha$ and $\beta$ are type parameters.

[3] Here, $\mathcal{R}$ is added as an argument of narrowing_step implicitly using a locale in Isabelle .

In the Narrowing directory (below the thys directory in IsaFoR), Narrowing.thy is concerned with narrowing without using equational terms, while Equational_Narrowing.thy is concerned with equational narrowing using equational terms. Note that $\mathcal{R}$ in the former file denotes the usual rewrite system with the condition that for each $\ell \to r \in \mathcal{R}$, $\mathcal{V}(\ell) \supseteq \mathcal{V}(r)$ and $\ell$ is not a variable, while $\mathcal{R}$ in the latter file additionally includes the rule $x \approx^? x \to \top$, written by a pair (Fun $\doteq$ [Var $x$, Var $x$], Fun $\top$ [ ]) in our formalization. We may also need to consider the original rewrite system from $\mathcal{R}$ excluding the rule $x \approx^? x \to \top$, where the binary symbol $\doteq$ and the constant $\top$ do not occur in the original rewrite system. We use the Isabelle's locale [5] to specify these in Equational_Narrowing.thy.

**locale** *equational_narrowing* = *narrowing* $\mathcal{R}$ + *additional_function_symbols DOTEQ TOP"* +
 **for** $\mathcal{R}$ :: "('f, 'v:: infinite) trs"
 ...
 **fixes** $\mathcal{R}'$ :: "('f, 'v:: infinite) trs"
  *and* $\mathcal{F}$ :: "'f sig"
  *and* $\mathcal{D}$ :: "'f sig"
  *and* $x$ :: "'v"
 **assumes** "$wf\_trs\ \mathcal{R}$"
  *and* "$\mathcal{R} = \mathcal{R}' \cup \{(Fun\ \doteq\ [Var\ x, Var\ x], Fun\ \top\ [\,])\}$"
  *and* "$funas\_trs\ \mathcal{R}' \subseteq \mathcal{F}$"
  *and* "$\mathcal{D} = \{(\doteq, 2), (\top, 0)\}$"
  *and* "$\mathcal{D} \cap \mathcal{F} = \{\}$"
  ...

Above, $\mathcal{R}'$ is the original rewrite system, while $\mathcal{R}$ is the rewrite system $\mathcal{R} = \mathcal{R}' \cup \{(Fun\ \doteq\ [Var\ x, Var\ x], Fun\ \top\ [\,])\}$. We assume that the function symbols of the original rewrite system $\mathcal{R}'$ is contained in $\mathcal{F}$, which is written as $funas\_trs\ \mathcal{R}' \subseteq \mathcal{F}$. Also, $\mathcal{D}$ is the set of fresh symbols $\{(\doteq, 2), (\top, 0)\}$, which should be disjoint from the original set of function symbols $\mathcal{F}$ (i.e., $\mathcal{D} \cap \mathcal{F} = \{\}$).

Note that the lifting lemma is a key lemma for narrowing, which states that a rewriting sequence can be "lifted" to a narrowing derivation. Our formalization includes four lifting lemmas, i.e., the lifting lemma for narrowing in Narrowing.thy, the lifting lemma for equational narrowing in Equational_Narrowing.thy, and the lifting lemma for multiset narrowing (see Lemma 22) and its slight variation in Multiset_Narrowing.thy, respectively. Here, we consider our formalization of the lifting lemma in equational narrowing, which is given as follows:

**lemma** *lifting_lemma:*
**fixes** $V$ :: "('v :: infinite) set" *and* $S$ :: "('f,'v)term" *and* $T$ :: "('f,'v)term"
 **assumes** "$normal\_subst\ \mathcal{R}\ \theta$"
  *and* "$wf\_equational\_term\ S$"
  *and* "$T = S \cdot \theta$"
  *and* "$vars\_term\ S \cup subst\_domain\ \theta \subseteq V$"
  *and* "$(T, T') \in rstep\ \mathcal{R})^*$"
  *and* "$finite\ V$"
 **shows** "$\exists \sigma\, \theta'\, S'. narrowing\_derivation\ S\ S'\ \sigma \wedge T' = S' \cdot \theta' \wedge wf\_equational\_term\ S' \wedge normal\_subst\ \mathcal{R}\ \theta' \wedge (\sigma \circ_s \theta') \mid_S V = \theta \mid_S V$"

There are slight differences between the formalization statement above and Lemma 2. Here, we use wf_equational_terms instead of ordinary terms. Each narrowing step transforms one wf_equational_term into another wf_equational_term. Also, we assume that $V$ is finite because we only consider finite wf_equational_terms and finite substitution domains for their associated substitutions. It is easier to rename the variables of the rules distinct from a finite $V$ instead of the infinite $V$. (For example, if $V$ is the universe of all variables of the given type, then we cannot rename the variables of the rules distinct from $V$.) Also, in the above formalization statement of the lifting lemma, $(T,\ T') \in (\mathsf{rstep}\ \mathcal{R})^*$ denotes the rewriting sequence from $T$ to $T'$, where the formalization of rstep is already available from IsaFoR [1] (see below):

**inductive_set** *rstep* :: "_ $\Rightarrow$ ('f,'v)term rel" **for** $\mathcal{R}$ :: "('f,'v) trs" **where**
"rstep : $\bigwedge C\ \sigma\ l\ r.\ (l,r) \in \mathcal{R} \Longrightarrow s = C\langle l \cdot \sigma\rangle \Longrightarrow t = C\langle r \cdot \sigma\rangle \Longrightarrow (s,t) \in rstep\ \mathcal{R}$"

Above, $(\sigma \circ_s \theta')\,|_S\, V$ (resp. $\theta\,|_S\, V$) denotes the restriction of a substitution $\sigma \circ_s \theta'$ (resp. $\theta$) to a set of variables $V$, where the restriction of a substitution subst_restrict is also available from IsaFoR (see below):

**definition** *subst_restrict* :: "('f,'v) subst $\Rightarrow$'v set $\Rightarrow$ ('f,'v) subst" (infix "|s" 67) **where**
"$\sigma\,|_S\, V = (\lambda x.\ if\ x \in V\ then\ \sigma(x)\ else\ Var\ x)$"

Similarly to the proof of the lifting lemma in [23], the proof of the formalization of the lifting lemma is proceeded by the induction on the length of the reduction sequence from $T$ to $T'$. To this end, from the assumption $(T, T') \in (\mathsf{rstep}\ R)^*$, we may obtain a chain and a number in such a way that

**obtain** $f\ n$ **where** "$f\ 0 = T$" and "$f\ n = T'$" and "$\forall i < n.(f\ i, f\ (Suc\ i)) \in rstep\ \mathcal{R}$"

Then we show the following statement using induction on $n$:

$\exists \sigma\, \theta'\, S'.$ narrowing_derivation_num $S\ S'\ \sigma\ n \wedge T' = S' \cdot \theta' \wedge wf\_equation\_term\ S' \wedge normal\_subst\ \mathcal{R}\ \theta' \wedge (\sigma \circ_s \theta')\,|_S\, V = \theta\,|_S\, V$.

Above, the narrowing_derivation_num is simply narrowing_derivation with the number of derivation steps being explicitly specified:

**definition** *narrowing_derivation_num* **where**
"$narrowing\_derivation\_num\ s\ s'\ \sigma\ n \longleftrightarrow (\exists f\ \tau.\ f\ 0 = s \wedge f\ n = s' \wedge (\forall i < n.((f\ i), (f\ (Suc\ i)), (\tau\ i)) \in narrowing\ step) \wedge (if\ n = 0\ then\ \sigma = Var\ else\ \sigma = compose\ (map\ (\lambda i.(\tau\ i))[0..< n])))$"

We leave it to our formalization for all the technical details of the proof of the lifting lemma.

## 4 $E$-unifiability

Narrowing is known to be a complete method of solving $E$-unification problems if $E$ can be represented by a semi-complete rewrite system [23]. The completeness of narrowing w.r.t. $E$-unification is derived from the lifting lemma using a semi-complete rewrite system representing $E$. The underlying idea of using narrowing is as follows (cf. [28]): A narrowing step from a term $s$ may represent many rewrite steps starting with instances of $s$. If $s\theta \to_{\mathcal{R}} t'$ is a rewrite step from $s\theta$ using a (fresh variant of) rule $\ell \to r$ at a non-variable position $p$ of $s$, then $s|_p$ and $\ell$ are unifiable. Then, using the most general unifier $\delta$ of $s|_p$ and $\ell$, we have a rewrite step $s\delta \to_{\mathcal{R}} t$ by applying the same rule $\ell \to r$ at the same position $p$ of $s$, where $t' = t\sigma$ for some substitution $\sigma$. Now, the narrowing step $s \rightsquigarrow_{\delta,\mathcal{R}} t$ may represent different rewriting steps for each unifier $\tau$ of $s|_p$ and $\ell$, where $s \rightsquigarrow_{\delta,\mathcal{R}} t$ implies $s\delta \to_{\mathcal{R}} t$. This can be extended to narrowing sequences in such a way that $s \rightsquigarrow^*_{\sigma,\mathcal{R}} t$ implies $s\delta \to^*_{\mathcal{R}} t$. The following lemma is used for both narrowing-based $E$-unification and the reachability analysis in the next section.

▶ **Lemma 4.**
**(i)** $s \rightsquigarrow^*_{\sigma,\mathcal{R}} t$ *implies* $s\sigma \to^*_{\mathcal{R}} t$.
**(ii)** $s \approx^? t \rightsquigarrow^*_{\sigma,\mathcal{R}} \top$ *implies* $s\sigma \approx^? t\sigma \to^*_{\mathcal{R}} \top$.

**Proof.** For the proof of (i), we proceed by induction on the length of the narrowing derivation $s \rightsquigarrow^*_{\sigma,\mathcal{R}} t$. The base case is immediate because we have $s = t$ and $\sigma = \varepsilon$ (i.e., the identity substitution). For the inductive case, we have some $u$ such that $s \rightsquigarrow^*_{\sigma_1,\mathcal{R}} u \rightsquigarrow_{\sigma_2,\mathcal{R}} t$, where the length of the narrowing derivation $s \rightsquigarrow^*_{\sigma_1,\mathcal{R}} u$ is one less than the length of the narrowing derivation in $s \rightsquigarrow^*_{\sigma,\mathcal{R}} t$ with $\sigma = \sigma_2 \circ \sigma_1$. The induction hypothesis yields $s\sigma_1 \to^*_{\mathcal{R}} u$. Also, by Definition 1, we see that $u\sigma_2 \to_{\mathcal{R}} t$ from $u \rightsquigarrow_{\sigma_2,\mathcal{R}} t$. Now, we have $(s\sigma_1)\sigma_2 \to^*_{\mathcal{R}} u\sigma_2 \to_{\mathcal{R}} t$, and thus the conclusion of (i) follows. We omit the proof of (ii), since it is almost identical to the proof of (i). ◀

Recall that we have the rule $x \approx^? x \to \top$ included in $\mathcal{R}$, where $\top$ is a fresh constant symbol. This means that if $s\theta \approx^? t\theta \to^*_{\mathcal{R}} \top$, then $\theta$ is an $\mathcal{R}$-unifier of $s$ and $t$ because $s\theta$ and $t\theta$ should be joined by $\mathcal{R}$. (Otherwise, no rewriting sequence by $\mathcal{R}$ from $s\theta \approx^? t\theta$ reaches $\top$.) Now, the following lemma directly follows from this observation using Lemma 4(ii).

▶ **Lemma 5** ( [23]). *Given a TRS $\mathcal{R}$, if $s \approx^? t \rightsquigarrow^*_{\sigma,\mathcal{R}} \top$ for some substitution $\sigma$, then $s$ and $t$ are $\mathcal{R}$-unifiable.*

In the above, given a set of equations $E$ represented by a rewrite system $\mathcal{R}$, $E$-unifiable is formalized in the following way, where *eq* is a pair of terms for representing an equation, and $\tau$ denotes an $E$-unifier.

**definition** "*E_unifiable eq* $\longleftrightarrow$ $(\exists\tau.((fst\ eq)\cdot\tau, (snd\ eq)\cdot\tau \in (rstep\ \mathcal{R})^{\leftrightarrow^*})$

▶ **Example 6.** Let $E = \{f(x,0) \approx g(x), g(b) \approx c\}$ and consider the unification problem $f(x,y) \approx^?_E c$. A rewrite system for $E$ is $\mathcal{R} = \{f(x,0) \to g(x), g(b) \to c, x \approx^? x \to \top\}$, where the rule $x \approx^? x \to \top$ is added using the fresh constant $\top$. We rename the rules in $\mathcal{R}$ whenever necessary, where variables with subscripts denote the renamed variables in this example. First, find the *mgu* of $f(x,y)$ and $f(x_1,0)$ in $f(x_1,0) \to g(x_1)$, which is $\sigma_1 = \{x \mapsto x_1, y \mapsto 0\}$. This yields the narrowing step $(f(x,y) \approx^? c) \rightsquigarrow_{\sigma_1} (g(x_1) \approx^? c)$. Next, find the *mgu* of $g(x_1)$ and $g(b)$, which is $\sigma_2 = \{x_1 \mapsto b\}$. This yields the narrowing step $(g(x_1) \approx^? c) \rightsquigarrow_{\sigma_2} (c \approx^? c)$. Then, find the *mgu* of $c \approx^? c$ and $x_2 \approx^? x_2$ in $x_2 \approx^? x_2 \to \top$, which is $\sigma_3 = \{x_2 \mapsto c\}$. This yields the narrowing step $c \approx^? c \rightsquigarrow_{\sigma_3} \top$.

We see that $\sigma := \sigma_3 \circ \sigma_2 \circ \sigma_1$ is an $\mathcal{R}$-unifier (or an $E$-unifier) of $f(x,y)$ and $c$, where $\sigma = \{x \mapsto b, y \mapsto 0, x_1 \mapsto b, x_2 \mapsto c\}$.

Now, given a semi-complete TRS $\mathcal{R}$, if $\theta$ is an $\mathcal{R}$-unifier of $s$ and $t$ (i.e., $s\theta \approx_{\mathcal{R}} t\theta$), then $s\theta \approx^? t\theta \to^*_{\mathcal{R}} \top$ because $\mathcal{R}$ is confluent. By the semi-completeness of $\mathcal{R}$, a normal substitution $\theta'$ of $\theta$ exists such that $s\theta' \approx^? t\theta' \to^*_{\mathcal{R}} \top$, and thus $\theta'$ is also an $\mathcal{R}$-unifier of $s$ and $t$. Applying the lifting lemma yields a narrowing sequence $s \approx^? t \rightsquigarrow^*_{\sigma,\mathcal{R}} \top$ such that $\sigma \leq \theta' [\mathcal{V}(s) \cup \mathcal{V}(t)]$. By Lemma 4(ii), we have $s\sigma \approx^? t\sigma \to^*_{\mathcal{R}} \top$, and thus $\sigma$ is also an $\mathcal{R}$-unifier of $s$ and $t$. Since we have $\theta \approx_{\mathcal{R}} \theta'$ and $\sigma \leq \theta' [\mathcal{V}(s) \cup \mathcal{V}(t)]$, we see that $\sigma \leq_{\mathcal{R}} \theta [\mathcal{V}(s) \cup \mathcal{V}(t)]$. This observation implies that for every $\mathcal{R}$-unifier of $s$ and $t$, a more general $R$-unifier can be found by narrowing. The completeness of narrowing for $E$-unification was originally proposed by Hullot [18], where $E$ is represented by a complete TRS. Later, it was shown that the semi-completeness of TRS suffices for the completeness of narrowing for $E$-unification [23].

▶ **Theorem 7** ( [23]). *Let $\mathcal{R}$ be a semi-complete TRS. If $s\theta \approx_{\mathcal{R}} t\theta$, then there is a narrowing derivation $s \approx^? t \rightsquigarrow^*_{\sigma,\mathcal{R}} \top$ such that $\sigma \leq_{\mathcal{R}} \theta [\mathcal{V}(s) \cup \mathcal{V}(t)]$.*

Unfortunately, the completeness of narrowing for $E$-unification alone does not imply $E$-unifiability by narrowing, which is also important in equational reasoning. In the remainder of this section, we show that given a semi-complete TRS $\mathcal{R}$, if narrowing terminates, then it provides a decision procedure for $E$-unifiability.

▶ **Lemma 8.** *Given a TRS $\mathcal{R}$, if there is no narrowing derivation $s \approx^? t \rightsquigarrow^*_{\sigma,\mathcal{R}} \top$ for any substitution $\sigma$, then there is no normal substitution $\theta$ satisfying $s\theta \approx^? t\theta \rightarrow^*_{\mathcal{R}} \top$.* ☑

**Proof.** Suppose to the contrary that there is a normal substitution $\theta$ satisfying $s\theta \approx^? t\theta \rightarrow^*_{\mathcal{R}} \top$. Let $V =: \mathcal{V}(S) \cup \mathcal{D}\theta$, where $S = s \approx^? t$. Then, by Lemma 2, there exists some substitution $\sigma$ such that $s \approx^? t \rightsquigarrow^*_{\sigma,\mathcal{R}} \top$, which is the required contradiction. ◀

▶ **Example 9.** Consider $\mathcal{R} = \{a \rightarrow b, f(a,b) \rightarrow c\}$ and $s = f(x,x)$ and $t = c$. Then $s \approx^? t$ is not narrowable, so there is no narrowing derivation $s \approx^? t \rightsquigarrow^*_{\sigma,\mathcal{R}} \top$ for any substitution $\sigma$. By Lemma 8, there is no normal substitution $\theta$ satisfying $s\theta \approx^? t\theta \rightarrow^*_{\mathcal{R}} \top$. However, there is a *non*-normal substitution $\delta := \{x \mapsto a\}$ satisfying $s\delta \approx^? t\delta \rightarrow^*_{\mathcal{R}} \top$, i.e., $f(a,a) \approx^? c \rightarrow_{\mathcal{R}} f(a,b) \approx^? c \rightarrow_{\mathcal{R}} c \approx^? c \rightarrow_{\mathcal{R}} \top$, where $s\delta = f(a,a)$ and $t\delta = c$.

The following lemma is immediate by observing that given a confluent TRS, $s \xleftrightarrow{*}_{\mathcal{R}} t$ implies that $s$ and $t$ are joinable.

▶ **Lemma 10.** *Given a confluent TRS $\mathcal{R}$, $s \xleftrightarrow{*}_{\mathcal{R}} t$ implies $s \approx^? t \rightarrow^*_{\mathcal{R}} \top$.* ☑

▶ **Lemma 11.** *Given a semi-complete TRS $\mathcal{R}$, if there is no narrowing derivation $s \approx^? t \rightsquigarrow^*_{\sigma,\mathcal{R}} \top$ for any substitution $\sigma$, then $s$ and $t$ have no R-unifier.* ☑

**Proof.** Assume that there is no narrowing derivation $s \approx^? t \rightsquigarrow^*_{\sigma,\mathcal{R}} \top$ for any substitution $\sigma$. Then, by Lemma 8, there is no normal substitution $\theta$ satisfying $s\theta \approx^? t\theta \rightarrow^*_{\mathcal{R}} \top$. Now, suppose, towards a contradiction, that $s$ and $t$ have an $\mathcal{R}$-unifier. Then, there is some substitution $\tau$ such that $s\tau \xleftrightarrow{*}_{\mathcal{R}} t\tau$. Since $\mathcal{R}$ is semi-complete, there is a normal substitution $\tau'$ of $\tau$ such that $s\tau' \xleftrightarrow{*}_{\mathcal{R}} t\tau'$. Now, we have $s\tau' \approx^? t\tau' \rightarrow^*_{\mathcal{R}} \top$ by Lemma 10, which is the required contradiction. ◀

From Lemmas 5 and 11, we have the following theorem of $E$-unifiability by narrowing.

▶ **Theorem 12.** *Given a semi-complete TRS $\mathcal{R}$, if all narrowing derivations starting from $s \approx^? t$ terminate (or simply $\rightsquigarrow$ terminates), then we can decide whether $s \approx^? t$ has an $\mathcal{R}$-unifier or not.* ☑

## 5 Reachability and Infeasibility

The *reachability problem* [14,27] is one of the fundamental problems in term rewriting systems, which originally has the following form: Given a TRS $\mathcal{R}$ and a source term $s$, does $s$ reach $t$ by a rewriting sequence, written $s \rightarrow^*_{\mathcal{R}} t$? This problem has the following generalization [21,27] for $s$ and $t$ containing variables: Is there a substitution $\sigma$ such that $s\sigma \rightarrow^*_{\mathcal{R}} t\sigma$? If there is no such substitution, then the problem is called *infeasible* [21,27]. In this paper, by the reachability problem, we mean the generalized reachability problem discussed above. In our Isabelle/HOL formalization, *reachable* and *infeasible* for a pair of terms are formalized as follows:

**definition** "*reachable eq* $\longleftrightarrow$ $(\exists\tau.((fst\ eq)\cdot\tau, (snd\ eq)\cdot\tau \in (rstep\ \mathcal{R})^*)$

**definition** "*infeasible eq* $\longleftrightarrow$ $(\neg(\exists\tau.((fst\ eq)\cdot\tau, (snd\ eq)\cdot\tau \in (rstep\ \mathcal{R})^*))$

The following lemma provides a sufficient condition of satisfying the reachability problem using narrowing, which requires neither the confluence nor the termination of the underlying TRS.

▶ **Lemma 13.**
**(i)** *If there is some substitution $\sigma$ such that $s \rightsquigarrow^*_{\sigma,\mathcal{R}} t\sigma$, then the reachability problem from $s$ to $t$ is satisfiable.* ☑
**(ii)** *If there is some substitution $\sigma$ such that $s \rightsquigarrow^*_{\sigma,\mathcal{R}} t'\sigma$ and $t'\sigma$ and $t\sigma$ are unifiable, then the reachability problem from $s$ to $t$ is satisfiable.* ☑

**Proof.** The proof of (i) is immediate using Lemma 4. For the proof of (ii), we have $s\sigma \rightarrow^*_{\mathcal{R}} t'\sigma$ from $s \rightsquigarrow^*_{\sigma,\mathcal{R}} t'\sigma$ using Lemma 4. Since $t'\sigma$ and $t\sigma$ are unifiable, there is some *mgu* $\delta$ such that $(t'\sigma)\delta = (t\sigma)\delta$. Then, we have $(s\sigma)\delta \rightarrow^*_{\mathcal{R}} (t'\sigma)\delta = (t\sigma)\delta$, and thus the reachability problem from $s$ to $t$ is satisfiable using substitution $\delta \circ \sigma$. ◀

▶ **Example 14.** Let $\mathcal{R} = \{f(x,x) \rightarrow g(x), a \rightarrow b\}$. For the reachability problem from $f(y,a)$ to $g(b)$, we have $f(y,a) \rightsquigarrow_{\sigma_1,\mathcal{R}} g(a)$, where $\sigma_1 = \{x \mapsto a, y \mapsto a\}$ is the *mgu* of $f(x,x)$ and $f(y,a)$. Then, we have $g(a) \rightsquigarrow_{\varepsilon,\mathcal{R}} g(b)$, so the reachability problem from $f(y,a)$ to $g(b)$ is satisfiable by Lemma 13(i) using substitution $\sigma_1 = \{x \mapsto a, y \mapsto a\}$.

Lemma 13(ii) provides a means to compute a solution of the reachability problem from $s$ to $t$ using a narrowing tree starting from $s$. Since a narrowing derivation along with its substitution are computed incrementally, a typical way of computing a solution of the reachability problem using a narrowing tree is to use the breadth-first search for each length of narrowing derivations and expand the narrowing tree (if it is possible) when a solution of the reachability problem cannot be found. (A more efficient way of solving reachability problems is considered in the next section.)

However, narrowing is known to be *weakly complete* [22] in reachability analysis in the sense that it may fail to find a solution of the reachability problem even if it exists. In particular, narrowing may fail to find a *non*-normalized solution of a reachability problem.

▶ **Example 15.** Given $\mathcal{R} = \{a \rightarrow b, a \rightarrow c, g(f(b), f(c)) \rightarrow a\}$, consider the reachability problem from $g(f(x), f(x))$ to $a$. The problem is satisfiable using substitution $\{x \mapsto a\}$ (i.e., $g(f(a), f(a)) \rightarrow_{\mathcal{R}} g(f(b), f(a)) \rightarrow_{\mathcal{R}} g(f(b), f(c)) \rightarrow_{\mathcal{R}} a$), but we may not apply Lemma 13(ii) because there is neither a narrowing step from $g(f(x), f(x))$ nor is it unifiable with $a$.

In what condition the reachability problem is shown to be either satisfiable or infeasible using narrowing? In the remainder of this section, if $\mathcal{R}$ is semi-complete and $t$ is a strongly-irreducible term (e.g. a constructor term), then we show that a narrowing derivation $s \approx^? t \rightsquigarrow^*_{\sigma,\mathcal{R}} \top$ for some substitution $\sigma$ implies the reachability from $s$ to $t$, while no narrowing derivation $s \approx^? t \rightsquigarrow^*_{\sigma,\mathcal{R}} \top$ for any substitution $\sigma$ implies the infeasibility of the reachability problem from $s$ to $t$, assuming that all narrowing derivations from $s \approx^? t$ terminates.

▶ **Lemma 16.** *Let $\mathcal{R}$ be a semi-complete TRS and $t$ be a strongly irreducible term. If there is some substitution $\sigma$ such that $s \approx^? t \rightsquigarrow^*_{\sigma,\mathcal{R}} \top$, then the reachability problem from $s$ to $t$ is satisfiable.* ☑

**Proof.** Suppose that there is some substitution $\sigma$ such that $s \approx^? t \rightsquigarrow^*_{\sigma,\mathcal{R}} \top$. Then, by Lemma 4(ii), we have $s\sigma \approx^? t\sigma \rightarrow^*_{\mathcal{R}} \top$. Since $\mathcal{R}$ is semi-complete, there is a normal substitution $\sigma'$ of $\sigma$ such that $s\sigma \approx^? t\sigma \rightarrow^*_{\mathcal{R}} s\sigma' \approx^? t\sigma'$ and $s\sigma' \approx^? t\sigma' \rightarrow^*_{\mathcal{R}} \top$. Also, $t\sigma'$ is a normal form of $\mathcal{R}$ because $t$ is strongly irreducible. Since $s\sigma' \approx^? t\sigma' \rightarrow^*_{\mathcal{R}} \top$ and $t\sigma'$ is normal form of $\mathcal{R}$, we may infer that $s\sigma' \rightarrow^*_{\mathcal{R}} t\sigma'$, and thus the conclusion follows. ◀

▶ **Lemma 17.** *Let $\mathcal{R}$ be a semi-complete TRS and $t$ be a strongly irreducible term. If there is no narrowing derivation $s \approx^? t \leadsto^*_{\sigma,\mathcal{R}} \top$ for any substitution $\sigma$, then the reachability problem from $s$ to $t$ is infeasible.*

**Proof.** Assume that there is no narrowing derivation $s \approx^? t \leadsto^*_{\sigma,\mathcal{R}} \top$ for any substitution $\sigma$. Then, by Lemma 8, there is no normal substitution $\theta$ satisfying $s\theta \approx^? t\theta \to^*_{\mathcal{R}} \top$. Now, suppose, towards a contradiction, that the reachability problem from $s$ to $t$ is satisfiable. Then, there is a substitution $\tau$ such that $\tau$ such that $s\tau \to^*_{\mathcal{R}} t\tau$. Since $\mathcal{R}$ is weakly normalizing, there is a normal substitution $\tau'$ of $\tau$ such that $s\tau' \ {}^*_{\mathcal{R}}\!\leftarrow s\tau \to^*_{\mathcal{R}} t\tau \to^*_{\mathcal{R}} t\tau'$. We see that $t\tau'$ is a normal form because $t$ is a strongly irreducible term and $\tau'$ is a normal substitution. Since $\mathcal{R}$ is confluent and $t\tau'$ is a normal form of $\mathcal{R}$, we have $s\tau' \to^*_{\mathcal{R}} t\tau'$, and thus $s\tau' \approx^? t\tau' \to^*_{\mathcal{R}} \top$, which is the required contradiction. ◀

From Lemmas 16 and 17, we have the following decidability result of the reachability problem using narrowing. (Note that Lemma 13 only provides a sufficient condition of satisfying the reachability problem using narrowing.)

▶ **Theorem 18.** *Let $\mathcal{R}$ be a semi-complete TRS and $t$ be a strongly irreducible term. If all narrowing derivations starting from $s \approx^? t$ terminate (or simply $\leadsto$ terminates), then we can decide whether the reachability problem from $s$ to $t$ is satisfiable or not (i.e., infeasible).*

## 6 Multiset Narrowing

In this section, we consider *multiset narrowing* for multiset reachability analysis and multiple goals in the reachability and $E$-unification problems. Our multiset narrowing[4] is adapted from *Narrowing Calculus* (NC) in [24], but it is also concerned with multisets of ordinary terms, equational terms, and pairs of terms. Note that a multiset is a generalization of a set, allowing elements in the multiset to occur more than once. It has an additional flexibility because identical elements (or states) in a multiset can reach different elements (or states).

Now, we consider multiset narrowing for multisets of terms (or equational terms). First, we consider *multiset rewriting* for multisets of terms (or equational terms).

▶ **Definition 19.** *Let $S$ and $T$ be multisets of terms. We write $S \to_{[\mathcal{R},M]} T$ if there exists a term $s \in S$ such that $s \to_{\mathcal{R}} t$ and $T = (S - \{s\}) \cup \{t\}$.*

▶ **Definition 20.** *Given a multiset of terms $S = \{t_1, \dots, t_n\}$, the* multiset reachability problem *is described as follows: is there a substitution $\sigma$ such that $S\sigma := \{t_1\sigma, \dots, t_n\sigma\}$ reaches the target multiset of terms $G = \{t'_1, \dots, t'_n\}$ using multiset rewriting, i.e., $S\sigma \to^*_{[\mathcal{R},M]} G$? If there is such a substitution $\sigma$, then we say that the multiset reachability problem from $S$ to $G$ is* satisfiable*. Otherwise, we say that it is* infeasible*.*

In the above definition, the source multiset $S$ and the target multiset $G$ are fixed for multiset reachability analysis which can be done using the following multiset narrowing.

▶ **Definition 21.** *A multiset of terms $S$ is* narrowable *into a multiset of terms $T$ if there exist a term $s \in S$ and a substitution $\sigma$ such that*

- *$s \leadsto_{\sigma,\mathcal{R}} t$,*
- *$T = ((S - \{s\})\sigma \cup \{t\}$.*

[4] Narrowing in a multiset environment is also considered in CHR [16], but it is considered in the context of logic programming, which does not consider multisets of ordinary terms.

*Then, we write* $S \rightsquigarrow_{\sigma,\mathcal{R},M} T$*. Also, we write* $S \rightsquigarrow^*_{\sigma,\mathcal{R},M} S'$ *if there exists a narrowing derivation* $S = S_1 \rightsquigarrow_{\sigma_1,\mathcal{R},M} S_2 \rightsquigarrow_{\sigma_2,\mathcal{R},M} \cdots \rightsquigarrow_{\sigma_{n-1},\mathcal{R},M} S_n = S'$ *such that* $\sigma = \sigma_{n-1} \circ \cdots \circ \sigma_2 \circ \sigma_1$*. If* $n = 1$*, then* $\sigma = \varepsilon$*.*

Intuitively speaking, $S \to_{[\mathcal{R},M]} T$ if $T$ is obtained by replacing one element (term) in $S$ using a rewriting step in $\mathcal{R}$, while $S \rightsquigarrow_{\sigma,\mathcal{R},M} T$ if $T$ is obtained by replacing one element (term) in $S$ using a narrowing step in Definition 1 and then applying the narrowing substitution to the remaining multiset $S - \{s\}$.

In our Isabelle/HOL formalization, we use finite multisets for multiset narrowing, where a finite multiset is a finite collection of elements, denoted by $\{\#x_1, \ldots, x_n\#\}$ in isabelle. Duplication is allowed and orders are irrelevant in multisets, i.e., $\{\#s, t, t, s\#\} = \{\#t, t, s, s\#\}$. Also, + denotes multiset sum and - denotes multiset difference. Now, the multiset reduction in Definition 19 can be used for multisets of both ordinary and equational terms. Then, $S \to_{[\mathcal{R},M]} T$ iff $(S, T) \in$ *multiset_reduction_step* (see below).

**inductive_set** *multiset_reduction_step* *where*
"$s \in\# S \wedge T = (S - \{\#s\#\} + \{\#t\#\}) \wedge (s,t) \in rstep\ \mathcal{R} \Rightarrow (S,\ T) \in multiset\_reduction\_step$"

The corresponding multiset narrowing in Definition 21 is formalized as follows, where $S \rightsquigarrow_{\sigma,\mathcal{R},M} T$ iff $(S, T, \sigma) \in$ *multiset_narrowing_step*.

**inductive_set** *multiset_narrowing_step* *where*
"$(s,t) \in\# S \wedge T = (subst\_term\_multiset\ \sigma\ (S - \{\#s\#\}) + \{\#t\#\}) \wedge (s,t,\sigma) \in narrowing\_step \Rightarrow (S,T,\sigma) \in multiset\_narrowing\_step$"

The lifting lemma for multisets of terms can be easily adapted from Lemma 2.[5]

▶ **Lemma 22.** *Let* $\mathcal{R}$ *be a TRS. Suppose we have two multisets of terms* $S$ *and* $T$*, a normalized substitution* $\theta$ *and a set of variables* $V$ *such that* $\mathcal{V}(S) \cup \mathcal{D}\theta \subseteq V$ *and* $T = S\theta$*. If* $T \to^*_{[\mathcal{R},M]} T'$*, then there exist a multiset of terms* $S'$ *and substitutions* $\theta'$*,* $\sigma$ *such that*

- $S \rightsquigarrow^*_{\sigma,\mathcal{R},M} S'$*,*
- $S'\theta' = T'$*,*
- $\theta' \circ \sigma = \theta[V]$*,*
- $\theta'$ *is normalized.* ☑

The following lemma can be proved by induction on the length of the multiset narrowing derivation $S \rightsquigarrow^*_{\sigma,\mathcal{R},M} T$ using the observation that $S'\sigma' \to_{[\mathcal{R},M]} T'$ whenever $S' \rightsquigarrow_{\sigma',\mathcal{R},M} T'$ (cf. Lemma 4).

▶ **Lemma 23.** *Let* $\mathcal{R}$ *be a TRS and* $S$ *be a multiset of terms (or equational terms). Then,* $S \rightsquigarrow^*_{\sigma,\mathcal{R},M} T$ *implies* $S\sigma \to^*_{[\mathcal{R},M]} T$*.* ☑

▶ **Lemma 24.** *If there are some substitutions* $\sigma$ *and* $\eta$ *such that* $S \rightsquigarrow^*_{\sigma,\mathcal{R},M} S'$ *and* $S'\eta = G$*, then the multiset reachability problem from* $S$ *to* $G$ *is satisfiable.* ☑

**Proof.** Suppose that there are some substitution $\sigma$ and $\eta$ such that $S \rightsquigarrow^*_{\sigma,\mathcal{R},M} S'$ and $S'\eta = G$. Then, by Lemma 23, we have $S\sigma \to^*_{[\mathcal{R},M]} S'$. By Definition 19 and easy induction on the length of multiset rewriting steps, we may infer that $\to^*_{[\mathcal{R},M]}$ is closed under substitutions. Now, we have $S\sigma\eta \to^*_{[\mathcal{R},M]} S'\eta = G$, and thus the conclusion follows. ◀

[5] The lifting lemma for multisets of equational terms is also a slight variation of the lifting lemma for multisets of terms, where Definition 3 needs to be checked.

▶ **Example 25.** We consider the multiset reachability problem introduced in Section 1. Let $S = \{f(x, y), f(x, y)\}$, the (renamed) rewrite system $\mathcal{R} = \{f(a, b) \to d, f(a, z_1) \to g(z_1), f(z_2, a) \to d, g(a) \to c\}$, and the target multiset $G = \{c, d\}$. Multiset narrowing starts with $S = \{f(x, y), f(x, y)\}$ and narrow into $S_1 = \{g(z_1), f(a, z_1)\}$ using the rule $f(a, z_1) \to g(z_1)$ with substitution $\sigma_1 = \{x \mapsto a, y \mapsto z_1\}$. Then, it narrows into $S_2 = \{c, f(a, a)\}$ using the rule $g(a) \to c$ with substitution $\sigma_2 = \{z_1 \mapsto a\}$. Finally, it narrows into $S_3 = \{c, d\}$ using the rule $f(z_2, a) \to d$, with substitution $\sigma_3 = \{z_2 \mapsto a\}$. Then by Lemma 24, the above multiset reachability problem is satisfied with substitution $\sigma = \sigma_3 \circ \sigma_2 \circ \sigma_1 = \{x \mapsto a, y \mapsto a, z_1 \mapsto a, z_2 \mapsto a\}$.

▶ **Lemma 26.** *If there is no multiset narrowing derivation $S \rightsquigarrow^*_{\sigma,\mathcal{R},M} S'$ for any substitution $\sigma$ and $\eta$ with $S'\eta = G$, then there is no normal substitution $\theta$ satisfying the multiset reachability problem from $S$ to $G$.* ☑

The above lemma describes the *weak completeness* of multiset narrowing w.r.t. multiset reachability analysis. For example, the multiset reachability problem from $\{g(f(x), f(x))\}$ to $\{a\}$ using $\mathcal{R}$ in Example 15 is satisfiable using substitution $\{x \mapsto a\}$, but there is no multiset narrowing step from $\{g(f(x), f(x))\}$ nor is there some substitution $\eta$ such that $\{g(f(x), f(x))\eta\} = \{a\}$.

▶ **Lemma 27.**
**(i)** *If $\mathcal{R}$ is strongly normalizing, then $\to_{[\mathcal{R},M]}$ is strongly normalizing.* ☑
**(ii)** *If $\mathcal{R}$ is complete, then $\to_{[\mathcal{R},M]}$ is confluent.* ☑

▶ **Lemma 28.** *Given a complete TRS $\mathcal{R}$, if there is no multiset narrowing derivation $S \rightsquigarrow^*_{\sigma,\mathcal{R},M} S'$ for any substitution $\sigma$ and $\eta$ with $S'\eta = G$ and $G$ is in normal form w.r.t. $\to_{[\mathcal{R},M]}$, then there is no substitution $\theta$ satisfying the multiset reachability problem from $S$ to $G$.* ☑

**Proof.** Assume that there is no multiset narrowing derivation $S \rightsquigarrow^*_{\sigma,\mathcal{R},M} S'$ for any substitution $\sigma$ and $\eta$ with $S'\eta = G$. Then, by Lemma 26, there is no normal substitution $\theta$ satisfying the multiset reachability problem from $S$ to $G$. Now, suppose to the contrary that there is some substitution $\theta$ satisfying the multiset reachability problem from $S$ to $G$, i.e., $S\theta \to^*_{[\mathcal{R},M]} G$. By Lemma 27, $\to_{[\mathcal{R},M]}$ is strongly normalizing and confluent. Now, we have $S\theta \to^*_{[\mathcal{R},M]} S\theta'$, where $\theta'$ is the normal substitution of $\theta$. (This can be shown using a straightforward induction on the size of $S\theta$.) Since $\to_{[\mathcal{R},M]}$ is strongly normalizing and confluent and $G$ is in normal form w.r.t. $\to_{[\mathcal{R},M]}$, we have $S\theta \to^*_{[\mathcal{R},M]} S\theta' \to^*_{[\mathcal{R},M]} G$, contradicting that there is no normal substitution satisfying the multiset reachability problem from $S$ to $G$. ◀

From Lemmas 24 and 28, we have the following decidability result of multiset reachability analysis using multiset narrowing.

▶ **Theorem 29.** *Let $\mathcal{R}$ be a complete TRS $\mathcal{R}$, $S$ and $G$ be multisets of terms, and $G$ be in normal form w.r.t. $\to_{[\mathcal{R},M]}$. If all multiset narrowing derivations starting from $S$ terminate, then we can decide whether the multiset reachability problem from $S$ to $G$ is satisfiable or not (i.e., infeasible).* ☑

Meanwhile, multiset narrowing can also be used for $E$-unification problems consisting of multiple goals. In the following, by a slight abuse of notation, we denote by $\top$ a finite multiset consisting only of $\top's$ or simply $\top$ in Definition 3. The next theorem provides the completeness of multiset narrowing for $E$-unification problems consisting of multiple goals.

▶ **Theorem 30.** *Let $\mathcal{R}$ be a complete TRS and $S = \{s_1 \approx^? t_1, \ldots, s_n \approx^? t_n\}$ be a multiset of equational terms. If there is some $\mathcal{R}$-unifier $\theta$ satisfying $s_k\theta \approx_\mathcal{R} t_k\theta$ for all $1 \le k \le n$, then there is some multiset narrowing derivation $S \rightsquigarrow^*_{\sigma,\mathcal{R},M} \top$ such that $\sigma \le_\mathcal{R} \theta\,[\mathcal{V}(S)]$.* ☑

Next, we consider $E$-unifiability consisting of multiple goals using multiset narrowing. The following lemma provides a sufficient condition of satisfying an $E$-unifiability problem (consisting of multiple goals) using multiset narrowing.

▶ **Lemma 31.** *Let $\mathcal{R}$ be a TRS and $S = \{s_1 \approx^? t_1, \ldots, s_n \approx^? t_n\}$ be a multiset of equational terms. If $S \rightsquigarrow^*_{\sigma,\mathcal{R},M} \top$ for some substitution $\sigma$, then $s_k$ and $t_k$ for all $1 \le k \le n$ are $\mathcal{R}$-unifiable.* ☑

**Proof.** Suppose $S \rightsquigarrow^*_{\sigma,\mathcal{R},M} \top$. Then, we have $S\sigma \to^*_{[R,M]} \top$ by Lemma 23. Also, $S\sigma \to^+_{[R,M]} \top$ because it needs at least one step including the step using the rule $x \approx x \to \top$. Now, observe that for any nonempty $S' \subset S\sigma$, we have $S' \to^+_{[R,M]} \top$. Therefore, for any $1 \le k \le n$, we have $\{s_k\sigma \approx t_k\sigma\} \to^+_{[R,M]} \top$. Now, we proceed by induction on the number of $\to^+_{[R,M]}$-steps in $\{s_k\sigma \approx t_k\sigma\} \to^+_{[R,M]} \top$ and show that $s_k\sigma \xleftrightarrow{*}_\mathcal{R} t_k\sigma$.

The base case is obvious, i.e., $s_k\sigma = t_k\sigma$. For the inductive case, consider $s'$ and $t'$, where $\{s_k\sigma \approx t_k\sigma\} \to_{[R,M]} \{s' \approx t'\}$ and $\{s' \approx t'\} \to^+_{[R,M]} \top$. The induction hypothesis yields $s' \xleftrightarrow{*}_\mathcal{R} t'$. Since $\{s_k\sigma \approx t_k\sigma\} \to_{[R,M]} \{s' \approx t'\}$, we see that either $s_k\sigma \to_\mathcal{R} s'$ with $t_k\sigma = t'$ or $t_k\sigma \to_\mathcal{R} t'$ with $s_k\sigma = s'$ by Definition 19, and thus the conclusion follows from $s_k\sigma \xrightarrow{*}_\mathcal{R} s' \xleftrightarrow{*}_\mathcal{R} t' \xleftarrow{*}_\mathcal{R} t_k\sigma$. ◀

▶ **Lemma 32.** *Let $\mathcal{R}$ be a complete TRS and $S = \{s_1 \approx^? t_1, \ldots, s_n \approx^? t_n\}$ be a multiset of equational terms. If there is no multiset narrowing derivation $S \rightsquigarrow^*_{\sigma,\mathcal{R},M} \top$ for any substitution $\sigma$, then there is no $\mathcal{R}$-unifier $\sigma$ satisfying $s_k\sigma \approx_\mathcal{R} t_k\sigma$ for all $1 \le k \le n$, where $\mathcal{R}$ is viewed as a set of equations.* ☑

From Lemmas 31 and 32, we have the following theorem of $E$-unifiability (consisting of multiple goals) by multiset narrowing.

▶ **Theorem 33.** *Let $\mathcal{R}$ be a complete TRS and $S = \{s_1 \approx^? t_1, \ldots, s_n \approx^? t_n\}$ be a multiset of equational terms. If all multiset narrowing derivation starting from $S$ terminate, then we can decide whether there is an $\mathcal{R}$-unifier $\sigma$ satisfying $s_k\sigma \approx_\mathcal{R} t_k\sigma$ for all $1 \le k \le n$.* ☑

Next, we adapt the narrowing discussed in [22] for (ordinary) reachability analysis using multisets of pairs of terms. Given a rewrite system $\mathcal{R}$ and pairs of terms $(s_1, t_1), \ldots, (s_n, t_n)$, the purpose of reachability analysis is to determine whether there is a substitution $\sigma$ such that $s_1\sigma \to^*_\mathcal{R} t_1\sigma \wedge \cdots \wedge s_n\sigma \to^*_\mathcal{R} t_n\sigma$. Here, the reachability problem is represented by the multiset $\{(s_k, t_k) \mid 1 \le k \le n\}$.

▶ **Definition 34.** *Let $S$ and $T$ be multisets of the pairs of terms. We write $S \to_{[\mathcal{R},M_p]} T$ if there exists a pair of terms $(s,t) \in S$ such that $s \to_\mathcal{R} u$ and $T = (S - \{(s,t)\}) \cup \{(u,t)\}$.*

▶ **Definition 35.** *A multiset of pairs of terms $S$ is* narrowable *into a multiset of pairs of terms $T$ if there exists a pair of terms $(s,t) \in S$ and a substitution $\sigma$ such that*

- *$s \rightsquigarrow_{\sigma,\mathcal{R}} u$, and*
- *$T = (S - \{(s,t)\})\sigma \cup \{(u, t\sigma)\}$.*

*Then, we write $S \rightsquigarrow_{\sigma,\mathcal{R},M_p} T$. Also, we write $S \rightsquigarrow^*_{\sigma,\mathcal{R},M_p} S'$ if there exists a narrowing derivation $S = S_1 \rightsquigarrow_{\sigma_1,\mathcal{R},M_p} S_2 \rightsquigarrow_{\sigma_2,\mathcal{R},M_p} \cdots \rightsquigarrow_{\sigma_{n-1},\mathcal{R},M_p} S_n = S'$ such that $\sigma = \sigma_{n-1} \circ \cdots \circ \sigma_2 \circ \sigma_1$. If $n = 1$, then $\sigma = \varepsilon$.*

Intuitively, $S \to_{[\mathcal{R},M_p]} T$ if $T$ is obtained by replacing one pair of elements $(s,t)$ in $S$ with $(u,t)$ using $s \to_{\mathcal{R}}$ u. Only the first element in a pair can be rewritten by $\mathcal{R}$, while the second element serves as a target and is intact for $\to_{[\mathcal{R},M_p]}$-steps. Meanwhile, $S \rightsquigarrow_{\sigma,\mathcal{R},M_p} T$ if $T$ is obtained by replacing one pair of elements $(s,t)$ in $S$ with $(u,t\sigma)$ from $s \rightsquigarrow_{\sigma,\mathcal{R}} u$ and then applying the narrowing substitution to the remaining multiset $S - \{(s,t)\}$.

In our Isabelle/HOL formalization, for the multiset reduction in Definition 34, we use the following inductive set in Isabelle such a way that $S \to_{[\mathcal{R},M_p]} T$ iff $(S,T) \in$ *multiset_pair_reduction_step*. (Here, $\mathcal{R}$ is implicitly included as a parameter of multiset_pair_reduction_step in the locale.)

**inductive_set** *multiset_pair_reduction_step* *where*
"$(s,t) \in\# S \wedge T = (S - \{\#(s,t)\#\} + \{\#(u,t)\#\}) \wedge (s,u) \in rstep\ \mathcal{R} \Rightarrow (S,T) \in$ *multiset_pair_reduction_step*"

Similarly, for the multiset narrowing in Definition 35, we use the following inductive set in such a way that $S \rightsquigarrow_{\sigma,\mathcal{R},M_p} T$ iff $(S,T,\sigma) \in$ *multiset_pair_narrowing_step*.

**inductive_set** *multiset_pair_narrowing_step* *where*
"$(s,t) \in\# S \wedge T = (subst\_pairs\_multiset\ \sigma\ (S - \{\#(s,t)\#\}) + \{\#(u, t \cdot \sigma)\#\}) \wedge (s,u,\sigma) \in$ *narrowing_step* $\Rightarrow (S,T,\sigma) \in$ *multiset_pair_narrowing_step*"

▶ **Definition 36.**

**(i)** *We say that a multiset of pairs of terms* $\{(s_k,t_k) \mid 1 \le k \le n\}$ *is* trivially unifiable *if* $s_k = t_k$ *for all* $1 \le k \le n$.

**(ii)** *We say that a multiset of pairs of terms* $\{(s_k,t_k) \mid 1 \le k \le n\}$ *is* syntactically unifiable with *a substitution* $\theta$ *if* $s_k\theta = t_k\theta$ *for all* $1 \le k \le n$.

**(iii)** *We say that a substitution* $\tau$ *is a* solution *of the reachability problem represented by* $S = \{(s_1,t_1),\ldots,(s_n,t_n)\}$ *if* $s_1\tau \to^*_{\mathcal{R}} t_1\tau \wedge \cdots \wedge s_n\tau \to^*_{\mathcal{R}} t_n\tau$.

▶ **Lemma 37.** *Let* $\mathcal{R}$ *be a TRS and* $S = \{(s_1,t_1),\ldots,(s_n,t_n)\}$ *be a multiset of pairs of terms. If* $S \to^*_{[\mathcal{R},M_p]} S'$ *and* $S'$ *is trivially unifiable, then* $s_1 \to^*_{\mathcal{R}} t_1 \wedge \cdots \wedge s_n \to^*_{\mathcal{R}} t_n$. ☑

**Proof.** We proceed by induction on the number of $\to^*_{[R,M_p]}$-steps in $S \to^*_{[\mathcal{R},M_p]} S'$. The base case is trivial, i.e., $S = S'$. For the inductive case, consider $S \to_{[\mathcal{R},M_p]} U$ and $U \to^*_{[\mathcal{R},M_p]} S'$. From $S \to_{[\mathcal{R},M_p]} U$, we have some $(s,t) \in S$, $s \to_{\mathcal{R}} u$, and $U = (S - \{(s,t)\}) \cup \{(u,t)\}$. By the induction hypothesis, for all pairs $(v,w)$ in $U$, we have $v \to^*_{\mathcal{R}} w$. This means that $u \to^*_{\mathcal{R}} t$ and for all pairs $(v',w') \in (S - \{(s,t)\})$, we have $v' \to^*_{\mathcal{R}} w'$. Therefore, it remains to show that $s \to^*_{\mathcal{R}} t$, which is obvious from $s \to_{\mathcal{R}} u$ and $u \to^*_{\mathcal{R}} t$. ◀

▶ **Proposition 38.** *Let* $\mathcal{R}$ *be a TRS and* $S = \{(s_1,t_1),\ldots,(s_n,t_n)\}$ *be a multiset of pairs of terms. If* $S \rightsquigarrow^*_{\sigma,\mathcal{R},M_p} S'$ *and* $S'$ *is syntactically unifiable with* $\theta$*, then* $\theta \circ \sigma$ *is a solution of the reachability problem represented by* $S = \{(s_1,t_1),\ldots,(s_n,t_n)\}$. ☑

**Proof.** Suppose $S \rightsquigarrow^*_{\sigma,\mathcal{R},M_p} S'$. Then, we have $S\sigma \to^*_{[\mathcal{R},M_p]} S'$ by adapting the proof of Lemma 4. Also, the relation $\to^*_{[R,M_p]}$ is closed under substitutions, which can be shown using induction on the number of $\to_{[\mathcal{R},M_p]}$-steps. Then, we have $(S\sigma)\theta \to^*_{[\mathcal{R},M_p]} S'\theta$, where $S'\theta$ is trivially unifiable. Thus, the conclusion follows by Lemma 37. ◀

The above proposition provides a sufficient condition of satisfying a reachability problem consisting of multiple goals using multiset narrowing on multisets of pairs of terms. However, it alone does not provide the decidability of a reachability problem consisting of multiple goals.

Next, we consider multiset narrowing on multisets of equational terms again (instead of multisets of pairs of terms) for ordinary reachability problems. Similarly to Definition 36(iii), we say that a substitution $\sigma$ is a *solution* of the reachability problem represented by a multiset $S = \{s_1 \approx^? t_1, \ldots, s_n \approx^? t_n\}$ if $s_1\sigma \to^*_{\mathcal{R}} t_1\sigma \wedge \cdots \wedge s_n\sigma \to^*_{\mathcal{R}} t_n\sigma$. If $\sigma$ is a *solution* of the reachability problem represented by $S$, then we say that the reachability problem represented by $S$ is *satisfiable*. Otherwise, if there is no solution of the reachability problem represented by $S$, then we say that the reachability problem represented by $S$ is *infeasible*.

▶ **Lemma 39.** *Let $\mathcal{R}$ be a TRS and $S = \{s_1 \approx^? t_1, \ldots, s_n \approx^? t_n\}$ be a multiset of equational terms such that $s_1 \to^*_{\mathcal{R}} t_1 \wedge \cdots \wedge s_n \to^*_{\mathcal{R}} t_n$ and each $t_k$, $1 \le k \le n$, is a normal form of $\mathcal{R}$. Then, $S \to^*_{[\mathcal{R},M]} \top$.* ☑

▶ **Lemma 40.** *Let $S = \{s_1 \approx^? t_1, \ldots, s_n \approx^? t_n\}$ be a multiset of equational terms. If there is no multiset narrowing derivation $S \rightsquigarrow^*_{\sigma,\mathcal{R},M} \top$ for any substitution $\sigma$, then there is no normal substitution $\theta$ satisfying $S\theta \to^*_{[R,M]} \top$.* ☑

▶ **Lemma 41.** *Let $\mathcal{R}$ be a semi-complete TRS and $S = \{s_1 \approx^? t_1, \ldots, s_n \approx^? t_n\}$ be a multiset of equational terms, where each $t_k$, $1 \le k \le n$, is a strongly irreducible term. If there is no multiset narrowing derivation $S \rightsquigarrow^*_{\sigma,\mathcal{R},M} \top$ for any substitution $\sigma$, then the reachability problem represented by $S$ is infeasible.* ☑

**Proof.** Assume that there is no multiset narrowing derivation $S \rightsquigarrow^*_{\sigma,\mathcal{R},M} \top$ for any substitution $\sigma$. Then, by Lemma 40, there is no normal substitution $\theta$ satisfying $S\theta \to^*_{[\mathcal{R},M]} \top$. Now, suppose, towards a contradiction, that the reachability problem represented by $S$ is satisfiable. Then, there is a substitution $\tau$ such that $s_1\tau \to^*_{\mathcal{R}} t_1\tau \wedge \cdots \wedge s_n\tau \to^*_{\mathcal{R}} t_n\tau$. Since $\mathcal{R}$ is weakly normalizing, there is a normal substitution $\tau'$ of $\tau$ such that $s_k\tau' \; {}^*_{\mathcal{R}}\!\!\leftarrow s_k\tau \to^*_{\mathcal{R}} t_k\tau \to^*_{\mathcal{R}} t_k\tau'$ for all $1 \le k \le n$. We see that each $t_k\tau'$, $1 \le k \le n$, is in normal form (w.r.t. $\mathcal{R}$) because $t_k$ is a strongly irreducible term and $\tau'$ is a normal substitution. Since $\mathcal{R}$ is confluent and each $t_k\tau'$, $1 \le k \le n$, is in normal form (w.r.t. $\mathcal{R}$), we have $s_k\tau' \to^*_{\mathcal{R}} t_k\tau'$ for all $1 \le k \le n$. Now, we have $S\tau' \to^*_{[\mathcal{R},M]} \top$ by Lemma 39, which is the required contradiction. ◀

▶ **Lemma 42.** *Let $\mathcal{R}$ be a semi-complete TRS and $S = \{s_1 \approx^? t_1, \ldots, s_n \approx^? t_n\}$ be a multiset of equational terms, where each $t_k$, $1 \le k \le n$, is a strongly irreducible term. If $S \rightsquigarrow^*_{\sigma,\mathcal{R},M} \top$ for some substitution $\sigma$, then the reachability problem represented by $S$ is satisfiable.* ☑

Now, we have the following decidability result of a reachability problem (consisting of multiple goals) using multiset narrowing on multisets of equational terms by Lemmas 41 and 42.

▶ **Theorem 43.** *Let $\mathcal{R}$ be a semi-complete TRS and $S = \{s_1 \approx^? t_1, \ldots, s_n \approx^? t_n\}$ be a multiset of equational terms, where each $t_k$, $1 \le k \le n$, is a strongly irreducible term. If all multiset narrowing derivations starting from $S$ terminate, then we can decide whether the reachability problem represented by $S$ is satisfiable or not (i.e., infeasible).* ☑

## 7 Related Work and Discussion

In this paper, we have focused on an Isabelle/HOL formalization of narrowing and multiset narrowing. There are other important narrowing techniques, such as *basic* [23], conditional [6], constrained [8], *nominal* [2], and *folding variant* [12] narrowing, which have not been discussed in this paper. For $E$-unification and reachability analysis, there are also existing narrowing-based computational tools (not using an Isabelle/HOL proof assistant); in particular, see the *Maude* system [11] using folding variant narrowing.

Meanwhile, multiset narrowing presented in this paper provides a natural method for multiset reachability analysis. Note that there are some limitations on simulating multiset rewriting (resp. multiset narrowing) using ordinary rewriting (resp. ordinary narrowing). Consider, for example, $S = \{s_1, s_2, s_3, s_4\}$ and $T = \{t_1, s_2, s_3, s_4\}$, where all $s_i$, $1 \leq i \leq 4$, are distinct, $s_1 \to_{\mathcal{R}} t_1$, and thus $S \to_{[\mathcal{R},M]} T$. If we simulate the multiset rewriting $S \to_{[\mathcal{R},M]} T$ using ordinary rewriting with a new function symbol $\bar{f}$, we have to consider the following cases: (1) $\bar{f}(s_1, s_2, s_3, s_4) \to_{\mathcal{R}} \bar{f}(t_1, s_2, s_3, s_4)$, (2) $\bar{f}(s_2, s_1, s_3, s_4) \to_{\mathcal{R}} \bar{f}(s_2, t_1, s_3, s_4)$, ..., (24) $\bar{f}(s_4, s_3, s_2, s_1) \to_{\mathcal{R}} \bar{f}(s_4, s_3, s_2, t_1)$. Here, $S \to_{[\mathcal{R},M]} T$ is a compact representation of the above 24 cases. Similarly, let $S = \{s_1, s_2, s_3, s_4\}$ as above and $U = \{u_1, u_2, u_3, u_4\}$, where all $u_i$, $1 \leq i \leq 4$, are distinct. Now, determining whether $S \rightsquigarrow^*_{\sigma,\mathcal{R},M} U$ using some $\sigma$ exists is a compact representation of determining whether one of the following 24 cases of ordinary narrowing using some $\sigma_i$ exists with a new function symbol $\bar{g}$: (1) $\bar{g}(s_1, s_2, s_3, s_4) \rightsquigarrow^*_{\sigma_1,\mathcal{R}} \bar{g}(u_1, u_2, u_3, u_4)$, (2) $\bar{g}(s_1, s_2, s_3, s_4) \rightsquigarrow^*_{\sigma_2,\mathcal{R}} \bar{g}(u_2, u_1, u_3, u_4)$, ..., (24) $\bar{g}(s_1, s_2, s_3, s_4) \rightsquigarrow^*_{\sigma_{24},\mathcal{R}} \bar{g}(u_4, u_3, u_2, u_1)$. Here, without using multiset narrowing, one may have to create 24 (ordinary) narrowing trees in the worst case (with possibly many duplicated narrowing steps) for the corresponding multiset reachability problem.

When considering multiset reachability problems by determining whether a substitution $\sigma$ exists such that $S\sigma \to^*_{[\mathcal{R},M]} U$, multiset narrowing provides a simple and compact sufficient condition of satisfying the multiset reachability problem, i.e., $S \rightsquigarrow^*_{\sigma,\mathcal{R},M} U$ using some $\sigma$.

## 8 Conclusion

Although narrowing plays an important role in equational unification and reachability analysis, formalization of narrowing and its related results on equational unification and reachability analysis has not been much done in the proof assistants. We have presented a new Isabelle/HOL formalization of narrowing and multiset narrowing for $E$-unifiability and (multiset) reachability analysis. The results discussed in this paper are built on IsaFoR (Isabelle/HOL Formalization of Rewriting) [1].

Given a semi-complete rewrite system $\mathcal{R}$ representing $E$ and two terms $s$ and $t$, we show a formalized correctness proof that if all narrowing derivations starting from $s \approx^? t$ terminate (or simply $\rightsquigarrow$ terminates), then we can decide whether $s$ and $t$ are $E$-unifiable.

We have also presented multiset narrowing and its formalization for multiset reachability analysis. Our multiset narrowing is generic in the sense that it encapsulates (the ordinary) rewriting and narrowing for multiset rewriting and multiset narrowing. It is also applicable to $E$-unifiability/$E$-unification and reachability problems consisting of multiple goals. In particular, given a complete rewrite system $\mathcal{R}$, it provides a complete method for $\mathcal{R}$-unifiability problems consisting of multiple goals, where $\mathcal{R}$ is viewed as a set of equations. Furthermore, if $\mathcal{R}$ is semi-complete and the right-hand sides of multiple goals in a reachability problem are strongly irreducible terms, then it provides a decision procedure for the reachability problem if it terminates. (Recall that if $\mathcal{R}$ is complete, then $\mathcal{R}$ is semi-complete, but not vice versa.)

Finally, much work still remains ahead. In particular, developing and formalizing parallel multiset rewriting/narrowing is a potential future research direction. It is also interesting to see whether multiset narrowing encapsulating other rewriting and narrowing strategies (such as *basic narrowing* [23]) can improve the multiset narrowing discussed in this paper.

## References

1 An Isabelle/HOL Formalization of Rewriting for Certified Tool Assertions. Computational Logic group at the University of Innsbruck, `http://cl-informatik.uibk.ac.at/isafor/`.

**2** Mauricio Ayala-Rincón, Maribel Fernández, and Daniele Nantes-Sobrinho. Nominal narrowing. In Delia Kesner and Brigitte Pientka, editors, *1st International Conference on Formal Structures for Computation and Deduction, FSCD 2016, June 22-26, 2016, Porto, Portugal*, volume 52 of *LIPIcs*, pages 11:1–11:17. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2016. doi:10.4230/LIPICS.FSCD.2016.11.

**3** Franz Baader and Tobias Nipkow. *Term Rewriting and All That*. Cambridge University Press, Cambridge, UK, 1998.

**4** Franz Baader and Wayne Snyder. Unification Theory. In *Handbook of Automated Reasoning*, Volume I, chapter 8, pages 445–532. Elsevier, Amsterdam, 2001.

**5** Clemens Ballarin. Tutorial to Locales and Locale Interpretation. URL: http://isabelle.in.tum.de/doc/locales.pdf.

**6** Alexander Bockmayr. *Contributions to the Theory of Logic-Functional Programming*. PhD thesis, Fakultät für Informatik, Universität Karlsruhe, 1990.

**7** Andrew Cholewa, Santiago Escobar, and José Meseguer. Constrained narrowing for conditional equational theories modulo axioms. *Sci. Comput. Program.*, 112:24–57, 2015. doi:10.1016/J.SCICO.2015.06.001.

**8** Hubert Comon and Claude Kirchner. Constraint Solving on Terms. In Gerhard Goos, Juris Hartmanis, Jan van Leeuwen, Hubert Comon, Claude Marché, and Ralf Treinen, editors, *Constraints in Computational Logics: Theory and Applications International Summer School, CCL '99 Gif-sur-Yvette, France, September 5–8, 1999 Revised Lectures*, pages 47–103. Springer Berlin Heidelberg, Berlin, Heidelberg, 2001. doi:10.1007/3-540-45406-3_2.

**9** Hubert Comon-Lundh and Stéphanie Delaune. The Finite Variant Property: How to Get Rid of Some Algebraic Properties. In Jürgen Giesl, editor, *Term Rewriting and Applications, 16th International Conference, RTA 2005, Nara, Japan, April 19-21, 2005, Proceedings*, volume 3467 of *Lecture Notes in Computer Science*, pages 294–307. Springer, 2005. doi:10.1007/978-3-540-32033-3_22.

**10** Nachum Dershowitz and David A. Plaisted. Rewriting. In John Alan Robinson and Andrei Voronkov, editors, *Handbook of Automated Reasoning (in 2 volumes)*, pages 535–610. Elsevier and MIT Press, 2001.

**11** Francisco Durán, Steven Eker, Santiago Escobar, Narciso Martí-Oliet, José Meseguer, Rubén Rubio, and Carolyn L. Talcott. Equational Unification and Matching, and Symbolic Reachability Analysis in Maude 3.2 (System Description). In Jasmin Blanchette, Laura Kovács, and Dirk Pattinson, editors, *Automated Reasoning - 11th International Joint Conference, IJCAR 2022, Haifa, Israel, August 8-10, 2022, Proceedings*, volume 13385 of *Lecture Notes in Computer Science*, pages 529–540. Springer, 2022. doi:10.1007/978-3-031-10769-6_31.

**12** Santiago Escobar, Ralf Sasse, and José Meseguer. Folding Variant Narrowing and Optimal Variant Termination. In Peter Csaba Ölveczky, editor, *Rewriting Logic and Its Applications - 8th International Workshop, WRLA 2010, Held as a Satellite Event of ETAPS 2010, Paphos, Cyprus, March 20-21, 2010, Revised Selected Papers*, volume 6381 of *Lecture Notes in Computer Science*, pages 52–68. Springer, 2010. doi:10.1007/978-3-642-16310-4_5.

**13** M. Fay. First-Order Unification in Equational Theories. In *Fourth International Workshopon Automated Deduction, Austin, Texas, Proceedings*, pages 161–167, 1979.

**14** Guillaume Feuillade, Thomas Genet, and Valérie Viet Triem Tong. Reachability Analysis over Term Rewriting Systems. *J. Autom. Reason.*, 33(3-4):341–383, 2004. doi:10.1007/S10817-004-6246-0.

**15** Jean H Gallier and Wayne Snyder. Complete sets of transformations for general E-unification. *Theoretical Computer Science*, 67(2-3):203–260, 1989. doi:10.1016/0304-3975(89)90004-2.

**16** Michael Hanus. CHR(Curry): Interpretation and Compilation of Constraint Handling Rules in Curry. In Enrico Pontelli and Tran Cao Son, editors, *Practical Aspects of Declarative Languages - 17th International Symposium, PADL 2015, Portland, OR, USA, June 18-19, 2015. Proceedings*, volume 9131 of *Lecture Notes in Computer Science*, pages 74–89. Springer, 2015. doi:10.1007/978-3-319-19686-2_6.

**17** Nao Hirokawa, Aart Middeldorp, and Christian Sternagel. A New and Formalized Proof of Abstract Completion. In Gerwin Klein and Ruben Gamboa, editors, *Interactive Theorem Proving - 5th International Conference, ITP 2014, Held as Part of the Vienna Summer of Logic, VSL 2014, Vienna, Austria, July 14-17, 2014. Proceedings*, volume 8558 of *Lecture Notes in Computer Science*, pages 292–307. Springer, 2014. doi:10.1007/978-3-319-08970-6_19.

**18** Jean-Marie Hullot. Canonical forms and unification. In Wolfgang Bibel and Robert A. Kowalski, editors, *5th Conference on Automated Deduction, Les Arcs, France, July 8-11, 1980, Proceedings*, volume 87 of *Lecture Notes in Computer Science*, pages 318–334. Springer, 1980. doi:10.1007/3-540-10009-1_25.

**19** Hélène Kirchner. On the Use of Constraints in Automated Deduction. In Andreas Podelski, editor, *Constraint Programming: Basics and Trends, Châtillon Spring School, Châtillon-sur-Seine, France, May 16 - 20, 1994, Selected Papers*, volume 910 of *Lecture Notes in Computer Science*, pages 128–146. Springer, 1994. doi:10.1007/3-540-59155-9_8.

**20** John W. Lloyd. *Foundations of Logic Programming, 3rd Edition.* Springer, 2012.

**21** Salvador Lucas and Raúl Gutiérrez. Use of logical models for proving infeasibility in term rewriting. *Inf. Process. Lett.*, 136:90–95, 2018. doi:10.1016/J.IPL.2018.04.002.

**22** José Meseguer and Prasanna Thati. Symbolic reachability analysis using narrowing and its application to verification of cryptographic protocols. *High. Order Symb. Comput.*, 20(1-2):123–160, 2007. doi:10.1007/S10990-007-9000-6.

**23** Aart Middeldorp and Erik Hamoen. Completeness results for basic narrowing. *Appl. Algebra Eng. Commun. Comput.*, 5:213–253, 1994. doi:10.1007/BF01190830.

**24** Aart Middeldorp, Satoshi Okui, and Tetsuo Ida. Lazy Narrowing: Strong Completeness and Eager Variable Elimination. *Theor. Comput. Sci.*, 167(1&2):95–130, 1996. doi:10.1016/0304-3975(96)00071-0.

**25** Robert Nieuwenhuis. Decidability and Complexity Analysis by Basic Paramodulation. *Inf. Comput.*, 147(1):1–21, 1998. doi:10.1006/INCO.1998.2730.

**26** Tobias Nipkow, Lawrence C. Paulson, and Markus Wenzel. *Isabelle/HOL - A Proof Assistant for Higher-Order Logic*, volume 2283 of *Lecture Notes in Computer Science*. Springer, 2002.

**27** Christian Sternagel and Akihisa Yamada. Reachability Analysis for Termination and Confluence of Rewriting. In Tomás Vojnar and Lijun Zhang, editors, *Tools and Algorithms for the Construction and Analysis of Systems - 25th International Conference, TACAS 2019, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2019, Prague, Czech Republic, April 6-11, 2019, Proceedings, Part I*, volume 11427 of *Lecture Notes in Computer Science*, pages 262–278. Springer, 2019. doi:10.1007/978-3-030-17462-0_15.

**28** Prasanna Thati and José Meseguer. Complete Symbolic Reachability Analysis Using Back-and-Forth Narrowing. In José Luiz Fiadeiro, Neil Harman, Markus Roggenbach, and Jan J. M. M. Rutten, editors, *Algebra and Coalgebra in Computer Science: First International Conference, CALCO 2005, Swansea, UK, September 3-6, 2005, Proceedings*, volume 3629 of *Lecture Notes in Computer Science*, pages 379–394. Springer, 2005. doi:10.1007/11548133_24.

**29** Emanuele Viola. E-unifiability via Narrowing. In Antonio Restivo, Simona Ronchi Della Rocca, and Luca Roversi, editors, *Theoretical Computer Science, 7th Italian Conference, ICTCS 2001, Torino, Italy, October 4-6, 2001, Proceedings*, volume 2202 of *Lecture Notes in Computer Science*, pages 426–438. Springer, 2001. doi:10.1007/3-540-45446-2_27.

**30** Akihiro Yamamoto. Completeness of extended unification based on basic narrowing. In Koichi Furukawa, Hozumi Tanaka, and Tetsunosuke Fujisaki, editors, *Logic Programming '88*, pages 1–10, Berlin, Heidelberg, 1989. Springer Berlin Heidelberg.