REST: Integrating Term Rewriting with Program Verification

Anonymous author
Anonymous affiliation

Abstract

We introduce REST, a novel term rewriting technique for theorem proving that uses online termination checking and can be integrated with existing program verifiers. REST enables flexible but terminating term rewriting for theorem proving by: (1) exploiting newly-introduced term orderings that are more permissive than standard rewrite simplification orderings; (2) dynamically and iteratively selecting orderings based on the path of rewrites taken so far; and (3) integrating external oracles that allow steps that cannot be justified with rewrite rules. Our REST approach is designed around an easily implementable core algorithm, parameterizable by choices of term orderings and their implementations; in this way our approach can be easily integrated into existing tools. We implemented REST as a Haskell library and incorporated it into Liquid Haskell's evaluation strategy, extending Liquid Haskell with rewriting rules. We evaluated our REST implementation by comparing it against both existing rewriting techniques and E-matching and by showing that it can be used to supplant manual lemma application in many existing Liquid Haskell proofs.

2012 ACM Subject Classification Theory of computation → Program verification

Keywords and phrases term rewriting, program verification, theorem proving

Digital Object Identifier 10.4230/LIPIcs...

1 Introduction

For all disjoint sets s_0 and s_1, the identity $(s_0 \cup s_1) \cap s_0 = s_0$ can be proven in many ways. Informally accepting this property is easy, but a machine-checked formal proof may require the instantiation of multiple set theoretic axioms. Analogously, further proofs relying on this identity may themselves need to apply it as a previously-proven lemma. For example, proving functional correctness of any program that relies on a set data structure typically requires the instantiation of set-related lemmas. Manual instantiation of such universally quantified equalities is tedious, and the burden becomes substantial for more complex proofs: a proof author needs to identify exactly which equalities to instantiate and with which arguments; in the context of program verification, a wide variety of such lemmas are typically available. Given this need, most program verifiers provide some automated technique or heuristics for instantiating universally quantified equalities.

For the wide range of practical program verifiers that are built upon SMT solvers (e.g., [33, 23, 49, 37, 45, 43]), quantified equalities can naturally be expressed in the SMT solver's logic. However, relying solely on such solvers' E-matching techniques [19] for quantifier instantiation (as the majority of these verifiers do) can lead to both non-termination and incompletenesses that may be unpredictable [32] and challenging to diagnose [7]. The theory of how to prove that an E-matching-based encoding of equality reasoning guarantees termination and completeness is difficult and relatively unexplored [21].

A classical alternative approach to automating equality reasoning is *term rewriting* [26], which can be used to encode lemma properties as (directed) rewrite rules, matching terms against the existing set of rules to identify potential rewrites; the termination of these systems is a well-studied problem [16]. Although SMT solvers often perform rewriting as an internal simplification step, verifiers built on top typically cannot access or customize these rules, e.g., to add previously-proved lemmas as rewrite rules. By contrast, many

Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

mainstream proof assistants (e.g., Coq [11], Isabelle/HOL [38], Lean [5]) provide automated, customizable term rewriting tactics. However, the rewriting functionalities of mainstream proof assistants either do not ensure the termination of rewriting (potentially resulting in divergence, for example Isabelle) or enforce termination checks that are overly restrictive in general, potentially rejecting necessary rewrite steps (for example, Lean).

In this paper, we present *REST (REwriting and Selecting Termination orderings)*: a novel technique that equips program verifiers with automatic lemma application facilities via term rewriting, enabling equational reasoning with complementary strengths to E-matching-based techniques. While term rewriting in general does not guarantee termination, our technique weaves together three key technical ingredients to automatically generate and explore guaranteed-terminating restrictions of a given rewriting system while typically retaining the rewrites needed in practice: (1) REST compares terms using well-quasi-orderings derived from (strict) simplification orderings; thereby facilitating common and important rules such as commutativity and associativity properties. (2) REST simultaneously considers an entire family of term orderings; selecting the appropriate term ordering to justify rewrite steps *during term rewriting itself*. (3) REST allows integration of an *external oracle* that generates additional steps outside of the term rewriting system. This allows the incorporation of reasoning steps awkward or impossible to justify via rewriting rules, all without compromising the termination and relative completeness guarantees of our overall technique.

Contributions and Overview We make the following contributions:

1. We design and present a new approach (REST) for applying term rewriting rules and simultaneously selecting appropriate term orderings to permit as many rewriting steps as possible while guaranteeing termination (Sec. 3).
2. We introduce ordering constraint algebras, an abstraction for reasoning effectively about multiple (and possibly infinitely many) term orderings simultaneously (Sec. 4).
3. We introduce and formalize recursive path quasi-orderings (RPQOs) derived from the well-known recursive path ordering [15] (Sec. 4.1.2). RPQOs are more permissive than classical RPOs, and so let us prove more properties.
4. We formalize and prove key results for our technique: soundness, relative completeness, and termination (Sec. 5).
5. We implement REST as a stand-alone library, and integrate the REST library into Liquid Haskell to facilitate automatic lemma instantiation (Sec. 6).
6. We evaluate REST by comparing it to other term rewriting tactics and E-matching-based axiomatization, and show that it can substantially simplify equational reasoning proofs (Sec. 7).

We discuss related work in Sec. 8; we begin (Sec. 2) by identifying five key problems that all need solving for a reliable and automatic integration of term rewriting into a program verification tool.

2 Five Challenges for Automating Term Rewriting

In this section, we describe *five key challenges* that naturally arise when term rewriting is used for program verification and outline how REST is designed to address them. To illustrate the challenges, we use simple verification goals that involve uninterpreted functions and the set operators ($\emptyset$, $\cup$, $\cap$) that satisfy the standard properties of Figure 1. The variables x, y, z

Name	*Formula*
idem-union	$x \cup x = x$
idem-inter	$x \cap x = x$
empty-union	$x \cup \emptyset = x$
empty-inter	$x \cap \emptyset = \emptyset$
commut-union	$x \cup y = y \cup x$
symm-inter	$x \cap y = y \cap x$
distrib-union	$(x \cup y) \cap z = (x \cap z) \cup (y \cap z)$
distrib-inter	$(x \cap y) \cup z = (x \cup z) \cap (y \cup z)$
assoc-union	$x \cup (y \cup z) = (x \cup y) \cup z$

Figure 1 Set identities used for examples in this section. Variables x, y, z are implicitly quantified. We write the binary functions $\cup, \cap$ infix; along with (nullary) $\emptyset$ these are fixed function symbols.

are implicitly quantified[1] in these rules. In formalizations of set theory, such properties may be assumed as (quantified) axioms, or proven as lemmas and then used in future proofs.

Term rewriting systems (defined formally in Sec. 5.1) are a standard approach for formally expressing and applying equational reasoning (rewriting terms via known identities). A term rewriting system consists of a finite set of *rewrite rules*, each consisting of a pair of a *source term* and a *target term*, representing that terms matching a rule's source can be replaced by corresponding terms matching its target. For example, the rewrite rule $x \cup \emptyset \to x$ can replace set unions of some set x and the empty set with the corresponding set x. Rewrite rules are applied to a term t by identifying some subterm of t which is equal to a rule's source under some substitution of the source's free variables (here, x, but not constants such as $\emptyset$); the subterm is then replaced with the correspondingly substituted target term. This rewriting step *induces an equality* between the original and new terms. For instance, the example rewrite rule above can be used to rewrite a term $f(s_0 \cup \emptyset)$ into $f(s_0)$, inducing an equality between the two.

Rewrite rules classically come with two restrictions: the free variables of the target must all occur in the source and the source must not be a single variable. This precludes rewrite rules which invent terms, such as $\emptyset \to x \cap \emptyset$, and those that trivially lead to infinite derivations. Under these restrictions, the first four identities induce rewrite rules from left-to-right (which we denote by e.g., *idem-inter*$\rightarrow$), while the remaining induce rewrite rules in both directions (e.g., *assoc-union*$\rightarrow$ vs. *assoc-union*$\leftarrow$).

Next, we present a simple proof obligation taken from [34] in the style of equational reasoning (*calculational proofs*) supported in the Dafny program verifier [33].

▶ **Example 1.** We aim to prove, for two sets s_0 and s_1 and some unary function f on sets, that, if the sets are disjoint (that is, $s_1 \cap s_0 = \emptyset$), then $f((s_0 \cup s_1) \cap s_0) = f(s_0)$.

$$\begin{array}{llll} \textit{Equational Proof:} & f((s_0 \cup s_1) \cap s_0) = & f((s_0 \cap s_0) \cup (s_1 \cap s_0)) & \textit{(distrib-union}\rightarrow\textit{)} \\ & = & f(s_0 \cup (s_1 \cap s_0)) & \textit{(idem-inter}\rightarrow\textit{)} \\ & = & f(s_0 \cup \emptyset) & \textit{(disjointness ass.}\rightarrow\textit{)} \\ & = & f(s_0) & \textit{(empty-union}\rightarrow\textit{)} \end{array}$$

(Possible Term Ordering, as explained shortly: RPO instance with $\cap > \cup$)

[1] over sets; we omit explicit types in such formulas, whose type-checking is standard.

This manual proof closely follows the user annotations employed in the corresponding Dafny proof [34]; the application of the function f serves only to illustrate equational reasoning on subterms. Every step of the proof could be explained by term rewriting, hinting at the possibility of an *automated* proof in which term rewriting is used to solve such proof obligations. In particular, taking the term rewriting system naturally induced by the set identities of Figure 1 *along with* the assumed equality expressing disjointness of s_0 and s_1 results in a term rewriting system in which the four proof steps are all valid rewriting steps.

In the remainder of the section, we consider what it would take to make term rewriting effective for reliably automating such verification tasks. Perhaps unsurprisingly, there are multiple problems with the simplistic approach outlined so far. The first and most serious is that term rewriting systems in general *do not guarantee termination*; a proof search may continue indefinitely by repeatedly applying rewrite rules. For example, the rules *distrib-union* and *distrib-inter* can lead to an infinite derivation $(s_0 \cup s_1) \cap s_2 \to (s_0 \cap s_2) \cup (s_1 \cap s_2) \to (s_0 \cup (s_1 \cap s_2)) \cap (s_2 \cup (s_1 \cap s_2)) \to \dots$

Challenge 1: Unrestricted term rewriting systems do not guarantee termination.

To ensure termination (as proved in Theorem 22) `REST` follows the classical approach of restricting a term-rewriting system to a variant in which sequences of term rewrites (*rewrite paths*) are allowed only if each consecutive pair of terms is *ordered* according to some term ordering which rules out infinite paths.

For example, *Recursive path orderings* (RPOs) [15] define well-founded orders $>_{\mathcal{T}}$ on terms $\mathcal{T}$ based on an underlying well-founded strict partial order $>$ on *function symbols*. Intuitively, such orderings use $>$ to order terms with different top-level function symbols, combined with the properties of a *simplification order* [14] (e.g., compatibility with the subterm relation). Different choices of the underlying $>$ parameter yield different RPO instances that order different pairs of terms; in particular, potentially allowing or disallowing certain rewrite paths.

In Example 1, the RPO based on the partial order $\cap > \cup$ and $\cap > \emptyset$ permits all the rewriting steps, that is, the left-hand-side of each equation is greater than the right-hand-side.

Sadly, this ordering will not permit the rewriting steps required by our next example.

▶ **Example 2.** We aim to prove, for two sets s_0 and s_1 and some unary function f on sets, that, if s_1 is a subset of s_0 (that is, $s_0 \cup s_1 = s_0$), then $f((s_0 \cap s_1) \cup s_0) = f(s_0)$.

$$\begin{array}{lcll} \textit{Equational Proof:}\ f((s_0 \cap s_1) \cup s_0) & = & f((s_0 \cup s_0) \cap (s_1 \cup s_0)) & \textit{(distrib-inter}\rightarrow\textit{)} \\ & = & f(s_0 \cap (s_1 \cup s_0)) & \textit{(idem-union}\rightarrow\textit{)} \\ & = & f(s_0 \cap (s_0 \cup s_1)) & \textit{(commut-union}\rightarrow\textit{)} \\ & = & f(s_0 \cap s_0) & \textit{(subset ass.}\rightarrow\textit{)} \\ & = & f(s_0) & \textit{(idem-inter}\rightarrow\textit{)} \end{array}$$

(Possible Term Ordering: RPQO instance, explained shortly, with $\cap > \cup$)

An RPO based on the function symbol ordering $\cap > \cup$ (as required by Example 1) will not permit the first step of this proof (since the RPO ordering first compares the top level function symbols). Instead, this step requires an RPO based on the ordering $\cup > \cap$. To accept *both* this proof step *and* the Example 1 we need different restrictions of the rewrite rules for different proofs; in particular, different rewrite paths may be ordered according to RPOs that are based on different function orderings.

To generalize this problem we will call RPOs a term ordering *family* that is *parametric* with respect to the underlying function ordering. Thus, a concrete RPO term ordering (called an *instance* of the family) is obtained after the parametric function ordering is instantiated. With this terminology, the next challenge can be stated as follows:

Challenge 2: Different proofs require different term orderings within a family.

Note that enumerating all term orderings in a term ordering family is typically impractical (this set is often very large and may be infinite). To address this challenge, REST uses a novel algebraic structure (Sec. 4.2) to allow for an abstract representation of sets of term orderings with which one can efficiently check whether any instance of a chosen term ordering family can orient the necessary rewrite steps to complete a proof.

Going back to Example 2, the RPO instance with $\cup > \cap$ will permit all the steps, apart from the commutativity axiom expressed by *(commut-union→)*. To permit this step we need an ordering for which $t_1 \cup t_2 >_{\mathcal{T}} t_2 \cup t_1$. But for RPO instances, as well as for many other term orderings, the terms $t_1 \cup t_2$ and $t_2 \cup t_1$ are equivalent and thus cannot be oriented; associativity axioms are also similarly challenging. Since many proofs require such properties, it is important in practice for rewriting to support them.

Challenge 3: Strict orderings restrict commutativity and associativity steps.

To address this challenge REST relaxes the strictness constraint by requiring the chosen term ordering family to consist (only) of *thin well-quasi-orderings* (Sec. 4). Intuitively, such orderings permit rewriting to terms which are *equal* according to the ordering, but such equivalence classes of terms must be guaranteed to be finite. In Sec. 4 we show how to lift well-known families of term orderings to analogous and more-permissive families of thin well-quasi-orders. In particular, we show how to lift RPOs to a particularly powerful family of term orderings that we call *recursive path quasi-orderings (RPQOs)*, whose instances allow us to accept Example 2.

Despite the permissiveness of RPQOs, there remain some rewrite derivations that will be rejected by all term orderings in the RPQO family. For example, consider the following proof that set union is monotonic with respect to the subset relation:

▶ **Example 3.** We aim to prove, for sets s_0, s_1, and s_2, that, if s_1 is a subset of s_0 (that is, $s_0 \cup s_1 = s_0$), then $(s_2 \cup s_1) \cup (s_2 \cup s_0) = s_2 \cup s_0$.

Equational Proof:
$$\begin{aligned}
(s_2 \cup s_1) \cup (s_2 \cup s_0) &= s_2 \cup (s_1 \cup (s_2 \cup s_0)) && \textit{(assoc-union}\leftarrow\textit{)}\\
&= s_2 \cup ((s_1 \cup s_2) \cup s_0) && \textit{(assoc-union}\rightarrow\textit{)}\\
&= s_2 \cup ((s_2 \cup s_1) \cup s_0) && \textit{(commut-union}\rightarrow\textit{)}\\
&= s_2 \cup (s_2 \cup (s_1 \cup s_0)) && \textit{(assoc-union}\leftarrow\textit{)}\\
&= s_2 \cup (s_2 \cup (s_0 \cup s_1)) && \textit{(commut-union}\rightarrow\textit{)}\\
&= s_2 \cup (s_2 \cup s_0) && \textit{(subset ass.}\rightarrow\textit{)}\\
&= (s_2 \cup s_2) \cup s_0 && \textit{(assoc-union}\rightarrow\textit{)}\\
&= s_2 \cup s_0 && \textit{(idem-union}\rightarrow\textit{)}
\end{aligned}$$

(Possible Term Ordering: any KBQO instance)

The above rewrite rule steps cannot be oriented by any RPQO, but are trivially oriented by a quasi-ordering that is based on the syntactic size of the term, e.g., a quasi-ordering based on the well-known Knuth-Bendix family of term orderings [29]. Yet, a Knuth-Bendix quasi-ordering (KBQO, defined in Sec. 4) cannot be used on our previous two examples; fixing even a single choice of term ordering *family* would still be too restrictive in general.

Challenge 4: Some proofs require different families of term orderings.

To address this challenge, REST (Sec. 3.2) is defined parametrically in the choice and representation of a term ordering family.

Finally, although equational reasoning is powerful enough for these examples, general verification problems usually require reasoning beyond the scope of simple rewriting. For example, simply altering Example 1 to express the disjointness hypothesis instead via cardinality as $|s_0 \cap s_1| = 0$ means that, to achieve a similar proof, reasoning within the theory of sets is necessary to deduce that this hypothesis implies the equality needed for the proof; this is beyond the abilities of term rewriting.

Challenge 5: Program verification needs proof steps not expressible by rewriting.

To address this challenge, our RESTapproach allows the integration of an external oracle that can generate equalities not justifiable by term rewriting, while still guaranteeing termination (Sec. 3.3).

3 The REST Approach

We develop REST to tackle the above five challenges and integrate a flexible, expressive, and guaranteed-terminating term rewriting system with a verification tool. REST consists of an interface for defining term orderings and an algorithm for exploring the rewrite paths supported by the term orderings. In Sec. 3.1 we describe the representation of term orderings in REST and how they address Challenges 2 and 4. In Sec. 3.2 we describe the REST algorithm that is parametric to these orderings and Sec. 3.3 describes the integration with external oracles (Challenge 5).

3.1 Representation of Term Orderings in REST

Rather than considering individual term orderings, REST operates on indexed sets (families) of term orderings (whose instances must all be thin well-quasi-orderings).

▶ **Definition 4** (Term Ordering Family). *A* term ordering family Γ *is a set of thin well-quasi-orderings on terms, indexed by some parameters P. An* instance *of the family is a term ordering obtained by a particular instantiation of P.*

For example, the concept of recursive path ordering is defined parametrically with respect to a precedence on function symbols, and therefore defines a term ordering family indexed by this choice of function symbol ordering.

A core concern of REST is determining whether any instance of a given term ordering family can orient a rewrite path. However, term ordering families cannot directly compare terms; doing so requires choosing an ordering inside the family. The root of Challenge 2 is that choosing an ordering in advance is too restrictive: different orderings are necessary to complete different proofs. The idea behind REST's search algorithm is to address this challenge by simultaneously considering all orderings in the family when considering rewrite paths and continuing the path so long as it can be oriented by *any* ordering.

To demonstrate the technique, we show how REST's approach can be derived from a naïve algorithm. The purpose of the algorithm is to determine if any ordering in a family Γ can orient a path $t_1 \to \ldots \to t_n$; i.e., if there is a $>_\mathcal{T} \in \Gamma$ such that $t_1 >_\mathcal{T} \ldots >_\mathcal{T} t_n$.

```
orients : (Set O × List 𝒯) → Bool
orients(Γ, ts) =
  os := Γ;                                        (1)
  for i ∈ 1 to |ts| − 1 {
    os := {>𝒯 ∈ os | ts_i >𝒯 ts_{i+1}};          (2)
    if (os = ∅)                                   (3)
      return false;
  }
  return true;
```

```
orients : (OCA × List 𝒯)) → Bool
orients(⟨⊤, refine, sat⟩, ts) =
  c := ⊤;
  for i ∈ 1 to |ts| − 1 {
    c := refine(c, ts_i, ts_{i+1});
    if (not(sat(c)))
      return false;
  }
  return true;
```

Figure 2 Two algorithms that determine if an ordering in the term ordering family Γ can orient a path of terms ts. **Left** presents the naïve, exhaustive algorithm. **Right** is using the ordering constraint algebra $\langle \top, \mathit{refine}, \mathit{sat} \rangle$ that returns true iff an ordering in Γ can orient ts without explicitly constructing any term orderings. $O$is the type of a term ordering.

The naïve algorithm is depicted on the left of Figure 2. The naïve algorithm works iteratively, computing the set of orderings os that can orient an increasingly-long path, short-circuiting if the set becomes empty. The algorithm enumerates each ordering in Γ and compares terms with each ordering (potentially multiple times). Unfortunately, this enumeration is not practical: some term ordering families have infinite or prohibitively large numbers of instances. `REST` avoids these issues by allowing the set of term orderings to be abstracted via a structure called an Ordering Constraint Algebra (OCA, Def. 14 of Sec. 4.2).

An OCA for a term ordering family Γ consists of a type C along with four parameters $\gamma : C \to \mathcal{P}(\Gamma)$, $\top : C$, $\mathit{refine} : C \to \mathcal{T} \to \mathcal{T} \to C$, and $\mathit{sat} : C \to \mathit{Bool}$. C is a type whose elements *represent* subsets of Γ. The function γ is the *concretisation function* of the OCA, not needed programmatically but instead defining the *meaning* of elements of C in terms of the subsets of the term ordering family they represent. The remaining three functions correspond to the operations on sets of term orderings used in lines (1), (2), and (3) of the naïve algorithm. $\top$ represents the set of all term orderings in Γ, $\mathit{refine}(c, t, u)$ filters the set of orderings represented by c to include only those where $t >_{\mathcal{T}} u$, and $\mathit{sat}(c)$ is a predicate that returns true if the set of orderings represented by c is nonempty. Figure 2 on the right shows how the ordering constraint algebra can be used to perform an equivalent computation to the naïve algorithm, without explicitly instantiating sets of term orderings. The OCA plays a role similar to abstract interpretation in a program analysis, where C is an abstraction over sets of term orderings, and the results of the abstract operations on C correspond to their concrete equivalents. Namely, we have $\gamma(\top) = \Gamma$, $\gamma(\mathit{refine}(c, t_l, t_r)) = \{\succcurlyeq \mid \succcurlyeq \in \gamma(c) \wedge t_l \succcurlyeq t_r\}$, and $\mathit{sat}(c) \Leftrightarrow \gamma(c) \neq \emptyset$.

The ordering constraint algebra enables three main advantages compared to direct computation with sets of term orderings:

1. The number of term orderings can be very large, or even infinite, thus making enumeration of the entire set intractable.
2. An OCA can provide efficient implementations for *refine* and *sat* by exploiting properties of the term ordering family. Comparing terms using the constituent term orderings requires repeating the comparison for each ordering, despite the fact that most orderings will differ in ways that are irrelevant for the comparison.
3. The OCA does not impose any requirements on the type of C or the implementation of $\top$, *refine*, and *sat*. For example, an OCA can use $\top$ and *refine* to construct logical

$\mathsf{REST} : (OCA \times \mathcal{R} \times \mathcal{T} \times (\mathcal{T} \to \text{Set } \mathcal{T})) \to \text{Set } \mathcal{T}$

```
REST(⟨⊤, refine, sat⟩, R, t_0, ℰ) =
  o := ∅;
  p := [([t_0], ⊤)];
  while (p is not empty){
    pop(ts, c) from p;
    t := last ts;
    o := o ∪ {t};
    foreach (t′ such that t′ ∉ ts ∧ (t →_R t′ ∨ t′ ∈ ℰ(t))){
      if (t′ ∈ ℰ(t) ∨ (t →_R t′ ∧ sat(refine(c, t, t′)))){
        push (ts ++ [t′], refine(c, t, t′)) to p
      }
    }
  }
  return o;
```

Figure 3 The REST algorithm.

formulas, with *sat* using an external solver to check their satisfiability. Alternatively, it could define C to be sets of term orderings that are reasoned about explicitly, and implement $\top$, *refine*, and *sat* as the operations of the naïve algorithm.

We now describe how the REST algorithm uses the OCA to explore rewrite paths.

3.2 The REST Algorithm

Figure 3 presents the REST algorithm. The algorithm takes four parameters. The first parameter is an OCA $\langle \top, \mathit{refine}, \mathit{sat} \rangle$, as discussed above. The algorithm's second parameter, R, is a finite set of term rewriting rules (not required to be terminating); for example, we could pass the oriented rewrite rules corresponding to Figure 1. The third parameter t_0 is the term from which term rewrites are sought. The final parameter $\mathcal{E}$ acts as an external oracle, generating additional rewrite steps that need *not* follow from the term rewriting rules R. To simplify the explanation, we will initially assume that $\mathcal{E} = \lambda t.\emptyset$, i.e., this parameter has no effect. Our algorithm produces a set of terms, each of which are reachable by *some* rewrite path beginning from t_0, and for which *some* ordering allows the rewrite path. The algorithm addresses Challenge 1 (termination; Theorem 22) because every path must be finite: no ordering could orient an infinite path.

Our algorithm operates in worklist fashion, storing in p a list of pairs (ts, c) where ts is a non-empty list of terms representing a rewrite path already explored (the head of which is always t_0) and c tracks the ordering constraints of the path so far. The set o records the output terms (initially empty): all terms discovered (down any rewrite path) equal to t_0 via the rewriting paths explored.

While there are still rewrite paths to be extended, i.e., p is not empty, a tuple (ts, c) is popped from p. REST puts t, i.e., the last term of the path, into the set of output terms o and considers all terms t' that are: (a) not *already* in the path and (b) reachable by a single rewrite step of R (or returned by the function $\mathcal{E}$ explained later). The crucial decision of whether or not to extend a rewrite path with the additional step $t \to t'$ is handled in the **if** check of REST. This check is to guarantee termination, by enforcing that we only add

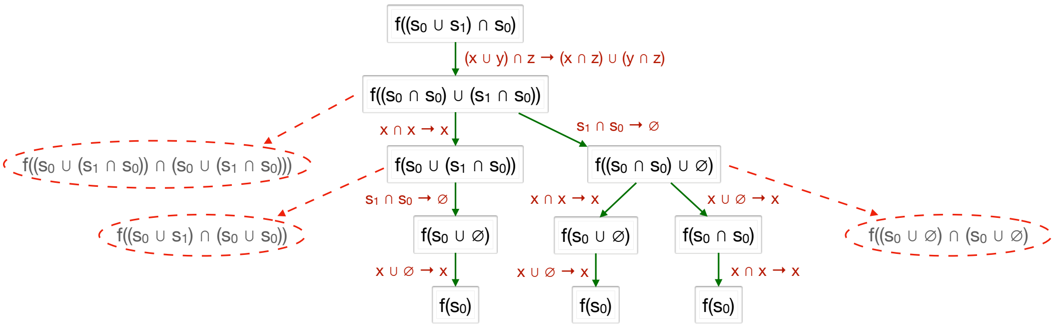

Figure 4 A visualization of REST running on the term from Example 1. Each path through the tree shown represents a rewrite path uncovered by our algorithm; the edge labels show the rewrite rule applied. The red dotted lines indicate rewrite steps rejected by REST.

rewrite steps which would leave the extended path still justifiable by *some* term ordering, as enforced by the *sat* check.

Figure 4 visualizes the rewrite paths explored by our algorithm for a run corresponding to the problem from Example 1, using the OCA for the recursive path quasi-ordering (Sec. 4.2.1)[2]. The manual proof in Example 1 corresponds to the right-most path in this tree; the other paths apply the same reasoning steps in different orders. In our implementation, we optimize the algorithm to avoid re-exploring the same term multiple times unless this could lead to further rewrites being discovered (cf. Sec. 6).

The arrow from the root of the tree to its child corresponds to the first rewrite REST applies: $f((s_0 \cup s_1) \cap s_0) \to f((s_0 \cap s_0) \cup (s_1 \cap s_0))$. This rewrite step can only be oriented by RPQOs with precedence $\cap > \cup$; therefore applying this rewrite constrains the set of RPQOs that REST must consider in subsequent applications. For example, the rewrite to the left child of $f((s_0 \cap s_0) \cup (s_1 \cap s_0))$ can only be oriented by RPQOs with precedence $\cup > \cap$. Since no RPQO can have both $\cap > \cup$ and $\cup > \cap$, no RPQO can orient the entire path from the root; REST must therefore reject the rewrite. On the other hand, the rewrite to the right child can be oriented by any RPQO where $s_0 > \emptyset$, $s_1 > \emptyset$, or $\cap > \emptyset$. The path from the root can thus continue down the right-hand side, as there are RPQOs that satisfy both $\cap > \cup$ and the other conditions. The subsequent rewrites down the right-hand side do not impose any new constraints on the ordering: $f((s_0 \cap s_0) \cup \emptyset) >_\mathcal{T} f(s_0 \cap s_0) >_\mathcal{T} f(s_0)$ in all RPQOs.

Similarly, REST will prove Example 2 but will reject Example 3 when the input OCA represents RPQO orderings. As shown in our benchmarks (Table 2 of Sec. 7), Example 3 is solved by REST with an OCA for the Knuth-Bendix term ordering family.

3.3 Integrating an External Oracle

Finally, to tackle Challenge 5, we turn to the (so far ignored) third parameter of the algorithm, the external oracle $\mathcal{E}$. In the example variant presented at the end of Sec. 2, such a function might supply the rewrite step $s_0 \cap s_1 \to \emptyset$ by analysis of the logical assumption $|s_0 \cap s_1| = 0$, which goes beyond term-rewriting. More generally, any external solver capable of producing rewrite steps (equal terms) can be connected to our algorithm via $\mathcal{E}$. In our implementation in Liquid Haskell, we use the pre-existing *Proof by Logical Evaluation (PLE)*

[2] We omit the commutativity rules from this run, just to keep the diagram easy to visualize, but our implementation handles the example easily with or without them.

technique [50], which complements rewriting with the expansion of program function definitions, under certain checks made via SMT solving. Our only requirements on the oracle $\mathcal{E}$ are that the binary relation on terms generated by calls to it is bounded (finitely-branching) and strongly normalizing (cf. Sec. 5).

Our algorithm therefore flexibly allows the interleaving of term rewriting steps and those justified by the external oracle; we avoid the potential for this interaction to cause non-termination by conditioning any further rewriting steps on the fact that the entire path (including the steps inserted by the oracle) can be oriented by at least one candidate term ordering.

The combination of our interfacing for defining term orderings via ordering constraint algebras, a search algorithm that effectively explores all rewrites enabled by the orderings, and the flexible possibility of combination with external solvers via the oracle parameter makes REST very adaptable and powerful in practice.

4 Well-Quasi-Orderings and the Ordering Constraint Algebra

Term orderings are typically defined as *strict well-founded* orderings; this requirement ensures that rewriting will obtain a normal form. However, as mentioned in Challenge 3, the restriction to strict orderings limits what can be achieved with rewriting. In this section we describe the derivation of well-quasi-orderings from strict orderings (Sec. 4.1) and introduce Knoth-Bendix quasi-orderings (Sec. 4.1.1) and recursive path quasi-orderings (Sec. 4.1.2), two novel term ordering families respectively based on the classical recursive path and Knoth-Bendix orderings. In addition, we formally introduce ordering constraint algebras (Sec. 4.2) and use them to develop an efficient ordering constraint algebra for RPQOs.

4.1 Well-Quasi-Orderings

We define well-quasi-orderings in the standard way.

▶ **Definition 5** (Well-Quasi-Orderings). *A relation $\geqslant$ is a quasi-order if it is reflexive and transitive. Given elements t and u in S, we say $t \approx u$ if $t \geqslant u$ and $u \geqslant t$. A quasi-order $\geqslant$ is also characterized as:*

1. *WQO, when for all infinite chains $x_1, x_2, \ldots$ there exists an $i, j, i < j$ such that $x_j \geqslant x_i$,*
2. *thin, when forall $t \in S$, the set $\{u \in S \mid t \approx u\}$ is finite, and*
3. *total, when for all $t, u \in S$ either $t \geqslant s$ or $s \geqslant t$.*

Well-quasi-orderings are not required to be antisymmetric, however the corresponding strict part of the ordering must be well-founded. Hence, a WQO derives a strict ordering over equivalence classes of terms; REST also requires that these equivalence classes are finite (i.e., the ordering is thin). With this requirement, REST guarantees termination by exploring only duplicate-free paths.

Many simplification orderings can be converted into more permissive WQOs. Intuitively, given an ordering $>_o$ its quasi-ordering derivation also accepts equal terms, so we denote it as $\geqslant_o$. We next present two such derivations.

4.1.1 Knuth-Bendix Quasi-Orderings (KBQO)

The Knuth-Bendix ordering [29] is a well-known simplification ordering used in the Knuth-Bendix completion procedure. Here, we present a simplified version of the ordering, used by REST that is using ordering to only compare ground terms.

▶ **Definition 6.** *A weight function w is a function $\mathcal{F} \to \mathbb{N}$, where $w(f) > 0$ for all nullary functions symbols, and $w(f) = 0$ for at most one unary function symbol. w is compatible with a quasi-ordering $\geqslant_{\mathcal{F}}$ on $\mathcal{F}$ if, for any unary function f such that $w(f) > 0$, we have $f >_{\mathcal{F}} g$ for all g. $w(t)$ denotes the weight of a term t, such that $w(f(t_1, \ldots, t_n)) = w(f) + \sum_{1 \leqslant i \leqslant n} w(t_i)$*

▶ **Definition 7** (Knuth-Bendix ordering (KBO) on ground terms)**.** *The Knuth-Bendix Ordering $>_{kbo}$ for a given weight function w and compatible precedence order $\geqslant_{\mathcal{F}}$ is defined as $f(t_1, \ldots, t_m) = t >_{kbo} u = g(u_1, \ldots, u_n)$ iff $w(t) \geqslant w(u)$, and:*

1. *$w(t) > w(u)$, or*
2. *$f >_{\mathcal{F}} g$, or*
3. *$f \geqslant_{\mathcal{F}} g$, and $(t_1, \ldots, t_m) >_{kbolex} (u_1, \ldots, u_n)$.*

Where $>_{kbolex}$ performs a lexicographic comparison using $>_{kbo}$ as the underlying ordering.

Intuitively, KBO compares terms by their weights, using $\geqslant_{\mathcal{F}}$ and the lexicographic comparison as "tie-breakers" for cases when terms have equal weights. However, as $\geqslant$ is already a well-quasi-ordering on $\mathbb{N}$, we can derive a more general ordering by removing these tie-breakers and the need for a precedence ordering at all.

▶ **Definition 8** (Knuth-Bendix Quasi-ordering (KBQO))**.** *Given a weight function w, the Knuth-Bendix quasi-ordering $\geqslant_{kbo}$ is defined as $t \geqslant_{kbo} u$ iff $w(t) \geqslant w(u)$.*

The resulting quasi-ordering is considerably simpler to implement and is more permissive: $t >_{kbo} u$ implies $t \geqslant_{kbo} u$; and also enables arbitrary associativity and commutativity axioms as rewrite rules, since it only considers the weights of the function symbols and no structural components of the term. However, one caveat is that REST operates on well-quasi-ordering that are thin (Def. 5) and therefore can only consider KBQOs where $w(f) > 0$ for all unary function symbols f.

However, the fact that KBO and KBQO largely ignore the structure of the term in their comparison has a corresponding downside: it is not possible to orient distributivity axioms, or many other axioms that increase the number of symbols in a term. Therefore, we have found that a WQO derived from the recursive path ordering [15] to be more useful in practice.

4.1.2 Recursive Path Quasi-Orderings (RPQO)

In this section, we define a particular family of orderings designed to be typically useful for term-rewriting via REST. Our family of orderings is a novel extension of the classical notion of RPO, designed to also be more compatible with symmetrical rules such as commutativity and associativity (cf. Challenge 3, Sec. 2).

Like the classical RPO notions, our *recursive path quasi-ordering* (RPQO) is defined in three layers, derived from an underlying ordering on function symbols:

- The input ordering $\succcurlyeq_{\mathcal{F}}$ can be any quasi-ordering over $\mathcal{F}$.
- The corresponding *multiset quasi-ordering* $\succcurlyeq_{M(X)}$ lifts an ordering $\succcurlyeq_X$ over X to an ordering $\succcurlyeq_{M(X)}$ over multisets of X. Intuitively $T \succcurlyeq_{M(X)} U$ when U can be obtained from T by replacing zero or more elements in T with the same number of equal (with respect to $\succcurlyeq_X$) elements, and replacing zero or more elements in T with a finite number of smaller ones (Def. 9).

- Finally, the corresponding *recursive path quasi-ordering* $\succcurlyeq_{rpo}$ is an ordering over terms. Intuitively $f(ts) \succcurlyeq_{rpo} g(us)$ uses $\succcurlyeq_{\mathcal{F}}$ to compare the function symbols f and g and the corresponding $\succcurlyeq_{M(rpo)}$ to compare the argument sets ts and us (Def. 10).

Below we provide the formal definitions of the multiset quasi-ordering and recursive path quasi-ordering respectively generalized from the multiset ordering of [18] and the recursive path ordering [15] to operate on quasi-orderings. For all the three orderings, we write $x_l < x_r \doteq x_l \not\succcurlyeq x_r$ and $x_l > x_r \doteq x_l \succcurlyeq x_r \wedge x_r \not\succcurlyeq x_l$.

▶ **Definition 9** (Multiset Ordering). *Given a ordering $\succcurlyeq_X$ over a set X, the derived multiset ordering $\succcurlyeq_{M(X)}$ over finite multisets of X is defined as $T \succcurlyeq_{M(X)} U$ iff:*

1. *$U = \emptyset$, or*
2. *$t \in T \wedge u \in U \wedge t \approx u \wedge (T - t) \succcurlyeq_{M(X)} (U - u)$, or*
3. *$t \in T \wedge (T - t) \succcurlyeq_{M(X)} (U \setminus \{u \in U \mid u <_X t\})$.*

▶ **Definition 10** (Recursive Path Quasi-Ordering). *Given a basic ordering $\succcurlyeq_{\mathcal{F}}$, the recursive path quasi-ordering (RPQO) is the ordering $\succcurlyeq_{rpo}$ over $\mathcal{T}$ defined as follows: $f(t_1, \ldots, t_m) \succcurlyeq_{rpo} g(u_1, \ldots, u_n)$ iff*

1. *$f >_{\mathcal{F}} g$ and $\{f(t_1, \ldots, t_m)\} >_{M(rpo)} \{u_1, \ldots, u_n\}$, or*
2. *$g >_{\mathcal{F}} f$ and $\{t_1, \ldots, t_m\} \succcurlyeq_{M(rpo)} \{g(u_1, \ldots, u_n)\}$, or*
3. *$f \approx g$ and $\{t_1, \ldots, t_m\} \succcurlyeq_{M(rpo)} \{u_1, \ldots, u_n\}$.*

▶ **Example 11.** As a first example, any RPQO $\succcurlyeq_{\mathcal{T}}$ used to restrict term rewriting will accept the rule $x + y \to y + x$, since $x + y \succcurlyeq_{\mathcal{T}} y + x$ always holds. Since the top level function symbol is the same $+ \approx +$, by Def. 10(3) we need to show $\{x, y\} \succcurlyeq_{M(rpo)} \{y, x\}$. By Def. 9(2) (choosing both t and u to be x), we can reduce this to $\{y\} \succcurlyeq_{M(rpo)} \{y\}$; the same step applied to y reduces this to showing $\emptyset \succcurlyeq_{M(rpo)} \emptyset$ which follows directly from Def. 9(3).

From this example, we can see that both $x + y \succcurlyeq_{rpo} y + x$ and $y + x \succcurlyeq_{rpo} x + y$ hold, in this case independently of the choice of input ordering $\succcurlyeq_{\mathcal{F}}$ on function symbols. In our next example, the choice of input ordering makes a difference.

▶ **Example 12.** As a next example, we compare the terms $s(x){+}y$ and $s\ (x{+}y)$. Now that the outer function symbols are *not* equal, the order relies on the ordering between $+$ and s. Let's assume that $+ >_{\mathcal{F}} s$. Now to get $s(x){+}y \succcurlyeq_{rpo} s\ (x{+}y)$, the 1st case of Definition 10 further requires $\{s(x){+}y\} >_{M(rpo)} \{x{+}y\}$, which holds if $s(x){+}y >_{rpo} x{+}y$. The outermost symbol for both expressions is $+$, so we must check the multiset ordering: $\{s(x), y\} >_{M(rpo)} \{x, y\}$, which holds because by case splitting on the relation between s and x, we can show that $s(x)$ is always smaller than x. In short, if $+ >_{\mathcal{F}} s$, then $s(x) + y \succcurlyeq_{rpo} s\ (x + y)$.

Developing on our *RPQO* notion (Def. 10), we consider the set of *all* such orderings that are generated by any total, well-quasi-ordering over the operators. We prove that such term orderings satisfy the termination requirements of Theorem 22. Concretely:

▶ **Theorem 13.** *If $\succcurlyeq_{\mathcal{F}}$ is a total, well-quasi-ordering, then*

1. *$\succcurlyeq_{rpo}$ is a well-quasi-ordering,*
2. *$\succcurlyeq_{rpo}$ is thin, and*
3. *$\succcurlyeq_{rpo}$ is thin well-founded.*

Proof. The detailed proofs can be found in App. B. (1) uses the well-foundedness theorem of Dershowitz [15] and the fact that $\succcurlyeq_{rpo}$ is a quasi-simplification ordering. (2) relies on the fact that a finite number of function symbols can only generate a finite number of equal terms. (3) is a corollary of (1) and (2) combined. ◀

4.2 Ordering Constraint Algebras

Ordering constraint algebras play a crucial role in the REST algorithm (Sec. 3.2), by enabling the algorithm to simultaneously consider an entire family of term orderings during the exploration of rewrite paths. In this section, we provide a formal definition for ordering constraint algebras and describe the construction of an algebra for the RPQO.

▶ **Definition 14** (Ordering Constraint Algebra). *An Ordering Constraint Algebra (OCA) $\mathcal{A}_{(T,\Gamma)}$ over a set of terms T and term ordering family Γ, is a five-tuple $\mathcal{A}_{(T,\Gamma)} \doteq \langle C, \gamma, \top, \mathit{refine}, \mathit{sat}\rangle$, where:*

1. *C, the constraint language, can be any non-empty set. Elements of C are called constraints, and are ranged over by c.*
2. *γ, the concretization function of $\mathcal{A}_{(T,\Gamma)}$, is a function from elements of C to subsets of Γ.*
3. *$\top$, the top constraint, is a distinguished constant from C, satisfying $\gamma(\top) = \Gamma$.*
4. *refine, the refinement function, is a function $C \to T \to T \to C$, satisfying (for all c, t_l, t_r) $\gamma(\mathit{refine}(c, t_l, t_r)) = \{\succcurlyeq \mid \succcurlyeq \in \gamma(c) \ \wedge\ t_l \succcurlyeq t_r\}$.*
5. *sat, the satisfiability function, is a function $C \to \mathit{Bool}$, satisfying (for all c) $\mathit{sat}(c) = \mathit{true} \Leftrightarrow \gamma(c) \neq \emptyset$.*

The functions $\top$, *refine*, and *sat* are all called from our REST algorithm (Figure 3), and must be implemented as (terminating) functions when implementing REST. Specifically, REST instantiates the initial path with constraints $c = \top$. When a path can be extended via a rewrite application $t_l \to_R t_r$, REST refines the prior path constraints c to $c' \doteq \mathit{refine}(c, t_l, t_r)$. Then, the new term is added to the path only if the new constraints are satisfiable ($\mathit{sat}(c')$ holds); that is, if c' admits an ordering that orients the generated path. The function γ need *not* be implemented in practice; it is purely a mathematical concept used to give semantics to the algebra.

Given terms T and a finite term ordering family Γ, a trivial OCA is obtained by letting $C = \mathcal{P}(\Gamma)$, and making γ the identity function; straightforward corresponding elements $\top$, *refine*, and *sat* can be directly read off from the constraints in the definition above.

However, for efficiency reasons (or in order to support potentially infinite sets of orderings, which our theory allows), tracking these sets symbolically via some suitably chosen constraint language can be preferable. For example, consider lexicographic orderings on pairs of constants, represented by a set T of terms of the form $p(q_1, q_2)$ for a fixed function symbol p and q_1, q_2 chosen from some finite set of constant symbols Q. We choose the term ordering family $\Gamma = \{\succcurlyeq_{lex(\succcurlyeq)} \mid \succcurlyeq \textit{ is a total order on } Q\}$ writing $\succcurlyeq_{lex(\succcurlyeq)}$ to mean the corresponding lexicographic ordering on $p(q_1, q_2)$ terms generated from an ordering $\succcurlyeq$ on Q.

A possible OCA over these T and Γ can be defined by choosing the constraint language C to be *formulas*: conjunctions and disjunctions of atomic constraints of the forms $q_1 > q_2$ and $q_1 = q_2$ prescribing conditions on the underlying orderings on Q. The concretization γ is given by $\gamma(c) = \{\succcurlyeq_{lex(\succcurlyeq)} \mid \succcurlyeq \textit{ satisfies } c\}$, i.e., a constraint maps to all lexicographic orders generated from orderings of Q that satisfy the constraints described by c, defined in the natural way. We define $\top$ to be e.g., $q = q$ for some $q \in Q$. A satisfiability function *sat* can be implemented by checking the satisfiability of c as a formula. Finally, by inverting the standard definition of lexicographic ordering, we define:

$$\mathit{refine}(c, p(q_1, q_2), p(r_1, r_2)) = c \wedge (q_1 > r_1 \vee (q_1 = r_1 \wedge q_2 > r_2))$$

Using this example algebra, suppose that REST explores two potential rewrite steps $p(a_1, a_2) \rightarrow p(b_1, a_2) \rightarrow p(a_1, a_1)$. Starting from the initial constraint $c_0 = \top$, the constraint for the first step $c_1 \doteq \mathit{refine}(c_0, p(a_1, a_2), p(b_1, a_2)) = a_1 > b_1 \vee (a_1 = b_1 \wedge a_2 > a_2)$ is satisfiable, e.g., for any total order for which $a_1 > b_1$. However, considering the subsequent step, the refined constraint $c_2 \doteq \mathit{refine}(c_1, p(b_1, a_2), p(a_1, a_1))$, computed as $c_2 = c_1 \wedge (a_2 > a_2 \vee (a_2 = a_2 \wedge b_1 > a_1))$ is no longer satisfiable. Note that this allows us to conclude that there is no lexicographic ordering allowing this sequence of two steps, even without explicitly constructing any orderings.

We now describe an OCA for RPQOs (Sec. 4.1.2), based on a compact representation of sets of these orderings.

4.2.1 An Ordering Constraint Algebra for $\succcurlyeq_{rpo}$

The OCA for RPQOs enables their usage in REST's proof search. One simple but computationally intractable approach would be to enumerate the entire set of RPQOs that orient a path; continuing the path so long as the set is not empty. This has two drawbacks. First, the number of RPQOs grows at an extremely fast rate with respect to the number of function symbols; for example there are $6,942$ RPQOs describing five function symbols, and $209,527$ over six. Second, most of these orderings differ in ways that are not relevant to the comparisons made by REST.

Instead, we define a language to succinctly describe the set of candidate RPQOs, by calculating the minimal constraints that would ensure orientation of the path of terms; REST continues so long as there is some RPQO that satisfies the constraints. Crucially the satisfiability check can be performed effectively using an SMT solver, as described in Sec. 6.2, without actually instantiating any orderings.

Before formally describing the language, we begin with some examples, showing how the ordering constraints could be constructed to guide the termination check of REST.

▶ **Example 15** (Satisfiability of Ordering Constraints). Consider the following rewrite path given by the rules $r_1 \doteq f(g(x), y) \rightarrow g(f(y, y))$ and $r_2 \doteq f(x, x) \rightarrow f(k, x)$:

$$f(g(h), k) \rightarrow_{r_1} g(f(h, h)) \rightarrow_{r_2} g(f(k, h))$$

To perform the first rewrite REST has to ensure that there exists an RPQO $\succcurlyeq_{rpo}$ such that $f(g(h), k) \succcurlyeq_{rpo} g(f(h, h))$. Following from Definition 10, we obtain three possibilities:

1. $f >_{\mathcal{F}} g$ and $\{f(g(h), k)\} >_{M(rpo)} \{f(h, h)\}$, or
2. $g >_{\mathcal{F}} f$ and $\{g(h), k\} \succcurlyeq_{M(rpo)} \{g(f(h, h))\}$, or
3. $f \approx g$ and $\{g(h), k\} \succcurlyeq_{M(rpo)} \{f(h, h)\}$.

We can further simplify these using the definition of the multiset quasi-ordering (Def. 9). Concretely, the multiset comparison of (1) always holds, while the multiset comparisons of (2) and (3) reduce to $k >_{\mathcal{F}} f \wedge k >_{\mathcal{F}} g \wedge k >_{\mathcal{F}} h$. Thus, we can define the exact constraints c_0 on $\succcurlyeq_{rpo}$ to satisfy $f(g(h), k) \succcurlyeq_{rpo} g(f(h, h))$ as

$$c_0 \doteq f >_{\mathcal{F}} g \vee (k >_{\mathcal{F}} f \wedge k >_{\mathcal{F}} g \wedge k >_{\mathcal{F}} h)$$

Since there exist many quasi-orderings satisfying this formula (trivially, the one containing the single relation $f >_{\mathcal{F}} g$), the first rewrite is satisfiable.

Similarly, for the second rewrite, the comparison $g(f(z, z)) \succcurlyeq_{rpo} g(f(k, z))$ entails the constraints $c_1 \doteq z \succcurlyeq_{\mathcal{F}} k$. To perform this second rewrite the conjunction of c_0 and c_1 must be satisfiable. Since the second disjunct of c_0 contradicts c_1, the resulting constraints $f >_{\mathcal{F}} g \wedge z \succcurlyeq_{\mathcal{F}} k$ is satisfiable by an RPQO, thus the path is satisfiable.

▶ **Example 16** (Unsatisfiable Ordering Constraint). As a second example, consider the rewrite rules $r_1 \doteq f(x) \to g(s(x))$ and $r_2 \doteq g(s(x)) \to f(h(x))$. These rewrite rules can clearly cause divergence, as applying rule r_1 followed by r_2 will enable a subsequent application of r_1 to a larger term. Now let's examine how our ordering constraint algebra can show the unsatisfiability of the diverging path:

$$f(z) \to_{r_1} g(s(z)) \not\to_{r_2} f(h(z))$$

$f(z) \succcurlyeq_{rpo} g(s(z))$ requires $c_0 \doteq f > g \wedge f > s$ which is satisfiable, but $g(s(z)) \succcurlyeq_{rpo} f(h(z))$ requires $c_1 \doteq (g \geqslant f \wedge g \geqslant h) \vee (g \geqslant f \wedge s \geqslant h) \vee (s > f \wedge s > h)$, which, although satisfiable on it's own, conflicts with c_0. Since no $RPQO$ can satisfy both c_0 and c_1, the rewrite path is not satisfiable.

Having primed intuition through the examples, we now present a way to compute such constraints. First, it is clear that we can define an RPQO based on the precedence over symbols $\mathcal{F}$. Therefore, we define our language of constraints to include the standard logical operators as well as atoms representing the relations between elements of $\mathcal{F}$, as:

$$C_{\mathcal{F}} \doteq f >_{\mathcal{F}} g \mid f \approx g \mid C_{\mathcal{F}} \wedge C_{\mathcal{F}} \mid C_{\mathcal{F}} \vee C_{\mathcal{F}} \mid \top \mid \bot$$

Next, we lift our definition of $RPQO$ and the multiset quasi-ordering to derive functions: $rpo : \mathcal{T} \to \mathcal{T} \to C_{\mathcal{F}}$, and $mul : (\mathcal{T} \to \mathcal{T} \to C_{\mathcal{F}}) \to M(\mathcal{T}) \to M(\mathcal{T}) \to C_{\mathcal{F}}$. rpo is derived by a straightforward translation of Def. 10:

$$\begin{aligned} rpo(f(t_1,\ldots,t_m), g(u_1,\ldots,u_n)) = \quad & f >_{\mathcal{F}} g \;\wedge\; mul'(rpo, \{f(t_1,\ldots,t_m)\}, \{u_1,\ldots,u_n\}) \vee \\ & g >_{\mathcal{F}} f \;\wedge\; mul(rpo, \{t_1,\ldots,t_m\}, \{g(u_1,\ldots,u_n)\}) \vee \\ & f \approx g \;\wedge\; mul(rpo, \{t_1,\ldots,t_m\}, \{u_1,\ldots,u_n\}) \end{aligned}$$

where mul' is the strict multiset comparison: $mul'(f,T,U) = mul(f,T,U) \wedge \neg mul(f,U,T)$. $\neg : C_{\mathcal{F}} \to C_{\mathcal{F}}$ inverts the constraints, with $\neg(f >_{\mathcal{F}} g) = f \approx g \vee g >_{\mathcal{F}} f$ and $\neg(f \approx g) = f >_{\mathcal{F}} g \vee g >_{\mathcal{F}} f$; the other cases are defined in the typical way.

The definition for mul is more complex. Recall that $T \succcurlyeq_{M(X)} U$ when U can be obtained from T by replacing zero or more elements in T with the same number of equal (with respect to $\succcurlyeq_X$) elements, and by replacing zero or more elements in T with a finite number of smaller ones. Therefore each justification for $\{t_1,\ldots,t_m\} \succcurlyeq_{M(X)} \{u_1,\ldots,u_n\}$ can be represented by a bipartite graph with nodes labeled $t_1,\ldots,t_m$ and $u_1,\ldots,u_n$, such that:

1. Each node u_i has exactly one incoming edge from some node t_j.
2. If a node t_i has exactly one outgoing edge, it is labeled either `GT` or `EQ`.
3. If a node t_i has more than one outgoing edge, it is labeled `GT`.

$mul(f, \{t_1,\ldots,t_m\}, \{u_1,\ldots,u_n\})$ generates all such graphs: for each graph converts each labeled edge $(t, u, \texttt{EQ})$ to the formula $f(t,u) \wedge f(u,t)$, each edge $(t, u, \texttt{GT})$ to the formula $f(t,u) \wedge \neg f(u,t)$, and finally joins the formulas for the graph via a conjunction. The resulting constraint is defined to be the disjunction of the formulas generated from all such graphs.

Having defined the lifting of recursive path quasi-orderings to the language of constraints, we define our ordering constraint algebra $\mathcal{A}_{(\mathcal{T},\Gamma)}$ as the tuple $\langle C_{\mathcal{F}}, \top, refine, \gamma, sat\rangle$ where:

- $refine(c,t,u) = c \wedge rpo(t,u)$,
- Γ is the set of all RPQOs,
- $\gamma(c)$ is the set of RPQOs derived from the underlying quasi-orders $\succcurlyeq_{\mathcal{F}}$ that satisfy c, and
- $sat(c) = true$ if and only if there exists a quasi-order $\succcurlyeq_{\mathcal{F}}$ satisfying c.

That $\mathcal{A}_{(\mathcal{T},\Gamma)}$ is an OCA, i.e., satisfies the requirements of Def. 14, follows by construction. Namely, the function $rpo(t, u)$ produces constraints c such that, for any RPQO $\succcurlyeq_{rpo}$, $t \succcurlyeq_{rpo} u$ if and only if its underlying ordering $\succcurlyeq_{\mathcal{F}}$ satisfies c. In Sec. 6.2 we further discuss how the satisfiability check is mechanized and implemented using an SMT solver.

Having shown that using RPQOs as a term ordering is useful for theorem proving, satisfies the necessary properties for REST, and admits an efficient ordering constraint algebra, we continue our formal work by stating and proving the metaproperties of REST.

5 REST Metaproperties: Soundness, Completeness, and Termination

We now present the metaproperties of the REST algorithm defined in Figure 3. We show correctness (Theorem 17), completeness (Theorem 19) relative to the input term ordering family (recall that its instances must all be thin well-quasi-orderings), and termination (Theorem 22) which requires that calls to the OCA functions used in the algorithm, as well as the external oracle function, themselves terminate. The property that the orderings are thin well-founded guarantees in particular that any *duplicate-free* path (such as those that REST generates) that can be oriented by any of these orderings is guaranteed to be finite. We provide here the key invariants and statements of the formal results, and relegate the detailed proofs to App. A.

5.1 Formal Definitions

Our formalism of rewriting is standard; based on the terminology of [28]. Our language consists of the following:

1. An infinite set of meta-variables (the variables for rewrite rules) $\mathcal{V}$ with elements X, Y, ...
2. A finite set of function symbols $\mathcal{F}$ with elements f, g, ..., x, y, ...
   Each operator is associated with a fixed numeric arity and types for its arguments and result (elided here, for simplicity). By convention, we use the variables x, y to range over zero-arity function symbols (constants).
3. A set of terms $\mathcal{T}$ with elements $t, u, \ldots$ inductively defined as follows: (a) $X \in \mathcal{V} \Rightarrow X \in \mathcal{T}$ and (b) $f \in \mathcal{F}$, f has arity n, $t_1, \ldots, t_n \in \mathcal{T} \Rightarrow f(t_1, \ldots, t_n) \in \mathcal{T}$.

We use $FV(t)$ to refer to the set of meta-variables in t. A term t is *ground* if $FV(t) = \emptyset$.

A *substitution* $\sigma \subseteq \mathcal{V} \times \mathcal{T}$ is a mapping from meta-variables to terms. We write $\sigma \cdot t$ to denote the simultaneous application of the substitution: namely, $\sigma \cdot t$ replaces each occurrence of each meta-variable X in t with $\sigma(X)$. A substitution σ *grounds* t if, for all $X \in FV(t)$, $\sigma(X)$ is a ground term. A substitution σ *unifies* two terms t and u if $\sigma \cdot t = \sigma \cdot u$.

A *context* E is a term-like object that contains exactly one term placeholder $\bullet$. If t is a term, then $E[t]$ is the term generated by replacing the $\bullet$ in E with t.

A *rewrite rule* r is a pair of terms $r \doteq (t, u)$ such that $FV(u) \subseteq FV(t)$ and $t \notin \mathcal{V}$. Each rewrite rule $r \doteq (t, u)$ defines a binary relation $\rightarrow_r$ which is the smallest relation such that, for all contexts E and substitutions σ grounding t (and therefore u), $E[\sigma \cdot t] \rightarrow_r E[\sigma \cdot u]$.

We use R to range over sets of rewrite rules. We write $v \rightarrow_R w$ iff $v \rightarrow_r w$ for some $r \in R$.

For oracle functions (from terms to sets of terms) $\mathcal{E}$, we write $t \rightarrow_{\mathcal{E}} t'$ *iff* $t' \in \mathcal{E}(t)$. We write $t \rightarrow_{R+\mathcal{E}} t'$ if $t \rightarrow_R t'$ or $t \rightarrow_{\mathcal{E}} t'$. For a relation $\rightarrow$ we write $\rightarrow^*$ for its reflexive, transitive closure. A *path* is a list of terms. A binary relation $\succcurlyeq$ *orients* a path $t_1, \ldots, t_n$ if $\forall i, 1 \leq i < n, t_i \succcurlyeq t_{i+1}$.

5.2 Soundness

Soundness of REST means that any term of the output ($u \in \text{REST}(\mathcal{A}, R, t_0, \mathcal{E})$) can be derived from the original input term by some combination of term rewriting steps from R and steps via the oracle function $\mathcal{E}$ (in other words, $t_0 \to^*_{R+\mathcal{E}} u$).

Our proof relies on the following simple invariant of REST: any path stored in the stack during the execution of the algorithm can be derived by the rewrite rules in R or the external oracle $\mathcal{E}$.

▶ **REST Invariant 1** (Path Invariant)**.** *For any execution of REST$(\mathcal{A}, R, t_0, \mathcal{E})$, at the start of any iteration of the main loop, for each $(ts, c) \in p$, the list ts is a path of $R + \mathcal{E}$ starting from t_0.*

Proof. (Sketch:) By straightforward induction on iterations of the main loop. ◀

▶ **Theorem 17** (Soundness of REST)**.** *For all R, u, and t_0, if $u \in$ REST$(\mathcal{A}, R, t_0, \mathcal{E})$, then $t_0 \to^*_{R+\mathcal{E}} u$.*

Proof. In each iteration of REST, the term t added to the output o is the last element of the list ts for the tuple $(ts, c) \in p$. By Invariant 1, t must be on the path of $R + \mathcal{E}$ starting from t_0. ◀

5.3 Completeness

A naïve completeness statement for REST might be that, for any terms t_0 and u, if $t_0 \to^*_{R+\mathcal{E}} u$ then u is in our output ($u \in \text{REST}(\mathcal{A}, R, t_0, \mathcal{E})$). This result doesn't hold in general by design, since REST explores only paths permitted by at least one candidate instance of its input term ordering family. We prove this *relative* completeness result in two stages. First (Theorem 18), we show that completeness always holds if all steps only involve the external oracle. Then (Theorem 19), we prove relative completeness of REST with respect to the provided term ordering family. We begin by stating another simple invariant of our algorithm: that any term appearing in a path in the stack p, will belong to the final output:

▶ **REST Invariant 2.** *For any execution of REST$(\mathcal{A}, R, t_0, \mathcal{E})$, at the start of any iteration of the main loop, if $t \in ts$ and $(ts, c) \in p$, then, when the algorithm terminates, we will have $t \in$ REST$(\mathcal{A}, R, t_0, \mathcal{E})$.*

Proof. (Sketch:) We can prove inductively that terms contained in any list in p either remain in p or end up in o; since p is empty on termination, the result follows. ◀

▶ **Theorem 18** (Completeness w.r.t. $\mathcal{E}$)**.** *For all R, u, and t_0, if $t_0 \to^*_{\mathcal{E}} u$, then $u \in$ REST$(\mathcal{A}, R, t_0, \mathcal{E})$.*

Proof. (Sketch:) Since, $\mathcal{E}$ is strongly normalizing, the path of terms $t_0 \to_{\mathcal{E}} \ldots \to_{\mathcal{E}} u$ will not contain any duplicates; REST will therefore insert each term in the path into ts. Since u is in that path, Invariant 2 ensures $u \in \text{REST}(\mathcal{A}, R, t_0, \mathcal{E})$. ◀

▶ **Theorem 19** (Relative Completeness)**.** *For all R, u, and t_0, if $t_0 \to^*_{R+\mathcal{E}} u$ and there exists an ordering $\succcurlyeq \in \gamma(\top)$ that orients the path justifying $t_0 \to^*_{R+\mathcal{E}} u$, then $u \in$ REST$(\mathcal{A}, R, t_0, \mathcal{E})$.*

Proof. (Sketch:) The proof structure is similar to Theorem 18; in this case the terms in the path are guaranteed to be in ts because some ordering in $\gamma(\top)$ can orient the path. ◀

5.4 Termination

Termination of REST requires appropriate conditions on the external oracle $\mathcal{E}$ and the ordering constraint algebra $\mathcal{A}$ employed. We formally define these requirements and then prove termination of REST.

▶ **Definition 20** (Well-Founded ordering constraint algebras)**.** *For ordering constraint algebras $\mathcal{A} = \langle C, \top, \mathit{refine}, \mathit{sat}, \gamma \rangle$, for $c, c' \in C$, we say c' strictly refines c (denoted $c' \sqsubset_{\mathcal{A}} c$) if $c' = \mathit{refine}(c, t, u)$ for some terms t and u, and $\gamma(c') \subset \gamma(c)$. Then, we say $\mathcal{A}$ is well-founded if $\sqsubset_{\mathcal{A}}$ is.*

Down every path explored by REST, the tracked constraint is only ever refined; well-foundedness of $\mathcal{A}$ guarantees that finitely many such refinements can be strict.

We note that if the OCA describes a finite set of orderings, then it is trivially well-founded: $\subset$ is well-founded on finite sets. For example, the ordering constraint algebra for RPQOs (Sec. 4.2.1) is well-founded when the set of functions symbols $\mathcal{F}$ is finite, as there are a only a finite number of possible RPQOs over a finite set of function symbols.

▶ **Definition 21.** *A relation $t_l \rightarrow t_r$ is normalizing if it does not admit an infinite path and bounded if for each t_l it only admits finite t_r.*

▶ **Theorem 22** (Termination of REST)**.** *For any finite set of rewriting rules R, if:*

1. *$\rightarrow_{\mathcal{E}}$ is normalizing and bounded,*
2. *The refine and sat functions from $\mathcal{A}$ are decidable (always-terminating, in an implementation),*
3. *$\mathcal{A}$ is well-founded,*

then, for all terms t_0, $\mathit{REST}(\mathcal{A}, R, t_0, \mathcal{E})$ terminates.

Proof. (Sketch:) The paths constructed by REST implicitly constructs a finitely branching tree, and the four restrictions ensures that all paths down the tree are finite. This ensures that the resulting tree is finite; and thus that REST's implicit construction of the tree will terminate. ◀

Note that any deterministic, terminating external oracle function satisfies the first requirement. Having completed the formalization, we now move on to the details of our implementation.

6 Implementation of REST

We implemented REST as a standalone library, comprising 2337 lines of Haskell code (Sec. 6.1). Our implementation includes the REST algorithm, several ordering constraint algebra implementations (including RPQOs [Def. 10]) and exposes the API for implementing ordering constraint algebras (Sec. 6.2). We integrated this library into the Liquid Haskell program verifier [49] (Sec. 6.3), where we chose the task of applying *lemmas* in Liquid Haskell proofs as a suitable target problem for automation via REST.

```
-- Interface of OC Algebra
data OC C T = OC
  { top    :: C
  , refine :: C → T → T → C
  , sat    :: C → IO Bool
  }

-- Language of Logical Formulas
data LF A = LTrue   | LFalse
        | A :>: A   | A :=: A
        | LF A :∧: LF A | LF A :∨: LF A
```

```
-- Implementation of OC Algebra
rpoOC :: OC (LF 𝓕) 𝓣
rpoOC = OC LTrue refine sat where

  refine :: LF 𝓕 → 𝓣 → 𝓣 → LF 𝓕
  refine c t u =
    c :∧: rpo t u -- As in Def 10

  sat :: LF 𝓕 → IO Bool
  sat = smtSat . toSMT -- SMT Interface
```

Figure 5 The implementation of our RPQO Ordering Constraint Algebra

6.1 The REST Library

Our REST implementation is developed in Haskell and can be used directly by other Haskell projects. The library is designed modularly; for example, a client of the library can decide to use REST only for comparing terms via an OCA, without also using the proof search algorithm of Sec. 3.2. In addition, our library has a small code footprint and can be used with or without external solvers, making it ideal for integration into existing program analysis tools and theorem provers.

Furthermore, we include in the library built-in helper utilities for encoding and solving constraints on term orderings. Although the library enables integration of arbitrary solvers; it provides several built-in solvers for constraints on finite WQOs and also provides an interface for solving constraints with external SMT solvers. These utilities comprise the majority of the code in the REST library (1369 out of the 2337 lines).

Our implementation defines the OCA interface of Sec. 4.2 and provides three built-in instances for RPQOs, LPQOs (derived from the Lexicographic path ordering), and KBQOs (Sec. 4.1.1). The helper utilities included in the library enable a concise implementation of these OCAs: the three OCA implementations consist of 200 lines of code in total.

To facilitate debugging and evaluation of OCAs, the library also provides a standalone executable that produces visualizations of the rewrite paths that REST explores when using the OCA to compute the rewrites paths from a given term. Figure 4 and Figure 8 were produced using this functionality; we also note that the visualization is also capable of displaying the accumulated constraints on the ordering at each node in the tree.

We now describe the interface for defining OCAs in our REST implementation, via a presentation of the RPQO algebra in the library.

6.2 Efficient Implementations of OCAs in REST

Figure 5 presents REST's library interface for ordering constraint algebras and the implementation of RPQOs. The interface `OC` is parametric in the language of constraints `C` and the type of terms `t`. The logical formulas `LF A` describe constraints on WQOs over `A`, in the case of RPQOs, `LF` $\mathcal{F}$ tracks constraints on the underlying precedence of function symbols.

Our implementation `rpoOC` defines the initial constraints `top` to be `LTrue`, (intuitively, permitting any RPQO). The function `refine c t u` conjoins the current constraints `c` with the constraints `rpo t u`, ensuring $t \succcurlyeq u$. Finally the `sat` function converts the constraints into an equisatisfiable SMT formula, by encoding each distinct function symbol as an SMT integer variable, encoding the logical operators as their SMT equivalent, and checking for

```
{-@ example1 :: s0 : Set → { s1 : Set | IsDisjoint s0 s1 } → f : (Set → a) → { f ((s0
      \/ s1) /\ s0)  = f s0 } @-}
example1 :: Set → Set → (Set → a) → Unit
example1 s0 s1 f =
       f ((s0 \/ s1) /\ s0)         ? distribUnion s0 s1 s0
   === f ((s0 /\ s0) \/ (s1 /\ s0)) ? idemInter s0
   === f (s0 \/ (s1 /\ s0))         ? symmInter s1 s0
   === f (s0 \/ (s0 /\ s1))         -- Disjoint
   === f (s0 \/ emptySet)           ? emptyUnion s0
   === f s0
   *** QED
```

Figure 6 Liquid Haskell version of the proof from Example 1.

satisfiability of the resulting formula.

REST's interface supports arbitrary implementations for ordering constraints and is not dependent on any particular ordering, constraint language, or solver. For example, the `sat` function for RPQOs could evaluate the formulas using an alternative solver; in fact REST includes a built-in solver for this purpose, although it does not achieve as high performance as the SMT-based approach.

6.3 Integration of REST in Liquid Haskell

We used REST to automate lemma application in Liquid Haskell. Here we provide a brief overview of Liquid Haskell (Sec. 6.3.1), how REST is used to automate lemma instantiations (Sec. 6.3.2) and how it mutually interacts with the existing Liquid Haskell automation (Sec. 6.3.3).

6.3.1 Liquid Haskell and Program Lemmas

Liquid Haskell performs program verification via *refinement types* for Haskell; function types can be annotated with refinements that capture logical/value constraints about the function's parameters, return value and their relation. For example, the function `example1` in Figure 6 ports the set example of Example 1 to Liquid Haskell, without any use of REST. User-defined lemmas amount to nothing more than additional program functions, whose refinement types express the logical requirements of the lemma. The first line of the figure is special comment syntax used in Liquid Haskell to introduce refinement types; it expresses that the first parameter `s0` is unconstrained, while the second `s1` is refined in terms of `s0`: it must be some value such that `IsDisjoint s0 s1` holds. The refinement type on the (unit) return value expresses the proof goal; the body of the function provides the proof of this lemma. The proof is written in equational style; the `?` annotations specify lemmas used to justify proof steps [48]. The penultimate step requires no lemma; the verifier can discharge it based on the refinement on the `s1` parameter.

Lemmas already proven can be used in the proof of further lemmas; as is standard for program verification, care needs to be taken to avoid circular reasoning. Liquid Haskell ensures this via well-founded recursion: lemmas can only be instantiated recursively with smaller arguments.

6.3.2 REST for Automatic Lemma Application in Liquid Haskell

We apply REST to automate the application of equality lemmas in the context of Liquid Haskell. The basic idea is to extract a set of rewrite rules from a set of refinement-typed functions, each of which must have a refinement type signature of the following shape:

```
{-@ rrule :: x₁:t₁ → ... → xₙ:tₙ → {v:() | eₗ = eᵣ } @-}
```

In particular, the equality $e_l = e_r$ refinement of the (unit) return value generates potential rewrite rules to feed to REST, in both directions. Let $FV(e)$ be the free variables of e, if $FV(e_r) \subseteq FV(e_l)$ and $e_l \notin \{x_1, \ldots, x_n\}$ then $e_l \rightarrow e_r$ is generated as a rewrite rule. Symmetrically, if $FV(e_l) \subseteq FV(e_r)$ and $e_r \notin \{x_1, \ldots, x_n\}$ then $e_r \rightarrow e_l$ is generated as a rewrite rule. These rewrite rules are fed to REST along with the current terms we are trying to equate in the proof goal; any rewrites performed by REST are fed back to the context of the verifier as assumed equalities.

Since the extracted rewrite rules are defined as refinement-typed expressions, our implementation technically goes beyond simple term rewriting, since instantiations of these rules in our implementation are also refinement-type-checked; i.e., it instantiates only the rules with expressions of the proper refined type, achieving a form of conditional rewriting [27].

Selective Activation of Lemmas: Local and Global Rewrite Rules In our Liquid Haskell extension, the user can activate a rewrite rule globally or locally, using the `rewrite` and `rewriteWith` pragmas, *resp.*. For example, with the below annotations

```
{-@ rewrite global @-}
{-@ rewriteWith theorem [local] @-}
```

the rule `global` will be active when verifying every function in the current Haskell module, while the rule `local` is used only when verifying `theorem`.

Preventing Circular Reasoning Our implementation finally ensures that rewrites cannot be used to justify circular reasoning, by checking that there are no cycles induced by our `rewrite` and `rewriteWith` pragmas. For example, the below, unsound, circular dependency will be rejected with a rewrite error by our implementation.

```
{-@ rewriteWith p1 [p2] @-}
{-@ rewriteWith p2 [p1] @-}
{-@ p1, p2 :: x:Int → { x = x + 1 } @-}
p1 _ = () ; p2 = p1
```

To prevent circular dependencies, we check that the dependency graph of the rewrite rules (which are made available for proving with) has no cycles. This simple restriction is stronger than strictly necessary; a more-complex termination check could allow rewrites to be mutually justified by ensuring that recursive rewrites are applied with smaller arguments. In practice, our coarse check isn't too restrictive: because Haskell's module system enforces acyclicity of imports, rewrite rules placed in their own module can be freely referenced by importing the library.

Lemma Automation Using our implementation, the same Example 1 proven manually in Figure 6 can be alternatively proven (with all relevant rewrite rules in scope) as follows:

```
{-@ example1 :: s0 : Set → { s1 : Set | IsDisjoint s0 s1 } →  f : (Set → a) →
                { f ((s0 \/ s1) /\ s0) = f s0 } @-}
example1 s0 s1 _ = ()
```

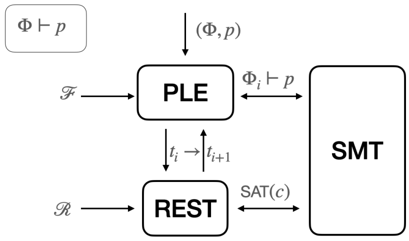

Figure 7 Interaction between PLE and REST.

The proof is fully automatic: no manual lemma calls are needed as these are all handled by REST. Integrating REST into Liquid Haskell required around 500 lines of code, mainly for surface syntax.

6.3.3 Mutual PLE and REST Interaction

Liquid Haskell includes a technique called *Proof by Logical Evaluation* (PLE) [50] for automating the expansion of terminating program function definitions. PLE expands function calls into single cases of their (possibly conditional) bodies exactly when the verifier can prove that a unique case definitely applies. This check is performed via SMT and so can condition on arbitrary logical information; in our implementation, this forms a natural complement to the term rewriting of REST, and plays the role of its external oracle (cf. Sec. 3). Since PLE is proven terminating [50], the termination of this collaboration is also guaranteed (cf. Sec. 5).

Figure 7 summarizes the mutual interaction between PLE and REST on a verification condition $\Phi \vdash p$, where Φ is an environment of assumptions. PLE also takes as input a set $\mathcal{F}$ of (provably) terminating, user-defined function definitions that it iteratively evaluates. Meanwhile, REST is provided with the rewrite rules extracted from in-scope lemmas in the program (cf. Sec. 6.3.2); these two techniques can then generate paths of equal terms including steps justified by each technique. For example, consider the following simple lemma `countPosExtra`, stating that the number of strictly positive values in `xs ++ [y]` is the number in `xs`, provided that `y <= 0`, and a lemma stating that *countPos* of two lists appended gives the same result if their orders are swapped.

```
{-@ lm :: xs : [Int] → ys : [Int] → { countPos (xs ++ ys) = countPos (ys ++ xs) } @-}

{-@ rewriteWith countPosExtra [lm] @-}
{-@ countPosExtra :: xs : [Int] → {y : Int | y <= 0 } →
                    { countPos (xs ++ [y]) = countPos xs } @-}
countPosExtra :: [Int] → Int → ()
countPosExtra _ _ = () -- proof is fully automatic!
```

The proof requires rewriting `countPos(xs ++ [y])` first via lemma `lm` (by REST), expanding the definition of `++` twice (via PLE) to give `countPos(y:xs)`, and finally one more PLE step evaluating `countPos`, using the logical fact that `y` is not positive. Note in particular that the first step requires applying an external lemma (out of scope for PLE) and the last requires SMT reasoning not expressible by term rewriting. The two techniques together allow for a fully automatic proof.

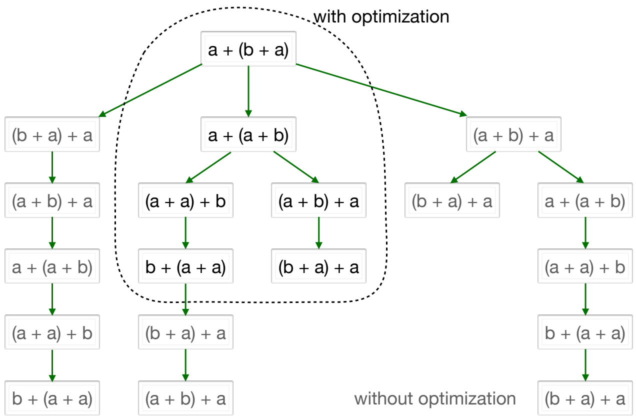

Figure 8 Associative-commutative rewrites of $a + (b + a)$ generated by REST. Paths explored by REST with the explored terms optimization are within the dashed line. Using the explored terms optimization, REST only considers each term once.

6.4 Further Optimizing the REST Algorithm

When a rewrite system is branching, REST may encounter different rewrite paths from an initial term t to an arbitrary term u. For example, in Figure 8, the term $(b + a) + a$ is explored in 5 different paths. In general, REST cannot always ignore the repeat encounters of u, as a new path from t to u may impose ordering constraints enabling more rewrites in the future. Nonetheless, reducing the number of explored paths naturally improves performance. Therefore, we optimize REST based on the following observations:

1. A term t does not need to be revisited if all of it's rewrites have already been visited.
2. If a term t was previously visited at constraints c, revisiting t at constraints c' is not necessary if c permits all orderings permitted by c', i.e., $\gamma(c') \subseteq \gamma(c)$.

To implement this optimization, REST maintains a mapping M from terms to the logical constraints c each term was explored with (initially mapping all terms to **top**). To explore a term t under logical constraints c, the algorithm checks that this term is *explorable*, formally defined by:

$$\text{explorable}(t, c) \doteq t \notin M \vee (\neg(c \Rightarrow M[t]) \wedge \exists u.(t \rightarrow_R u \wedge \text{explorable}(u, c)))$$

This predicate ensures that either this term was not explored before or it comes with weaker constraints that can derive at least one new term in the path.

After exploring a new term, REST weakens the mapping M for this term to the disjunction of the constraints under which it was newly explored and those previously mapped to in M. With this optimization, a term will appear in more than one path in the REST graph only when it can lead to different terms in the path. This optimization critically reduces the number of explored terms even for small examples: as shown in Figure 8 where 19 vertices of the REST graph shown reduced to only the 6 in the dotted region.

7 Evaluation

Our evaluation seeks to answer three research questions:

§ 7.1: How does REST compare to existing rewriting tactics?

Property	LH+	Coq	Agda	Lean	Isabelle	Zeno	Isa+
Diverge	OK	loop	loop	fail	loop	OK	OK
Plus AC	OK	loop	loop	fail	fail	OK	OK
Congruence	OK	OK	OK	OK	OK	fail	OK

Table 1 Comparison of REST with existing theorem provers. LH+ is Liquid Haskell with rewriting. The potential outcomes are **OK** when the property is proved; **loop** when no answer is returned after 300 sec; and **fail** when the property cannot be proven. Isa+ is Isabelle/HOL with Sledgehammer.

§ 7.2: How does REST compare to E-matching based axiomatization?
§ 7.3: Does REST simplify equational proofs?

We evaluate REST using the Liquid Haskell implementation described in Sec. 6. In Sec. 7.1, we compare our implementation's rewriting functionality with that of other theorem provers, with respect to the challenges mentioned in Sec. 2. In Sec. 7.2, we compare against Dafny [33] by porting Dafny's calculational proofs to Liquid Haskell, using rewriting to handle axiom instantiation. Finally, in Sec. 7.3, we port proofs from various sources into Liquid Haskell both with and without rewriting, and compare the performance and complexity of the resulting proofs.

7.1 Comparison with Other Theorem Provers

To compare REST with the rewriting functionality of other theorem provers, we developed three examples to test the five challenges described in Sec. 2 and compare our implementation to that of other solvers. We chose to evaluate against Agda [39], Coq [11], Lean [5], Isabelle/HOL [38], and Zeno [44], as they are widely known theorem provers that either support a rewrite tactic, or use rewriting internally. Agda, Lean, and Isabelle/HOL allow user-defined rewrites. In Lean and Isabelle/HOL, the tactic for applying rewrite rules multiple times is called simp; for simplification. Agda, Coq, and Isabelle/HOL's implementation of rewriting can diverge for nonterminating rewrite systems [11, 1, 38]. On the other hand, Lean enforces termination, at least to some degree, by ensuring that associative and commutative operators can only be applied according to a well-founded ordering [4]. Zeno [44] does not allow for user-defined rewrite rules, rather it generates rewrites internally based on user-provided axioms. Sledgehammer [36, 42, 41] is a powerful tactic supported by Isabelle/HOL that (on top of the built-in rewriting) dispatches proof obligations to various external provers and succeeds when any of the external provers succeed; this tactic operates under a built-in (customizable) timeout.

1. Diverge tests how the prover handles the challenges 1 and 5: restricting the rewrite system to ensure termination and integrating external oracle steps. This example encodes a single (terminating) rewrite rule $f(x) \to g(s(s(x)))$ and terminating, mutually recursive function definitions for f and g. However, the combination of the rules and function expansions can cause divergence. This test also requires a simple proof that follows directly from the function definitions.

2. Plus AC tests the challenges 2 and 3 by encoding a task that requires a permissive term ordering. This example encodes p, q, and r, user-defined natural numbers, and requires that expressions such as (p + q) + r can be rewritten into different groupings such as (r + q) + p, via associativity and commutativity rules.

3. Congruence is an additional test to ensure that the implementation of the rewrite system is permissive enough to generate the expected result. This test evaluates a basic

expected property, that the expressions $f(g(x))$ and $f(g'(x))$ can be proved equal if there exists a rewrite rule of the form $g(x) \to g'(x)$.

We present our results in Table 1. As expected, `Coq`, `Agda`, and `Isabelle/HOL` diverge on the first example, as they do not ensure termination of rewriting. `Lean` does not diverge, but it also fails to prove the theorem. Unsurprisingly, the commutativity axiom of `Plus AC` causes theorem provers that don't ensure termination of rewriting to loop. Although `Lean` ensures termination, it does not generate the necessary rewrite application in every case, because it orients associative-commutative rewriting applications according to a fixed order. With the exception of `Zeno`, all of the theorem provers tested were able to prove the necessary theorem for the final example. Our implementation succeeds on these three examples by implementing a permissive termination check based on non-strict orderings.

For this selection of simple but illustrative examples, the only tools to succeed on all cases are our implementation, and Isabelle's Sledgehammer. The latter combines a great many techniques which go beyond term rewriting. Nonetheless, we note that our novel approach provides a clear and general formal basis for incorporation with a wide variety of verifiers and reasoning techniques (due to its generic definition and formal requirements), and provides strong formal guarantees for such combinations. In particular, `REST` provides general termination and relative completeness guarantees, which Sledgehammer (via its timeout mechanism) does not.

7.2 Comparison with E-matching

To evaluate `REST` against the E-matching based approach to axiom instantiation, we compared with Dafny [33], a state-of-the-art program verifier. Dafny supports equational reasoning via calculational proofs [34] and calculation with user-defined functions [2]. We ported the calculational proofs of [34] to Liquid Haskell, using rewriting to automatically instantiate the necessary axioms.

7.2.1 List Involution

Figure 9 shows an example taken directly from Dafny [34], proving that the reverse operation on lists is an involution, i.e., $\forall \text{xs}.\textit{reverse(reverse(xs))} = \textit{xs}$. In this example, both Liquid Haskell and Dafny operate on inductively defined lists with user-defined functions $+\!\!+$ and `reverse`. The original Dafny proof goes through via the combination of a manual application of a lemma called `ReverseAppendDistrib` (stating that for all lists xs and ys, $reverse(xs +\!\!+ ys) = reverse(ys) +\!\!+ reverse(xs)$) and induction on the size of the list.

Using term rewriting as enabled by our `REST` library, Liquid Haskell is able to simplify the proof, with PLE expanding the function definitions for `reverse` and `append`, and `REST` applying the necessary equality `reverse (reverse xs ++ [x]) = reverse [x] ++ reverse (reverse xs)`.

In Dafny, a similar simplification of the calculational proof is not possible; the proof fails if the manual equality steps are simply removed. We experimented further and found that the lemma `ReverseAppendDistrib` can be alternatively encoded as a user-defined axiom which, by itself, does not appear to cause trouble for E-matching, and with this change alone the proof succeeds without the need for this single lemma call. On the other hand, the equalities must still be mentioned for the calculational proof to succeed. Perhaps surprisingly, removing these intermediate equality steps caused Dafny to stall[3]; analysis with the Axiom

[3] We include this version in App. D

```
lemma LemmaReverseTwice(xs: List)
    ensures reverse(reverse(xs)) == xs;
{
  match xs {
    case Nil =>
    case Cons(x, xrest) =>
      calc {
        reverse(reverse(xs));
        reverse(append(reverse(xrest), Cons(x, Nil)));
        { ReverseAppendDistrib(reverse(xrest), Cons(x, Nil)); }
        append(reverse(Cons(x, Nil)), reverse(reverse(xrest)));
        { LemmaReverseTwice(xrest); }
        append(reverse(Cons(x, Nil)), xrest);
        append(Cons(x, Nil), xrest);
        xs;
      }
  }
}
```

(a) Calculation-style proof in Dafny, from [34].

```
{-@ involutionP :: xs:[a] → {reverse (reverse xs) == xs } @-}
{-@ rewriteWith involutionP [distributivityP] @-}
involutionP []     = ()
involutionP (x:xs) = involutionP xs
```

(b) An equivalent proof implemented in Liquid Haskell extended with REST

Figure 9 List Involution proofs in Liquid Haskell and Dafny

Profiler [7] indicated the presence of a (rather complex) matching loop involving the axiom ReverseAppendDistrib in combination with axioms internally generated by the verifier itself. This illustrates that achieving further automation of such E-matching-based proofs is not straightforward, and can easily lead to performance difficulties due to matching loops which can be hard to predict and understand, even in this state-of-the-art verifier. By contrast, REST can automatically provide the necessary equality steps for this proof without introducing any risk of non-termination.

7.2.2 Set Properties

Figure 10 shows the Dafny and Liquid Haskell proofs for the implication $s_0 \cap s_1 = \emptyset \implies f((s_0 \cup s_1) \cap s_0) = f(s_0)$.

Dafny uses a calculational proof to show the equality $(s_0 \cup s_1) \cap s_0 = s_0$, seemingly by applying distributivity. In fact, the distributivity aspect is not relevant to the proof; rather, the set equality in the proof syntax causes Dafny to instantiate the set extensionality axiom discharging the proof. It is for this reason that Dafny requires an extra proof step to prove $f((s_0 \cup s_1) \cap s_0) = f(s_0)$, as this term does not include an equality on sets, but rather on applications of f. Dafny's set axiomatization does not include the distributivity axiom, as such an axiom could easily lead to matching loops.

Using REST, it is safe to encode arbitrary lemmas as rewrite rules, as the termination is guaranteed; in this case the distributivity lemma can be used to complete the proof (and is permitted as a rewrite rule with the precedence $\cap > \cup$).

In conclusion, we have shown that using REST to apply rewrites could be used as an

```
lemma Proof<a>(s0: set<int>, s1: set<int>, f: set<int> → a)
  requires s0 * s1 == {}
  ensures f((s0 + s1) * s0) == f(s0) {
    calc {
      (s0 + s1) * s0; (s0 * s0) + (s1 * s0);
      s0;
    }
  }
```

(a) Proof in Dafny using built-in set axiomatization

```
{-@ assume unionEmpty :: ma : Set → {v : () | ma \/ emptySet = ma } @-}
{-@ assume intersectComm :: ma : Set → mb : Set → {v : () | ma /\ mb = mb /\ ma } @-}
{-@ assume intersectSelf :: s0 : Set → { s0 /\ s0 = s0 } @-}
{-@ assume unionIntersect :: s0 : Set → s1 : Set → s2 : Set →
                          { (s0 \/ s1) /\ s2 = (s0 /\ s2) \/ (s1 /\ s2) } @-}
{-@ rwDisjoint :: s0 : Set → {s1 : Set | IsDisjoint s0 s1} → { s0 /\ s1 = emptySet } @-}

{-@ example1 :: s0 : Set → { s1 : Set | IsDisjoint s0 s1 } →  f : (Set → a) →
                { f ((s0 \/ s1) /\ s0) = f s0 } @-}
example1 s0 s1 _ = ()
```

(b) An equivalent proof implemented in Liquid Haskell, with a user-defined axiomatization of sets.

Figure 10 Set Proofs in Liquid Haskell and Dafny

alternative to E-matching based axiomatization. Furthermore, the termination guarantee of `REST` enables axioms that may give rise to matching loops to, instead, be encoded as rewrite rules.

7.3 Simplification of Equational Proofs

Finally, we evaluate how `REST` can simplify equational proofs. We chose to include the set example from [34] (described in Sec. 7.2.2), data structure proofs from [48], examples from the Liquid Haskell test suite, as well as our own case study. We developed each example in Liquid Haskell both with and without rewriting, and compared the timing and proof complexity. Each proof using rewriting was evaluated using each different ordering constraint algebras built-in to our Haskell `REST` library. The proofs in [48] were selected because they require induction, expansion of user-defined functions, and equational reasoning steps to prove properties about trees and lists. The examples from the Liquid Haskell test suite were taken to evaluate the rewriting across a range of representative proofs.

Our `DSL` case study evaluates the performance of our implementation using a larger set of rewrite rules, by verifying optimizations for a simple programming language, containing statements (i.e., print, sequence, branches, repeats and no-ops) and expressions (i.e., constants, variables, arithmetic and boolean expressions) using 23 rewrite rules. Our rewriting technique can prove the kind of equivalences used in techniques such as supercompilation [8, 52, 46], by encoding the basic equality axioms as rewrite rules and using them to prove more complicated theorems. A full list of the axioms and proved theorems are available in App. C. We note that we encoded arithmetic operations as uninterpreted SMT functions, so that the built-in arithmetic theory of the SMT does not aid proof automation.

We present our results in Table 2. By using rewriting, we were able to eliminate all but two of the non-inductive axiom instantiations, while maintaining a reasonable verification

Name	Orig.	Cut	Rules	Time				
				Orig.	RPQO	LPQO	KBQO	Fuel
Set-Dafny	4	4	5	1.11s	✓1.15s	✓1.19s	✗1.13s	✓1.22s
Set-Mono	7	7	4	1.16s	✗1.40s	✗1.41s	✓1.47s	✓1.60s
List	3	3	3	2.46s	✓3.17s	✗4.21s	✗2.24s	✓3.54s
Tree	3	3	3	1.61s	✓2.64s	✓3.40s	✓3.08s	✓3.12s
DSL	43	43	23	2.89s	✓5.46s	✗3.85s	✗4.19s	✓6.54s
LH-FingerTree	2	1	1	5.55s	✓5.60s	✓5.57s	✓5.64s	✓5.95s
LH-T1013	1	1	1	1.11s	✓1.06s	✓1.00s	✓1.02s	✓1.06s
LH-T1025	2	2	2	1.03s	✓1.05s	✓1.08s	✓1.07s	✓1.13s
LH-T1548	1	1	1	1.45s	✓1.33s	✓1.38s	✓1.32s	✓1.45s
LH-T1660	1	1	1	1.09s	✓1.12s	✓1.12s	✓1.12s	✓1.20s
LH-MapReduce	4	3	2	14.38s	✓29.50s	✓518.91s	✓28.49s	✗Timeout

Table 2 Results from simplification of proofs with rewriting. **Set-Dafny** is the set example from[34], **Set-Mono** describes a similar property. **List** and **Tree** are equational proofs from [48]. **DSL** is the program equivalence case study. The remaining proofs are from the Liquid Haskell test suite folder `tests/pos`, excluding those using only inductive or mutually inductive lemmas. **Orig.** is the number of non-inductive lemma applications in the original proof. **Cut** is the number of lemma applications that were removed by rewriting. **Rules** is the number of axioms encoded as rewrite rules. **Time (Orig.)** is verification time in seconds for the original proof. **LPQO** and **KBQO** are OCAs derived from the Lexicographic Path Ordering and Knuth-Bendix ordering respectively, and **Fuel** is an OCA allowing up to 5 rewrite applications per proof goal.

time. As expected, no ordering constraint algebra was able to complete all the proofs using rewriting; however, each proof could be verified with at least one of them.

The test cases `LH-FingerTree` and `LH-MapReduce` required manual axiom instantiations because the structure of the term did not match the rewrite rule for the axiom. `LH-MapReduce`, requires proving the identity `op (f (take n is)) (mapReduce n f op (drop n is)) = f is`. An inductive lemma application generates the background equality `mapReduce n f op (drop n is) = f (drop n is)`, and a rewrite matching the term `op (f (take n is)) (f (drop n is))` must be instantiated to complete the proof. However, since the background equality is neither a rewrite rule nor an evaluation step, the necessary term `op (f (take n is)) (f (drop n is))` never appears. Therefore, it is necessary to manually instantiate the lemma. As future work, a limited form of E-matching [12] could be used to address this issue in the general case.

In conclusion, we've shown that extending Liquid Haskell to use `REST` enables rewriting functionality not subsumed by existing theorem provers, that `REST` is effective for axiom instantiation, and that `REST` can simplify equational proofs.

8 Related Work

Theorem Provers & Rewriting Term rewriting is an effective technique to automate theorem proving [25] supported by most standard theorem provers. § 7.1 compares, by examples, our technique with `Coq`, `Agda`, `Lean`, and `Isabelle/HOL`. In short, our approach is different because it uses user-specified rewrite rules to derive, in a terminating way, equalities that strengthen the SMT-decidable verification conditions generated during program verification.

SMT Verification & Rewriting Our rewrite rules could be encoded in SMT solvers as universally quantified equations and instantiated using *E-matching* [12], i.e., a common algorithm for quantifier instantiation. E-matching might generate matching loops leading to unpredictable divergence. [32] refers to this unpredictable behavior of E-matching as the “the butterfly effect” and partially addresses it by detecting formulas that could give rise to matching loops. Our approach circumvents unpredictability by using the terminating REST algorithm to instantiate the rewrite rules outside of the SMT solver.

Z3 [13] and CVC4 [6] are state-of-the-art SMT solvers; both support theory-specific rewrite rules internally. Recent work [40] enables user-provided rewrite rules to be added to CVC4. However, using the SMT solver as a rewrite engine offers little control over rewrite rule instantiation, which is necessary for ensuring termination.

Rewriting in Haskell Haskell itself has used various notions of rewriting for program verification. GHC supports the RULES pragma with which the user can specify unchecked, quantified expression equalities that are used at compile time for program optimization. [10] proposes Inspection Testing as a way to check such rewrite rules using runtime execution and metaprogramming, while [22] prove rewrite rules via metaprogramming and user-provided hints. In a work closely related to ours, Zeno [44] is using rewriting, induction, and further heuristics to provide lemma discovery and fully automatic proof generation of inductive properties. Unlike our approach, Zeno’s syntax is restricted (e.g., it does not allow for existentials) and it does not allow for user-provided hints when automation fails. HALO [51] enables Haskell verification by converting Haskell into logic and using an SMT solver to verify user-defined formulas. However, this approach relies on SMT quantifiers to encode user functions, thus the solver can diverge and verification becomes unpredictable.

Termination of Rewriting and Runtime Termination Checking Early work on proving termination of rewriting using simplification orderings is described in [15]. More recent work involves dependency pairs [3] and applying the size-change termination principle [31] in the context of rewriting [47]. Tools like AProVE [24] and NaTT [54] can statically prove the termination of rewriting.

In contrast, REST is not focused on statically proving termination of rewriting; rather it uses a well-founded ordering to ensure termination at runtime. This approach enables integration of arbitrary external oracles to produce rewrite applications, as a static analysis is not possible in principle. Furthermore, our approach enables nonterminating rewriting systems to be useful: REST will still apply certain rewrite rules to satisfy a proof obligation, even if the rewrite rules themselves cannot be statically shown to terminate.

We choose to use a well-quasi-ordering [30] because it enables rewriting to terms that are not strictly decreasing in a simplification ordering. WQOs are commonly used in online termination checking [35], especially for program optimization techniques such as supercompilation [9].

Equality Saturation In our implementation, REST passes equalities to the SMT environment, ultimately used for *equality saturation* via an E-graph data structure [20]. Equality saturation has also been used for supercompilation[46]. REST does not currently exploit equality saturation (unless indirectly via its oracle). However, as future work we might explore local usage of efficient E-graph implementations. (e.g., [53]) for caching the equivalence classes generated via rewrite applications.

Associative-Commutative Rewriting Traditionally, enforcing a strict ordering on terms prevents the application of rewrite rules for associativity or commutativity (AC); this problem motivates REST’s use of well-quasi orders. However, another solution is to omit the rules and instead perform the substitution step of rewriting modulo AC. Termination of the resulting system can be proved using an AC ordering [17]; the essential requirement is that the ordering also respects AC, such that $t > u$ implies $t' > u'$ for all terms t' AC-equivalent to t and u' AC-equivalent to u.

REST's use of well-quasi-orderings enables AC axioms to be encoded as rewrite rules, guaranteeing completeness if the AC-equivalence class of a term is a subset of the equivalence class induced by the ordering. This is a significant practical benefit as it does not require REST to identify AC symbols and treat them differently for unification.

However, we note that treating AC axioms as rewrite rules can lead to an explosion in the number of terms obtained via rewriting. As future work, it could be possible to extend REST to support AC rewriting and unification in order to reduce the number of explicitly instantiated terms.

9 Conclusion

We have presented REST, a novel approach to rewriting that uses an online termination check that simultaneously considers entire families of term orderings via a newly introduced Ordering Constraint Algebra. We defined our algebra on well-quasi orderings that are more permissive than standard simplification orderings, and demonstrated how to derive well-quasi orderings from well-known simplification orderings. In addition, we proved correctness, relative completeness, and (online) termination of our algorithm. Our REST approach is designed, via a generic core algorithm and the pluggable abstraction of our OCAs, to be simple to (re-)implement and adapt to different programming languages or efficient implementations of term ordering families. We demonstrated this by writing an implementation of REST as a small Haskell library suitable for integration with existing verification tools. To evaluate REST we used our library to extend Liquid Haskell, and showed that the resulting system compares well with existing rewriting techniques, it can be used as an alternative to E-matching based axiomatizations approaches without risking non-termination, and can substantially simplify proofs requiring equational reasoning steps.

References

1 Agda Developers. *The Agda Language Reference, version 2.6.1*, 2020. Available electronically at https://agda.readthedocs.io/en/v2.6.1/language/index.html.

2 Nada Amin, K Rustan M Leino, and Tiark Rompf. Computing with an smt solver. In *International Conference on Tests and Proofs*, pages 20–35. Springer, 2014.

3 Thomas Arts and Jürgen Giesl. Termination of term rewriting using dependency pairs. *Theoretical Computer Science*, 236(1):133–178, April 2000. URL: http://www.sciencedirect.com/science/article/pii/S0304397599002078, doi:10.1016/S0304-3975(99)00207-8.

4 Jeremy Avigad, Leonardo de Moura, and Soonho Kong. *Theorem Proving in Lean, Release 3.20.0*, September 2020. p 73. URL: https://leanprover.github.io/theorem_proving_in_lean/theorem_proving_in_lean.pdf.

5 Jeremy Avigad, Gabriel Ebner, and Sebastian Ullrich. *The Lean Reference Manual, Release 3.3.0*, 2018. URL: https://leanprover.github.io/reference/lean_reference.pdf.

6 Clark Barrett, Christopher L. Conway, Morgan Deters, Liana Hadarean, Dejan Jovanovi'c, Tim King, Andrew Reynolds, and Cesare Tinelli. CVC4. In Ganesh Gopalakrishnan and Shaz Qadeer, editors, *Proceedings of the 23rd International Conference on Computer Aided Verification (CAV '11)*, volume 6806 of *Lecture Notes in Computer Science*, pages 171–177. Springer, July 2011. Snowbird, Utah. URL: http://www.cs.stanford.edu/~barrett/pubs/BCD+11.pdf.

7 N. Becker, P. Müller, and A. J. Summers. The axiom profiler: Understanding and debugging smt quantifier instantiations. In *Tools and Algorithms for the Construction and Analysis of Systems (TACAS) 2019*, LNCS, pages 99–116. Springer-Verlag, 2019.

8 Maximilian Bolingbroke and Simon Peyton Jones. Supercompilation by evaluation. *SIGPLAN Not.*, 45(11):135146, September 2010. doi:10.1145/2088456.1863540.

9 Maximilian Bolingbroke, Simon Peyton Jones, and Dimitrios Vytiniotis. Termination combinators forever. In *Proceedings of the 4th ACM symposium on Haskell*, pages 23–34, 2011.

10 Joachim Breitner. A promise checked is a promise kept: inspection testing. In Nicolas Wu, editor, *Proceedings of the 11th ACM SIGPLAN International Symposium on Haskell, Haskell@ICFP 2018, St. Louis, MO, USA, September 27-17, 2018*, pages 14–25. ACM, 2018. doi:10.1145/3242744.3242748.

11 The Coq Development Team. *The Coq Reference Manual, version 8.11.2*, 2020. Available electronically at http://coq.inria.fr/refman.

12 Leonardo de Moura and Nikolaj Bjørner. Efficient e-matching for smt solvers. In Frank Pfenning, editor, *Automated Deduction – CADE-21*, pages 183–198, Berlin, Heidelberg, 2007. Springer Berlin Heidelberg.

13 Leonardo De Moura and Nikolaj Bjørner. Z3: An efficient smt solver. In *International conference on Tools and Algorithms for the Construction and Analysis of Systems*, pages 337–340. Springer, 2008.

14 Nachum Dershowitz. A note on simplification orderings. *Information Processing Letters*, 9(5):212–215, 1979. doi:https://doi.org/10.1016/0020-0190(79)90071-1.

15 Nachum Dershowitz. Orderings for term-rewriting systems. *Theoretical computer science*, 17(3):279–301, 1982.

16 Nachum Dershowitz. Termination of rewriting. *Journal of symbolic computation*, 3(1-2):69–115, 1987.

17 Nachum Dershowitz, Jieh Hsiang, N Alan Josephson, and David A Plaisted. Associative-commutative rewriting. In *IJCAI*, pages 940–944, 1983.

18 Nachum Dershowitz and Zohar Manna. Proving termination with multiset orderings. In Hermann A. Maurer, editor, *Automata, Languages and Programming*, Lecture Notes in Computer Science, pages 188–202, Berlin, Heidelberg, 1979. Springer. doi:10.1007/3-540-09510-1_15.

19 David Detlefs, Greg Nelson, and James B. Saxe. Simplify: A theorem prover for program checking. *J. ACM*, 52(3):365–473, May 2005. URL: http://doi.acm.org/10.1145/1066100.1066102, doi:10.1145/1066100.1066102.

20 David Detlefs, Greg Nelson, and James B Saxe. Simplify: a theorem prover for program checking. *Journal of the ACM (JACM)*, 52(3):365–473, 2005.

21 Claire Dross, Sylvain Conchon, Johannes Kanig, and Andrei Paskevich. Adding decision procedures to smt solvers using axioms with triggers. *Journal of Automated Reasoning*, 56(4):387–457, 2016.

22 Andrew Farmer, Neil Sculthorpe, and Andy Gill. Reasoning with the hermit: Tool support for equational reasoning on ghc core programs. In *Proceedings of the 2015 ACM SIGPLAN Symposium on Haskell*, Haskell 15, page 2334, New York, NY, USA, 2015. Association for Computing Machinery. doi:10.1145/2804302.2804303.

23 Jean-Christophe Filliâtre and Andrei Paskevich. Why3—where programs meet provers. In Matthias Felleisen and Philippa Gardner, editors, *Programming Languages and Systems (ESOP)*, volume 7792 of *Lecture Notes in Computer Science*, pages 125–128. Springer, 2013.

24 Jürgen Giesl, Cornelius Aschermann, Marc Brockschmidt, Fabian Emmes, Florian Frohn, Carsten Fuhs, Jera Hensel, Carsten Otto, Martin Plücker, Peter Schneider-Kamp, et al. Analyzing program termination and complexity automatically with aprove. *Journal of Automated Reasoning*, 58(1):3–31, 2017.

25 Jieh Hsiang, Hélène Kirchner, Pierre Lescanne, and Michaël Rusinowitch. The term rewriting approach to automated theorem proving. *The Journal of Logic Programming*, 14(1):71–99,

October 1992. URL: http://www.sciencedirect.com/science/article/pii/0743106692900477, doi:10.1016/0743-1066(92)90047-7.

26 Gerard Huet. Confluent reductions: Abstract properties and applications to term rewriting systems. In *Proceedings of the 18th Annual Symposium on Foundations of Computer Science*, SFCS '77, page 3045, USA, 1977. IEEE Computer Society. doi:10.1109/SFCS.1977.9.

27 Stéphane Kaplan. Conditional rewrite rules. *Theoretical Computer Science*, 33(2):175–193, 1984. URL: https://www.sciencedirect.com/science/article/pii/0304397584900872, doi:https://doi.org/10.1016/0304-3975(84)90087-2.

28 J. W. Klop. *Term Rewriting Systems*, page 1116. Oxford University Press, Inc., USA, 1993.

29 Donald E Knuth and Peter B Bendix. Simple word problems in universal algebras. In *Automation of Reasoning*, pages 342–376. Springer, 1983.

30 Joseph B Kruskal. The theory of well-quasi-ordering: A frequently discovered concept. *Journal of Combinatorial Theory, Series A*, 13(3):297–305, November 1972. URL: http://www.sciencedirect.com/science/article/pii/0097316572900635, doi:10.1016/0097-3165(72)90063-5.

31 Chin Soon Lee, Neil D. Jones, and Amir M. Ben-Amram. The size-change principle for program termination. In *Proceedings of the 28th ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages*, POPL 01, page 8192, New York, NY, USA, 2001. Association for Computing Machinery. doi:10.1145/360204.360210.

32 K. R. M. Leino and Clément Pit-Claudel. Trigger Selection Strategies to Stabilize Program Verifiers. In Swarat Chaudhuri and Azadeh Farzan, editors, *Computer Aided Verification*, Lecture Notes in Computer Science, pages 361–381, Cham, 2016. Springer International Publishing. doi:10.1007/978-3-319-41528-4_20.

33 K. Rustan M. Leino. Dafny: An automatic program verifier for functional correctness. In *Proceedings of the 16th International Conference on Logic for Programming, Artificial Intelligence, and Reasoning*, LPAR10, page 348370, Berlin, Heidelberg, 2010. Springer-Verlag.

34 K Rustan M Leino and Nadia Polikarpova. Verified calculations. In *Working Conference on Verified Software: Theories, Tools, and Experiments*, pages 170–190. Springer, 2013.

35 Michael Leuschel. Homeomorphic Embedding for Online Termination of Symbolic Methods. In Gerhard Goos, Juris Hartmanis, Jan van Leeuwen, Torben Æ. Mogensen, David A. Schmidt, and I. Hal Sudborough, editors, *The Essence of Computation*, volume 2566, pages 379–403. Springer Berlin Heidelberg, Berlin, Heidelberg, 2002. Series Title: Lecture Notes in Computer Science. URL: http://link.springer.com/10.1007/3-540-36377-7_17, doi:10.1007/3-540-36377-7_17.

36 Jia Meng and Lawrence C Paulson. Translating higher-order clauses to first-order clauses. *Journal of Automated Reasoning*, 40(1):35–60, 2008.

37 P. Müller, M. Schwerhoff, and A. J. Summers. Viper: A verification infrastructure for permission-based reasoning. In *VMCAI*, 2016.

38 Tobias Nipkow, Markus Wenzel, and Lawrence C. Paulson. *Isabelle/HOL: A Proof Assistant for Higher-Order Logic*. Springer-Verlag, 2020.

39 Ulf Norell. Dependently typed programming in agda. In *Proceedings of the 6th International Conference on Advanced Functional Programming*, AFP08, page 230266, Berlin, Heidelberg, 2008. Springer-Verlag.

40 Andres Nötzli, Andrew Reynolds, Haniel Barbosa, Aina Niemetz, Mathias Preiner, Clark Barrett, and Cesare Tinelli. Syntax-guided rewrite rule enumeration for smt solvers. In Mikoláš Janota and Inês Lynce, editors, *Theory and Applications of Satisfiability Testing – SAT 2019*, pages 279–297, Cham, 2019. Springer International Publishing.

41 Lawrence C Paulson and Kong Woei Susanto. Source-level proof reconstruction for interactive theorem proving. In *International Conference on Theorem Proving in Higher Order Logics*, pages 232–245. Springer, 2007.

42 Lawrence C Paulsson and Jasmin C Blanchette. Three years of experience with sledgehammer, a practical link between automatic and interactive theorem provers. In *Proceedings of the 8th*

International Workshop on the Implementation of Logics (IWIL-2010), Yogyakarta, Indonesia. EPiC, volume 2, 2012.

43 Julien Signoles, Pascal Cuoq, Florent Kirchner, Nikolai Kosmatov, Virgile Prevosto, and Boris Yakobowski. Frama-c: a software analysis perspective. volume 27, 10 2012. doi:10.1007/s00165-014-0326-7.

44 William Sonnex, Sophia Drossopoulou, and Susan Eisenbach. Zeno: An automated prover for properties of recursive data structures. In Cormac Flanagan and Barbara König, editors, *Tools and Algorithms for the Construction and Analysis of Systems*, pages 407–421, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg.

45 Nikhil Swamy, Catalin Hritcu, Chantal Keller, Aseem Rastogi, Antoine Delignat-Lavaud, Simon Forest, Karthikeyan Bhargavan, Cédric Fournet, Pierre-Yves Strub, Markulf Kohlweiss, Jean-Karim Zinzindohoué, and Santiago Zanella-Béguelin. Dependent types and multi-monadic effects in F*. In *43rd ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages (POPL)*, pages 256–270. ACM, January 2016. URL: https://www.fstar-lang.org/papers/mumon/.

46 Ross Tate, Michael Stepp, Zachary Tatlock, and Sorin Lerner. Equality saturation: a new approach to optimization. In *POPL '09: Proceedings of the 36th annual ACM SIGPLAN-SIGACT symposium on Principles of Programming Languages*, pages 264–276, New York, NY, USA, 2009. ACM. URL: http://www.cs.cornell.edu/~ross/publications/eqsat/, doi:http://doi.acm.org/10.1145/1480881.1480915.

47 René Thiemann and Jürgen Giesl. Size-Change Termination for Term Rewriting. volume 2706, pages 264–278, March 2007. doi:10.1007/3-540-44881-0_19.

48 Niki Vazou, Joachim Breitner, Rose Kunkel, David Van Horn, and Graham Hutton. Theorem proving for all: equational reasoning in liquid Haskell (functional pearl). In *Proceedings of the 11th ACM SIGPLAN International Symposium on Haskell*, Haskell 2018, pages 132–144, St. Louis, MO, USA, September 2018. Association for Computing Machinery. doi:10.1145/3242744.3242756.

49 Niki Vazou, Eric L. Seidel, Ranjit Jhala, Dimitrios Vytiniotis, and Simon Peyton-Jones. Refinement types for haskell. In *Proceedings of the 19th ACM SIGPLAN International Conference on Functional Programming*, ICFP 14, page 269282, New York, NY, USA, 2014. Association for Computing Machinery. doi:10.1145/2628136.2628161.

50 Niki Vazou, Anish Tondwalkar, Vikraman Choudhury, Ryan G. Scott, Ryan R. Newton, Philip Wadler, and Ranjit Jhala. Refinement reflection: Complete verification with smt. *Proc. ACM Program. Lang.*, 2(POPL), December 2017. doi:10.1145/3158141.

51 Dimitrios Vytiniotis, Simon Peyton Jones, Koen Claessen, and Dan Rosén. Halo: Haskell to logic through denotational semantics. In *Proceedings of the 40th annual ACM SIGPLAN-SIGACT symposium on Principles of programming languages*, pages 431–442, 2013.

52 Philip Wadler. Deforestation: transforming programs to eliminate trees. *Theoretical Computer Science*, 73(2):231 – 248, 1990. URL: http://www.sciencedirect.com/science/article/pii/030439759090147A, doi:https://doi.org/10.1016/0304-3975(90)90147-A.

53 Max Willsey, Chandrakana Nandi, Yisu Remy Wang, Oliver Flatt, Zachary Tatlock, and Pavel Panchekha. Egg: Fast and extensible equality saturation. *Proceedings of the ACM on Programming Languages*, 5(POPL):1–29, 2021.

54 Akihisa Yamada, Keiichirou Kusakari, and Toshiki Sakabe. Nagoya termination tool. In *Rewriting and Typed Lambda Calculi*, pages 466–475. Springer, 2014.

A Metaproperty Proofs

▶ **REST Invariant 1** (Path Invariant). *For any execution of* REST$(\mathcal{A}, R, t_0, \mathcal{E})$*, at the start of any iteration of the main loop, for each* $(ts, c) \in p$*, the list* ts *is a path of* $R + \mathcal{E}$ *starting from* t_0*.*

Proof. By induction on the loop iterations of the algorithm. p is initialized with the single element $([t_0], c)$. $[t_0]$ is a valid path of $R+\mathcal{E}$, because it only contains a single term; clearly this path also starts with t_0.

At each loop iteration, new elements are potentially pushed to p. Suppose the path ts is popped from p at the beginning of the loop. The element to be pushed is a pair $(ts +\!\!+ [t'], c)$ where $last(ts) \rightarrow_{R+\mathcal{E}} t'$. This exactly satisfies the inductive hypothesis: if ts is a path of $R+\mathcal{E}$, then $ts +\!\!+ [t']$ is also a path of $R+\mathcal{E}$. Furthermore, this operation preserves the head of the list: t_0 is still the first element. ◀

▶ **Theorem 18** (Completeness w.r.t. $\mathcal{E}$). *For all* R*,* u*, and* t_0*, if* $t_0 \rightarrow^*_{\mathcal{E}} u$*, then* $u \in$ REST$(\mathcal{A}, R, t_0, \mathcal{E})$*.*

Proof. The proof goes by induction on the number of steps of the path.

Assume the path has n steps: $t_0 \rightarrow_{\mathcal{E}} t_1 \rightarrow_{\mathcal{E}} \ldots \rightarrow_{\mathcal{E}} t_{n-1} \rightarrow_{\mathcal{E}} t_n \equiv u$.

For the base case, $n = 0$ and $u \equiv t_0$. Since p is initialized with $([t_0], \top)$, by the Invariant 2, $t \in$ REST$(\mathcal{A}, R, t_0, \mathcal{E})$.

For the inductive case, assume that $t_0 \rightarrow^*_{\mathcal{E}} t_{n-1} \rightarrow_{\mathcal{E}} t_n$. By inductive hypothesis, $t_{n-1} \in$ REST$(\mathcal{A}, R, t_0, \mathcal{E})$. When t_{n-1} was added in the result, it was the last element of a path ts that was popped from the stack p. Since $t_{n-1} \rightarrow_{\mathcal{E}} t_n$, we split cases on whether or not $t_n \in ts$. If $t_n \in ts$, then by Invariant 2 $t_n \in$ REST$(\mathcal{A}, R, t_0, \mathcal{E})$. Otherwise, $(ts +\!\!+ [t_n], c)$ will be pushed into p and, again, by Invariant 2 it will appear in the output. ◀

▶ **Theorem 19** (Relative Completeness). *For all* R*,* u*, and* t_0*, if* $t_0 \rightarrow^*_{R+\mathcal{E}} u$ *and there exists an ordering* $\succcurlyeq \in \gamma(\top)$ *that orients the path justifying* $t_0 \rightarrow^*_{R+\mathcal{E}} u$*, then* $u \in$ REST$(\mathcal{A}, R, t_0, \mathcal{E})$*.*

First, we observe the (somewhat standard) property that if any path justifies $t_0 \rightarrow^*_{R+\mathcal{E}} u$, there is a *duplicate-free variant* of such path (intuitively, obtained by cutting out all subpaths leading from a term to itself).

Below, we prove that if $t_0 \rightarrow^*_{R+\mathcal{E}} u$ and the ordering $\succcurlyeq$ orients the path, then a duplicate-free variant path ts belongs in the stack p with some constraints c and $\succcurlyeq \in \gamma(c)$.

▶ **REST Invariant 3.** *For any execution of* REST$(\mathcal{A}, R, t_0, \mathcal{E})$*, if* $t_0 \rightarrow^*_{R+\mathcal{E}} t_n$ *and* $\succcurlyeq \in \gamma(\top)$ *is an ordering that orients* $t_0 \rightarrow^*_{R+\mathcal{E}} u$*, then at some iteration of the main loop, a duplicate-free variant path* ts *of this path is stored in* p*, with some ordering constraints* c *and* $\succcurlyeq \in \gamma(c)$*.*

Proof. The proof goes by strong induction on the length $n+1$ of the path justifying $t_0 \rightarrow^*_{R+\mathcal{E}} t_n$.

First, consider the case $n = 0$, where the path is $[t_0]$ and the constraints $\top$. $([t_0], \top) \in p$ by initialization and trivially $\succcurlyeq \in \gamma(\top)$.

Otherwise, when $n > 0$, assume that $t_0 \rightarrow^*_{R+\mathcal{E}} t_{n-1} \rightarrow_{R+\mathcal{E}} t_n$. If there are any duplicate terms in this path, a duplicate-free variant exists of shorter length, and we can conclude by our induction hypothesis. Otherwise, consider this path with the last element t_n removed. Being already duplicate-free, by our induction hypothesis we must have that, at some iteration of our main loop, this path is contained in p along with a constraint c_{n-1} such that $\succcurlyeq \in \gamma(c_{n-1})$. By the assumption that $\succcurlyeq$ orients the original path, in particular we must have $t_{n-1} \succcurlyeq t_n$, and so, by Def. 14, $\succcurlyeq \in \gamma(\mathit{refine}(c_{n-1}, t, t'))$ and therefore $\mathit{refine}(c_{n-1}, t, t')$

is satisfiable. Therefore, the original path will be pushed to p with this constraint in this loop iteration. ◀

Proof. The proof is similar to Theorem 18, but now we need to also show that the relation that orients the path satisfies all the ordering constraints generated by the respective REST path. By Invariant 3, at some iteration of the main loop, there must be some path ending in u contained in p. Then, by Invariant 2 it follows that all the elements of the path, thus also u, belong in the result.

◀

▶ **Theorem 22** (Termination of REST). *For any finite set of rewriting rules R, if:*

1. *$\to_{\mathcal{E}}$ is normalizing and bounded,*
2. *The refine and sat functions from $\mathcal{A}$ are decidable (always-terminating, in an implementation),*
3. *$\mathcal{A}$ is well-founded,*

then, for all terms t_0, REST$(\mathcal{A}, R, t_0, \mathcal{E})$ terminates.

Proof. At every iteration of REST, a path with length n is popped off the stack and due to Requirement 1, and the fact that only a finite number of new terms can be generated by single applications of the rules R to an arbitrary term, a finite number of paths with length $n+1$ is pushed on. Therefore, REST implicitly builds (via its set of paths p) a *finitely-branching* tree starting from t_0. For REST to not terminate, there must be an infinite path down the tree (note that Requirement 2 eliminates the possibility that the operations called from the ordering constraint algebra cause non-termination).

Consider an arbitrary path down the tree explored by REST, represented by the (ts, c) pairs iteratively generated. Firstly, due to the first condition in the `foreach` of REST (cf. Figure 3), this path will remain duplicate-free. By Requirement 3, at only finitely many steps is the constraint tracked *strictly* refined. Consider then, the postfix of the path after the last time that this happens; at every step, the constraint c remains identical. The normalization assumption (Requirement 1) of $\mathcal{E}$ entails that this path contains no infinite sequence of steps all justified by $\mathcal{E}$. However, for each step justified instead by a rewriting step from R, the additional condition $sat(c)$ must hold; by Def. 14 this means that there is some $\succcurlyeq \in \gamma(c)$ which orients all of these steps. As $\succcurlyeq$ must be an instance of a term ordering family (Def. 4), it is a thin well-quasi-order. Therefore $\succcurlyeq$ can only orient a finite number of steps, and the path down the tree must be finite.

Since every path in the finitely-branching tree explored is finite, the algorithm (always) terminates. ◀

B Proofs on Orderings

▶ **Lemma 23.** *If $T \succcurlyeq_{M(X)} U$, then $T \succcurlyeq_{M(X)} U'$ for all $U' \subset U$.*

Proof. It is sufficient to show that $T \succcurlyeq_{M(X)} U$ implies $T \succcurlyeq_{M(X)} (U - u')$, for any $u' \in U$, since the subset can be obtained by removing a finite number of elements. That is, if U' was obtained by removing elements $u_1, \ldots, u_n$ from U, we can show that $T \succcurlyeq_{M(X)} (U \setminus \{u_1\})$ implies $T \succcurlyeq_{M(X)} (U \setminus \{u_1, u_2\})$ and so on.

The proof goes by induction on the size of T and case analysis on $T \succcurlyeq_{M(X)} U$.

For case one there are no u' in U, so the proof holds vacuously.

For case two, we have either $u = u'$ or $u \neq u'$. If $u = u'$, a proof of $T \succcurlyeq_{M(X)} (U - u)$ can be made by modifying the proof of $(T - t) \succcurlyeq_{M(X)} (U - u)$. The base case of that proof must be of the form $T' \succcurlyeq_{M(X)} \emptyset$. We modify the base case to be $(T' + t) \succcurlyeq_{M(X)} \emptyset$. Each recursive case is also modified to replace T' with $(T' + t)$, yielding $(T' + t) \succcurlyeq_{M(X)} (U - u)$ $= T \succcurlyeq_{M(X)} (U - u)$, as required. The proof that $T \succcurlyeq_{M(X)} (U - u')$ for all other $u' \in U$ is obtained by induction. By the inductive hypothesis, we have $(T - t) \succcurlyeq_{M(X)} (U - u - u')$, since $u \neq u'$, we also have $u \in (U - u')$. Therefore applying case two we get $T \succcurlyeq_{M(X)} (U - u')$.

For case three, we have either $u' < t$ or $u' \not< t$. If $u' < t$, then the proof $(T - t) \succcurlyeq_{M(X)}$ $(U \setminus \{u \in U \mid u < t\})$ is also a proof of $(T - t) \succcurlyeq_{M(X)} ((U - u') \setminus \{u \in U \mid u < t\})$, thus we obtain obtain the proof directly. The proof for all other $u' \in U$ is obtained by induction. By the inductive hypothesis we have $(T - t) \succcurlyeq_{M(X)} ((U \setminus \{u \in U \mid u < t\}) - u')$. Then, applying the same top-level proof yields $T \succcurlyeq_{M(X)} (U - u')$, since u' is not in the set $\{u \in U \mid u < t\}$.

◀

▶ **Lemma 24.** *If $\succcurlyeq_X$ is a quasi-order, then the multiset extension $\succcurlyeq_{M(X)}$ is also a quasi-order.*

Proof. To show that $\succcurlyeq_{M(X)}$ is a quasi-order, we define a single-step version $\geqslant_{mul}$, and show that $T \succcurlyeq_{M(X)} U$ if and only if $T \geqslant_{mul*} U$, where $\geqslant_{mul*}$ is the reflexive transitive closure of U.

We define $\geqslant_{mul}$ as:

1. For all elements t, u if $t \in T$ and $u \approx t$, then $T \geqslant_{mul} (T - t + u)$
2. For all elements $t \in T$ and finite multisets U, if $t > u$ for all $u \in U$, then $T \geqslant_{mul}$ $((T - t) \cup U)$

First, observe that $\geqslant_{mul*}$ is monotonic with respect to multiset union: for all multisets T, U, and V, $T \geqslant_{mul*} U$ implies $(T \cup V) \geqslant_{mul*} (U \cup V)$.

The reflexive case is given by $T \cup V = T \cup V$; we show the transitive case by showing there is a correspondence for each single-step. The proof for each case assumes an arbitrary multiset V.

In case one we must show for all $t, u \in T$, $T \geqslant_{mul*} (T - t + u)$ implies $(T \cup V) \geqslant_{mul*}$ $((T - t + u) \cup V)$. t and u are also in $T \cup V$, therefore we have $(T \cup V) \geqslant_{mul*} ((T \cup V) - t + u)$. We have $(T \cup V) - t + u = (T - t + u) \cup V$, giving us the desired result. Case two is similar: $t \in T$ implies $t \in (T \cup V)$, and $((T \cup V) - t) \cup U = ((T - t) \cup U) \cup V$ for all U, V.

Now we show the if direction by case analysis.

Case 1: $U = \emptyset$.

If $T = \emptyset$, then we have $T \geqslant_{mul*} U$ via reflexivity. Otherwise we can select an arbitrary t

to remove from T, and by definition of $\geqslant_{mul}$ we have $T \geqslant_{mul} ((T-t) \cup \emptyset)$. Then by induction on the size of T we have $((T-t)\cup\emptyset) \geqslant_{mul*} \emptyset$. Then $T \geqslant_{mul} ((T-t)\cup\emptyset) \geqslant_{mul*} \emptyset$, as required.

Case 2: $t \in T \land u \in U \land t \approx u \land (T-t) \geqslant_{mul} (U-u)$.
Let $T' = T - t$ and $U' = U - u$. Then we have $(T'+t) \geqslant_{mul} (T'+u)$ by definition and $T' \succcurlyeq_{M(X)} U'$ implies $T'+u \geqslant_{mul*} U'+u$ via the inductive hypothesis and monotonicity. Thus $T = (T'+t) \geqslant_{mul} (T'+u) \geqslant_{mul*} (U'+u) = U$ as required.

Case 3: $t \in T \land (T-t) \geqslant_{mul} (U \setminus \{u \in U \mid u < t\})$
Partition U into two sets U_1 and U_2 where $U_1 = \{u \in U \mid u \not< t\}$ and $U_2 = \{u \in U \mid u < t\}$. By definition we have $U = U_1 \cup U_2$. As before $T' = T - t$. Then we have $(T'+t) \geqslant_{mul} (T' \cup U_2)$. $T' \succcurlyeq_{M(X)} U_1$ implies $(T' \cup U_2) \geqslant_{mul*} (U_1 \cup U_2)$ via monotonicity and induction. Thus $T = (T'+t) \geqslant_{mul} (T' \cup U_2) \geqslant_{mul*} (U_1 \cup U_2) = U$ as required.

Now the only-if direction. First we have that $\succcurlyeq_{M(X)}$ is reflexive via induction on size with base case $T = U = \emptyset$ handled by case 1, and recursive case by case 2, similar to above, we remove an arbitrary t from T. Now we show how to handle one or more steps from $\geqslant_{mul}$ in a single step of $\succcurlyeq_{M(X)}$.

The key observation is that all elements u of U, have exactly one "responsible" element t in T that justifies $T \geqslant_{mul*} U$: we must have either $t > u$ or $t \approx u$ (in which case t is uniquely responsible for u and no other elements of U). To prove $T \succcurlyeq_{M(X)} U$, for each t in T, we recursively build a tuple (T', U', p) where T', and U' are multisets and p is the proof that $T' \succcurlyeq_{M(X)} U'$. The tuple is initialized to $(\emptyset, \emptyset, U = \emptyset)$.

For each t uniquely responsible for one u, we update the tuple to $(T'+t, U'+u, t \in T \land u \in U \land t \approx u \land p)$. The new proof state is valid because by induction we have p being a proof of $T' \succcurlyeq_{M(X)} U'$, as required.

Now consider each $t \in T$ where t justified some multiset U''. By induction, we have a proof of $T' \succcurlyeq_{M(X)} U'$; we need a proof that $T' \succcurlyeq_{M(X)} ((U' \cup U'') \setminus \{u \in (U' \cup U'') \mid u < t\})$. Since we have $t > u$ for all $u \in U''$, this simplifies to: $T' \succcurlyeq_{M(X)} (U' \setminus \{u \in U' \mid u < t\})$, which we can obtain via the hypothesis $T' \succcurlyeq_{M(X)} U'$ and lemma 23.

◀

▶ **Lemma 25.** *If $\succcurlyeq_X$ is a well-quasi-order, the strict part of it's multiset extension defined as $t >_{M(X)} u$ if $t \succcurlyeq_{M(X)} u$ and $u \not\succcurlyeq_{M(X)} t$ is a well-founded order.*

Proof. This proof operates on the single-step relation defined in 24. Proving the well-founded property is done by showing that an infinite descent in $>_{M(X)}$ would correspond to an infinite descent in the underlying ordering.

Now, consider a tree built from an infinite path $T_1, T_2, \ldots$ of multisets related by $\succcurlyeq_{M(X)}$. With the exception of special nodes $\top$ and $\bot$, each node in the tree represents an element in a multiset, and the vertices connect the elements to the smaller ones they were replaced with via an application of $\geqslant_{mul}$. Crucially, every edge represents an descent in a well-founded order.

The tree is constructed as follows: let $\top$ be the root of the tree, and let the elements of T_1 be the children of $\top$. Then, for each T_i in the infinite list, it was either obtained by replacing some element in T_{i-1} with a same-sized element, or by removing some element t and replacing it with a finite number of smaller elements *ts*.

In the former case, the tree is not modified.

In the latter case, if $ts = \emptyset$, add a single child $\bot$ to the t in the tree. Otherwise, let ts be the children of t.

Now, we note that the case one of $\geqslant_{mul}$ is symmetric. Therefore, each pair of terms related by $>_{M(X)}$ must correspond to at least one step in case two of $\geqslant_{mul}$, Therefore in an infinite path of terms related by $>_{M(X)}$ contains an infinite number of applications of case two in $\geqslant_{mul}$.

Therefore, an infinite number of vertices will be added to the tree. Since the tree is finitely branching, it must have an infinitely descending path. However, this infinitely descending path would correspond to an infinite descent in the underlying ordering, contradicting that hypothesis that $\succcurlyeq_X$ is a WQO. ◀

▶ **Lemma 26.** *If $\succcurlyeq_{\mathcal{F}}$ is a total quasi-ordering, then $\succcurlyeq_{rpo}$ is a quasi-simplification ordering.*

Proof. We must show that $\succcurlyeq_{rpo}$ is a quasi-ordering, i.e it is reflexive and transitive; and also that it satisfies the replacement, subterm, and deletion properties.

Reflexivity occurs via case 3 and 24.Replacement and deletion follow from case 3 of RPO and the definition of the multiset ordering.

To prove the subterm property, we show a slightly stronger property: for all terms $t = f(t_1, \ldots, t_m)$ and (not necessarily immediate) subterms $u = g(u_1, \ldots, u_n)$, $t >_{rpo} u$. The proof goes by induction on the term size, where terms are bigger than their subterms, and by case analysis on the relationship between f and g. Because $\succcurlyeq_{\mathcal{F}}$ is total, we have either $f >_{\mathcal{F}} g$, $f \approx g$, or $g >_{\mathcal{F}} f$.

If $f >_{\mathcal{F}} g$, then to get $t \succcurlyeq_{rpo} u$ we must show $\{t\} >_{M(rpo)} \{u_1, \ldots, u_n\}$. Via induction, we have $t >_{rpo} u_i$ for all $1 \leqslant i \leqslant n$, as each u_i is a subterm of u. To show $u \not\succcurlyeq_{rpo} t$, observe that we need $\{u_1, \ldots, u_n\} \succcurlyeq_{M(rpo)} \{t\}$. This is impossible via the inductive hypothesis and the definition of $\succcurlyeq_{M(rpo)}$: we already have $t >_{rpo} u_i$ for all u_i.

If $f \approx g$, then we must show $\{t_1, \ldots, t_m\} >_{M(rpo)} \{u_1, \ldots, u_n\}$. If u is a direct subterm of t, then $u = t_i$ for some i. By the inductive hypothesis we have $t_i \approx u >_{rpo} u_j$ for all u_j, which implies $\{t_1, \ldots, t_m\} >_{M(rpo)} \{u_1, \ldots, u_n\}$. If u is a nested subterm, then we have some $t_i >_{rpo} u_j$ for all u_j via the induction hypothesis: all u_j are subterms of t_i.

If $g >_{\mathcal{F}} f$, to get $t \succcurlyeq_{rpo} u$ then we must show $\{t_1, \ldots, t_m\} \succcurlyeq_{M(rpo)} \{u\}$. If u was a direct subterm, then $t_i = u$ gives us the desired result; otherwise we have $t_i >_{rpo} u$ via the inductive hypothesis. To show $u \not\succcurlyeq_{rpo} t$, observe that showing $u \succcurlyeq_{rpo} t$ would require $\{u\} >_{M(rpo)} \{t_1, \ldots, t_m\}$. However we already have some $t_i \approx u$, which prevents this possibility.

Transitivity is also proven via induction on size. Assume we have $s = f(s_1, \ldots, s_m) \succcurlyeq_{rpo} t = g(t_1, \ldots, t_n)$ and $t \succcurlyeq_{rpo} u = h(u_1, \ldots, u_p)$. We proceed to show $s \succcurlyeq_{rpo} u$ by for each relationship between f, g, and h.

1. $f >_{\mathcal{F}} g >_{\mathcal{F}} h$, or $f >_{\mathcal{F}} g > h$: Via transitivity of $>_{\mathcal{F}}$ we have $f >_{\mathcal{F}} h$, therefore we must show $\{s\} >_{M(rpo)} \{u_1, \ldots, u_p\}$. $\{s\} \succcurlyeq_{M(rpo)} \{t\}$ follows from our assumption $s \succcurlyeq_{rpo} t$, and $\{t\} >_{M(rpo)} \{u_1, \ldots, u_p\}$ follows from $t \succcurlyeq_{rpo} u$. By the inductive hypothesis, we have $s \succcurlyeq_{rpo} t \succcurlyeq_{rpo} u_i$ for all u_i, and therefore $\{s\} \succcurlyeq_{M(rpo)} \{t\} >_{M(rpo)} \{u_1, \ldots, u_p\}$.
2. $h >_{\mathcal{F}} g$: There must exist some subterm t_i such that $t_i \succcurlyeq_{rpo} u$. Therefore we have $s \succcurlyeq_{rpo} t_i$ and $t_i \succcurlyeq_{rpo} u$, the inductive hypothesis gives us $s \succcurlyeq_{rpo} t_i \succcurlyeq_{rpo} u$.
3. $g >_{\mathcal{F}} f$: There must exist some subterm s_i such that $s_i \succcurlyeq_{rpo} t$. As above, using the induction hypothesis allows us to show $s_i \succcurlyeq_{rpo} u$, by the subterm property we have $s \succcurlyeq_{rpo} s_i$. We show $s \succcurlyeq_{rpo} u$ by the definition of $\succcurlyeq_{rpo}$.
4. $f \approx g \approx h$. We clearly have $f \approx h$, we need to show $\{s_1, \ldots, s_m\} \succcurlyeq_{M(rpo)} \{u_1, \ldots, u_p\}$, which we have via 24.

◀

▶ **Theorem 27.** *If $\succcurlyeq_{\mathcal{F}}$ is a total WQO, then $\succcurlyeq_{rpo}$ is a WQO.*

Proof. To show that $\succcurlyeq_{rpo}$ is WQO, via the well-foundedness theorem of Dershowitz [15], which states that a quasi-simplification ordering $\geqslant'$ is WQO if there exists a well-quasi ordering $\geqslant$ such that $f \geqslant g$ implies $f(t_1, \ldots, t_n) \geqslant' g(t_1, \ldots, t_n)$.

By 26 we have that $\succcurlyeq_{rpo}$ is a quasi-simplification ordering, and there exists an ordering over function symbols to satisfy the condition of the well-foundedness theorem: namely the underlying order $\succcurlyeq_{\mathcal{F}}$ from which $\succcurlyeq_{rpo}$ is constructed.

◀

▶ **Theorem 28.** *If $\succcurlyeq_{\mathcal{F}}$ is a total WQO, then $\succcurlyeq_{rpo}$ is thin*

Proof. We show that for any term $t = f(t_1, \ldots, t_m)$, the set of terms $\{u \mid t \approx u = g(u_1, \ldots, u_m)\}$ is finite.

If $t \approx u$, then we must have $t \succcurlyeq_{rpo} u$ and $u \succcurlyeq_{rpo} t$. Assume we have $t \succcurlyeq_{rpo} u$.

First, we show that if $f > g$ then $u \not\succcurlyeq_{rpo} t$. Assume $u \succcurlyeq_{rpo} t$, then there must have some u_i such that $u_i \succcurlyeq_{rpo} t$. But via the subterm property, we have $u >_{rpo} u_i \succcurlyeq_{rpo} t$, contradicting $t \succcurlyeq_{rpo} u$.

Likewise, if $g > f$, then there is some $t_i \succcurlyeq_{rpo} u$. Then $t >_{rpo} t_i \succcurlyeq_{rpo} u$. Therefore we also have $u \not\succcurlyeq_{rpo} t$.

Therefore, $t \approx u$ only if $f \approx g$. Since there are only a finite number of function symbols, then to show thinness we must show that only a finite number of multisets $\{u_1, \ldots, u_n\}$ such that $\{t_1, \ldots, t_m\} \succcurlyeq_{M(rpo)} \{u_1, \ldots, u_n\}$ and $\{u_1, \ldots, u_n\} \succcurlyeq_{M(rpo)} \{t_1, \ldots, t_m\}$. If $\{t_1, \ldots, t_m\} = \emptyset$, then the only such set is $\emptyset$. Otherwise, only such multisets are those where $\{u_1, \ldots, u_n\}$ is obtained from $\{t_1, \ldots, t_m\}$ by removing zero or more terms t_i and replacing them the same number of terms u_j where $t_i \approx u_j$. If $\{t_1, \ldots, t_m\} \succcurlyeq_{M(rpo)} \{u_1, \ldots, u_n\}$ was justified by removing t_i from $\{t_1, \ldots, t_m\}$ and removing smaller terms $\{u' \mid u' < t_i\}$ from $\{u_1, \ldots, u_n\}$, then we would have $\{t_1, \ldots, t_m\} >_{M(rpo)} \{u_1, \ldots, u_n\}$: this corresponds to the irreflexive single-step operation shown to form a well-founded order in lemma 25.

Since the multisets contain a finite number of elements, and each term only has a finite number of equivalent terms (by induction on term size), there are only a finite number of such multisets. ◀

C Basic Equalities and Proved Theorems in the Program Equivalence Case Study

	Name	Formula
1.	addDist	$(x * y) + (z * y) = (x + z) * y$
2.	subDist	$(x * y) - (z * y) = (x - z) * y$
3.	times2Plus	$x * 2 = x + x$
4.	plus0	$x + 0 = x$
5.	mul0	$x * 0 = 0$
6.	mul1	$x * 1 = x$
7.	subSelf	$x - x = 0$
8.	divSelf	$x/x = 1$
9.	subAdd	$x - y = x + (-y)$
10.	mulSym	$e * e' = e' * e$
11.	addSym	$e + e' = e' + e$
12.	mulAssoc	$(x * y) * z = x * (y * z)$
13.	addAssoc	$(x + y) + z = x + (y + z)$
14.	ifT	`if True then lhs else rhs = lhs`
15.	ifF	`if False then lhs else rhs = rhs`
16.	seqNop	`seq lhs nop = lhs`
17.	seqNop'	`seq nop rhs = rhs`
18.	repeatNop	`repeat 0 body = nop`
19.	repeatN1	`repeat (S n) body = seq body (repeat n body)`
20.	ifJoin	`if c1 then (if c2 then op else nop) else nop` `= if (c1 and c2) then op else nop`
21.	mapFusion	`map g (map f xs) = map (g . f) xs`
22.	foldMap	`(foldr f e) . (map g) = foldr (f . g) e`
23.	foldFusion	$\forall$ `x y . h (f x y) = f' x (h y)` $\Longrightarrow$ `h . (foldr f e) xs = foldr f' (h e) xs`

Table 3 Basic Equality Axioms used in our Program Equivalence Case Study

Formula	Rewrites
$(-(x+x)) + (x+x) = 0$	7, 11, 9
$(x*2)*2 = (x+x+x+x)$	3, 13
$(x*y) + (y*x) = (x*2*y)$	3, 10, 12
$(x*y) + (y*z) - ((x+z)*y) = 0$	1, 7, 10
$(x*y) - (0*y) = x*y$	2, 9, 7, 4
$x*(1-(x/x)) = 0$	5, 7, 8
$x*1 = x+0$	4, 6
`if true then (seq nop hw) else nop = hw`	17, 14
`repeat (S (S Z)) hw = seq hw hw`	16, 18, 19
`if True then (if False then hw else nop) else nop`	
` = if (True and False) then hw else nop`	20
`map p1 (map p2 list) = map p3 list`	21
`( (foldr add 0) .  (map p1)) list = foldr addP1 0 list`	22
`double .  (foldr add 0) list = foldr twicePlus 0 list`	23

Table 4 Theorems Proved via Rewriting using the Basic Equality axioms in 3

1476

D Dafny Matching Loop Example

```
datatype List = Nil | Cons(head: int, tail: List)

function append(xs: List, ys: List): List
{
    match xs
    case Nil => ys
    case Cons(x, xrest) => Cons(x, append(xrest, ys))
}

function reverse(xs: List): List
{
    match xs
    case Nil => Nil
    case Cons(x, xrest) => append(reverse(xrest), Cons(x, Nil))
}

lemma AppendNil(xs: List)
    ensures append(xs, Nil) == xs; {}

lemma AppendAssoc(xs: List, ys: List, zs: List)
  ensures append(xs, append(ys, zs)) == append(append(xs, ys), zs); {}

lemma ReverseAppendDistrib(xs: List, ys: List)
  ensures reverse(append(xs, ys)) == append(reverse(ys), reverse(xs));
{
    forall xs : List {AppendNil(xs);}
    forall xs : List, ys: List, zs: List {AppendAssoc(xs, ys, zs);}
}

lemma ReverseInvolution(xs: List)
    ensures reverse(reverse(xs)) == xs;
{
    // Axiom definition inserted here
    { forall (xs, ys) { ReverseAppendDistrib(xs, ys); } }

    match xs {
      case Nil =>
      case Cons(x, xrest) =>
        calc { // Equational reasoning steps removed here
            reverse(reverse(xs));
            {ReverseInvolution(xrest);}
            xs;
         }
    }
}
```

Figure 11 A version of the reverse involution proof (Figure 9) from [34] with intermediate equality steps removed. Attempting to verify this code causes a matching loop when using Dafny version 3.3.0 and Z3 version 4.8.5