Noname manuscript No.
(will be inserted by the editor)

Refutation-Based Synthesis in SMT

**Andrew Reynolds · Viktor Kuncak ·
Cesare Tinelli · Clark Barrett ·
Morgan Deters**

Received: date / Accepted: date

Abstract We introduce the first program synthesis engine implemented inside an SMT solver. We present an approach that extracts solution functions from unsatisfiability proofs of the negated form of synthesis conjectures. We also discuss novel counterexample-guided techniques for quantifier instantiation that we use to make finding such proofs practically feasible. A particularly important class of specifications are single-invocation properties, for which we present a dedicated algorithm. To support syntax restrictions on generated solutions, our approach can transform a solution found without restrictions into the desired syntactic form. As an alternative, we show how to use evaluation function axioms to embed syntactic restrictions into constraints over algebraic datatypes, and then use an algebraic datatype decision procedure to drive synthesis. Our experimental evaluation on syntax-guided synthesis benchmarks shows that our implementation in the CVC4 SMT solver is competitive with state-of-the-art tools for synthesis.

Keywords Synthesis · Satisfiability Modulo Theories · Automated Deduction

Andrew Reynolds
École Polytechnique Fédérale de Lausanne (EPFL), Switzerland
E-mail: andrew.reynolds@epfl.ch

Viktor Kuncak
École Polytechnique Fédérale de Lausanne (EPFL), Switzerland
E-mail: viktor.kuncak@epfl.ch

Cesare Tinelli
Department of Computer Science, The University of Iowa
E-mail: cesare-tinelli@uiowa.edu

Clark Barrett
Department of Computer Science, New York University
E-mail: barrett@cs.nyu.edu

Morgan Deters
Department of Computer Science, New York University

1 Introduction

The synthesis of functions that meet a given specification is a long-standing fundamental goal that has received great attention recently. This functionality directly applies to the synthesis of functional programs [22,23] but also translates to imperative programs through techniques that include bounding input space, verification condition generation, and invariant discovery [38–40]. Function synthesis is also an important subtask in the synthesis of protocols and reactive systems, especially when these systems are infinite-state [4, 35]. The SyGuS format and competition [2,3,30], inspired by the success of the SMT-LIB and SMT-COMP efforts [6], has significantly improved and simplified the process of rigorously comparing different solvers on synthesis problems.

The connection between synthesis and theorem proving was established already in early work on the subject [16,25]. It is notable that early research [25] found that the capabilities of theorem provers were the main bottleneck for synthesis. Taking lessons from automated software verification, recent work on synthesis has made use of advances in theorem proving, particularly in SAT and SMT solvers. However, that work avoids formulating the overall synthesis task as a theorem proving problem directly. Instead, existing work typically builds custom loops outside of an SMT or SAT solver, often using numerous variants of counterexample-guided synthesis. A typical role of the SMT solver has been to validate candidate solutions and provide counterexamples that guide subsequent search, although approaches such as symbolic term exploration [20] also use an SMT solver to explore a representation of the space of solutions. In existing approaches, SMT solvers thus receive a large number of separate queries, with limited communication between these different steps.

In this paper, which is an extended and improved version of [31], we revisit the formulation of the overall synthesis task as a theorem proving problem. We observe that SMT solvers already have some of the key functionality for synthesis; we show how to improve existing algorithms and introduce new ones to make SMT-based synthesis competitive. Specifically, we do the following.

- We show how to formulate an important class of synthesis problems as the problem of disproving universally quantified formulas and how to synthesize functions automatically from selected instances of these formulas.
- We present counterexample-guided techniques for quantifier instantiation, which are crucial to obtain competitive performance on synthesis tasks.
- We discuss techniques to simplify the synthesized functions to help ensure that they are small and adhere to specified syntactic requirements.
- We show how to encode syntactic restrictions using theories of algebraic datatypes and axiomatizable evaluation functions.
- We show that for an important class of single-invocation properties, the synthesis of functions from relations, the implementation of our approach in CVC4 significantly outperforms leading tools from the SyGuS competition.

1.1 Preliminaries.

Since synthesis involves finding (and so proving the existence of) functions, we use notions from many-sorted *second-order* logic to define the general problem.

Signatures We fix a large enough set $\mathbf{S}$ of *sort symbols* and an (infix) equality predicate $\approx$ of type $\sigma \times \sigma$ for each $\sigma \in \mathbf{S}$, which we always interpret as the identity relation over (the set denoted by) σ. For every non-empty sort sequence $\boldsymbol{\sigma} \in \mathbf{S}^+$ with $\boldsymbol{\sigma} = \sigma_1 \cdots \sigma_n \sigma$, we fix an infinite set $\mathbf{X}_{\boldsymbol{\sigma}}$ of *variables* $x^{\sigma_1 \cdots \sigma_n \sigma}$ *of type* $\sigma_1 \times \cdots \times \sigma_n \to \sigma$. For each sort σ we identity the type $() \to \sigma$ with σ and call it a *first-order type*. We assume the sets $\mathbf{X}_{\boldsymbol{\sigma}}$ are pairwise disjoint and let $\mathbf{X}$ be their union. A *signature* Σ consists of a set $\Sigma^{\mathrm{s}} \subseteq \mathbf{S}$ of sort symbols and a set Σ^{f} of *function symbols* $f^{\sigma_1 \cdots \sigma_n \sigma}$ *of type* $\sigma_1 \times \cdots \times \sigma_n \to \sigma$, where $n \geq 0$ and $\sigma_1, \ldots, \sigma_n, \sigma \in \Sigma^{\mathrm{s}}$. When n above is 0, we call f a *constant symbol*. Note that we consider only *first-order* function symbols, that is, functions symbols whose input and output types are all first-order types. We drop the sort superscript from variables or function symbols when it is clear from context or unimportant. We assume that, unless stated otherwise, signatures always include a Boolean sort Bool and constants tt and ff of type Bool (respectively, for true and false). The *union* $\Sigma_1 \cup \Sigma_2$ of a signature Σ_1 and a signature Σ_2 is the signature Σ such that $\Sigma^{\mathrm{s}} = \Sigma_1^{\mathrm{s}} \cup \Sigma^{\mathrm{s}}$ and $\Sigma^{\mathrm{f}} = \Sigma_1^{\mathrm{f}} \cup \Sigma^{\mathrm{f}}$. A signature Σ_1 is a *subsignature* of a signature Σ if $\Sigma = \Sigma_1 \cup \Sigma_2$ for some signature Σ_2.

Term and formulas Given a many-sorted signature Σ together with quantifiers and lambda abstractions $\lambda x_1^{\sigma_1} \ldots x_n^{\sigma_n}\, t$, the notion of well-sorted (Σ-)term, atom, literal, clause, and formula with variables in $\mathbf{X}$ are defined as usual in second-order logic. Note that all atoms have the form $s \approx t$. Having $\approx$ as the only predicate symbol causes no loss of generality since we can model other predicate symbols as function symbols with return sort Bool. We will, however, often write just t in place of the atom $t \approx \mathsf{tt}$, to simplify the notation. A Σ-term/formula is *ground* if it has no variables, and it is *first-order* if it has only *first-order variables*, that is, variables of first-order type. Free and bound occurrences of a variable in a formula are also defined as usual. A *(Σ-)sentence* is a (Σ-)formula with no free variables.

When $\boldsymbol{x} = (x_1, \ldots, x_n)$ is a tuple of variables and Q is either $\forall$ or $\exists$, we write $Q\boldsymbol{x}\, \varphi$ as an abbreviation of $Qx_1 \cdots Qx_n\, \varphi$. If s is a Σ-term or formula and $\boldsymbol{x} = (x_1, \ldots, x_n)$ has no repeated variables, we write $s[\boldsymbol{x}]$ to denote that all of s's free variables are from $\boldsymbol{x}$; if $\boldsymbol{t} = (t_1, \ldots, t_n)$ is a term tuple, we write $s[\boldsymbol{t}]$ for the term or formula obtained from s by simultaneously replacing, for all $i = 1, \ldots, n$, every occurrence of x_i in s by t_i. When convenient, we will treat tuples like $\boldsymbol{x}$ and $\boldsymbol{t}$ as the set of their elements.

Interpretations A *Σ-interpretation $\mathcal{I}$* maps: each $\sigma \in \Sigma^{\mathrm{s}}$ to a non-empty set $\sigma^{\mathcal{I}}$, the *domain* of σ in $\mathcal{I}$, with $\mathsf{Bool}^{\mathcal{I}} = \{\mathsf{tt}, \mathsf{ff}\}$, $\mathsf{ff}^{\mathcal{I}} = \mathsf{ff}$ and $\mathsf{tt}^{\mathcal{I}} = \mathsf{tt}$; each $u^{\sigma_1 \cdots \sigma_n \sigma} \in \mathbf{X} \cup \Sigma^{\mathrm{f}}$ to a total function $u^{\mathcal{I}} : \sigma_1^{\mathcal{I}} \times \cdots \times \sigma_n^{\mathcal{I}} \to \sigma^{\mathcal{I}}$ when $n > 0$ and to an element of $\sigma^{\mathcal{I}}$ when $n = 0$. If $x_1^{\sigma_1}, \ldots, x_n^{\sigma_n}$ are distinct variables and $e_1, \ldots, e_n$ are domain elements with $e_1 \in \sigma_1^{\mathcal{I}}, \ldots, e_n \in \sigma_n^{\mathcal{I}}$, we denote by $\mathcal{I}[x_1 \mapsto e_1, \ldots, x_n \mapsto e_n]$ the Σ-interpretation that maps each x_i to e_i and is otherwise identical to $\mathcal{I}$. The interpretation $\mathcal{I}$ induces a mapping from terms t of sort σ to elements $t^{\mathcal{I}}$ of $\sigma^{\mathcal{I}}$ as expected. A satisfiability relation between Σ-interpretations and Σ-formulas or sets thereof is defined inductively as usual. A satisfying interpretation for a Σ-formula φ *models (or is a model of)* φ. A Σ-interpretation $\mathcal{I}$ is *term-generated* if each of its domain elements is denoted by a ground Σ-term, that is, if for all $\sigma \in \Sigma^{\mathrm{s}}$ and all $e \in \sigma^{\mathcal{I}}$ there is a ground Σ-term t such that $e = t^{\mathcal{I}}$.

If Ω is a subsignature of a signature Σ, the *Ω-reductt* of a Σ-interpretation $\mathcal{I}$ is an Ω-interpretation that interprets its sort and function symbols exactly as $\mathcal{I}$.

An *isomorphism* from a Σ-interpretation $\mathcal{I}$ to a Σ-interpretation $\mathcal{J}$ is a family of bijective mappings $\{h_\sigma : \sigma^{\mathcal{I}} \to \sigma^{\mathcal{J}} \mid \sigma \in \Sigma^{\mathrm{s}}\}$ such that for every $f \in \Sigma^{\mathrm{f}}$ of type $\sigma_1 \times \cdots \times \sigma_n \to \sigma_{n+1}$, and every $(v_1, \ldots, v_n) \in \sigma_1^{\mathcal{I}} \times \cdots \times \sigma_n^{\mathcal{I}}$,

$$h_{\sigma_{n+1}}(f^{\mathcal{I}}(v_1, \ldots, v_n)) = f^{\mathcal{J}}(h_{\sigma_1}(v_1), \ldots, h_{\sigma_n}(v_n))$$

Two Σ-interpretations $\mathcal{I}$ and $\mathcal{J}$ are *isomorphic*, written $\mathcal{I} \equiv \mathcal{J}$, if there is an isomorphism from one to the other. Among other things, it is possible to show that $\equiv$ is an equivalence relation and that two isomorphic interpretations satisfy exactly the same Σ-sentences.

Theories A *theory* is a pair $T = (\Sigma, \mathbf{I})$ where Σ is a signature and $\mathbf{I}$ is a non-empty class of Σ-interpretations, the *models* of T, that is closed under isomorphism.[1] A Σ-formula φ is *T-satisfiable* (resp., *T-unsatisfiable*) if it is satisfied by some (resp., no) interpretation in $\mathbf{I}$. A formula φ is *T-valid*, written $\models_T \varphi$, if every model of T is a model of φ. A set Γ of formulas *T-entails* a Σ-formula φ, written $\Gamma \models_T \varphi$, if every interpretation in $\mathbf{I}$ that satisfies all formulas in Γ satisfies φ as well. Two Σ-formulas are *T-equivalent* if the T-entail each other. Two Σ-terms s and t are *T-equivalent* if $s \approx t$ is T-valid.

If T_1 is a Σ_1-theory and T_2 is a Σ_2-theories, the *union* $T_1 \cup T_2$ of T_1 and T_2, when it exists, is the $\Sigma_1 \cup \Sigma_2$ theory whose set of models consists of all the $(\Sigma_1 \cup \Sigma_2)$-interpretations whose Σ_i-reduct is a model of T_i for $i = 1, 2$.[2]

The *theory of a Σ-structure* $\mathcal{I}$ is the theory $T = (\Sigma, \mathbf{I})$ where $\mathbf{I}$ consists of all the Σ-structures that are isomorphic to $\mathcal{I}$.

Sometimes we will extend the signature Σ of a theory T with *free* or *Skolem* constants, (first-order) constant symbols that do not occur in Σ. The set of models is extended correspondingly as follows. If k is one of the new constants and has type σ, for every model $\mathcal{I}$ of T and every $v \in \sigma^{\mathcal{I}}$, the extended theory has a model that interprets k as v and is otherwise identical to $\mathcal{I}$. It is not difficult to see that, for satisfiability purposes, free constants in a formula effectively behave like free variables.

2 Synthesis with SMT Solvers

We are interested in synthesizing computable functions automatically from formal logical specifications stating properties of these functions. As we show later, under the right conditions, we can formulate a version of the synthesis problem in *first-order logic* alone, which allows us to tackle the problem using SMT solvers.

We consider the synthesis problem in the context of some theory T of signature Σ that allows us to provide the function's specification as a Σ-formula. Specifically, we consider *synthesis conjectures* expressed as (well-sorted) formulas of the form

$$\exists f^{\sigma_1 \cdots \sigma_n \sigma}\, \forall x_1^{\sigma_1} \cdots \forall x_n^{\sigma_n}\, P[f, x_1, \ldots, x_n] \tag{1}$$

[1] That is, every Σ-interpretation isomorphic to an interpretation in $\mathbf{I}$ is also in $\mathbf{I}$.

[2] For $T_1 \cup T_2$ to exist there must be a model $\mathcal{I}_1$ of T_1 and a model $\mathcal{I}_2$ of T_2 that agree on the interpretation they give to the sort and function symbols shared by Σ_1 and Σ_2.

or $\exists f\, \forall \boldsymbol{x}\, P[f, \boldsymbol{x}]$, for short, where the second-order variable f represents the function to be synthesized and P is a Σ-formula encoding properties that f must satisfy for all possible values of the input tuple $\boldsymbol{x} = (x_1, \ldots, x_n)$. In this setting, finding a witness for this satisfiability problem amounts to finding a function of type $\sigma_1 \times \cdots \times \sigma_n \to \sigma$ in some model of T that satisfies $\forall \boldsymbol{x}\, P[f, \boldsymbol{x}]$. Since we are interested in automatic synthesis, we restrict ourselves here to methods that search over a subspace S of solutions representable syntactically as Σ-terms. We will say then that a synthesis conjecture is *solvable* if it has a syntactic solution in S.

In this paper we present two approaches that work with classes $\mathbf{L}$ of synthesis conjectures $\exists f\, \forall \boldsymbol{x}\, P[f, \boldsymbol{x}]$ and Σ-theories T where the T-validity of formulas of the form $\forall \boldsymbol{x}\, P[\lambda \boldsymbol{x}\, t[\boldsymbol{x}], \boldsymbol{x}]$ is decidable. In both approaches, we solve the synthesis conjecture by relying on quantifier-instantiation techniques to produce a first-order Σ-term $t[\boldsymbol{x}]$ of sort σ such that $\forall \boldsymbol{x}\, P[\lambda \boldsymbol{x}\, t, \boldsymbol{x}]$ is T-valid. When this t is found, the synthesized function is precisely $\lambda \boldsymbol{x}\, t$.

In principle, under the right assumptions on T, to determine the solvability of the conjecture $\exists f\, \forall \boldsymbol{x}\, P[f, \boldsymbol{x}]$ an SMT solver supporting the theory T can consider the satisfiability of the (open) formula $\forall \boldsymbol{x}\, P[f, \boldsymbol{x}]$ by treating f as an uninterpreted function symbol. This sort of Skolemization is not usually a problem for SMT solvers since many of them can process formulas with uninterpreted symbols. The real challenge is the universal quantification over $\boldsymbol{x}$ because it requires the solver to construct internally (a finite representation of) an interpretation of f that is guaranteed to satisfy $P[f, \boldsymbol{x}]$ for every possible value of $\boldsymbol{x}$ [15, 33].

More traditional SMT solver designs to handle universally quantified formulas have focused on instantiation-based methods to show *un*satisfiability. They generate ground instances of those formulas until a refutation is found at the ground level [13]. While these techniques are incomplete in general, they have been shown to be quite effective in practice [27, 34]. For this reason, we advocate approaches to synthesis geared toward establishing the *unsatisfiability of the negation* of the synthesis conjecture:

$$\forall f\, \exists \boldsymbol{x}\, \neg P[f, \boldsymbol{x}] \tag{2}$$

We show in this paper how a syntactic solution $\lambda \boldsymbol{x}\, t$ for (1) can be constructed from a refutation of (2), as opposed to being extracted from the valuation of f in a model of $\forall \boldsymbol{x}\, P[f, \boldsymbol{x}]$.

Two synthesis methods. Proving (2) unsatisfiable poses its own challenge to current SMT solvers, namely, dealing with the second-order universal quantification of f. To our knowledge, no SMT solvers so far have had direct support for higher-order quantification. In the following, however, we describe two specialized methods to refute negated synthesis conjectures like (2) that build on existing capabilities of these solvers.

The first method applies to a restricted, but fairly common, case of synthesis problems $\exists f\, \forall \boldsymbol{x}\, P[f, \boldsymbol{x}]$ where every occurrence of f in P is in terms of the form $f(\boldsymbol{x})$. In this case, we can express the problem in the first-order form $\forall \boldsymbol{x}\, \exists y\, Q[\boldsymbol{x}, y]$ and then tackle its negation using appropriate quantifier instantiation techniques.

The second method applies to theories T with term-generated interpretations. This class of theories includes the majority of theories of interest in synthesis such as (versions of) the theories of integer, rational and floating point arithmetic, bit vectors, strings, constructible arrays, algebraic datatypes, finite sets, and their combinations.[3] The method is well suited for the *syntax-guided synthesis* paradigm [2, 3] where the synthesis conjecture is accompanied by an explicit syntactic restriction on the space of possible solutions. It is based on encoding the syntax of terms as first-order values. We use a deep embedding into an extension of the background theory T with a theory of algebraic data types, encoding the restrictions of a syntax-guided synthesis problem.

For the rest of the paper, we fix a Σ-theory T and a class $\mathbf{P}$ of quantifier-free Σ-formulas $P[f, \boldsymbol{x}]$ such that the T-satifiabilty of formulas of the form $(\neg)P[\lambda\boldsymbol{x}\, t, \boldsymbol{x}]$ with t a Σ-term is decidable. We will consider synthesis conjectures for this theory from the set $\mathbf{L} := \{\exists f\, \forall\boldsymbol{x}\, P[f, \boldsymbol{x}] \mid P \in \mathbf{P}\}$.

Our canonical example of T throughout the paper will be the theory of integers. The signature of this theory contains only two sorts Int and Bool and the usual operators: the numerals, $+$, $*$, unary and binary $-$, $<$, and $\leq$ (with the last two of type $\mathsf{Int} \times \mathsf{Int} \to \mathsf{Bool}$). Let $\mathcal{I}$ be the Σ-interpretation that interprets the sort Int as the integers and interprets the various operators as expected. It is easy to see that this interpretation is term-generated. The models of T are all the Σ-interpretations that are isomorphic to $\mathcal{I}$. For this theory, $\mathbf{P}$ is the class of all Boolean combinations of linear equations and inequations. Based on results originally by Presburger, one can easily argue that the T-satisfiability of formulas of the form $(\neg)P[\lambda\boldsymbol{x}\, t, \boldsymbol{x}]$ where $P \in \mathbf{P}$ and t is a linear term is decidable. One can do that with current SMT solvers after eliminating from $P[\lambda\boldsymbol{x}\, t, \boldsymbol{x}]$ all occurrences of $\lambda\boldsymbol{x}\, t$ by β-reduction.

3 Refutation-Based Synthesis

When axiomatizing properties of a desired function f of type $\sigma_1 \times \cdots \times \sigma_n \to \sigma$, a particularly well-behaved class are *single-invocation properties* (see, e.g., [17]). These properties include, in particular, standard function contracts, so they can be used to synthesize a function implementation given its postcondition as a relation between the arguments and the result of the function. This is also the form of the specification for synthesis problems considered in complete functional synthesis [21–23]. Note that, in our case, we aim to prove that the output exists for all inputs, as opposed to, more generally, computing the set of inputs for which the output exists.

A *single-invocation property* is any formula of the form $Q[\boldsymbol{x}, f(\boldsymbol{x})]$ obtained by replacing y with $f(\boldsymbol{x})$ in a quantifier-free formula $Q[\boldsymbol{x}, y]$ not containing f. Note that the only occurrences of f in $Q[\boldsymbol{x}, f(\boldsymbol{x})]$ are in subterms of the form $f(\boldsymbol{x})$ with the *same* tuple $\boldsymbol{x}$ of *pairwise distinct* variables.[4] The conjecture $\exists f\, \forall\boldsymbol{x}\, Q[\boldsymbol{x}, f(\boldsymbol{x})]$ is logically equivalent to the *first-order* formula

$$\forall\boldsymbol{x}\, \exists y\, Q[\boldsymbol{x}, y] \quad . \tag{3}$$

The equivalence is easy to seem by observing that the conjecture is the Skolemized version of (3). By the semantics of $\forall$ and $\exists$, finding a model $\mathcal{I}$ for either formula

[3] These are versions that either do not have function symbols denoting partial functions or make those functions total in some way.

[4] An example of a property that is *not* single-invocation is $\forall x_1\, x_2\, f(x_1, x_2) \approx f(x_2, x_1)$.

amounts (under the axioms of choice) to finding a function $h : \sigma_1^{\mathcal{I}} \times \cdots \times \sigma_n^{\mathcal{I}} \to \sigma^{\mathcal{I}}$ such that for all $\boldsymbol{e} \in \sigma_1^{\mathcal{I}} \times \cdots \times \sigma_n^{\mathcal{I}}$, the interpretation $\mathcal{I}[\boldsymbol{x} \mapsto \boldsymbol{e}, y \mapsto h(\boldsymbol{e})]$ satisfies $Q[\boldsymbol{x}, y]$. This section considers the case when $\mathbf{P}$ consists of single-invocation properties and describes a general approach for determining the satisfiability of formulas like (3) while computing a syntactic representation of a function like h in the process. For the latter, it will be convenient to assume that the language of terms contains an if-then-else operator ite of type $\mathsf{Bool} \times \sigma \times \sigma \to \sigma$ for each sort σ, with the usual semantics.

If (3) belongs to a fragment that admits quantifier elimination in T, such as the linear fragment of integer arithmetic, determining its satisfiability can be achieved using a method for quantifier elimination [8, 26]. Such cases have been examined in the context of software synthesis [22]. Here we propose instead an alternative instantiation-based approach aimed at establishing the *un*satisfiability of the *negated* form of (3):

$$\exists \boldsymbol{x} \, \forall y \, \neg Q[\boldsymbol{x}, y] \tag{4}$$

or, equivalently, of a Skolemized version $\forall y \, \neg Q[\mathsf{k}, y]$ of (4) for some tuple k of fresh free constants of the right sort. Finding a T-unsatisfiable finite set Γ of ground instances of $\neg Q[\mathsf{k}, y]$, which is what an SMT solver would do to prove the unsatisfiability of (4), suffices to solve the original synthesis problem. The reason is that, then, a solution for f can be constructed directly from Γ, as indicated by the following result.

Proposition 1 Given $\boldsymbol{x} = (x_1^{\sigma_1}, \ldots, x_n^{\sigma_n})$, let $Q[\boldsymbol{x}, y^{\sigma}]$ be a Σ-formula such that $\exists f \, \forall \boldsymbol{x} \, Q[\boldsymbol{x}, f(\boldsymbol{x})] \in \mathbf{L}$ for some $f^{\sigma_1 \cdots \sigma_n \sigma}$ not occurring in $Q[\boldsymbol{x}, y]$. Suppose some set $\Gamma = \{\neg Q[\boldsymbol{x}, t_1[\boldsymbol{x}]], \ldots, \neg Q[\boldsymbol{x}, t_p[\boldsymbol{x}]]\}$ where $t_1[\boldsymbol{x}]$, ..., $t_p[\boldsymbol{x}]$ are Σ-terms of sort σ is T-unsatisfiable. Then, one solution for $\exists f \, \forall \boldsymbol{x} \, Q[\boldsymbol{x}, f(\boldsymbol{x})]$ is

$$\lambda \boldsymbol{x} \, \mathsf{ite}(Q[\boldsymbol{x}, t_p], t_p, (\cdots \mathsf{ite}(Q[\boldsymbol{x}, t_2], t_2, t_1) \cdots)) \; .$$

Proof Let ℓ be the solution specified above. Let $\mathcal{J}$ be any model of T and let $\boldsymbol{v} = (v_1, \ldots, v_n)$ be an arbitrary element of $\sigma_1^{\mathcal{J}} \times \cdots \times \sigma_n^{\mathcal{J}}$. It is enough to show that $\mathcal{I} \models Q[\boldsymbol{x}, \ell(\boldsymbol{x})]$ where $\mathcal{I} = \mathcal{J}[\boldsymbol{x} \mapsto \boldsymbol{v}]$. Suppose first that $\mathcal{I} \models Q[\boldsymbol{x}, t_i[\boldsymbol{x}]]$ for some $i \in \{2, \ldots, p\}$ and let m be the greatest such i. Then, by construction, $\ell(\boldsymbol{x})^{\mathcal{I}} = (t_m[\boldsymbol{x}])^{\mathcal{I}}$, and thus $\mathcal{I} \models Q[\boldsymbol{x}, \ell(\boldsymbol{x})]$. If, on the other hand, $\mathcal{I} \models \neg Q[\boldsymbol{x}, t_i[\boldsymbol{x}]]$ for all $i = 2, \ldots, p$, then $\ell(\boldsymbol{x})^{\mathcal{I}} = (t_1[\boldsymbol{x}])^{\mathcal{I}}$. Since Γ is T-unsatisfiable, we have that $\neg Q[\boldsymbol{x}, t_2[\boldsymbol{x}]], \ldots, \neg Q[\boldsymbol{u}, t_p[\boldsymbol{x}]] \models_T Q[\boldsymbol{x}, t_1[\boldsymbol{x}]]$. It follows that $\mathcal{I} \models Q[\boldsymbol{x}, \ell(\boldsymbol{x})]$. □

Example 1 Let T be the theory of integer arithmetic as described in Section 2. Now consider the single-invocation property

$$P[f, \boldsymbol{x}] := f(\boldsymbol{x}) \geq x_1 \wedge f(\boldsymbol{x}) \geq x_2 \wedge (f(\boldsymbol{x}) \approx x_1 \vee f(\boldsymbol{x}) \approx x_2) \tag{5}$$

with f of type $\mathsf{Int} \times \mathsf{Int} \to \mathsf{Int}$ and $\boldsymbol{x} = (x_1, x_2)$ where x_1 and x_2 are of type Int. The synthesis problem $\exists f \, \forall \boldsymbol{x} \, P[f, \boldsymbol{x}]$ is solved exactly by the function that returns the maximum of its two inputs. Since P is single-invocation, we can solve that problem by proving the T-unsatisfiability of the conjecture $\exists \boldsymbol{x} \, \forall y \, \neg Q[\boldsymbol{x}, y]$ where

$$Q[\boldsymbol{x}, y] := y \geq x_1 \wedge y \geq x_2 \wedge (y \approx x_1 \vee y \approx x_2) \tag{6}$$

```
Synth_{SI}($\exists f\, \forall \boldsymbol{x}\, Q[\boldsymbol{x}, f(\boldsymbol{x})]$):
1. Let $\Gamma := \{\}$, and let $\mathsf{k}$, $\mathsf{e}$ be distinct fresh free constants
2. While $\Gamma$ is $T$-satisfiable
       If there is a $T$-model $\mathcal{I}$ of $\Gamma$ satisfying $Q[\mathsf{k}, \mathsf{e}]$ then
           let $\Gamma := \Gamma \cup \{\neg Q[\mathsf{k}, t[\mathsf{k}]]\}$ for some $\Sigma$-term $t[\boldsymbol{x}]$ such that $t[\mathsf{k}]^{\mathcal{I}} = \mathsf{e}^{\mathcal{I}}$
       else
           return "no solution found"
3. Let $\{\neg Q[\mathsf{k}, t_1[\mathsf{k}]], \ldots, \neg Q[\mathsf{k}, t_p[\mathsf{k}]]\}$ be a $T$-unsatisfiable subset of $\Gamma$
4. Return $\lambda \boldsymbol{x}\, \mathsf{ite}(Q[\boldsymbol{x}, t_p[\boldsymbol{x}]], t_p[\boldsymbol{x}], (\cdots \mathsf{ite}(Q[\boldsymbol{x}, t_2[\boldsymbol{x}]], t_2[\boldsymbol{x}], t_1[\boldsymbol{x}]) \cdots))$ for $f$
```

Fig. 1 A refutation-based synthesis procedure Synth$_{SI}$ for single-invocation property $\exists f\, \forall \boldsymbol{x}\, Q[\boldsymbol{x}, f(\boldsymbol{x})]$.

After Skolemization, the conjecture becomes $\forall y\, \neg Q[\mathsf{a}, y]$ for fresh constants $\mathsf{a} = (\mathsf{a}_1, \mathsf{a}_2)$. When asked to determine the satisfiability of that conjecture, an SMT solver may, for instance, instantiate it with a_1 and then a_2 for y, producing the T-unsatisfiable set $\{\neg Q[\mathsf{a}, \mathsf{a}_1], \neg Q[\mathsf{a}, \mathsf{a}_2]\}$. Since a_1 and a_2 are fresh, it follows that $\{\neg Q[\boldsymbol{x}, x_1], \neg Q[\boldsymbol{x}, x_2]\}$ is T-unsatisfiable as well. By Proposition 1, one solution for $\forall \boldsymbol{x}\, P[f, \boldsymbol{x}]$ is $f = \lambda \boldsymbol{x}\, \mathsf{ite}(Q[\boldsymbol{x}, x_2], x_2, x_1)$, which simplifies to $\lambda \boldsymbol{x}\, \mathsf{ite}(x_2 \geq x_1, x_2, x_1)$, representing the desired maximum function. $\square$

3.1 Synthesis by Counterexample-Guided Quantifier Instantiation

Given Proposition 1, the main question is how to get the SMT solver to generate the necessary ground instances from $\forall y\, \neg Q[\mathsf{k}, y]$. Typically, SMT solvers that reason about quantified formulas use heuristic quantifier instantiation techniques based on E-matching [27], which instantiates universal quantifiers with terms occurring in some current set of ground terms built incrementally from the input formula. Using E-matching-based heuristic instantiation alone is unlikely to be effective in synthesis, where required terms need to be synthesized based on the semantics of the input specification. This is confirmed by our preliminary experiments, even for simple conjectures. We have developed instead a specialized new technique, which we refer to as *counterexample-guided quantifier instantiation*, that allows the SMT solver to quickly converge in many cases to the instantiations that refute the negated synthesis conjecture (4).

The new technique is similar to a popular scheme for synthesis known as counterexample-guided inductive synthesis, implemented in various synthesis approaches (e.g., [18, 39]), but with the major difference of being built directly into the SMT solver. The technique is illustrated by the procedure in Figure 1, which grows a set Γ of ground instances of $\neg Q[\mathsf{k}, y]$. The procedure, which may not terminate in general, terminates either when Γ becomes unsatisfiable, in which case it has found a solution, or when Γ is satisfiable but all of its models falsify $Q[\mathsf{k}, \mathsf{e}]$. When this is the case, the search for a solution is inconclusive. The procedure is not *solution-complete*, that is, it is not guaranteed to return a solution whenever there is one. However, thanks to Proposition 1, it is *solution-sound*: every λ-term it returns is indeed a solution of the original synthesis problem.

Iteration	Γ unsat?	$\Gamma \cup \mathsf{Q}[\mathbf{a}, \mathsf{e}]$ unsat?	$\mathcal{I} \models \Gamma \cup \mathsf{Q}[\mathbf{a}, \mathsf{e}]$	Add to Γ
1	no	no	$\{\mathsf{e} \mapsto 0, \mathsf{a}_1 \mapsto 0, \mathsf{a}_2 \mapsto 0\}$	$\neg Q[\mathbf{a}, \mathsf{a}_1]$
2	no	no	$\{\mathsf{e} \mapsto 1, \mathsf{a}_1 \mapsto 0, \mathsf{a}_2 \mapsto 1\}$	$\neg Q[\mathbf{a}, \mathsf{a}_2]$
3	yes			

Fig. 2 A run of the procedure Synth$_{SI}$ from Figure 1 on input $\exists f\, \forall \boldsymbol{x}\, Q[\boldsymbol{x}, f(\boldsymbol{x})]$, where $Q[\mathbf{a}, \mathsf{e}] := \mathsf{e} \geq \mathsf{a}_1 \wedge \mathsf{e} \geq \mathsf{a}_2 \wedge (\mathsf{e} \approx \mathsf{a}_1 \vee \mathsf{e} \approx \mathsf{a}_2)$.

3.2 Finding instantiations

The choice of the term $t[\boldsymbol{x}]$ in Step 2 of the procedure is intentionally left underspecified because it can be done in a number of ways. Having a good heuristic for such instantiations is, however, critical to the effectiveness of the procedure in practice. In a Σ-theory T, like integer arithmetic, with a fixed interpretation for symbols in Σ and a distinguished set of ground Σ-terms denoting the elements of a sort, a simple, if naive, choice for t in Figure 1 is the distinguished term denoting the element $\mathsf{e}^{\mathcal{I}}$. For instance, if σ is Int in integer arithmetic, t could be a concrete integer constant $(0, \pm 1, \pm 2, \ldots)$. This choice amounts to testing whether points in the codomain of the sought function f satisfy the original specification P.

More sophisticated choices for $t[\boldsymbol{x}]$, in particular where t contains the variables $\boldsymbol{x}$, may increase the generalization power of this procedure and hence its ability to find a solution. Our implementation in the CVC4 solver relies on the fact that the model $\mathcal{I}$ in Step 2 is constructed from a set of equivalence classes over terms computed by the solver during its search. The procedure selects a term t among those in the equivalence class of e, other than e itself, which by construction of $\mathcal{I}$ is such that $t^{\mathcal{I}} = \mathsf{e}^{\mathcal{I}}$.

Example 2 Consider the single invocation synthesis conjecture $\exists f\, \forall \boldsymbol{x}\, Q[\boldsymbol{x}, f(\boldsymbol{x})]$ where Q is defined in Equation (6) from Example 1. The procedure from Figure 1 on this input is shown in Figure 2. In Step 1, Γ is initially empty. and we introduce the fresh free constants a_1, a_2, and e of sort Int. The columns of the table show details of the internal state of the procedure on iterations of Step 2. In the first iteration of Step 2 of the procedure, we determine that Γ is satisfiable, and a model $\mathcal{I}$ exists for $\Gamma \cup Q[\mathbf{a}, \mathsf{e}]$. Since $\mathcal{I}$ must satisfy the third conjunct of $Q[\mathbf{a}, \mathsf{e}]$, it must be the case that either $\mathsf{e}^{\mathcal{I}} = \mathsf{a}_1^{\mathcal{I}}$, $\mathsf{e}^{\mathcal{I}} = \mathsf{a}_2^{\mathcal{I}}$, or both. Assume the model $\mathcal{I}$ is such that $\mathsf{e}^{\mathcal{I}} = \mathsf{a}_1^{\mathcal{I}} = \mathsf{a}_2^{\mathcal{I}} = 0$ on this iteration. Assuming our heuristic for instantiation selects a term whose interpretation in $\mathcal{I}$ is the same as e, we choose to add the formula $\neg Q[\mathbf{a}, \mathsf{a}_1]$ to Γ on this step. On the second iteration of Step 2 of the procedure, we discover both Γ and $\Gamma \cup Q[\mathbf{a}, \mathsf{e}]$ are still satisfiable. The model $\mathcal{I}$ on this iteration must satisfy $\neg Q[\mathbf{a}, \mathsf{a}_1]$, which is $\neg \mathsf{a}_1 \geq \mathsf{a}_1 \vee \neg \mathsf{a}_1 \geq \mathsf{a}_2 \vee (\mathsf{a}_1 \not\approx \mathsf{a}_1 \wedge \mathsf{a}_1 \not\approx \mathsf{a}_2)$ and simplifies to $\neg \mathsf{a}_1 \geq \mathsf{a}_2$. Notice that the solver can no longer choose to interpret $\mathsf{e}^{\mathcal{I}} = \mathsf{a}_1^{\mathcal{I}}$ since it must satisfy $\mathsf{e} \geq \mathsf{a}_2$ and $\mathsf{a}_2 > \mathsf{a}_1$. Hence, it must be the case that $\mathsf{e}^{\mathcal{I}} = \mathsf{a}_2^{\mathcal{I}}$ on this iteration. Subsequently, we add the formula $\neg Q[\mathbf{a}, \mathsf{a}_2]$ to Γ on this step, which simplifies to $\neg \mathsf{a}_2 \geq \mathsf{a}_1$ and together with $\neg Q[\mathbf{a}, \mathsf{a}_1] \in \Gamma$ is unsatisfiable. □

The development of more sophisticated criteria for selecting instantiations that are both complete and efficient and practice is a subject of ongoing work [32]. Quantifier elimination techniques [8, 26] and approaches currently used to infer invariants from templates [11, 24] are likely to be informative for devising such

$$\begin{array}{ll}
\forall x\,y\;\mathsf{ev}(\mathsf{x}_1, x, y) \approx x & \forall s_1\,s_2\,x\,y\;\mathsf{ev}(\mathsf{leq}(s_1, s_2), x, y) \approx (\mathsf{ev}(s_1, x, y) \leq \mathsf{ev}(s_2, x, y)) \\
\forall x\,y\;\mathsf{ev}(\mathsf{x}_2, x, y) \approx y & \forall s_1\,s_2\,x\,y\;\mathsf{ev}(\mathsf{eq}(s_1, s_2), x, y) \approx (\mathsf{ev}(s_1, x, y) \approx \mathsf{ev}(s_2, x, y)) \\
\forall x\,y\;\mathsf{ev}(\mathsf{zero}, x, y) \approx 0 & \forall c_1\,c_2\,x\,y\;\mathsf{ev}(\mathsf{and}(c_1, c_2), x, y) \approx (\mathsf{ev}(c_1, x, y) \wedge \mathsf{ev}(c_2, x, y)) \\
\forall x\,y\;\mathsf{ev}(\mathsf{one}, x, y) \approx 1 & \forall c\,x\,y\;\mathsf{ev}(\mathsf{not}(c), x, y) \approx \neg\mathsf{ev}(c, x, y) \\
\multicolumn{2}{l}{\forall s_1\,s_2\,x\,y\;\mathsf{ev}(\mathsf{plus}(s_1, s_2), x, y) \approx \mathsf{ev}(s_1, x, y) + \mathsf{ev}(s_2, x, y)} \\
\multicolumn{2}{l}{\forall s_1\,s_2\,x\,y\;\mathsf{ev}(\mathsf{minus}(s_1, s_2), x, y) \approx \mathsf{ev}(s_1, x, y) - \mathsf{ev}(s_2, x, y)} \\
\multicolumn{2}{l}{\forall c\,s_1\,s_2\,x\,y\;\mathsf{ev}(\mathsf{if}(c, s_1, s_2), x, y) \approx \mathsf{ite}(\mathsf{ev}(c, x, y), \mathsf{ev}(s_1, x, y), \mathsf{ev}(s_2, x, y))}
\end{array}$$

Fig. 3 Axiomatization of the evaluation operators in grammar R from Section 4.1.

criteria. The advantage of developing these techniques within an SMT solver is that they directly benefit both synthesis and verification in the presence of quantified conjectures, thus fostering cross-fertilization between different fields.

4 Refutation-Based Syntax-Guided Synthesis

In syntax-guided synthesis, the functional specification is strengthened by an accompanying set of syntactic restrictions on the form of the expected solutions. In a recent line of work [2, 3, 30], these restrictions are expressed by a grammar R (augmented with a kind of *let* binder) defining the language of solution terms, or *programs*, for the synthesis problem. In this section, we present a variant of the approach in the previous section that incorporates the syntactic restriction R directly into the SMT solver via a deep embedding [42, 44] of the terms meeting the restriction into the solver's logic. The main idea is to represent R as a set of algebraic datatypes and build into the solver an interpretation of these datatypes in terms of the original theory T.

Our approach is limited to theories T with term-generated interpretations and to restrictions R that can be expressed as algebraic datatypes, but is generic with respect to such theories and restrictions.

For simplicity, but without loss of generality, we assume that T has a set of function symbols $\approx^{\sigma\,\sigma\,\mathsf{Bool}}$ for all $\sigma \in \Sigma^{\mathrm{s}}$, $\neg^{\mathsf{Bool}\,\mathsf{Bool}}$, $\wedge^{\mathsf{Bool}\,\mathsf{Bool}\,\mathsf{Bool}}$, etc., corresponding to the various logical connectives and interpreted as expected in every model.[5] It is not difficult to prove that then every quantifier-free formula φ can be written equivalently as an equation $t_\varphi \approx \mathsf{tt}$ where t_φ is a Boolean term—that is, $\models_T \varphi \Leftrightarrow (t_\varphi \approx \mathsf{tt})$ for some term t_φ of type Bool. This allows us to treat function symbols and logical connectives uniformly. We will then sometimes abuse the notation and not distinguish between terms of type Bool and quantifier-free formulas.

Before defining T in its full generality, we introduce it with a concrete example.

4.1 An example

Consider again the synthesis conjecture (6) from Example 1, where T is the theory of linear integer arithmetic, but now with a syntactic restriction R for the solution

[5] That is, $\neg^{\mathsf{Bool}\,\mathsf{Bool}}$ interpreted as Boolean negation, $\wedge^{\mathsf{Bool}\,\mathsf{Bool}\,\mathsf{Bool}}$ as Boolean conjunction, $=^{\sigma\,\sigma\,\mathsf{Bool}}$ as the equality function, and so on.

space expressed by these algebraic datatypes:

$$\begin{aligned}\mathsf{I} &:= \mathsf{x}_1 \mid \mathsf{x}_2 \mid \mathsf{zero} \mid \mathsf{one} \mid \mathsf{plus}(\mathsf{I},\mathsf{I}) \mid \mathsf{minus}(\mathsf{I},\mathsf{I}) \mid \mathsf{if}(\mathsf{B},\mathsf{I},\mathsf{I})\\ \mathsf{B} &:= \mathsf{leq}(\mathsf{I},\mathsf{I}) \mid \mathsf{eq}(\mathsf{I},\mathsf{I}) \mid \mathsf{and}(\mathsf{B},\mathsf{B}) \mid \mathsf{not}(\mathsf{B})\end{aligned}$$

The datatypes are meant to encode a term signature that includes nullary constructors for the integer variables x_1 and x_2 of (6), and constructors for the symbols of the arithmetic theory T. Terms of sort I (resp., B) refer to theory terms of sort Int (resp., Bool).

Instead of T, we now consider its combination T_{ev} with the theory of the datatypes above extended with two *evaluation operators*, that is, two function symbols $\mathsf{ev}^{\mathsf{I}\ \mathsf{Int}\ \mathsf{Int}\ \mathsf{Int}}$ and $\mathsf{ev}^{\mathsf{B}\ \mathsf{Int}\ \mathsf{Int}\ \mathsf{Bool}}$ respectively embedding I in Int and B in Bool. We define T_{ev} so that all of its models satisfy the formulas in Figure 3. The evaluation operators effectively define an interpreter for programs (i.e., terms of sort I and B) with input parameters x_1 and x_2.

If the syntactic restriction R is expressed in the SyGuS language [30], it is possible to instrument an SMT solver that supports (user-defined) algebraic datatypes, quantifiers and linear arithmetic so that it constructs automatically from R both the datatypes I and B and the two evaluation operators. Reasoning about I and B can be done with the datatype subsolver as long as the solver is able to decide the satisfiability of ground formulas as well as *enumerate* their models. Moreover, for each model $\mathcal{I}$ of a formula φ and each free varibale x in φ, the solver must be able to present the value $x^{\mathcal{I}}$ as a *constructor term*, that is, a term containing only constructor symbols. Reasoning about the evaluation operators is achieved by reducing ground terms of the form $\mathsf{ev}(d, t_1, t_2)$ where d is a constructor term to smaller terms by using the axioms from Figure 3 as rewrite rules, orienting each equation/equivalence from left to right.

For instance, the formula $P[f, \boldsymbol{x}]$ in Equation (5) from Example 1 can be restated in T_{ev} as the formula below where g is a variable of type I:

$$P_{\mathsf{ev}}[g, \boldsymbol{x}] := \mathsf{ev}(g, \boldsymbol{x}) \geq x_1 \wedge \mathsf{ev}(g, \boldsymbol{x}) \geq x_2 \wedge (\mathsf{ev}(g, \boldsymbol{x}) \approx x_1 \vee \mathsf{ev}(g, \boldsymbol{x}) \approx x_2)$$

In contrast to $P[f, \boldsymbol{x}]$, the new formula $P_{\mathsf{ev}}[g, \boldsymbol{x}]$ (equivalent to $P[\lambda \boldsymbol{x}\, \mathsf{ev}(g, \boldsymbol{x}), \boldsymbol{x}]$) is first-order, with the role of the second-order variable f now played by the first-order variable g.

When asked for a solution for (5) under the restriction R, the instrumented SMT solver will try to determine instead the T_{ev}-unsatisfiability of $\forall g\, \exists \boldsymbol{x}\, \neg P_{\mathsf{ev}}[g, \boldsymbol{x}]$. Instantiating g in the latter formula with $s := \mathsf{if}(\mathsf{leq}(\mathsf{x}_1, \mathsf{x}_2), \mathsf{x}_2, \mathsf{x}_1)$, say, produces a formula that the solver can prove to be T_{ev}-unsatisfiable. This suffices to show that the program $\mathsf{ite}(x_1 \leq x_2, x_2, x_1)$, the analogue of s in the language of T, is a solution of the synthesis conjecture (5) under the syntactic restriction R.

In the following we provide a general and formal definition of the theory T_{ev} mentioned in the example above, as well as a description of the synthesis procedure that we use with this theory.

4.2 From T to T_{ev}

Let T be the background Σ-theory fixed in Section 2 but with the additional assumptions that

1. T is the theory of some term-generated interpretation, and
2. its associated satisfiability procedure can also find a model for each formula φ it determines to be T-satisfiable, and output the values of φ's free variables in that model as ground Σ-terms.

The theory T_{ev} is an extension of T. Its construction depends on the syntax restriction R and on a tuple $\boldsymbol{x} = (x_1^{\sigma_1}, \dots, x_k^{\sigma_k})$ of variables with $\sigma_1, \dots, \sigma_k \in \Sigma^{\mathrm{s}}$. We fix such a tuple and consider conjectures of the form $\exists f\, \forall \boldsymbol{x}\, P[f, \boldsymbol{x}]$ from $\mathbf{L}$.

For generality, we do not discuss here any concrete language to specify the restriction R. Instead, we assume that it is possible to express R by a context-free grammar whose production rules can be faithfully encoded as a set of algebraic datatype definitions.[6] We use these datatypes to construct T_{ev}.

The datatype restriction

Let D be a theory of algebraic datatypes [7] with signature Ω and set of constructor symbols Ω^{c} which shares no sort and no function symbols with T. We denote the constructors of D by c, possibly with subscripts. We call *constructor term* any term, possibly with variables, all of whose functions symbols are constructors.

We make the following assumptions on Ω:

1. There is a ground constructor term of type σ for each $\sigma \in \Omega^{\mathrm{s}}$.
2. The number of ground constructor terms of size n is finite for each $n > 0$.
3. There is an injective mapping $s : \Omega^{\mathrm{c}} \to \Sigma^{\mathrm{f}} \cup \boldsymbol{x}$, a mapping $m : \Omega^{\mathrm{s}} \to \Sigma^{\mathrm{s}}$, and an injective mapping $m : \mathbf{X} \to \mathbf{X}$ such that
   - for all $x \in \boldsymbol{x}$, there is a $c_x \in \Omega^{\mathrm{c}}$ with $m(c_x) = x$;
   - for all $c^{\delta_1 \cdots \delta_n \delta} \in \Omega^{\mathrm{c}}$, $m(c)$ and has type $s(\delta_1) \times \cdots \times s(\delta_n) \to s(\delta)$ in Σ;
   - for all $\delta \in \Omega^{\mathrm{s}}$ and $x \in \mathbf{X}_\delta$, $m(x) \in \mathbf{X}_{s(\delta)}$.

Intuitively, D captures the symbols and the syntax of some fragment of the language of quantifier-free Σ-formulas with variables in $\boldsymbol{x}$. Some sorts of T, possibly including Bool, are represented by one or more sorts of D; some function symbols f of T are represented in D by a constructor c_f whose type maps to the type of f; and every variable x in $\boldsymbol{x}$ is represented in D by a (constant) constructor c_x whose sort maps to the sort of x. Finally, every variable of datatype δ is mapped to a variable of the corresponding type $s(\delta)$. Note that the correspondence between sorts in D and sorts in T needs not be one-to-one; this allows for greater flexibility in expressing language fragments. The fact that neither s nor m needs to be surjective corresponds to restricting the sorts and the function symbols in the fragment.

We will use m also to denote the homomorphic extension of the union of the two m's to constructor terms. Intuitively, this extension maps each constructor term to the term in T it represents.

Example 3 Looking back at the example in Section 4.1 we have that

$$\begin{aligned}
\Omega^{\mathrm{s}} &= \{\mathsf{I}, \mathsf{B}\} \\
\Omega^{\mathrm{c}} &= \{\mathsf{x}_1^{\mathsf{I}}, \mathsf{x}_2^{\mathsf{I}}, \mathsf{zero}^{\mathsf{I}}, \mathsf{one}^{\mathsf{I}}, \mathsf{plus}^{\mathsf{I\,I\,I}}, \mathsf{minus}^{\mathsf{I\,I\,I}}, \mathsf{if}^{\mathsf{B\,I\,I\,I}}, \mathsf{leq}^{\mathsf{I\,I\,B}}, \mathsf{eq}^{\mathsf{I\,I\,B}}, \mathsf{and}^{\mathsf{B\,B\,B}}, \mathsf{not}^{\mathsf{B\,B}}\} \\
s &= \{\mathsf{I} \mapsto \mathsf{Int}, \mathsf{B} \mapsto \mathsf{Bool}\} \\
m &= \{\mathsf{x}_1 \mapsto x_1, \mathsf{x}_2 \mapsto x_2, \mathsf{zero} \mapsto 0, \dots, \mathsf{if} \mapsto \mathsf{ite}, \dots, \mathsf{not} \mapsto \neg\}
\end{aligned}$$

[6] We make this assumption to simplify the presentation. The SyGuS language [30] defines a more general class of restrictions R for which this is not necessarily possible.

```
Synth_{SG}($\exists g\, \forall \boldsymbol{x}\, P_{\mathsf{ev}}[g, \boldsymbol{x}]$):
1. Let $\Gamma = \emptyset$, let $n = 1$
2. Loop
   (a) While $\Gamma$ is not $n$-satisfiable in $T_{\mathsf{ev}}$
         If the $n$-unsatisfiability proof shows that $\Gamma$ is actually $T_{\mathsf{ev}}$-unsatisfiable then
           return "no solution"
         else
           increase $n$ by some positive amount
   (b) Let $\mathcal{I}$ be a model of $T_{\mathsf{ev}}$ satisfying $\Gamma$ with $size(g^{\mathcal{I}}) \leq n$
   (c) If there is a model $\mathcal{J}$ of $T_{\mathsf{ev}}$ satisfying $\neg P_{\mathsf{ev}}[g^{\mathcal{I}}, \boldsymbol{x}]$ then
         let $\Gamma = \Gamma \cup \{P_{\mathsf{ev}}[g, \boldsymbol{t}]\}$ where $\boldsymbol{t}$ are ground $\Sigma$-terms such that $\boldsymbol{t}^{\mathcal{J}} = \boldsymbol{x}^{\mathcal{J}}$
       else
         return $g^{\mathcal{I}}$ as a solution
```

Fig. 4 A refutation-based syntax-guided synthesis procedure Synth_{SG} for $\exists g\, \forall \boldsymbol{x}\, P_{\mathsf{ev}}[g, \boldsymbol{x}]$. The expression $size(d)$ denotes the size of constructor term d.

If d is the constructor term $\mathsf{leq}(\mathsf{plus}(\mathsf{x}_1, z), \mathsf{plus}(\mathsf{z}, \mathsf{one}))$, say, where z is a variable of type I and $y = m(z)$, then $m(d) = x_1 + y \leq x_2 + 1$. □

The theory T_{ev}

The theory T_{ev} has a signature Σ_{ev} that extends the union of Σ and Ω with the following families of new function symbols:

$$\{\mathsf{ev}^{\delta\, \sigma_1 \cdots \sigma_k\, m(\delta)} \mid \delta \in D^{\mathrm{s}}\}$$
$$\{\approx^{\delta\, \delta\, \mathsf{Bool}} \mid \delta \in D^{\mathrm{s}}\}$$

T_{ev} is the union[7] of T, D and the theory consisting of all Σ_{ev}-interpretations that interpret the symbols $\mathsf{eq}^{\delta\delta\,\mathsf{Bool}}$ like the equality predicate and satisfy the following formulas for each $c^{\delta_1 \cdots \delta_n \delta} \in \Omega^{\mathrm{c}}$:

$$\forall z_1 \cdots z_k\, \mathsf{ev}(c, z_1, \cdots z_k) \approx z_i \text{ where } m(c) = x_i \text{ for some } i = 1, \ldots, k \tag{7}$$
$$\forall \boldsymbol{z}\, \mathsf{ev}(c(s_1, \ldots, s_n), \boldsymbol{z}) \approx m(c)(\mathsf{ev}(s_1, \boldsymbol{z}), \ldots, \mathsf{ev}(s_n, \boldsymbol{z})) \text{ where } m(c) \notin \boldsymbol{x} \tag{8}$$

Our synthesis procedure applies to conjectures in the class

$$\mathbf{L}_2 := \{\exists g\, \forall \boldsymbol{x}\, P[\lambda \boldsymbol{z}\, \mathsf{ev}(g, \boldsymbol{z}),\, \boldsymbol{x}] \mid P[f, \boldsymbol{x}] \in \mathbf{P}\}$$

where instead of terms of the form $f(t_1, \ldots, t_k)$ in $\mathbf{P}$ we have, modulo β-reductions, terms of the form $\mathsf{ev}(g, t_1, , \ldots, t_k)$.

Let $\exists g\, \forall \boldsymbol{x}\, \varphi[g, \boldsymbol{x}] \in \mathbf{L}_2$ and observe that g is the only term of type $\delta \in \Omega^{\mathrm{s}}$ in φ and it occurs there only in terms of the form $\mathsf{ev}(d, \boldsymbol{t})$. We say that a set $\{\varphi[g, \boldsymbol{t}_1], \ldots, \varphi[g, \boldsymbol{t}_n]\}$ of instances of a formula φ like the above is *n-satisfiable in* T_{ev} if the set is satisfied by a model of T_{ev} that interprets g as a ground constructor term of size at most n. Our procedure relies on the following fact.

Proposition 2 *For any $n > 0$, n-satisfiability in T_{ev} is decidable.*

[7] We omit the technical and somewhat tedious proof that such union theory exists.

Iteration	$\mathcal{I}$	$\mathcal{J}$	Added Formula
1	$\{g \mapsto \mathsf{x}_1, \ldots\}$	$\{x_1 \mapsto 0, x_2 \mapsto 1, \ldots\}$	$P_{\mathsf{ev}}[g, 0, 1]$
2	$\{g \mapsto \mathsf{x}_2, \ldots\}$	$\{x_1 \mapsto 1, x_2 \mapsto 0, \ldots\}$	$P_{\mathsf{ev}}[g, 1, 0]$
3	$\{g \mapsto \mathsf{one}, \ldots\}$	$\{x_1 \mapsto 2, x_2 \mapsto 0, \ldots\}$	$P_{\mathsf{ev}}[g, 2, 0]$
4	$\{g \mapsto \mathsf{plus}(x_1, \mathsf{x}_2), \ldots\}$	$\{x_1 \mapsto 1, x_2 \mapsto 1, \ldots\}$	$P_{\mathsf{ev}}[g, 1, 1]$
5	$\{g \mapsto \mathsf{if}(\mathsf{leq}(\mathsf{x}_1, \mathsf{one}), \mathsf{one}, \mathsf{x}_1), \ldots\}$	$\{x_1 \mapsto 1, x_2 \mapsto 2, \ldots\}$	$P_{\mathsf{ev}}[g, 1, 2]$
6	$\{g \mapsto \mathsf{if}(\mathsf{leq}(\mathsf{x}_1, \mathsf{x}_2), \mathsf{x}_2, \mathsf{x}_1), \ldots\}$	none	

Fig. 5 A run of the procedure Synth$_{SG}$ from Figure 4 on input $\exists g\, \forall \boldsymbol{x}\, P_{\mathsf{ev}}[g, \boldsymbol{x}]$, where $P_{\mathsf{ev}}[g, \boldsymbol{x}]$ is $\mathsf{ev}(g, \boldsymbol{x}) \geq x_1 \wedge \mathsf{ev}(g, \boldsymbol{x}) \geq x_2 \wedge (\mathsf{ev}(g, \boldsymbol{x}) \approx x_1 \vee \mathsf{ev}(g, \boldsymbol{x}) \approx x_2)$.

Proof Let $\{\varphi[g, \boldsymbol{t}_1], \ldots, \varphi[g, \boldsymbol{t}_n]\}$ be as in the definition of n-satisfiability. By our assumptions, there is a finite number of ground constructor terms of size at most n. For each such term d, each formula $\varphi[d, \boldsymbol{t}_i]$ can be effectively reduced to an T_{ev}-equivalent quantifier-free Σ-formula $\psi[\boldsymbol{t}_i]$ by using the axioms (7) and (8) as rewrite rules oriented from left to right. The satisfiability of $\{\psi[\boldsymbol{t}_1], \ldots, \psi[\boldsymbol{t}_n]\}$ can then be checked by the decision procedure for T. $\square$

Given a synthesis conjecture $\exists g\, \forall \boldsymbol{x}\, P_{\mathsf{ev}}[g, \boldsymbol{x}] \in \mathbf{L}_2$ where g is the datatype variable standing for the program to be synthesized, we use a procedure analogous to that in Section 3 to extract a solution for g from a refutation of $\forall g\, \exists \boldsymbol{x}\, \neg P_{\mathsf{ev}}[g, \boldsymbol{x}]$. The main difference, of course, is that now g ranges over the datatype representing the restricted solution space.

The procedure is described in Figure 4. It maintains a set Γ of ground instances of $P_{\mathsf{ev}}[g, \boldsymbol{x}]$ and an integer n representing an upper bound on the size of g. In each iteration of the main loop, it first determines if Γ is n-unsatisfiable. If it is, it might be because Γ itself is T_{ev}-unsatisfiable. In that case, the procedure has determined that the conjecture has no solution in the solution space defined by the restriction R and so it quits with failure. Otherwise, it has determined only that there are no solutions of size at most n and so it increments n. If the procedure finds that Γ is n-satisfiable in a model $\mathcal{I}$, it checks the T_{ev}-satisfiability of $\neg P_{\mathsf{ev}}[g^{\mathcal{I}}, \boldsymbol{x}]$. Note that since $g^{\mathcal{I}}$ is a ground constructor term, this check is effective, as explained in the proof of Proposition 2. If that formula has a model $\mathcal{J}$ this means that the candidate solution $g^{\mathcal{I}}$ has a *counterexample*, a tuple of input values, denoted by the terms $\boldsymbol{t}$, for which $g^{\mathcal{I}}$ fails to satisfy its specification P_{ev}. The procedure then adds the clause $P_{\mathsf{ev}}[g, \boldsymbol{t}^{\mathcal{J}}]$ to Γ, and repeats the main loop. That addition has the effect of requiring explicitly that all subsequent candidate solutions for g satisfy the specification P_{ev} for that counterexample for $g^{\mathcal{I}}$. Note that the terms $\boldsymbol{t}$ exist because T is the theory of a term-generated interpretation. If instead $\neg P_{\mathsf{ev}}[g^{\mathcal{I}}, \boldsymbol{x}]$ is T_{ev}-unsatisfiable, the procedure has determined that $g^{\mathcal{I}}$ satisfies $P_{\mathsf{ev}}[g^{\mathcal{I}}, \boldsymbol{x}]$ for all possible values of $\boldsymbol{x}$ and so it returns it as a solution.

The returned term $g^{\mathcal{I}}$ is a ground constructor term. A solution of the original conjecture $\exists f\, \forall \boldsymbol{x}\, P[f,\, \boldsymbol{x}]$ is then, by construction, the term $m(g^{\mathcal{I}})$.

We implemented the procedure Synth$_{SG}$ in the CVC4 solver [5]. Figure 5 shows a run of the procedure for the conjecture from Section 4.1. We show the relevant values for g in the model $\mathcal{I}$ from Step (2a), and for x_1 and x_2 in Step (2c) for 6 iterations. Each successive model $\mathcal{I}$ interprets g as a term that meets the specification P_{ev} for all inputs $\boldsymbol{x}^{\mathcal{J}}$ found for previous candidates. After the sixth iteration, the procedure finds the candidate $\mathsf{if}(\mathsf{leq}(\mathsf{x}_1, \mathsf{x}_2), \mathsf{x}_2, \mathsf{x}_1)$ For this term, the procedure cannot find an input value that falsifies P_{ev}, indicating that it is a solution for the synthesis conjecture.

Proposition 3 The procedure Synth_{SG} has the following properties:

1. (Solution Soundness) If it returns a term d as a solution, then $\forall \boldsymbol{x}\, P_{\mathsf{ev}}[d, \boldsymbol{x}]$ is T_{ev}-satisfiable.
2. (Refutation Soundness) If it answers "no solution", then $\exists g\, \forall \boldsymbol{x}\, P_{\mathsf{ev}}[g, \boldsymbol{x}]$ is T_{ev}-unsatisfiable.
3. (Solution Completeness) it terminates (with a solution) whenever its input conjecture is satisfiable.

Proof (1) For solution soundness, suppose the procedure returns some ground constructor term d as a solution. We have that $\neg P_{\mathsf{ev}}[d, \boldsymbol{x}]$ is T_{ev}-unsatisfiable. Thus, $\exists \boldsymbol{x}\, \neg P_{\mathsf{ev}}[d, \boldsymbol{x}]$ is T_{ev}-unsatisfiable and so $\forall \boldsymbol{x}\, P_{\mathsf{ev}}[d, \boldsymbol{x}]$ is T_{ev}-satisfiable.

(2) For refutation soundness, suppose the procedure returns "no solution". Then, there is a $\Gamma = \{P_{\mathsf{ev}}[g, \boldsymbol{t}_1], \ldots, P_{\mathsf{ev}}[g, \boldsymbol{t}_n]\}$ that is T_{ev}-unsatisfiable, where $\boldsymbol{t}_1, \ldots, \boldsymbol{t}_n$ are tuples of ground terms. This means that $\forall \boldsymbol{x}\, P_{\mathsf{ev}}[g, \boldsymbol{x}]$ is T_{ev}-unsatisfiable, making $\exists g\, \forall \boldsymbol{x}\, P_{\mathsf{ev}}[g, \boldsymbol{x}]$ T_{ev}-unsatisfiable as well.

(3) For solution completeness, suppose that $\exists g\, \forall \boldsymbol{x}\, P_{\mathsf{ev}}[g, \boldsymbol{x}]$ is T_{ev}-satisfiable. Then there is a ground constructor term d_0 of minimal size n_0 such that $\forall \boldsymbol{x}\, P_{\mathsf{ev}}[d_0, \boldsymbol{x}]$ is T_{ev}-satisfiable. Observe first that all the satisfiability tests in the procedure are effective. In particular, the one in Step (2a) is so by Proposition 2; the one in Step (2c) because $g^{\mathcal{I}}$ contains no variables. Hence it is enough to show that both loops in the procedure are terminating.

The while loop in Step (2a) terminates because the n counter is incremented at every iteration of the while loop. If the loop does not terminate for another reason, the value of n will eventually reach some value $k \geq n_0$, after which the loop condition will fail. This is because Γ is a consequence of $\exists g\, \forall \boldsymbol{x}\, P_{\mathsf{ev}}[g, \boldsymbol{x}]$, and hence by assumption is satisfied by a model $\mathcal{I}$ where $g^{\mathcal{I}}$ is d_0, whose size is at most k, and hence by definition Γ is k-satisfiable in T_{ev}.

The main loop terminates because, once $n = k$, by our assumptions on the datatype signature Ω, the set S of ground constructor terms with size $\leq k$ is finite. The subset S_0 of S of terms such that $\forall \boldsymbol{x}\, P_{\mathsf{ev}}[d, \boldsymbol{x}]$ is T_{ev}-satisfiable is non-empty. The value of $g^{\mathcal{I}}$ on each iteration in Step (2b) is unique. This is because each d chosen on prior iterations are such that Γ contains $P_{\mathsf{ev}}[g, \boldsymbol{t}]$ for some $\boldsymbol{t}$ where $\neg P_{\mathsf{ev}}[d, \boldsymbol{t}]$ is T_{ev}-satisfiable. By our assumptions on T, this implies all models $\mathcal{I}$ of Γ are such that $g^{\mathcal{I}} \neq d$. Hence, a model $\mathcal{I}$ where $g^{\mathcal{I}}$ with $g^{\mathcal{I}} \in S_0$ will be eventually chosen in Step (2b). At that point, the test in Step (2c) will fail and $g^{\mathcal{I}}$ will be returned. □

Note the procedure is not refutation complete: it can diverge if, but only if, the input synthesis conjecture has no solution.

4.3 Early pruning by symmetry breaking

The procedure Synth_{SG} in Figure 4 effectively considers multiple candidate solutions for the negated synthesis conjecture $\forall \boldsymbol{x}\ \neg P_{\mathsf{ev}}[g, \boldsymbol{x}]$. A concrete implementation of the procedure can be obtained through a form of branch and bound search, where the bound is on the size of the the ground constructor terms representing candidate solutions. For efficiency, it is important for such an implementation to avoid spending time considering candidate solutions that are equivalent to one another. Our implementation of the procedure within CVC4 uses *blocking clauses* to express at some point that all remaining candidate solutions for g in the current branch of the search are equivalent to previously considered solutions, forcing the search of that branch to be abandoned. These clauses may significantly increase performance.

This is typical in SMT solvers based on the DPLL(T) architecture which check the satisfiability of a quantifier-free formula φ in a background Σ-theory T by combining a SAT solver, used to check the propositional satisfiability φ, with a (set of) *theory solvers* for checking the T-satisfiability of an evolving set M of Σ-literals representing a (partial) truth value assignment for the atoms of φ. If at any time a theory solver determines a subset C of M to be T-unsatisfiable, it adds the *blocking clause* $\bigvee_{l \in C} \neg l$ to the SAT solver, indicating that at least one literals in M must be retracted. Blocking clauses, in addition to expressing that the current set M is T_{ev}-unsatisfiable, can also be used to prune future search branches by prohibiting a certain combination of literals to be reasserted again in the set M. The same mechanism is used in some advanced approaches to SAT and SMT also to break *symmetries* [1,12]. We can understand symmetries informally here as the existence of several solutions for a subproblem which are, however, equivalent for all interesting purposes.

Symmetries arise in procedure Synth_{SG} because distinct constructor terms in T_{ev} typically represent terms in T that are T-equivalent. For instance, in the example of Section 4.1 the arithmetic terms $x_1 + x_2$, $x_2 + x_1$, and $x_1 + (0 + x_2)$, corresponding respectively to the constructor terms $\mathsf{plus}(\mathsf{x}_1, \mathsf{x}_2)$, $\mathsf{plus}(\mathsf{x}_2, \mathsf{x}_1)$, and $\mathsf{plus}(\mathsf{x}_1, \mathsf{plus}(\mathsf{zero}, \mathsf{x}_2))$ are all T-equivalent. This means that once, $\mathsf{plus}(\mathsf{x}_1, \mathsf{x}_2)$, say, is considered as a candidate solution the others should not because they represent essentially the same solution.

Breaking symmetries. To show how we break symmetries in our implementation of the procedure Synth_{SG} let us start with the following observation. First, for each constructor $c^{\delta_1 \cdots \delta_n \delta} \in \Sigma_{\mathsf{ev}}^{\mathsf{c}}$ and $i = 1, \ldots, n$ let $\mathsf{s}_{c,i}^{\delta \delta_i}$ be the *selector* for the i-the argument of c. Also, let $\mathsf{is}_c^{\delta\,\mathsf{Bool}}$ denote the corresponding *tester*.[8] Then, a set of Σ_{ev}-literals M entails part of the shape that the values of a datatype variable g must have in in all models of M. For instance, if $M = \{\mathsf{is}_{\mathsf{plus}}(g), \mathsf{is}_{\mathsf{zero}}(\mathsf{s}_{\mathsf{plus},1}(g))\}$, then all the T_{ev}-models of M must interpret g as a constructor term of the form $\mathsf{plus}(\mathsf{zero}, d)$ for some d. We will write $\mathsf{d}_{g,M}[\boldsymbol{z}]$ to denote the smallest constructor term such that each variable of $\boldsymbol{z}$ occurs in the term at most once and $M \models_{T_{\mathsf{ev}}} \exists \boldsymbol{z}\, \mathsf{d}_{g,M} \approx g$. Intuitively, $\mathsf{d}_{g,M}$ is a template that expresses everything that M entails about the shape of g, and nothing more. The idea is to consider only terms $\mathsf{d}_{g,M}$ whose analogues $m(\mathsf{d}_{M,g})$ in T are unique up to *equivalence*, for some suitable equivalence relation $\equiv$.

To define $\equiv$ we assume that every Σ-term t can be effectively reduced to some T-equivalent, unique and irreducible term, which we denote by $t\!\downarrow$ and call a

[8] The selector $\mathsf{s}_{c,i}$ is such that $\mathsf{s}_{c,i}(c(d_1, \ldots, d_n)) = d_i$ for all constructor terms $d_1, \ldots, d_n$ of respective type $\delta_1, \ldots, \delta_n$. The tester is_c is such that, for all terms d of type δ, $\mathsf{is}_c(d) = \mathsf{tt}$ iff the top symbol of d is c.

$m(\mathsf{d}_{g,M_0}) \in \mathcal{S}$	$m(\mathsf{d}_{g,M_0}){\downarrow} \in \mathcal{N}$	$m(\mathsf{d}_{g,M})$	$m(\mathsf{d}_{g,M}){\downarrow}$	Blocking Clause
$(z'_1 + z'_2) + z'_3$	$z'_1 + z'_2 + z'_3$	$z_1 + (z_2 + z_3)$	$z_1 + z_2 + z_3$	$\neg(\mathsf{is}_{\mathsf{plus}}(g) \wedge \mathsf{is}_{\mathsf{plus}}(\mathsf{s}_{\mathsf{plus},2}(g)))$
z'_1	z'_1	$0 + z_1$	z_1	$\neg(\mathsf{is}_{\mathsf{plus}}(g) \wedge \mathsf{is}_{\mathsf{zero}}(\mathsf{s}_{\mathsf{plus},1}(g)))$
$x_1 + x_2$	$x_1 + x_2$	$x_2 + x_1$	$x_1 + x_2$	$\neg(\mathsf{is}_{\mathsf{plus}}(g) \wedge \mathsf{is}_{\mathsf{x}_2}(\mathsf{s}_{\mathsf{plus},1}(g)) \wedge \mathsf{is}_{\mathsf{x}_1}(\mathsf{s}_{\mathsf{plus},2}(g)))$
$\mathsf{ite}(z'_1, z'_2, z'_3)$	$\mathsf{ite}(z'_1, z'_2, z'_3)$	$\mathsf{ite}(\neg z_1, z_2, z_3)$	$\mathsf{ite}(z_1, z_3, z_2)$	$\neg(\mathsf{is}_{\mathsf{ite}}(g) \wedge \mathsf{is}_{\mathsf{not}}(\mathsf{s}_{\mathsf{ite},1}(g)))$

Fig. 6 Symmetry breaking clauses for the grammar from the example in Section 4.1. Assuming there are no models for M_0, we conclude there are no models for M.

normal form of t.[9] Then two Σ-terms t and u are $\equiv$-equivalent if there is a bijective renaming ρ of the variables of $u{\downarrow}$ such that $t{\downarrow}$ and $\rho(u{\downarrow})$ are T-equivalent.

When checking T_{ev}-satisfiability, we maintain a set of terms $\mathcal{S}$ indicating the shapes of terms we have considered (or are currently considering) for a datatype variable g so far, and a set $\mathcal{N}$ containing the normal forms for all terms in $\mathcal{S}$. At any time we are asked to check the T_{ev}-satisfiability of a set M with a free datatype variable g and $m(\mathsf{d}_{g,M})$ is not in $\mathcal{S}$, we add that term to $\mathcal{S}$ and construct the term $u = m(\mathsf{d}_{g,M}){\downarrow}$. If u is $\equiv$-equivalent to a term $m(\mathsf{d}_{g,M_0}){\downarrow}$ in $\mathcal{N}$ for some subset M_0 of M, then we produce the blocking clause $\bigvee_{l \in C} \neg l$, for some minimal $C \subseteq M$ such that $\mathsf{d}_{g,C} = \mathsf{d}_{g,M}$. Otherwise, we add u to $\mathcal{N}$.

In this method, we use the same normal form $t{\downarrow}$ for terms t that is used by theory solver for T in CVC4, making this method parametric in the background theory considered.

Examples of blocking clauses are shown in Figure 6. The first clause assumes that we have already determined that there are no solutions where g is interpreted as a term of the form $(z'_1 + z'_2) + z'_3$ for any z'_1, z'_2, z'_3. It is not useful to consider the case where g is a term of the form $z_1 + (z_2 + z_3)$, by noting their normal forms are equivalent up to renaming of variables. More generally, this scheme restricts our search to consider only left-associated chains of applications of associative operators. The other cases are similar, ensuring that the solver avoids solutions involving addition with 0, sums of monomials with identical normal forms, and terms where Boolean simplification results in a previously considered term.

5 Single Invocation Techniques for Syntax-Guided Problems

In this section, we considered the combined case of *single-invocation synthesis conjectures with syntactic restrictions* R, giving two alternative solving methods. The method uses the deep embedding introduced in Section 4 to encode the first-order variant of the synthesis conjecture. This approach handles the exploration of different cases *natively* within the SMT solver while explicitly modeling the syntax of terms returned in each case using the embedding. The second method is based on dividing the synthesis process into two steps, where we first solve the synthesis conjecture while ignoring its associated syntactic restrictions, and afterwards reconstruct an equivalent solution meeting them.

To simplify the exposition we restrict ourselves to sets R of syntactic restrictions expressed by a datatype I for programs and a datatype B for Boolean expressions. The general case of syntactic restrictions with more datatypes is similar.

[9] Note that this assumption is pragmatic and can be made arbitrarily mild. For instance, depending on the theory, in the worst case one can always consider every term to be irreducible.

Iteration	Γ unsat?	$\Gamma \cup Q_{\mathsf{B}}[\boldsymbol{a}, \mathsf{e}]$ unsat?	$\mathcal{I} \models \Gamma \cup Q_{\mathsf{B}}[\boldsymbol{a}, \mathsf{e}]$	Add to Γ
1	no	no	$\{\mathsf{e} \mapsto \mathsf{x}_1, a_1 \mapsto 0, a_2 \mapsto 0\}$	$\neg Q_{\mathsf{B}}[\mathbf{a}, \mathsf{x}_1]$
2	no	no	$\{\mathsf{e} \mapsto \mathsf{x}_2, a_1 \mapsto 0, a_2 \mapsto 1\}$	$\neg Q_{\mathsf{B}}[\mathbf{a}, \mathsf{x}_2]$
3	yes			

Fig. 7 A run of the procedure Synth$_{SI}$ from Figure 1 on input $\exists g \, \forall \boldsymbol{x} \, Q_{\mathsf{B}}[\boldsymbol{x}, g(\boldsymbol{x})]$, where $Q_{\mathsf{B}}[\boldsymbol{a}, \mathsf{e}]$ is $\mathsf{ev}(\mathsf{and}(\mathsf{leq}(\mathsf{x}_1, \mathsf{e}), \mathsf{and}(\mathsf{leq}(\mathsf{x}_2, \mathsf{e}), \mathsf{or}(\mathsf{eq}(\mathsf{e}, \mathsf{x}_1), \mathsf{eq}(\mathsf{e}, \mathsf{x}_2)))), \boldsymbol{a})$.

5.1 Method 1: Encode Single Invocation Property in T_{ev}

We consider first the case in which both of the following two conditions hold:

1. S contains the constructor $\mathsf{if}^{\mathsf{B\,I\,I\,I}}$, mapped to the if-then-else logical operator ite, and
2. the function to be synthesized is specified by a single-invocation property that can be expressed as a term of sort B.

This is the case for the syntactic restrictions R from the example in Section 4.1. To solve this conjecture, we may encode the synthesis conjecture into the language of R, and use the procedure Synth$_{SI}$ from Figure 1, since by the above properties of R it is guaranteed to find solutions meeting our syntactic requirements.

Example 4 Consider the single invocation property $\exists f \, \forall \boldsymbol{x} \, Q[\boldsymbol{x}, f(\boldsymbol{x})]$ where Q is defined in Equation (6) from Example 1, where f has type $\mathsf{Int} \times \mathsf{Int} \to \mathsf{Int}$. Assume the syntactic restrictions R from Section 4.1 are given for solutions of this conjecture. We may rephrase this conjecture as $\exists g \, \forall \boldsymbol{x} \, Q_{\mathsf{B}}[\boldsymbol{x}, g(\boldsymbol{x})]$, where

$$Q_{\mathsf{B}}[\boldsymbol{x}, y] := \mathsf{ev}(\mathsf{and}(\mathsf{leq}(\mathsf{x}_1, y), \mathsf{and}(\mathsf{leq}(\mathsf{x}_2, y), \mathsf{or}(\mathsf{eq}(y, \mathsf{x}_1), \mathsf{eq}(y, \mathsf{x}_2)))), \boldsymbol{x}) \quad (9)$$

g is a function of type $\mathsf{Int} \times \mathsf{Int} \to \mathsf{I}$, and y is of type I where I is a datatype that encodes integer terms that meet the syntactic restrictions R. A run of the procedure Synth$_{SI}$ from Figure 1 on input $\exists g \, \forall \boldsymbol{x} \, Q_{\mathsf{B}}[\boldsymbol{x}, g(\boldsymbol{x})]$ is shown in Figure 7. In Step 1, we initialize Γ to $\emptyset$, and introduce the fresh free constants a_1, a_2 of sort Int, and e of sort I. On the first iteration of Step 2 of the procedure, we find that Γ is satisfiable, and that $\Gamma \cup Q_{\mathsf{B}}[\mathbf{a}, \mathsf{e}]$ has a model $\mathcal{I}$. In contrast to the run described in Figure 2 where e was of sort Int, the e in Figure 7 is of sort I. As such, the heuristic used in Figure 2, which considered the relation between terms of sort Int, is no longer applicable for choosing an instance to add to Γ. Instead, assume we use a heuristic that adds instances to Γ based on the *value* of e in model $\mathcal{I}$, which interprets the datatype I as the set of terms built from its constructors. One such model $\mathcal{I}$ interprets e as the term x_1. Assuming this model, we add the instance $\neg Q_{\mathsf{B}}[\mathbf{a}, \mathsf{x}_1]$ to Γ, after which we discover both Γ and $\Gamma \cup Q_{\mathsf{B}}[\mathbf{a}, \mathsf{e}]$ are satisfiable. Assuming the model on the next iteration interprets e as x_2, we add the instance $\neg Q_{\mathsf{B}}[\mathbf{a}, \mathsf{x}_2]$ to Γ, which together with the previous instance are unsatisfiable. This tells us that the solution

$$\lambda \boldsymbol{x} \, \mathsf{ite}(\mathsf{ev}(\mathsf{and}(\mathsf{leq}(\mathsf{x}_1, \mathsf{x}_1), \mathsf{and}(\mathsf{leq}(\mathsf{x}_2, \mathsf{x}_1), \mathsf{or}(\mathsf{eq}(\mathsf{x}_1, \mathsf{x}_1), \mathsf{eq}(\mathsf{x}_1, \mathsf{x}_2)))), \boldsymbol{x}), \mathsf{x}_1, \mathsf{x}_2) \quad (10)$$

is a solution for g in $\exists g \, \forall \boldsymbol{x} \, Q_{\mathsf{B}}[\boldsymbol{x}, g(\boldsymbol{x})]$. The analogue of this solution

$$\lambda \boldsymbol{x} \, \mathsf{ite}(x_1 \leq x_1 \wedge x_2 \leq x_1 \wedge (x_1 \approx x_1 \vee x_1 \approx x_2), x_1, x_2) \quad (11)$$

is a solution for f in $\exists f \, \forall \boldsymbol{x} \, Q[\boldsymbol{x}, f(\boldsymbol{x})]$ that meets the syntactic restrictions R. $\square$

```
reconstruct($t, \delta$):
1. $A := \emptyset$ ; $t' := t{\downarrow}$
2. for $i = 1, 2, \ldots$
   (a) $(u, U) := \text{rcon}(t', \delta, A)$;
   (b) if $U$ is empty, return $u$; otherwise, for each datatype $\delta_j$ occurring in $U$
           let $d_i$ be the $i^{th}$ term in a fair enumeration of the elements of $\delta_j$
           add $(m(d_i){\downarrow}, m(d_i), \delta_j)$ to $A$

rcon($t, \delta, A$):
   if $(t, u, \delta) \in A$, return $(u, \emptyset)$; otherwise, do one of the following:
     (1)  choose a $f(t_1, \ldots, t_n)$ s.t. $f(t_1, \ldots, t_n){\downarrow} = t$ and $m(c) = f$ for some $c^{\delta_1 \cdots \delta_n \delta}$ in $\delta$
          let $(u_i, U_i) = \text{rcon}(t_i{\downarrow}, \delta_i, A)$ for $i = 1, \ldots, n$
          return $(f(u_1, \ldots, u_n), U_1 \cup \ldots \cup U_n)$
     (2)  return $(t, \{(t, \delta)\})$
```

Fig. 8 A procedure reconstruct for finding a term equivalent to t that meets the syntactic restrictions specified by datatype δ.

In contrast to the procedure Synth_{SG}, the advantage of running the procedure Synth_{SI} in the way described in this example is that only the outputs of a solution need to be synthesized and not conditions in ite-terms. Moreover, it maintains the benefits of the theoretical properties stated in Proposition 3, given a fair strategy for selecting candidate terms for instantiation. However, the selection criteria for instantiation in Figure 7 is much weaker than the one used in Figure 2. In Figure 7, we assumed the procedure interpreted e as x_1 and x_2 on the two iterations of the run. However, it may have interpreted e as zero on the first iteration, or one on the second iteration instead. Assuming a fair strategy for enumerating solutions for e_1, the number of iterations of Step 2 of the procedure could have been up to four. Since the efficiency of our approach for synthesis is highly dependent upon having a good method for selecting instances to add to Γ, we describe an alternative method in the following.

5.2 Method 2: Solve Single Invocation Property, Reconstruct Solution in R

An alternative approach to solve single-invocation synthesis conjectures with syntactic restrictions R is to run the procedure Synth_{SI} from Figure 1 as is, ignoring the restrictions, and subsequently reconstructs from its returned solution one that satisfies them. This has two significant advantages over the method described in Section 5.1. First, reasoning about conjectures directly allows us more powerful criteria for selecting instantiations, as seen in the differences between the runs in Figure 2 and Figure 7. Second, our experimental evaluation found that the overhead of solving ground satisfiability problems that involve an embedding into datatypes for syntax-guided problems is significant with respect to the performance of the solver on problems with no syntactic restrictions.

Figure 8 presents a procedure, called reconstruct, that takes as input a term t and a datatype I, and attempts to construct a term that is equivalent to t and meets the syntactic restrictions specified by datatype I. This procedure maintains an evolving set A of triples of the form $(u{\downarrow}, u, \delta)$, where δ is the datatype I or B, u is a term satisfying the restrictions specified by δ, and $\downarrow$ is the normalization operator as described in Section 4.3. The procedure incrementally makes calls to the (non-deterministic) subprocedure rcon, which takes a normal form term t, a datatype δ and the set A above, and returns a pair (u, U) where u is a term equivalent to the input t to reconstruct, and U is a set of pairs of the form (u', δ') where u' is a subterm of u that fails to satisfy the syntactic restriction expressed by datatype δ'. The procedure rcon may either try to match t to a term whose top symbol f has an analogue c in δ, or simply return the set $\{(t, \delta)\}$, indicating that it failed to match t to the syntactic restriction given by δ. Overall, the procedure alternates between calling rcon and adding triples to A until $\text{rcon}(t, \delta, A)$ returns a pair of the form $(s, \emptyset)$, indicating that u is a term T-equivalent to t that satisfies the syntactic restrictions embodied by I and B.

Example 5 Consider the single invocation property $\exists f\, \forall \boldsymbol{x}\, Q[\boldsymbol{x}, f(\boldsymbol{x})]$ where Q is defined in Equation (6) from Example 1, assume the syntactic restrictions R from Section 4.1 are given for solutions of this conjecture. Say we use the procedure Synth_{SI} from Figure 1 for finding a solution to this conjecture, and that Synth_{SI} returns the solution $\lambda \boldsymbol{x}\, u$ for f, where $u = \mathsf{ite}((-1 * x_1) + x_2 \leq 0, x_1, x_2)$. Note that this solution is indeed a solution for our conjecture, but it does not meet the syntactic restrictions given by the datatype I in R, since it contains the multiplication operator $*$ and unary minus $-$. To construct from that a solution that meets the syntactic restrictions represented by datatype I, we run the procedure reconstruct from Figure 8 on u and I. We let $A = \emptyset$, and call $\text{rcon}(u{\downarrow}, \mathsf{I}, A)$. The intermediate calls of a run of $\text{rcon}(u, \mathsf{I}, \emptyset)$ are shown below, where we assume that $u'{\downarrow} = u'$ for all subterms u' of u.

t	δ	return
$\mathsf{ite}((-1 * x_1) + x_2 \leq 0, x_1, x_2)$	I	$(\mathsf{ite}((-1 * x_1) + x_2 \leq 0, x_1, x_2), \{(-1 * x_1, \mathsf{I})\})$
$(-1 * x_1) + x_2 \leq 0$	B	$((-1 * x_1) + x_2 \leq 0, \{(-1 * x_1, \mathsf{I})\})$
x_1	I	$(x_1, \emptyset)$
x_2	I	$(x_2, \emptyset)$
0	I	$(0, \emptyset)$
$(-1 * x_1) + x_2$	I	$((-1 * x_1) + x_2, \{(-1 * x_1, \mathsf{I})\})$
$(-1 * x_1)$	I	$(-1 * x_1, \{(-1 * x_1, \mathsf{I})\})$

In more detail, on the initial call to rcon, since A is empty and ite is the analogue of constructor $\mathsf{if}^{\mathsf{BII}}$ in I, the procedure rcon may choose to return a pair based on the result of calling $\text{rcon}((-1 * x_1) + x_2 \leq 0, \mathsf{B}, A)$, $\text{rcon}(x_1, \mathsf{I}, A)$, and $\text{rcon}(x_2, \mathsf{I}, A)$. We may similarly traverse the function symbols of all subterms of u, with the exception of $(-1 * x_1)$. On the recursive call to $\text{rcon}(-1 * x_1, \mathsf{I}, \emptyset)$, the procedure chooses to returns a pair whose second component contains $(-1 * x_1, \mathsf{I})$, indicating it failed to produce a solution that met our syntactic restrictions. Overall, $\text{rcon}(u, \mathsf{I}, \emptyset)$ returns the pair $(u, \{(-1 * x_1, \mathsf{I})\})$. Since the second component is non-empty, in the procedure reconstruct, we enumerate the first element of I, x_1 say, whose analogue $m(\mathsf{x}_1)$ in T is x_1, and add the triple (x_1, x_1, I) to A. This indicates that there is a term in grammar I (as witnessed by x_1) that is equivalent to x_1. The call to $\text{rcon}(u, \mathsf{I}, A)$ will again return the same value, after which we pick another triple to add to A and repeat. This process continues until we enumerate the term $\mathsf{minus}(\mathsf{x}_2, \mathsf{x}_1)$ say, whose analogue $m(\mathsf{minus}(\mathsf{x}_2, \mathsf{x}_1))$ in T is $x_2 - x_1$. Assume $(x_2 - x_1){\downarrow} = (-1 * x_1) + x_2$. We add the pair $((-1 * x_1) + x_2, x_2 - x_1, \mathsf{I})$ to A.

After doing so, the subcall to rcon$((-1 * x_1) + x_2, \mathsf{I}, A)$ returns $(x_2 - x_1, \emptyset)$, and hence rcon(s, I, A) may return $(\mathsf{ite}(x_2 - x_1 \leq 0, x_1, x_2), \emptyset)$. This indicates that $\lambda x_1 x_2\, \mathsf{ite}(x_2 - x_1 \leq 0, x_1, x_2)$ is equivalent to $\lambda x_1 x_2\, \mathsf{ite}((-1 * x_1) + x_2 \leq 0, x_1, x_2)$ and thus is a solution for our conjecture, and moreover meets the restrictions specified by I. $\square$

The procedure reconstruct depends upon a normal forms for terms. Since the top symbol of t is generally ite, this normalization includes both low-level rewriting of literals within t, but also includes high-level rewriting techniques such as ite simplification, redundant subterm elimination and destructive equality resolution. Also, notice that we are not insisting that every two T-equivalent terms have the same normal form, and thus normal forms only under-approximate T-equivalence between terms. Having a stronger term reduction mechanism that reduces larger sets of terms to the same normal form allows us to compute a tighter under-approximation of T-equivalence, thus improving the performance of the reconstruction. This is the case for theories such as linear arithmetic whose normal form for terms is a sorted list of monomials, as opposed for instance to theories such as bitvectors, where distinct terms in normal form may be still equivalent to one another.

We use several optimizations, omitted in the description of the procedure in Figure 8, to increase the likelihood that the procedure terminates in a reasonable amount of time. For instance, in our implementation the return value of rcon is not recomputed every time A is updated. Instead, we maintain an evolving directed acyclic graph whose nodes are a triples of the form (t, δ, f), where t is a term δ is a datatype, and f is either a function (in which case we call it a *complete* node) or a distinguished symbol $\oslash$ (in which case we call it an *incomplete* node). In this graph, complete nodes have children of the form $(t_1, \delta_1, f_1), \ldots, (t_n, \delta_n, f_n)$, where $f(t_1, \ldots, t_n)\!\downarrow = t$ and there is a $c^{\delta_1 \cdots \delta_n \delta}$ in δ such that $m(c) = f$, and incomplete nodes have no children. We then enumerate datatype terms d_i for all datatypes δ that occur in incomplete nodes $(t, \delta, \oslash)$, replacing such a node with a corresponding graph containing only complete nodes if we find a d_i such that $m(d_i)\!\downarrow = t$. We succeed in finding a term t that meets the syntactic restrictions specified by datatype δ if we construct a graph of this form with root node (t, δ, f) that contains no incomplete nodes, where the solution can be extracted by traversing this graph.

Another important optimization is that our implementation simultaneously tries multiple alternatives considering a term t with restrictions δ. For instance, in Example 5, when calling rcon on $(-1 * x_1) + x_2 \leq 0$ and B, we chose to match it to $(-1 * x_1) + x_2 \leq 0$ whose top symbol $\leq$ has an analogue leq in B. However, we may have choosen to match it to $\neg((-1 * x_1) + x_2 > 0)$ instead, noting not is also in B. In terms of the directed acyclic graph described above, we introduce multiple nodes of the form $(t, \delta, f_1), \ldots, (t, \delta, f_n)$ when considering a term t with restrictions δ. These nodes are chosen in a way such that the size of the graph does not diverge. Whenever the graph whose root node is (t, δ, f_i) contains only complete nodes for *any* i, we remove the nodes (t, δ, f_j) for $j \neq i$ from the graph, direct their inward edges to (t, δ, f_i), and prune their children accordingly. Again, we succeed for t and δ when we construct a graph with root node (t, δ, f) containing no incomplete nodes.

	array (31)		**bv** (13)		**hd** (44)		**icfp** (50)		**int** (90)		**Total** (228)	
	#	time	#	time	#	time	#	time	#	time	#	time
cvc4+si	**31**	63.0	**6**	0.1	**44**	0.9	0	0.0	**52**	21.6	**133**	85.6
cvc4+si-r	(31)	56.9	(7)	14.0	(44)	0.9	(50)	5426.8	(89)	22.8	(221)	5521.4
cvc4+sg	1	3.4	0	0.0	26	1253.8	**1**	0.5	27	1378.0	55	2635.7
enum	3	15.4	2	28.3	38	197.5	0	0.0	36	2482.0	79	2723.1
stoch	1	0.4	0	0.0	32	655.9	0	0.0	35	702.0	68	1358.3
sketch	9	2993.6	0	0.0	35	2829.7	0	0.0	17	146.4	61	5969.8
stoast	0	0.0	0	0.0	25	2384.3	**1**	1.2	6	36.6	32	2422.1

Fig. 9 Results for single-invocation synthesis conjectures with syntactic restrictions, showing times (in seconds) and number of benchmarks solved by each solver and configuration over 5 benchmark classes with a 1800s timeout. The number of benchmarks solved by configuration **cvc4+si-r** are in parentheses because its solutions do not necessarily satisfy the given syntactic restrictions.

Although the overhead of this procedure can be significant when large subterms do not meet the syntactic restrictions, we found that in practice it quickly terminates successfully for a majority of the solutions we considered where reconstruction was possible, as we discuss in the next section. Furthermore, it makes our implementation more robust, since it effectively treats in the same way different properties that are equal modulo normalization (which is parametric in the built-in theories we consider).

6 Experimental Evaluation

We implemented the techniques from the previous sections in the SMT solver CVC4 [5], which has support for quantified formulas and a wide range of theories including arithmetic, bitvectors, and algebraic datatypes. We considered multiple configurations of CVC4 corresponding to the techniques mentioned in this paper. Configuration **cvc4+sg** executes the syntax-guided procedure from Section 4, even in cases where the synthesis conjecture is single-invocation. Configuration **cvc4+si-r** executes the procedure from Section 3 on all benchmarks having conjectures that it can deduce are single-invocation. This configuration simply ignores any syntax restrictions on the expected solution. Finally, configuration **cvc4+si** uses the same procedure used by **cvc4+si-r** but then attempts to reconstruct any found solution as a term in required syntax, as described in Section 5.2.

6.1 Benchmarks with syntactic restrictions

We evaluated our implementation on 308 benchmarks taken from the 2014 SyGuS competition [2] and the general track of the 2015 competition. The benchmarks are in a new format for specifying syntax-guided synthesis problems [30], which is supported in the latest parser of CVC4. These 308 benchmarks have an associated grammar specifying syntactic restrictions on the possible solutions.

All experiments were run on the StarExec cluster [41].[10] We provide comparative results of CVC4 against entrants of the general track of the 2015 SyGuS competition, which was won by CVC4. The other entrants of this competition were the enumerative CEGIS solver ESolver [43] (denoted **enum**), a solver based

[10] A detailed summary can be found at `http://lara.epfl.ch/w/cvc4-synthesis`.

on Stochastic search techniques [37] (denoted **stoch**), as well as Sketch [39] and Sosy Toast. We further divide our set of benchmarks into single-invocation and non-single-invocation properties. In total, CVC4 discovered that 228 of the 308 benchmarks could be rewritten into a form that was single-invocation.

Benchmarks with single-invocation synthesis conjectures. The results for benchmarks with single-invocation properties are shown in Figure 9. Configuration **cvc4+si-r** found a solution (although not necessarily in the required language) very quickly for a majority of benchmarks. It terminated successfully for 221 of 228 benchmarks, and in less than a second for 186 of those. Not all solutions found using this method met the syntactic restrictions. Nevertheless, our methods for reconstructing these solutions into the required grammar, implemented in configuration **cvc4+si**, succeeded in 133 cases, or 60% of the total. This is 54 more benchmarks than the 79 solved by the next best solver, ESolver. In total, **cvc4+si** solved 66 benchmarks that ESolver did not, while ESolver solved 12 that **cvc4+si** did not. For each of these 12 benchmarks, CVC4 was able to find a solution, but timed out trying to reconstruct the solution into the required syntax.

The solutions returned by **cvc4+si-r** were often large, having on the order of 10K subterms for harder benchmarks. However, after exhaustively applying simplification techniques during reconstruction with configuration **cvc4+si**, we found that the size of those solutions was comparable to that of the solutions produced by other solvers. For instance, among the 66 benchmarks solved by both ESolver and **cvc4+si**, the former produced a smaller solution in 25 cases. However, only in 4 cases did **cvc4+si** produce a solution that had 20 more subterms than the solution produced by ESolver. This indicates that in addition to having a high precision, the techniques from Section 5 used for solution reconstruction are generally effective at producing succinct solutions for this benchmark library.[11]

Configuration **cvc4+sg** does not take advantage of the fact that a synthesis conjecture is single-invocation. However, it was able to solve 55 of these benchmarks. In addition to being solution complete, **cvc4+sg** always produces solutions of minimal term size, something not guaranteed by the other solvers and CVC4 configurations. We found the solutions returned by **cvc4+sg** were no larger than those returned by ESolver on the 50 benchmarks they both solved. This provides an experimental confirmation that the fairness techniques for term size described in Section 4 ensure minimal size solutions.

We observe that for certain classes of benchmarks, configuration **cvc4+si** scales significantly better than state-of-the-art synthesis tools. For instance, for benchmarks from the **array** class[12], whose solutions are loop-free programs that compute the first instance of an element in a sorted array, **cvc4+si** was able to reconstruct solutions for arrays of size 15 (the largest benchmark in the class) in 0.3 seconds, and solved each of the benchmarks in the class but 8 within 1 second. In contrast, the next best tool of those shown in Figure 11 was Sketch, which solved a problem for an array of length 7 in approximately 2 minutes, but timed out for larger benchmarks. Similarly, for the parametric class of problems for synthesizing

[11] As an exception, for the **array** class of benchmarks, **cvc4+si** found solutions that were rewritten into exponentially larger ones during solution reconstruction in order to meet the given syntactic restrictions.

[12] These benchmarks, as contributed to the SyGuS benchmark set, use integer variables only; they were generated by expanding fixed-size arrays and contain no operations on arrays.

n	2	3	4	5	6	7	8	9	10	11	12	13	14	15
cvc4+si	0.01	0.01	0.02	0.03	0.1	0.1	0.3	0.4	0.6	1.2	1.9	2.6	5.9	9.9
cvc4+sg	1.6	–	–	–	–	–	–	–	–	–	–	–	–	–
enum	0.01	1033.0	–	–	–	–	–	–	–	–	–	–	–	–
stoch	0.3	4.4	997.7	–	–	–	–	–	–	–	–	–	–	–
sketch	3.1	284.0	–	–	–	–	–	–	–	–	–	–	–	–
stoast	28.7	–	–	–	–	–	–	–	–	–	–	–	–	–

Fig. 10 Results for parametric benchmarks class encoding the maximum of n integers. The columns show the run time for solvers with a 1800s timeout.

	int (8)		**invg** (26)		**invgu** (26)		**MP** (12)		**vctrl** (8)		**Total** (80)	
	#	time	#	time	#	time	#	time	#	time	#	time
cvc4+sg	2	0.1	20	2601.0	21	908.1	4	925.5	**5**	807.7	52	5242.4
enum	**3**	1.3	**23**	7.8	**23**	9.5	**9**	67.6	4	262.8	**62**	349.0
stoch	**3**	5.2	10	4.6	21	40.7	2	65.4	3	317.3	39	433.2
sketch	**3**	15.6	9	68.1	8	65.5	0	0.0	0	0.0	20	149.2
stoast	2	1.1	0	0.0	14	2805.8	0	0.0	1	21.4	17	2828.2

Fig. 11 Results for non-single-invocation synthesis conjectures with syntactic restrictions, showing times (in seconds) and numbers of benchmarks solved over 5 benchmark classes with a 1800s timeout.

a function that computes the maximum of n integer inputs, **cvc4+si** outperforms the other tools shown here by an order of magnitude or more. Figure 10 shows the comparison with other tools on such benchmarks. The Stochastic solver is able to solve the problem for $n = 3$ in approximately 15 minutes, whereas CVC4 scales to $n = 15$ and more.

Benchmarks with non-single-invocation synthesis conjectures. The configuration denoted **cvc4+sg** is the only CVC4 configuration that can process benchmarks with synthesis conjectures that are not single-invocation. The results for CVC4 and the other entrants of the 2015 SyGuS competition on the 80 non-single-invocation benchmarks from our set are shown in Figure 11. Configuration **cvc4+sg** solved 52 of them, while ESOLVER solved 62. In more detail, ESOLVER solved 12 benchmarks that **cvc4+sg** did not, while **cvc4+sg** solved 2 benchmarks (from the **vctrl** class) that ESOLVER could not solve. In terms of precision, **cvc4+sg** is competitive with the state of the art on these benchmarks, solving less than **esolver**, but nevertheless noticeably more than the other solvers of the 2015 competition.

Overall results. In total, over the entire SyGuS 2014 benchmark set, 185 benchmarks can be solved by a configuration of CVC4 that, whenever possible, runs the methods for single-invocation properties described in Section 3, and otherwise runs the method described in Section 4. This number is 44 higher than the 141 benchmarks solved in total by ESOLVER. Running both configuration **cvc4+sg** and **cvc4+si** in parallel[13] solves 193 benchmarks, indicating that CVC4 is highly competitive with state-of-the-art tools for syntax guided synthesis. CVC4's performance is noticeably better than all other solvers on single-invocation properties, where our new quantifier instantiation techniques give it a distinct advantage.

[13] CVC4 has a *portfolio* mode that allows it to run multiple configurations at the same time.

	cvc4+si-r		**AlchemistCSDT**		**AlchemistCS**	
	#	time	#	time	#	time
Total (73)	**73**	30.1	43	2340.8	33	1460.0

Fig. 12 Results for conditional linear arithmetic synthesis conjectures without syntactic restrictions. Table shows runtime and number solved for a 1800s timeout.

6.2 Benchmarks without syntactic restrictions

We also evaluated our implementation on 73 benchmarks from the conditional linear integer arithmetic track of the 2015 competition that do not have associated syntactic restrictions. We found that 71 of these 73 benchmarks had conjectures that were single-invocation. This track was won by CVC4. Here, we compare against the other entrants of this track, two versions of the Alchemist tool [36].

The results over the 73 benchmarks are shown in Figure 12. The most recent version of CVC4 is able to solve every benchmark in this set with a total time of 30.1 seconds, solving 66 of the 73 within 1 second. By contrast, the second best solver configuration, **AlchemistCSDT**, is able to solve only 43 of the 73 benchmarks. This shows that the techniques in CVC4 for handling single invocation properties are able to scale significantly better than existing approaches for synthesis for linear arithmetic.

All 43 benchmarks solved by **AlchemistCSDT** were also solved by CVC4. For these 43 benchmarks, the average term size of solutions produced by Alchemist was 116.4, and the average term size of solutions produced by CVC4 was 317.8. For 37 benchmarks, Alchemist produced a solution smaller or the same size as CVC4, and for 13 benchmarks, CVC4 produced a solution smaller or the same size as Alchemist. For 23 benchmarks, the solution produced by CVC4 did not have more than 10 more subterms than the solution produced by Alchemist. Thus, we conclude that CVC4 is able to find solutions much faster than existing approaches, but could benefit from additional techniques to minimize solution size. For instance, such techniques could be run as a post-processor to the solutions that CVC4 produces, which can be large but are produced in a highly efficient manner. The development of such techniques is the subject of future work.

7 Related Work

Early work on synthesis established a connection with automated theorem proving [16, 25], including the formulation of certain synthesis problems using $\forall\exists$ formulas. This early work used resolution-based techniques. It was noted early on that the capabilities of theorem provers were a bottleneck for effective synthesis.

The work on software synthesis procedures [21–23] introduces a particular notion of synthesis equivalent to decision procedures. Whereas the proposed framework of synthesis procedures is more general, the reported instances of that framework are based on modified quantifier elimination procedures; they were implemented and have theoretical completeness properties but do not have the efficiency and scalability of SMT solvers.

Recent work on synthesis has made use of advances in theorem proving, particularly in SAT and SMT solvers, which have already proven successful in related domains of hardware and software verification. The traditional strength of

SAT and SMT solvers has been on non-quantified formulas, so many algorithms implement the quantifier alternation outside of the solver. These approaches typically generate a sequence of increasingly more precise non-quantified queries to the solver and make use of counterexamples returned to refine the query. This approach is often called Counterexample-Guided Inductive Synthesis (CEGIS). A tool that pioneered recent interest in synthesis from complex specifications is SKETCH [38, 38, 39], which has been applied in a number of domains and uses a SAT solver. Other approaches use SMT solvers combined with dedicated implementations of search outside of the solver, which can take domain-specific constraints into account, such as synthesis of bit-manipulating programs [18, 28]. A typical role of the SMT solver has been to validate candidate solutions and provide counterexamples that guide subsequent search. SMT solvers thus receive a large number of separate queries, with limited communication between these different steps. Approaches such as symbolic term exploration [19, 20] also use an SMT solver to explore a representation of the space of solutions, although they are also implemented outside an SMT solver.

Constructive logic has recognized the usefulness of explicit witnesses, taking the (arguably extreme) viewpoint of requiring all quantifiers to be backed up by witnesses. The efforts have largely evolved around logical foundations, interactive provers [10], and type systems. For decidable theories such as integer linear arithmetic the constructive requirements are less relevant; what is more important are algorithmic aspects of automated reasoners. Insights from constructive logic are more likely to be relevant for more general synthesis tasks, such as synthesis of recursive and higher-order functions, which fall outside of the current SyGuS format.

Reactive synthesis considers a more general scenario of synthesizing a transition system that potentially interacts with a specified environment in each step of the execution. Often the specifications are given in linear temporal logic [29] or its fragments [9]. The unbounded and infinite execution traces present a source of infinity for such a model, even if the transition system itself is finite. The original techniques for this problem are based on automata theory and appear to have very high lower bounds on complexity. Recent breakthroughs in bounded synthesis, however, leverage SMT solvers [14], reducing a number of classes of reactive and distributed synthesis problems to quantified constraints over bounded linear order, providing further evidence that SMT solvers with support for quantifiers are a natural underlying technology for synthesis.

8 Conclusion

We have shown that SMT solvers, instead of just acting as subroutines for automated software synthesis tasks, can be instrumented to perform synthesis themselves. We have presented a few approaches for enabling SMT solvers to construct solutions for the broad class of syntax-guided synthesis problems and discussed their implementation in CVC4. This is, to the best of our knowledge, the first implementation of synthesis inside an SMT solver and it already shows considerable promise. Using a novel quantifier instantiation technique and a solution enumeration technique for the theory of algebraic datatypes, our implementation is highly competitive with the state of the art represented by the systems that participated in the 2015 syntax-guided synthesis competition. Moreover, for the important class of single-invocation problems when syntax restrictions permit the if-then-else operator, our implementation significantly outperforms those systems.

Acknowledgements We would like to thank Liana Hadarean for helpful discussions on the normal form used in CVC4 for bit vector terms, and Tim King for his contributions to the ground linear arithmetic solver in CVC4.

References

1. Aloul, F.A., Ramani, A., Markov, I.L., Sakallah, K.A.: Solving difficult sat instances in the presence of symmetry. In: Proceedings of the 39th annual Design Automation Conference, pp. 731–736. ACM (2002)
2. Alur, R., Bodik, R., Dallal, E., Fisman, D., Garg, P., Juniwal, G., Kress-Gazit, H., Madhusudan, P., Martin, M.M.K., Raghothaman, M., Saha, S., Seshia, S.A., Singh, R., Solar-Lezama, A., Torlak, E., Udupa, A.: Syntax-guided synthesis. To Appear in Marktoberdrof NATO proceedings (2014). `http://sygus.seas.upenn.edu/files/sygus_extended.pdf`, retrieved 2015-02-06
3. Alur, R., Bodík, R., Juniwal, G., Martin, M.M.K., Raghothaman, M., Seshia, S.A., Singh, R., Solar-Lezama, A., Torlak, E., Udupa, A.: Syntax-guided synthesis. In: FMCAD, pp. 1–17. IEEE (2013)
4. Alur, R., Martin, M.M.K., Raghothaman, M., Stergiou, C., Tripakis, S., Udupa, A.: Synthesizing finite-state protocols from scenarios and requirements. In: E. Yahav (ed.) Haifa Verification Conference, *LNCS*, vol. 8855, pp. 75–91. Springer (2014). DOI 10.1007/978-3-319-13338-6_7
5. Barrett, C., Conway, C., Deters, M., Hadarean, L., Jovanovic, D., King, T., Reynolds, A., Tinelli, C.: CVC4. In: Proceedings of CAV'11, *LNCS*, vol. 6806, pp. 171–177. Springer (2011)
6. Barrett, C., Deters, M., de Moura, L.M., Oliveras, A., Stump, A.: 6 years of SMT-COMP. JAR **50**(3), 243–277 (2013). DOI 10.1007/s10817-012-9246-5
7. Barrett, C., Shikanian, I., Tinelli, C.: An abstract decision procedure for satisfiability in the theory of inductive data types. Journal on Satisfiability, Boolean Modeling and Computation **3**, 21–46 (2007)
8. Bjørner, N.: Linear quantifier elimination as an abstract decision procedure. In: J. Giesl, R. Hähnle (eds.) IJCAR, *LNCS*, vol. 6173, pp. 316–330. Springer (2010). DOI 10.1007/978-3-642-14203-1_27
9. Bloem, R., Jobstmann, B., Piterman, N., Pnueli, A., Sa'ar, Y.: Synthesis of reactive(1) designs. J. Comput. Syst. Sci. **78**(3), 911–938 (2012). DOI 10.1016/j.jcss.2011.08.007. URL `http://dx.doi.org/10.1016/j.jcss.2011.08.007`
10. Constable, R.L., Allen, S.F., Bromley, M., Cleaveland, R., Cremer, J.F., Harper, R.W., Howe, D.J., Knoblock, T.B., Mendler, N.P., Panangaden, P., Sasaki, J.T., Smith, S.F.: Implementing mathematics with the Nuprl proof development system. Prentice Hall (1986)
11. Cousot, P.: Proving program invariance and termination by parametric abstraction, lagrangian relaxation and semidefinite programming. In: R. Cousot (ed.) VMCAI, *LNCS*, vol. 3385, pp. 1–24. Springer (2005). DOI 10.1007/978-3-540-30579-8_1
12. Déharbe, D., Fontaine, P., Merz, S., Paleo, B.W.: Exploiting symmetry in smt problems. In: Automated Deduction–CADE-23, pp. 222–236. Springer (2011)
13. Detlefs, D., Nelson, G., Saxe, J.B.: Simplify: A theorem prover for program checking. Tech. rep., J. ACM (2003)
14. Finkbeiner, B., Schewe, S.: Bounded synthesis. STTT **15**(5-6), 519–539 (2013). DOI 10.1007/s10009-012-0228-z. URL `http://dx.doi.org/10.1007/s10009-012-0228-z`
15. Ge, Y., de Moura, L.: Complete instantiation for quantified formulas in satisfiability modulo theories. In: Proceedings of CAV'09, *LNCS*, vol. 5643, pp. 306–320. Springer (2009). DOI http://dx.doi.org/10.1007/978-3-642-02658-4_25
16. Green, C.C.: Application of theorem proving to problem solving. In: D.E. Walker, L.M. Norton (eds.) IJCAI, pp. 219–240. William Kaufmann (1969)

17. Jacobs, S., Kuncak, V.: Towards complete reasoning about axiomatic specifications. In: Verification, Model Checking, And Abstract Interpretation, pp. 278–293. Springer Berlin Heidelberg (2011)
18. Jha, S., Gulwani, S., Seshia, S.A., Tiwari, A.: Oracle-guided component-based program synthesis. In: J. Kramer, J. Bishop, P.T. Devanbu, S. Uchitel (eds.) ICSE, pp. 215–224. ACM (2010). DOI 10.1145/1806799.1806833
19. Kneuss, E., Koukoutos, M., Kuncak, V.: Deductive program repair. In: D. Kroening, C.S. Pasareanu (eds.) CAV, *LNCS*, vol. 9207, pp. 217–233. Springer (2015). DOI 10.1007/978-3-319-21668-3_13
20. Kneuss, E., Kuraj, I., Kuncak, V., Suter, P.: Synthesis modulo recursive functions. In: A.L. Hosking, P.T. Eugster, C.V. Lopes (eds.) OOPSLA, pp. 407–426. ACM (2013). DOI 10.1145/2509136.2509555
21. Kuncak, V., Mayer, M., Piskac, R., Suter, P.: Complete functional synthesis. In: B.G. Zorn, A. Aiken (eds.) PLDI, pp. 316–329. ACM (2010). DOI 10.1145/1806596.1806632
22. Kuncak, V., Mayer, M., Piskac, R., Suter, P.: Software synthesis procedures. CACM **55**(2), 103–111 (2012). DOI 10.1145/2076450.2076472
23. Kuncak, V., Mayer, M., Piskac, R., Suter, P.: Functional synthesis for linear arithmetic and sets. STTT **15**(5-6), 455–474 (2013). DOI 10.1007/s10009-011-0217-7
24. Madhavan, R., Kuncak, V.: Symbolic resource bound inference for functional programs. In: A. Biere, R. Bloem (eds.) CAV, *LNCS*, vol. 8559, pp. 762–778. Springer (2014). DOI 10.1007/978-3-319-08867-9_51
25. Manna, Z., Waldinger, R.J.: A deductive approach to program synthesis. TOPLAS **2**(1), 90–121 (1980). DOI 10.1145/357084.357090
26. Monniaux, D.: Quantifier elimination by lazy model enumeration. In: T. Touili, B. Cook, P. Jackson (eds.) CAV, *LNCS*, vol. 6174, pp. 585–599. Springer (2010). DOI 10.1007/978-3-642-14295-6_51
27. de Moura, L.M., Bjørner, N.: Efficient e-matching for SMT solvers. In: F. Pfenning (ed.) CADE, *LNCS*, vol. 4603, pp. 183–198. Springer (2007). DOI 10.1007/978-3-540-73595-3_13
28. Perelman, D., Gulwani, S., Grossman, D., Provost, P.: Test-driven synthesis. In: M.F.P. O'Boyle, K. Pingali (eds.) PLDI, p. 43. ACM (2014). DOI 10.1145/2594291.2594297
29. Pnueli, A., Rosner, R.: On the synthesis of a reactive module. In: Conference Record of the Sixteenth Annual ACM Symposium on Principles of Programming Languages, Austin, Texas, USA, January 11-13, 1989, pp. 179–190 (1989). DOI 10.1145/75277.75293
30. Raghothaman, M., Udupa, A.: Language to specify syntax-guided synthesis problems. CoRR **abs/1405.5590** (2014). URL `http://arxiv.org/abs/1405.5590`
31. Reynolds, A., Deters, M., Kuncak, V., Barrett, C.W., Tinelli, C.: Counterexample guided quantifier instantiation for synthesis in CVC4. In: Computer Aided Verification (CAV). Springer (2015)
32. Reynolds, A., King, T., Kuncak, V.: An instantiation-based approach for solving quantified linear arithmetic. CoRR **abs/1510.02642** (2015)
33. Reynolds, A., Tinelli, C., Goel, A., Krstić, S., Deters, M., Barrett., C.: Quantifier instantiation techniques for finite model finding in SMT. In: M.P. Bonacina (ed.) Proceedings of the 24th International Conference on Automated Deduction (Lake Placid, NY, USA), *Lecture Notes in Computer Science*, vol. 7898, pp. 377–391. Springer (2013)
34. Reynolds, A., Tinelli, C., Moura, L.D.: Finding conflicting instances of quantified formulas in SMT. In: Formal Methods in Computer-Aided Design (FMCAD) (2014)
35. Ryzhyk, L., Walker, A., Keys, J., Legg, A., Raghunath, A., Stumm, M., Vij, M.: User-guided device driver synthesis. In: J. Flinn, H. Levy (eds.) OSDI, pp. 661–676. USENIX Association (2014)
36. Saha, S., Garg, P., Madhusudan, P.: Alchemist: Learning guarded affine functions. In: D. Kroening, C.S. Psreanu (eds.) Computer Aided Verification, *Lecture Notes in Computer Science*, vol. 9206, pp. 440–446. Springer International Publishing (2015). DOI 10.1007/978-3-319-21690-4_26. URL `http://dx.doi.org/10.1007/978-3-319-21690-4_26`
37. Schkufza, E., Sharma, R., Aiken, A.: Stochastic superoptimization. SIGPLAN Not. **48**(4), 305–316 (2013). DOI 10.1145/2499368.2451150. URL `http://doi.acm.org/10.1145/2499368.2451150`
38. Solar-Lezama, A.: Program sketching. STTT **15**(5-6), 475–495 (2013). DOI 10.1007/s10009-012-0249-7
39. Solar-Lezama, A., Tancau, L., Bodík, R., Seshia, S.A., Saraswat, V.A.: Combinatorial sketching for finite programs. In: J.P. Shen, M. Martonosi (eds.) ASPLOS, pp. 404–415. ACM (2006). DOI 10.1145/1168857.1168907

40. Srivastava, S., Gulwani, S., Foster, J.S.: Template-based program verification and program synthesis. STTT **15**(5-6), 497–518 (2013). DOI 10.1007/s10009-012-0223-4
41. Stump, A., Sutcliffe, G., Tinelli, C.: Starexec: a cross-community infrastructure for logic solving. In: Proceedings of the 7th International Joint Conference on Automated Reasoning, Lecture Notes in Artificial Intelligence. Springer-Verlag (2014)
42. Svenningsson, J., Axelsson, E.: Combining deep and shallow embedding for EDSL. In: Trends in Functional Programming - 13th International Symposium, TFP 2012, St. Andrews, UK, June 12-14, 2012, Revised Selected Papers, pp. 21–36 (2012). DOI 10.1007/978-3-642-40447-4_2
43. Udupa, A., Raghavan, A., Deshmukh, J.V., Mador-Haim, S., Martin, M.M., Alur, R.: Transit: Specifying protocols with concolic snippets. In: PLDI, pp. 287–296. ACM (2013). DOI 10.1145/2491956.2462174. URL `http://doi.acm.org/10.1145/2491956.2462174`
44. Wildmoser, M., Nipkow, T.: Certifying machine code safety: Shallow versus deep embedding. In: Theorem Proving in Higher Order Logics, 17th International Conference, TPHOLs 2004, Park City, Utah, USA, September 14-17, 2004, Proceedings, pp. 305–320 (2004). DOI 10.1007/978-3-540-30142-4_22