Enumerating Choice Terms in Model-Based Quantifier Instantiation

Lydia Kondylidou[1], Andrew Reynolds[2], Jasmin Blanchette[1], and Cesare Tinelli[2]

[1] Ludwig-Maximilians-Universität München, Munich, Germany
l.kondylidou@lmu.de, jasmin.blanchette@lmu.de
[2] The University of Iowa, Iowa City, United States
andrew-reynolds@uiowa.edu, cesare-tinelli@uiowa.edu

Abstract. Satisfiability modulo theories (SMT) solvers are widely used for determining the satisfiability of logical formulas with respect to background theories. SMT solvers are traditionally based on first-order logic, but some also support higher-order logic. Recently, Kondylidou et al. introduced model-based quantifier instantiation with fast enumeration (MBQI-Enum), a quantifier instantiation strategy that works for both logics. A weakness of MBQI-Enum is that it does not find refutations when Hilbert choice terms are necessary. In this work, we present an extension of MBQI-Enum that enables it to reason effectively about Hilbert's choice operator. The extended strategy substantially increases the success rate of the SMT solver cvc5 on higher-order benchmarks.

1 Introduction

Satisfiability modulo theories (SMT) solvers combine a Boolean satisfiability (SAT) solver with decision procedures for interpreted theories. They work by refutation: They assume the negation of the conjecture as an axiom and try to establish the unsatisfiability (i.e., provability) of the input problem. Several SMT solvers, including Bitwuzla [25], Boolector [26], CVC4 [5], cvc5 [3], veriT [8], and Z3 [23], support quantifiers via Skolemization and instantiation.

Most SMT solvers are based on first-order logic, but a few support higher-order logic. Specifically, CVC4, its successor cvc5, and a prototype version of veriT have been extended to parse and solve higher-order problems [4]. In principle, higher-order logic is often more convenient than first-order logic for expressing problems, particularly those involving binders (e.g., λx, $\sum_i$, $\prod_i$, $\int_x$).

A crucial aspect of higher-order SMT solvers is their use of quantifier instantiation strategies. Barbosa et al. [4] partly extended the first-order E-matching strategy [22] to a higher-order setting. Recently, Kondylidou et al. [19] introduced model-based quantifier instantiation with fast enumeration (MBQI-Enum), a strategy that extends model-based quantifier instantiation (MBQI) [15] with syntax-guided synthesis (SyGuS) [27] techniques. While traditional MBQI relies only on ground terms from the MBQI model to instantiate quantified variables, MBQI-Enum broadens this approach by generating a wider range of candidate instantiations. It does so by enumerating terms guided by a SyGuS grammar, which enables it to construct more complex instantiations involving, for instance, identity functions and uninterpreted symbols. MBQI-Enum was found to be the most successful strategy [19, Sect. 5] for solving higher-order problems from the TPTP [39] library.

A weakness of MBQI-Enum is that it does not attempt instantiations containing Hilbert's choice operator. A Hilbert choice expression has the form $\varepsilon x.\ \varphi$, where variable x is bound in formula φ. The entire expression denotes some value of x that satisfies φ, if such a value exists; otherwise, it denotes an arbitrary value from the type of x. Sometimes the ε operator occurs in the input problem, in which case the SMT solver must reason about it. Even if ε is absent, an ε term might be useful in quantifier instantiations. The ε operator is reputed to be difficult to reason about because it is characterized by a higher-order axiom.

In this paper, we extend MBQI-Enum so that it instantiates quantifiers with terms that include Hilbert's choice operator. This extension requires three main modifications to MBQI-Enum. First, we augment the SyGuS grammar to include ε terms. Second, we introduce fresh Skolem symbols that represent these ε terms, since this is simpler than implementing ε terms throughout the SMT solver. Finally, we generate lemmas corresponding to instances of ε's characteristic axiom and add them to the solver.

Example 1. We compare our extended MBQI-Enum with the original on TPTP benchmark `SEV431^1`. Let u be an uninterpreted sort, and let f be an uninterpreted symbol of type $u \to u$. Consider the second-order problem consisting of the injectivity axiom $\forall x, y.\ \mathsf{f}\ x = \mathsf{f}\ y \Longrightarrow x = y$ for f and the negated conjecture $\forall g.\ \exists z.\ g\ (\mathsf{f}\ z) \neq z$, where g has type $u \to u$ and x, y, z have type u. MBQI-Enum generates the instantiations $\lambda y.\ y$, $\lambda y.\ \mathsf{f}\ y$, $\lambda y.\ \mathsf{f}\ (\mathsf{f}\ y)$, $\ldots$ for g. As a result, the SMT solver does not terminate. By contrast, our strategy instantiates g with the term $\lambda y.\ \varepsilon x.\ y = \mathsf{f}\ x$ using the grammar. This term denotes the left inverse of f. Then the strategy introduces a fresh Skolem symbol h that represents this term. The lemma $\forall y.\ \neg\,(\exists x.\ y = \mathsf{f}\ x) \vee y = \mathsf{f}\ (\mathsf{h}\ y)$ characterizing h is added to the problem. After instantiating g with h, the solver can derive a contradiction with the injectivity axiom. This implies that the unnegated conjecture $\exists g.\ \forall z.\ g\ (\mathsf{f}\ z) = z$ is entailed by the axiom. ■

We implemented the extended MBQI-Enum strategy in the cvc5 solver. We evaluated it on the higher-order benchmarks of the TPTP library [39]. The results show that Hilbert's choice is often useful for establishing the unsatisfiability of an input problem without introducing much overhead in problems where it is not needed. Remarkably, our extension solves a strict superset of the problems solved by the original MBQI-Enum.

We also compare our approach with other cvc5 strategies and find that it outperforms them. Finally, we compare cvc5 equipped with our strategy with the provers Satallax [10], Vampire [7], and Zipperposition [40], representing the state of the art in higher-order automated theorem proving, and find that it is highly competitive. The raw evaluation data are available online [21]. Our source code, along with instructions for reproducing the experiments, is also available online [20].

2 Preliminaries

Our approach builds on higher-order logic with Hilbert's choice, SMT with quantifiers, and MBQI-Enum. Below, we briefly introduce these concepts.

Higher-Order Logic with Hilbert's Choice. Monomorphic higher-order logic [1, 16], also called simple type theory [11], generalizes classical first-order logic by allowing quantification over functions. The syntax distinguishes between types and terms. Types τ are either base types, or sorts, κ or applications of the function type constructor $\to$ to two types: $\tau_1 \to \tau_2$. The type of Booleans is denoted by o.

The term language is based on the simply typed λ-calculus, where terms $t, e, \ldots$ are inductively defined as variables $g, x, y, z, \ldots$, symbols $\mathsf{a}, \mathsf{b}, \mathsf{f}, \mathsf{h}, \mathsf{p}, \mathsf{r} \ldots$ (possibly of function type), term applications $t\ t'$, and λ-abstractions $\lambda x.\ t$, where x is the bound variable and t is the abstraction's body. A variable is free in a term if it is not bound by an enclosing λ-abstraction. Terms are syntactically equal modulo α-, β-, and η-conversion, meaning, for example, that $(\lambda x.\ \mathsf{f}\ x\ x)\ \mathsf{c}$ is syntactically equal to $\mathsf{f}\ \mathsf{c}\ \mathsf{c}$. Terms of type $\tau_1 \to \cdots \to \tau_n \to o$ are called predicates. A term of type o is called a formula.

A common extension of higher-order logic is *Hilbert's choice operator* ε [17,18, 37]: If x is a variable of type τ and φ is a formula that contains x, then $\varepsilon\ (\lambda x.\ \varphi)$, abbreviated as $\varepsilon x.\ \varphi$, is a term of type τ. Intuitively, the term $\varepsilon x.\ \varphi$ denotes *some* element x of type τ for which the formula φ holds if such an element exists. If no such element exists, it denotes an arbitrary element of type τ. Unlike existential quantification, which only asserts the existence of such an element, the choice operator yields the element. Hilbert's choice operator is characterized by the higher-order axiom $\forall y.\ (\exists x.\ y\ x) \Longrightarrow y\ (\varepsilon x.\ y\ x)$. The operator makes it possible to denote functions that exist according to the standard semantics of higher-order logic but that otherwise cannot be expressed (cf. Example 1). Some treatments of Hilbert's choice include an axiom $\forall y, z.\ y = z \Longrightarrow (\varepsilon x.\ y\ x) = (\varepsilon x.\ z\ x)$, but this is a special case of congruence, which is built into higher-order logic.

Quantifier Instantiation in SMT. To reason about the satisfiability of quantified formulas, SMT solvers typically rely on a combination of Skolemization and heuristic instantiation: Universal quantifiers occurring positively in a formula are instantiated, while those occurring negatively are Skolemized. Dually, existential quantifiers occurring positively are Skolemized, whereas those occurring negatively are instantiated.

Let $\mathcal{T}$ be a theory, or combination of theories, over a set of interpreted symbols, and let F be a (conjunctively understood) set of formulas, or *problem*, over the signature of $\mathcal{T}$. To determine the satisfiability of F in $\mathcal{T}$, typical SMT solvers rely on a *quantifier-free subsolver* for reasoning about quantifier-free formulas and an *instantiation module* for reasoning about quantified formulas. The former, which combines a SAT solver and a specialized solver for the theory $\mathcal{T}$, enumerates potential models of an abstraction of F in which quantified subformulas are treated as propositional variables. The SAT solver proposes candidate models for F at the propositional level—truth assignments to atomic formulas in F that propositionally satisfy F—while the theory solvers check these models for theory-specific conflicts. If all candidate models generate conflicts, F is declared unsatisfiable. Otherwise, the instantiation module handles the quantified subformulas in F by generating selected instances of them and adding them to F. The added instances preserve the satisfiability of F. Instances with quantifiers will be possibly instantiated further, whereas those with no quantifiers will be directly processed by the quantified-free subsolver.

Instantiation is performed following user-selectable strategies that compute substitutions σ for the top-level variables in quantified formulas of F, mapping those variables to ground (i.e., variable-free) terms. More precisely, for selected quantified formulas q of the form $\forall x_1, \ldots, x_n.\ \psi$ in F and selected substitutions σ that map each variable x_i to a ground term t_i, the instantiation module produces *lemmas* of the form $q \Longrightarrow \psi\{x_1 \mapsto t_1, \ldots, x_n \mapsto t_n\}$ that are then added to the set F. Existentially quantified formulas in F are instantiated by Skolemization: for each selected formula q of the form $\exists x_1, \ldots, x_n.\ \psi$ in F, n fresh Skolem symbols $\mathsf{sk}_1, \ldots, \mathsf{sk}_n$ are introduced and a corresponding lemma of the form $q \Longrightarrow \psi\{x_1 \mapsto \mathsf{sk}_1, \ldots, x_n \mapsto \mathsf{sk}_n\}$ is added to F. The main SMT loop is then reentered with the new F.

This iterative process continues until the SMT solver finds a model for F with respect to $\mathcal{T}$ (which is itself a difficult task given the presence in F of quantified formulas [14, 15, 32, 34]) or the quantifier-free subsolver determines F to be unsatisfiable in $\mathcal{T}$. Of course, at any point, the SMT solver may run out of time or memory.

To generate relevant ground terms for instantiating universal quantifiers, SMT solvers rely on *instantiation strategies*. For first-order logic with theories, some strategies are refutationally complete and enable SMT solvers to offer semidecision procedures: They will find a proof of unsatisfiability if one exists, but they may not terminate otherwise. Higher-order logic—which allows terms to include functions as arguments, λ-abstractions, and partial applications—remains semidecidable under Henkin semantics but is more challenging to support in practice than first-order logic. The strategies implemented in CVC4, cvc5, and veriT are pragmatic and do not seek refutational completeness [4, 19].

For higher-order logic, cvc5 initially supported only an extension of the E-matching [22] instantiation strategy. New strategies that also partly work on higher-order problems were later added to the solver. Specifically, MBQI [15] is a strategy that iteratively refines a candidate model constructed from the quantifier-free part of the problem and then attempts to refute the model by generating ground instances of the quantified formulas that may be falsified by the model. It has been used on higher-order benchmarks and has proved somewhat effective on these [19, Sect. 5]. Common techniques such as conflict-guided instantiation [35], enumerative instantiation [30], and counterexample-guided instantiation [33] can also be used on higher-order problems, but they have generally proved less effective [19, Sect. 5], which is unsurprising given that they were originally designed for first-order reasoning. SyQI [27], a syntax-guided synthesis approach, works well on some higher-order problems since it can generate terms by synthesizing candidate expressions with the help of a grammar [19, Sect. 5].

MBQI-Enum. MBQI-Enum [19] is a recent addition to cvc5 that integrates a SyGuS-based enumerator directly into MBQI. For each quantified variable in the formula, MBQI-Enum first constructs a grammar. It then iteratively enumerates candidate terms generated from the grammar and checks each resulting instance against the current model. If one instance refutes the current model, the strategy builds a substitution by mapping the quantified variable to the current enumerated term. If the instance fails to refute the model, the enumerator proceeds to the next candidate term. If all candidates are exhausted without success, the strategy reverts to standard MBQI. This fallback mechanism guarantees that MBQI-Enum can, in principle, solve any problem that MBQI can solve.

A key aspect of MBQI-Enum is the construction of the SyGuS grammar that guides the enumeration. The grammar is created from a set of symbols including uninterpreted symbols extracted from the entire formula and bound variables that have not yet been instantiated. For higher-order variables, the grammar incorporates rules for generating λ-abstractions. The enumerator then systematically constructs λ-expressions by generating terms over the abstractions' argument variables.

Example 2. We consider a modified version of the TPTP benchmark `SYO288^5`, which is over the theory $\mathcal{T}$ of equality with uninterpreted sorts and functions. Let u be an uninterpreted sort. Let f be a function symbol $\mathsf{f} : u \to u$ and consider the second-order conjecture $\exists y.\ \forall z.\ y\ z = \mathsf{f}\ z$, where y is a variable of type $u \to u$ and z is a variable of type u. To prove its validity in $\mathcal{T}$, we ask the SMT solver to prove the unsatisfiability of its negation, or, equivalently, of the input problem:

$$F = \{\forall y.\ \exists z.\ y\ z \neq \mathsf{f}\ z\}$$

MBQI-Enum first constructs a SyGuS grammar for y to guide the enumeration process. The set of symbols is empty. The grammar consists of the rules

$$\begin{aligned}
g_0 &::= \lambda x.\ g_1 \\
g_1 &::= x \mid \mathsf{f}\ g_1 \mid \mathrm{ite}(g_2, g_1, g_1) \\
g_2 &::= \mathsf{true} \mid \mathsf{false} \mid g_1 = g_1 \mid \neg\, g_2 \mid g_2 \wedge g_2 \mid g_2 \vee g_2
\end{aligned}$$

In the first iteration, MBQI-Enum enumerates terms generated by the grammar and tries them in sequence. When the enumeration reaches the term $\lambda x.\ x$, MBQI-Enum generates the substitution $\sigma = \{y \mapsto \lambda x.\ x\}$. The strategy then determines whether the instantiation of the quantified formula, obtained by applying σ to the subformula $\exists z.\ y\ z \neq \mathsf{f}\ z$, refutes the current model. This is achieved by checking whether the negation of that subformula, under the current substitution σ, is $\mathcal{T}$-satisfiable. Now, if we negate the subformula, instantiate y with $\lambda x.\ x$ in it and β-reduce, we obtain a formula equivalent to $\forall z.\ z = \mathsf{f}\ z$, which is satisfiable in $\mathcal{T}$ since in that theory the symbol f can be interpreted as the identity function. The substitution σ is then returned.

Back in the SMT loop, the instantiation lemma $(\forall y.\ \exists z.\ y\ z \neq \mathsf{f}\ z) \Longrightarrow \exists z.\ z \neq \mathsf{f}\ z$ is then added to F. In essence, this causes the solver to also add the propositionally entailed formula $\exists z.\ z \neq \mathsf{f}\ z$ to F. Since that formula has an existential quantifier prefix, it is Skolemized into $\mathsf{sk}_1 \neq \mathsf{f}\ \mathsf{sk}_1$ where sk_1 is a fresh symbol of type u. Then, the Skolemization lemma $(\exists z.\ z \neq \mathsf{f}\ z) \Longrightarrow \mathsf{sk}_1 \neq \mathsf{f}\ \mathsf{sk}_1$ added to F.

In the next iteration, MBQI-Enum tries the term $\lambda x.\ \mathsf{f}\ x$ generated by the grammar and produces the substitution $\sigma = \{y \mapsto \lambda x.\ \mathsf{f}\ x\}$, leading to the addition of the instantiation lemma $(\forall y.\ \exists z.\ y\ z \neq \mathsf{f}\ z) \Longrightarrow \exists z.\ \mathsf{f}\ z \neq \mathsf{f}\ z$ to F. Then, the formula $\exists z.\ \mathsf{f}\ z \neq \mathsf{f}\ z$ is Skolemized, with variable z replaced by a fresh Skolem symbol sk_2, and the Skolemization lemma $(\exists z.\ \mathsf{f}\ z \neq \mathsf{f}\ z) \Longrightarrow \mathsf{f}\ \mathsf{sk}_2 \neq \mathsf{f}\ \mathsf{sk}_2$ is added to F. Now, from the formulas $\forall y.\ \exists z.\ y\ z \neq \mathsf{f}\ z$, $(\forall y.\ \exists z.\ y\ z \neq \mathsf{f}\ z) \Longrightarrow \exists z.\ \mathsf{f}\ z \neq \mathsf{f}\ z$, and $(\exists z.\ \mathsf{f}\ z \neq \mathsf{f}\ z) \Longrightarrow \mathsf{f}\ \mathsf{sk}_2 \neq \mathsf{f}\ \mathsf{sk}_2$ in F, the quantifier-free subsolver deduces the propositionally entailed formula $\mathsf{f}\ \mathsf{sk}_2 \neq \mathsf{f}\ \mathsf{sk}_2$, which it then determines to be unsatisfiable. The SMT solver then concludes that the original F is unsatisfiable in $\mathcal{T}$, and hence that the conjecture is valid. ■

3 The Extended Strategy

Our strategy is based on MBQI-Enum, which has proved effective for higher-order reasoning. It adds three main ingredients:

- the grammar guiding the enumeration is augmented to include ε terms;
- fresh Skolem symbols are introduced to represent these terms; and
- term filtering is applied to eliminate redundant or invalid candidate terms during enumeration.

In addition to these extensions, our strategy generalizes MBQI-Enum by returning not only a substitution but also a set of lemmas. These lemmas are added to the input problem F, extending it iteratively.

The generalized strategy is detailed in Algorithm 3. It takes as input a quantified formula q of the form $\forall y_1, \ldots, y_m.\, \psi$ occurring in F, where ψ does not start with a universal quantifier. It starts by computing an initial substitution σ from the model (line 3). If no such substitution exists, it returns the empty set, indicating that no instantiation could be found.

After computing σ, the strategy iterates over the quantified variables $y_1, \ldots,$ y_m in the formula q to instantiate. For each y_i where $i \in \{1, \ldots, m\}$, it constructs a SyGuS grammar G_i based on uninterpreted symbols appearing in the entire input formula (line 7). The grammar is further augmented to include terms of the form $\varepsilon x.\ \varphi$, where x is a variable whose type is the return type of y_i and φ is a formula over x and other free variables and symbols from the formula q.

Algorithm 3 Generalized MBQI-Enum

1: **function** MBQI_ENUM_WITH_CHOICE(q)
2:     **assume** q is $\forall y_1, \ldots, y_n.\ \psi$
3:     **let** $\sigma \leftarrow$ INSTS_MBQI(q)
4:     **if** σ does not exist **then**
5:         **return** $\emptyset$
6:     **for each** $i \in \{1, \ldots, n\}$ **do**
7:         **let** $G_i \leftarrow$ MAKE_GRAMMAR(q, y_i)
8:         **for each** $j \in \{1, 2, \ldots\}$ **do**
9:             **let** $e \leftarrow$ GET_ENUM_TERM(G_i, j)
10:             **if** e does not exist **then**
11:                 **break**
12:             $\sigma' \leftarrow \sigma[y_i \mapsto e]$
13:             **if** $\neg \psi\sigma'$ is satisfiable **then**
14:                 $\sigma \leftarrow \sigma'$
15:                 **break**
16:     **return** GET_LEMMAS(σ)

This grammar guides the enumeration of candidate terms that may be used to instantiate y_i. As a result, the SyGuS enumerator generates both ε and non-ε terms as potential substitutions for y_i.

For each candidate term e generated by the enumerator, a new substitution σ' is formed by mapping y_i to the current term e in the initial substitution σ (line 12). The strategy then checks whether the negation of the body of the quantified formula under σ' is satisfiable in the current model (line 13). This satisfiability check serves as a semantic filter, pruning out terms that do not yield counterexamples. If the check succeeds, σ is updated and the enumeration process proceeds to the next quantified variable. Otherwise, it continues with the next candidate term. Once all quantified variables have been considered, our strategy returns both the final substitution σ and a set of lemmas, which include those introduced during the handling of ε terms and quantifier instantiations. These lemmas are subsequently added to the input formula F, extending it iteratively.

Augmented Grammar. Our strategy augments the base grammar used for term enumeration in MBQI-Enum by incorporating ε terms, allowing the generation of arbitrary formulas. Our approach is detailed in Algorithm 4.

MBQI-Enum constructs a base grammar for each universally quantified variable y_i in the formula q using a set of symbols S from the entire input formula, such as uninterpreted symbols and bound variables (line 4). A grammar G is a triple (g_0, G, R_G), where g_0 is the initial symbol, G is a set of nonterminal symbols with $g_0 \in G$, and R_G is a set of production rules. Each rule has the form $g ::= t$, where $g \in G$ and t is a term built from symbols in the set S, nonterminals in G, free variables, and symbols from the signature of a background theory $\mathcal{T}$. The symbols in S, along with the free variables and the theory symbols from $\mathcal{T}$, act as terminal symbols in the grammar. Each grammar generates

Algorithm 4 Augmented grammar

1: **function** MAKE_GRAMMAR(q, y_i)
2:     **let** F be the input formula
3:     $S \leftarrow \text{symbols}(F) \cup \text{symbols}(q) \cup \{y_{i+1}, \ldots, y_n\}$
4:     **let** $(g_0, G, R_G) \leftarrow$ MAKE_BASE_GRAMMAR(S, y_i)
5:     **if** choice is enabled **then**
6:         **let** x be a variable of type τ
7:         $(p_0, P, R_P) \leftarrow$ MAKE_PREDICATE_GRAMMAR(S, x, y_i)
8:         **extend** $G \leftarrow G \cup P$
9:         **extend** $R_G \leftarrow \{g_0 \leftarrow \varepsilon x.\ p_0\} \cup R_G \cup R_P$
10:     **return** (g_0, G, R_G)
11: **function** MAKE_PREDICATE_GRAMMAR(S, x, y_i)
12:     $S \leftarrow S \cup \{x\}$
13:     **let** $\bar{z} \leftarrow$ bound variables in y_i
14:     **define** p_0 nonterminals
15:     **let** $P \leftarrow \{\bar{z}\} \cup S \cup \{\mathsf{true}, \mathsf{false}, =, \neg, \wedge\}$
16:     **define** $R_P \leftarrow$ production rules
17:     **return** (p_0, P, R_P)

terms whose type matches the return type τ of y_i. For higher-order variables, the grammars include rules for generating λ-abstractions. In this case, the initial symbol g_0 expands to λ-terms. Whereas earlier versions of MBQI-Enum considered only terms in η-long β-normal form [19], the current cvc5 implementation allows partial application. The variables bound by λ-abstractions are terminal symbols in the grammar that generates the abstraction's body. For example, if y_i is a function variable of arity n whose arguments and result are all of the same type, the grammar includes rules such as

$$g_0 ::= \lambda x_1, \ldots, x_n.\ g_1 \qquad\qquad g_1 ::= x_1 \mid \cdots \mid x_n$$

To include ε terms of the form $\varepsilon x.\ \varphi$, our strategy extends the grammar as follows. For the nonterminal g_1 in the base grammar corresponding to y_i's return type τ, a variable x of type τ is introduced (line 6) and added to the symbol set S (line 12). A new grammar is then built to generate Boolean formulas over the extended symbol set $S \cup \{x\}$ (line 7). This predicate grammar is defined as the triple (p_0, P, R_P), where p_0 is the initial nonterminal symbols and P is a set of nonterminal symbols that includes p_0. The set R_P contains production rules of the form $p ::= t$, where $p \in P$ and t is a term. Each term t is built from symbols in S, nonterminals in P, variables that were bound in y_i but now appear free, and symbols from the signature of the Boolean background theory (line 15). These elements act as terminal symbols in the grammar. The new grammar includes the rules

$$p_0 ::= \mathsf{true} \mid \mathsf{false} \mid p_1 = p_1 \mid \neg\, p_0 \mid p_0 \wedge p_0 \qquad\qquad p_1 ::= x_1 \mid \cdots \mid x_n \mid x$$

The nonterminal p_0 directly defines the body of the ε term. We define a new production rule of the form $g_1\ ::=\ \varepsilon x.\ p_0$. The resulting grammar is constructed by extending the base grammar (g_0, G, R_G) with the predicate grammar (p_0, P, R_P) and the new production rule (lines 8–9). The final grammar is defined as $(p_0, G \cup P, \{g_1 ::= \varepsilon x.\ p_0\} \cup R_G \cup R_P)$. It can be used to enumerate both ε and non-ε terms and is thus more expressive than the original grammar. For our example, the final grammar consists of the rules

$$g_0 ::= \lambda x_1, \dots, x_n.\ g_1 \qquad g_1 ::= \varepsilon x.\ p_0 \mid x_1 \mid \cdots \mid x_n$$
$$p_0 ::= \mathsf{true} \mid \mathsf{false} \mid p_1 = p_1 \mid \neg\, p_0 \mid p_0 \wedge p_0 \qquad p_1 ::= x_1 \mid \cdots \mid x_n \mid x$$

Skolem Symbols. The most important part of the strategy is where we take formulas containing ε terms generated by the grammar, abstract these terms into fresh Skolem symbols, and generate corresponding lemmas with triggers to guide efficient higher-order quantifier instantiation. Our approach is detailed in Algorithm 5.

Algorithm 5 Lemmas for Skolem symbols

1: **function** GET_LEMMAS(σ)
2:     **let** $L \leftarrow \emptyset$
3:     **let** $\psi' \leftarrow \psi$
4:     **for all** $i \in \{1, \dots, m\}$ **do**
5:         **assume** $y_i\sigma = \lambda(x_1, \dots, x_n).\ t$ ▷ y_i is the variable being instantiated
6:         **let** $t' \leftarrow t$
7:         **for all** subterms e of t of the form $\varepsilon x : \tau.\ \varphi(\bar{z}, x)$ **do**
8:             **let** $\bar{z} = (z_1 : \tau_1, \dots, z_k : \tau_k) \subseteq \{x_1, \dots, x_n\}$ be φ's free variables
9:             **let** $h : (\tau_1, \dots, \tau_k) \to \tau$ be a fresh Skolem symbol
10:            **let** $\mathit{lemma} \leftarrow \forall \bar{z}.\ \neg\,(\exists x.\ \varphi(\bar{z}, x)) \vee \varphi(\bar{z}, \mathsf{h}\ \bar{z})$
11:            **extend** $L \leftarrow L \cup \{\mathit{lemma}\}$
12:            **replace** e **in** t' **with** $\mathsf{h}\ \bar{z}$
13:        $\psi' \leftarrow \psi'\{y_i \mapsto \lambda(x_1, \dots, x_n).\ t'\}$
14:    **let** $\mathit{inst\_lemma} \leftarrow (\forall y_1, \dots, y_m.\ \psi) \Longrightarrow \psi'$
15:    **extend** $L \leftarrow L \cup \{\mathit{inst\_lemma}\}$
16:    **return** L

Our strategy generalizes MBQI-Enum by iterating over the variables in $\forall y_1, \dots, y_m.\, \psi$ and incrementally building a fully instantiated formula before generating the instantiation lemmas. It then returns the set of generated lemmas L instead of only the substitutions.

Given a substitution $\sigma = \{y_i \mapsto \lambda x_1, \dots, x_n.\ t\}$ for each variable y_i, consider any subterm e of t of the form $\varepsilon x : \tau.\ \varphi(\bar{z}, x)$, where $\bar{z} = (z_1 : \tau_1, \dots, z_k : \tau_k)$, the free variables of e, are included in $\{x_1, \dots, x_n\}$. We replace all occurrences of e in t with an application of a fresh Skolem symbol h to the variables $\bar{z}$ (lines 5–10). In other words, the subterm $\varepsilon x.\ \varphi(\bar{z}, x)$ is converted to $\mathsf{h}\ \bar{z}$, where h has type $\tau_1 \to \cdots \to \tau_k \to \tau$. The introduction of these Skolems abstracts the ε binder into symbols understood by the rest of the SMT solver. Our strategy constructs the lemma $\forall \bar{z}.\ \neg\,(\exists x.\ \varphi(\bar{z}, x)) \vee \varphi(\bar{z}, \mathsf{h}\ \bar{z})$ (line 10) and adds it to the lemma set L. When later added to the set F, this lemma makes the strategy less incomplete while preserving F's satisfiability. It guarantees that if there exists an x satisfying $\varphi(\bar{z}, x)$, then $\mathsf{h}\ \bar{z}$ satisfies φ as well.

Additionally, we mark the Skolem application $\mathsf{h}\ \bar{z}$ as a *trigger* [12,13] in the lemma. This trigger guides E-matching, ensuring that instantiations are driven by terms that match the structure of the Skolem application, which is applied to exactly the same set of variables as in the original quantified formula.

After all ε-subterms in t have been abstracted, the modified substitution $y_i \mapsto \lambda x_1, \ldots, x_n.\ t'$, where t' is obtained by replacing the ε-subterms of t with the corresponding Skolem applications $\mathsf{h}\ \bar{z}$ (line 12), is applied to the partially instantiated body ψ of the formula q (line 13). This procedure is repeated for each variable y_i, yielding a fully instantiated formula ψ' in which all substitutions have been applied to the universally quantified variables and all Skolem symbols have been introduced. Finally, the instantiation lemma $(\forall y_1, \ldots, y_m.\ \psi) \Longrightarrow \psi'$ is generated (line 14) and added to the lemma set L. The complete set L is eventually returned and merged with the current set F.

Filtering. To avoid redundant instantiations, term filtering is applied during enumeration. Enumerated terms already present in the set of candidate terms are discarded, ensuring that terms equivalent under rewriting are considered only once. For terms of the form $\varepsilon x.\ \varphi$, the bound variable x must occur in the body φ. Terms in which x does not occur are excluded, to prevent the introduction of trivial instantiations. Since cvc5 already performs extensive theory-specific filtering by default, no additional filtering is required in our setting.

Example 6. We return to Example 1 and present it in more detail. Let $\mathcal{T}$ be the same theory as in the previous examples. Let u be an uninterpreted sort and let f be an uninterpreted function of type $u \to u$. Consider the conjecture

$$(\forall x, y.\ \mathsf{f}\ x = \mathsf{f}\ y \Longrightarrow x = y) \Longrightarrow \exists g.\ \forall z.\ g\ (\mathsf{f}\ z) = z$$

where g has type $u \to u$ and x, y, z have type u. The formula states that if f is injective, then it has a left inverse. We prove that the formula holds in $\mathcal{T}$ by contradiction. That is, we negate it, obtaining the following input problem

$$F = \{\forall x, y.\ \mathsf{f}\ x = \mathsf{f}\ y \Longrightarrow x = y,\ \forall g.\ \exists z.\ g\ (\mathsf{f}\ z) \neq z\}\ ,$$

and prove that F is unsatisfiable in $\mathcal{T}$.

MBQI-Enum generates a grammar for g using the set of symbols $S = \{\mathsf{f}\}$. The grammar consists of the rules

$$g_0 ::= \lambda y.\ g_1 \qquad\qquad g_1 ::= y \mid \mathsf{f}\ g_1$$

Based on this grammar, MBQI-Enum generates the substitutions $\{g \mapsto \lambda y.\ y\}$, $\{g \mapsto \lambda y.\ \mathsf{f}\ y\}$, $\{g \mapsto \lambda y.\ \mathsf{f}\ (\mathsf{f}\ y)\}$, $\ldots$ for g. Since this enumeration does not capture inverse functions, the SMT solver does not terminate.

Our strategy addresses this challenge by constructing a richer grammar. We introduce a variable x of type u and extend the symbol set to $S \cup \{x\}$. We then define the following grammar:

$$p_0 ::= \mathsf{true} \mid \mathsf{false} \mid p_1 = p_1 \mid \neg\, p_0 \mid p_0 \wedge p_0 \qquad\qquad p_1 ::= y \mid x \mid \mathsf{f}\ p_1$$

We extend the MBQI-Enum grammar by adding a rule to produce $\varepsilon x.\ p_0$. The result consists of the rules for p_0 and p_1 above and the following rules:

$$g_0 ::= \lambda y.\ g_1 \qquad\qquad g_1 ::= \varepsilon x.\ p_0 \mid y \mid \mathsf{f}\ g_1$$

Using this grammar, our strategy generates the term $\lambda y.\ \varepsilon x.\ y = \mathsf{f}\ x$, which denotes a left inverse of f, leading to the substitution $\{g \mapsto \lambda y.\ \varepsilon x.\ y = \mathsf{f}\ x\}$.

Corresponding to this term, we introduce a fresh Skolem symbol $\mathsf{h} : u \to u$. Our strategy asserts the following lemma:

$$\forall y.\ \neg\, (\exists x.\ y = \mathsf{f}\ x) \vee y = \mathsf{f}\ (\mathsf{h}\ y) \tag{ℓ_1}$$

The lemma states that for any y, if there exists an x such that $y = \mathsf{f}\ x$ (i.e., y is in the image of f), then applying f to $\mathsf{h}\ y$ must yield y. The original enumerated term $\lambda y.\ \varepsilon x.\ y = \mathsf{f}\ x$ is thus rewritten to the term $\lambda y.\ \mathsf{h}\ y$. The substitution $\sigma = \{g \mapsto \lambda y.\ \mathsf{h}\ y\}$ is considered. Next, the instantiation lemma

$$(\forall g.\ \exists z.\ g\ (\mathsf{f}\ z) \neq z) \Longrightarrow \exists z.\ \mathsf{h}\ (\mathsf{f}\ z) \neq z \tag{ℓ_2}$$

is generated by applying the substitution σ followed by β-reduction. The set of lemmas $L = \{(\ell_1), (\ell_2)\}$ is constructed. Each lemma in L is then added to set F. After that, Skolemization introduces a fresh symbol sk corresponding to the existentially quantified variable z. The lemma

$$(\exists z.\ \mathsf{h}\ (\mathsf{f}\ z) \neq z) \Longrightarrow \mathsf{h}\ (\mathsf{f}\ \mathsf{sk}) \neq \mathsf{sk} \tag{ℓ_3}$$

is also added to F.

The strategy proceeds to instantiate the universally quantified formulas in F. Quantified variables of type u are instantiated using terms from F. First, the strategy instantiates x and y in $\forall x, y.\ \mathsf{f}\ x = \mathsf{f}\ y \Longrightarrow x = y$ with sk and $\mathsf{h}\ (\mathsf{f}\ \mathsf{sk})$, respectively. Then it instantiates x and y in the lemma $\forall y.\ \neg\, (\exists x.\ y = \mathsf{f}\ x) \vee y = \mathsf{f}\ (\mathsf{h}\ y)$ with sk and $\mathsf{f}\ \mathsf{sk}$, respectively. These instantiations arise from E-matching, which matches each lemma's pattern against the terms already present in the equivalence classes maintained by the SMT solver. The instantiation lemmas

$$(\forall y.\ \neg\, (\exists x.\ y = \mathsf{f}\ x) \vee y = \mathsf{f}\ (\mathsf{h}\ y)) \Longrightarrow \mathsf{f}\ \mathsf{sk} \neq \mathsf{f}\ \mathsf{sk} \vee \mathsf{f}\ \mathsf{sk} = \mathsf{f}\ (\mathsf{h}\ (\mathsf{f}\ \mathsf{sk})) \tag{ℓ_4}$$

$$(\forall x, y.\ \mathsf{f}\ x = \mathsf{f}\ y \Longrightarrow x = y) \Longrightarrow \mathsf{f}\ \mathsf{sk} \neq \mathsf{f}\ (\mathsf{h}\ (\mathsf{f}\ \mathsf{sk})) \vee \mathsf{sk} = \mathsf{h}\ (\mathsf{f}\ \mathsf{sk}) \tag{ℓ_5}$$

are then added to F. At this point, the quantifier-free subsolver finds F unsatisfiable in $\mathcal{T}$ by determining that the conclusions of the lemmas ℓ_3, ℓ_4, and ℓ_5 are jointly unsatisfiable in $\mathcal{T}$. This implies that the conjecture is valid. ■

Example 7. We consider a simplified version of TPTP benchmark `SYO268^5`. Let u_1, u_2 be uninterpreted sorts, and let r be a function symbol of type $u_1 \to u_2 \to o$. Consider the conjecture

$$(\forall x.\ \exists y.\ \mathsf{r}\ x\ y) \Longrightarrow \exists g.\ \forall x.\ \mathsf{r}\ x\ (g\ x)$$

where g has type $u_1 \to u_2$, x has type u_1, and y has type u_2. The formula states that for every total relation r on $u_1 \times u_2$, there exists a corresponding function from u_1 to u_2. To prove the formula by contradiction, we negate it and obtain the input problem $F = \{\forall x.\ \exists y.\ \mathsf{r}\ x\ y,\ \forall g.\ \exists x.\ \neg\, \mathsf{r}\ x\ (g\ x)\}$.

MBQI-Enum generates a grammar for g using the set of symbols $S = \{\mathsf{r}\}$. The grammar consists of the rules

$$g_0 ::= \lambda y.\ g_1 \qquad\qquad g_1 ::= t$$

Based on this grammar, MBQI-Enum generates the substitution $\{g \mapsto \lambda y.\ t\}$, where t is a ground term of type u_2. As a result, the SMT solver terminates with an unknown status.

By contrast, with our extension, the solver behaves as follows. Let x be a variable of type u_2. We extend the symbol set to $S \cup \{x\}$. Consider the rules

$$\begin{aligned} p_0 &::= \mathsf{true} \mid \mathsf{false} \mid p_1 = p_1 \mid p_2 = p_2 \mid \mathsf{r}\ p_1\ p_2 \mid \neg\, p_2 \mid p_2 \wedge p_2 \\ p_1 &::= y \\ p_2 &::= x \end{aligned}$$

We augment the MBQI-Enum grammar by adding a rule to produce $\varepsilon x.\ p_0$. The resulting grammar consists of the rules for p_0, p_1, and p_2 above together with

$$g_0 ::= \lambda y.\ g_1 \qquad\qquad g_1 ::= \varepsilon x.\ p_0 \mid t$$

Using this grammar, our strategy generates the term $\lambda y.\ \varepsilon x.\ \mathsf{r}\ y\ x$, leading to the substitution $\{g \mapsto \lambda y.\ \varepsilon x.\ \mathsf{r}\ y\ x\}$.

Corresponding to this term, we introduce a fresh Skolem symbol $\mathsf{h} : u_1 \to u_2$. Our strategy asserts the following lemma:

$$\forall y.\ \neg\, (\exists x.\ \mathsf{r}\ y\ x) \vee \mathsf{r}\ y\ (\mathsf{h}\ y) \tag{ℓ_1}$$

The original enumerated term $\lambda y.\ \varepsilon x.\ \mathsf{r}\ y\ x$ is thus rewritten to the term $\lambda y.\ \mathsf{h}\ y$. The substitution $\sigma = \{y \mapsto \lambda y.\ \mathsf{h}\ y\}$ is considered. Next, the instantiation lemma

$$(\forall g.\ \exists x.\ \neg\, \mathsf{r}\ x\ (g\ x)) \Longrightarrow \exists x.\ \neg\, \mathsf{r}\ x\ (\mathsf{h}\ x) \tag{ℓ_2}$$

is generated by applying the substitution σ followed by β-reduction. The set of lemmas $L = \{(\ell_1), (\ell_2)\}$ is constructed and then merged with F. Afterward, Skolemization introduces a fresh symbol sk corresponding to the existentially quantified variable x. The lemma

$$(\exists x.\ \neg\, \mathsf{r}\ x\ (\mathsf{h}\ x)) \Longrightarrow \neg\, \mathsf{r}\ \mathsf{sk}\ (\mathsf{h}\ \mathsf{sk}) \tag{ℓ_3}$$

is added to F.

The strategy proceeds to instantiate the universally quantified formulas of F. Quantified variables of nonfunction type are instantiated using terms from F. First, the strategy instantiates x and y in the lemma $\forall y.\ \neg\,(\exists x.\ \mathsf{r}\ y\ x) \lor \mathsf{r}\ y\ (\mathsf{h}\ y)$ with $\mathsf{h}\ \mathsf{sk}$ and sk, respectively. Then it instantiates x and y in the axiom $\forall x.\ \exists y.\ \mathsf{r}\ x\ y$ with sk, and $\mathsf{h}\ \mathsf{sk}$, respectively. These instantiations arise from E-matching, which matches each lemma's pattern against the terms already present in the equivalence classes maintained by the solver. The instantiation lemmas

$$(\forall y.\ \neg\,(\exists x.\ \mathsf{r}\ y\ x) \lor \mathsf{r}\ y\ (\mathsf{h}\ y)) \Longrightarrow \neg\,\mathsf{r}\ \mathsf{sk}\ (\mathsf{h}\ \mathsf{sk}) \lor \mathsf{r}\ \mathsf{sk}\ (\mathsf{h}\ \mathsf{sk}) \qquad (\ell_4)$$
$$(\forall x.\ \exists y.\ \mathsf{r}\ x\ y) \Longrightarrow \mathsf{r}\ \mathsf{sk}\ (\mathsf{h}\ \mathsf{sk}) \qquad (\ell_5)$$

are added to F. At this point, the quantifier-free subsolver finds F unsatisfiable. Hence the conjecture is valid. ■

4 Implementation

We developed our strategy as an extension of cvc5's implementation of MBQI-Enum. Our extension refines the selection of grammars for term enumeration, introduces a mechanism for managing choice terms generated by these grammars, and incorporates a filtering step to discard redundant or unsuitable terms.

For each quantified variable, MBQI-Enum constructs a SyGuS grammar that specifies the space of candidate instantiations (Algorithm 4, line 4). Our implementation augments these grammars with terminal rules for generating choice terms. Specifically, for each nonterminal type that qualifies, we add a corresponding expression of the form $\varepsilon x.\ \varphi$, which selects a value satisfying a predicate over that type (lines 5–9). This allows the enumerator to synthesize Skolem-style instantiations directly from the grammar.

We use the existing implementation of the fast SyGuS-based enumerator for term generation. This enumerator considers only terms that are unique up to rewriting [31]. During the enumeration process, ε terms are constructed within cvc5's internal abstract syntax tree representation, thereby exploiting cvc5's built-in rewriter. For example, $\varepsilon x.\ x = t$, where x does not occur free in t, is rewritten to t.

Beyond these rewrite-based simplifications, our implementation applies additional heuristics for filtering ε terms. In particular, it explicitly discards any term of the form $\varepsilon x.\ \varphi$, where x does not occur free in φ. These heuristics are implemented via a callback from the SyGuS enumerator.

The enumerator constructs ε terms that are immediately purified by our implementation. Recall that each ε term is replaced with a fresh Skolem symbol, and an accompanying lemma is generated to relate the Skolem to the original ε term (Algorithm 5, lines 6–10). For example, given a term $\varepsilon x.\ \varphi$, the procedure introduces a fresh Skolem h and asserts that either $\varphi(\mathsf{h})$ holds or the quantified condition that motivated the ε term is already satisfied. An important heuristic is that our approach designates the Skolem application itself as a trigger for the generated quantified formula, thereby ensuring that instantiations during solving are guided by the intended terms.

During instantiation, terms containing ε terms are first normalized through the purification and filtering procedures described above before being considered as candidate substitutions. The generalized strategy, shown in Algorithm 3, iteratively builds substitutions by enumerating terms from the augmented grammars (lines 6–12), tests them in the current model (line 13), and commits suitable assignments (line 14). Notably, unlike the standard MBQI-Enum procedure, our strategy returns not only substitutions but also auxiliary lemmas (line 16).

5 Evaluation

We extensively evaluated our cvc5 implementation of the extended MBQI-Enum on higher-order benchmarks.

Since the publication of Kondylidou et al. [19], MBQI-Enum has been developed further independently of our work on choice. In our evaluation, by "original MBQI-Enum" we mean the most recent version of the strategy excluding our modifications related to ε. This strategy includes the following enhancements over Kondylidou et al.:

- MBQI-Enum now supports partial applications, allowing the strategy to consider candidate terms that partially apply functions during enumeration.
- Timeouts have been added to subsolver invocations used for checking candidate instantiations, preventing individual candidates from stalling the overall SMT loop. In addition, there is an option that completely prohibits nested MBQI calls, which avoids potential nontermination caused by nested subsolver invocations.
- MBQI-Enum now incorporates a guess-and-test loop with a progress-aware instantiation cache: When a proposed instantiation fails to refute the current model, it is cached to avoid rechecking the same substitution later, thereby ensuring that the strategy always makes progress.
- We now incorporate a lookahead approach that preemptively considers Skolem symbols for quantified formulas that are not introduced during preprocessing. These symbols are incorporated into the SyGuS grammars, thus enriching the set of terms we consider for instantiation with MBQI-Enum.

Setup. We denote our configuration of MBQI-Enum extended with choice by cvc5[C]. We compare our strategy against several established quantifier instantiation strategies in cvc5: cvc5[s], which uses SyQI [27]; cvc5[m], which implements MBQI [15]; cvc5[M], which implements the original MBQI-Enum [19]; cvc5[hoelim], which eagerly rewrites higher-order constraints into first-order form and applies full quantifier saturation; and cvc5[foinst], which relies on exhaustive first-order-style instantiation and axiomatic handling of higher-order applications.

Table 1. Extended MBQI-Enum vs. other provers on TPTP TH0 benchmarks

	Satallax	Vampire	Zipperposition	cvc5[C]
Satisfiable	196	14	0	**204**
Unsatisfiable	2162	**2303**	2083	2178
Total	2358	2317	2083	**2382**
Unknown	15	16	0	154
Timeouts	1384	1424	1674	1221

We also include a comparison with the state-of-the-art provers Satallax [10], Vampire [7], and Zipperposition [40]. Satallax was run with its default settings. Vampire was run in portfolio mode, as its higher-order reasoning configuration automatically enables this mode and disabling it makes the prover much less competitive. For Zipperposition, we enabled several options that were extensively evaluated by Bozec and Blanchette [9] and shown to achieve the best overall performance on higher-order benchmarks.

All experiments were conducted on a machine with a 40-core Intel Xeon Silver 4114 CPU running at 2.20 GHz, equipped with 192 GB of RAM and running Debian 12 (Bookworm). Each benchmark was executed with a timeout of 60 seconds.

Benchmarks. The experiments were carried out on monomorphic higher-order problems (TH0) from version 9.1.0 of the TPTP library [39]. The benchmark set consists of 3757 problems. From the 3962 TH0 problems, we excluded 205 benchmarks that one or more systems could not parse (e.g., because of the use of interpreted arithmetic).

To support reasoning with choice, we implemented external parser support for the TPTP operators for ε and the definite description operator ι. Although the TPTP library contains very few problems with ε and ι, they might arise in users' problems.

Results. The results are shown in Tables 1 and 2, where bold indicates the most successful system. Notably, our approach achieves the highest total count of solved benchmarks, surpassing the nearest competitor by 24 solved problems and the baseline MBQI-Enum by 23.

Table 1 compares our strategy (cvc5[C]) with external state-of-the-art provers. Our approach achieves the highest total number of solved benchmarks, with 2382 problems, surpassing Satallax, Vampire, and Zipperposition. It outperforms the nearest competitor by 24 benchmarks. In particular, our strategy solves 204 satisfiable problems, while Satallax solves 196, Vampire solves 14, and Zipperposition solves none. For unsatisfiable problems, Vampire performs the best, solving 2303 benchmarks, whereas our strategy solves 2178, Satallax 2162, and Zipperposition 2083. Overall, our strategy demonstrates strong performance across

Table 2. Extended MBQI-Enum vs. other cvc5 strategies on TPTP TH0 benchmarks

	cvc5[hoelim]	cvc5[foinst]	cvc5[s]	cvc5[m]	cvc5[M]	cvc5[C]
Satisfiable	14	37	88	188	204	**204**
Unsatisfiable	2082	2126	1657	2074	2155	**2178**
Total	2096	2163	1745	2262	2359	**2382**
Unknown	0	125	45	289	158	154
Timeouts	1661	1469	1967	1206	1240	1221

both satisfiable and unsatisfiable problems compared with the state-of-the-art provers.

Table 2 compares cvc5[C] with other cvc5 quantifier instantiation strategies. Our configuration solves 23 unsatisfiable benchmarks that the baseline MBQI-Enum cannot, with no benchmarks lost, while solving the same number of satisfiable problems (204). This shows that extending MBQI-Enum with choice consistently improves solver performance. Overall, when compared against all other available quantifier instantiation strategies, our strategy achieves the strongest performance, demonstrating that it is considerably more effective on higher-order benchmarks than any other individual technique in cvc5. It may seem as though 23 additional benchmarks are not that many; however, we should bear in mind that the vast majority of higher-order TPTP benchmarks do not contain choice and can likely be solved without reasoning about choice.

In addition, our technique solves several problems not solved by its competitors: 143 problems not solved by Vampire, 296 not solved by Zipperposition, 235 not solved by Satallax, and 21 not solved by any other cvc5 strategy. When considering all systems, cvc5[C] solves exactly one benchmark that no other system can solve.

6 Related Work

Besides MBQI-Enum, other quantifier instantiation strategies work on higher-order problems, notably E-matching, MBQI, and SyQI. In higher-order E-matching, quantified variables are instantiated by matching their patterns against ground terms available in the current context—that is, the set of ground terms introduced during proof search. This technique has been extended to support higher-order features, such as functional arguments [3]. Higher-order MBQI constructs candidate models that provide interpretations for functions. This enables the solver to generate ground instantiations of quantified formulas in an attempt to refute the current model [15]. Higher-order SyQI further improves the instantiation process by selecting candidate terms using syntactic templates and heuristics derived from the problem's structure, guided by a fixed grammar [27]. None of these approaches provides special support for Hilbert's choice.

Other higher-order provers support Hilbert's choice in various ways. Satallax implements specialized tableau inference rules that directly reason about ε

terms in a sound and refutationally complete way [2]. Leo-III integrates ε into its extensional higher-order paramodulation calculus and supports Henkin semantics with choice [38]. Zipperposition includes a dedicated choice axiom within its higher-order superposition calculus, allowing it to reason about ε terms as needed [6]. Vampire supports Hilbert's choice either through a Leo-III-style inference or by introducing the choice axiom [7].

Hilbert's choice is also present in proof assistants such as Isabelle/HOL [28], Lean [24], and Rocq [36] (formerly known as Coq). In Isabelle/HOL, the choice operator is introduced as an axiom as part of the system's core libraries. In systems based on dependent type theory, such as Lean and Rocq, choice is introduced via classical axioms, which should be avoided in constructive definitions and proofs. Even in Isabelle/HOL, some users, such as Paulson [29], prefer to avoid choice:

> Pragmatists may argue that in verification nobody cares whether choice is used or not. However, pragmatists should be concerned that reasoning about ε-terms is tricky.

Reasoning about choice is indeed tricky, but cvc5 can do it reasonably well now.

7 Conclusion

We presented an SMT quantifier instantiation strategy that extends the MBQI-Enum strategy so as to reason natively about Hilbert's choice operator. It can be used to solve problems that contain Hilbert's choice but also problems that do not (cf. Example 1). We implemented the approach in the SMT solver cvc5. The empirical results on higher-order TPTP benchmarks show that our strategy is clearly superior to the original MBQI-Enum.

We believe there is potential for further improving the strategy. We noticed that the ε terms we need often occur too late in the term enumeration (which is by increasing term size), since they tend to be relatively complex. Better term enumeration heuristics and term filtering could help the SMT solver generate the desired terms earlier—before the timeout.

Acknowledgments. We thank Mark Summerfield and the anonymous reviewers for their helpful comments. We also thank Haniel Barbosa, who provided many comments on an earlier draft.

This research was cofunded by the European Union (ERC, Nekoka, 101083038). Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.

This research was also partially funded by the Defense Advanced Research Projects Agency (DARPA) under contract FA8750-24-2-1001. Any opinions, findings, and conclusions or recommendations expressed here are those of the authors and do not necessarily reflect the views of DARPA.

Data-Availability Statement. The artifact associated with this paper, including the software and scripts used for experiments, is publicly available on Zenodo [20]. The raw evaluation data are also available on Zenodo [21]. The instructions on the first Zenodo page [20] explain how to reproduce the experimental evaluation.

References

1. Andrews, B.: An Introduction to Mathematical Logic and Type Theory: To Truth Through Proof, APLS, vol. 27. Springer (2022)
2. Backes, J., Brown, C.E.: Analytic tableaux for higher-order logic with choice. Journal of Automated Reasoning **47**, 451–479 (2011)
3. Barbosa, H., Barrett, C., Brain, M., Kremer, G., Lachnitt, H., Mann, M., Mohamed, A., Mohamed, M., Niemetz, A., Nötzli, A., Ozdemir, A., Preiner, M., Reynolds, A., Sheng, Y., Tinelli, C., Zohar, Y.: cvc5: A versatile and industrial-strength SMT solver. In: Fishman, D., Rosu, G. (eds.) TACAS 2022. LNCS, vol. 13243, pp. 415–442. Springer (2022)
4. Barbosa, H., Reynolds, A., Ouraoui, D.E., Tinelli, C., Barrett, C.: Extending SMT solvers to higher-order logic. In: Fontaine, P. (ed.) CADE 2019. LNCS, vol. 11716, pp. 35–54. Springer (2019)
5. Barrett, C.W., Conway, C.L., Deters, M., Hadarean, L., Jovanovic, D., King, T., Reynolds, A., Tinelli, C.: CVC4. In: Gopalakrishnan, G., Qadeer, S. (eds.) CAV 2011. LNCS, vol. 6806, pp. 171–177. Springer (2011)
6. Bentkamp, A., Blanchette, J., Tourret, S., Vukmirović, P.: Superposition for higher-order logic. Journal of Automated Reasoning **67** (2023)
7. Bhayat, A., Suda, M.: A higher-order Vampire (short paper). In: Benzmüller, C., Heule, M.J.H., Schmidt, R.A. (eds.) IJCAR 2024. LNCS, vol. 14739, pp. 75–85. Springer (2024)
8. Bouton, T., de Oliveira, D.C.B., Déharbe, D., Fontaine, P.: veriT: An open, trustable and efficient SMT-solver. In: Schmidt, R.A. (ed.) CADE 2009. LNCS, vol. 5663, pp. 151–156. Springer (2009)
9. Bozec, T., Blanchette, J.: Iterative monomorphisation. In: Thiemann, R., Weidenbach, C. (eds.) FroCoS 2025. LNCS, vol. 15979, pp. 269–286. Springer (2025)
10. Brown, C.E.: Satallax: An automatic higher-order prover. Journal of Automated Reasoning **7364**, 111–117 (2012)
11. Church, A.: A formulation of the simple theory of types. Journal of Symbolic Logic **5**, 56–68 (1940)
12. Detlefs, D., Nelson, G., Saxe, J.B.: Simplify: A theorem prover for program checking. Journal of the ACM **52**(3), 365–473 (2005)
13. Dross, C., Conchon, S., Paskevic, A.: Reasoning with triggers. In: Fontaine, P., Goel, A. (eds.) SMT 2012. EPiC Series in Computing, vol. 20, pp. 22–31 (2012)
14. Ge, Y., Barrett, C.W., Tinelli, C.: Solving quantified verification conditions using satisfiability modulo theories. Annals of Mathematics and Artificial Intelligence **55**(1-2), 101–122 (2009)
15. Ge, Y., de Moura, L.: Complete instantiation for quantified formulas in satisfiabiliby modulo theories. In: Bouajjani, A., Maler, O. (eds.) CAV 2009. LNCS, vol. 5643, pp. 306–320. Springer (2009)
16. Gordon, M.J.C.: Introduction to the HOL system. In: Archer, M., Joyce, J.J., Levitt, K.N., Windley, P.J. (eds.) HOL 1991. pp. 2–3. IEEE (1991)

17. Hilbert, D., Ackermann, W.: Grundzüge der Theoretischen Logik, Grundlehren der mathematischen Wissenschaften, vol. 27. Springer (1928)
18. Hilbert, D., Bernays, P.: Grundlagen der Mathematik I, Grundlehren der mathematischen Wissenschaften, vol. 40. Springer (1934)
19. Kondylidou, L., Reynolds, A., Blanchette, J.: Augmenting model-based instantiation with fast enumeration. In: Gurfinkel, A., Heule, M. (eds.) TACAS 2025. LNCS, vol. 15696, pp. 85–103. Springer (2025)
20. Kondylidou, L., Reynolds, A., Blanchette, J., Tinelli, C.: Artifact for paper "Enumerating choice terms in model-based quantifier instantiation". In: TACAS 2026. Zenodo (2026), https://doi.org/10.5281/zenodo.18304537
21. Kondylidou, L., Reynolds, A., Blanchette, J., Tinelli, C.: Enumerating choice terms in model-based quantifier instantiation [data set]. In: TACAS 2026. Zenodo (2026), https://doi.org/10.5281/zenodo.18212605
22. de Moura, L., Bjørner, N.: Efficient E-matching for SMT solvers. In: Pfenning, F. (ed.) CADE 2007. LNCS, vol. 4603, pp. 183–198. Springer (2007)
23. de Moura, L., Bjørner, N.: Z3: An efficient SMT solver. In: Ramakrishnan, C.R., Rehof, J. (eds.) TACAS 2008. LNCS, vol. 4963, pp. 337–340. Springer (2008)
24. de Moura, L., Ullrich, S.: The Lean 4 theorem prover and programming language. In: Platzer, A., Sutcliffe, G. (eds.) CADE 2021. LNCS, vol. 12699, pp. 625–635. Springer (2021)
25. Niemetz, A., Preiner, M.: Bitwuzla. In: Enea, C., Lal, A. (eds.) CAV 2023. LNCS, vol. 13965, pp. 3–17. Springer (2023)
26. Niemetz, A., Preiner, M., Biere, A.: Boolector 2.0. Journal on Satisfiability, Boolean Modeling and Computation **9**, 53–58 (2014)
27. Niemetz, A., Preiner, M., Reynolds, A., Barrett, C., Tinelli, C.: Syntax-guided quantifier instantiation. In: Groote, J.F., Larsen, K.G. (eds.) TACAS 2021. LNCS, vol. 12652, pp. 145–163. Springer (2021)
28. Nipkow, T., Paulson, L.C., Wenzel, M.: Isabelle/HOL: A Proof Assistant for Higher-Order Logic, LNCS, vol. 2283. Springer (2002)
29. Paulson, L.C.: Defining functions on equivalence classes. In: Guruswami, V., Glazener, T. (eds.) CoRR 2019. vol. 7, pp. 658–675. Association for Computing Machinery (2019)
30. Reynolds, A., Barbosa, H., Fontaine, P.: Revisiting enumerative instantiation. In: Beyer, D. (ed.) TACAS 2018. LNCS, vol. 10806, pp. 112–131. Springer (2018)
31. Reynolds, A., Barbosa, H., NÖtzli, A., Barrett, C.W., Tinelli, C.: cvc4sy: Smart and fast term enumeration for syntax-guided synthesis. In: Dillig, I., Tasiran, S. (eds.) CAV 2019, Part II. LNCS, vol. 11562, pp. 74–83. Springer (2019)
32. Reynolds, A., Blanchette, J.C., Cruanes, S., Tinelli, C.: Model finding for recursive functions in SMT. In: Olivetti, N., Tiwari, A. (eds.) IJCAR 2016. LNCS, vol. 9706, pp. 133–151. Springer (2016)
33. Reynolds, A., Deters, M., Kuncak, V., Tinelli, C., Barrett, C.: Counterexample-guided quantifier instantiation for synthesis in SMT. In: Kroening, D., Păsăreanu, C.S. (eds.) CAV 2015. LNCS, vol. 9207, pp. 198–216. Springer (2015)
34. Reynolds, A., Tinelli, C., Goel, A., Krstić, S.: Finite model finding in SMT. In: Sharygina, N., Veith, H. (eds.) CAV 2013. LNCS, vol. 8044, pp. 640–655. Springer (2013)
35. Reynolds, A., Tinelli, C., de Moura, L.: Finding conflicting instances of quantified formulas in SMT. In: FMCAD 2014. pp. 195–202. IEEE (2014)
36. Rocq Development Team: Rocq Prover. https://zenodo.org/records/15149629 (2025), Zenodo

37. Simpson, S.G.: The epsilon calculus and conservative extensions. Notre Dame Journal of Formal Logic **25**(3), 193–202 (1984)
38. Steen, A.: Extensional Paramodulation for Higher-Order Logic and its Effective Implementation Leo-III. Ph.D. thesis, Freie Universität Berlin (2020)
39. Sutcliffe, G.: The TPTP problem library and associated infrastructure—from CNF to TH0, TPTP v6.4.0. Journal of Automated Reasoning **59**, 483–502 (2017)
40. Vukmirović, P., Bentkamp, A., Blanchette, J., Cruanes, S., Nummelin, V., Tourret, S.: Making higher-order superposition work. Journal of Automated Reasoning **66**, 541–564 (2022)