

# Automatic Termination Verification for Higher-Order Functional Programs\*

Takuya Kuwahara<sup>1</sup>, Tachio Terauchi<sup>2</sup>, Hiroshi Unno<sup>3</sup>, and Naoki Kobayashi<sup>4</sup>

<sup>1</sup> University of Tokyo  
kuwahara@is.s.u-tokyo.ac.jp  
<sup>2</sup> Nagoya University  
terauchi@is.nagoya-u.ac.jp  
<sup>3</sup> University of Tsukuba  
uhiro@cs.tsukuba.ac.jp  
<sup>4</sup> University of Tokyo  
koba@is.s.u-tokyo.ac.jp

**Abstract.** We present an automated approach to verifying termination of higher-order functional programs. Our approach adopts the idea from the recent work on termination verification via transition invariants (a.k.a. binary reachability analysis), and is fully automated. Our approach is able to soundly handle the subtle aspects of higher-order programs, including partial applications, indirect calls, and ranking functions over function closure values. In contrast to the previous approaches to automated termination verification for functional programs, our approach is sound and complete, relative to the soundness and completeness of the underlying reachability analysis and ranking function inference. We have implemented a prototype of our approach for a subset of the OCaml language, and we have confirmed that it is able to automatically verify termination of some non-trivial higher-order programs.

## 1 Introduction

Recent years have witnessed a dramatic progress in automated verification of higher-order functional programs [27,31,24,11,16,33,37]. The line of work takes the ideas from the recent advances in the verification of non-functional programs, such as predicate abstraction, counterexample-guided abstraction refinement, and interpolation [1,9,21,10], to the verification of higher-order functional programs by way of refinement (dependent) types [36] and higher-order model checking [23,14].

However, except for the case when the base-type data is of a finite domain [15,20], the above line of work (to the extent of software model checking techniques for higher-order programs) has been limited to the verification of safety (i.e., reachability) properties. In particular, it is unable to verify liveness properties, such as termination.

For automated termination verification of higher-order programs, popular methods have been based on size-change termination [12,29,28] or TRS (term rewriting systems) techniques [6]. (Besides them, Xi [35] has proposed termination analysis based on dependent types, but his technique is not fully automated in the sense that users

---

\* This work was supported by MEXT Kakenhi 23220001, 23700026, 25280023, and 25730035.

```

let rec app f x () =
  if x>0 then app f (x-1) () else f x () in
let id () = () in
let rec g x = if x=0 then id else app g x in
let t = * in g t ()

```

**Fig. 1.** A non-terminating higher-order program

have to provide dependent types of recursive functions as witness of termination.) The current methods based on those approaches are not completely satisfactory, especially in terms of precision. Roughly, these techniques first construct a finite graph (called a static call graph [12,29,28] or a termination graph [6]) that over-approximates certain dependencies on termination, and then use techniques for first-order programs [19,7] to show that there is no cyclic dependency. In these two-phase approaches to termination, information is often lost in the first phase; see Section 5 for more details.

In the present paper, we follow an approach based on transition invariants [26,4,5], and extend it to deal with higher-order functional programs. The transition-invariant-based approach has emerged as a powerful technique for verifying termination of first-order imperative programs [26,4,5]. The technique iteratively reduces the termination verification problem to the problem of checking the *binary* reachability of program transition relations. It then delegates the binary reachability checking to a reachability checker by encoding the problem as a plain reachability problem via a program transformation. Advantages of this approach are that termination arguments can be flexibly adjusted for each program by choosing an appropriate binary reachability relation, and that precise flow information can be taken into account in the plain reachability verification phase. The latter advantage is particularly important since the termination property often depends on safety properties. For example, consider the program:

```
let f x = if p(x) then () else loop_forever()
```

Then, a call of  $f$  terminates if and only if  $p(x)$  is true, and the latter condition can be checked during the reachability verification. In the higher-order case, termination verification would be even more complicated since the condition can be passed as a parameter of  $f$ . This shows the advantage of reducing termination verification to reachability verification, where all the relevant information (such as value-dependent control flow and size change) is put together and precisely analyzed by taking advantage of the recent advance of reachability verification tools for higher-order programs.

The extension of the transition invariant-based technique for higher-order programs is non-trivial. To see why, let us consider the OCaml program  $P_0$  shown in Figure 1. (Here,  $*$  denotes a non-deterministic integer.) The program is non-terminating for any non-deterministic choice of  $t$  such that  $t < 0$ . For example, for  $t = -1$ , the program exhibits the following infinite reduction sequence.

$$g(-1)() \rightarrow^* \text{app } g(-1)() \rightarrow^* g(-1)() \rightarrow^* \text{app } g(-1)() \rightarrow^* \dots$$

Note here that  $g$  is passed to and indirectly called by  $\text{app}$ , and there is no direct call to  $g$  in the definition of  $g$ . Moreover,  $g$  itself does not (totally) call  $\text{app}$  but returns a

partially applied closure of the form  $\text{app } g \ n$ . Therefore, a termination verifier must soundly handle indirect calls and function closures to avoid incorrectly reporting that the program is terminating.

For a terminating example, let us consider a variation  $P_1$  of the above program, obtained by replacing the branching condition  $x = 0$  in  $g$  with  $x \leq 0$ . To prove that  $P_1$  is terminating for any non-deterministically chosen  $t$  (and any non-deterministic choice of the integers chosen inside  $g$ ), we need to know that the sequence of the third arguments passed to the recursive calls to  $\text{app}$  is strictly decreasing and is bounded below by 0.

Our technique consists of: (i) a method to find an appropriate (disjunctively) well-founded relation, including those over function closure values, that over-approximates the binary reachability relation (the relation between two recursive function calls), (ii) a program transformation that reduces the binary reachability problem to the plain reachability problem, and (iii) a plain reachability analysis for higher-order programs. For (iii), we employ off-the-shelf reachability verification tools for higher-order programs [27,31,11,16,33,37]. For (i), we adopt the previous technique [33] for automatically inserting implicit integer parameters that represent information about function closures. We can then adopt the existing techniques to find ranking functions (on integer arguments) from counterexamples [25,3,5]. The most subtle part is (ii): how to reduce the binary reachability analysis to plain reachability analysis. Actually, Ledesma-Garza and Rybalchenko [18] has recently tackled this problem, but (as admitted in [18], Section 8), their solution does not work quite well in the presence of partial applications and indirect function calls. In fact, their method cannot properly deal with the programs  $P_0$  and  $P_1$  above (cf. Section 5 of this paper for more details). By contrast, the reduction from binary reachability to plain reachability presented in this paper is *sound* and *complete*.

Our contributions are: (i) The first sound approach to the termination verification of higher-order functional programs that is based on the transition invariant / binary reachability technique. The approach is also complete relative to the completeness of the backend reachability checker and the ranking function inference process. A notable aspect of our approach is an inference of ranking functions over closure values via the automatic inference of implicit parameter instantiations. (ii) A prototype implementation to show the effectiveness of the proposed approach.

The rest of the paper is organized as follows. We define the target functional language of termination verification in Section 2. Section 3 formalizes our termination verification method. Section 4 reports on a preliminary implementation and experiment results. We compare our method with related work in Section 5 and conclude the paper in Section 6. The extended report [17] contains extra materials and proofs of the theorems.

## 2 Preliminaries

In this section, we introduce a higher-order functional language  $L$ , which is the target of our termination verification. Figure 2 shows the syntax of  $L$ . Here,  $\tilde{x}$  is an abbreviation for a (non-empty) variable sequence  $x_1 \ x_2 \ \dots \ x_k$ . The meta-variables  $f$ ,  $x$ ,  $c$  and  $op$  range over the sets of function symbols, variables, constants, and binary operators respectively. We write  $|\tilde{x}|$  for the length of  $\tilde{x}$ . The arity of function  $f_i$ , written  $\text{arity}(f_i)$ , is the number of formal parameters, i.e.,  $|\tilde{x}_i|$  in the function definition  $f_i \ \tilde{x}_i = e_i$ .

*Programs*  $P ::= \{f_1 \tilde{x}_1 = e_1, \dots, f_n \tilde{x}_n = e_n\}$   
*Expressions*  $e ::= v \mid x \mid \text{let } x = e_1 \text{ in } e_2 \mid e_1 \text{ op } e_2 \mid e_1 \ e_2 \mid \text{if } e_1 \text{ then } e_2 \text{ else } e_3 \mid *_{\text{int}}$   
*Values*  $v ::= c \mid f \mid f \tilde{v}$  (where  $|\tilde{v}| < \text{arity}(f)$ )

Fig. 2. Syntax of  $L$ 

*Eval. contexts*  $E ::= [] \mid E \text{ op } e \mid v \text{ op } E \mid E e \mid v E \mid \text{let } x = E \text{ in } e$   
 $\mid \text{if } E \text{ then } e_1 \text{ else } e_2$

$$E [c_1 \text{ op } c_2] \rightarrow_P E [[\text{op}]](c_1, c_2)$$

$$E [\text{let } x = v \text{ in } e] \rightarrow_P E [[v/x] e]$$

$$E [\text{if true then } e_1 \text{ else } e_2] \rightarrow_P E [e_1]$$

$$E [\text{if false then } e_1 \text{ else } e_2] \rightarrow_P E [e_2]$$

$$\frac{n \in \mathbb{Z}}{E [*_{\text{int}}] \rightarrow_P E [n]}$$

$$\frac{f \tilde{x} = e \in P \quad |\tilde{x}| = |\tilde{v}|}{E [f \tilde{v}] \rightarrow_P E [[\tilde{v}/\tilde{x}] e]}$$
Fig. 3. Operational semantics of  $L$ 

We assume that the set of constants includes  $()$ , `true`, `false` and (unbounded) integers, and that the set of binary operators includes comparators:  $>$ ,  $<$ ,  $\geq$ ,  $\leq$ ,  $=$  and boolean operators:  $\wedge$ ,  $\vee$ ,  $\Rightarrow$ . We also assume that a program has a special one-arity function named `main` that does not occur in the body of a function definition.

A program is a set of top-level function definitions; note that this does not lose generality because any functional program can be transformed to this form via  $\lambda$ -lifting. In the definition of the expression,  $*_{\text{int}}$  evaluates to some integer in a non-deterministic manner. Note that the non-deterministic boolean,  $*_{\text{bool}}$ , can be defined as  $*_{\text{int}} = 0$ .

The set of evaluation contexts and the reduction relation are given in Figure 3. Here,  $[[\text{op}]]$  denotes the binary operation on constants denoted by  $\text{op}$ . Note that the evaluation is call-by-value and non-deterministic (because of  $*_{\text{int}}$ ). We write  $\rightarrow_P^*$  for the reflexive and transitive closure of  $\rightarrow_P$ , and  $\rightarrow_P^+$  for the transitive closure of  $\rightarrow_P$ . When it is clear from the context, we omit the subscript  $P$  from the relations.

*Example 1.* The following program `fib` chooses an integer  $n$  non-deterministically and computes the  $n$ -th Fibonacci number.

$$\left\{ \begin{array}{l} \text{fib } n = \text{if } n < 2 \text{ then } 1 \text{ else fib}(n-1) + \text{fib}(n-2), \\ \text{main } () = \text{fib } *_{\text{int}} \end{array} \right\}$$

The following is a possible reduction sequence of the program:

$$\begin{aligned} \text{main } () &\rightarrow_{\text{fib}} \text{fib } *_{\text{int}} \rightarrow_{\text{fib}} \text{fib}(2) \rightarrow_{\text{fib}}^* \text{fib}(1) + \text{fib}(0) \\ &\rightarrow_{\text{fib}}^* 1 + \text{fib}(0) \rightarrow_{\text{fib}}^* 1 + 1 \rightarrow_{\text{fib}}^* 2 \end{aligned}$$

For readability, we often write a program in the OCaml-like syntax as shown below:

```
let rec fib n = if n < 2 then 1 else fib (n-1) + fib (n-2)
let main () = fib *int
```

*Example 2.* The following program `indirect` is a simplified variant of the program  $P_1$  in Section 1, obtained by removing the then branch of `app` and moving the decrement operation (i.e.,  $x - 1$ ) to inside `g`.

$$\left\{ \begin{array}{l} \text{app f x u} = \text{f x u} \\ \text{id u} = \text{u} \\ \text{g x} = \text{if } x \leq 0 \text{ then id else app g (x - 1)} \\ \text{main ()} = \text{g *int ()} \end{array} \right\}$$

The following is a possible reduction sequence of the program.

$$\begin{aligned} \text{main ()} &\rightarrow^* \text{g 2 ()} \rightarrow^* \text{app g 1 ()} \rightarrow^* \text{g 1 ()} \\ &\rightarrow^* \text{app g 0 ()} \rightarrow^* \text{g 0 ()} \rightarrow^* \text{id ()} \rightarrow^* () \end{aligned}$$

Note that, although `g` is applied to two arguments in the reduction above,  $\text{arity}(\text{g}) = 1$  in our definition (because  $x$  is the only formal parameter in the definition of `g`).

We define termination as follows.

**Definition 1.** A program  $P$  is terminating, if there is no infinite reduction sequence  $\text{main ()} \rightarrow_P e_1 \rightarrow_P e_2 \rightarrow_P \dots$ .

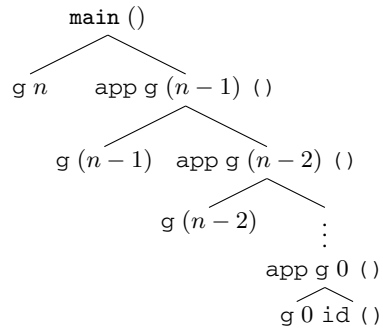
*Remark 1.* The language  $L$  is untyped and therefore, a program evaluation may get stuck. We consider a reduction sequence that ends with a stuck expression as terminating. Our approach is sound and complete even for untyped languages. But, our implementation currently supports only the typed subset because it delegates the reachability checking to a higher-order program model checker for a typed language.

### 3 Termination Verification via Binary Reachability

This section describes our termination verification method. We give an informal overview of the whole process in Subsection 3.1, and discuss each step in a more detail in the later subsections.

#### 3.1 Overview

We use `indirect` from Example 2 as a running example in this subsection. Our termination verification method is based on the observation that a functional program is terminating if and only if each of its *call tree*,<sup>1</sup> which expresses how the functions are called in an execution of the program, is finite. Figure 4 shows a call tree for `indirect`. Each node expresses a fully applied function call (a



**Fig. 4.** A call tree of program `indirect`

<sup>1</sup> The call tree in this paper roughly corresponds to the *dynamic call graph* of [29].

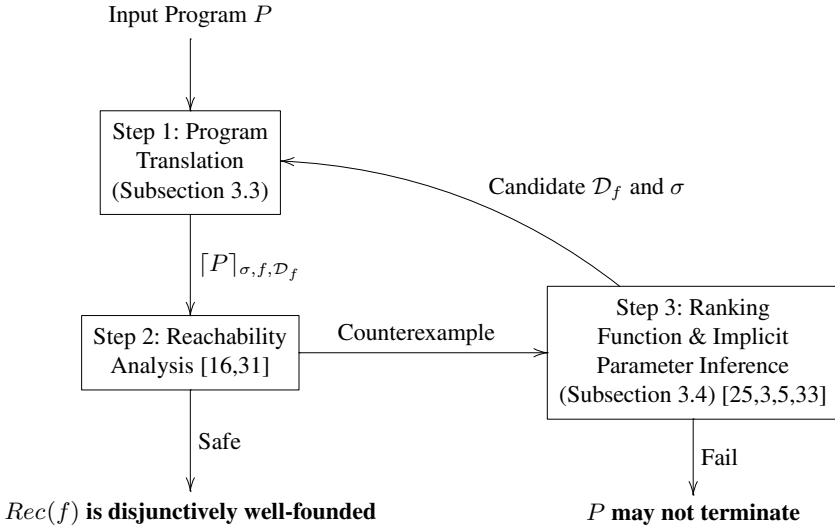


Fig. 5. Overview of the termination verification process

call of the form  $f \tilde{v}$  with  $\text{arity}(f) = |\tilde{v}|$ , and an edge represents that the function call represented by the target node is made from the function call of the source node. For example, the edge from “app g (n - 1) ()” to “g (n - 1)” means that g (n - 1) is called during evaluation of app g (n - 1) () (i.e.,  $\text{app g (n - 1) ()} \rightarrow^+ E[\text{g (n - 1)}]$  for some evaluation context  $E$ ). For every (possibly infinite) reduction sequence, the corresponding call tree is finitely branching. Therefore, by König’s lemma, to show that every call tree of the program is finite and therefore the program is terminating, it is sufficient to show that no call tree has an infinite path from the root. The latter is equivalent to showing that for every function  $f$ , there is no infinite path  $f \tilde{v}_1 \triangleright f \tilde{v}_2 \triangleright f \tilde{v}_3 \triangleright \dots$  where  $f \tilde{v} \triangleright g \tilde{w}$  means that  $f \tilde{v}$  is an ancestor of  $g \tilde{w}$  in the call tree. In the running example, the only non-trivial sequence is  $\text{app g } n_1 \text{ ()} \triangleright \text{app g } n_2 \text{ ()} \triangleright \text{app g } n_3 \text{ ()} \triangleright \dots$  and this sequence must be finite since  $n_i$  is decreasing and bounded below by 0. Thus, we may conclude that `indirect` is terminating. We refer to Subsection 3.2 for the formal exposition of the above argument.

By the above argument, to show that the program terminates, it suffices to show that, for every function  $f$ , the relation  $\text{Rec}(f) = \{(\tilde{v}, \tilde{v}') \mid f \tilde{v} \triangleright f \tilde{v}'\}$  is well-founded, or disjointively well-founded<sup>2</sup> [26] because  $\text{Rec}_P(f)$  is transitive. Thus, as mentioned in Section 1, our termination verification method proceeds as follows.

- (i) For each function  $f$ , guess a disjointively well-founded relation  $\mathcal{D}_f$  that over-approximates  $\text{Rec}(f)$ , and reduce the termination verification problem to the so called the *binary reachability analysis* problem of showing  $\text{Rec}(f) \subseteq \mathcal{D}_f$ .

<sup>2</sup> A binary relation is *disjointively well-founded* if it is a finite union of well-founded sets [26].

- (ii) Use a program transformation to reduce the binary reachability problem to a plain reachability problem; and
- (iii) Solve the plain reachability problem by using an off-the-shelf software model checker.

We express the disjunctively well-founded relation  $\mathcal{D}_f$  by using a set of ranking functions, and gradually refine the set by using the technique of counterexample-guided abstraction refinement [25,3,5,33].

Figure 5 shows the overall flow of the process. We start with an empty set of ranking functions (i.e., with  $\mathcal{D}_f = \emptyset$ ), and apply a program transformation to reduce the binary reachability analysis problem of deciding  $Rec(f) \subseteq \mathcal{D}_f$  to the plain reachability analysis problem of deciding if an assertion failure is reachable in the translated program  $[P]_{\sigma, f, \mathcal{D}_f}$  (Step 1 in Figure 5). Here,  $\sigma$  is a *candidate implicit parameter instantiation* and is used for ranking functions over higher-order values (i.e., function closures). For the purpose of the exposition, we focus on the case where the ranking functions are only over the first-order values in Subsections 3.2 and 3.3, and defer implicit parameters and ranking functions over higher-order values to Subsection 3.4. For simplicity, we write  $[P]_{f, \mathcal{D}_f}$  for  $[P]_{\sigma, f, \mathcal{D}_f}$  when implicit parameters are irrelevant to the discussion.

Informally, the idea of the program transformation is to pass around the argument  $\widetilde{pv}$  of an ancestor call as an *extra argument*, and assert in the body of  $f$  that  $\widetilde{pv}$  and the current argument  $\widetilde{v}$  are related by  $\mathcal{D}_f$ . For the running example, the definition of `app` would be transformed as follows.

```
app s1 f s2 x s3 u = assert((s3, (f, x, u)) ∈  $\mathcal{D}_{app}$ ); ...
```

Here, `s1`, `s2`, and `s3` are the extra arguments that carry the arguments of ancestor calls to `app`. Because a function body does not evaluate until the function is fully applied, only `s3` is relevant and is checked with the current arguments  $(f, x, u)$  for the candidate disjunctively well-founded relation  $\mathcal{D}_{app}$ . We use non-determinism to ensure that the extra parameter can be bound to the argument of an *arbitrary* ancestor call, so that the transformed program is assertion safe if and only if  $Rec(f) \subseteq \mathcal{D}_f$ . Because a function can be called indirectly in a higher-order program, we pass extra arguments at *all* call sites, including indirect calls, to account for all possible ancestor relations. Therefore, for the running example, we also transform the functions `g` and `id` to take extra parameters (but only track and check the arguments of `app`). This makes our approach sound and complete even in the presence of higher-order functions.<sup>3</sup> In this manner, the binary reachability problem is reduced to a plain reachability problem. We refer to Subsection 3.3 for a more formal exposition.

We proceed to the plain reachability analysis (Step 2 in Figure 5), to check if the inserted assertion may fail. If the assertion cannot fail, then we conclude that  $Rec(f) \subseteq \mathcal{D}_f$  holds (and if that is the case for every function  $f$  in the program, we conclude that the program is terminating). Even if the program is terminating, however, an assertion failure may occur, because of a wrongly guessed  $\mathcal{D}_f$ . For our running example, in the initial

<sup>3</sup> This is a crucial difference with the previous approaches to automated termination verification for higher-order programs [30,28,29,6,12] that handle higher-order functions by approximating the call graph up front (e.g., via a control flow analysis), which can lose precision when the calls depend on non-trivial safety conditions.

iteration, we set  $\mathcal{D}_{\text{app}} = \emptyset$ , and when the transformed program is given to a reachability verifier (e.g., MoChi [16]), the verifier would return a concrete counterexample to the assertion safety. Suppose the following counterexample is returned.

$$\text{main } () \rightarrow^* \text{app } s_1 \text{ g } s_2 \text{ 0 } (\text{g}, 1, ()) () \rightarrow^* \text{fail}$$

which corresponds to the following reduction sequence in the original program:

$$\text{main } () \rightarrow^* \text{g } 2 () \rightarrow^* \text{app } \text{g } 1 () \rightarrow^* \text{app } \text{g } 0 ()$$

We analyze the counterexample, and infer a new ranking function to refine  $\mathcal{D}_f$  (Step 3 in Figure 5). For the running example, we look for a ranking function  $r$  such that  $r(\text{g}, 1, ()) > r(\text{g}, 0, ()) \geq 0$  and update  $\mathcal{D}_{\text{app}}$  to  $\mathcal{D}_{\text{app}} \cup \{((pf, px, pu), (f, x, u)) \mid r(pf, px, pu) > r(f, x, u) \geq 0\}$ . To find a ranking function, we adopt the existing techniques for inferring ranking functions from counterexamples of first-order programs [25,5].<sup>4</sup> If no ranking function can be found, then we report that the program may not be terminating. This may happen either because the program is indeed non-terminating, or the method used for the ranking function synthesis is incomplete. We refer to Subsection 3.4 for a more formal exposition on this step.

Upon refining  $\mathcal{D}_f$ , we go back to Step 1 and repeat the process. If the program is terminating, and if the underlying reachability analysis tool and the ranking function synthesis *were* complete, the loop eventually terminates and we conclude that the program is terminating.

### 3.2 Termination and Binary Reachability

We discuss how termination verification is reduced to binary reachability analysis (to show  $\text{Rec}(f) \subseteq \mathcal{D}_f$ ) more formally. First, we define the relations  $\triangleright$  and  $\text{Rec}(f)$ .

**Definition 2.** *The call relation  $\triangleright_P$  is the binary relation defined by:*

$$\triangleright_P := \{(f \tilde{v}, g \tilde{w}) \mid (\text{main } () \rightarrow_P^* E_1 [f \tilde{v}]) \wedge (f \tilde{v} \rightarrow_P^+ E_2 [g \tilde{w}]) \\ \wedge \text{arity}(f) = |\tilde{v}| \wedge \text{arity}(g) = |\tilde{w}|\}$$

We often use the infix notation  $f \tilde{v} \triangleright_P g \tilde{w}$  for  $(f \tilde{v}, g \tilde{w}) \in \triangleright_P$ .

**Definition 3.**  *$\text{Rec}_P(f)$ , the recursion relation of  $f$  in  $P$ , is the binary relation defined by:*

$$\text{Rec}_P(f) := \{(\langle \tilde{v}_1 \rangle, \langle \tilde{v}_2 \rangle) \mid f \tilde{v}_1 \triangleright_P f \tilde{v}_2\}$$

When it is clear from contexts, we omit the subscript  $P$ . Note that the relations  $\triangleright_P$  and  $\text{Rec}_P(f)$  are transitive.

*Example 3.* Recall program `fib` in Example 1.  $\text{Rec}(\text{fib})$  is:

$$(\{(n, n-1) \mid n > 1\} \cup \{(n, n-2) \mid n > 2\})^+ = \{(m, n) \mid m > n \geq 0\}$$

<sup>4</sup> To infer a ranking relation over function closure values, we use implicit parameters and infer sufficient instantiations for them to represent the closures. This is done by adopting the counterexample-guided technique from the previous work [33] (cf. Subsection 3.4).



As shown in the example below,  $Rec(f)$  may be non-empty even if  $f$  is not recursively defined, and  $Rec(f)$  may be empty even if  $f$  is recursively defined.

*Example 4.* Recall the program in Example 2. The recursion relations are:

$$\begin{aligned} Rec(\mathbf{app}) &= \{(\langle \mathbf{g}, m, () \rangle, \langle \mathbf{g}, n, () \rangle) \mid m > n \geq 0\} \\ Rec(\mathbf{id}) &= Rec(\mathbf{g}) = \emptyset \end{aligned}$$

Note that  $\mathbf{g}$  is defined recursively but does not cause a recursive call to  $\mathbf{g}$ . Therefore, the relation  $Rec(\mathbf{g})$  is empty. On the other hand,  $\mathbf{app}$  is not recursively defined but  $Rec(\mathbf{app}) \neq \emptyset$ . This shows that we must check the disjunctive well-foundedness of  $Rec(f)$  for each function  $f$ , regardless of whether  $f$  is recursively defined or not.

Next, we show the termination verification problem can be reduced soundly and completely to the problem of showing that  $Rec(f)$  is disjunctively well-founded for every  $f$ .

We state the soundness and the completeness of the reduction.

**Theorem 1 (Soundness).** *A program  $P$  is terminating if  $Rec_P(f)$  is disjunctively well-founded for every  $f$  defined in  $P$ .*

**Theorem 2 (Completeness).** *If a program  $P$  is terminating, then  $Rec_P(f)$  is disjunctively well-founded for every  $f$  defined in  $P$ .*

### 3.3 From Binary Reachability to Plain Reachability

This subsection presents the reduction from the binary reachability problem of deciding  $Rec_P(f) \subseteq \mathcal{D}_f$  to a plain reachability problem. As remarked in Subsection 3.1, we transform  $P$  to the program  $\lceil P \rceil_{f, \mathcal{D}_f}$  that simulates the program  $P$  and asserts the property  $(\tilde{v}', \tilde{v}) \in \mathcal{D}_f$  whenever a recursive call relation  $f \tilde{v}' \triangleright_P f \tilde{v}$  is detected. Then,  $Rec_P(f) \subseteq \mathcal{D}_f$  holds if and only if an assertion in  $\lceil P \rceil_{f, \mathcal{D}_f}$  may fail.

The target language of the program transformation is an extension of  $L$  with tuples  $\langle e_1, \dots, e_k \rangle$ , assertions  $\mathbf{assert}(e_1); e_2$ , and a special value  $\perp$ . The semantics of assertions is defined by:

$$E[\mathbf{assert}(\mathbf{true}); e] \rightarrow_P E[e] \quad E[\mathbf{assert}(\mathbf{false}); e] \rightarrow_P \mathbf{fail}$$

where the evaluation contexts are extended accordingly:  $E ::= \dots \mid \mathbf{assert}(E); e$ . The special value  $\perp$  is used in place of the argument of an ancestor call, when there is no tracked ancestor call; see examples below.

Before we give the formal definition of the transformation,  $\lceil \cdot \rceil_{f, \mathcal{D}_f}$ , we informally describe the idea. The program  $\lceil P \rceil_{f, \mathcal{D}_f}$  is obtained by adding extra function arguments that represent the arguments of past calls to  $f$ .

First we consider the simple case where the program only contains first-order functions. For example, let  $P$  be the Fibonacci program from Example 1. Let  $\mathcal{D}_{\mathbf{fib}} = \{(\mathbf{pn}, \mathbf{n}) \mid \mathbf{pn} > \mathbf{n} \geq 0\}$ . Then,  $\lceil P \rceil_{\mathbf{fib}, \mathcal{D}_{\mathbf{fib}}}$  would be as follows.

```

1: let rec fib pn n =
2:   assert(pn>n && n≥0);
3:   let pn' = if *bool then pn else n in
4:   if n < 2 then 1 else fib pn' (n-1) + fib pn' (n-2)
5: let main () = fib ⊥ *int

```

We have added the formal argument  $pn$  that represents the argument of an ancestor call (i.e., a call that corresponds to an ancestor node in the call tree) for  $fib$ , and inserted the assertion that checks  $(pn, n) \in \mathcal{D}_{fib}$  (lines 1-2). Accordingly, we have also inserted an extra parameter to each function call (calls in line 4). In line 3, we non-deterministically “update” the tracked argument to the current argument in order to compare the argument of the current call with a future call of  $fib$ . The extra parameter is initially set to  $\perp$ , to indicate that there is no ancestor call (line 5). We assume that  $>$  is extended so that  $\perp > n$  holds for every  $n$ . For the two recursive calls to  $fib$  in the definition of  $fib$ ,  $pn$  or  $n$  is passed in a non-deterministic manner. Below is a possible reduction of the transformed program.

$$\begin{aligned}
& \text{main} () \rightarrow^* \text{fib } \perp \ 2 \rightarrow^* \text{assert}(\perp > 2 \ \&\& \ 2 \geq 0); \dots \\
& \rightarrow^* \text{fib } 2 \ 1 + \text{fib } 2 \ 0 \\
& \rightarrow^* (\text{assert}(2 > 1 \ \&\& \ 1 \geq 0); \dots) + \text{fib } 2 \ 0 \\
& \rightarrow^* 1 + \text{fib } 2 \ 0 \\
& \rightarrow^* 1 + (\text{assert}(2 > 0 \ \&\& \ 0 \geq 0); \dots) \\
& \rightarrow^* 1 + 1 \rightarrow^* 2
\end{aligned}$$

The subexpressions  $fib \ 2 \ 1$  and  $fib \ 2 \ 0$  capture the fact that  $fib \ 1$  and  $fib \ 0$  are called from  $fib \ 2$  in the original program. It is easy to see that  $Rec_P(fib) \subseteq \mathcal{D}_{fib}$  if and only if  $[P]_{fib, \mathcal{D}_{fib}}$  does not cause an assertion failure.

The transformation is more subtle for higher-order programs with partial applications and indirect calls. For example, let  $P$  be the program `indirect` from Example 2 and 4. Suppose that we wish to show  $Rec_P(\text{app}) \subseteq \mathcal{D}_{\text{app}}$  where  $\mathcal{D}_{\text{app}} = \{((ph, pv, pu), (h, v, u)) \mid pv > v \geq 0\}$ . Then,  $[P]_{\text{app}, \mathcal{D}_{\text{app}}}$  would be as follows.

```

1: let app _ h _ v (ph, pv, pu) u =
2:   assert(pv>v && v≥0);
3:   let (ph, pv, pu) = if *bool then (ph, pv, pu) else (h, v, u)
4:   in
5:     h (ph, pv, pu) v (ph, pv, pu) u
6: let id (ph, pv, pu) u = u
7: let rec g (ph, pv, pu) x =
8:   if x≤0 then id
9:   else app (ph, pv, pu)  $\tilde{g}$  (ph, pv, pu) (x-1)
10: let main () = g ⊥ *int ⊥ ()

```

Here,  $\tilde{\perp}$  denotes  $(\perp, \perp, \perp)$ . As before, the extra parameter  $(ph, pv, pu)$  is inserted to represent the arguments of an ancestor call to  $\text{app}$  (line 1). Note that the extra parameter for this program takes a tuple of values, so that all three arguments of  $\text{app}$  can be tracked (i.e.,  $h$ ,  $v$ , and  $u$ ). As before, the assertion is inserted at the beginning

of the body of `app` to check  $((ph, pv, pu), (h, v, u)) \in \mathcal{D}_{fib}$  (line 2), and we update the tracked past arguments with the current arguments non-deterministically, in order to compare the arguments of the current call with a future call of the function (line 3).

To soundly and completely track the extra parameters through partial applications and indirect calls, we pass the extra parameter at *every* function application site (lines 5, 9, and 10). To be able to do this, note that we have also transformed the definitions of `id` and `g` to take the extra parameters and pass them at the applications that occur in their body (lines 6 and 7), and also transformed the definition of `app` to take an extra parameter argument just before `h` and `v` as well as just before `u`, even though `app` only checks the well-foundedness against the one passed just before `u`. (Note that `_` in the definition of `app` is an unused argument.) This is needed, because, in general, we cannot statically decide which (indirect) function call is a fully applied function call, nor which is a call to the target function (i.e., `app` in the example). We note that it is possible to soundly eliminate some of the redundancy via a static analysis (see Example 5 below), but it is in general impossible to completely decide *a priori* which function is called in what context. In effect, the idea of our transformation is to delegate such tasks to the backend reachability checker.

*Example 5.* By using useless code elimination [34,13], we can simplify the above program to:

```
let app h v (ph,pv,pu) u =
  assert(pv>v && v≥0);
  let (ph,pv,pu) = if *bool then (ph,pv,pu) else (h,v,u) in
  h v (ph,pv,pu) u
let id (ph,pv,pu) u = u
let rec g x = if x≤0 then id else app g (x-1)
let main () = g *int 1 ()
```

Below is a possible reduction of  $[P]_{app, \mathcal{D}_{app}}$ . (For simplicity, we use a reduction sequence from the optimized version in Example 5.)

$$\begin{aligned} \text{main} () &\longrightarrow^* g\ 2\ \tilde{1} () \longrightarrow^* \text{app}\ g\ 1\ \tilde{1} () \\ &\longrightarrow^* g\ 1\ (g, 1, ()) () \longrightarrow^* \text{app}\ g\ 0\ (g, 1, ()) () \end{aligned}$$

Note that the reached state `app g 0 (g, 1, ()) ()` captures the recursion relation `app g 1 () ▷P app g 0 ()` of the original program  $P$ .

### The Formal Definition of $[\cdot]_{f, \mathcal{D}}$

We now define the transformation formally.  $[P]_{f, \mathcal{D}}$  is obtained by transforming each function definition:

$$[P]_{f, \mathcal{D}} = \{[g\ \tilde{x} = e]_{f, \mathcal{D}} \mid g\ \tilde{x} = e \in P\}$$

where the function definition transformation is defined as follows.

$$[g x_1 \cdots x_k = e]_{f, \mathcal{D}} = \begin{cases} g s_1 x_1 \cdots s_k x_k = \\ \quad \text{let } s = \mathbf{check\&upd}(\mathcal{D}, s_k, \langle x_1, \dots, x_k \rangle) \text{ in } [e]_s & \text{if } g = f \\ g s_1 x_1 \cdots s_k x_k = [e]_{s_k} & \text{if } g \neq f \end{cases}$$

Note that an extra parameter  $s_i$  is added before every original parameter  $x_i$  of a function. Therefore, as opposed to the informal examples given above, the main function in the target program now takes *two* arguments, the first of which is always instantiated to  $\perp$ . The code that checks the candidate well-foundedness and non-deterministically updates the arguments is inserted at the beginning of the body of the target function  $f$ . Here,  $\mathbf{check\&upd}(\mathcal{D}, s_k, \langle x_1, \dots, x_k \rangle)$  denotes the expression

$$\mathbf{assert}(\mathcal{D}^\#(s_k, \langle x_1, \dots, x_k \rangle)); \text{if } *_{\mathbf{bool}} \text{ then } s_k \text{ else } \langle x_1, \dots, x_k \rangle.$$

where the relation  $\mathcal{D}^\#$  is the extension of  $\mathcal{D}$  defined by:

$$\mathcal{D}^\# = \{(\tilde{\perp}, \langle v_1, \dots, v_k \rangle) \mid v_1, \dots, v_k \text{ are values}\} \cup \{(\langle v'_1, \dots, v'_k \rangle, \langle v_1, \dots, v_k \rangle) \mid (\langle [v'_1], \dots, [v'_k] \rangle, \langle [v_1], \dots, [v_k] \rangle) \in \mathcal{D}\}$$

Here,  $[v]$  is the value obtained by removing all the extra arguments from partial applications; see below for the definition. We assume that  $\mathcal{D}^\#$  is represented by a formula of some logic. (In the implementation, we use the first-order logic with linear arithmetic.)

Note that the body  $e$  of each function definition is transformed by  $[e]_s$  where  $s$  is the extra parameter passed just before the last argument (non-deterministically updated to the current arguments in the case of the target function). The expression transformation  $[e]_s$  passes  $s$  at each application site in  $e$ , and is formally defined as follows.

$$\begin{aligned} [c]_s &= c & [*_{\mathbf{int}}]_s &= *_{\mathbf{int}} & [f]_s &= f & [x]_s &= x \\ [\mathbf{let } x = e_1 \text{ in } e_2]_s &= \mathbf{let } x = [e_1]_s \text{ in } [e_2]_s \\ [e_1 \text{ op } e_2]_s &= [e_1]_s \text{ op } [e_2]_s \\ [\mathbf{if } e_1 \text{ then } e_2 \text{ else } e_3]_s &= \mathbf{if } [e_1]_s \text{ then } [e_2]_s \text{ else } [e_3]_s \\ [e_1 e_2]_s &= [e_1]_s s [e_2]_s \end{aligned}$$

The operation  $[e]$  for removing extra arguments (used in the definition of  $\mathcal{D}^\#$ ) is defined as follows.

$$\begin{aligned} [c] &= c & [*_{\mathbf{int}}] &= *_{\mathbf{int}} & [f] &= f & [x] &= x \\ [\mathbf{let } x = e_1 \text{ in } e_2] &= \mathbf{let } x = [e_1] \text{ in } [e_2] \\ [e_1 \text{ op } e_2] &= [e_1] \text{ op } [e_2] & [e_1 s e_2] &= [e_1] [e_2] \\ [\mathbf{if } e_1 \text{ then } e_2 \text{ else } e_3] &= \mathbf{if } [e_1] \text{ then } [e_2] \text{ else } [e_3] \end{aligned}$$

Note that  $[[e]_s] = e$ .

We prove the soundness and the completeness of the transformation. The following theorem states the soundness of the transformation. It says that the target program reaches an assertion failure when the recursion relation is not a subset of  $\mathcal{D}$ .

**Theorem 3 (Soundness of  $[\cdot]_{f, \mathcal{D}_f}$ )**

Suppose that  $\text{main}() \rightarrow_P^* E_1[f v_1 \cdots v_k]$ ,  $f v_1 \cdots v_k \rightarrow_P^+ E_2[f w_1 \cdots w_k]$ , and  $(\langle v_1, \dots, v_k \rangle, \langle w_1, \dots, w_k \rangle) \notin \mathcal{D}$ . Then,  $\text{main} \perp () \rightarrow_{[P]_{f, \mathcal{D}}}^* \text{fail}$ .

The theorem below states the completeness. It says that the target program reaches an assertion failure only when the recursion relation is not a subset of  $\mathcal{D}$ .

**Theorem 4 (Completeness of  $[\cdot]_{f, \mathcal{D}_f}$ )**

If  $\text{main} \perp () \rightarrow_{[P]_{f, \mathcal{D}}}^* \text{fail}$ , then  $\text{main}() \rightarrow_P^* E_1[f \tilde{v}]$  and  $f \tilde{v} \rightarrow_P^+ E_2[f \tilde{w}]$ , and  $(\langle \tilde{v} \rangle, \langle \tilde{w} \rangle) \notin \mathcal{D}$  for some  $E_1, E_2, \tilde{v}, \tilde{w}$ .

**3.4 Ranking Function Inference**

This subsection details how we refine the candidate disjunctively well-founded relation  $\mathcal{D}_f$ . As remarked in Subsection 3.1, we actually infer both  $\mathcal{D}_f$  and the *implicit parameter instantiation*  $\sigma$ . The implicit parameters are used to assert and check well-founded relation over function closure values.

We first describe the case where only  $\mathcal{D}_f$  is inferred. (This happens, for example, when  $f$  is first-order and does not take function closures as arguments.) Recall that the inference is invoked when  $\text{Rec}(f) \not\subseteq \mathcal{D}_f$  (cf. Step 3 of Figure 5), and in such a case, the reachability checker returns a counterexample of the form:

$$\begin{aligned} \text{main}() &\rightarrow^* E[f s_1 v_1 s_2 v_2 \cdots (v'_1, \dots, v'_n) v_n] \\ &\rightarrow \text{assert}((v'_1, \dots, v'_n), (v_1, \dots, v_n)) \in \mathcal{D}_f; \dots \rightarrow \text{fail} \end{aligned}$$

As remarked in Subsection 3.3, this implies that  $f v'_1 \cdots v'_n \triangleright f v_1 \cdots v_n$ , and we have that  $((v'_1, \dots, v'_n), (v_1, \dots, v_n)) \in \text{Rec}(f)$  and  $((v'_1, \dots, v'_n), (v_1, \dots, v_n)) \notin \mathcal{D}_f$ .

The goal of ranking function inference is to obtain a refined disjunctively well-founded relation  $\mathcal{D}'_f$  such that

$$\mathcal{D}_f \cup \{((v'_1, \dots, v'_n), (v_1, \dots, v_n))\} \subseteq \mathcal{D}'_f.$$

To this end, we infer a new ranking function  $r(x_1, \dots, x_n)$  such that  $r(v'_1, \dots, v'_n) > r(v_1, \dots, v_n) \geq 0$  and let  $\mathcal{D}'_f = \mathcal{D}_f \cup \{(\tilde{x}', \tilde{x}) \mid r(\tilde{x}') > r(\tilde{x}) \geq 0\}$ . We adopt the constraint-based technique [25,5] to infer  $r(\tilde{x})$ .

We overview the inference process. We prepare a ranking function *template*  $c_0 + c_1 x_1 + \cdots + c_n x_n$ . Here,  $c_i$ 's are fresh variables, serving as *unknowns*. Then, we solve for the assignments to  $c_i$ 's that satisfy the constraint

$$\forall \tilde{x}. \llbracket \pi \rrbracket \Rightarrow c_0 + c_1 v'_1 + \cdots + c_n v'_n > c_0 + c_1 v_1 + \cdots + c_n v_n \geq 0$$

where  $\tilde{x}$  are the free variables in  $\llbracket \pi \rrbracket$  and  $v_1, \dots, v_n, v'_1, \dots, v'_n$ . Here,  $\llbracket \pi \rrbracket$  is the strongest postcondition of the given counterexample  $\pi$ .<sup>5</sup> Finally, we set  $r(x_1, \dots, x_n) = \alpha_0 + \alpha_1 x_1 + \cdots + \alpha_n x_n$  where each  $\alpha_i$  is the assignment obtained for  $c_i$ .

Next, we extend the above process with *implicit parameters* to infer ranking functions over higher-order values. We illustrate the need for ranking functions over higher-order values with the following program `indirectHO`.

<sup>5</sup> More precisely, we construct a corresponding straightline program from the counterexample, and take its strongest postcondition (cf. [16,33] and the extended report [17]).

```

let app h v = h () v
let id x = x
let rec g x u =
  if x <= 0 then id else app (g (x-1))
let main () = g *int () ()

```

The program is similar to `indirect` from Example 2 and 4, except that `app` no longer takes an integer argument and instead has the “decreasing” integer value captured inside the function closure passed as `h`. To show that this program is terminating, we need to show that the recursion relation for `app` is disjunctively well-founded. However, because `app` only takes function-type arguments (besides unit), ranking functions over first-order values are insufficient for this.

To this end, we adopt the idea from the previous work [33] and systematically add an integer-type *implicit parameter* just before each function-type parameter.<sup>6</sup> For `indirectHO`, we add an implicit parameter `h_IMPPARAM` before `h` so that the program is now the following. (The added parts are underlined.)

```

let app h_IMPPARAM h () = h () ()
let id x = x
let rec g x () =
  if x <= 0 then id else app σ(ℓ) (g (x-1))
let main () = g *int () ()

```

Here,  $\sigma$  is the *candidate implicit parameter instantiation* that maps each *instantiation site*  $\ell$  to an arithmetic expression over the variables bound in the context of  $\ell$ . Formally, an instantiation site is at an application of a function-type argument, and is syntactically determined (i.e., between  $e_1$   $e_2$  where  $e_2$  is function-type). Clearly, the addition of implicit parameters and their instantiations do not affect the termination of the program, and so we may check the termination of the program with the implicit parameters added to check the termination of the original. The verification process starts by initializing the candidate instantiations to some arithmetic expression (e.g., 0), and refine them iteratively via a counterexample analysis (cf. Figure 5).

As remarked above, in the presence of implicit parameters, we infer both  $\sigma$  and  $\mathcal{D}_f$  when given a counterexample. To this end, the above inference process is extended as follows. We prepare templates for the instantiation expressions in addition to the templates for the ranking functions. Then, when generating the constraints, we use the template instantiation expressions in the strongest postcondition of the counterexample, and solve for both the unknowns in the ranking function templates and the instantiation expression templates.

More formally, let  $\Pi = \{\pi_1, \dots, \pi_m\}$  be the set of counterexamples we have seen so far for  $f$ . We prepare a template instantiation map  $\Delta$  that maps each  $\ell$  to an expression of the form  $c_0 + c_1x_1 + \dots + c_nx_n$  where  $c_i$ 's are fresh unknowns and  $x_i$ 's are the

---

<sup>6</sup> Implicit parameters are called “extra parameters” in [33]. We call them implicit parameters here to avoid confusion with the extra parameters in this paper which are used for a different purpose.

**Table 1.** Experiment results

program	ord	time	program	ord	time	program	ord	time
Ackermann	1	5.85	alias_partial	1	0.32	churchNum	4	3.13
Fibonacci	1	0.15	quicksort	2	timeout	CE-Jones_Bohr	4	0.71
McCarthy91	1	4.95	indirectIntro	2	4.76	up_down	2	0.65
loop2	1	0.61	indirect	2	1.36	map	2	1.59
append	1	0.14	indirectHO	2	7.75	toChurch	2	0.69
zip	1	0.15	CE-0CFA	2	0.14	x_plus_2^n	2	2.02
binomial	1	0.70	CE-1CFA	2	0.24	foldr	2	1.19

integer-type variables that are bound in the context of  $\ell$  (which may include implicit parameters). Then, we form the following constraint

$$\bigwedge_{\pi \in \Pi} \forall \tilde{x}. \llbracket \pi \Delta \rrbracket \Rightarrow c_{\pi,0} + c_{\pi,1}v'_{\pi,1} + \dots + c_{\pi,n}v'_{\pi,n} > c_{\pi,0} + c_{\pi,1}v_{\pi,1} + \dots + c_{\pi,n}v_{\pi,n} \geq 0$$

where  $c_{\pi,i}$ 's are fresh unknowns, and  $\tilde{x}$  are the free non-unknown variables in  $\llbracket \pi \Delta \rrbracket$  and  $v_1, \dots, v_n, v'_1, \dots, v'_n$ . (Here, we assume that the counterexamples  $\pi$  explicitly use the instantiation sites  $\ell$  as expressions.) We solve for the unknowns that satisfy the constraint to obtain implicit parameter instantiations and ranking functions that refute the counterexample. We obtain the new candidate disjunctive well-founded relation from the ranking functions:  $\mathcal{D} = \bigcup_{\pi \in \Pi} \{(\tilde{x}', \tilde{x}) \mid r_{\pi}(\tilde{x}') > r_{\pi}(\tilde{x}) \geq 0\}$  where  $r_{\pi}(\tilde{x}) = \alpha_{\pi,0} + \alpha_{\pi,1}x_1 + \dots + \alpha_{\pi,n}x_n$  and each  $\alpha_{\pi,i}$  is the obtained assignment for  $c_{\pi,i}$ . And, we substitute the assignments to the unknowns in  $\Delta$  to obtain the new candidate implicit parameter instantiation. The extended report [17] contains details of the inference process applied to `indirectHO`.

## 4 Implementation and Experiments

We have implemented a prototype of the termination verifier for a subset of OCaml. We use MoChi [16] as the backend reachability checker, and Z3 [22] as a constraint solver for ranking function inference. As an optimization, we have extended the ranking function inference described in Section 3.4 to also infer lexicographic linear ranking functions [5] whenever possible. The extended report [17] contains details of the lexicographic linear ranking function inference process.

We have tested our tool on various termination verification benchmark programs in literature, taken mostly from the previous work on termination verification of higher-order programs, as well as some synthetic but non-trivial examples. We ran the experiment on a machine with 3.20GHz CPU and 16GB of memory, with timeout of 600 seconds. The web interface of the verification tool and the programs used in the experiments are available online [17].

Table 1 summarizes the experiment results. The column “program” shows the names of programs, and the column “ord” shows the order of the program (where order-1 functions take only base type values, order-2 functions may take order-1 functions as

arguments, etc., and the order of a program is the maximum order of the functions in the program). The column “time” shows the running time in seconds.

We briefly describe the benchmark programs. The seven programs in the left column and `alias_partial` are first-order (i.e., order-1) programs. The programs `append`, `zip`, and `binomial` are from [2]. `Ackermann` is the Ackermann’s function, and is also used as examples in [2,35]. `McCarthy91` is the McCarthy’s 91 function (used as a benchmark program in, e.g., [18]<sup>7</sup>). `Fibonacci` is the Fibonacci number function from Example 1. The program `alias_partial` is from Section 8 of [18] and is given as an example on which their approach fails.

The program `quicksort` is from [35],<sup>8</sup> and is a second-order program where the list sorting function is parametrized by the “compare” function. We check the termination of a program that passes the sorting function a terminating compare function and an arbitrary list. (Our tool currently does not directly support lists, and so a list is represented by the integer denoting its length.) Our tool fails to verify the program within the time limit due to the underlying reachability checker MoCHI failing to verify the necessary assertion safety. This is not a fundamental limitation with our termination verification approach, and we expect further advances in reachability verification to allow our approach to verify instances like `quicksort`.

The rest of the programs are higher-order programs whose termination depends non-trivially on the functions passed as the arguments, and precise reasoning about the function arguments is required for proving termination. They are mostly from the examples and benchmarks in [12,28,29]. Many of these are difficult examples that the previous approaches cannot verify. (We have selected the ones given as examples where their approaches fail). We refer to the extended report [17] for further description of these programs. As seen in Table 1, the benchmark results are promising and show that our tool is able to automatically verify the difficult instances quickly, except `quicksort` whose reason for the failure is elaborated above.

## 5 Related Work

There have been three major approaches to automated termination verification for first-order programs: *transition invariants* [26,3,4], *size-change termination* [19], and *term rewriting* [7] (see also [32,8] for relationships between those approaches). The approaches have recently been extended to the termination verification of higher-order programs [18,30,28,29,6,12]. Below, we compare them with our approach.

### 5.1 Transition Invariants

Closest to our work is the work by Ledesma-Garza and Rybalchenko [18]. Similar to our work, they propose a program transformation to reduce the transition invariant verification problem (i.e., binary reachability analysis) to a plain reachability problem via a program transformation. Unfortunately, as also admitted in their paper (Section

<sup>7</sup> [18] is not fully automated and requires the user to provide the sufficient ranking functions as well as the predicates to be used for reachability checking.

<sup>8</sup> [35] is not automated.



8 of [18]), their approach has a limited applicability to the verification of higher-order programs because it does not correctly handle indirect calls and is actually unsound. For example, their approach would incorrectly report the program  $P_0$  from Section 1 to be terminating. Moreover, their approach is not fully automated and requires a sufficient well-foundedness relation to be provided manually, and it also cannot handle well-foundedness relations over function closure values.

By contrast, we have proposed the first sound and (relatively) complete approach to termination verification of higher-order programs via binary reachability analysis. A key idea of our approach is the novel program transformation that precisely tracks the call-tree ancestor's arguments values through the higher-order control flow without a priori approximation. We have also presented a method to infer well-foundedness relations (including those over function closure values) from counterexamples returned by a higher-order program verifier, thus realizing a fully automated verification.

## 5.2 Size-Change Analysis

The size-change approach [19] to termination verification involves the following two steps: (1) an analysis of the program to construct a *size-change graph*, and (2) an analysis of the obtained graph to decide if the program is terminating. For functional programs, the size-change graph is a graph comprising functions in the program where the edges express the changes in the values that may be passed as arguments. Step (1) constructs the graph by statically approximating the possible calls that the program would make in its actual execution.

To apply size-change termination verification to higher-order programs, a control flow analysis (CFA) is employed to statically approximate the possible call relations as a call graph and construct a sound size-change graph from the call graph [30,28,29,12]. Therefore, the approach involves *a priori* approximation of the program, and can lead to loss in a precision when a precise graph cannot be constructed by the static analysis. For example, the approach may fail on cases where a non-terminating call depends on a safety property (recall the simple example from Section 1 where a non-terminating function is called if and only if the condition  $p(x)$  is met). By contrast, our approach suffers no a priori loss in precision and is sound and complete.

Like our approach, the size-change approach to higher-order programs [30,28,29,12] can prove termination of programs that require well-foundedness relation over function closure values. For example, Jones and Bohr [12] and Sereni [28] order closure arguments by using the subtree relation on their tree representations. By contrast, we have presented a generic approach that uses implicit parameters and counterexample analysis to infer the appropriate instantiations for the implicit parameters. Our approach is more general in the sense that it is not fixed to one pattern of closure information to be used for the well-foundedness relation. For example, the subtree relation used in [12,28] can be expressed by inferring instantiations that encode the depth of the closures, and our prototype implementation automatically verifies examples in their paper that require such information (cf. Section 4 and the extended report [17]).

On the other hand, we employ counterexample analysis and constraint-based inference to automatically infer the instantiations, and so the approach of Jones, Bohr, and

Sereni that fixes the closure information to a pre-determined pattern may be more efficient on instances that are known to be verifiable with such information.

### 5.3 Term Rewriting

Similar to the size-change approach, the application of termination verification techniques for term rewriting systems to higher-order programs is done in a two-step process [6]. There, in the first step, a static analysis is employed to construct a term rewriting system that soundly approximates the given program such that the program is terminating if the constructed rewriting system is terminating. Then, the second step applies a termination verifier for term rewriting systems [7] to verify termination.

As with the size-change approach, this two-step approach can introduce a loss in precision because of the approximation in the first step. For example, Giesl et al. [6] show a simple program on which their approach fails because of this limitation (Example 4.12 in [6]).

## 6 Conclusion

We have presented a new automated approach to termination verification of higher-order functional programs. In stark contrast to the previous approaches, our approach is sound and complete relative to the soundness and completeness of the underlying reachability analysis and ranking function inference. Our approach is the first sound binary reachability analysis based approach to the termination verification of higher-order programs. The key features of our approach are the novel program transformation that correctly tracks the call-tree ancestor's arguments through the higher-order control flow, and the inference method for ranking functions over higher-order values via implicit parameter instantiation inference.

## References

1. Ball, T., Rajamani, S.K.: The SLAM project: debugging system software via static analysis. In: POPL, pp. 1–3 (2002)
2. Chin, W.N., Khoo, S.C.: Calculating sized types. *Higher-Order and Symbolic Computation* 14(2-3), 261–300 (2001)
3. Cook, B., Podelski, A., Rybalchenko, A.: Abstraction refinement for termination. In: Hankin, C., Siveroni, I. (eds.) SAS 2005. LNCS, vol. 3672, pp. 87–101. Springer, Heidelberg (2005)
4. Cook, B., Podelski, A., Rybalchenko, A.: Termination proofs for systems code. In: PLDI, pp. 415–426. ACM (2006)
5. Cook, B., See, A., Zuleger, F.: Ramsey vs. lexicographic termination proving. In: Piterman, N., Smolka, S.A. (eds.) TACAS 2013. LNCS, vol. 7795, pp. 47–61. Springer, Heidelberg (2013)
6. Giesl, J., Raffelsieper, M., Schneider-Kamp, P., Swiderski, S., Thiemann, R.: Automated termination proofs for Haskell by term rewriting. *ACM Transactions on Programming Languages and Systems* 33(2), 7:1–7:39 (2011)
7. Giesl, J., Thiemann, R., Schneider-Kamp, P., Falke, S.: Automated termination proofs with aProVE. In: van Oostrom, V. (ed.) RTA 2004. LNCS, vol. 3091, pp. 210–220. Springer, Heidelberg (2004)

8. Heizmann, M., Jones, N.D., Podelski, A.: Size-change termination and transition invariants. In: Cousot, R., Martel, M. (eds.) SAS 2010. LNCS, vol. 6337, pp. 22–50. Springer, Heidelberg (2010)
9. Henzinger, T.A., Jhala, R., Majumdar, R., Sutre, G.: Lazy abstraction. In: POPL, pp. 58–70 (2002)
10. Jhala, R., Majumdar, R.: Software model checking. *ACM Comput. Surv.* 41(4) (2009)
11. Jhala, R., Majumdar, R., Rybalchenko, A.: HMC: Verifying functional programs using abstract interpreters. In: Gopalakrishnan, G., Qadeer, S. (eds.) CAV 2011. LNCS, vol. 6806, pp. 470–485. Springer, Heidelberg (2011)
12. Jones, N.D., Bohr, N.: Call-by-value termination in the untyped lambda-calculus. *Logical Methods in Computer Science* 4(1) (2008)
13. Kobayashi, N.: Type-based useless-variable elimination. *Higher-Order and Symbolic Computation* 14(2-3), 221–260 (2001)
14. Kobayashi, N.: Model checking higher-order programs. *Journal of the ACM* 60(3) (2013)
15. Kobayashi, N., Ong, C.H.L.: A type system equivalent to the modal mu-calculus model checking of higher-order recursion schemes. In: LICS, pp. 179–188. IEEE Computer Society (2009)
16. Kobayashi, N., Sato, R., Unno, H.: Predicate abstraction and CEGAR for higher-order model checking. In: PLDI, pp. 222–233. ACM (2011)
17. Kuwahara, T., Terauchi, T., Unno, H., Kobayashi, N.: Automatic termination verification for higher-order functional programs (2013), <http://www-kb.is.s.u-tokyo.ac.jp/~kuwahara/termination>
18. Ledesma-Garza, R., Rybalchenko, A.: Binary reachability analysis of higher order functional programs. In: Miné, A., Schmidt, D. (eds.) SAS 2012. LNCS, vol. 7460, pp. 388–404. Springer, Heidelberg (2012)
19. Lee, C.S., Jones, N.D., Ben-Amram, A.M.: The size-change principle for program termination. In: POPL, pp. 81–92. ACM (2001)
20. Lester, M.M., Neatherway, R.P., Ong, C.H.L., Ramsay, S.J.: Model checking liveness properties of higher-order functional programs. In: Proceedings of ML Workshop 2011 (2011)
21. McMillan, K.L.: Lazy abstraction with interpolants. In: Ball, T., Jones, R.B. (eds.) CAV 2006. LNCS, vol. 4144, pp. 123–136. Springer, Heidelberg (2006)
22. de Moura, L.M., Björner, N.: Z3: An efficient SMT solver. In: Ramakrishnan, C.R., Rehof, J. (eds.) TACAS 2008. LNCS, vol. 4963, pp. 337–340. Springer, Heidelberg (2008)
23. Ong, C.H.L.: On model-checking trees generated by higher-order recursion schemes. In: LICS, pp. 81–90. IEEE Computer Society (2006)
24. Ong, C.H.L., Ramsay, S.: Verifying higher-order programs with pattern-matching algebraic data types. In: Proceedings of POPL 2011, pp. 587–598. ACM (2011)
25. Podelski, A., Rybalchenko, A.: A complete method for the synthesis of linear ranking functions. In: Steffen, B., Levi, G. (eds.) VMCAI 2004. LNCS, vol. 2937, pp. 239–251. Springer, Heidelberg (2004)
26. Podelski, A., Rybalchenko, A.: Transition invariants. In: LICS, pp. 32–41. IEEE Computer Society (2004)
27. Rondon, P.M., Kawaguchi, M., Jhala, R.: Liquid types. In: PLDI, pp. 159–169. ACM (2008)
28. Sereni, D.: Termination analysis of higher-order functional programs. Ph.D. thesis, Magdalen College (2006)
29. Sereni, D.: Termination analysis and call graph construction for higher-order functional programs. In: ICFP, pp. 71–84. ACM (2007)
30. Sereni, D., Jones, N.D.: Termination analysis of higher-order functional programs. In: Yi, K. (ed.) APLAS 2005. LNCS, vol. 3780, pp. 281–297. Springer, Heidelberg (2005)
31. Terauchi, T.: Dependent types from counterexamples. In: POPL, pp. 119–130. ACM (2010)

32. Thiemann, R., Giesl, J.: The size-change principle and dependency pairs for termination of term rewriting. *Appl. Algebra Eng. Commun. Comput.* 16(4), 229–270 (2005)
33. Unno, H., Terauchi, T., Kobayashi, N.: Automating relatively complete verification of higher-order functional programs. In: *POPL*, pp. 75–86. ACM (2013)
34. Wand, M., Siveroni, I.: Constraint systems for useless variable elimination. In: *Proceedings of POPL 1999*, pp. 291–302 (1999)
35. Xi, H.: Dependent types for program termination verification. In: *LICS 2001*, pp. 231–242. IEEE (2001)
36. Xi, H., Pfenning, F.: Dependent types in practical programming. In: *POPL*, pp. 214–227 (1999)
37. Zhu, H., Jagannathan, S.: Compositional and lightweight dependent type inference for ML. In: Giacobazzi, R., Berdine, J., Mastroeni, I. (eds.) *VMCAI 2013*. LNCS, vol. 7737, pp. 295–314. Springer, Heidelberg (2013)