Automating Relatively Complete Verification of Higher-Order Functional Programs

Hiroshi Unno
University of Tsukuba
uhiro@cs.tsukuba.ac.jp

Tachio Terauchi
Nagoya University
terauchi@is.nagoya-u.ac.jp

Naoki Kobayashi
University of Tokyo
koba@is.s.u-tokyo.ac.jp

Abstract
We present an automated approach to relatively completely verifying safety (i.e., reachability) property of higher-order functional programs. Our contribution is two-fold. First, we extend the refinement type system framework employed in the recent work on (incomplete) automated higher-order verification by drawing on the classical work on relatively complete “Hoare logic like” program logic for higher-order procedural languages. Then, by adopting the recently proposed techniques for solving constraints over quantified first-order logic formulas, we develop an automated type inference method for the type system, thereby realizing an automated relatively complete verification of higher-order programs.

Categories and Subject Descriptors D.2.4 [Software Engineering]: Software/Program Verification: F.3.1 [Logics and Meaning of Programs]: Specifying and Verifying and Reasoning about Programs

Keywords Relative Completeness, Higher-Order Programs, Software Model Checking, Type Inference

1. Introduction
Recently, motivated by the success of software model checkers [12] for the automated verification of first-order programs, researchers have proposed “model checkers” for the automated verification of higher-order programs [13, 18, 23, 26, 27]. Interestingly, they have all been formulated as a form of refinement type inference.¹

The refinement type systems employed in the automated higher-order program verification have some important differences from the ones from the original, non-automated approaches like DML [30], making them more amenable to automation (such as the lack of implicit II types). But, they follow the non-automated ones in that the types embed first-order logic (FOL) formulas, called refinement predicates, over program values (as in dependent types) that are used to express and enforce detailed properties of the program.

For example, consider the OCaml program shown in Figure 1. Here, * denotes a non-deterministic choice. Given an integer argument i to main, the program recursively calls app non-deterministically to apply the closure check i to i + j where j is the number of times the then branch is taken in app. The program is safe in that assert false is unreachable for any argument i. The program is difficult to verify via first-order program verification methods because of the higher-order recursive function app.

Recent advances in higher-order program verification have enabled automated verification of such programs via refinement type inference. For the program above, the following refinement types may be automatically inferred to verify its safety.²

```
let rec app x f = if * then app (x+1) f
  else f x in
let check x y = if x <= y then ()
  else assert false in
let main i = app i (check i)
```

Figure 1. A simple higher-order program.

¹ In fact, the equivalence of model checking and refinement type inference has been shown for the finite domain base-type data case [16, 17].

² The type syntax is borrowed from Augustsson [1].

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

POPL’13, January 23–25, 2013, Rome, Italy.
Copyright © 2013 ACM 978-1-4503-1832-7/13/01…$10.00
We informally describe the incompleteness issue by showing how previous automated approaches fail to type the programs shown in Figures 2 and 3.

First, let us try to type Figure 2. Here, the goal is to show that `main` can be given the type `i:int -> unit`, that is, `main` is safe to be called with any integer `i`. (For a base type `B`, we often abbreviate the refinement type `{u:B | \tau}` as `B`. E.g., `int = {u:int | \top}`.) Therefore, we try to type the body of `main` under the assumption that the type of `i` is `{u:int | \top}`. The most precise type for the partial application `check i` is `{u:int | u \geq i} -> unit` (i.e., functions that can take any integer at least as large as `i`), and the most precise type for `i` is `{u:int | u = i}` (i.e., integers equal to `i`). (Intuitively, `{u:B | \theta}` expresses values of the base type `B` satisfying the refinement predicate `\theta`, and `x: \tau \rightarrow \sigma` expresses functions that return a value of the type `\sigma[e/x]` when given the argument `e` of the type `\tau`.)

We show that the type systems fail to give the higher-order function `app` a sufficiently precise type to verify the program's safety. The type of `app` must be of the form

\[ f:{\{u:int | \theta\} \rightarrow unit} \rightarrow x:{\{u:int | \phi\} \rightarrow unit} \]

where the refinement predicates `\theta` and `\phi` are FOL formulas in the theory of base-type data.\(^3\) In addition, as discussed below, to prevent degenerate types, the refinement type systems enforce an important well-formedness condition that restricts the variables that can appear free in a refinement predicate. (Here, `\theta` and `\phi` are restricted so that `fv(\theta) \subseteq \{u\}` and `fv(\phi) \subseteq \{f,u\}`, where `fv(\theta)` denotes the free variables of `\theta`.)

From the application `f x` in the body of `app`, the type system asserts that `\phi \Rightarrow \theta`, that is, `f` must be safe to be given the argument `x`. In the body of `main`, `app` is applied to `check i` and `i`, and the type system asserts that the type of `check i` is a subtype of `{u:int | \theta} -> unit` and the type of `i` is a subtype of `{u:int | \phi}`. This leads to the constraints `u = i \Rightarrow \phi \land \theta \Rightarrow u \geq i`, accurately expressing the fact that `i` is passed as the second argument to `app` (i.e., `check i`) expects a value at least as large as `i`.

A solution to the set of constraints is `\theta \equiv u \geq i` and `\phi \equiv u = i`, but this is degenerate because `main`'s variable `i` would appear in `app`'s type. (Recall that `\theta`, `\phi` are refinement predicates of `app`'s type.) Intuitively, it means that the function `app` could "see" values from a specific context of its use (i.e., `main`). Indeed, as remarked above, the refinement type systems disallow such degenerate types via the well-formedness condition, and only allow `\theta`, `\phi` such that `fv(\theta) \subseteq \{u\}` and `fv(\phi) \subseteq \{f,u\}`. That is, the refinement predicates are allowed to only mention the variables in their respective scopes. Consequently, the program is (conservatively) rejected as untypable. It is worth noting that the program of Figure 1 does not have this issue as `app`'s arguments are conveniently ordered so that `app`'s function-type argument could depend on the base-type argument to allow the following sufficiently precise type:

\[ x:int \rightarrow f:{\{u:int | u \geq x\} \rightarrow unit} \rightarrow unit \]

Unfortunately, as exemplified by Figure 3, the incompleteness issue is not just a matter of choosing the right order of function arguments. Here, the program is untypable even if we were allowed to change the order of the arguments, and the untypability comes from not being able to give a precise enough type to `app3`. Ideally, we would like to express via `app3`'s type the fact that `g` is safe when applied to `f` (and also when applied to `app2` for any `j > 0`). But, this requires a type that is parametric in `f` and `g`'s behavior, and that cannot be given because the only parameters of `app3` are `f` and `g`, and they are both function-type arguments. Note that refinement predicates of the form "`g` is safe to be called with `f`" are prohibited as they are required to be FOL formulas over the base-type data.

\(^3\) We could also give non-T refinement predicates to `unit`, but it does not affect the example.
2.1 Our Approach

Our approach to solving the incompleteness issue is inspired by the research on relatively complete program logic for higher-order procedural languages [6–8]. The main idea is to add extra dummy base-type parameters that are instantiated appropriately so that the refinement type of a higher-order function can depend on the parameters.

For example, for the program of Figure 2, we add an extra base-type parameter \( a \) to the program and obtain the following program:

\[
\begin{align*}
&\text{let rec app a f x = if * then app[a] f (x+1)} \\
&\quad \text{else f x in} \\
&\text{let check x y = if x <= y then ()} \\
&\quad \text{else assert false in} \\
&\text{let main i = app[i] (check i) i}
\end{align*}
\]

Here, for clarity, the parameter passings for the extra parameters are written as quantifier instantiations, but they may be understood as ordinary function applications (e.g., \( \text{app} [a] \) is \( \text{app} \)). With the addition of the extra parameter, it becomes possible for the existing refinement type systems designed for automated verification to encode the research on relatively complete program logic for higher-order software model checker (i.e., refinement type inference system) MoChI [18]. Section 6 discusses related work, and Section 7 concludes. Appendix contains the materials omitted from the main body of the paper.

2.2 Paper Organization

The rest of the paper is organized as follows. Section 3 presents the refinement type system along with the target functional language, and proves its relative completeness under the extra parameter addition strategy. Section 4 presents the type inference algorithm. Section 5 presents the preliminary experience with the type inference algorithm, implemented as an extension to the higher-order software model checker (i.e., refinement type inference system) MoChI [18]. Section 6 discusses related work, and Section 7 concludes. Appendix contains the materials omitted from the main body of the paper.

3. Language and Type System

We formalize the target programming language. Figure 4 shows the syntax. The language is essentially the simply-typed lambda calculus with recursion and primitives for integer arithmetic operations. For simplicity the only base-type data in this language is integers, but other base types and their operations can be encoded in the standard way: for example, \( \text{true} = 1 \) and \( \text{false} = 0 \).

A program, \( d \), is a finite set of function definitions, \( F \), defining a function named \( F \) with the formal parameters \( \overline{x} \) and the body \( e \). The notation \( \overline{x} \) denotes a possibly empty sequence. The functions are mutually recursive in that the body of a function may refer to other functions, including itself. Each function is closed except for the free function names (i.e., functions are lambda lifted [15]). We also assume that \( d \) contains a function named \( \text{main} \) that only takes base-type (i.e., integer) arguments.

Expressions, \( e \), comprise non-deterministic branches, let expressions, constants \( c \), function names \( F \), variables \( x \), and (constant or user-defined function) applications \( e \). Constants include integer constants such as \( -1, 0, 1, 2 \), and integer operations such as \( + \) and \( \leq \) (recall that we model booleans via integers). We assume that there are unary constant operators named \( \text{assert} \) and \( \text{assume} \).

We restrict expressions to be in continuation passing style (CPS) so that they are non-returning, except when they occur let-bound (i.e., occurs as \( c_1 \) in \( \text{let} \ x = c_1 \in e_2 \)). By contrast, let-bound expressions are restricted to be non-CPS (i.e., value returning) expressions, which are total applications of constant operators including...
the assume and assertion expressions \( \text{assume} \) and \( \text{assert} \), and partial applications (of user-defined functions). We restrict constant operator applications to be total. CPS/A-normal-form-style is used for simplicity. Direct-style syntax can be supported by CPS conversion.

The rest of the syntax is straightforward. As usual, applications associate to the left so that \( c_0 \cdot c_1 \cdot c_2 = (c_0 \cdot (c_1 \cdot c_2)) \). We write \( c_0 \text{app} \) for the series of applications \( c_0 \cdot c_1 \cdot c_2 \ldots c_n \) where \( \text{app} \) is \( e_1, e_2, \ldots, e_n \). We write \( e_0 \text{app} \) for \( \text{let } x = e_1 \text{ in } e_2 \) such that \( x \notin \text{fv}(e_2) \), where \( \text{fv}(e) \) denotes the free variables of \( e \). Without loss of generality, we assume that bound variables are distinct. Note that, while the language only has non-deterministic branches, a conditional branch if \( x \) then \( e_1 \) else \( e_2 \) can be encoded as

\[
\text{if } \neg \text{assume } x; e_1 \text{ else let } y = \neg x \text{ in } \text{assume } y; e_2
\]

which is equivalent for assertion safety.

We define the operational semantics of the language as a small-step reduction relation from states to states. A \textit{state} is a run-time expression \( e \) that extends the source expressions with non-variable arguments at function applications, a special failure state, and a special safe state. (We overload the symbol \( e \) to range over run-time expressions when it is clear from the context.)

The reduction relation \( \rightarrow \) is defined by the rules shown in Figure 5. Here, \( e \rightarrow e' \) denotes the evaluation of the value returning expression \( e \) to the value \( e' \), and is defined as \( F \text{app} \rightarrow \text{app} F \text{app} \) and \( e \rightarrow e' \) where \( \text{app} \) is the evaluation denoting the semantics of a non-assume/assert constant \( c \), so that, for example \( \text{app} (i+j, i) = i+j \) for integers \( i \) and \( j \). (An integer constant is represented by a 0-ary constant.) We let partial applications (i.e., \( F \text{app} \)) represent function closures. Formally, a value is either an integer constant or a function closure.

The semantics of a program \( d \) is defined as a series of reductions starting from an initial state \( \varepsilon_{\text{main}}[\downarrow / \text{app}] \) where \( \varepsilon_{\text{main}}[\downarrow / \text{app}] \) and \( \overline{\tau} \) are integer arguments for \( \text{main} \) (i.e., \( [\overline{\tau}] = [\downarrow / \text{app}] \)). Note that, because of CPS, reductions only occur at the top level. We write \( e \rightarrow^* e' \) for zero or more reductions from \( e \) to \( e' \). We say that a program is \textit{assertion} safe if its evaluation does not cause an assertion failure, that is, if \( \varepsilon_{\text{main}}[\downarrow / \text{app}] \not\rightarrow^* \text{fail} \) fail for any arguments \( \overline{\tau} \) of \( \text{main} \).

We assume that a program is typed under the standard simple type system whose type grammar is shown below.

\[
s ::= \star \mid \text{int} \mid s \rightarrow s'
\]

Here, the type \( \star \) represents the type of a CPS expression. For each expression \( e \) in the program, we write \( \text{sty}(e) \) to denote its simple type.

The typability in the simple type system assures that the program does not “get stuck”, for example, by trying to use an integer as a function, but it does not guarantee its safety. Therefore, a program either runs forever safely (due to CPS), a program cannot return, stops safely in the state safe by reaching a false \text{assume}, or aborts with an assertion failure. The typing rules for the simple type system are standard and are deferred to Appendix A.
\[
\begin{align*}
\Gamma \vdash \sigma_2 \leq \sigma_1 & \quad \Gamma, x : \sigma_2 \vdash \tau_1 \leq \tau_2 \\
\Gamma \vdash x : \sigma_1 & \rightarrow \tau_1 \leq x : \sigma_2 \rightarrow \tau_2 \\
\text{Sf} & \\
\Gamma \vdash \star \leq \star & \\
\Gamma \vdash u \notin \text{fv}([\Gamma]) & \quad ([\Gamma] \land \theta_1) \Rightarrow \theta_2 \\
\text{Sb} & \\
\end{align*}
\]

Figure 8. The subtyping rules.

Figure 7 shows the typing rules. The judgements for the expressions are of the form \(\Gamma \vdash e : \tau\) where \(\Gamma\) is a type environment mapping variables and function names to types.

We discuss each typing rule. VB types base-type variables. Note that the rule ignores the environment. Expressibility is not reduced, however, because the assumption about \(x\) in the environment gets discharged at subtyping. VF types function-type variables and function names by looking up the environment. Here, \(\rightarrow\) denotes the set of simple function types. (We use the meta variable \(\kappa\) to range over both variables and function names.) CST types constants. Here, \(ty(c)\) denotes the pre-assigned type of the constant \(c\) such that

\[
ty(\text{assume}) = x : \text{int} \rightarrow \{u \mid x = 1 \land u = 1\} \\
ty(\text{assert}) = x : u = 1 \rightarrow \{u \mid u = 1\}
\]

and \(ty(c)\) is the precise type for a non-assume/assert constant \(c\) (e.g., \(ty(+) = x : \text{int} \rightarrow y : \text{int} \rightarrow \{u \mid u = x + y\}\)).\footnote{See Appendix B for the definition of a precise constant type. Also, non-precise but sound constant types, that over-approximate the actual semantics, are sufficient for soundness.} Let is self-explanatory. APP types applications. Here, \(\gamma(x/y)\) is the usual capture-avoiding substitution. IF types branch expressions.

Sub is the subsumption rule. The subtyping relation is defined in Figure 8. In SB, \([\Gamma]\) is the FOL formula denoting the assumptions about the base-type variables in \(\Gamma\), and is defined as follows.

\[
[\Gamma] = \bigwedge_{\forall x : \{u \mid u \in \mathcal{U}\} \in \text{cl}} 0^{[\theta/u]}
\]

Like the rule of consequence of the Hoare logic [10], SB asks the client verifier (i.e., the type inference system) to decide the validity of arbitrary FOL arithmetic formulas. As expected, our completeness result is relative to the hypothetical completeness of deciding this. In verification practice, one may settle for incomplete theorem proving or a decidable theory subset to make type checking decidable.\footnote{Also, type checking being decidable is different from type inference (i.e., typability) being decidable. The previous research has proposed various incomplete inference approaches [13, 18, 26, 27].}

Sub enforces well-formedness of the introduced type by asserting that its free variables appear bound in the environment. As shown below, the top-level types (i.e., the types of the recursive functions) are required not to contain free variables.

**Definition 3.1 (Well-formed type).** We say that a type is closed if it has no free variables. Let \(\Delta\) be a top-level type environment mapping function names to types. We say that \(\sigma\) is a well-formed type for \(F\) if \(\sigma\) is closed and \(\text{shape}(\sigma) = \text{ty}(F)\). We say that \(\Delta\) is a well-formed top-level type environment if \(\Delta(F)\) is well-formed for each \(F\).

Let us write \(\overline{x : \tau} \rightarrow \tau\) to abbreviate the function type \(x_1 : \sigma_1 \rightarrow \cdots \rightarrow x_n : \sigma_n \rightarrow \tau\) where \(\overline{x : \tau} = x_1 : \tau_1, \ldots, x_n : \tau_n\). We define the notion of a well-typed program.

**Definition 3.2 (Well-typed program).** We write \(\Delta \vdash d\) if the following conditions hold.

1. \(\Delta\) is a well-formed top-level type environment.

2. For each function \(F \overline{x : \tau} \rightarrow e \in \Delta\), we have \(\Delta, \overline{x : \tau} \vdash e : \tau\) where \(\Delta(F) = \overline{x : \tau} \rightarrow \tau\).

3. \(\Delta(\text{main})\) is of the form \(x : \text{int} \rightarrow \tau\). (I.e., the refinement predicates for the arguments of main are all \(\top\).

A program \(d\) is said to be well-typed (equivalently, typable) if there exists \(\Delta\) such that \(\Delta \vdash d\).

The condition (2) says that \(\Delta\) contains fixed-point types for the recursive functions comprising \(d\), and (3) says that \(\text{main}\) is safe to be called with any arguments (recall that \(\text{main}\) is a function over integer arguments).

The type system is sound in that it ensures that a well-typed program does not cause an assertion failure.

**Theorem 3.3 (Soundness).** If \(\Delta \vdash d\) then \(d\) is safe.

The theorem follows from the soundness of the refinement type system extended with extra parameter additions (Theorem 3.6).

**Example 3.4.** Let the program \(d_1\) consist of the following functions.

\[
\begin{align*}
\text{app} & = f \Rightarrow \text{if } \text{app}(x + 1) \text{ then } f \text{ else } f \\
\text{check} & = x \Rightarrow \text{if } \text{assume}(x \leq y) \text{ then } \text{check} \text{ else } \text{assume}(-x \leq y) \Rightarrow \text{assert} 0 \Rightarrow \text{check} \text{ x x} \\
\text{main} & = \text{app} i \Rightarrow \text{check} i
\end{align*}
\]

The program is the one from Figure 1 translated into the target language. (We elide \(\lambda\)-normalization for readability.) Let \(\Delta\) be the following type environment.

\[
\begin{align*}
\Delta(\text{app}) & = x : \text{int} \rightarrow f : (\{u \mid u \geq x\} \rightarrow \ast) \rightarrow \ast \\
\Delta(\text{check}) & = x : \text{int} \rightarrow y : (\{u \mid u \geq x\} \rightarrow \ast) \rightarrow \ast \\
\Delta(\text{main}) & = i : \text{int} \rightarrow \ast
\end{align*}
\]

It is routine to check that \(\Delta \vdash d_1\). Therefore, \(d_1\) is typable and is safe.

Next, consider \(d_2\) shown below.

\[
\begin{align*}
\text{app} & = f \Rightarrow \text{if } \text{app}(x + 1) \text{ else } f \\
\text{check} & = x \Rightarrow \text{if } \text{assume}(x \leq y) \text{ then } \text{check} \text{ else } \text{assume}(-x \leq y) \Rightarrow \text{assert} 0 \Rightarrow \text{check} \text{ x x} \\
\text{main} & = \text{app} (\text{check} i) i
\end{align*}
\]

The program is a translation of the program from Figure 2. The program is safe but untypable (as remarked in Section 2, also under the previous refinement type systems [13, 18, 23, 26, 27]). Note that we cannot simply assign the type \(\Delta(F)\) app from the above but with the order of \(x\) and \(f\) reversed so that

\[
\begin{align*}
\text{app} & = f \Rightarrow \text{if } \text{app}(x + 1) \text{ then } f \text{ else } f \\
\text{check} & = x \Rightarrow \text{if } \text{assume}(x \leq y) \text{ then } \text{check} \text{ else } \text{assume}(-x \leq y) \Rightarrow \text{assert} 0 \Rightarrow \text{check} \text{ x x} \\
\text{main} & = \text{app} (\text{check} i) (\text{app} i)
\end{align*}
\]

The type is not closed and therefore is not a well-typed top-level type (cf. Definition 3.1). The well-formedness condition forces the refinement predicates to only refer to the values passed earlier. Well-formedness is not a superficial restriction: In the presence of higher-order functions and function closures (i.e., partial applications), we cannot generally determine “up front” in the program logic what will be passed later to a closure.

The issue is apparent in the following program \(d_3\), which is a translation of Figure 3.

\[
\begin{align*}
\text{app3} & = g \Rightarrow \text{if } \text{app3} (\text{succ} f) \text{ g else } g \\
\text{app} & = f \Rightarrow \text{if } \text{app} f \text{ f else f} \\
\text{check} & = x \Rightarrow \text{if } \text{assume}(x \leq y) \text{ then } \text{check} \text{ else } \text{assume}(-x \leq y) \Rightarrow \text{assert} 0 \Rightarrow \text{check} \text{ x x} \\
\text{main} & = \text{app3} (\text{check} i) (\text{app} i)
\end{align*}
\]

As remarked in Section 2, the program is safe but untypable, even if we were allowed to change the order of the function arguments. Here, it is not possible to determine what will be passed to the
We extend the typing rules to (non-A-normal-form) side-effect-free closures. This is incomplete by itself. While a complete checking of safety is represented as universally bound variables so that they can be easily distinguished from ordinary parameters.

We extend the notion of well-typed program to accommodate universal quantifier bindings.

**Definition 3.5 (Well-typed program – Extended).** We write \( \Delta \vdash \gamma \) if the following conditions hold.

1. \( \Delta \) is a well-formed top-level type environment.
2. For each function \( F \vdash \gamma = \epsilon \in d \), we have \( \Delta, \langle \gamma \rangle \vdash \gamma : \epsilon \)
   where \( \Delta(F) = \gamma \rightarrow \epsilon \).
3. \( \Delta(\text{main}) \) is of the form \( x:\text{int} \rightarrow \epsilon \).

A program \( d \) is said to be well-typed if \( \exists \Delta.\Delta \vdash \gamma \vdash d \).

Note that the only differences from Definition 3.2 are \( \langle \gamma \rangle \) in (2) which allow universal quantifiers in the type of the functions. Also, note that the well-formedness condition now allows the refinement predicates to refer to the universally quantified variables in their scope.

We state the soundness of the extended refinement type system.

**Theorem 3.6 (Soundness – Extended).** If \( \Delta \vdash \gamma \vdash d \) then \( d \) is safe.

The proof is standard [23, 27, 29] and is deferred to Appendix D. As remarked in Section 2, for relative completeness, it suffices to limit the position of extra parameters (i.e., universal quantifiers) to one before each function-type parameter. To this end, we define \( \text{ushape}(\sigma) \) to be a simple type such that, for a function of the simple type \( \text{ushape}(\sigma) \), \( \sigma \) has just the sufficient extra parameters. Formally, \( \text{ushape}(\sigma) \) is defined by the rules below.

\[
\text{ushape}([u \theta]) = \text{int} \quad \text{ushape}(\epsilon) = \star
\]

\[
\text{ushape}(x : \sigma \rightarrow \tau) = \text{ushape}(\sigma) \rightarrow \text{ushape}(\tau)
\]

\[
\text{ushape}(\forall \sigma : x : \sigma \rightarrow \tau) = \text{ushape}(\sigma) \rightarrow \text{ushape}(\tau)
\]

Let \( \text{ushape}(\sigma) \) be defined only for \( \sigma \) that has one extra parameter just before a function-type parameter. Also, for \( \sigma \) such that \( \text{ushape}(\sigma) \) is defined, we have \( \text{ushape}(\sigma) = \text{tshape}(\sigma) \). We are now ready to state the relative completeness theorem, which says that if a program is safe then it can be typed with extra parameter additions (even when their positions are restricted to the pattern above).

**Theorem 3.7 (Relative Completeness).** If \( d \) is safe, then there exists \( \Delta \) such that \( \Delta \vdash \gamma \vdash d \) and \( \text{ushape}(\Delta(F)) = \text{sty}(F) \) for each \( F \vdash \gamma = \epsilon \in d \).

We defer the proof to Appendix E. The proof adopts the ideas from the work on relatively complete program logics for higher-order procedural languages [6–8] that instantiate the extra parameters by the base-type encoding of the function closures. It is worth noting that, as a corollary of Theorem 3.7, it follows that the refinement type system is relatively complete for first-order programs (i.e., programs without function-type parameters) even without the extra parameter extension.

**Example 3.8.** We show how \( \vdash \gamma \) types \( d_2 \) and \( d_3 \) from Example 3.4. First, we have \( \Delta_2 \vdash \gamma \vdash d_2 \) where

\[
\Delta_2(\text{app}) = \forall u : \text{int}. f : ([u \mid u \geq a] \rightarrow \star) \rightarrow x : [u \mid u \geq a] \rightarrow \star
\]

\[
\Delta_2(\text{check}) = x : \text{int} \rightarrow y : [u \mid u \geq x] \rightarrow \star
\]

\[
\Delta_2(\text{main}) = i : \text{int} \rightarrow \star
\]
Note that the types correspond to the ones used to type Figure 2 in Section 2.1. Also, note that \( \text{ustype}(\Delta_2(\text{app})) = \text{sty}(\text{app}) \), \( \text{ustype}(\Delta_2(\text{check})) = \text{sty}(\text{check}) \), and \( \text{ustype}(\Delta_3(\text{main})) = \text{sty}(\text{main}) \). It is easy to see that check can be given the type \( \Delta_3(\text{check}) \). We show that app can be given the type \( \Delta_2(\text{app}) \). We type the body of app under the type environment \( \Delta_2, a \text{ int. f} : \{ u | u \geq a \} \rightarrow \star, x : \{ u | u \geq a \} \). Therefore, the application \( f \times \text{x -type checks to give the type } \star \). To type the other branch, \( \text{app } f (x + 1) \), we instantiate \( \Delta_2(\text{app}) \) with \( a \times \text{app} \) to give the type \( \{ u | u \geq a \} \rightarrow \star \rightarrow \{ u | u \geq a \} \rightarrow \star \). The type is then used to give the application \( \text{app } f \times \text{the } x : \{ u | u \geq a \} \rightarrow \star, x : \{ u | u \geq a \} \). Finally, \( \text{main} \) can be given the type \( \Delta_2(\text{main}) \) by instantiating \( \Delta_2(\text{app}) \) with \( i \).

Likewise, let \( \Delta_3 \) be the following type environment.

\[
\begin{align*}
\Delta_3(\text{app}) &= \forall a : \text{int. f} : \tau_1 \rightarrow \forall b : \text{int. g} : \tau_2 \rightarrow \star \\
\Delta_3(\text{check}) &= \forall a : \text{int. f} : \{ u | u \geq a \} \rightarrow \star \\
\Delta_3(\text{main}) &= \forall i : \text{int} \\
\end{align*}
\]

such that \( \tau_1 = \{ u | u \geq a \} \rightarrow \star, \tau_2 = \forall c : \text{int.} \tau_1 \rightarrow \star, \) and \( \sigma = \{ u | u \leq b \} \). Note that \( \text{ustype}(F) = \text{sty}(\Delta_3(F)) \) for each \( F = \text{app3, app, succ, check, and main} \). We show that \( \Delta_3 \vdash \forall \text{d3} \) by following the same instantiation scheme used to type Figure 3 in Section 2.1. We type the then branch of \( \text{app3} \) by instantiating \( \text{app3} \) and \( \text{succ} \) with \( a \), and instantiating the resulting \( \text{app3} \) \( \text{succ} \) \( \text{f} \) with \( a \) (or with an arbitrary side-effect-free integer-type expression, because \( b \) does not appear free in its scope in \( \Delta_3(\text{app3}) \)). The else branch can be typed by instantiating \( g \) with an arbitrary side-effect-free integer-type expression (because \( c \) does not appear free in its scope in \( \Delta_3(\text{app3}) \)). Then, to type \( \text{main} \), we instantiate \( \text{app3} \) with \( i \), and instantiate \( \text{app3} \) \( \text{check} i \) with an arbitrary side-effect-free integer-type expression.

The examples show that instantiating via simple expressions are often sufficient for verifying safety, and not all quantifiers may even be needed. (Contrast this with the instantiation via Gödel numbering used in the proof of Theorem 3.7.) We take advantage of the observation in the type inference method described in Section 4.

4. Type Inference

The type inference framework is based on, and extends the recent work on the counterexample-guided approach to refinement type inference [18, 26] (but, the idea may be adopted to extend the back-end of other refinement type inference systems like [13, 27]). To allow smooth adoption and use the existing type inference algorithms mostly as a blackbox, we implement the extra parameter addition as a program translation as in Section 2, instead of modifying the underlying type system to model it by universally quantified types as in Section 3.2.

Figure 10 shows the overview of the type inference process. In Step 1, we translate the given program by adding extra parameters as done in Section 2. Here, we maintain and use a parameter substitution for instantiating the extra parameters. We initialize the instantiation expressions to arbitrary constants (e.g., 0), and then perform a counterexample-guided refinement type inference over the translated program.

Following the counterexample-guided abstraction refinement (CEGAR) scheme popularized in model checking [3, 12], in the counterexample-guided refinement type inference, we maintain a set of candidate refinement types (or the refinement predicates that comprise them). Then, we try to find a type assignment for the program within the candidates via the standard fixed-point type inference process (Step 2 in Figure 10). The fixed-point type inference uses an automated FOL theorem prover, such as a SMT solver, to decide the FOL implications discharged at the subtyping rule \( \text{SB} \) We refer interested readers to the previous work [18, 26] for details.

A counterexample is generated when the candidates are found insufficient, and we analyze the reason for the failure to either validate the counterexample (Step 3), or update the candidate types or the parameter substitution to eliminate the spurious counterexample (Step 4). In the refinement step, the previous work [18, 26] only generated new candidate types (i.e., the arrow from Step 4 to Step 2). But, because of the incompleteness of the underlying type system, some spurious counterexamples are impossible to refute by just adding new candidate types. The key component of our new type inference method is generating new parameter substitutions in such a situation (i.e., the arrow from Step 4 to Step 1).

Then, the CEGAR process is repeated with the updated candidate types or the updated parameter substitutions. The former case is identical to the previous work, and in the latter case, we repeat from Step 1 to re-translate the program with the updated parameter instantiations.

4.1 Adding Extra Parameters

We add an extra parameter just before each function-type parameter. Also, we assign a unique label to each sub-expression of the program where a parameter instantiation occurs. (Note that such places are syntactically determined.) Let \( L \) be the set of the labels and \( \text{bvs}(l) \) denote the base-type variables that can occur free in the scope of the sub-expression with the label \( l \in L \).

The type inference process maintains the parameter substitution \( P \) that maps each \( l \) to a side-effect-free integer-type expression \( o \) such that \( \text{f}(o) \subseteq \text{bvs}(l) \). We use \( P(l) \) as the instantiation expression at \( l \). We restrict the range of \( P \), and therefore, the instantiation expressions, to linear arithmetic expressions. That is, \( P(l) \) is restricted to be of the form

\[
c_0 + c_1 x_1 + \cdots + c_n x_n
\]

where \( \text{bvs}(l) = \{ x_1, \ldots, x_n \} \) and \( c_0, \ldots, c_n \) are integer constants. This is an important design choice motivated both by the desire to minimize the complexity of the type inference, and also by the ob-
To infer a new parameter substitution $P_T$ for the SHP so that the SHP with each of its instantiation site $i$ instantiated with $P_T(i)$ is typable. Then, we update the parameter substitution with $P_T$, that is, we set $P := P_R \cup P_T$. where $L_{\text{SHP}}$ is the labels of the SHP. And, the CEGAR process returns to the parameter addition phase to re-translate the program (i.e., Step 1 of Figure 10).

The parameter substitution inference proceeds as follows. First, we assign labels to the SHP such that every copy of a sub-expression gets the same label as the one in the original. (Note that because of function copying, the same expression in the original program can have multiple copies in the SHP.) Next, we introduce a parameter substitution template, $P_T$, that maps each $i \in L_{\text{SHP}}$ to a linear arithmetic expression template $P_0 + p_1 x_1 + \cdots + p_n x_n$, where $bvs(l) = \{x_1, \ldots, x_n\}$ and $p_0, \ldots, p_n$ are fresh integer variables. This, in turns, induces a SHP template that has $P_T(i)$ as the instantiation expression at each $i \in L_{\text{SHP}}$.

Now, the problem of inferring $P_T$ is reduced to that of inferring an appropriate integer substitution for the integer variables $p_0^{\prime}$ in the template. We face the problem to constraint solving by generating a FOL constraint over $\mathcal{P}_T$ of the form $\forall \mathcal{P}_T. \theta$ (for some non-linear FOL formula $\theta$ over the variables $p_0^{\prime}$ and $\mathcal{P}_T$). That is, the SHP is typable with the instantiation expression $P_T(i)$ at each $i$ if $\forall \mathcal{P}_T. \theta$ holds where $\theta$ is a substitution that maps $p_0^{\prime}$ to integer constants. Although the problem of solving quantified integer non-linear FOL constraints is undecidable in general, the recently proposed constraint solving techniques based on Farkas’ lemma [4, 9, 24] have shown effective for our application in many cases. We describe the constraint generation process in Section 4.2.1 and the constraint solving process in Section 4.2.2.

### 4.2 Counterexample-Guided Refinement

In MoCHi [18], a counterexample is a straightline higher-order program (SHP) that is untypable with the current candidate types. SHP is defined to be a recursion-free slice of the target program obtained by copying functions and removing branches so that it contains no branches and every function occurrence is “linear” (i.e., each function is called only once). Intuitively, a counterexample corresponds to the (abstract) program path taken to reach the assertion failure.

We check if the counterexample SHP is feasible. This part is identical to the previous work [18] and is done by symbolically evaluating the SHP. If it is feasible, then the program is determined unsafe and we are done. Otherwise, we attempt to infer refinement types for the SHP. If the SHP is found typable, then the inferred refinement types (or the refinement predicates embedded therein) are added to the candidates and the CEGAR iteration returns to the fixed-point type inference phase (i.e., Step 2 of Figure 10).

We infer types for counterexamples by using the techniques from the previous work [18, 26, 27] that reduce the inference problem to solving constraints over FOL formulas and predicate variables such that the predicate variables serve as placeholders of the refinement predicates to be inferred. We refer to the previous work for details.

#### 4.2.1 Constraint Generation

The constraint generation algorithm is essentially the same as the ones from the previous work [18, 26, 27] used for candidate type inference. We give a brief review of the algorithm. For each function in the program, we prepare type templates containing predicate variables that serve as placeholders of the refinement predicates to be inferred. Then, we generate constraints over the predicate variables in the standard way. That is, we apply the typing rules to the SHP but restricting the application of the subsumption rule $\text{Sub}$ to just the argument position of function applications (cf. Section 3.1). This generates constraints over the type templates, which, in turn, reduce to Horn-clause-like constraints over the predicate variables. (See Appendix F for the formal definition of the constraint generation rules.) Because SHP is recursion-free and linear, the generated Horn clauses are non-recursive, and we can obtain an equivalent FOL formula.

#### 4.2.2 Constraint Solving

To infer a new instantiation expression, we prepare the parameter substitution template $P_T$ for the instantiation sites of $d_S$, and
from it, obtain the SHP template $d_T$ shown below.

\[
\text{app1 } a \ x \ f \ x \Rightarrow \text{app2 } (p_0 + p_1a + p_2x) \ f \ (x + 1)
\]
\[
\text{app2 } a \ x \ f \ x
\]
\[
\text{check } x \ y \ \Rightarrow \text{assume } -(x \leq y); \text{assert } 0
\]
\[
\text{main } = \text{app1}(p_1 + p_4)(\text{check } i)
\]

where $p_0, \ldots, p_4$ are free integer variables. Next, we generate the constraints by preparing the type template \( \Delta \) such that

\[
\Delta(\text{app1}) = \alpha: \{u \mid P_1(u) \Rightarrow \{u \mid P_2(a, u) \Rightarrow \star\} \Rightarrow \{u \mid P_3(a, u) \Rightarrow \star\}
\]
\[
\Delta(\text{app2}) = \alpha: \{u \mid P_1(u) \Rightarrow \{u \mid P_2(a, u) \Rightarrow \star\} \Rightarrow \{u \mid P_3(a, u) \Rightarrow \star\}
\]
\[
\Delta(\text{check}) = x : \{u \mid P_4(u) \Rightarrow \{u \mid P_5(x, u) \Rightarrow \star\}
\]
\[
\Delta(\text{main}) = i : \text{int} \Rightarrow \star
\]

This generates the following non-recursive set of Horn-clause-like constraints on the predicate variables $P_1, \ldots, P_k$.

\[
\forall a, x, P_1(a) \land P_2(a, x) \Rightarrow P_3(p_0 + p_1a + p_2x)
\]
\[
\forall x, u, P_1(a) \land P_2(a, x) \land P_3(p_0 + p_1a + p_2x, u) \Rightarrow P_4(a, u)
\]
\[
\forall x, y, P_1(a) \land P_2(a, x, y) \Rightarrow P_4(a, x)
\]
\[
\forall a, x, \text{check } x \Rightarrow P_5(x, a)
\]
\[
\forall a, x, \text{check } x \Rightarrow P_5(x, a, i)
\]
\[
\forall i, \text{check } i \Rightarrow P_6(p_3 + p_4, i)
\]

Systematically simplifying the constraints by computing the least solutions for the predicate variables in a bottom-up manner, we obtain an equisatisfiable FOL formula on the variables $p_0, \ldots, p_4$ shown below.

\[
\forall x, y, z.
\]
\[
p_{14}x = p_{14}z \land p_{14}p_{14}x + p_{14}z = (p_{14}p_{14} + p_{14})(y - 1) \Rightarrow x \leq y
\]

### 4.2.2 Constraint Solving

The constraint generation process above returns a constraint of the form $\forall \theta \exists \phi$, where $\theta$ is a quantifier-free non-linear FOL formula over the variables $\overline{P}$ and $\overline{\theta}$. (More precisely, $\theta$ is linear over $\overline{P}$ with coefficients over $\overline{\theta}$.)\(^{11}\) The goal of the constraint solving phase is to find an assignment $\rho$ for $\overline{\theta}$ satisfying $\forall \theta \exists \phi$.

We solve the constraints by adopting Gulwani et al.’s approach [9]. We give a brief overview of the idea. First, we use Farkas’ lemma to remove the universal quantifications in the constraint $\forall \theta \exists \phi$, and obtain a constraint of the form $\exists \exists \phi$, where $\phi$ is a quantifier-free non-linear FOL formula on the variables $\overline{P}$ and $\overline{\phi}$. Then, we translate the constraint to a SAT formula by modeling integer variables as bit-vectors and integer operations such as addition, multiplication, and comparison as boolean operations. Finally, we use a state-of-the-art SAT solver to find a substitution $\rho$ for $\overline{\phi}$ that satisfies $\phi$.

We describe the approach in more detail. Given a universally quantified constraint $\forall \theta \exists \phi$, we convert $\theta$ to an equivalent formula of the form $\bigwedge_{i=1}^{m} \theta_i$, where each $\theta_i$ is of the form $-(\bigwedge_{j=1}^{n} c_{i,j} \geq 0)$ such that each $c_{i,j} \geq 0$ is a linear inequality on $\overline{\phi}$ with polynomials on $\overline{\theta_i}$ as coefficients. We translate each $\theta_i$ by applying Farkas’ lemma.

**Theorem 4.4** (Farkas’ lemma). Consider the following system of inequalities over real-valued variables $x_1, \ldots, x_n$.

\[
\begin{bmatrix}
c_{1,0} + c_{1,1}x_1 + \ldots + c_{1,n}x_n \geq 0 \\
\vdots \\
c_{m,0} + c_{m,1}x_1 + \ldots + c_{m,n}x_n \geq 0
\end{bmatrix}
\]

The system is unsatisfiable iff there exist non-negative reals $r_0, r_1, \ldots, r_m$ such that

- $r_0 + r_1c_{1,0} + \ldots + r_mc_{m,0} = -1$, and
- $r_1c_{1,1} + \ldots + r_mc_{m,1} = 0$ for each $j = 1, \ldots, n$.

Farkas’ lemma is incomplete for integers because the “only if” direction does not always hold. (Trivially, the “if” direction holds even for integers.) The incompleteness, however, has not affected the experiments in Section 5. (See Section 4.3 for the list of limitations with our approach.)

**Example 4.5.** Recall the constraints generated from the SHP template $d_T$ in Example 4.3. We translate the constraints to the equivalent form below.

\[
\forall x, y, z.
\]
\[
p_{14}x - p_{14}z \geq 0 \\
-p_{14}x + p_{14}y \geq 0 \\
-p_{14}p_{14} + p_{14}p_{14}x + (p_{14}p_{14} + p_{14})y - p_{14}z \geq 0 \\
-1 + x - y \geq 0
\]

Applying Farkas’ lemma, the constraint is further translated to the following form.

\[
\exists r_0, r_1, r_2, r_3, r_4 \geq 0.
\]
\[
r_0 + (p_{14}p_{14} + p_{14})r_3 - r_2 = -1 \\
p_{14}r_1 - r_3 + p_{14}p_{14}r_4 - r_4 = 0 \\
(p_{14}p_{14} + p_{14})(r_3 - r_4) - r_0 = 0 \\
p_{14}p_{14}p_{14} + p_{14}r_3 - r_4 = 0
\]

We reduce the constraints to SAT by modeling integers as bit-vectors, and apply SAT solving to find a satisfying assignment. A possible solution is

\[
\begin{align*}
p_0 &= 0, & p_1 &= 1, & p_2 &= 0, & p_3 &= 0, & p_4 &= 1 \\
r_0 &= 1, & r_1 &= 0, & r_2 &= 0, & r_3 &= 0, & r_4 &= 1, & r_5 &= 1
\end{align*}
\]

And from the solution, we obtain the substitution

\[
\rho = \{p_0 \mapsto 0, p_1 \mapsto 1, p_2 \mapsto 0, p_3 \mapsto 0, p_4 \mapsto 1\}
\]

Note that applying the inferred parameter substitution $\rho$ to $d_T$ results in the program equivalent to $d_S$ from Example 4.2. Finally, we update the parameter substitution $P$ via $P_{\rho}$ to translate the input program, which results in the program $d_{p'}$ from Example 4.1, and is typable by the underlying type system. While the example only require a variable as the instantiation expression, this is not the case in general (even when the instantiations are restricted to linear arithmetic expressions). Section 5 shows examples needing non-variable instantiation expressions.

We use the iterative approach of Gulwani et al. [9] to model integers as bit-vectors: we start with a low number of bits, and gradually increase the number of bits until a solution is found. Because of the incompleteness limitation, it is possible that a solution is not found even though the process is given a feasible SHP. (cf. Section 4.3.) Section 5 shows heuristics for coping with the issue.

**New candidate types with new parameter substitutions** It is often desirable to infer new candidate types along with the new parameter substitutions. For instance, in Example 4.5, rather than returning to Step 1 of Figure 10 after the updated parameter substitution $P_{\rho}$ is inferred, it is better to infer candidate types for the translated counterexample SHP $d_T$ so that the counterexample is immediately eliminated from the future CEGAR iterations.

**Inferring parameter substitution for prior counterexamples** For simplicity, we have shown how to infer a parameter substitution for one counterexample SHP. In practice, it is better to record the counterexamples that have been encountered (or, the constraints
generated from them), so that the inferred parameter substitution is able to also refute the past counterexamples. This can be done by a minor modification to the algorithm because constraints from multiple counterexamples simply add up as conjunctions. We take this approach in the implementation described in Section 5.

### 4.3 Limitations

Our type inference algorithm is incomplete as it is not able to decide the typability of all programs. This is expected because the underlying problem (i.e., safety verification) is undecidable. Here, we list the source of incompleteness.

1. Like the predicate abstraction in first-order software model checking [12], we restrict the refinement predicates to an efficiently decidable first-order theory, such as the quantifier-free theory of linear arithmetic. Consequently, there are programs, even ones that are first-order, that we cannot verify. Also, predicate discovery uses various heuristics such as interpolation, and is in general incomplete even for the theory subset (but, there are approaches to make this part complete [14]).

2. As remarked in Section 4.1, we restrict instantiation expressions to linear arithmetic expressions. While this has been sufficient for many cases, as we show in Section 5, there are programs that we cannot verify because of the restriction.\(^{12}\) (But, as we show there, simple heuristics can often be used to handle such cases.)

3. The parameter substitution inference is invoked only when the counterexample is unsatisfiable with the current substitutions, because otherwise it may simply infer the same substitutions. For some programs, the parameter substitution inference becomes forever blocked because the program generates infinitely many counterexamples each of which can be refuted by just updating the candidate types.\(^{13}\) (See Section 5 for an example.)

4. As remarked in Section 4.2.2, Farkas’ lemma is incomplete for integers.

### 5. Implementation and Experiments

We have implemented a prototype of the type inference algorithm as an extension to the higher-order model checker MoChI [18]. It takes (direct-style and CPS) OCaml programs. We use CVVC [2] for the SAT-based non-linear constraint solving described in Section 4.2.2. A web interface of the implementation and the benchmark programs from the experiments are available online [28].

We ran experiments on small but representative higher-order programs. Table 1 summarizes the results. The column S is the size measured in the number of words in the program. O is the largest order of the functions in the program. (Base-type values are order-0, and order-n functions only take arguments of order less than n.) The column \(|\overline{p}|\) is the total number of the variables \(\overline{p}\) that occur in the range of the parameter substitution template \(P_T\) (cf. Section 4.2). \#E is the number of extra parameters added before each function-type parameter. \(#E > 1\) is used in the heuristics described below.) \#I is the number of CEGAR iterations. The column T is the running time in seconds, and P/T is the fraction of the running time spent on the parameter substitution inference. NV marks the programs that required non-variable instantiation expressions. The experiments were conducted on a machine with Intel Xeon E5620 2.4GHz CPU and 4GB RAM.

Note that the programs are higher-order (i.e., O > 1). The programs are safe, but none of them can be verified by the previous refinement type systems proposed for automated higher-order program verification [13, 18, 23, 26, 27]. We describe each program.

- \(d_2\) and \(d_3\) are from Example 3.4.
- \(fhnn\) compares the return value of two function arguments:
  ```ocaml
  let f x y = assert (x () = y ()) in
  let h x () = x in let main n = f (h n) (h n)
  ```
- repeat-add makes the closure add \(n\) for some \(n\) ≥ 0, applies it to 0 for \(k\)-times for some \(k > 0\), and asserts that the result is not less than \(n\):
  ```ocaml
  let add x1 x2 = x1 + x2 in
  let rec repeat f k x = if k <= 0 then x
     else f (repeat f (k - 1) x) in
  let main n k = if n >= 0 && k > 0 then
      assert (repeat (add n) k 0 >= n) in
  ```
- app-leq invokes the following definition:
  ```ocaml
  app-leq (resp. app-lin-ord2) is the program below with e ≡ (a*x=b) (resp. e ≡ (4*a*x+2*b)).
  ```
- app-succ0 is app-succ with \(n\) in the body of main fixed to 0.
- l-xxx, we model and check array manipulating programs.

### Table 1. Experiment results.

<table>
<thead>
<tr>
<th>program</th>
<th>S</th>
<th>O</th>
<th></th>
<th>p</th>
<th></th>
<th>#E</th>
<th>#I</th>
<th>T</th>
<th>P/T</th>
<th>NV</th>
</tr>
</thead>
<tbody>
<tr>
<td>(d_2)</td>
<td>40</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0.43s</td>
<td>40%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(d_3)</td>
<td>56</td>
<td>3</td>
<td>10</td>
<td>1</td>
<td>3</td>
<td>1.76s</td>
<td>1.5%</td>
<td>26</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>(fhnn)</td>
<td>26</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.15s</td>
<td>12%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>repeat-add</td>
<td>51</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>3</td>
<td>1.05s</td>
<td>7%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>app-leq</td>
<td>29</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.26s</td>
<td>41%</td>
<td>√</td>
<td></td>
<td></td>
</tr>
<tr>
<td>app-lin-ord2</td>
<td>37</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.22s</td>
<td>38%</td>
<td>√</td>
<td></td>
<td></td>
</tr>
<tr>
<td>app-lin-ord3</td>
<td>43</td>
<td>5</td>
<td>6</td>
<td>1</td>
<td>1</td>
<td>0.96s</td>
<td>84%</td>
<td>√</td>
<td></td>
<td></td>
</tr>
<tr>
<td>app-succ</td>
<td>50</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>0.36s</td>
<td>37%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>app-succ0</td>
<td>50</td>
<td>2</td>
<td>5</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>a-test-upd</td>
<td>80</td>
<td>2</td>
<td>9</td>
<td>2</td>
<td>7</td>
<td>2.06s</td>
<td>32%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a-checksum</td>
<td>76</td>
<td>2</td>
<td>6</td>
<td>2</td>
<td>8</td>
<td>3.38s</td>
<td>29%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a-max</td>
<td>78</td>
<td>2</td>
<td>8</td>
<td>1</td>
<td>4</td>
<td>7.32s</td>
<td>26%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>l-forall-leq</td>
<td>62</td>
<td>2</td>
<td>8</td>
<td>1</td>
<td>2</td>
<td>2.68s</td>
<td>51%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>l-len-append</td>
<td>126</td>
<td>2</td>
<td>22</td>
<td>1</td>
<td>1</td>
<td>0.73s</td>
<td>35%</td>
<td>√</td>
<td></td>
<td></td>
</tr>
<tr>
<td>l-isort</td>
<td>170</td>
<td>2</td>
<td>39</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>
The array benchmarks extend the ones from the previous work [18] that demonstrate the application of higher-order program verification to the verification of data structure properties.

- a-test-upd tests array updates. It creates an array of size \( n \) and updates the \( i \)-th element to \( x \) for some \( i \in \{0, \ldots, n - 1\} \). The program then reads the element and checks that the result is equal to \( x \) as expected:

\[
\begin{align*}
\text{let test } (n, ar) & \ i x = \text{assert } (ar i = x) \text{ in} \\
\text{let main } n & \ i x = \\
& \text{if } 0 \leq i \land i < n \text{ then} \\
& \text{test } (\text{upd } (\text{make_array } n) i x) \text{ i x}
\end{align*}
\]

Note that the program also checks for bounds violation.

- a-checksum creates an array of size 2, updates the first and the second elements to \( a \) and \( b \) respectively, and asserts that the sum of the two array elements equals \( a + b \):

\[
\begin{align*}
\text{let checksum } (n, ar) & \ x = \\
& \text{assert } ((ar 0) + (ar 1) = x) \text{ in} \\
\text{let main } & \ a b = \\
& \text{checksum} \\
& (\text{upd } (\text{make_array } 2) 0 a) \ 1 b) \ (a + b)
\end{align*}
\]

- a-max creates an array of size \( n \) whose elements are from the set \( \{0, \ldots, x\} \) where \( x > 0 \), computes the maximum element \( m \) of the array, and asserts that the sum of \( m \leq x \).

- The programs 1-xxx model and check list operations. We model a list by a pair of its length and a function from a non-negative integer \( i \) to the \( i \)-th element. We model the core list operations as follows:

\[
\begin{align*}
\text{let nil} & = (0, \text{fun } i \rightarrow \text{assert false}) \\
\text{let cons } a & \ (\text{len}, 1) = \\
& (\text{len} + 1, \text{fun } i \rightarrow \text{if } i=0 \text{ then } \text{a else } 1 (i-1)) \\
\text{let hd } (\text{len}, 1) & = 1 0 \\
\text{let tl } (\text{len}, 1) & = (\text{len} - 1, \text{fun } i \rightarrow 1 (i + 1)) \\
\text{let is nil } (\text{len}, 1) & = \text{len=0}
\end{align*}
\]

- 1-forall-leq creates a list of the length \( n \) whose \( i \)-th element is \( n - i \), and asserts that \( x \leq n \) holds for all element \( x \) of the list.

- 1-len-append appends two lists of the length \( len1 \) and \( len2 \), computes the length of the resulting list, and asserts that it is not greater than \( len1 + len2 \).

- 1-isort creates a list, sorts the list via an insertion sort, and then checks that the result is actually sorted.

The implementation was able to successfully verify the programs by automatically inferring appropriate extra parameter substitutions, except for app-suucc0 and 1-isort. Before we discuss the reason for failure on these two, we describe the heuristic that was needed for verifying a-test-upd and a-checksum.

Recall that we limit instantiation expressions to linear arithmetic expressions. With this restriction, it can be shown that a-test-upd and a-checksum (and also 1-isort) are actually untypable when we are allowed to add only one extra parameter before a function-type parameter. However, a-test-upd and a-checksum can be typed if we are allowed to add two extra parameters before a function-type parameter, even with the linear-instantiation-expression restriction. (Note that \( \#E \) is 2 for these programs in Table 1.)

For instance, in a-checksum, to show the safety of the call to checksum, we need to infer the property that the second argument equals the sum of the first and the second elements of the function-encoded array passed as the first argument. Because the two elements of the array are independent (i.e., the arguments \( a \) and \( b \) of main), it is not possible to express such a fact via a refinement type with just one (linearly instantiatable) extra parameter. Nonetheless, the property can be expressed even under the linear-instantiation-expression restriction by using two extra parameters. For example, the following type is sufficient:

\[
\forall u: \text{int}. \forall v: \text{int}. \\
(\text{int} \times (x: \text{int} \rightarrow \{u | \theta\})) \rightarrow \{u | u = a + b \} \rightarrow \star
\]

where \( \theta \equiv (x = 0 \Rightarrow u = a) \land (x = 1 \Rightarrow u = b) \), and \( x \) is the pair type constructor. We use a simple heuristic to progressively increase the number of extra parameters that are added before each function-type parameter: we increment the number by one when the parameter substitution inference fails to find a solution to the constraints, with some pre-defined threshold on the number of bits used in the bit-vector modeling (cf. Section 4.2.2). With the heuristic, the implementation was able to verify a-test-upd and a-checksum. Note that the increased extra parameter heuristic is sound because the soundness theorem (Theorem 3.6) holds for any pattern of extra parameter additions.

The program 1-isort remains untypable even with the heuristic. In fact, it can be shown that no finite number of extra parameter additions can type the program when the instantiation expressions are restricted to be linear.

The implementation fails to verify app-suucc0 for a different reason. To verify the program, we need to infer the following type for app with the extra parameter \( i \):

\[
\forall i: \text{int}. f: \{u | u = i \} \rightarrow \star \rightarrow x: \{u | u = i \} \rightarrow \star
\]

However, each counterexample SHP generated from app-suucc0 is actually typable by assigning (the copies of) app the types of the following form whose refinement predicates do not mention the extra parameter:

\[
\forall i: \text{int}. f: \{u | u = c \} \rightarrow \star \rightarrow x: \{u | u = c \} \rightarrow \star
\]

where \( c \) is some integer constant. Therefore, the type inference system generates infinitely many counterexample SHPs that are refuted by the types of the above form each time, and the parameter substitution inference is never invoked to infer the necessary type that uses the extra parameter. We leave the issue for future work.

6. Related Work

6.1 Refinement Type Systems

The relatively complete refinement type system of this paper can be seen as DML [30] but with quantifiers (i.e., introduction of \( \Pi \) types) restricted to one before each function-type parameter. Hence, our result shows that DML is actually relatively complete (even when the quantification pattern is restricted). In relation to DML, this paper’s contribution is in actually proving relative completeness, and in proposing an automated type inference method.

The previous refinement type systems proposed for automated higher-order program verification [13, 18, 23, 26, 27] lack relative completeness, except for the case when the base-type data domain is finite [16, 17]. We have shown how the type inference systems underlying the verifiers may be extended to attain relative completeness.

6.2 Hoare Logic for Higher-order Languages

There is a long line of research on relatively complete Hoare logic like proof systems for higher-order functional (or procedural) languages. One way to achieve relative completeness is to have a higher-order logic as the interpretation logic (i.e., the logic for the
data part) so that the formulas can directly refer to higher-order functions as data (see, e.g., [5, 11, 21, 22]). However, such an approach is difficult to automate as the client verifier must rely on a higher-order logic theorem prover. The dilemma is similar to that arising from explicitly encoding functions as base-type data in the program logic.

The approach of using base-type encoding of functions to achieve relative completeness was first proposed by German et al. [6, 7]. (A similar approach is taken in [8].) Crucial to their approach is the idea of avoiding the explicit encoding in the program logic via quantification, allowing the client verifier to choose appropriate quantifier instantiations. To our knowledge, no actual verifier was built based on their program logic.

The encoding approach is also used in a recent work on relatively complete dependent (revisited) system [19] where a DML-like system is extended with linear intersection types, and is shown sound and relatively complete (for terminating programs). Their type system is designed for a type-based complexity analysis, and the paper does not discuss automated verification.

6.3 Quantified First-order Logic Constraint Solving

Our type inference method utilizes an algorithm for deciding the validity of a FOIL formula of the form \( \exists \vec{x} \forall \vec{y} \theta \) where \( \theta \) is a quantifier-free arithmetic formula over the variables \( \vec{x} \) and \( \vec{y} \), and it valid, finding appropriate instantiations for \( \vec{x} \) so that the formula can be reduced to \( \forall \vec{y} \theta[\vec{y}/\vec{x}] \).

The problem is studied in the context of constraint-based (or, template-based) program verification [4, 9, 24] where the existentially quantified variables are used to represent the unknowns in the invariant template. The technique has found wide applicability, including hybrid system verification [20] and program synthesis [25]. Our work adds higher-order program verification to the list.

7. Conclusion

We have presented an automated approach to a relatively complete verification of higher-order functional programs. Our work extends the recent research on refinement type inference for automated program verification. We have extended the underlying refinement type system by adopting the classical result on relatively complete Hoare logic like proof systems for higher-order procedural languages, which shows that certain extra base-type parameter additions are sufficient for relative completeness. Then, we have extended the type inference system by utilizing the techniques from the recent work on quantified FOIL constraint solving over template expressions to infer appropriate instantiations for the extra base-type parameters.

Acknowledgments We thank the anonymous reviewers for useful comments. This work is partially supported by Kakenhi 23220001 and 23700026.

References


A. Simple Type System

The syntax of simple types is defined by the following grammar: (Also shown in Section 3.)

\[ s ::= ∗ | \text{int} | s → s' \]
The proof is by the usual syntactic approach [29] and is similar to those of the previous work on refinement type systems [23, 27].
We extend the typing rules to type the stack augmented intermediate expressions. (We equate $st \varepsilon e$ with $e$.)

$$\frac{\Gamma \vdash \forall \gamma \tau \sigma}{\Gamma \vdash st (a = \ell, b \in \ell) \vdash e : \star} \quad St$$

We also add the following so that safe is always typable.

$$\Gamma \vdash \forall \gamma \tau \sigma \vdash st (a = \ell, b \in \ell) \vdash e : \star$$

The following is immediate from the property of sound constant types (Appendix B).

**Lemma D.2.** Suppose $o \in \text{pureExp}$ and $\Gamma \vdash o : \{u \mid \theta\}$. Then, $[\Gamma] = \exists u. \theta$.

The following lemma states that to we can derive $\Gamma \vdash e : *$ by binding a fresh variable bound to the type of a side-effect-free integer type-expression.

**Lemma D.3.** Suppose $\Gamma \vdash o : \tau$ and $\Gamma, x : \tau \vdash e : *$ such that $o \in \text{pureExp}$ and $x \notin \text{fv}(e)$. Then, $\Gamma \vdash e : *$. (Note that we assume $x \notin \text{dom}(\Gamma)$.)

**Proof:** By inspection of the $\Gamma \vdash \gamma$ typing rules, the binding of the non-occurring variable $x$ can affect the type derivation of $e$ in only two ways: (1) by using $x$ as a sub-expression of the instantiating expression at $\text{Inst}$, or (2) by having $x : \tau$ in the left hand side of $\Rightarrow$ within the FOL formula denoting the assumptions about the base-type variables in the environment at the subtyping judgement $\text{Sub}$. In the case (1), any such use of $x$ can be substituted by $o$ to attain the same effect. In the case (2), suppose $\gamma = \{u \mid \theta\}$. Then, in $\text{Sub}$, $\theta$ may only occur as $\phi_i \land \theta[x/u] \Rightarrow \phi_k$ such that $x$ does not occur free in $\phi_i, \phi_k, \theta$, or $\theta$. Therefore, the formula is logically equivalent to $\phi_i \land \exists u. \theta \Rightarrow \phi_k$. By inspection of the typing rules, $\phi_i$ must be such that $\phi_i = [[\Gamma]]$. Also because $\Gamma \vdash o : \{u \mid \theta\}$ and $o \in \text{pureExp}$, it follows that $[[\Gamma]] = \exists u. \theta$ by Lemma D.2. And, we can derive the instance of $\text{Sub}$ without the binding. □

We state a couple of substitution lemmas.

**Lemma D.4 (Substitution I).** Suppose $\Gamma \vdash y : \sigma$ and $\Gamma, x : \sigma \vdash e : *$. Then, $\Gamma \vdash e[y/x] : *$.

**Proof:** By induction on the derivation. □

**Lemma D.5 (Substitution II).**

* Suppose $\Gamma \vdash y : \sigma_1$, and $\Gamma, x : \sigma_1 \vdash \sigma_2 \leq \sigma_3$. Then, $\Gamma \vdash \sigma_2[y/x] \leq \sigma_3$.

* Suppose $\Gamma \vdash o : \sigma_1$, $o \in \text{pureExp}$, and $\Gamma, x : \sigma_1 \vdash \sigma_2 \leq \sigma_3$. Then, $\Gamma \vdash \sigma_2[a/x] \leq \sigma_3$.

**Proof:** By induction on the derivation. □

Given type environments $\Gamma_1$ and $\Gamma_2$, let $\Gamma_1, \Gamma_2$ denote the type environment formed by appending the two in sequence in that order. It suffices show the following preservation lemma to prove the theorem.

**Lemma D.6 (Preservation).** Suppose $\Delta \vdash e : *$ and $e \rightarrow_d e'$. Then, $\Delta \vdash e'$.

**Proof:** By case analysis on the reduction relation.

Suppose $e$ is of the form $st \ell$ in $let x = e_1 in e_2$ and $e_1$ is neither an assume nor assertion expression. Then, $\Delta \vdash e : *$ and $e \rightarrow_d e'$ imply that we must have $\Delta, \Gamma \vdash v : \sigma$ and $\Delta, \Gamma, x : \sigma \vdash e_2 : *$ where $\ell(e_1) \leq \sigma$ and $\Gamma$ is the bindings from the stack $\ell$ formed by $\text{St}$. Therefore, by Lemma D.4 and $\text{St}$, we have $\Delta \vdash \forall \gamma \tau \sigma \vdash st (b = \ell) \in e_2[b/x] : *$.

Suppose $e$ is of the form $st \ell$ in $let x = \text{assert} a$ in $e_1$. Then, $\Delta \vdash e : *$ and $e \rightarrow_d e'$ imply that we must have $\Delta, \Gamma \vdash a : \{u \mid u = 1\}$ and $\Delta, \Gamma, x : \{u \mid u = 1\} \vdash e_1 : *$ where $\Gamma$ is the bindings from the stack $\ell$ formed by $\text{St}$. Therefore, it must be the case that $(a = 1) \in \ell$, and we have $e \rightarrow_d st \ell, (b = 1) \in e_1[b/x]$. Hence, by Lemma D.4 and $\text{St}$, we have $\Delta \vdash \forall \gamma \tau \sigma \vdash st \ell, (b = 1) \in e_1[b/x] : *$.

Suppose $e$ is of the form $let x = a \in e_1$. Then, $\Delta \vdash e : *$ and $e \rightarrow_d e'$ imply that we must have $\Delta, \Gamma, x : \{u \mid a = 1 \land u = 1\} \vdash e_1, e_2 : *$ where $\Gamma$ is the bindings from the stack $\ell$ formed by $\text{St}$. Therefore, if $(a = 1) \in \ell$, then $e \rightarrow_d at \ell, (b = 1) \in e_1[b/x]$. Therefore, by Lemma D.4 and $\text{St}$, we have $\Delta \vdash at \ell, (b = 1) \in e_1[b/x] : *$. Otherwise, $(a = i) \in \ell$ such that $i \neq 1$. Then, $e \rightarrow_d safe$ and we have $\Delta \vdash safe : *$.

Suppose $e$ is of the form $if \ then \ else \ e_1 \ else \ e_2$ and $st \ell$ in $if \ then \ else \ e_1 \ else \ e_2$ in $\ell$. Then, $\Delta \vdash e : *$ imply that we must have $\Delta, \Gamma \vdash e : *$ and sequences of pure integer-type expressions $\sigma_0, \ldots, \sigma_n$ such that

$$\forall \gamma \rightarrow \gamma'$$

for and for $0 \leq i \leq n$ and $1 \leq j \leq n_i$, $\Delta, \Gamma \vdash a_i : \sigma_i \pi_i$, $\Delta, \Gamma \vdash \alpha_{ij} : \sigma_{ij} \pi_{ij}$, $\pi_{ij} = e \Rightarrow \pi_{ij} = \pi_{ij}[\sigma_{nj}/y_{nj}]$, $\pi_{ij} = \pi_{ij-1}[\alpha_{ij}/y_{ij-1}]$ $\forall \gamma \rightarrow \gamma'$

where $\Gamma$ is the bindings from the stack $\ell$ formed by $\text{St}$, and $[\gamma]$ $= n_i$ for $0 \leq i \leq n$. Let $F x^\gamma = \text{eval}_d \in d$. Then, we must have $\gamma' \rightarrow *$ such that

$$\forall \gamma \rightarrow \gamma'$$

From the rule of subtyping, we have $\Delta, \Gamma, \{\gamma'_i\} \vdash \tau_{ij} \leq \tau'_j$ for each $i, j$ where $\gamma'_i$ is the prefix of $\gamma'$ up the binding just before $y_{ij} : \tau_{ij}$, and $\Delta, \Gamma, \{\gamma'_i\} \vdash \sigma_i \leq \sigma'_j$ for each $i$ where $\gamma'_i$ is the prefix of $\gamma'$ up the binding just before $x_i : \sigma_i$. Therefore, by $\text{Sub}$ and Lemma D.5, we have $\Delta, \Gamma \vdash o_{ij} : \tau_{ij}$ and $\Delta, \Gamma \vdash a_i : \sigma_i$ for each $i, j$. Then, by Lemma D.4, Lemma D.3, and $\text{St}$, we have $\Delta \vdash st \ell \in \text{eval}_d (d^\gamma) : *$.

The case $e$ is of the form $at \ell$ in $b a$ is analogous. □

Finally, Theorem 3.6 follows from Lemma D.6 because, given $\Delta$ such that $\Delta \vdash d$, we have $\Delta \vdash \text{eval} : *$ for any initial state $\text{eval} = a = i \in \text{eval}_d (d^\gamma) : *$ and $\Delta \not\vdash fail : *$.

**E. Proof of Theorem 3.7**

**Theorem 3.7:** If $d$ is safe, then there exists $\Delta$ such that $\Delta \vdash d$ and $\text{ushtype}(\Delta (F)) = \text{style}(F)$ for each $F d = e \in e$.

Following [6–8], we encode functions as integers via Gödel numbering. We define concrete value (simply value, henceforth), $v$, to be a simply-typed expression defined by the following grammar.

$$v ::= F \mid i \mid v_1 \cdot v_2$$

Here, $i$ is an integer constant. We define applicable term (simply term, henceforth), $t$, to be a simply-typed expression defined by the following grammar.

$$t ::= x \mid F \mid t_1 \cdot t_2$$

Note that this is different from the notion of “value” used in the dynamic semantics (cf. Section 3 and Appendix D).
where $c \overline{t}$ is a total application of a non-assume/assert constant. Note that any value is a term. Given a term $t$ with free variables $\overline{x}$, we define the evaluation relation $t(\overline{v}/\overline{x}) \Downarrow v$ in the obvious way: $v$ is $t(\overline{v}/\overline{x})$ with constant operator applications in the subexpression reduced to integer constants.

For a value $v$, we write $[v]$ to denote the G"odel number that encodes $v$. We let $[i]$ be for an integer constant $i$. (Collisions between function-type and integer-type values can always be avoided in our setting because the program is assumed to be simply typed.) For each variable $x$, we introduce a distinct base-type (i.e., integer-type) variable $\overline{x}$. We write $\overline{x}$ for the sequence $x_1, \ldots, x_n$ where $\overline{x} = x_1, \ldots, x_n$. We write $wpre(\overline{t}, \psi)$ for the set of variables $\overline{z}$, where $wpre(\overline{t}) = \overline{z}$.

For a term $t$, we define $\langle t \rangle$ to be the integer-type term of order $t$ that encodes $t$. More formally, $\langle t \rangle$ is the integer term that $\langle t \rangle = x_1, \ldots, x_n$.

\begin{align*}
\forall v, \overline{v}. (\langle t(\overline{v}/\overline{x}) \Downarrow v \rangle \equiv (\langle t \rangle(\overline{v}/\overline{x}) = [v]))
\end{align*}

where $\text{maxargs}$ be the maximum number of function arguments of $d$, that is, $\text{maxargs} = \max\{|\overline{t}| : F \overline{t} = d \in d\}$. Without loss of generality, we assume that each function of $d$ has the same function argument names, that is, there exists $\overline{\beta}$ such that $|\overline{\beta}| = \text{maxargs}$ and for each $F \overline{t} = d \in d$, $\overline{t}$ is a prefix of $\overline{\beta}$. (Note that the variables of the same name in different functions need not have the same simple type.) In the rest of the proof, we fix the variables $\overline{\beta}$ to be the sequence of function arguments of the program $d$. Let $\overline{x}$ be $\overline{\beta}$. We also introduce a special base-type variable $\overline{x}_{\text{hd}}$ that is distinct from $\overline{\alpha}$.

Let $\psi$ be a formula such that $\forall \overline{x}. (wpre(\overline{t}, \psi) \Rightarrow (\langle t \rangle(\overline{v}/\overline{x}) = [v]))$. We intend to use $\psi(\overline{x}_{\text{hd}}, \overline{\alpha})$ to express the condition for which it is safe to call the function encoded by $\overline{x}_{\text{hd}}$ via the arguments encoded by $\overline{\alpha}$. We say that an expression is a CPS expression if it is given the simple type $\star$ (e.g., a function body). Figure 13 shows the rules for deriving $wpre(\overline{t}, \psi)$ for a CPS expression $\overline{t}$. Intuitively, $wpre(\overline{t}, \psi)$ is the condition safe to evaluate $\overline{t}$ assuming that $\psi$ is the condition safe to call the functions. Note that $\forall \overline{x}. (wpre(\overline{t}, \psi)) \subseteq \langle \overline{\alpha} \rangle$.

For a value $v$, we define the weakest precondition of the program by taking the fixed point of the following predicate

\begin{align*}
\text{step}(\psi) &\equiv \forall t, \overline{v}. (\overline{t} = x_{\text{hd}} \Rightarrow wpre(e_F, \psi)(\overline{v}/\overline{x}_{\text{hd}}) = \overline{v}/\overline{y})
\end{align*}

where $e_F = e_F \in d$, $\overline{x} = \overline{x}_{\text{maxargs}-1}$, and $\overline{y} = \overline{x}_{\text{maxargs}+1}$. Note that $\forall n. (\text{step}(\psi) \subseteq (x_{\text{hd}}, \overline{\alpha}))$.

Next, we show (2) by induction on $n$. The case $n = 0$ is trivial. Suppose that (1) and (2) hold for $n - 1$. Let $F \overline{t} = e_F \in d$. Then, (1), $F \overline{t} = e_F \overline{\alpha} \overline{\beta}$ fail if $\forall \overline{\theta}. (\overline{\theta} \overline{\alpha} \overline{\beta})$ is valid where $\theta = wpre(e_F, \psi)$ and $\overline{\psi} = \text{step}^{n-1}(\overline{\psi})$. Therefore, by the definition of $\theta$, $F \overline{t} = e_F \overline{\alpha} \overline{\beta}$ fail if $\forall \overline{\theta}. (\overline{\theta} \overline{\alpha} \overline{\beta})$ is valid.

Next, we show (2) by induction on $n$. The case $n = 0$ is trivial. Suppose that (1) and (2) hold for $n - 1$. Let $F \overline{t} = e_F \in d$. Then, (1), $F \overline{t} = e_F \overline{\alpha} \overline{\beta}$ fail if $\forall \overline{\theta}. (\overline{\theta} \overline{\alpha} \overline{\beta})$ is valid where $\theta = wpre(e_F, \psi)$ and $\overline{\psi} = \text{step}^{n-1}(\overline{\psi})$. Therefore, by the definition of $\theta$, $F \overline{t} = e_F \overline{\alpha} \overline{\beta}$ fail if $\forall \overline{\theta}. (\overline{\theta} \overline{\alpha} \overline{\beta})$ is valid.

The least fixed point of $\text{step}$ is the following formula. (Least in the order of logical implication.)

\begin{align*}
\Psi_{\text{wp}} &\equiv \forall t, \text{step}^n(\overline{\psi})
\end{align*}

Note that $\Psi_{\text{wp}}$ can be expressed as a FOL formula by using G"odel numbering to encode the sequence of step applications. This requires encoding the FOL formulas in the sequence.

The following is immediate from the definition of $\Psi_{\text{wp}}$.

\begin{align*}
\Psi_{\text{wp}}(t) &\equiv \forall t, \text{step}^n(\overline{\psi})
\end{align*}

\begin{align*}
\Psi_{\text{wp}}(t) &\equiv \forall t, \text{step}^n(\overline{\psi})
\end{align*}
From Lemma E.1, it follows that \( \Psi_{\text{wp}} \) expresses the weakest precondition of the program. That is,

**Corollary E.3.** Let \( F \vdash \Psi \vdash \theta \) be a simply-typed total application. Then, \( F \vdash \Psi \vdash \theta \) if and only if \( \Psi_{\text{wp}} \vdash \theta \). \( \square \)

In particular, the weakest precondition of main of a safe program must be a tautology, that is, it is safe to execute for all of its inputs.

**Corollary E.4.** \( \Psi_{\text{wp}}([\text{main}] / x_{\text{hd}}) = \top \) if and only if \( d \) is safe.

In the rest of the proof, we assume that \( d \) is safe. We use \( \Psi_{\text{wp}} \) to build refinement types for \( d \)'s functions. Specifically, for each function-type term, we define its weakest precondition refinement type by the rules shown in Figure 14.

Here, \( \text{args}(\overrightarrow{\sigma} \rightarrow \star) \) is the prefix of \( \overrightarrow{\sigma} \) such that \( |\text{args}(\overrightarrow{\sigma} \rightarrow \star)| = \lceil \overrightarrow{\sigma} \rceil \). We define \( \text{args}^{-1}(s) \) to be the set of variables \( \{ \overrightarrow{\sigma} \} \setminus \{ \text{args}(s) \} \). Then, given a function-type term \( t \), we define

\[
\text{wp}(t) \equiv \text{wp}(\{t\}, \text{args}(\text{sty}(t)), \text{args}^{-1}(\text{sty}(t)), \text{sty}(t))
\]

The following is immediate from the rules of constructing \( \text{wp}(t) \).

**Lemma E.5.** Let \( \tau = \text{wp}(t) \). Then, we have \( \text{fv}(\tau) \subseteq \text{fv}(d) \) and \( \text{ushape}(\tau) = \text{sty}(t) \).

**Proof:** By induction on \( \text{sty}(t) \). \( \square \)

**Corollary E.6.** Let \( \tau = \text{wp}(F) \). Then, \( \tau \) is a well-formed type of \( F \) such that \( \text{ushape}(\tau) = \text{sty}(F) \).

Therefore, to prove Theorem 3.7, it suffices to show \( \Delta \vdash \rightarrow d \) for \( \Delta \) such that \( \Delta(F) = \text{wp}(F) \) for each \( F \).

**Example E.7.** For example, the program \( d_2 \) from Example 3.4 has the following weakest precondition. (Here, we set \( \beta = x, y, \).

\[
\begin{align*}
\Psi_{\text{wp}}([\text{main}] / x_{\text{hd}}) & = \top \\
\forall_i. \Psi_{\text{wp}}([\text{main}] / x_{\text{hd}}) & = \top \\
\forall_i. \Psi_{\text{wp}}([\text{check}] / x_{\text{hd}}) & = \frac{\leq}{y} \\
\forall_i. \Psi_{\text{wp}}([\text{check \ i}] / x_{\text{hd}}) & = \frac{i \leq x}{y} \\
\forall_i. \Psi_{\text{wp}}([\text{check \ i}] / y_{\text{hd}}) & = \frac{j}{y} \\
\forall_i. \text{wp}([\text{app \ i}] / x_{\text{hd}}) & = \frac{\leq}{y} \\
\forall_i. \text{wp}([\text{app \ i}] / y_{\text{hd}}) & = \frac{\leq}{y} \\
\forall_i. \text{wp}([\text{app \ i}] / y_{\text{hd}}) & = \frac{\leq}{y}
\end{align*}
\]

And, the corresponding weakest precondition types are shown below.

\[
\begin{align*}
\text{wp}(\text{main}) & = x : \text{int} \rightarrow \star \\
\text{wp}(\text{check}) & = x : \text{int} \rightarrow y : \{ u \mid x \leq u \} \rightarrow \star \\
\text{wp}(\text{app}) & = \forall x'. \text{int} \cdot x : \sigma \rightarrow y : \{ u \mid i \leq u \} \rightarrow \star
\end{align*}
\]

where \( \sigma = x : \{ u \mid i \leq u \} \rightarrow \star \).

We prove a few auxiliary lemmas.

**Lemma E.8.** Let \( t_1 t_2 \) be a simply-typed partial application. Let \( \tau_1 = \text{wp}(t_1) \).

1. If \( \text{sty}(t_2) = \text{int} \), then \( \tau_1 = x : \text{int} \rightarrow \sigma \) such that \( \sigma[[t_2]/x] = \text{wp}(t_2) \).
2. Otherwise, \( \text{sty}(t_2) \neq \text{int} \) and \( \text{wp}(t_1) = \forall x : \text{int}. y : \tau \rightarrow \sigma \) such that \( \tau[[t_2]/x] = \text{wp}(t_2) \).

**Proof:** We show (1). By the construction of \( \text{wp}(t) \) and the fact that \( t_1 t_2 \) is a partial application, it follows that \( \tau_1 \) is of the form \( x : \text{int} \rightarrow \sigma \) (i.e., because the \( \text{wp}(t) \) construction sets the refinement predicate of a non-last base-type argument to \( \top \)). Let \( \tau = \text{wp}(t_1 t_2) \).

Suppose the last argument of \( \tau \) and \( \sigma[[t_2]/x_1] \) is a base type.

From the construction, it is easy to see that \( \tau \) and \( \sigma[[t_2]/x_1] \) are alpha-equivalent, except for the refinement predicate in the last base-type argument. Let \( \sigma \) have \( n \) arguments so that \( \sigma = x_1 : \overrightarrow{x_i} \rightarrow \star \) with \( i = 2, \ldots, n + 1 \), and \( x_1 : \overrightarrow{x_{n+1}} \) is a prefix of \( \overrightarrow{x_i} \). Let \( \overrightarrow{x_i} = x_1, \ldots, x_{n+1} \). Let \( \overrightarrow{x_i} \) also be the arguments of \( \tau \) (i.e., via alpha-renaming). Let \( \theta_\tau \) and \( \theta_\sigma[[t_2]/x_1] \) be the refinement predicates in the last argument of \( \tau \) and \( \sigma[[t_2]/x_1] \), respectively (i.e., \( \theta_\tau \) is the refinement predicate in the last argument of \( \sigma \)). Then, we have

\[
\begin{align*}
\theta_\tau & = (\forall Y_1. \Psi_{\text{wp}})[[t_1 t_2]/x_{\text{hd}}] / [u / \alpha] \\
\theta_\sigma[[t_2]/x_1] & = (\forall Y_2. \Psi_{\text{wp}})[[t_1 t_2]/x_{\text{hd}}] / [u / \alpha_1][[t_2]/x_1]
\end{align*}
\]

where \( Y_1 = \text{args}^{-1}(\text{sty}(t_1 t_2)) \), and \( Y_2 = \text{args}^{-1}(\text{sty}(t_1)) \).

Therefore, by Lemma E.2, it follows that

\[
\theta_\tau[[t_2]/x_1] \Rightarrow \theta_\sigma[[t_2]/x_1]
\]

Therefore, \( \sigma[[t_2]/x] = \text{wp}(t_1 t_2) \) (via alpha-renaming).

Otherwise, the last argument of \( \tau \) and \( \sigma[[t_2]/x_1] \) is a function type (of the same simple type), and an analogous argument holds for the refinement predicate in the type of the universally bound variable just before the argument.

We show (2). By the construction of \( \text{wp}(t) \) and the fact that \( t_1 t_2 \) is a partial application, it follows that \( \tau_1 \) is of the form \( \forall x : \text{int}. y : \tau \rightarrow \sigma \) (i.e., because the \( \text{wp}(t) \) construction sets the refinement predicate of a universally bound variable just before a non-last function-type argument to \( \top \)). From the construction, it follows that \( \tau[[t_2]/x] = \text{wp}(t_2) \). And, \( \sigma[[t_2]/x] = \text{wp}(t_1 t_2) \) follows from the argument similar to (1). \( \square \)

**Lemma E.9.** Suppose \( \Gamma \vdash e : \tau \) and \( [\Gamma] \Rightarrow x = t \). Then, \( \Gamma \vdash e : \tau[[t]/x] \).
Proof: By Sub and the fact that if $\theta \Rightarrow \psi$ and $\theta \Rightarrow x = t$ then $\theta \Rightarrow \psi[t/x]$.

Note that for any term $t$, $[t] \in \text{pureExps}$. We are now ready to state the main lemma.

Lemma E.10. Let $e$ be a CPS-sub-expression. Let $\rho$ be the substitution such that $\rho(x) = \text{if } x \text{ is a base-type variable in } e \text{ and } \rho(\underline{e}) = \text{if } e \text{ is a function-type variable in } e$. Let $\Gamma$ be a type environment for $e$ such that $\Gamma \vdash \text{wp}(e, \Psi_{wp})\rho$, and for each function-type variable or function name $\kappa \in \text{dom}(\Gamma)$, there exists a term $t$ such that $\text{fv}(\underline{t}) \subseteq \text{dom}(\Gamma)$, and $\Gamma(\kappa) = \text{wp}(t)$. Then,

(A) For each integer-type sub-term $t$ of $e$, $\Gamma \vdash_{\tau} t : \{u \mid [t]_\rho = u\}$. (B) For each function-type sub-term $t$ of $e$, $\Gamma \vdash_{\tau} \text{wp}(t)$. (C) $\Gamma \vdash_{\tau} e : \star$.

Proof: (A) follows from the definition of precise constant types (cf. Appendix B).

We show (B) by induction on $t$. The base case (i.e., $t = x$ or $t = F$) is immediate from the assumption. So, let $t = t_1 t_2$. By induction hypothesis, we have $\Gamma \vdash_{\tau} t_1 : \text{wp}(t_1)$. If $t_2$ is an integer-type term, then the result follows from (A), Lemma E.8 (1), and Lemma E.9. Otherwise, $t_2$ is a function-type term, and the result follows from Lemma E.8 (2), applying $\text{Inst}$ with the instantiating expression $[t_2]$ to instantiate $t_1$'s type to $\text{wp}(t_2) \rightarrow \text{wp}(t_2 t_2)$.

We show (C) by induction on $e$. Suppose $e$ is let $x = t$ in $e_1$. If $e_1$ is a base type, then by (A), we have

$$\Gamma, x : \tau \vdash t : \tau$$

Therefore, by induction hypothesis, $\Gamma, x : \tau \vdash e : \star$ by induction hypothesis, and by $\text{Let}$, we have $\Gamma \vdash_{\tau} e : \star$. Otherwise, $e_1$ is a function type (i.e., a partial application), and we have $\Gamma, x : \sigma, \Gamma, \underline{\tau} \vdash \text{wp}(e_1, \Psi_{wp})\rho$ where $\sigma = \text{wp}(e_1)$ and $\tau = \{u \mid [t]_\rho = u\}$. Then, by induction hypothesis, we have $\Gamma, \tau \vdash \tau_1 e_1 : \star$. Therefore, by (B), Lemma D.3, and $\text{Let}$, we have $\Gamma \vdash_{\tau} e : \star$.

Suppose $e$ is let $x = \text{assert} y$ in $e_1$. We have

$$\Gamma, x : \{u \mid u = 1\} \vdash \text{wp}(e_1, \Psi_{wp})\rho$$

Therefore, by induction hypothesis, $\Gamma, x : \{u \mid u = 1\} \vdash e_1 : \star$. Also, $\Gamma, \Gamma \vdash \text{wp}(e_1, \Psi_{wp})\rho$ $\Rightarrow y = 1$. Therefore, by $\text{Let}$ and the type of $\text{assert}$ (i.e., $x : \{u \mid u = 1\} \Rightarrow \{u \mid u = 1\}$), we have $\Gamma \vdash_{\tau} e : \star$. The case $\text{let} x = \text{assume} y$ in $e_1$ is similar.

Suppose $e$ is if $*$ then $e_1$ else $e_2$. We have

$$\Gamma \vdash \text{wp}(e_1, \Psi_{wp})\rho$$

Therefore, by induction hypothesis, we have $\Gamma \vdash_{\tau} e_1 : \star$ and $\Gamma \vdash_{\tau} e_2 : \star$. Then, by $\text{If}$, $\Gamma \vdash_{\tau} e : \star$.

Suppose $e$ is a total application $t \rightarrow t_1 \ldots t_n$. By (B), we have $\Gamma \vdash_{\tau} t' : \text{wp}(t')$. Suppose $\text{sty}(t_n) = \text{int}$. Then, by (A), we have that $\Gamma \vdash_{\tau} t_n : \{u \mid [\alpha]_\rho = u\}$. By the construction of $\text{wp}(t')$, we have

$$\theta = (\forall Y, \Psi_{wp})[[t']/x_{\alpha}]_{\alpha}$$

where $\text{wp}(t') = x : \{u \mid \theta\} \Rightarrow \star$, $Y = \text{args}^{-1}(\text{sty}(t')) = \{\alpha\} \setminus \{x\}$, and $x$ is the first element of $\alpha$ (i.e., by alpha-renaming). Therefore, by Lemma E.2, $\text{wp}(e, \Psi_{wp})\rho = (\{\alpha\} \setminus \{u\}_\rho = u \Rightarrow \theta)$, and so $\Gamma \vdash \{u\}_\rho = u \Rightarrow \theta$. Hence, it follows that $\Gamma \vdash_{\tau} e : \star$.

The case $\text{sty}(t_n) = \star$ is similar.

Finally, the following is clear from the construction of $\text{wp}(t')$.

Lemma E.11. Let $F \rightarrow d = e_F \in d$ and $\text{wp}(F) = \Downarrow \rightarrow \star$. Then,

- For each function type variable $x \in \text{dom}(\Downarrow)$, $(\Downarrow)(x) = \text{wp}(x)$ with $\text{fv}(\text{wp}(x)) = \{x\} \subseteq \text{dom}(\Downarrow)$. (Namely, we select $\Downarrow$ to be the one bound universally just before $x$.)

Let $\Delta$ be such that $\Delta(F) = \text{wp}(F)$ for each $F$. By Lemma E.10 (C) and Lemma E.11, we have $\Delta, (\Downarrow) \vdash_{\tau} e_F : \star$ where $\Delta(F) = \Downarrow \rightarrow \star$, for each $F \rightarrow d = e_F \in d$. Because $\Delta$ is well-formed (cf. Corollary E.6) and $\Delta(\text{main}) = \text{int} \rightarrow \star$ (cf. Corollary E.4), we have $\Delta \vdash_{\tau} d$.

F. Constraint Generation Process

\[
\begin{array}{c}
sty(x) = \text{int} \\
cgen(\Gamma \vdash e) = (\{u \mid u = x\}, \Gamma)
\end{array}
\]

\[
\begin{array}{c}
sty(\kappa) = \kappa \\
cgen(\Gamma \vdash \kappa) = (\Gamma(\kappa), \Gamma)
\end{array}
\]

\[
\begin{array}{c}
cgen(\Gamma \vdash e) = (\Psi(c), \Gamma)
\end{array}
\]

\[
\begin{array}{c}
(\sigma_1 \cdot \theta_1) = cgen(\Gamma \vdash e_1) \\
(\star, \theta_1) = cgen(\Gamma \vdash e_1)
\end{array}
\]

\[
\begin{array}{c}
(\sigma_1 \cdot \theta_1) = cgen(\Gamma \vdash e_1) \\
(\star, \theta_1) = cgen(\Gamma \vdash e_1)
\end{array}
\]

\[
\begin{array}{c}
(\sigma_1 \cdot \theta_1) = cgen(\Gamma \vdash e_1) \\
(\star, \theta_1) = cgen(\Gamma \vdash e_1)
\end{array}
\]

Figure 15. The constraint generation rules.

Figure 15 shows the constraint generation rules. Here, $cgen(\Gamma \vdash e)$ returns a pair of a refinement type $\tau$ and a constraint $\theta$ such that $\Gamma \vdash e : \tau$ is derivable if and only if $\theta$ is valid. Similarly, $cgen(\Gamma \vdash \tau_1 \leq \tau_2)$ returns a constraint $\theta$ such that $\Gamma \vdash \tau_1 \leq \tau_2$ is derivable if and only if $\theta$ is valid.