

System FR: Formalized Foundations for the Stainless Verifier

JAD HAMZA, EPFL, Switzerland
NICOLAS VOIROL, EPFL, Switzerland
VIKTOR KUNČAK, EPFL, Switzerland

We present the design, implementation, and foundation of a verifier for higher-order functional programs with generics and recursive data types. Our system supports proving safety and termination using preconditions, postconditions and assertions. It supports writing proof hints using assertions and recursive calls. To formalize the soundness of the system we introduce System FR, a calculus supporting System F polymorphism, dependent refinement types, and recursive types (including recursion through contravariant positions of function types). Through the use of sized types, System FR supports reasoning about termination of lazy data structures such as streams. We formalize a reducibility argument using the Coq proof assistant and prove the soundness of a type-checker with respect to call-by-value semantics, ensuring type safety and normalization for typeable programs. Our program verifier is implemented as an alternative verification-condition generator for the Stainless tool, which relies on the Inox SMT-based solver backend for automation. We demonstrate the efficiency of our approach by verifying a collection of higher-order functional programs comprising around 14000 lines of polymorphic higher-order Scala code, including graph search algorithms, basic number theory, monad laws, functional data structures, and assignments from popular Functional Programming MOOCs.

CCS Concepts: • **Theory of computation** → **Logic and verification; Programming logic; Type theory;**
• **Software and its engineering** → **Software verification;** • **Computing methodologies** → *Theorem proving algorithms.*

Additional Key Words and Phrases: software verification, dependent types, recursive types, System F, SMT

ACM Reference Format:

Jad Hamza, Nicolas Voirol, and Viktor Kunčak. 2019. System FR: Formalized Foundations for the Stainless Verifier. *Proc. ACM Program. Lang.* 3, OOPSLA, Article 166 (October 2019), 30 pages. <https://doi.org/10.1145/3360592>

1 INTRODUCTION

Automatically verifying the correctness of higher-order programs is a long-standing problem that arises in most programming languages and proof assistants. Despite extensive research in program verifiers and proof assistants [Abel 2010; Bertot and Castéran 2004a; Brady 2013; Gordon and Melham 1993; Harrison 2009, 2017; Leino 2010; Nipkow et al. 2002a; Norell 2007; Swamy et al. 2013; Vazou et al. 2014] there remain significant challenges and trade-offs in checking safety and termination. A motivation for our work are implementations that verify polymorphic functional programs using SMT solvers [Suter et al. 2011; Vazou et al. 2014]. To focus on foundations, we look at simpler verifiers that do not perform invariant inference and are mostly based on unfolding recursive definitions and encoding of higher-order functions into SMT theories [Blanc and Kuncak 2015; Suter et al. 2011; Voirol et al. 2015]. A recent implementation of such a verifier is the Stainless

Authors' addresses: Jad Hamza, LARA, EPFL, Switzerland, jad.hamza@epfl.ch; Nicolas Voirol, LARA, EPFL, Switzerland, nicolas.voirol@epfl.ch; Viktor Kunčak, EPFL IC LARA - INR 318 - Station 14, EPFL, Switzerland, viktor.kuncak@epfl.ch.

Permission to make digital or hard copies of part or all 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. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

© 2019 Copyright held by the owner/author(s).

2475-1421/2019/10-ART166

<https://doi.org/10.1145/3360592>

system [LARA 2019], which handles a subset of Scala [Odersky et al. 2008]. The goal of Stainless is to verify that function contracts hold and that all functions terminate. It was shown [Hupel and Kuncak 2016] how to map certain patterns of specified Scala programs into Isabelle/HOL. Whereas this approach ensures soundness, it does not reuse the reasoning of Stainless and it can verify only some of the programs that Stainless verifier can prove. The present paper seeks to provide direct foundations for verification and termination checking of functional programs with a rich set of features for purely functional programming including non-monotonic data types. On the other hand, our calculus does not aspire to directly support effects for which there exist excellent other systems [Swamy et al. 2013].

The subtleties of ensuring function termination have been an initial impetus for the calculus we present. Termination is desirable for many executable functions in programs and is even more important in formal specifications. A non-terminating function definition such as $f(x) = 1 + f(x)$ could be easily mapped to a contradiction and violate the conservative extension principle for definitions. Yet termination in the presence of higher-order functions and data types is challenging to ensure. For example, when using non-monotonic recursive types, terms can diverge even without the explicit use of recursive functions, as illustrated by the following snippet of Scala code:

```
case class D(f: D ⇒ Unit) // non-monotonic recursive type
def g(d: D): Unit = d.f(d) // non-recursive function definition
g(D(g)) // diverging term, reduces to D(g).f(D(g)) and then again to g(D(g))
```

Furthermore, even though the concept of termination for all function inputs is an intuitively clear property, its modular definition is subtle: a higher order function g taking another function f as an argument should terminate when given any terminating function f , which, in turn, can be applied to expressions involving further calls to g . The quest for solid foundations for termination led us to type theoretic techniques, where *reducibility* method has long been used to show strong normalization of expressive calculi [Tait 1967], [Girard 1990, Chapter 6], [Harper 2016]. As a natural framework for analyzing support for first-class functions with preconditions and post-conditions we embraced the ideas of refinement dependent types similar to those in Liquid Haskell [Vazou et al. 2014] with refinement-based notion of subtyping. To explain proof obligation generation in the higher-order case (including the question of which assumptions should be visible when checking a given assertion), we resorted to well-known dependent (Π) function types. To support parametric polymorphism we incorporated type quantifiers, as in System F [Girard 1971, 1990]. We found that the presence of refinement types allowed us to explain soundness of well-founded recursion based on user-defined measures. The recursion in programs is thus not syntactically restricted as in, e.g., System F. To provide expressive support for iterative unfolding of recursive functions, we introduced rules to make function bodies available while type checking of recursive functions. For recursive type definitions, many existing systems introduce separate notions of inductive and co-inductive definitions. We found this distinction less natural for developers and chose to support expressive recursive types (without a necessary restriction to positive recursion) using sized types [Abel 2010]. We draw inspiration from a number of existing systems, yet our solution is a new sound combination of features that work nicely together.

We combined these features into a new type system, System FR, which we present as a bidirectional type checking algorithm. The algorithm generates type checking and type inference goals by traversing terms and types, until it reaches a point where it has to check that a given term evaluates to true. This typically arises when we want to check that a term t has a refinement type $\{x : T \mid b\}$, which is the case when t has type T , and when the term b evaluates to true in the context where x equals t . Following the tradition of SMT-based verifiers [Barnett et al. 2004; Detlefs et al. 1998], we use the term *verification condition* (VC) to refer to a term that should evaluate to true.

We prove the soundness of our type system using a reducibility interpretation of types. The goal of our verification system is to ensure that a given term belongs to the semantic denotation of a given type. For simple types such as natural numbers, this denotation is the set of untyped lambda calculus terms that evaluate, in a finite number of steps, to a non-negative integer. For function types the denotation are, as is typical in reducibility approaches, terms that, when applied to terms in denotation of argument type, evaluate to terms in the denotation of the result type. Such denotation gives us a unified framework for function contracts expressed as refinement types. The approach ensures termination of programs because the semantics of types only contain terms that are terminating in call-by-value semantics.

We have formally proven using the Coq proof assistant [Bertot and Castéran 2004a] the soundness of our typing algorithm, implying that when verification conditions generated for checking that a term t belongs to a type T are semantically valid, the term t belongs to the semantic denotation of the type T . The bidirectional typing algorithm handles the expressive types in a deterministic and predictable way, which enables good and localized error reporting to the user. To solve generated verification conditions, we use existing implementation invoking the Inox solver¹ that reduces higher-order queries to the first-order language of SMT solvers [Voirol et al. 2015]. Our semantics of types provides a definition of soundness for such solvers; any solver that respects the semantics can be used with our verification condition generator. Our bidirectional type checking algorithm thus becomes a new, trustworthy verification condition generator for Stainless. We were successful in verifying many existing Stainless benchmarks using the new approach.

We summarize our contributions as follows:

- We present a rich type system, called System FR, that combines System F with dependent types, refinements, equality types, and recursive types (Sections 3 and 4).
- We define a bidirectional type-checking algorithm for System FR (Section 5). Our algorithm generates verification conditions that are solved by the (existing) SMT-based solver Inox.
- We prove² soundness of our bidirectional type-checking algorithm that reduces program correctness to proving that certain formulas always evaluate to true (Section 6). Our formalization also supports additional expressive notions, such as infinite intersections and unions as well as refinement conditions given by non-emptiness of an arbitrary type.
- We built a verification condition generator based on these foundations and evaluated it on around 14k lines of benchmarks (Section 7), showing that generating proof obligations using type checking is effective in practice.

```

def f(x: τ1): τ2 = {
  require(pre[x])
  decreases(m[x])
  E[x, f]
} ensuring { res ⇒ post[x, res] }

```

Fig. 1. Template of a recursive function with user-given contracts and a decreasing measure.

2 EXAMPLES OF PROGRAM VERIFICATION AND TERMINATION CHECKING

Our goal is to verify correctness and termination of pure Scala functions written as in Figure 1. $pre[x]$ is the precondition of the function f , and is written by the user in the same language as

¹<https://github.com/epfl-lara/inox>

²<https://github.com/epfl-lara/SystemFR/tree/oopsla2019>

the body of f . The precondition may contain arbitrary expressions and calls to other functions. Similarly, the user specifies in *post* the property that the results of the function should satisfy. To ensure termination of f (which might call itself recursively), the user may also provide a measure using the **decreases** keyword, which is also an expression (of type Nat , the type of natural numbers) written in the same language. τ_1 and τ_2 may be arbitrary types, including function types or algebraic data types. Informally, the function is terminating and correct, if, for every value v of type τ_1 such that $\text{pre}[v]$ evaluates to true , $f(v)$ returns (in a finite number of steps) a value res of type τ_2 such that $\text{post}[v, \text{res}]$ evaluates to true . By using dependent and refinement types, this can be summarized by saying that the function f has type: $\Pi x : \{x : \tau_1 \mid \text{pre}[x]\}. \{res : \tau_2 \mid \text{post}[x, res]\}$.

```
sealed abstract class List
case object Nil extends List
case class Cons(head: Z, tail: List) extends List
```

```
def filter(l: List, p: Z => Boolean): List = {
  decreases(size(l))
  | match {
    case Nil => Nil
    case Cons(h, t) if p(h) => Cons(h, filter(t, p))
    case Cons(_, t) => filter(t, p) } }

def count(l: List, x: Z): Z = {
  decreases(size(l))
  | match {
    case Nil => 0
    case Cons(h, t) => (if (h == x) 1 else 0)
                      + count(t, x) }
```

Fig. 2. The function `filter` filters elements of a list based on a predicate p , and `count` counts the number of occurrences of x in a list.

```
def partition(
  l: List[Z],
  p: X => Bool
): (List[Z], List[Z]) = {
  decreases(size(l))
  | match {
    case Nil => Nil
    case x :: xs =>
      val (l1, l2) = partition(xs, p)
      if (p(x)) (x :: l1, l2)
      else (l1, x :: l2) }
  } ensuring { res =>
    res._1 == filter(l, p) &&
    res._2 == filter(l, x => !p(x)) }
```

Fig. 3. A partition function specified using `filter` and with termination measure is given with `size`.

As an example, consider the list type as defined in Figure 2. We use \mathbb{Z} to denote the type of integers (corresponding to Scala’s `BigInt` in actual source code). The function `filter` filters elements from a list, while `count` counts the number of occurrences of an integer in the list. These two functions have no pre- or postconditions. The `decreases` clauses specify that the functions terminate because the *size* of the list decreases at each recursive call.

Using these functions we define `partition` in Figure 3, which takes a list l of integers and partitions it according to a predicate $p: \mathbb{Z} \Rightarrow \text{Bool}$. We prove in the postcondition that partitioning coincides with applying `filter` to the list with p and its negation.

```
def partitionMultiplicity(@induct l: List[Z], p: Z => Boolean, x: Z): Boolean = {
  val (l1, l2) = partition(l, p)
  count(l, x) == count(l1, x) + count(l2, x)
} holds
```

Fig. 4. A proof (by induction on l) that partitioning a list preserves the multiplicity of each element.

```

def isSorted(l: List): Boolean = {
  decreases(size(l))
  l match {
    case Nil() => true
    case Cons(x, Nil()) => true
    case Cons(x, Cons(y, ys)) =>
      x ≤ y && isSorted(Cons(y, ys))
  }
}

def merge(l1: List, l2: List): List = {
  require(isSorted(l1) && isSorted(l2))
  decreases(size(l1) + size(l2))
  (l1, l2) match {
    case (Cons(x, xs), Cons(y, ys)) =>
      if (x ≤ y) Cons(x, merge(xs, l2))
      else Cons(y, merge(l1, ys))
    case (Cons(_, _), Nil) => l1
    case _ => l2
  }
  ensuring { res => isSorted(res) }
}

```

Fig. 5. A function that checks whether a list is sorted and a function that merges two sorted lists

Figure 4 shows a theorem that `partition` also preserves the multiplicity of each element. We use here `count` to state the property, but we could have used `Multisets` instead (a type which is natively supported in Stainless). The `holds` keyword is a shorthand for `ensuring { res => res }`. The `@induct` annotation instructs the system to add a recursive call to `partitionMultiplicity` on the tail of `l` when `l` is not empty. This gives us access to the multiplicity property for the tail of `l`, which the system can then use automatically to prove that the property holds for `l` itself. This corresponds to a proof by induction on `l`.

Figure 5 shows a function `isSorted` that checks whether a list is sorted, and a function `merge` that combines two sorted lists in a sorted list. When given the above input, the system proves the termination of all functions, establishes that postconditions of functions hold, and shows that the theorem holds, without any user interaction or additional annotations. For the `merge` function, the postcondition might seem too weak to establish that e.g. `Cons(x, merge(xs, l2))` is sorted just based on the fact that `merge(xs, l2)` is sorted. However, since we put the definition of `merge` *while* type-checking the body of `merge` in the context, it is possible to establish that `x` is smaller than the head of `merge(xs, l2)`. We give more details on this feature of our system, called *body-visible recursion*, in Section 5.3.

2.1 Reasoning about Streams

Our system also supports reasoning about infinite data structures, including streams that are computed on demand. These data structures are challenging to deal with because even defining termination of an infinite stream is non-obvious, especially in absence of a concrete operation that uses the stream. Given some type `X`, `Stream[X]` represents the type of infinite streams containing elements in `X`. In a mainstream call-by-value language such as Scala, this type can be defined as:

```
case class Stream[X](head: X, tail: () => Stream[X])
```

For the sake of concise syntax, we typeset a function taking unit, $(u: \text{Unit}) \Rightarrow e$, using Scala's syntax $() \Rightarrow e$ for a function of zero parameters. Given a stream `s: Stream[X]`, we can call `s.head` to get the head of the stream (which is of type `X`), or `s.tail` to get the tail of the stream (which is of type $() \Rightarrow \text{Stream}[X]$). We can use recursion to define streams, as shown in figures 6, 8, 7. The `@erasable` annotation is used to mark the erasable parameters `n` of these functions. These parameters are used as annotations to guide our type-checker, but they do not influence the computation and can be erased at runtime. For instance, an erased version of `constant` (without erasable code and without type annotation) looks like:

```
def constant(x) = Stream(x, () => constant(x))
```

```

def constant[X](@erasable n: Nat, x: X): Stream[X](n) = {
  decreases(n)
  Stream(n)(x, () $\Rightarrow$  constant[X](n-1,x)) }

```

Fig. 6. Constant stream

```

def zipWith[X,Y,Z](@erasable n: Nat,
  f: X  $\Rightarrow$  Y  $\Rightarrow$  Z, s1: Stream[X](n), s2: Stream[Y](n)): Stream[Z](n) = {
  decreases(n)
  Stream[Z](n)(f (s1.head) (s2.head), () $\Rightarrow$  zipWith[X,Y,Z](n-1, f, s1.tail(), s2.tail())) }

```

Fig. 7. Zip function that combines elements of two streams using a two-argument function f

```

def fib(@erasable n: Nat): Stream[Z](n) = {
  decreases(n)
  Stream[Z](n)(0, () $\Rightarrow$  Stream[Z](n-1)(1, () $\Rightarrow$ 
  zipWith[Z,Z,Z](n-2, plus, fib(n-2), fib(n-1).tail())) }

```

Fig. 8. Fibonacci stream defined using zipWith

Informally, we can say that the constant stream is *terminating*. Indeed, it has the interesting property that, despite the recursion, for every $n \in \mathbb{N}$, we can take the first n elements in finite time (no divergence in the computation). We say that $\text{constant}(x)$ is an *n -non-diverging stream*. Moreover, when a stream is n -non-diverging for every $n \in \mathbb{N}$, we simply say that it is *non-diverging*, which means that we can take as many elements as we want without diverging, which is the case for $\text{constant}(x)$. Note that non-divergence of constant cannot be shown by defining a measure on its argument x that strictly decreases on each recursive call, because constant is called recursively on the exact same argument x . Instead, we define a measure on the erasable argument n of the annotated version. This corresponds to using type-based termination [Abel 2007, 2008; Barthe et al. 2008], where the type of the function for the recursive call is smaller than the type of the caller. We expand on that technique in Section 5.2.

In the annotated version of constant from Figure 6, the notation $\text{Stream}[X](n)$ stands for streams of elements in X which are n -non-diverging. The type of constant then states that constant can be called with any (erasable) parameter n to build an n -non-diverging stream. Since parameter n is computationally irrelevant, this proves that the erased version of constant returns a non-diverging stream.

We now give a variant of constant which is *diverging*:

```

def badConstant(x) = Stream(x, () $\Rightarrow$  badConstant(x).tail())

```

Indeed, given some x , a call to $\text{badConstant}(x).\text{tail}()$ ends up evaluating $\text{badConstant}(x).\text{tail}()$ again, and would diverge.

The zipWith function in Figure 7 takes two streams and a function f . It creates a new stream by applying f to pairs of elements taken from each stream. For zipWith , we can verify that as long as f terminates on every input, and $s1$ and $s2$ are non-diverging streams, then zipWith returns a non-diverging stream. We can then use zipWith to define the well-known Fibonacci stream (Figure 8), an infinite stream containing the Fibonacci sequence: 0, 1, 1, 2, 3, 5, 8, 13, etc. We make use of a function $\text{plus}:\mathbb{Z} \Rightarrow \mathbb{Z} \Rightarrow \mathbb{Z}$ that computes the sum of two integers. Just like the constant

streams, `fib` is a non-diverging stream. For instance, calling `fib.tail().tail().tail().head` returns (in finite time) the number 2 from the Fibonacci sequence.

The important property that the type signature of `zipWith` ensures is that, for every $n \in \text{Nat}$, if `s1` and `s2` are n -non-diverging streams, then `zipWith f s1 s2` is n -non-diverging as well. Our type system can check this property and then use it to make sure that the definition of `fib` type-checks. We can also prove further properties of interest, e.g., that zipping two streams `s1` and `s2` with the function $(x : \text{Nat}) \Rightarrow (y : \text{Nat}) \Rightarrow x$ returns a stream that behaves as the stream `s1`.

$$t ::= x \mid () \mid \lambda x. t \mid t \ t \mid (t, t) \mid \pi_1 t \mid \pi_2 t \mid$$

$$\text{left}(t) \mid \text{right}(t) \mid \text{either_match}(t, x \Rightarrow t, x \Rightarrow t) \mid \text{true} \mid \text{false} \mid \text{if } t \text{ then } t \text{ else } t \mid$$

$$\text{zero} \mid \text{succ}(t) \mid \text{rec}(t, t, (n, y) \Rightarrow t) \mid \text{fix}(y \Rightarrow t) \mid \text{match}(t, t, n \Rightarrow t) \mid$$

$$\text{fold}(t) \mid \text{unfold } t \text{ in } x \Rightarrow t \mid \Lambda t \mid t[] \mid \text{err} \mid \text{let } x = t \text{ in } t \mid \text{size}(t)$$

Fig. 9. Grammar for untyped lambda calculus terms

3 SYNTAX AND OPERATIONAL SEMANTICS

We now give a formal syntax for terms and show call-by-value operational semantics. This untyped lambda calculus with pairs, tagged unions, integers, booleans, and error values models programs that our verification system supports. It is Turing complete and rather conventional.

3.1 Terms of an Untyped Calculus

Let \mathcal{V} be a set of *variables*. We let *Terms* be the set of all (untyped) terms (see Figure 9) which includes the unit term `()`, pairs, booleans, natural numbers, a recursor `rec` for iterating over natural numbers, a pattern matching operator `match` for natural numbers, a recursion operator `fix`, an error term `err` to represent crashes. The recursor `rec` can be simulated using `fix` and `match` but we keep it in the paper for presenting examples.

The terms `fold(t)` and `unfold t1 in x ⇒ t2` are used to represent data structures (such as lists or streams), where ‘`fold()`’ plays the role of a constructor, and ‘`unfold in`’ the role of a deconstructor. The terms `Λt` and `t[]` are used to represent the erasure of type abstractions and type instantiation terms (for polymorphism) of the form $\Lambda \alpha. t$ and $t[\tau]$, where α is a type variable and τ is a type. These annotated terms will be introduced in a further section.

The term `size(t)` is a special term to internalize the sizes of syntax trees of values (ignoring lambdas) of our language. It is used for measure of recursive functions such as the `map` examples on lists shown in Section 2.

Given a term t we denote $\text{fv}(t)$ the set of all *free variables* of t . Terms are considered up to renaming of locally bound variables (*alpha-renaming*).

3.2 Call-by-Value Operational Semantics

The set *Val* of *values* of our language is defined (inductively) to be `zero`, `()`, `true`, `false`, every variable x , every lambda term $\lambda x. t$ or Λt , the terms of the form `succ(v)` or `fold(v)` where $v \in \text{Val}$, and the terms of the form (v_1, v_2) where $v_1, v_2 \in \text{Val}$.

The call-by-value small-step relation between two terms $t_1, t_2 \in \text{Terms}$, written $t_1 \hookrightarrow t_2$, is standard for the most part and given in Figure 10. Given a term t and a value v , $t[x \mapsto v]$ denotes the term t where every free occurrence of x has been replaced by v .

$$\begin{array}{c}
\frac{v_1 \in \text{Val} \quad v_2 \in \text{Val}}{\pi_1(v_1, v_2) \hookrightarrow v_1} \quad (\beta_1) \qquad \frac{v_1 \in \text{Val} \quad v_2 \in \text{Val}}{\pi_2(v_1, v_2) \hookrightarrow v_2} \quad (\beta_2) \\
\frac{v \in \text{Val}}{(\lambda x. t) v \hookrightarrow t[x \mapsto v]} \quad (\beta_3) \qquad \frac{}{(\Lambda t)[] \hookrightarrow t} \quad (\beta_4) \\
\frac{}{\text{if true then } t_1 \text{ else } _ \hookrightarrow t_1} \quad (\beta_5) \qquad \frac{}{\text{if false then } _ \text{ else } t_2 \hookrightarrow t_2} \quad (\beta_6) \\
\frac{}{\text{rec}(\text{zero}, t_0, _) \hookrightarrow t_0} \quad (\beta_7) \\
\frac{v \in \text{Val}}{\text{rec}(\text{succ}(v), t_0, (n, y) \Rightarrow t_s) \hookrightarrow t_s[n \mapsto v, y \mapsto \lambda u. \text{rec}(v, t_0, (n, y) \Rightarrow t_s)]} \quad (\beta_8) \\
\frac{}{\text{fix}(y \Rightarrow t) \hookrightarrow t[y \mapsto \lambda u. \text{fix}(y \Rightarrow t)]} \quad (\beta_9) \\
\frac{}{\text{match}(\text{zero}, t_0, _) \hookrightarrow t_0} \quad (\beta_{10}) \qquad \frac{v \in \text{Val}}{\text{match}(\text{succ}(v), _, n \Rightarrow t_s) \hookrightarrow t_s[n \mapsto v]} \quad (\beta_{11}) \\
\frac{v \in \text{Val}}{\text{either_match}(\text{left}(v), x \Rightarrow t, _) \hookrightarrow t[x \mapsto v]} \quad (\beta_{12}) \\
\frac{v \in \text{Val}}{\text{either_match}(\text{right}(v), _, x \Rightarrow t) \hookrightarrow t[x \mapsto v]} \quad (\beta_{13}) \\
\frac{v \in \text{Val}}{\text{let } x = v \text{ in } t \hookrightarrow t[x \mapsto v]} \quad (\beta_{14}) \qquad \frac{v \in \text{Val}}{\text{unfold fold}(v) \text{ in } x \Rightarrow t \hookrightarrow t[x \mapsto v]} \quad (\beta_{15}) \\
\frac{v \in \text{Val}}{\text{size}(v) \hookrightarrow \text{buildNat}(\text{size_sem}(v))} \quad (\beta_{16}) \qquad \frac{t \hookrightarrow t'}{\mathcal{E}[t] \hookrightarrow \mathcal{E}[t']} \quad (\text{Evaluation Context})
\end{array}$$

Fig. 10. Small-step call-by-value operational semantics of untyped terms

$$\begin{array}{ll}
\text{size_sem}(\text{succ}(v)) = 1 + \text{size_sem}(v) & \text{size_sem}(\text{fold}(v)) = 1 + \text{size_sem}(v) \\
\text{size_sem}(\text{right}(v)) = 1 + \text{size_sem}(v) & \text{size_sem}(\text{left}(v)) = 1 + \text{size_sem}(v) \\
\text{size_sem}(v_1, v_2) = \text{size_sem}(v_1) + \text{size_sem}(v_2) & \text{size_sem}(v) = 0, \text{ in all other cases}
\end{array}$$

Fig. 11. Definition of $\text{size_sem}(v)$ for a value v .

To evaluate the fixpoint operator fix , we use the rule $\text{fix}(y \Rightarrow t) \hookrightarrow t[y \mapsto \lambda(). \text{fix}(y \Rightarrow t)]$, which substitutes the fix under a lambda with unit argument. We do this wrapping of fix in a lambda term because we wanted all substitutions to be values for our call-by-value semantics, and fix is not. This also means that, to make a recursive call within t , one has to use $y()$ instead of y .

To define the semantics of $\text{size}()$, we use a (mathematical) function $\text{size_sem}()$ that returns the size of a value, ignoring lambdas for which it returns 0. The precise definition is given in Figure 11.

We make use in the operational semantics of an *evaluation context* \mathcal{E} , which specifies through a *hole* \mathcal{H} the next place where reduction can occur in a term. Inductively, an evaluation context \mathcal{E} must be of one of the following forms:

$$\begin{array}{l}
\mathcal{H} \mid \mathcal{E} e \mid v \mathcal{E} \mid (\mathcal{E}, e) \mid (v, \mathcal{E}) \mid \pi_1 \mathcal{E} \mid \pi_2 \mathcal{E} \mid \text{succ}(\mathcal{E}) \mid \\
\text{rec}(\mathcal{E}, _, _) \mid \text{match}(\mathcal{E}, _, _) \mid \text{if } \mathcal{E} \text{ then } t \text{ else } t \mid \text{let } x = \mathcal{E} \text{ in } t \mid \text{size}(\mathcal{E}) \mid \\
\text{fold}(\mathcal{E}) \mid \text{unfold } \mathcal{E} \text{ in } _ \mid \mathcal{E}[] \mid \text{left}(\mathcal{E}) \mid \text{right}(\mathcal{E}) \mid \text{either_match}(\mathcal{E}, _, _)
\end{array}$$

Given a term t , we denote by $\mathcal{E}[t]$ the context \mathcal{E} where the hole \mathcal{H} has been replaced by t .

We denote by \hookrightarrow^* the reflexive and transitive closure of \hookrightarrow . A term t is *normalizing* if there exists a value v such that $t \hookrightarrow^* v$.

$$\tau ::= \text{Unit} \mid \text{Bool} \mid \text{Nat} \mid \top \quad | \quad \Pi x : \tau. \tau \mid \Sigma x : \tau. \tau \mid \tau + \tau \quad | \quad \forall x : \tau. \tau \mid \forall \alpha : \text{Type}. \tau \\ \text{Rec}(n)(\alpha \Rightarrow \tau) \quad | \quad \{x : \tau \mid t\} \mid t \equiv t$$

Fig. 12. Grammar for types τ , where $x \in \mathcal{V}$ is a term variable, $\alpha \in \mathcal{V}$ is a type variable (t denotes annotated terms of Figure 16 that complete the mutually recursive definition)

4 TYPES, SEMANTICS AND REDUCIBILITY

We give in Figure 12 the grammar for the *types* τ that our verification system supports. Given two types τ_1 and τ_2 , we use the notation $\tau_1 \rightarrow \tau_2$ for $\Pi x : \tau_1. \tau_2$ when x is not a free variable of τ_2 . Similarly, we use the notation $\tau_1 \times \tau_2$ for $\Sigma x : \tau_1. \tau_2$ when x is not a free variable of τ_2 .

For recursive types, we introduce the notation:

$$\text{Rec}(\alpha \Rightarrow \tau) \triangleq \forall n : \text{Nat}. \text{Rec}(n)(\alpha \Rightarrow \tau)$$

Then, the type of (non-diverging) streams informally introduced in Section 2 can be understood as a notation, when X is a type, for: $\text{Stream}[X] \triangleq \text{Rec}(\alpha \Rightarrow X \times (\text{Unit} \rightarrow \alpha))$. Similarly, for a natural number n , the type of n -non-diverging streams $\text{Stream}_n[X]$ is a notation for $\text{Rec}(n)(\alpha \Rightarrow X \times (\text{Unit} \rightarrow \alpha))$. Using this notation, we can also define finite data structures such as lists of elements from X , as follows: $\text{List}[X] \triangleq \text{Rec}(\alpha \Rightarrow \text{Unit} + X \times \alpha)$.

We show in Section 4.3 that these types indeed correspond to streams and lists respectively.

Let Type be the set of all types. We define a (unary) logical relation on types to describe terms that do not get stuck (e.g. due to the error term `err`, or due to an ill-formed application such as ‘true zero’) and that terminate to a value of the given type. Our definition is inspired by the notion of *reducibility* or *hereditary termination* (see e.g. [Girard 1990; Harper 2016; Tait 1967]), which we use as a guiding principle for designing the type system and its extensions.

4.1 Reducibility for Closed Terms

For each type τ , we define in Figure 13 mutually recursively the sets of *reducible values* $\llbracket \tau \rrbracket_{\mathcal{V}}^{\theta}$ and *reducible terms* $\llbracket \tau \rrbracket_{\mathcal{T}}^{\theta}$. In that sense, a type τ can be understood as a *specification* that some terms satisfy (and some do not).

These definitions require an environment θ , called an *interpretation*, to give meaning to type variables. Concretely, an interpretation is a partial map from type variables to sets of terms. An interpretation θ has the constraint that for every type variable $\alpha \in \text{dom}(\theta)$, $\theta(\alpha)$ is a *reducibility candidate* C , which, in our setting, means that all terms in $\theta(\alpha)$ are (erased) values. The set of all reducibility candidates is denoted by $\text{Candidates} \subseteq 2^{\text{Terms}}$, and an interpretation θ is therefore a partial map in $\mathcal{V} \mapsto \text{Candidates}$.

When the interpretation has no influence on the definition, we may omit it. For instance, for every $\theta \in (\mathcal{V} \mapsto \text{Candidates})$, we have $\llbracket \text{Nat} \rrbracket_{\mathcal{V}}^{\theta} = \{\text{zero}, \text{succ}(\text{zero}), \text{succ}(\text{succ}(\text{zero})), \dots\}$, so we can just denote this set by $\llbracket \text{Nat} \rrbracket_{\mathcal{V}}$.

By construction, $\llbracket \tau \rrbracket_{\mathcal{V}}^{\theta}$ only contains (erased) *values* (of type τ), while $\llbracket \tau \rrbracket_{\mathcal{T}}^{\theta}$ contains (erased) *terms* that reduce to a value in $\llbracket \tau \rrbracket_{\mathcal{V}}^{\theta}$. For example, a term in $\llbracket \text{Nat} \rightarrow \text{Nat} \rrbracket_{\mathcal{T}}^{\theta}$ is not only normalizing as a term of its own, but also normalizes whenever applied to a value in $\llbracket \text{Nat} \rrbracket_{\mathcal{V}}^{\theta}$.

The type $\{x : \tau \mid b\}$ represents the values v of type τ for which $b[x \mapsto v]$ evaluates to true. We use this type as a building block for writing specifications (pre and postconditions).

The type $\forall x : \tau_1. \tau_2$ represents the values that are in the intersection of the types $\tau_2[x \mapsto a]$ when a ranges over values of type τ_1 . This type differs from $\Pi x : \tau_1. \tau_2$ in the sense that a value in $\forall x : \tau_1. \tau_2$ belongs to every $\tau_2[x \mapsto a]$ for a in τ_1 , while a value in $\Pi x : \tau_1. \tau_2$ is a function that,

$$\begin{aligned}
\llbracket \alpha \rrbracket_v^\theta &\triangleq \theta(\alpha) \\
\llbracket \top \rrbracket_v^\theta &\triangleq \text{Val} \\
\llbracket \text{Unit} \rrbracket_v^\theta &\triangleq \{()\} \\
\llbracket \text{Bool} \rrbracket_v^\theta &\triangleq \{\text{true}, \text{false}\} \\
\llbracket \text{Nat} \rrbracket_v^\theta &\triangleq \{\text{zero}, \text{succ}(\text{zero}), \text{succ}(\text{succ}(\text{zero})), \dots\} \\
\llbracket \Pi x : \tau_1. \tau_2 \rrbracket_v^\theta &\triangleq \{f \in \text{Val} \mid \forall a \in \llbracket \tau_1 \rrbracket_v^\theta. f a \in \llbracket \tau_2[x \mapsto a] \rrbracket_t^\theta\} \\
\llbracket \forall x : \tau_1. \tau_2 \rrbracket_v^\theta &\triangleq \{b \in \text{Val} \mid \forall a \in \llbracket \tau_1 \rrbracket_v^\theta. b \in \llbracket \tau_2[x \mapsto a] \rrbracket_v^\theta\} \\
\llbracket \Sigma x : \tau_1. \tau_2 \rrbracket_v^\theta &\triangleq \{(a, b) \mid a \in \llbracket \tau_1 \rrbracket_v^\theta \wedge b \in \llbracket \tau_2[x \mapsto a] \rrbracket_v^\theta\} \\
\llbracket \{x : \tau \mid b\} \rrbracket_v^\theta &\triangleq \{a \in \llbracket \tau \rrbracket_v^\theta \mid b[x \mapsto a] \hookrightarrow^* \text{true}\} \\
\llbracket \tau_1 + \tau_2 \rrbracket_v^\theta &\triangleq \{\text{left}(v) \mid v \in \llbracket \tau_1 \rrbracket_v^\theta\} \cup \{\text{right}(v) \mid v \in \llbracket \tau_2 \rrbracket_v^\theta\} \\
\llbracket t_1 \equiv t_2 \rrbracket_v^\theta &\triangleq \{()\} \text{ if } t_1 \approx t_2, \text{ and } \emptyset \text{ otherwise, where} \\
&\quad (t_1 \approx t_2) \triangleq (\forall v \in \text{Val}. (t_1 \hookrightarrow^* v \text{ iff } t_2 \hookrightarrow^* v)) \\
\llbracket \forall \alpha : \text{Type}. \tau \rrbracket_v^\theta &\triangleq \{v \in \text{Val} \mid \forall C \in \text{Candidates}. v[] \in \llbracket \tau \rrbracket_t^{\theta[\alpha \mapsto C]}\} \\
\llbracket \text{Rec}(t)(\alpha \Rightarrow \tau) \rrbracket_v^\theta &\triangleq \{\text{fold}(v) \mid (t \hookrightarrow^* \text{zero} \wedge v \in \llbracket \text{basetype}_\alpha(\tau) \rrbracket_v^\theta) \vee \\
&\quad \exists n \in \llbracket \text{Nat} \rrbracket_v. t \hookrightarrow^* \text{succ}(n) \wedge v \in \llbracket \tau \rrbracket_v^{\theta[\alpha \mapsto \llbracket \text{Rec}(n)(\alpha \Rightarrow \tau) \rrbracket_v^\theta]}\} \\
\llbracket \tau \rrbracket_t^\theta &\triangleq \{t \mid \exists v \in \llbracket \tau \rrbracket_v^\theta. t \hookrightarrow^* v\} \\
&\quad \text{basetype}_\alpha(\Sigma x : \tau_1. \tau_2) \triangleq \Sigma x : \text{basetype}_\alpha(\tau_1). \text{basetype}_\alpha(\tau_2) \\
&\quad \text{basetype}_\alpha(\tau_1 + \tau_2) \triangleq \text{basetype}_\alpha(\tau_1) + \text{basetype}_\alpha(\tau_2) \\
&\quad \text{basetype}_\alpha(\tau) \triangleq \text{if } \alpha \in \text{fv}(\tau) \text{ then } \top \text{ else } \tau
\end{aligned}$$

Fig. 13. Definition of reducibility for values and for terms for each type. The function `basetype()` is an auxiliary function, used in the base case of the definition for recursive types.

when applied to some a in τ_1 , produces a value in $\tau_2[x \mapsto a]$. From a value b in $\forall x : \tau_1. \tau_2$, we can build a value in $\Pi x : \tau_1. \tau_2$ (namely, $\lambda x. b$), while the other way around is not always possible.

The sum type $\tau_1 + \tau_2$ represents values that are either of the form `left(v)` where v is a reducible value of τ_1 , or of the form `right(v)` where v is a reducible value of τ_2 .

The set of reducible values for the equality type $\llbracket t_1 \equiv t_2 \rrbracket_v^\theta$ makes use of a notion of equivalence on terms which is based on operational semantics. More specifically, we say that t_1 and t_2 are *equivalent*, denoted $t_1 \approx t_2$, if for every value v , we have $t_1 \hookrightarrow^* v$ iff $t_2 \hookrightarrow^* v$. Note that this equivalence relation is defined even if we do not know anything about the types of terms t_1 and t_2 , and it ensures that if one of the terms reduces to a value, then so does the other.

The type $\forall \alpha : \text{Type}. \tau$ is the polymorphic type from System F. The set $\llbracket \forall \alpha : \text{Type}. \tau \rrbracket_v^\theta$ is defined by using the environment θ to bind the type variable α to an arbitrary reducibility candidate.

We use the *recursive type* $\text{Rec}(n)(\alpha \Rightarrow \tau)$ as a building block for representing data structures such as lists of streams. The definition of reducibility for the recursive type makes use of an auxiliary

function `basetype()` that can be seen as an (upper) approximation of the recursive type. Note that `basetypeα(τ)` (defined at the bottom of Figure 13) removes the type variable α from τ .

Our reducibility definition respects typical lemmas that are needed to prove the soundness of typing rules, such as the following substitution lemma (see [Girard 1971] for the lemma on System F), which we have formally proven (see also Section 6 below).

LEMMA 4.1. *Let τ_1 and τ_2 be two types, and let α be a type variable that may appear in τ_1 but not in τ_2 . Let θ be a type interpretation. Then, we have:*

$$\llbracket \tau_1 \rrbracket_v^{\theta[\alpha \mapsto \llbracket \tau_2 \rrbracket_v^\theta]} = \llbracket \tau_1[\alpha \mapsto \tau_2] \rrbracket_v^\theta$$

4.1.1 Well-Foundedness of the Reducibility Definition. We can show that the definition given Figure 13 is well-founded by defining a lexicographic measure $(\text{typeNodes}(\tau), \text{index}(\tau)) \in \mathbb{N} \times \text{Terms}$ on types τ . The function `typeNodes(τ)` returns the size of the syntactic tree of type τ , ignoring the terms that appear inside. This size roughly corresponds to the number of (top-level) type constructors in the tree of τ . For example, `typeNodes({x : τ | p})` = 1 + `typeNodes(τ)` (ignoring p), `typeNodes(τ1 + τ2)` = 1 + `typeNodes(τ1)` + `typeNodes(τ2)`, and `typeNodes(t1 ≡ t2)` = 0. Ignoring the size of terms inside types ensures that given a type τ , a term variable x , and a term a , we have: `typeNodes(τ[x ↦ a])` = `typeNodes(τ)`. As a result, the measure `typeNodes(τ)` in the definition of reducibility decreases for indexed types such as $\Sigma x : \tau_1. \tau_2$ or $\Pi x : \tau_1. \tau_2$.

The number `typeNodes(τ)` decreases in every case of Figure 13, *except* for recursive types `Rec(t)(α ⇒ τ)` where the measure stays the same in the recursive call to the denotation on `Rec(n)(α ⇒ τ)` with $t \hookrightarrow^* \text{succ}(n)$. This is where we use the second component of the lexicographic measure, `index(τ)`. We define `index(Rec(t)(α ⇒ τ))` = t and `index(τ)` = `err` for every other type. Then, given $t_1, t_2 \in \text{Terms}$, we consider t_1 to be (strictly) smaller than t_2 if there exist $v_1, v_2 \in \llbracket \text{Nat} \rrbracket_v$, such that $t_1 \hookrightarrow^* v_1, t_2 \hookrightarrow^* v_2$, and v_1 is strictly smaller than v_2 when seen as a natural number. Therefore, the second component `index(Rec(n)(α ⇒ τ))` = n is strictly smaller than `index(Rec(t)(α ⇒ τ))` = t in the definition (because $t \hookrightarrow^* \text{succ}(n)$), which ensures that the overall lexicographic measure decreases.

4.2 Reducibility for Open Terms

Having defined reducibility for closed terms, we now define what it means for a term t with free term and type variables to be reducible for a type τ . Informally, we want to ensure that for every interpretation of the type variables, and for every substitution of values for the term variables, the term t reduces in a finite number of steps to a value in type τ . This is formalized by a (semantic) typing relation $\Theta; \Gamma \models_{\text{Red}} t : \tau$ which is defined as follows.

First, a *context* $\Theta; \Gamma$ is made of a finite set $\Theta \subseteq \mathcal{V}$ of type variables and of a sequence Γ of pairs in $\mathcal{V} \times \text{Type}$. The *domain* of Γ , denoted `dom(Γ)` is the list of variables (in \mathcal{V}) appearing in the left-hand-sides of the pairs. We implicitly assume throughout the paper that all variables appearing in the domains are distinct. This enables us to use Γ as a partial map from \mathcal{V} to Type . We use a sequence to represent Γ as the order of variables is important, since a variable may have a (dependent) type which refers to previous variables in the context.

Given a partial map $\gamma \in \mathcal{V} \mapsto \text{Terms}$, we write $\gamma(t)$ for the term t where every variable x is replaced by $\gamma(x)$. We use the same notation $\gamma(\tau)$ for applying a substitution to a type τ .

Given a context $\Theta; \Gamma$, a *reducible substitution for* $\Theta; \Gamma$ is a pair of partial maps $\theta \in \mathcal{V} \mapsto \text{Candidates}$ and $\gamma \in \mathcal{V} \mapsto \text{Terms}$ where: `dom(θ)` = Θ , `dom(γ)` = `dom(Γ)`, and $\forall x \in \text{dom}(\Gamma). \gamma(x) \in \llbracket \gamma(\Gamma(x)) \rrbracket_v^\theta$.

Note that the substitution γ is also applied to the type $\Gamma(x)$, since $\Gamma(x)$ may be a dependent type with free term variables. The set of all pairs of reducible substitutions for $\Theta; \Gamma$ is denoted $\llbracket \Theta; \Gamma \rrbracket_v$.

Finally, given a context $\Theta; \Gamma$, a term t and a type τ , we say that $\Theta; \Gamma \models_{\text{Red}} t : \tau$ holds when for every pair of substitutions θ, γ for the context $\Theta; \Gamma$, $\gamma(t)$ belongs the reducible values at type $\gamma(\tau)$. Formally, $\Theta; \Gamma \models_{\text{Red}} t : \tau$ is defined to hold when:

$$\forall \theta, \gamma \in \llbracket \Theta; \Gamma \rrbracket_v. \gamma(t) \in \llbracket \gamma(\tau) \rrbracket_t^\theta$$

Our bidirectional type checking and inference algorithm in Section 5 is a sound (even if incomplete) procedure to check $\Theta; \Gamma \models_{\text{Red}} t : \tau$.

4.3 Recursive Types

We explain in this section how to interpret the type $\text{Rec}(n)(\alpha \Rightarrow \tau)$ (see reducibility definition in Figure 13) and how the $\text{Stream}[X]$ and $\text{List}[X]$ types represent streams and lists.

4.3.1 Infinite Streams. For a natural number n , consider the type $S_n \triangleq \text{Stream}_n[\text{Nat}] \triangleq \text{Rec}(n)(\alpha \Rightarrow \text{Nat} \times (\text{Unit} \rightarrow \alpha))$. Let us first see what S_n represents for small values of n . As a shortcut, we use the notations $0, 1, 2, \dots$ for zero, $\text{succ}(\text{zero})$, $\text{succ}(\text{succ}(\text{zero}))$, \dots

The definition $\llbracket S_0 \rrbracket_v$ refers to $\text{basetype}_\alpha(\text{Nat} \times (\text{Unit} \rightarrow \alpha))$, which is $\text{Nat} \times \top$ by definition. This means that $\llbracket S_0 \rrbracket_v$ is the set of values of the form $\text{fold}((a, v))$, where $a \in \llbracket \text{Nat} \rrbracket_v$, and $v \in \text{Val}$.

By unrolling the definition, we get that $\llbracket S_1 \rrbracket_v$ is the set of values of the form $\text{fold}(v)$ where v is in $\llbracket \text{Nat} \times (\text{Unit} \rightarrow \alpha) \rrbracket_v^{\alpha \mapsto \llbracket S_0 \rrbracket_v}$, which is the same (by Lemma 4.1) as $\llbracket \text{Nat} \times (\text{Unit} \rightarrow S_0) \rrbracket_v$. Therefore, $\llbracket S_1 \rrbracket_v$ is the set of values of the form $\text{fold}(a, f)$ where $a \in \llbracket \text{Nat} \rrbracket_v$ and $f \in \llbracket \text{Unit} \rightarrow S_0 \rrbracket_v$. This means that when it is applied to $()$, f terminates and returns a value in $\llbracket S_0 \rrbracket_v^\theta$. Similarly, $\llbracket S_2 \rrbracket_v$ is the set of values of the form $\text{fold}(a, f)$ where $n \in \llbracket \text{Nat} \rrbracket_v$ and $f \in \llbracket \text{Unit} \rightarrow S_1 \rrbracket_v$.

To summarize, we can say that for every $n \in \llbracket \text{Nat} \rrbracket_v$, S_n represents values of the language that behave as streams of natural numbers, as long as they are unfolded at most $n+1$ times. This matches the property we mentioned in Section 2, as S_n represents the streams that are $n+1$ -non-diverging. We can show that as n grows, S_n gets more and more constraints: $\llbracket S_0 \rrbracket_v \supseteq \llbracket S_1 \rrbracket_v \supseteq \llbracket S_2 \rrbracket_v \supseteq \dots$. In the limit, a value $v \in \llbracket \forall n : \text{Nat}. S_n \rrbracket_v$ (which is in every S_n for $n \in \llbracket \text{Nat} \rrbracket_v$), represents a stream of natural numbers, that, regardless of the number of times it is unfolded, does not diverge, i.e. a non-diverging stream. Equivalently, we have $v \in \llbracket \text{Stream}[\text{Nat}] \rrbracket_v$.

4.3.2 Finite Lists. Types of the form $\text{Rec}(\alpha \Rightarrow \tau)$ can also be used to represent finite data structures such as lists. We let $\text{List}_n[X]$ be a notation for $\text{Rec}(n)(\alpha \Rightarrow \text{Unit} + X \times \alpha)$, so that:

$$\text{List}[X] \triangleq \forall n : \text{Nat}. \text{List}_n[X].$$

Here are some examples to show how lists are encoded:

- The empty list is $\text{fold}(\text{left}())$,
- A list with one element n is $\text{fold}(\text{right}(n, \text{fold}(\text{left}())))$,
- Given an element n and a list l , we can construct the list $n :: l$ by writing: $\text{fold}(\text{right}(n, l))$.

Let us now see why $\text{List}[X]$ represents the type of all finite lists of elements in X . The first thing to note is that given $n \in \llbracket \text{Nat} \rrbracket_v$, $\text{List}_n[X]$ does *not* represent the lists of size n . For instance, we know that $\llbracket \text{List}_0[X] \rrbracket_v$ is the set of values of the form $\text{fold}(v)$ where $v \in \llbracket \text{basetype}_\alpha(\text{Unit} + X \times \alpha) \rrbracket_v$, i.e. $v \in \llbracket \top \rrbracket_v = \text{Val}$. Therefore, $\text{List}_0[X]$ contains lists of all sizes (and also all values that do not represent lists, such as $\text{fold}(\text{zero})$ or $\text{fold}(\lambda x. (()))$).

Instead, $\text{List}_n[X]$ can be understood as the values that, as long as they are unfolded no more than n times, behave as lists. As for streams, we have: $\llbracket \text{List}_0[X] \rrbracket_v \supseteq \llbracket \text{List}_1[X] \rrbracket_v \supseteq \llbracket \text{List}_2[X] \rrbracket_v \supseteq \dots$ where the monotonicity follows because α only appears in positive positions in the definitions of the recursive types for streams and lists. In the limit, we can show that $\text{List}[X]$ contains all finite lists, and nothing more.

LEMMA 4.2. *Let $v \in \text{Val}$ be a value and X be some type. Then, $v \in \llbracket \text{List}[X] \rrbracket_v$ if and only if there exists $k \geq 0$ and $a_1, \dots, a_k \in \llbracket X \rrbracket_v$ such that $v = \text{fold}(\text{right}(a_1, \dots \text{fold}(\text{right}(a_k, \text{fold}(\text{left}(\dots))))))$.*

It may seem surprising that the type of streams $\text{Rec}(\alpha \Rightarrow X \times (\text{Unit} \rightarrow \alpha))$ contains infinite streams while the type of lists $\text{Rec}(\alpha \Rightarrow \text{Unit} + X \times \alpha)$ only contains finite lists. The reason is that, in a call-by-value language, a value representing an infinite list would need to have an infinite syntax tree, with infinitely many $\text{fold}()$'s (which is not possible). On the other hand, we can represent infinite streams by hiding recursion underneath a lambda term as shown in Section 2.

5 A BIDIRECTIONAL TYPE-CHECKING ALGORITHM

In this section, we give procedures for inferring a type τ for a term t in a context $\Theta; \Gamma$, denoted $\Theta; \Gamma \vdash t \uparrow \tau$, as well as for checking that the type of a term t is τ , denoted $\Theta; \Gamma \vdash t \downarrow \tau$. We introduce rules of our procedures throughout this section; the full set of rules is given in figures 14 and 15.

Our inference and checking rules give rise to conditions of the form $\Theta; \Gamma \vdash t_1 \equiv t_2$. We call such checks *verification conditions* (in the rules, they are boxed and appear in blue color). The \equiv sign is part of the judgment form, and does not describe a formula. We rely on an external solver to perform these checks, and assume that when the verification condition is considered valid by the solver, then: $\forall \theta, \gamma \in \llbracket \Theta; \Gamma \rrbracket_v. \gamma(\text{erase}(t_1)) \approx \gamma(\text{erase}(t_2))$. This is an equivalent way of saying that $\llbracket \gamma(\text{erase}(t_1)) \equiv \gamma(\text{erase}(t_2)) \rrbracket_v^\theta$ is non-empty. Under these conditions, we have the following theorem.

THEOREM 5.1 (SOUNDNESS OF THE BIDIRECTIONAL TYPE-CHECKER). *If $\Theta; \Gamma \vdash t \uparrow \tau$ holds or if $\Theta; \Gamma \vdash t \downarrow \tau$ holds, then $\Theta; \text{erase}(\Gamma) \models_{\text{Red}} \text{erase}(t) : \text{erase}(\tau)$ holds.*

5.1 Annotated Terms

In order to guide our type-checking algorithm, we require terms to be annotated. We give in Figure 16 the grammar for *annotated terms*. The term $\text{inst}(t_1, t_2)$ is used to instantiate a term t_1 which has a type of the form $\forall x : \tau_2. \tau$ to a particular term t_2 of type τ_2 , in the (*Infer Forall Instantiation*) type inference rule of Figure 14. The term $\text{unfold_pos } t_1 \text{ in } t_2$ is an annotated variant of $\text{unfold } t_1 \text{ in } t_2$ (see rules (*Infer Unfold*) and (*Infer Unfold Positive*) in Figure 14). We discuss the difference between these rules in Section 5.6.

The type $\text{Let } x = t_2 \text{ in } \tau$ represents the type τ where the variable x is bound to t_2 by using let 's in each term that appears in τ . The formal definition is given in Section 5.

Annotations such as $\lambda x : \tau. t$ or $\text{inst}(t_1, t_2)$ have no runtime influence and are erased (respectively to $\lambda x. t$ and t_1). We write $\text{erase}(t)$ to refer to the *erasure* of t , where every annotation has been erased. The full definition is given in the long version of our paper [Hamza et al. 2019].

When a type τ has annotated terms inside, we write $\text{erase}(\tau)$ to erase their annotations. For instance $\text{erase}(\{x : \tau \mid b\})$ refers to $\{x : \tau \mid \text{erase}(b)\}$. Moreover, for a context Γ , we write $\text{erase}(\Gamma)$ to refer to the context Γ where each type τ has been replaced by $\text{erase}(\tau)$.

5.2 Contracts and Measures

The syntax we support in our verification tool translates into our core calculus presented above. In our tool we support named functions with contracts and measures which are desugared into fix terms. To compare natural numbers and express the fact that measures decrease, we use functions ' $<$ ', ' $<=$ ' and ' $==$ ' on natural numbers. These functions can be defined using the recursor rec (see the long version of our paper for definitions [Hamza et al. 2019]).

Figure 17 shows how, thanks to refinement types, the fix term can encode recursive functions (such as the one given in Section 2) that feature user-defined *pre-* and *post-conditions* and whose

$$\begin{array}{c}
\frac{\Gamma(x) = \tau}{\Theta; \Gamma \vdash x \uparrow \tau} \text{ (Infer Var)} \quad \frac{}{\Theta; \Gamma \vdash \text{true} \uparrow \text{Bool}} \text{ (Infer True)} \quad \frac{}{\Theta; \Gamma \vdash \text{false} \uparrow \text{Bool}} \text{ (Infer False)} \\
\frac{}{\Theta; \Gamma \vdash () \uparrow \text{Unit}} \text{ (Infer Unit)} \quad \frac{}{\Theta; \Gamma \vdash \text{zero} \uparrow \text{Nat}} \text{ (Infer Zero)} \quad \frac{\Theta; \Gamma \vdash t \Downarrow \text{Nat}}{\Theta; \Gamma \vdash \text{succ}(t) \uparrow \text{Nat}} \text{ (Infer Succ)} \\
\frac{\Theta; \Gamma \vdash t_1 \Downarrow \text{Bool} \quad \Theta; \Gamma, p : t_1 \equiv \text{true} \vdash t_2 \uparrow \tau_2 \quad \Theta; \Gamma, p : t_1 \equiv \text{false} \vdash t_3 \uparrow \tau_3}{\Theta; \Gamma \vdash \text{if } t_1 \text{ then } t_2 \text{ else } t_3 \uparrow \text{If } t_1 \text{ Then } \tau_2 \text{ Else } \tau_3} \text{ (Infer If)} \\
\frac{\Theta; \Gamma \vdash t \Downarrow \tau_1}{\Theta; \Gamma \vdash \text{left}[t_1 + t_2](t) \uparrow \tau_1 + \tau_2} \text{ (Infer Left)} \quad \frac{\Theta; \Gamma \vdash t \Downarrow \tau_2}{\Theta; \Gamma \vdash \text{right}[t_1 + t_2](t) \uparrow \tau_1 + \tau_2} \text{ (Infer Right)} \quad \frac{\Theta; \Gamma \vdash t \uparrow \tau}{\Theta; \Gamma \vdash \text{size}(t) \uparrow \text{Nat}} \text{ (Infer Size)} \\
\frac{\Theta; \Gamma \vdash t_n \Downarrow \text{Nat} \quad \Theta; \Gamma, p : t_n \equiv \text{zero} \vdash t_0 \uparrow \tau_1 \quad \Theta; \Gamma, n : \text{Nat}, p : t_n \equiv \text{succ}(n) \vdash t_s \uparrow \tau_2}{\Theta; \Gamma \vdash \text{match}(t_n, t_0, n \Rightarrow t_s) \uparrow \text{Match}(t_n, \tau_1, n \Rightarrow \tau_2)} \text{ (Infer Match)} \\
\frac{\Theta; \Gamma \vdash t \uparrow \tau_1 + \tau_2 \quad \Theta; \Gamma, x : \tau_1, p : t \equiv \text{left}(x) \vdash t_1 \uparrow \tau'_1 \quad \Theta; \Gamma, x : \tau_2, p : t \equiv \text{right}(x) \vdash t_2 \uparrow \tau'_2}{\Theta; \Gamma \vdash \text{either_match}(t, x \Rightarrow t_1, x \Rightarrow t_2) \uparrow \text{Either_Match}(t, x \Rightarrow \tau'_1, x \Rightarrow \tau'_2)} \text{ (Infer Either Match)} \\
\frac{\Theta; \Gamma, n : \text{Nat}, y : \text{Unit} \rightarrow \tau, p : y \equiv \lambda u : \text{Unit}. \text{rec}[n \rightarrow \tau](n, t_0, (n, y) \Rightarrow t_s) \vdash t_s \Downarrow \tau[n \mapsto \text{succ}(n)]}{\Theta; \Gamma \vdash \text{rec}[n \Rightarrow \tau](t_n, t_0, (n, y) \Rightarrow t_s) \uparrow \text{Let } n = t_n \text{ in } \tau} \text{ (Infer Rec)} \\
\frac{\Theta; \Gamma, n : \text{Nat}, y : \text{Unit} \rightarrow \forall m : \{m : \text{Nat} \mid m < n\}. \tau[n \mapsto m], \quad p : y \equiv \lambda u : \text{Unit}. \text{fix}[n \Rightarrow \tau]((n, y) \Rightarrow t) \quad \vdash t \Downarrow \tau}{\Theta; \Gamma \vdash \text{fix}[n \Rightarrow \tau]((n, y) \Rightarrow t) \uparrow \forall n : \text{Nat}. \tau} \text{ (Infer Fix)} \\
\frac{\Theta; \Gamma \vdash t_1 \uparrow \tau_1 \quad \Theta; \Gamma \vdash t_2 \uparrow \tau_2}{\Theta; \Gamma \vdash (t_1, t_2) \uparrow \Sigma x : \tau_1. \tau_2} \text{ (Infer Pair)} \quad \frac{\Theta; \Gamma \vdash t_1 \uparrow \tau_1 \quad \Theta; \Gamma, x : \tau_1, p : x \equiv t_1 \vdash t_2 \uparrow \tau_2}{\Theta; \Gamma \vdash \text{let } x = t_1 \text{ in } t_2 \uparrow \text{Let } x = t_1 \text{ in } \tau_2} \text{ (Infer Let)} \\
\frac{\Theta; \Gamma \vdash t \uparrow \Sigma x : \tau_1. \tau_2}{\Theta; \Gamma \vdash \pi_1 t \uparrow \tau_1} \text{ (Infer Proj1)} \quad \frac{\Theta; \Gamma \vdash t \uparrow \Sigma x : \tau_1. \tau_2}{\Theta; \Gamma \vdash \pi_2 t \uparrow \text{Let } x = \pi_1 t \text{ in } \tau_2} \text{ (Infer Proj2)} \\
\frac{\Theta; \Gamma, x : \tau_1 \vdash t \uparrow \tau_2}{\Theta; \Gamma \vdash \lambda x : \tau_1. t \uparrow \Pi x : \tau_1. \tau_2} \text{ (Infer Lambda)} \quad \frac{\Theta; \Gamma \vdash t_1 \uparrow \Pi x : \tau_2. \tau \quad \Theta; \Gamma \vdash t_2 \Downarrow \tau_2}{\Theta; \Gamma \vdash t_1 t_2 \uparrow \text{Let } x = t_2 \text{ in } \tau} \text{ (Infer App)} \\
\frac{\Theta, \alpha; \Gamma \vdash t \uparrow \tau}{\Theta; \Gamma \vdash \Lambda \alpha. t \uparrow \forall \alpha : \text{Type}. \tau} \text{ (Infer Type Abs)} \quad \frac{\Theta; \Gamma \vdash t \uparrow \forall \alpha : \text{Type}. \tau_1}{\Theta; \Gamma \vdash t[\tau_2] \uparrow \tau_1[\alpha \mapsto \tau_2]} \text{ (Infer Type App)} \\
\frac{\Theta; \Gamma \vdash t_1 \uparrow \forall x : \tau_2. \tau \quad \Theta; \Gamma \vdash t_2 \Downarrow \tau_2}{\Theta; \Gamma \vdash \text{inst}(t_1, t_2) \uparrow \text{Let } x = t_2 \text{ in } \tau} \text{ (Infer Forall Instantiation)} \\
\frac{\Theta; \Gamma \vdash n \Downarrow \text{Nat} \quad \Theta; \Gamma, p : n \equiv \text{zero} \vdash t \Downarrow \text{basetype}_\alpha(\tau) \quad \Theta; \Gamma, n' : \text{Nat}, p : n \equiv \text{succ}(n') \vdash t \Downarrow \tau[\alpha \mapsto \text{Rec}(n')(\alpha \Rightarrow \tau)]}{\Theta; \Gamma \vdash \text{fold}[\text{Rec}(n)(\alpha \Rightarrow \tau)](t) \uparrow \text{Rec}(n)(\alpha \Rightarrow \tau)} \text{ (Infer Fold)} \\
\frac{\Theta; \Gamma \vdash t_1 \uparrow \text{Rec}(n)(\alpha \Rightarrow \tau) \quad \Theta; \Gamma, x : \text{basetype}_\alpha(\tau), p_1 : t_1 \equiv \text{fold}(x), p_2 : n \equiv \text{zero} \vdash t_2 \uparrow \tau' \quad \Theta; \Gamma, x : \tau[\alpha \mapsto \text{Rec}(\text{pred}(n))(\alpha \Rightarrow \tau)], p : t_1 \equiv \text{fold}(x) \vdash t_2 \uparrow \tau'}{\Theta; \Gamma \vdash \text{unfold } t_1 \text{ in } x \Rightarrow t_2 \uparrow \tau'} \text{ (Infer Unfold)} \\
\frac{\Theta; \Gamma \vdash t_1 \uparrow \text{Rec}(n)(\alpha \Rightarrow \tau) \quad \boxed{\Theta; \Gamma \vdash \text{lessThan } 0 \ n \equiv \text{true}}}{\Theta; \Gamma, x : \tau[\alpha \mapsto \text{Rec}(\text{pred}(n))(\alpha \Rightarrow \tau)], p : t_1 \equiv \text{fold}(x) \vdash t_2 \uparrow \tau'} \text{ (Infer Unfold Positive)} \\
\frac{\boxed{\Theta; \Gamma \vdash \text{true} \equiv \text{false}}}{\Theta; \Gamma \vdash \text{err}[\tau] \uparrow \tau} \text{ (Infer Err)} \quad \frac{\boxed{\Theta; \Gamma \vdash t_1 \equiv t_2}}{\Theta; \Gamma \vdash \text{refl}[t_1, t_2] \uparrow t_1 \equiv t_2} \text{ (Infer Refl)} \\
\frac{\Theta; \Gamma \vdash t \uparrow \{x : \tau \mid p\}}{\Theta; \Gamma \vdash t \uparrow \tau} \text{ (Infer Drop Refinement)}
\end{array}$$

Fig. 14. $\Theta; \Gamma \vdash t \uparrow \tau$ infers a type τ for t in context $\Theta; \Gamma$ based on the shape of t . The *(Infer Drop Refinement)* rule is applied with low priority, only if no other rule is applicable, keeping type checking deterministic.

$$\begin{array}{c}
\frac{\Theta; \Gamma \vdash t_1 \Downarrow \text{Bool} \quad \Theta; \Gamma, p : t_1 \equiv \text{true} \vdash t_2 \Downarrow \tau \quad \Theta; \Gamma, p : t_1 \equiv \text{false} \vdash t_3 \Downarrow \tau}{\Theta; \Gamma \vdash \text{if } t_1 \text{ then } t_2 \text{ else } t_3 \Downarrow \tau} \text{ (Check If)} \\
\frac{\Theta; \Gamma \vdash t_n \Downarrow \text{Nat} \quad \Theta; \Gamma, p : t_n \equiv \text{zero} \vdash t_0 \Downarrow \tau \quad \Theta; \Gamma, n : \text{Nat}, p : t_n \equiv \text{succ}(n) \vdash t_s \Downarrow \tau}{\Theta; \Gamma \vdash \text{match}(t_n, t_0, n \Rightarrow t_s) \Downarrow \tau} \text{ (Check Match)} \\
\frac{\Theta; \Gamma \vdash t \Uparrow \tau_1 + \tau_2 \quad \Theta; \Gamma, x : \tau_1, p : t \equiv \text{left}(x) \vdash t_1 \Downarrow \tau \quad \Theta; \Gamma, x : \tau_2, p : t \equiv \text{right}(x) \vdash t_2 \Downarrow \tau}{\Theta; \Gamma \vdash \text{either_match}(t, x \Rightarrow t_1, x \Rightarrow t_2) \Downarrow \tau} \text{ (Check Either Match)} \\
\frac{\Theta; \Gamma \vdash t_1 \Uparrow \tau_1 \quad \Theta; \Gamma, x : \tau_1, p : x \equiv t_1 \vdash t_2 \Downarrow \tau}{\Theta; \Gamma \vdash \text{let } x = t_1 \text{ in } t_2 \Downarrow \tau} \text{ (Check Let)} \quad \frac{\Theta; \Gamma, x : \text{Nat} \vdash t[x] \Downarrow \tau}{\Theta; \Gamma \vdash t \Downarrow \forall x : \text{Nat}. \tau} \text{ (Check Forall)} \\
\frac{\Theta; \Gamma, x : \tau_1 \vdash t \Downarrow \tau_2}{\Theta; \Gamma \vdash t \Downarrow \prod x : \tau_1. \tau_2} \text{ (Check Pi)} \quad \frac{\Theta; \Gamma \vdash \pi_1 t \Downarrow \tau_1 \quad \Theta; \Gamma, x : \tau_1, p : x \equiv \pi_1 t \vdash \pi_2 t \Downarrow \tau_2}{\Theta; \Gamma \vdash t \Downarrow \sum x : \tau_1. \tau_2} \text{ (Check Sigma)} \\
\frac{\Theta; \Gamma \vdash t \Downarrow \tau \quad \boxed{\Theta; \Gamma, x : \tau, p : x \equiv t + b \equiv \text{true}}}{\Theta; \Gamma \vdash t \Downarrow \{x : \tau \mid b\}} \text{ (Check Refinement)} \\
\frac{\Theta, \alpha; \Gamma \vdash t[\alpha] \Downarrow \tau}{\Theta; \Gamma \vdash t \Downarrow \forall \alpha : \text{Type}. \tau} \text{ (Check Type Abs)} \quad \frac{\Theta; \Gamma \vdash t \Uparrow \text{Rec}(n_2)(\alpha \Rightarrow \tau) \quad \boxed{\Theta; \Gamma \vdash n_1 \equiv n_2}}{\Theta; \Gamma \vdash t \Downarrow \text{Rec}(n_1)(\alpha \Rightarrow \tau)} \text{ (Check Recursive)} \\
\frac{v \in \text{Val}}{\Theta; \Gamma \vdash v \Downarrow \top} \text{ (Check Top 1)} \quad \frac{\Theta; \Gamma \vdash t \Uparrow \tau}{\Theta; \Gamma \vdash t \Downarrow \top} \text{ (Check Top 2)} \quad \frac{\Theta; \Gamma \vdash t \Uparrow \tau}{\Theta; \Gamma \vdash t \Downarrow \tau} \text{ (Check Reflexive)}
\end{array}$$

Fig. 15. $\Theta; \Gamma \vdash t \Downarrow \tau$ checks that term t indeed has type τ under context $\Theta; \Gamma$. When multiple rules are applicable, they are applied from a priority order from top to bottom, left to right. The *(Check Forall)* rule can be generalized to (non-empty) types other than Nat , but we only need it for Nat (in type $\text{Rec}(\alpha \Rightarrow \tau)$).

$t ::= x \mid () \mid \lambda x : \tau. t \mid t t \mid (t, t) \mid \pi_1 t \mid \pi_2 t \mid \text{inst}(t, n) \mid$
 $\text{left}[\tau + \tau](t) \mid \text{right}[\tau + \tau](t) \mid \text{either_match}(t, x \Rightarrow t, x \Rightarrow t) \mid$
 $\text{true} \mid \text{false} \mid \text{if } t \text{ then } t \text{ else } t \mid$
 $\text{zero} \mid \text{succ}(t) \mid \text{rec}[x \Rightarrow \tau](t, t, (n, y) \Rightarrow t) \mid \text{fix}[n \Rightarrow \tau]((n, y) \Rightarrow t) \mid \text{match}(t, t, n \Rightarrow t) \mid$
 $\text{fold}[\tau](t) \mid \text{unfold } t \text{ in } x \Rightarrow t \mid \text{unfold_pos } t \text{ in } x \Rightarrow t \mid$
 $\Lambda \alpha. t \mid t[\tau] \mid \text{err}[\tau] \mid \text{refl}[t, t] \mid \text{let } x = t \text{ in } t \mid \text{size}(t)$

Fig. 16. Grammar for annotated terms t , where x, y and n are term variables and α is a type variable.

termination arguments relies on a user-defined *measure* function. The `fix` term shown on the right corresponds to the desugaring of the recursive function on the left whose contracts are given by the **require** and **ensuring** keywords, and whose measure is given by the **decreases** keyword. The contract terms *pre* and *post* are such that $x : \tau_1 \vdash \text{pre} : \text{Bool}$ and $x : \tau_1, \text{res} : \tau_2 \vdash \text{post} : \text{Bool}$, and the measure function *measure* satisfies $x : \tau_1 \vdash \text{measure} : \text{Nat}$. The term `pred` is a function of type $\{n : \text{Nat} \mid \text{zero} < n\} \rightarrow \text{Nat}$ that returns the predecessor of numbers greater than zero.

We now explain how our type-checking algorithm ensures termination of such a function. Our type inference rule for `fix` is *(Infer Fix)*. The side condition $n \notin \text{fv}(\text{erase}(t))$ ensures that n only appears in type annotations in t , and is not part of the computation. The other check corresponds to a proof by strong induction (over n) that the `fix` term has type $\forall n : \text{Nat}. \tau$. Indeed, we have to check that t , the body of the `fix` term, has type τ (for some $n : \text{Nat}$), under the assumption that y (which is the variable representing the recursion) has type $\tau[n \mapsto m]$ for all $m < n$. The ‘Unit \rightarrow ’


```

def f(x:  $\tau_1$ ):  $\tau_2$  = {
  require(pre[x])           f  $\equiv$ 
  decreases(measure[x])    fix[n  $\Rightarrow$   $\Pi x : \{x : \tau_1 \mid pre\} \mid measure \leq n\}. \{res : \tau_2 \mid post\}$ ](
  E[x, f]                   (n, f)  $\Rightarrow \lambda x : \{x : \tau_1 \mid pre\} \mid measure \leq n\}.
  } ensuring {              E[x, inst(f, pred(n))]()
  res  $\Rightarrow post[x, res]$  }$ 
```

Fig. 17. Encoding named function with pre- and post-conditions are given by the **require** and **ensuring**, and measure given by the **decreases** keyword (left) into a terminating fixpoint recursion (right).

part of the type of y corresponds to the fact that the operational semantics of `fix` replaces variable y by the `fix` term under a lambda (as explained in Section 3).

The variable p is a witness that the variable y is equal to the `fix` term (under a lambda). This feature is useful for *body-visible* recursion, and is explained in Section 5.3.

Back to the encoding presented in Figure 17, we explain how the (*Infer Fix*) rule ensures that *measure* decreases at each recursive call of function f . Assume that the premise of the (*Infer Fix*) rule holds, and that f is called with some value v of type τ_1 , such that $measure[v]$ evaluates to some (term representing a) natural number n . By instantiating the premise of the (*Infer Fix*) rule for that particular n , we get that $E[x, \text{inst}(f, \text{pred}(n))()]$ is well-typed under the condition that f has type:

$$\forall m : \{m : \text{Nat} \mid m < n\}. \text{Unit} \rightarrow \Pi x : \{x : \tau_1 \mid pre\} \mid measure \leq m\}. \{res : \tau_2 \mid post\}$$

First, in order for `pred` to be applied to n , we have to check that n is non-zero, meaning that the measure of v is strictly positive in the places where the recursive calls happen. This is ensured by the (*Check Refinement*) rule for checking refinement types (see Figure 15), which generates a verification condition.

Second, the rule (*Infer Forall Instantiation*) ensures that $\text{inst}(f, \text{pred}(n))()$ takes arguments of type $\{x : \{x : \tau_1 \mid pre\} \mid measure \leq \text{pred}(n)\}$. Therefore, if f is applied recursively to an argument v' , the rule (*Check Refinement*) ensures that $measure[v'] \leq \text{pred}(n)$ holds. Overall, we get $measure[v'] \leq \text{pred}(n) < n = measure[v]$, which ensures that the measures of arguments always decrease on recursive calls to f .

In our implementation, we do not go through the encoding with `fix` and `forall` types, but instead directly generate the verification conditions that correspond to the measure decreasing by using the left-hand-side form of Figure 17. Our system also supports mutually recursive functions (by requiring that the measure decreases for each call to a mutually recursive function), which can be encoded in the usual way by defining a `fix` term that returns a tuple of functions.

In the end, if the body of the function is well-typed, the (*Infer Fix*) rule infers the type:

$$\forall n : \text{Nat}. \Pi x : \{x : \{x : \tau_1 \mid pre\} \mid m \leq n\}. \{res : \tau_2 \mid post\}$$

One should note that this encoding imposes a scoping restriction on the original program, namely precondition, postcondition, and measure of a function f cannot contain calls to f . This restriction has not proved limiting in our experience with benchmarks.

5.2.1 Lexicographic Orderings. Functions whose termination arguments require lexicographic orderings can be encoded by using two levels of recursions, which is a known technique that shows expressive power of System T [Girard 1990, Section 7.3.2]. We review how this encoding works in our system in the long version of our paper [Hamza et al. 2019] and show an example of Ackerman’s function and its simple lexicographic measure. In our implementation, we support lexicographic measures directly.

5.3 Body-Visible Recursion

In this section, we give more details about the (*Infer Fix*) and (*Infer Rec*) typing rules for recursion. They allow *body-visible recursion* which gives the type-checker access to the definition of a recursive function while type-checking the body of the recursive function itself.

The first thing to note is that we introduce an equality type containing the definition (in the type of p) in the context, while we do not know yet whether the body of the recursion is well-typed. Since our equality type is defined (in Section 4) for all terms, regardless of whether they are well-typed, this is perfectly legal. We show in the merge example of Figure 5 how body-visible recursion relieves the user from writing excessive specification annotations.

Assume we want to prove, on paper, that merge indeed returns a sorted list when given two sorted lists $l1$ and $l2$, by induction over $\text{size}(l1) + \text{size}(l2)$. Consider the first branch of the `if then else` statement, where we return `Cons(x, merge(xs, l2))`. By the induction hypothesis, we know that the recursive call `merge(xs, l2)` is sorted, but this mere fact is not enough to conclude that `Cons(x, merge(xs, l2))` is sorted. By unfolding the definition of `isSorted`, we see that we need in addition to know that x is smaller than the head of the result `merge(xs, l2)`.

Therefore, the property we prove by induction needs to be strengthened by saying that the head of the result, if non-empty, is equal to the (smallest) head of one the input lists. From that, we will know by the induction hypothesis that the head h of `merge(xs, l2)` (if non-empty) is either the head of xs or the head of $l2$. In the first case, we can deduce that x is smaller than h by using the fact that `l1 = Cons(x, xs)` is sorted. In the second case, we have $h = y$, and we know from the condition of the `if then else` statement that $x \leq y$. In both cases, we can conclude that the whole list `Cons(x, merge(xs, l2))` is sorted.

If we are to type-check the program above, and if we only know the return type of `merge(xs, l2)`, that is $\{ l : \text{List} \mid \text{isSorted}(l) \}$, we will run into the same problem, and will not be able to conclude that `Cons(x, merge(xs, l2))` is sorted. In our type system, we get in addition access to the definition of `merge` while type-checking it, thanks to the p variable of equality type in the (*Infer Fix*) rule. By unfolding the definition of `merge(xs, l2)`, we conclude by case analysis that the head h of `merge(xs, l2)` (if non-empty) is either the head of xs or the head of $l2$ (which is y).

Without body-visible recursion, the developer would need to strengthen the postcondition to:

$$\begin{aligned} \text{isEmpty}(res) \parallel (!\text{isEmpty}(l1) \ \&\& \ \text{head}(res) == \text{head}(l1)) \\ \parallel (!\text{isEmpty}(l2) \ \&\& \ \text{head}(res) == \text{head}(l2)) \end{aligned}$$

In `Inox`, the external solver we use for verification conditions, definitions of recursive functions are unfolded automatically. `Inox` does incremental queries to SMT solvers. It first sends a query without unfolding at all, then a new query after unfolding once, and so on until a query succeeds or a timeout. Thanks to this approach, `Inox` does not rely on universal quantifiers to encode recursive functions. This feature is crucial to have such examples be verified without user intervention, and is here required to get the bodies of the calls to `merge` and `isSorted` when verifying the postcondition of `merge`.

5.4 Unification in Type Inference

In order to perform type inference, we expect a certain structure on the inferred types. For example, given term $t_1 t_2$, we expect term t_1 to have a function type when inferring the type of an application. Furthermore, type inference must perform least upper bound computation for `if`, `match` and `either_match` terms, which adds more complexity to the system. We handle these considerations in a generalized manner by performing *type unifications*, defined in Figure 18 by using the type-level notations `If Then Else`, `Let`, `Either_Match`, and `Match`.

$$\begin{aligned}
\text{unify}(E, \text{Unit}, \text{Unit}) &\triangleq \text{Unit} \\
\text{unify}(E, \text{Bool}, \text{Bool}) &\triangleq \text{Bool} \\
\text{unify}(E, \text{Nat}, \text{Nat}) &\triangleq \text{Nat} \\
\text{unify}(E, \top, \top) &\triangleq \top \\
\text{unify}(E, \alpha, \alpha) &\triangleq \alpha \\
\text{unify}(E, \Pi x : A_1. B_1, \Pi x : A_2. B_2) &\triangleq \Pi x : \text{unify}(E, A_1, A_2). \text{unify}(E, B_1, B_2) \\
\text{unify}(E, \forall x : A_1. B_1, \forall x : A_2. B_2) &\triangleq \forall x : \text{unify}(E, A_1, A_2). \text{unify}(E, B_1, B_2) \\
\text{unify}(E, \Sigma x : A_1. B_1, \Sigma x : A_2. B_2) &\triangleq \Sigma x : \text{unify}(E, A_1, A_2). \text{unify}(E, B_1, B_2) \\
\text{unify}(E, \text{Rec}(n_1)(\alpha \Rightarrow \tau_1), \text{Rec}(n_2)(\alpha \Rightarrow \tau_2)) &\triangleq \text{Rec}(E[n_1, n_2])(\alpha \Rightarrow \text{unify}(E, \tau_1, \tau_2)) \\
\text{unify}(E, \{x : A_1 \mid p_1\}, \{x : A_2 \mid p_2\}) &\triangleq \{x : \text{unify}(E, A_1, A_2) \mid E[p_1, p_2]\} \\
\text{unify}(E, t_{1,1} \equiv t_{1,2}, t_{2,1} \equiv t_{2,2}) &\triangleq E[t_{1,1}, t_{2,1}] \equiv E[t_{1,2}, t_{2,2}]
\end{aligned}$$

$$\text{If } t_1 \text{ Then } \tau_2 \text{ Else } \tau_3 \triangleq \text{unify}(\text{if } t_1 \text{ then } _ \text{ else } _, \tau_2, \tau_3)$$

$$\text{Let } x = t \text{ in } \tau \triangleq \text{unify}(\text{let } x = t \text{ in } _, \tau, \tau)$$

$$\text{Either_Match}(t, x \Rightarrow \tau_1, x \Rightarrow \tau_2) \triangleq \text{unify}(\text{either_match}(t, x \Rightarrow _, x \Rightarrow _), \tau_1, \tau_2)$$

$$\text{Match}(t, \tau_1, n \Rightarrow \tau_2) \triangleq \text{unify}(\text{match}(t, _, n \Rightarrow _), \tau_1, \tau_2)$$

Fig. 18. Recursion schema for unification of two types, where E is a context with two holes.

$$\frac{\Theta; \Gamma \vdash t_1 \uparrow \uparrow \text{Rec}(\alpha \Rightarrow \tau) \quad \Theta; \Gamma, x : \text{Rec}(\alpha \Rightarrow \tau), p : t_1 \equiv \text{fold}(x) \vdash t_2 \uparrow \uparrow \tau' \quad \text{spos}(\alpha, \tau)}{\Theta; \Gamma \vdash \text{unfold } t_1 \text{ in } x \Rightarrow t_2 \uparrow \uparrow \tau'} \text{ (Infer Unfold Gen)}$$

$$\frac{\Gamma \vdash t \Downarrow \tau[\alpha \mapsto \text{Rec}(\alpha \Rightarrow \tau)] \quad \text{spos}(\alpha, \tau)}{\Gamma \vdash \text{fold}[\text{Rec}(\alpha \Rightarrow \tau)](t) \uparrow \uparrow \text{Rec}(\alpha \Rightarrow \tau)} \text{ (Infer Fold Gen)}$$

Fig. 19. Folding and unfolding without indices for strictly positive recursive types.

We draw attention here to the fact that our type checking and inference procedures are *syntax directed* and *predictable*, which enables a natural verification process through term-level hints. The algorithmic nature of our type checking precludes the verification of certain well-typed programs. However, our experience has shown that this limitation is largely inconsequential in practice and is outweighed by the predictable nature of the algorithm.

5.5 Hiding Recursive Type Indices

Beyond general rules of Figure 14, in Figure 19 we show two additional rules for ‘fold’ and ‘unfold in’ that ignore the indices hidden under the Rec type for strictly positive recursive types (which is the case for the List and Stream types). We write $\text{spos}(\alpha, \tau)$ when a type variable appears only strictly positively in type τ , meaning only to the right of Π and \forall types (see the long version of our paper [Hamza et al. 2019] for the precise definition). This enables us, under some conditions, to fold and unfold a strictly positive recursive type without worrying about indices.

Practically, given an element l of type $\text{List}[X]$ (resp. $\text{Stream}[X]$), we can unfold it (with the *Infer Unfold Gen* rule) to get its head and its tail of type $\text{List}[X]$ (resp. $\text{Unit} \rightarrow \text{Stream}[X]$). Conversely, we use the rule *Infer Fold Gen* to build a list or a stream from an element and a tail.

Strict positivity gives us the following key lemma that ensures the soundness of our rules with respect to our reducibility definition. This lemma states that when a type variable α appears only strictly positively in τ , then quantifying with a forall type outside τ or inside a substitution for α is

the same (as long as we are quantifying over a non-empty type τ_1). This property is similar to the notions *lim sup-pushable* and *lim inf-pullable* [Abel 2008] and implies the soundness of our rules.

LEMMA 5.2. *Let τ and $\forall x : \tau_1. \tau_2$ be two types. Let α be a type variable that appears strictly positively in τ . Let θ be a type interpretation such that $\llbracket \tau_1 \rrbracket_v^\theta$ is not empty. We have:*

$$\llbracket \tau[\alpha \mapsto \forall x : \tau_1. \tau_2] \rrbracket_v^\theta = \llbracket \forall x : \tau_1. \tau[\alpha \mapsto \tau_2] \rrbracket_v^\theta$$

5.6 Type-Checking Algorithm Examples: Streams

5.6.1 *Constant Stream.* The `fix` term and associated typing rules can also be used to express the kind of recursion used to define the streams in Section 2. We start by revisiting the constant stream, which in our notations can be written as an untyped term:

$$\text{constant} \triangleq \text{fix}(\text{constant} \Rightarrow \Lambda. \lambda x. \text{fold}(x, \lambda u. \text{constant}())[])(x))$$

Assume we want to prove, on paper, that for any value x , `constant[]`(x) produces a non-diverging stream, i.e. a stream which is n -non-diverging for every $n \in \mathbb{N}$. A natural proof could be done by induction on n , as follows:

- ($n = 0$) `constant[]`(x) is 0-non-diverging, meaning that it reduces to a value of the form `fold`(x, v) where x and v are values. This is clear from the code of `constant`, as this expression evaluates in a few steps to `fold`($x, \lambda u. (\lambda u'. \text{constant}())[])(x)$.
- ($n = n' + 1$) Assume by induction that `constant[]`(x) is n' -non-diverging. By definition of n -non-diverging, we get that `fold`($x, \lambda u. \text{Unit}. \text{constant}[])(x)$ is n -non-diverging. Since this term is equivalent to the term to which `constant[]`(x) evaluates, we conclude that `constant[]`(x) is n -non-diverging as well.

Our type system and type-checking algorithm can be used to simulate this proof by using an annotated version of `constant`:

$$\text{constant} \triangleq \text{fix}[n \Rightarrow \forall X : \text{Type}. X \rightarrow \text{Stream}_n[X]]((n, \text{constant}) \Rightarrow \text{body}(n, \text{constant}))$$

where `body`($n, \text{constant}$) is a shorthand for

$$\Lambda X. \lambda x : X. \text{fold}[\text{Stream}_n[X]](x, \lambda u : \text{Unit}. \text{inst}(\text{constant}, \text{pred}(n))() [X])(x))$$

By applying the (*Infer Fix*) rule presented above, we get the type

$$\vdash \text{constant} \uparrow \forall n : \text{Nat}. \forall X : \text{Type}. X \rightarrow \text{Stream}_n[X].$$

The (*Infer Fix*) rule of our algorithm generates a check that corresponds to a (strong) induction that shows that for every $n \in \llbracket \text{Nat} \rrbracket_v$, `constant[X]`(x) in $\llbracket \text{Stream}_n[X] \rrbracket_v$, assuming that it is in $\llbracket \text{Stream}_m[X] \rrbracket_v$ for all $m < n$:

$$n : \text{Nat}, \text{constant} : \forall m : \{m : \text{Nat} \mid m < n\}. \text{Unit} \rightarrow \forall X : \text{Type}. X \rightarrow \text{Stream}_n[X] \vdash \\ \text{body}(n, \text{constant}) \Downarrow \forall X : \text{Type}. X \rightarrow \text{Stream}_n[X]$$

After applying standard rules related to λ and Λ , our algorithm will attempt to infer, using the (*Infer Fold*) rule, a type for the term:

$$\text{fold}[\text{Stream}_n[X]](x, \lambda u : \text{Unit}. \text{inst}(\text{constant}, \text{pred}(n))() [X])(x))$$

In addition to the type-check that n has type `Nat`, this rule generates two checks to cover the cases where n is zero or non-zero. These correspond to the informal proof by induction given above for the non-divergence of `constant`. The first check reduces (after applying some straightforward rules) to checking that $\lambda u : \text{Unit}. \text{inst}(\text{constant}, \text{pred}(n))() [X](x)$ has type \top (remember that $\text{basetype}_\alpha(X \times (\text{Unit} \rightarrow \alpha)) = X \times \top$), which goes through easily thanks to the rule (*Check Top 1*).

```

zipWith  $\triangleq$  fix[n  $\Rightarrow$   $\forall X, Y, Z : \text{Type}. (X \rightarrow Y \rightarrow Z) \rightarrow \text{Stream}_n[X] \rightarrow \text{Stream}_n[Y] \rightarrow \text{Stream}_n[Z]$ ](
  (n, zipWith)  $\Rightarrow$ 
     $\Lambda X, Y, Z. \lambda f : X \rightarrow Y \rightarrow Z. \lambda s1 : \text{Stream}_n[X]. \lambda s2 : \text{Stream}_n[Y].$ 
      fold[Streamn[Z]](
        unfold s1 in xs1  $\Rightarrow$ 
        unfold s2 in xs2  $\Rightarrow$ 
          f ( $\pi_1$  xs1) ( $\pi_1$  xs2),
         $\lambda u : \text{Unit}.$ 
          unfold_pos s1 in xs1  $\Rightarrow$ 
          unfold_pos s2 in xs2  $\Rightarrow$ 
            inst(zipWith, pred(n))([X][Y][Z] (( $\pi_2$  xs1) ()) (( $\pi_2$  xs2) ())))

```

Fig. 20. An annotated term of our calculus to define zipWith.

The second check amounts to checking that $\text{inst}(\text{constant}, \text{pred}(n))()([X](x))$ has type $\text{Stream}_{n'}[X]$ under the assumption that $n \equiv \text{succ}(n')$. By the context we know that $\text{inst}(\text{constant}, \text{pred}(n))()([X](x))$ has type $\text{Stream}_{\text{pred}(n)}[X]$. Since n' and $\text{pred}(n)$ are equivalent, we can use the rule (*Check Recursive*) (given in Figure 15) to convert between the two types.

Attempting to type-check the badConstant example from Section 2:

```
fix[n  $\Rightarrow$   $\forall X : \text{Type}. X \rightarrow \text{Stream}_n[X]$ ](n, badConstant)  $\Rightarrow$  badBody(n, badConstant)
```

where badBody(n, badConstant) stands for:

```
 $\Lambda X. \lambda x : X. \text{fold}[\text{Stream}_n[X]](x, \lambda u : \text{Unit}. \text{inst}(\text{badConstant}, \text{pred}(n))()([X](x)).\text{tail}())$ 
```

will lead to an error in the second check (corresponding to the inductive case), since the extra call to tail() decreases the index of the stream by one.

5.6.2 ZipWith Function on Streams. We now revisit the zipWith function from Figure 7. We said in Section 2 that for every $n \in \mathbb{N}$, when s1 and s2 are n -non-diverging streams, then so is zipWith(f, s1, s2). On paper, we can check by induction over $n \in \mathbb{N}$ that when s1 and s2 are n -non-diverging streams and f is terminating, then zipWith(f, s1, s2) (as written in Figure 7, and ignoring type annotations and the erasable parameter for the moment) is also an n -non-diverging stream.

- ($n = 0$) It is indeed the case that we can access the head of zipWith(f, s1, s2) as long as we can access the heads s1 and s2 (and as long as f terminates).
- ($n = n' + 1$) Let s1 and s2 be two n -non-diverging streams. By definition of non-diverging, we know that s1.tail() and s2.tail() are n' -non-diverging. By induction hypothesis, zipWith(f, s1.tail(), s2.tail()) is n' -non-diverging as well. This means that zipWith(f, s1, s2) is n -non-diverging, which concludes the proof.

Accordingly, our type-checking algorithm is able to infer the type

```
 $\forall n : \text{Nat}. \forall X, Y, Z : \text{Type}. (X \rightarrow Y \rightarrow Z) \rightarrow \text{Stream}_n[X] \rightarrow \text{Stream}_n[Y] \rightarrow \text{Stream}_n[Z].$ 
```

for the annotated zipWith term given in Figure 20. We use the term ‘unfold in’ to access the heads (with π_1). For the tails, we use instead ‘unfold_pos in’. To understand the difference, we must first look at the (*Infer Fold*) rule for the fold term, which will generate two subgoals, one with $n = 0$ and one with $n > 0$. In the subgoal with $n = 0$, we must check that the lambda term ‘ $\lambda u : \text{Unit}.$ [...]’ has type \top , which goes through directly thanks to (*Check Top 1*). Therefore, when

$$\begin{aligned}
\text{fib} &\triangleq \text{fix}[n \Rightarrow \text{Stream}_n[\text{Nat}]][(n, \text{fib}) \Rightarrow \\
&\quad \text{fold}[\text{Stream}_n[\text{Nat}]](0, \lambda u : \text{Unit}. \\
&\quad \quad \text{fold}[\text{Stream}_{\text{pred}(n)}[\text{Nat}]](1, \lambda u : \text{Unit}. \\
&\quad \quad \quad \text{unfold_pos inst}(\text{fib}, \text{pred}(n))() \text{ in } \text{xfib} \Rightarrow \\
&\quad \quad \quad \text{inst}(\text{zipWith}, \text{pred}(\text{pred}(n)))[\text{Nat}][\text{Nat}][\text{Nat}] \\
&\quad \quad \quad \text{plus (inst}(\text{fib}, \text{pred}(\text{pred}(n)))() ((\pi_2 \text{xfib}) ())))))
\end{aligned}$$

Fig. 21. An annotated term of our calculus to define the Fibonacci stream.

$$\begin{aligned}
\text{zipWith_fst} &\equiv \lambda n : \text{Nat}. \text{rec}[n \Rightarrow \tau](n, \\
&\quad \lambda X. \lambda s1 : \text{Stream}[X]. \lambda s2 : \text{Stream}[X]. \\
&\quad \quad \text{refl}[\text{nth } 0 \text{ s1}, \text{nth } 0 (\text{zipWith}(\lambda x. \lambda y. x, s1, s2))], \\
&\quad \quad (n, \text{zipWith_fst}) \Rightarrow \\
&\quad \lambda X. \lambda s1 : \text{Stream}[X]. \lambda s2 : \text{Stream}[X]. \\
&\quad \quad \text{unfold } s1 \text{ in } \text{xs1} \Rightarrow \\
&\quad \quad \text{unfold } s2 \text{ in } \text{xs2} \Rightarrow \\
&\quad \quad \text{let pr} : \text{nth } n (\pi_2 \text{xs1}) \equiv \text{nth } n (\text{zipWith}(\lambda x. \lambda y. x, \pi_2 \text{xs1}, \pi_2 \text{xs2})) = \\
&\quad \quad \quad \text{zipWith_fst}()[X] (\pi_2 \text{xs1}) (\pi_2 \text{xs2}) \text{ in} \\
&\quad \quad \text{refl}[\text{nth} (\text{succ}(n)) \text{ s1}, \text{nth} (\text{succ}(n)) (\text{zipWith}(\lambda x. \lambda y. x, s1, s2))]
\end{aligned}$$
Fig. 22. An annotated term of our calculus to prove that zipping two streams with $\lambda x. \lambda y. x$ returns a stream equivalent to the first.

we are type-checking the body of the lambda, we know that $n > 0$. We can thus access the tails using (*Infer Unfold Positive*), a variant of (*Infer Unfold*) that discards a subgoal with $n = 0$ but requires proving $n > 0$ instead.

5.6.3 Fibonacci Stream. We now consider the Fibonacci stream function from Figure 8. We can prove by induction that for all $n \in \mathbb{N}$, `fib` is an n -non-diverging stream. In our view, this corresponds to writing `fib` as in Figure 21.

When type-checking `fib`, we use the type of `zipWith` and instantiate it to `pred(pred(n))`, and get that: `inst(zipWith, pred(pred(n)))` has type:

$$\forall X, Y, Z : \text{Type}. (X \rightarrow Y \rightarrow Z) \rightarrow \text{Stream}_{\text{pred}(\text{pred}(n))}[X] \rightarrow \text{Stream}_{\text{pred}(\text{pred}(n))}[Y] \rightarrow \text{Stream}_{\text{pred}(\text{pred}(n))}[Z]$$

This is where we use the fact that `zipWith` returns an n -non-diverging stream when given n -non-diverging streams.

5.7 Verification of Properties

Finally, we show how to verify the property mentioned in Section 2: that zipping two streams `s1` and `s2` with the function $\lambda x. \lambda y. x$ returns a stream equivalent to `s1`.

We define a function `zipWith_fst` with the following type:

$$\Pi n : \text{Nat}. \forall X : \text{Type}. \Pi s1 : \text{Stream}[X]. \Pi s2 : \text{Stream}[X]. \text{nth } n \text{ s1} \equiv \text{nth } n (\text{zipWith}(\lambda x. \lambda y. x, s1, s2))$$

For short, we denote this type as $\Pi n : \text{Nat}. \tau$. The code of `zipWith_fst` is given in Figure 22 and makes use of the recursor `rec`. Our (*Infer Rec*) rule for inferring the type of `rec` is similar to `fix`. The difference is that `rec` allows references to n in the code, so the computation is allowed to

$$\begin{aligned}
\llbracket \perp \rrbracket_v^\theta &\triangleq \{ \} & \llbracket \{t\} \rrbracket_v^\theta &\triangleq \{v \in \text{Val} \mid t \hookrightarrow^* v\} \\
\llbracket \tau_1 \cap \tau_2 \rrbracket_v^\theta &\triangleq \llbracket \tau_1 \rrbracket_v^\theta \cap \llbracket \tau_2 \rrbracket_v^\theta & \llbracket \tau_1 \cup \tau_2 \rrbracket_v^\theta &\triangleq \llbracket \tau_1 \rrbracket_v^\theta \cup \llbracket \tau_2 \rrbracket_v^\theta \\
\llbracket \exists x : \tau_1. \tau_2 \rrbracket_v^\theta &\triangleq \{b \mid \exists a \in \llbracket \tau_1 \rrbracket_v^\theta. b \in \llbracket \tau_2[x \mapsto a] \rrbracket_v^\theta\} & & \text{(existential)} \\
\llbracket \{x : \tau_1 \mid \tau_2\} \rrbracket_v^\theta &\triangleq \{a \in \llbracket \tau_1 \rrbracket_v^\theta \mid \exists b \in \llbracket \tau_2[x \mapsto a] \rrbracket_v^\theta\} & & \text{(refinement by type)} \\
\llbracket \text{Let } x = t \text{ in } \tau \rrbracket_v^\theta &\triangleq \{v \mid \exists a \in \text{Val}. t \hookrightarrow^* a \wedge v \in \llbracket \tau[x \mapsto a] \rrbracket_v^\theta\}
\end{aligned}$$

Fig. 23. Definition of reducibility for intersection, union, singleton types, and refinement by types.

depend on n . Moreover, `rec` does not use \forall types. Finally, `rec` is analogous to a simple induction, while `fix` is analogous to a strong induction.

In the `zipWith_fst` example, applying this rule corresponds to a proof by induction over n , that for any two streams s_1 and s_2 , the n th element of s_1 is equal to the n th element of `zipWith`($\lambda x. \lambda y. x, s_1, s_2$).

In the base case, to type-check the term `refl[nth 0 s1, nth 0 (zipWith($\lambda x. \lambda y. x, s_1, s_2$))]`, we must make sure that the following equality holds: $\text{nth } 0 \ s_1 \equiv \text{nth } 0 \ (\text{zipWith}(\lambda x. \lambda y. x, s_1, s_2))$.

The corresponding type inference rule is (*Infer Refl*), which generates a verification condition.

In the inductive case, we explicitly instantiate our inductive hypothesis on the tails of s_1 and s_2 by using the `let` binding on variable `pr`. The equality given by the type of `pr` is then sufficient to prove what we wanted: $\text{nth}(\text{succ}(n)) \ s_1 \equiv \text{nth}(\text{succ}(n)) \ (\text{zipWith}(\lambda x. \lambda y. x, s_1, s_2))$.

This example also illustrates that developers can express the desired versions of what others might call extensional equality (here: same results when calling `nth`) using Π types and equi-reducibility \equiv ; our type system does not impose any preferred form of extensional equality.

6 FORMALIZATION IN THE COQ PROOF ASSISTANT

We here give more details about our formalization of Theorem 5.1 in Coq 8.9.1 (including the rules from Section 5.5). Our proofs are available from

<https://github.com/epfl-lara/SystemFR/tree/oopsla2019>

We represent terms and types using a *locally nameless representation* [Charguéraud 2012], where free variables are named, and where local variables are bound using De Bruijn indices. Using this representation, lambdas and other binders can be seen as terms with holes, which can be filled with other terms (typically values). We use the Coq Equations library [Sozeau 2010] to define the reducibility logical relation $\llbracket \cdot \rrbracket_v$. This library facilitates the use of functions which are defined recursively based on a well-founded measure.

We give an overview of our files (containing around 20k lines of code):

- `Trees.v` contains the definitions of types and terms,
- `Typing.v` gives all typing rules (containing rules from the paper and more),
- `SmallStep.v` contains the operational semantics of the language,
- `ReducibilityDefinition.v` contains the definition of reducibility,
- The `Reducibility*.v` files contain lemmas for the soundness of the rules from `Typing.v`,
- `Reducibility.v` contains the proof that all typing rules from `Typing.v` are sound with respect to the reducibility definition (which implies that Theorem 5.1 holds).

6.1 Extensions of Formalization

Our approach to formalization proved flexible and allowed us to extend the system with additional types. In addition to formalizing the types and rules for our type checking algorithm (figures 15 and 15), our Coq formalization also defines (Figure 23) reducibility for the bottom type, singleton types,

$t :: \tau$	\triangleq	$\exists x : \tau. x \equiv t$	(judgement-as-type)
τ^c	\triangleq	$\{\{x : \top \mid (x :: \tau) \rightarrow \perp\}\}$	(set complement)
$f[\tau]$	\triangleq	$\{\{y : \top \mid \exists x : \tau. f x \equiv y\}\}$	(type image)
$f^{-1}[\tau]$	\triangleq	$\{\{x : \top \mid \exists y : \tau. f x \equiv y\}\}$	(type pre-image)
if b then τ_1 else τ_2	\triangleq	$\{x : \tau_1 \mid b\} \cup \{x : \tau_2 \mid \text{not}(b)\}$	
ifNonempty τ then τ_1 else τ_2	\triangleq	$\{\{x : \tau_1 \mid \tau\}\} \cup \{\{x : \tau_2 \mid \tau \rightarrow \perp\}\}$	
match($t, \tau_1, x \Rightarrow \tau_2$)	\triangleq	$\{\{x : \tau_1 \mid t \equiv \text{zero}\}\} \cup \text{Let } x = \text{pred}(t) \text{ in } \tau_2$	
either_match($t, x \Rightarrow \tau_1, x \Rightarrow \tau_2$)	\triangleq	$(\text{Let } x = \text{unfold_left}(t) \text{ in } \tau_1) \cup$ $(\text{Let } x = \text{unfold_right}(t) \text{ in } \tau_2)$	
where		$\text{not}(b)$	\triangleq if b then false else true
		unfold_left(t)	\triangleq either_match($t, x \Rightarrow x, \text{err}$)
		unfold_right(t)	\triangleq either_match($t, \text{err}, x \Rightarrow x$)

Fig. 24. Syntactic sugar for further expressive types

union and intersection types, and an existential type more abstract than dependent pair (and dual to $\forall x : \tau_1. \tau_2$). We also formalized *refinement by type* $\{\{x : \tau_1 \mid \tau_2\}\}$, which generalizes refinements from computable terminating terms of our language to the condition of non-emptiness of arbitrary types in our system. Existential type and refinement by type can be viewed as second and first projections of dependent pairs. Together with existential and universal types, refinement by type allows us to refine types by quantified propositions, providing support for quantified preconditions and postconditions. Some of the type forms we present can be expressed using others; the set of type forms is not minimal.

We define further syntactic sugar, including precise if then else and match types (Figure 24), as an alternative to the If t_1 Then t_2 Else t_3 that relies on simplifications of Figure 18. We rely on Let type of Figure 23, and we use the ill-typed `err` term in the computations of (respectively) `pred`, `unfold_left` and `unfold_right`, to make sure that the denotations of the Let types are empty when t does not reduce (respectively) to a strictly positive natural number, `left(t')`, or `right(t')` for some t' . We proved soundness of typing rules that show that these types behave as expected. For instance, an if then else expression can be assigned an if then else type of the branches.

These types can be used to encode type level computation reminiscent to one present in Scala `HList`³, as illustrated by the following example. Given a list $L : \text{List}[Bool]$ and types τ_1 and τ_2 , we define a type `TypeFromList(L, τ_1, τ_2)` that represents nested pairs from τ_1 and τ_2 according to the boolean values (false for τ_1 and true for τ_2) and with nesting given by the length of L . For example, when L is the list `true, false, true` then `TypeFromList($L, \text{Nat}, \text{Nat} \rightarrow \text{Nat}$)` represents nested tuples w of the form $(a_0, (a_1, (a_2, \dots)))$ where $a_0 : \text{Nat}$, $a_1 : \text{Nat} \rightarrow \text{Nat}$, $a_2 : \text{Nat}$.

$$\text{TypeFromList}(L, \tau_1, \tau_2) \triangleq \{\{w : \top \mid \forall n : \{n : \text{Nat} \mid n < \text{size}(L)\}. \text{get}(w, n) :: (\text{if } \text{nth}(L, n) \text{ then } \tau_1 \text{ else } \tau_2)\}\}$$

The function `nth(L, n)` returns the n th element of the list L (defined as a recursive data type). The function `get(w, n)` is also a term-level function that operates on nested pairs w such that e.g. `get($w, 2$)` expands to $\pi_1(\pi_2(\pi_2(w)))$. The type $x \equiv \text{get}(w, n)$ in the expansion of refinement by judgement-as-type `_ :: _` (Figure 24) does not require type checking the `get` function; its semantics (see \approx in Figure 13) is that `get(w, n)` reduces to x in the (untyped) operational semantics.

Furthermore, we prove soundness of multiple rules (such as congruence rules) for establishing the equality judgment $\Theta; \Gamma \vdash t_1 \equiv t_2$ described in Section 5. We also prove soundness of a rule that unfolds the definition of a recursive function in the context, which is what is required by the solver in Section 5.3 to unfold the definition of `merge`. Such rules are a step towards justifying not only the verification condition generation but also verification condition solving.

³<https://github.com/milessabin/shapeless/blob/master/core/src/main/scala/shapeless/hlists.scala>

7 IMPLEMENTATION AND EVALUATION

We have implemented our bidirectional type checking procedure by writing an alternative verification-condition generator for Stainless⁴ [LARA 2019]. The code is merged into master and available in e.g. [release version 0.4.0](#). Its functionality can be invoked with the `--type-checker` command line option in both `scalac 2.12` and Dotty front end pipeline for Stainless. Thanks to these frontends, which provide a form of type inference, the type annotation burden is lessened for the user. Precise types such as indexed recursive types (which are not supported out of the box by these frontends) still need to be annotated manually.

The implementation was evaluated on benchmarks shown in Figure 25 totalling 14k LoC, collected from existing Stainless test suites and case studies. The proof/code ratio depends on the properties being proven. It can be 0 or close to 0 as in the merge example of Section 2, or higher than 1 when writing lemmas for proving detailed specifications (as in the [Huffman coding example](#)). The benchmarks reside in the `frontends/benchmarks/typechecker/valid` directory of Stainless. The streams benchmark relies on more expressive annotations of the Dotty compiler that we use to write recursive types, and is available in `frontends/benchmarks/dotty-specific/typechecker`. We use the following syntax to represent indexed types:

	notation in this paper	encoding in Dotty front-end
type	$\text{Stream}_n[X]$	<code>Stream[X] @indexedAt(n)</code>
constructor	<code>fold[Stream_n[X]]((hd, tl))</code>	<code>indexedAt(n, Stream[X](hd,tl))</code>

The suite of all benchmarks verifies in ~6 minutes in total when using implementation based on our type checker. The table shows the number of lines of codes within each benchmark, and the number of verification conditions which were generated by the type checker and verified by the SMT-backed Inox solver. The time given includes the time for generating the verification conditions and solving them, but not the parsing and compilation which is done by the Scala compiler `scalac` (which we use for some initial type inference and to obtain a tree representation of the program), nor the transformations which are internal to Stainless and which happen before using the type-checker. To understand the impact of these time measurements, note that 11% of time of total verification time is spent in the bidirectional type-checking algorithm that generates verification conditions and that is the focus of this paper, 66% in checking the verification conditions using a SMT-backed solver Inox for recursive and higher-order functions, 13% for the parsing, name resolution and most of the type checking pipeline of Scala's `scalac` compiler, and 10% for extraction of `scalac` abstract syntax trees to Stainless trees. All times were measured on a Lenovo X1 Carbon laptop with an Intel i7-7500U and 16GB RAM.

To illustrate the diversity of benchmarks, we note that `InsertionSort`, `QuickSorts`, `MergeSorts` and `StableSort` feature various implementations of the sorting algorithms with the typical properties shown. `ListMonad` and `OptionMonad` show that the monadic laws hold for the `List[X]` and `Option[X]` types. The `List` and `ListWithSize` benchmarks feature a collection of common higher-order functions on lists such as `map`, `filter`, `forall`, etc., as well as many properties shown about the implementations. In the `Gode1Numbering` benchmark, we prove that the pairing function $2^x(2y+1)-1$ is a bijection between natural numbers and pairs of natural numbers using a series of lemmas about linear and non-linear arithmetic. Similar non-linearity is featured in the `MoreExtendedEuclidGCD` benchmark where we show that an implementation of the extended Euclid's algorithm indeed computes the greatest common divisor and the coefficients of Bézout's identity. Benchmarks also include persistent data structures [Okasaki 1998], `ConcTree` and `ConcRope` [Prokopec and Odersky 2015], explicit state model checker `ReachabilityChecker`, small interpreters, dynamic programming

⁴<https://github.com/epfl-lara/stainless>

Name	LoC	VCs	Time (s)	Name	LoC	VCs	Time	Name	LoC	VCs	Time
AbstractRefinementMap	14	8	0.80	FunSets	134	19	1.11	NatMonoidLaws	75	39	0.71
Acc	26	1	0.14	GodelNumbering	486	389	84.00	NaturalBuiltin	25	8	0.11
Ackermann	12	12	0.29	H0Invocations	19	2	0.07	NestedLoop	19	4	0.08
AliasPartial	19	2	0.10	HammingMemoized	67	46	2.60	NotEquals	22	1	0.09
AmortizedQueue	155	64	3.92	Heaps	166	90	3.59	Numeric1	18	4	0.06
AnyDown	24	3	0.21	Huffman	520	228	9.52	OddEven	77	24	0.39
AssociativeFold	102	36	2.14	IgnoredField	48	2	0.07	OptionMonad	47	9	0.26
AssociativeList	66	28	0.80	IndirectHO	19	2	0.04	Overrides	25	9	0.24
BalancedParentheses	410	214	25.76	Induction	415	150	3.04	PackratParsing	137	31	1.06
BasicReal	28	6	0.18	InsertionSort	85	31	0.46	ParBalance	250	88	1.60
Basics	442	137	3.60	IntSet	65	29	0.51	PartialCompiler	66	38	10.67
BestRealTypes	27	1	0.14	IntSetProp	66	21	0.57	PartialKVTrace	77	19	3.69
BigIntMonoidLaws	38	9	1.96	IntSetUnit	60	23	0.51	Patterns	28	10	0.78
BinarySearch	18	7	6.61	Justify	79	46	0.90	Peano	36	12	0.28
BinomialHeap	189	22	0.95	Knapsack	73	23	0.30	PositiveMap	43	15	0.41
BitsTricks	103	35	0.72	LazyNumericalRep	158	29	1.15	PreflSpecs	27	6	0.09
BooleanOps	50	12	0.41	LazySelectionSort	69	21	0.40	PropositionalLogic	89	114	16.04
BottomUpMergeSort	127	62	2.80	LeftPad	65	25	1.10	Queue	28	9	0.25
Bytes	33	7	0.12	List	1006	529	16.07	QuickSorts	220	131	5.65
ChurchNum	21	0	0.03	ListMonad	86	29	0.97	ReachabilityChecker	561	282	15.34
Client	20	1	0.09	ListMonoidLaws	47	22	0.46	RealTimeQueue	77	9	0.33
ConcRope	493	254	49.21	ListWithSize	169	105	1.56	RedBlackTree	115	49	1.52
ConcTree	335	180	9.44	LiteralMaps	25	0	0.05	SearchLinkedList	59	22	0.27
CountTowardsZero	16	7	0.08	Longs	36	8	3.53	Shorts	33	7	0.10
Countable	29	17	0.49	Map	21	3	0.11	SimplInterpret	71	18	1.18
Deque	241	80	3.73	MapGetOrElse2	19	2	0.14	StableSorter	129	52	1.51
DivisionByZero	21	3	0.06	MapGetPlus	20	10	0.41	Streams	87	147	4.41
EffectfulPost	18	1	0.08	McCarthy91	24	8	0.17	Termination_passing1	37	10	0.16
EitherLaws	34	5	0.15	Mean	13	2	0.27	Theorem	28	1	0.05
Factorial	47	23	0.60	MergeSorts	301	167	4.51	ToChurch	27	4	0.07
Fibonacci	19	7	0.10	Methods	32	13	0.28	Trees1	29	11	0.43
FiniteSort	61	14	1.02	MicroTests	2246	505	15.59	TweetSet	494	253	11.91
FiniteStreams	42	13	0.30	Monoid	37	12	0.49	UpDown	49	11	0.25
FlatMap	60	20	0.66	MoreExtendedEuclidGCD	74	39	7.18	Viterbi	120	19	0.33
Foldr	19	3	0.09	MySet	18	2	0.03	XPlus2N	19	4	0.07
FoolProofAdder	13	1	0.04	NNF	135	177	17.77	example	95	70	1.74
Formulas	55	30	2.30	NNFSimple	107	103	5.27	recfun	66	38	0.89
Total LOC:			13842	Total VCs:			5815	Total Time (s):			387.94

Fig. 25. Summary of evaluation results for our verification tool, featuring lines of code, number of verification conditions (VCs), and type-checking times (including generation of VCs and checking of VCs by Inox).

algorithms such as Viterbi and Knapsack, and benchmarks solving assignments from Scala MOOCs (see <https://courseware.epfl.ch/> or <https://www.coursera.org/specializations/scala>).

8 RELATED WORK

In this paper, we tackle program verification using a type-theoretic approach. Other techniques, such as symbolic execution for verification (see e.g. [Hallahan et al. 2019; Nguyen et al. 2017]) or using term rewriting systems (see e.g. [Giesl et al. 2011, 2006, 2004]) can be used instead or in complement to our approach. The use of type checking in verification appears in program verifiers such as F^* [Ahman et al. 2017; Swamy et al. 2016, 2013] and Liquid Haskell [Vazou et al. 2013, 2014, 2018]. Like Dafny [Leino 2010] and Stainless, these systems rely on SMT solvers to automatically discharge verification conditions.

F^* is a dependently-typed programming language which supports a rich set of effects (such as divergence, mutation, etc.). To the best of our knowledge F^* does not support proving termination of functions operating on infinite data structures such as streams. Stainless has a desugaring pass that handles local state and certain forms of unique references. Furthermore, for semantic modeling of global state, a user can use monad syntax (Scala's `for` comprehensions) explicitly in the surface language, which Scala compiler desugars into higher-order functions that our system can handle. Overall, F^* supports a more expressive class of effects than what Stainless currently handles. The goal of Liquid Haskell [Vazou et al. 2018] is to add refinement types to Haskell, a call-by-need language (while we focus here on call-by-value languages). The system supports a sound

verification procedure which relies on decidable theories of SMT. The supported language imposes certain restrictions on refinement occurrences and the metatheory does not feature recursive types (although they are supported by the implementation). Similarly to our body-visible recursion, Liquid Haskell allows the type checker to access the body of recursive functions (for recursive calls) while type-checking the function itself using a technique called *refinement reflection* [Vazou et al. 2018]. Dafny verifier [Leino 2010] supports imperative and object-oriented as well as functional programming, including inductive and coinductive types, recursion using ordinals, quantifiers. We are not aware of metatheory to justify the soundness of Dafny. In contrast, we provide a mechanized proof of soundness for System FR, which increases confidence in VC generation soundness. This is important because it is easy to construct paradoxes when combining features such as impredicativity and contravariant recursion. Being based on a reducibility relation, our proof can be used as a basis to add new language features in a compositional way. Several other systems have been recently formalized, such as MetaCoq [Sozeau et al. 2019].

Proof assistants with great expressive power include Isabelle [Blanchette et al. 2017; Nipkow et al. 2002b], Coq [Barras 2010; Bertot and Castéran 2004b; Sacchini 2013], Idris [Brady 2013], Agda [Abel 2010; Norell 2007], PML₂ [Lepigre 2017], Lean [de Moura 2016] and Zombie [Casinghino et al. 2014]. When coinduction is supported in these systems (which is not always the case), it is typically through special constructs for coinductive types and their corresponding cofixpoint operators, dual to inductive types and their fixpoint operators. In System FR, we instead treat induction and coinduction in a uniform way (see e.g. Section 5.6). Our type system features a single kind of recursive types which allow uniform definition of inductive, coinductive and mixed recursive types. Our operational semantics further rely on the `fix` operator which roughly corresponds to general recursion in mainstream programming call-by-value languages. An alternative approach to uniform handling of recursion and corecursion is given in [Gianantonio and Miculan 2002]. We were able to encode the first example from this work (in our *Streams benchmark*) which generates a stream and whose termination proof requires a combination of inductive and coinductive reasoning. Using our system, we expressed that using a simple lexicographic measure combining the index of the type of the produced stream with the inductive measure. In a private communication, Andreas Abel showed us that Agda can also handle this example.

Our system follows Nuprl [Constable et al. 1986] (and other computational type theories) style of starting out with an untyped calculus and then introducing various types to classify untyped terms based on their behaviors. Nuprl supports a very expressive type system, which covers the types we present in this paper, except impredicative polymorphism. In a large development, Nuprl's metatheory has been formalized in Coq [Anand and Rahli 2014]. One part of that formalization relies on the use of the Coq axiom of functional choice `FunctionalChoice_on`, which gives a function f from A to B when a formula of the form $\forall a : A. \exists b : B. \phi(a, b)$ holds. Nuprl does not use SMT solvers for automation, but relies instead on built-in and user-defined tactics similarly to Coq.

Our reducibility definition for recursive type is inspired from step-indexed logical relations [Ahmed 2006]. The main difference is that the indices in step-indexed logical relations do not appear at the level of types, but at the level of the logical relation that gives meaning to types. In System FR we internalize the indices at the level of recursive types in order to give more expressive power to the users, and let them specify decreasing measures for recursive functions that manipulate infinite data structures. This treatment of recursive types is similar to the TORES [Jacob-Rao et al. 2018] type system, where recursive types can be indexed by an arbitrary index language. We only support recursive types indexed by natural numbers. TORES provides a decision procedure for a rich type system with inductive types, coinductive types, and indexed recursive types, yet its metatheory does not handle polymorphism nor refinement types.

Our termination criterion (rule for `fix`) is inspired by type-based termination [Abel 2004, 2007, 2008, 2012; Abel and Pientka 2013; Barthe et al. 2004]. Such work also typically uses two different kinds of recursion, one for induction and one for coinduction. In type-based termination, instead of requiring that a measure on the arguments of recursive calls decreases, we require that recursive functions are called at a *type* which is strictly smaller than the type of the caller. Our `fix` operator produces a term with a forall type, which is similar to the implicit function type of the implicit calculus of constructions [Miquel 2001]. Our termination measure on types can then be understood as a measure on the implicit argument of a function with an implicit function type. Earlier work on type-based termination checking includes simpler type systems such as DML [Xi 2001].

A proof of concept of extension of Dotty compiler by a restricted set of (mostly numerical) refinement predicates was presented in [Schmid and Kunčák 2016], without soundness proof. Developments specific to Scala include dependent object types [Amin 2016] which focus on path dependency instead of predicate refinement and, by design, admit possibly non-terminating programs. System FR does not attempt to support records, let alone path dependency. Stainless relies on Scala compiler front ends to obtain a symbol-resolved syntax tree. For example, implicit parameters in Scala [Odersky et al. 2018] become ordinary explicit parameters by the time Stainless processes them. Conversely, many of the types we define in System FR do not have their counterparts in Scala. An upcoming PhD thesis of Nicolas Voirol [Voirol 2019] presents a system related to System FR that more closely follows the expressive power of Inox solver on which Stainless relies.

9 CONCLUSION

We have presented System FR, a formalized type system and a bidirectional type checking algorithm that can serve as a basis of a verifier for higher-order functional programs. We were able to verify correctness and safety of a wide range of benchmarks amounting to 14k lines of code and proofs. Our formalization suggests that lazy data structures and non-covariant recursion are tractable and that explicit indices required in the general framework can be eliminated in commonly occurring cases. Our type system incorporates Π and Σ types, yet it also supports their variants (\forall and \exists) that correspond to infinite intersections and unions. Along with support for singleton types and refinements, we obtained a rich framework to approximate program semantics, which can also help further type, invariant, and measure inference algorithms. Our experience confirms an advantage of the semantic-based soundness proof: once we adopt an approach of interpreting types as sets of terms, we are less dependent on a particular choice of syntactic rules; we can introduce new classes of types and new rules for verification condition generation and solving, as long as we can justify them semantically. When viewing types in our system as propositions, we obtain an expressive quantified logic. We have proven the soundness of this logic in Coq using countable interpretation of ground System FR types. The success in verification of our benchmarks suggests that the rules of System FR work well for many properties of functional programs. Furthermore, refinement by type along with intersections and unions (Section 6.1) allows System FR to describe types whose interpretations are undecidable countable sets from higher levels of arithmetical hierarchy.

ACKNOWLEDGMENTS

Work supported in part by Swiss National Science Foundation project 200021_175676 “Scaling Predicate Types”. We thank Ravichandhran Kandhadai Madhavan for numerous discussions about termination for higher-order functional programs; Romain Ruetschi for his valuable contributions to Stainless implementation; Paolo G. Giarrusso for interesting discussions and very relevant pointers on type theory; Andreas Abel for explaining us how to mix induction and co-induction in Agda. Thanks to Yoan Gérard who implemented another prototype of our algorithm, found formalization discrepancies and proposed improvements to the typing rules.

REFERENCES

- Andreas Abel. 2004. Termination checking with types. *ITA* 38, 4 (2004), 277–319. <https://doi.org/10.1051/ita:2004015>
- Andreas Abel. 2007. *Type-based termination: a polymorphic lambda-calculus with sized higher-order types*. Ph.D. Dissertation. Ludwig Maximilians University Munich. <http://d-nb.info/984765581>
- Andreas Abel. 2008. Semi-Continuous Sized Types and Termination. *Logical Methods in Computer Science* 4, 2 (2008). [https://doi.org/10.2168/LMCS-4\(2:3\)2008](https://doi.org/10.2168/LMCS-4(2:3)2008)
- Andreas Abel. 2010. MiniAgda: Integrating Sized and Dependent Types. In *Partiality and Recursion in Interactive Theorem Provers, PAR@ITP 2010, Edinburgh, UK, July 15, 2010 (EPIc Series)*, Ekaterina Komendantskaya, Ana Bove, and Milad Niqui (Eds.), Vol. 5. EasyChair, 18–32. <https://doi.org/10.29007/322q>
- Andreas Abel. 2012. Type-Based Termination, Inflationary Fixed-Points, and Mixed Inductive-Coinductive Types. In *Proceedings 8th Workshop on Fixed Points in Computer Science, FICS 2012, Tallinn, Estonia, 24th March 2012. (EPTCS)*, Dale Miller and Zoltán Ésik (Eds.), Vol. 77. 1–11. <https://doi.org/10.4204/EPTCS.77.1>
- Andreas Abel and Brigitte Pientka. 2013. Wellfounded recursion with copatterns: a unified approach to termination and productivity. In *ACM SIGPLAN International Conference on Functional Programming, ICFP'13, Boston, MA, USA - September 25 - 27, 2013*, Greg Morrisett and Tarmo Uustalu (Eds.). ACM, 185–196. <https://doi.org/10.1145/2500365.2500591>
- Danel Ahman, Catalin Hritcu, Kenji Maillard, Guido Martínez, Gordon D. Plotkin, Jonathan Protzenko, Aseem Rastogi, and Nikhil Swamy. 2017. Dijkstra monads for free. In *Proceedings of the 44th ACM SIGPLAN Symposium on Principles of Programming Languages, POPL 2017, Paris, France, January 18-20, 2017*, Giuseppe Castagna and Andrew D. Gordon (Eds.). ACM, 515–529. <http://dl.acm.org/citation.cfm?id=3009878>
- Amal J. Ahmed. 2006. Step-Indexed Syntactic Logical Relations for Recursive and Quantified Types. In *Programming Languages and Systems, 15th European Symposium on Programming, ESOP 2006, Held as Part of the Joint European Conferences on Theory and Practice of Software, ETAPS 2006, Vienna, Austria, March 27-28, 2006, Proceedings (Lecture Notes in Computer Science)*, Peter Sestoft (Ed.), Vol. 3924. Springer, 69–83. https://doi.org/10.1007/11693024_6
- Nada Amin. 2016. *Dependent Object Types*. Ph.D. Dissertation. EPFL. <https://doi.org/10.5075/epfl-thesis-7156>
- Abhishek Anand and Vincent Rahli. 2014. Towards a Formally Verified Proof Assistant. In *Interactive Theorem Proving, Gerwin Klein and Ruben Gamboa (Eds.)*. Springer International Publishing, Cham, 27–44. https://doi.org/10.1007/978-3-319-08970-6_3
- Mike Barnett, K. Rustan M. Leino, and Wolfram Schulte. 2004. The Spec# Programming System: An Overview. In *Construction and Analysis of Safe, Secure, and Interoperable Smart Devices, International Workshop, CASSIS 2004, Marseille, France, March 10-14, 2004, Revised Selected Papers (Lecture Notes in Computer Science)*, Gilles Barthe, Lilian Burdy, Marieke Huisman, Jean-Louis Lanet, and Traian Muntean (Eds.), Vol. 3362. Springer, 49–69. https://doi.org/10.1007/978-3-540-30569-9_3
- Bruno Barras. 2010. Sets in Coq, Coq in Sets. *J. Formalized Reasoning* 3, 1 (2010), 29–48. <https://doi.org/10.6092/issn.1972-5787/1695>
- Gilles Barthe, Maria João Frade, Eduardo Giménez, Luís Pinto, and Tarmo Uustalu. 2004. Type-based termination of recursive definitions. *Mathematical Structures in Computer Science* 14, 1 (2004), 97–141. <https://doi.org/10.1017/S0960129503004122>
- Gilles Barthe, Benjamin Grégoire, and Colin Riba. 2008. A Tutorial on Type-Based Termination. In *Language Engineering and Rigorous Software Development, International LerNet ALFA Summer School 2008, Piriapolis, Uruguay, February 24 - March 1, 2008, Revised Tutorial Lectures*. 100–152. https://doi.org/10.1007/978-3-642-03153-3_3
- Yves Bertot and Pierre Castéran. 2004a. *Interactive Theorem Proving and Program Development - Coq'Art: The Calculus of Inductive Constructions*. Springer. <https://doi.org/10.1007/978-3-662-07964-5>
- Yves Bertot and Pierre Castéran. 2004b. *Interactive Theorem Proving and Program Development - Coq'Art: The Calculus of Inductive Constructions*. Springer. <https://doi.org/10.1007/978-3-662-07964-5>
- Régis Blanc and Viktor Kunčák. 2015. Sound reasoning about integral data types with a reusable SMT solver interface. In *Proceedings of the 6th ACM SIGPLAN Symposium on Scala, Scala@PLDI 2015, Portland, OR, USA, June 15-17, 2015*. 35–40. <https://doi.org/10.1145/2774975.2774980>
- Jasmin Christian Blanchette, Aymeric Bouzy, Andreas Lochbihler, Andrei Popescu, and Dmitriy Traytel. 2017. Friends with Benefits - Implementing Corecursion in Foundational Proof Assistants. In *Programming Languages and Systems - 26th European Symposium on Programming, ESOP 2017, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2017, Uppsala, Sweden, April 22-29, 2017, Proceedings (Lecture Notes in Computer Science)*, Hongseok Yang (Ed.), Vol. 10201. Springer, 111–140. https://doi.org/10.1007/978-3-662-54434-1_5
- Edwin Brady. 2013. Idris: general purpose programming with dependent types. In *Proceedings of the 7th Workshop on Programming languages meets program verification, PLPV 2013, Rome, Italy, January 22, 2013*, Matthew Might, David Van Horn, Andreas Abel, and Tim Sheard (Eds.). ACM, 1–2. <https://doi.org/10.1145/2428116.2428118>
- Chris Casinghino, Vilhelm Sjöberg, and Stephanie Weirich. 2014. Combining proofs and programs in a dependently typed language. In *The 41st Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages, POPL '14, San Diego, CA, USA, January 20-21, 2014*, Suresh Jagannathan and Peter Sewell (Eds.). ACM, 33–46. <https://doi.org/10.1145/2535838.2535883>

- Arthur Charguéraud. 2012. The Locally Nameless Representation. *J. Autom. Reasoning* 49, 3 (2012), 363–408. <https://doi.org/10.1007/s10817-011-9225-2>
- Robert L. Constable, Stuart F. Allen, Mark Bromley, Rance Cleaveland, J. F. Cremer, R. W. Harper, Douglas J. Howe, Todd B. Knoblock, N. P. Mendler, Prakash Panangaden, James T. Sasaki, and Scott F. Smith. 1986. *Implementing mathematics with the Nuprl proof development system*. Prentice Hall. <http://dl.acm.org/citation.cfm?id=10510>
- Leonardo de Moura. 2016. Formalizing Mathematics using the Lean Theorem Prover. In *International Symposium on Artificial Intelligence and Mathematics, ISAIM 2016, Fort Lauderdale, Florida, USA, January 4-6, 2016*. http://isaim2016.cs.virginia.edu/papers/ISAIM2016_Proofs_DeMoura.pdf
- David L. Detlefs, K. Rustan M. Leino, Greg Nelson, and James B. Saxe. 1998. *Extended Static Checking*. Technical Report 159. COMPAQ Systems Research Center.
- Pietro Di Gianantonio and Marino Miculan. 2002. A Unifying Approach to Recursive and Co-recursive Definitions. In *Types for Proofs and Programs, Second International Workshop, TYPES 2002, Berg en Dal, The Netherlands, April 24-28, 2002, Selected Papers (Lecture Notes in Computer Science)*, Herman Geuvers and Freek Wiedijk (Eds.), Vol. 2646. Springer, 148–161. https://doi.org/10.1007/3-540-39185-1_9
- Jürgen Giesl, Matthias Raffelsieper, Peter Schneider-Kamp, Stephan Swiderski, and René Thiemann. 2011. Automated Termination Proofs for Haskell by Term Rewriting. *ACM Trans. Program. Lang. Syst.* 33, 2 (Feb. 2011), 7:1–7:39. <https://doi.org/10.1145/1890028.1890030>
- J Giesl, S Swiderski, P Schneider-Kamp, and R Thiemann. 2006. Automated termination analysis for Haskell: From term rewriting to programming languages. In *Rewriting techniques and Applications*, Vol. 4098. 297–312.
- Jürgen Giesl, René Thiemann, Peter Schneider-Kamp, and Stephan Falke. 2004. Automated Termination Proofs with AProVE. In *Rewriting Techniques and Applications, 15th International Conference, RTA 2004, Aachen, Germany, June 3-5, 2004, Proceedings*. 210–220. https://doi.org/10.1007/978-3-540-25979-4_15
- Jean-Yves Girard. 1971. Une Extension De L'Interpretation De Gödel a L'Analyse, Et Son Application a L'Elimination Des Coupures Dans L'Analyse Et La Theorie Des Types. *Studies in Logic and the Foundations of Mathematics* 63 (1971), 63–92.
- Jean-Yves Girard. 1990. *Proofs and types*. Cambridge University Press. <http://www.paultaylor.eu/stable/prot.pdf>
- M. J. C. Gordon and T. F. Melham. 1993. *Introduction to HOL, a theorem proving environment for higher-order logic*. Cambridge University Press, Cambridge, England.
- William T. Hallahan, Anton Xue, Maxwell Troy Bland, Ranjit Jhala, and Ruzica Piskac. 2019. Lazy counterfactual symbolic execution. In *Proceedings of the 40th ACM SIGPLAN Conference on Programming Language Design and Implementation, PLDI 2019, Phoenix, AZ, USA, June 22-26, 2019*, Kathryn S. McKinley and Kathleen Fisher (Eds.). ACM, 411–424. <https://doi.org/10.1145/3314221.3314618>
- Jad Hamza, Nicolas Voirol, and Viktor Kunčák. 2019. System FR as Foundations for Stainless. arXiv:cs.LO/1904.03482
- Robert Harper. 2016. *Practical foundations for programming languages*. Cambridge University Press.
- John Harrison. 2009. *Handbook of Practical Logic and Automated Reasoning*. Cambridge University Press.
- John Harrison. 2017. HOL Light Tutorial. <https://www.cl.cam.ac.uk/~jrh13/hol-light/tutorial.pdf>. Retrieved 19 July 2019.
- Lars Hupel and Viktor Kunčák. 2016. Translating Scala Programs to Isabelle/HOL (System Description). In *International Joint Conference on Automated Reasoning (IJCAR)*.
- Rohan Jacob-Rao, Brigitte Pientka, and David Thibodeau. 2018. Index-Stratified Types. In *3rd International Conference on Formal Structures for Computation and Deduction, FSCD 2018, July 9-12, 2018, Oxford, UK (LIPICs)*, Hélène Kirchner (Ed.), Vol. 108. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 19:1–19:17. <https://doi.org/10.4230/LIPICs.FSCD.2018.19>
- EPFL IC LARA. 2019. Stainless: Formal Verification for Scala. <http://stainless.epfl.ch/>.
- K Rustan M Leino. 2010. Dafny: An Automatic Program Verifier for Functional Correctness. In *Logic for Programming, Artificial Intelligence, and Reasoning*. 348–370. https://doi.org/10.1007/978-3-642-17511-4_20
- Rodolphe Lepigre. 2017. *Semantics and Implementation of an Extension of ML for Proving Programs. (Sémantique et Implantation d'une Extension de ML pour la Preuve de Programmes)*. Ph.D. Dissertation. Grenoble Alpes University, France. <https://tel.archives-ouvertes.fr/tel-01590363>
- Alexandre Miquel. 2001. The Implicit Calculus of Constructions. In *TLCA*. 344–359. https://doi.org/10.1007/3-540-45413-6_27
- Phuc C. Nguyen, Sam Tobin-Hochstadt, and David Van Horn. 2017. Higher order symbolic execution for contract verification and refutation. *J. Funct. Program.* 27 (2017), e3. <https://doi.org/10.1017/S0956796816000216>
- Tobias Nipkow, Lawrence C. Paulson, and Markus Wenzel. 2002a. *Isabelle/HOL - A Proof Assistant for Higher-Order Logic*. Lecture Notes in Computer Science, Vol. 2283. Springer. <https://doi.org/10.1007/3-540-45949-9>
- Tobias Nipkow, Lawrence C. Paulson, and Markus Wenzel. 2002b. *Isabelle/HOL - A Proof Assistant for Higher-Order Logic*. Lecture Notes in Computer Science, Vol. 2283. Springer. <https://doi.org/10.1007/3-540-45949-9>
- Ulf Norell. 2007. *Towards a practical programming language based on dependent type theory*. Ph.D. Dissertation. Department of Computer Science and Engineering, Chalmers University of Technology, SE-412 96 Göteborg, Sweden.
- Martin Odersky, Olivier Blanvillain, Fengyun Liu, Aggelos Biboudis, Heather Miller, and Sandro Stucki. 2018. Simplicity: foundations and applications of implicit function types. *PACMPL* 2, POPL (2018), 42:1–42:29.

- Martin Odersky, Lex Spoon, and Bill Venners. 2008. *Programming in Scala: a comprehensive step-by-step guide*. Artima Press.
- Chris Okasaki. 1998. *Purely Functional Data Structures*. Cambridge University Press.
- Aleksandar Prokopec and Martin Odersky. 2015. Conc-Trees for Functional and Parallel Programming. In *Languages and Compilers for Parallel Computing, LCPC*. 254–268. https://doi.org/10.1007/978-3-319-29778-1_16
- Jorge Luis Sacchini. 2013. Type-Based Productivity of Stream Definitions in the Calculus of Constructions. In *28th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2013, New Orleans, LA, USA, June 25-28, 2013*. IEEE Computer Society, 233–242. <https://doi.org/10.1109/LICS.2013.29>
- Georg Stefan Schmid and Viktor Kunčák. 2016. SMT-Based Checking of Predicate-Qualified Types for Scala. In *Scala Symposium*.
- Matthieu Sozeau. 2010. Equations: A Dependent Pattern-Matching Compiler. In *Interactive Theorem Proving, First International Conference, ITP 2010, Edinburgh, UK, July 11-14, 2010. Proceedings (Lecture Notes in Computer Science)*, Matt Kaufmann and Lawrence C. Paulson (Eds.), Vol. 6172. Springer, 419–434. https://doi.org/10.1007/978-3-642-14052-5_29
- Matthieu Sozeau, Abhishek Anand, Simon Boulier, Cyril Cohen, Yannick Forster, Fabian Kunze, Gregory Malecha, Nicolas Tabareau, and Théo Winterhalter. 2019. The MetaCoq Project. (June 2019). <https://hal.inria.fr/hal-02167423> working paper or preprint.
- Philippe Suter, Ali Sinan Köksal, and Viktor Kuncak. 2011. Satisfiability Modulo Recursive Programs. In *Static Analysis Symposium (SAS)*.
- Nikhil Swamy, Catalin Hritcu, Chantal Keller, Aseem Rastogi, Antoine Delignat-Lavaud, Simon Forest, Karthikeyan Bhargavan, Cédric Fournet, Pierre-Yves Strub, Markulf Kohlweiss, Jean Karim Zinzindohoue, and Santiago Zanella Béguelin. 2016. Dependent types and multi-monadic effects in F. In *Proceedings of the 43rd Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages, POPL 2016, St. Petersburg, FL, USA, January 20 - 22, 2016*, Rastislav Bodík and Rupak Majumdar (Eds.). ACM, 256–270. <https://doi.org/10.1145/2837614.2837655>
- Nikhil Swamy, Joel Weinberger, Cole Schlesinger, Juan Chen, and Benjamin Livshits. 2013. Verifying higher-order programs with the dijkstra monad. *ACM SIGPLAN Conference on Programming Language Design and Implementation* (2013), 387. <https://doi.org/10.1145/2491956.2491978>
- W. W. Tait. 1967. Intensional interpretations of functionals of finite type I. *Journal of Symbolic Logic* 32, 2 (1967), 198–212. <https://doi.org/10.2307/2271658>
- Niki Vazou, Patrick Maxim Rondon, and Ranjit Jhala. 2013. Abstract Refinement Types. In *European Symposium on Programming, ESOP*. https://doi.org/10.1007/978-3-642-37036-6_13
- Niki Vazou, Eric L. Seidel, Ranjit Jhala, Dimitrios Vytiniotis, and Simon Peyton-Jones. 2014. Refinement Types for Haskell. In *International Conference on Functional Programming, ICFP*. 269–282. <https://doi.org/10.1145/2628136.2628161>
- Niki Vazou, Anish Tondwalkar, Vikraman Choudhury, Ryan G. Scott, Ryan R. Newton, Philip Wadler, and Ranjit Jhala. 2018. Refinement reflection: complete verification with SMT. *PACMPL* 2, POPL (2018), 53:1–53:31. <https://doi.org/10.1145/3158141>
- Nicolas Voirol. 2019. *Verifying Functional Programs*. Ph.D. Dissertation. EPFL.
- Nicolas Voirol, Etienne Kneuss, and Viktor Kuncak. 2015. Counter-example complete verification for higher-order functions. In *Symposium on Scala, Scala 2015*. 18–29. <https://doi.org/10.1145/2774975.2774978>
- Hongwei Xi. 2001. Dependent Types for Program Termination Verification. In *16th Annual IEEE Symposium on Logic in Computer Science, Boston, Massachusetts, USA, June 16-19, 2001, Proceedings*. 231–242. <https://doi.org/10.1109/LICS.2001.932500>