

# PROGRAM-ing Finger Trees in COQ

Matthieu Sozeau

Univ. Paris Sud, CNRS, Laboratoire LRI, UMR 8623, Orsay, F-91405  
INRIA Futurs, ProVal, Parc Orsay Université, F-91893  
sozeau@lri.fr

## Abstract

Finger Trees (Hinze and Paterson 2006) are a general purpose persistent data structure with good performance. Their genericity permits developing a wealth of structures like ordered sequences or interval trees on top of a single implementation. However, the type systems used by current functional languages do not guarantee the coherent parameterization and specialization of Finger Trees, let alone the correctness of their implementation. We present a certified implementation of Finger Trees solving these problems using the PROGRAM extension of COQ. We not only implement the structure but also prove its invariants along the way, which permit building certified structures on top of Finger Trees in an elegant way.

**Categories and Subject Descriptors** D.2.4 [Software/Program Verification]: Correctness proofs, Formal methods; F.3.1 [Specifying and Verifying and Reasoning about Programs]: Mechanical verification, Specification techniques

**General Terms** Algorithms, Languages, Verification

**Keywords** Coq, Dependent Types, Finger Trees, Certification

## 1. Introduction

Finger Trees are based on an optimization of 2-3 trees which gives constant or in the worst case logarithmic amortized time to the usual sequence operations, from adding an element at either end to splitting at an arbitrary location. Suffice to say that it has been integrated as the `Data.Sequence` implementation in `HASKELL 6.6` to witness the fact that it is a practical, useful data structure. Our first contribution will be to show that the implementation given by Hinze & Paterson is correct, which means that all functions are terminating and that the invariants of Finger Trees are respected. In fact, as the invariants are part of our data structure we will also carry the properties to the client side, which lead us to our second contribution, a certified implementation of random-access sequences built on top of Finger Trees. This development<sup>1</sup> was made possible using the PROGRAM extension of COQ. It is based on a high-level programming language (RUSSELL) which permits writing functional programs with rich specifications and a translation to COQ generating proof obligations. This paper also serves as a gentle introduction to PROGRAM-ing in COQ.

<sup>1</sup>[www.lri.fr/~sozeau/research/russell/fingertrees.en.html](http://www.lri.fr/~sozeau/research/russell/fingertrees.en.html)

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

ICFP'07, October 1–3, 2007, Freiburg, Germany.  
Copyright © 2007 ACM 978-1-59593-815-2/07/0010...\$5.00

The remaining of this paper is organized as follows: in sections 2 and 3 we give a short introduction to COQ and RUSSELL respectively, walking through simple examples. Then we move on to the implementation of Finger Trees, in section 4, their specialization in section 5 and their extraction in section 6. We finally discuss our method and results in section 7. No prior knowledge of the COQ tool is assumed, but practice of a typed functional language is recommended. Familiarity with Finger Trees is not required either.

**Notations** The bulk of this paper consists of a set of literate COQ files processed by the `coqdoc` tool. COQ code is typeset with the usual mathematical style; variable and definition names are typeset in *italic*, inductive types and constructor names are typeset in sans-serif and language keywords are typeset in `typewriter` font.

## 2. Preliminaries

In this section we present COQ's underlying calculus and some features of the tool we use in the rest of the paper. It can be safely skipped by COQ users.

### 2.1 CIC power

COQ is a proof assistant based on the Calculus of Inductive Constructions (CIC), a very powerful type system and logical language. It relies on the Curry-Howard isomorphism between proofs and programs, types and logical statements to provide an environment in which programming and proving are seemingly intermingled. CIC is an extension of the Calculus of Constructions (CC), the richest type system of Barendregt's  $\lambda$ -cube, which integrates polymorphic, dependent types with highly expressive (co-)inductive types. The environment built around this kernel allows users to do interactive theorem proving of mathematical developments and certified programming using the underlying  $\lambda$ -calculus. It provides facilities to define datatypes and functions like a regular functional programming language but it also gives the ability to write rich specifications and construct proofs in the same language viewed as a higher-order logic.

We present the typing rules of CC in figure 1 as a reminder. We will extend them when defining RUSSELL's type system in section 3. The basic calculus is church-style  $\lambda$ -calculus plus a set of sorts  $S = \{\text{Prop}\} \cup \{\text{Type}(i) \mid i \in \mathbb{N}\}$  and the dependent product  $\Pi x : A, B$  which types abstractions (rules PROD, ABS, APP). We do not detail the relation  $\mathcal{R}$  which allows an impredicative Prop sort and a predicative Type hierarchy. The sort  $\text{Type}(0)$  represents computational types (e.g. lists, naturals) and Prop logical propositions. Both sorts are typed by  $\text{Type}(1)$  which is itself typed using the rule TYPE. An essential rule is CONV which internalizes the fact that  $\beta$ -convertible types can be considered equal during type-checking ( $\equiv$  is the  $\beta$ -equivalence relation). Here lies the true power of the system: computation can be interleaved with reasoning anywhere. We do not treat the full kernel of COQ, which also contain sort subtyping and inductive datatypes as they are not essential to

$$\begin{array}{c}
\text{WF-EMPTY} \frac{}{\vdash \square \mathbf{wf}} \quad \text{WF-VAR} \frac{\Gamma \vdash A : s}{\vdash \Gamma, x : A \mathbf{wf}} \quad s \in \mathcal{S}, x \notin \Gamma \\
\\
\text{PROP} \frac{\vdash \Gamma \mathbf{wf}}{\Gamma \vdash \text{Prop} : \text{Type}(1)} \\
\\
\text{TYPE} \frac{\vdash \Gamma \mathbf{wf}}{\Gamma \vdash \text{Type}(i) : \text{Type}(i+1)} \quad i \in \mathbb{N} \\
\\
\text{VAR} \frac{\vdash \Gamma \mathbf{wf} \quad x : A \in \Gamma}{\Gamma \vdash x : A} \\
\\
\text{PROD} \frac{\Gamma \vdash T : s_1 \quad \Gamma, x : T \vdash U : s_2 \quad (s_1, s_2, s_3) \in \mathcal{R}}{\Gamma \vdash \Pi x : T.U : s_3} \\
\\
\text{ABS} \frac{\Gamma, x : T \vdash M : U \quad \Gamma \vdash \Pi x : T.U : s}{\Gamma \vdash \lambda x : T.M : \Pi x : T.U} \\
\\
\text{APP} \frac{\Gamma \vdash f : \Pi x : V.W \quad \Gamma \vdash u : V}{\Gamma \vdash (fu) : W[u/x]} \\
\\
\text{CONV} \frac{\Gamma \vdash t : U \quad \Gamma \vdash U \equiv T : s}{\Gamma \vdash t : T}
\end{array}$$

**Figure 1.** Calculus of Constructions

describe the core features RUSSELL will add; inductives will be described next by example.

Important properties of this system are Subject Reduction and Strong Normalization, i.e.: all well-typed terms reduce to a value in a finite number of reductions, which ensures that typing is decidable in the CONV rule. With such expressive power and strong metatheoretical properties comes verbosity and inflexibility. We will present in the following sections what mechanisms help overcome these difficulties in the COQ tool. But first we will present the salient features of inductives in COQ.

**Inductive types** Inductive types are a generalization of the usual algebraic datatypes of ML. For example we may define the standard polymorphic lists using the declaration:

```

Inductive list (A : Type) : Type :=
| nil : list A
| cons : A → list A → list A.

```

Here  $A$  is an *inductive parameter* representing the type of elements of the list and each constructor is declared with a product type from its arguments types to the inductive type. This presentation as a product comes from the fact that constructor arguments may depend on each other and that each constructor may target a different instance of the family. For example we may declare vectors (lists of fixed length) as:

```

Inductive vector (A : Type) : nat → Type :=
| vnil : vector A 0
| vcons : A → ∀ n, vector A n → vector A (S n).

```

The inductives of COQ (Paulin-Mohring 1993) are in fact parameterized, possibly nested, mutually recursive inductive families. An inductive family is a family of types indexed by a type or a value. Here for example, vector is a family parameterized by a type  $A$  of elements and indexed by the naturals  $\text{nat}$ . The `vnil` constructor is the empty vector of length 0 and the `vcons` constructor builds a vector of length  $S n$  from an object of type  $A$  and a vector of length  $n$ .

Finally, we can define nested inductive types using non-uniform parameters; this feature is needed to be able to define Finger Trees.

```

Fixpoint pow_map (A B : Type) (f : A → B)
(l : pow_list A) { struct l } : pow_list B :=
match l with
| pow_nil ⇒ pow_nil B
| pow_cons hd tl ⇒ pow_cons B (f hd)
  (pow_map (A × A) (B × B)
   (fun x : A × A ⇒ (f (fst x), f (snd x)))) tl
end.

```

**Figure 2.** `pow_map` definition

A classical example of a nested datatype is the power list of  $A$  where the tail of a list is a list of pairs of  $A$ . The important point is that we can reinstantiate the parameter  $A$  in the constructors.

```

Inductive pow_list (A : Type) : Type :=
| pow_nil : pow_list A
| pow_cons : A → pow_list (A × A) → pow_list A.

```

We can of course recurse on elements of these inductive types and in the last case we will even use polymorphic recursion, a feature the type system of ML lacks. For example the `map` combinator on power lists can be defined as in figure 2.

To ensure the aforementioned properties are preserved when adding inductive families, the kernel implements a relatively inflexible syntactic criterion to discard non-structurally recursive definitions. In the `pow_map` example, it checks that `tl` is indeed a subterm of `l` which the user marked as the structurally decreasing argument using the `{struct id}` annotation. In this definition we also face the foremost difficulty in using CIC as a programming language: verbosity. Clearly, writing all these types is not viable from a software development point of view, so we must find a way to overcome this problem while still retaining decidability of inference and type-checking. COQ's solution to this problem is to use a higher-level language GALLINA which elaborates/compiles into CIC. This approach does not change the kernel's language *at all*, so we keep all the nice properties once our definitions have compiled into it.

## 2.2 COQ surface: GALLINA

We will now present a few of COQ's surface language features which deal with the verbosity and abstraction power of CIC: implicit arguments to omit some type information, notations to abbreviate and beautify the otherwise obfuscated definitions and sections to easily parameterize whole developments.

### 2.2.1 Recovering inference

Implicit arguments systems permit writing terms omitting as much inferable information as possible. Its effect is best described by an example application. Suppose we want to apply the `pow_map` function to a function  $f : A \rightarrow B$  and a power list  $l : \text{pow\_list } A$ . Without implicit arguments, we must give the type of  $f$  as the first arguments of the function, like this: `pow_map A B f l`. In implicit arguments mode, the first two arguments become implicit as the system knows that they will be inferable at application time from the *type* of the functional argument  $f$ , hence we can write `pow_map f l` to achieve the same effect. This permits users to write definitions with a syntax very similar to languages based on Hindley-Milner type inference, e.g. the `pow_map` definition comes down to the code in figure 3. We do not consider partial applications here for which this mechanism is incomplete.

### 2.2.2 Notations: pretty parsing/printing

Notations are kind of hygienic macros which serve to abbreviate or beautify specifications and generally give a more mathematical feel to writing COQ. For example we can define the type representing

```

Fixpoint pow_map (A B : Type) (f : A → B)
  (l : pow_list A) { struct l } : pow_list B :=
  match l with
  | pow_nil ⇒ pow_nil
  | pow_cons hd tl ⇒ pow_cons (f hd)
    (pow_map (fun x ⇒ (f (fst x), f (snd x))) tl)
  end.

```

**Figure 3.** *pow\_map* definition with implicit arguments

the subset of elements of a type  $A$  having a particular property  $P$  as in figure 4.

```

Inductive sig (A : Type) (P : A → Prop) : Type :=
| exist : ∀ a : A, P a → sig A P.
Notation "{ x : A | P }" := (sig A (fun x : A ⇒ P))

```

**Figure 4.** Subset type in Coq

An element of the type  $\text{sig } A \ P$  is a tuple  $\text{exist } A \ P \ x \ p$  where  $x$  is a witness of type  $A$  and  $p$  is of type  $P \ a$ , i.e.  $p$  is a proof that the predicate  $P$  holds for  $a$ . However it is much clearer to write this type as  $\{x : A \mid P\}$  rather than  $\text{sig } A \ P$  and this is what the notation system permits.

An alternative encoding of the vector inductive structure may be the lists of fixed lengths which we can define using this notation (the types of  $A$  and  $n$  can be inferred automatically in this case):

**Definition**  $\text{vector } A \ n := \{x : \text{list } A \mid \text{length } x = n\}$ .

One can define arbitrary notations, possibly for binding constructs, and assign them priorities (levels) and associativities. For example we may define composition using the symbol  $\circ$  as a notation:

**Notation** " $g \circ f$ " :=  $(\text{fun } x \Rightarrow g (f \ x))$   
(at level 20, left associativity).

Armed with these tools we can start our first development.

### 2.2.3 Sections: implementing the monoid typeclass

As a gentle introduction to COQ, we will build a monoid datatype corresponding to the Monoid typeclass of HASKELL. We recall its signature:

```

class Monoid m where
  () :: m
  (⊙) :: m → m → m

```

Instances should make sure the following properties hold:

```

left identity   () ⊙ x = x
right identity  x ⊙ () = x
associativity   x ⊙ (y ⊙ z) = (x ⊙ y) ⊙ z

```

We first introduce the monoid laws we just described. We will use the section mechanism of COQ extensively in the following. Sections permit to write a bunch of definitions which are parameterized by some variables, e.g. for types and operations. When closing a section, every definition inside it is quantified over the variables it used.

**Section** *Monoid\_Laws*.

The carrier  $m$  type may be any type.

**Variable**  $m : \text{Type}$ .

The identity element *empty* and the operation *mappend*.

**Variables** (*empty* :  $m$ ) (*mappend* :  $m \rightarrow m \rightarrow m$ ).

We define fancy notations for the element and operation.

**Notation**  $\varepsilon := \text{empty}$ .

**Infix** "·" := *mappend* (right associativity, at level 20).

We now state the properties.

**Definition** *monoid\_id\_l\_t* :  $\text{Prop} := \forall x, \varepsilon \cdot x = x$ .

**Definition** *monoid\_id\_r\_t* :=  $\forall x, x \cdot \varepsilon = x$ .

**Definition** *monoid\_assoc\_t* :=  $\forall x \ y \ z, (x \cdot y) \cdot z = x \cdot y \cdot z$ .

**End** *Monoid\_Laws*.

Every variable in *Monoid\_Laws* has been discharged by now, so we must apply each definition to particular *empty* and *mappend* objects. We can finally define the dependent record which represents monoids on  $m$ :

```

Record monoid (m : Type) : Type := mkMonoid
  { empty : m ; mappend : m → m → m
  ; monoid_id_l : monoid_id_l_t empty mappend
  ; monoid_id_r : monoid_id_r_t empty mappend
  ; monoid_assoc : monoid_assoc_t mappend }.

```

We will come back to monoids when implementing Finger Trees. We will now present the RUSSELL language which was the essential tool we needed to develop them.

## 3. RUSSELL in COQ

The lack of distinction between proofs and programs in Type Theory helps having a clear, unified semantics for the complete system but hinders users and implementers alike, because sometimes you do not want to have proofs polluting computations (e.g. when evaluating a term, the way proofs are constructed should not matter, only the algorithm) and you certainly do not want to write complex programs using a procedural language with automation like the language used in proof scripts of COQ but directly inputting terms and have them interpreted in COQ's kernel language. A solution to the first problem is given by the Prop/Type distinction in COQ. If Prop is used as the type of propositions (e.g. conjunction, negation) and Type as the type of computational types (e.g. naturals, lists) then we can immediately see whether something is a proof or a program just by looking at its sort. This is used by the extraction mechanism (Letouzey 2002) to extract *only* the algorithmic code from a COQ term. However, the extraction is not used directly inside the proof-checker, so computations in COQ may freeze because of irreducible proofs in the term, even when they are irrelevant; we will discuss this later issue in section 7. For now we will concentrate on the problem of writing richly specified programs in COQ by taking advantage of this distinction between Prop and Type.

Usually, as soon as you have propositions appearing in the type of a function, you will write its body in COQ using tactics instead of using the functional language because otherwise you would have to directly manipulate proof terms. This is a showstopper because it becomes intractable as soon as you have a complex proof to do, especially since there is not so much support for incremental refinement in COQ, unlike in AGDA (Coquand 2006) or EPIGRAM (McBride and McKinna 2004), although the *ref ine tactic* can give the same flavor of program development (Chlipala 2006).

RUSSELL is a programming language built on top of COQ which permits writing only the algorithmic code of strongly specified functions and forgetting about the required proofs which COQ needs to establish correctness (i.e. typecheck the term). Its main purpose is to act as a convenient, permissive source language which elaborates into CIC, which can be type-checked by the safe kernel of COQ. Here, permissive means the user is not restricted to using structural recursion, complete pattern-matching and total functions only, as we have seen up to now. Obviously, you have to give evidence that your source term can be seen as a legitimate COQ

term, which generally means that what you left out can be proved. This is done by solving obligations which are generated *after* the type-checking procedure of RUSSELL terms by an interpretation into CIC. For a detailed presentation of this procedure, the reader is directed to (Sozeau 2007). The generation of obligations is done in a similar way as in PVS (Owre and Shankar 1997), using subsets to carry propositions. We will describe this essential feature of the RUSSELL type system and its extension to the richer CIC calculus next, then move on gradually to our motivating example: Finger Trees.

### 3.1 Subset equivalence

Subset equivalence is a simple idea: two subsets over the same carrier may be considered equal, given a proof that they contain the same elements. We can leave the properties out and just consider the carrier of subsets. That is, if we have an object  $t$  of type  $T$  where we expect an object of type  $\{x : T \mid P\}$ , we may say “ok, you will have to give me a proof that  $P[t/x]$  holds but we are good for now”. Conversely, considering an object of type  $\{x : T \mid P\}$  as an object to type  $T$  is alright. This is the essential idea behind the type system of RUSSELL: it checks type equivalence on the carriers only, ignoring if the properties really hold. We need only enrich the type conversion relation of CIC with the rule given figure 5 to formalize this idea.

$$\frac{\Gamma \vdash T \equiv U : \text{Type} \quad \Gamma, x : U \vdash P : \text{Prop}}{\Gamma \vdash T \equiv \{x : U \mid P\} : \text{Type}}$$

Figure 5. Subset equivalence rule

As an example, we can derive:

$$\frac{\Gamma \vdash \mathbb{N} \equiv \mathbb{N} : \text{Type} \quad \Gamma, x : \mathbb{N} \vdash x = 0 : \text{Prop}}{\Gamma \vdash 0 : \mathbb{N} \quad \Gamma \vdash \mathbb{N} \equiv \{x : \mathbb{N} \mid x = 0\} : \text{Type}} \\ \Gamma \vdash 0 : \{x : \mathbb{N} \mid x = 0\}$$

After a term has been typed in this system, we can automatically generate the obligations and produce a full-fledged, well-typed COQ term once these are solved. Following the example, a single obligation demanding that  $0 = 0$  be proved will be generated which will easily be solved by reflexivity of equality. But we could have derived  $\Gamma \vdash 0 : \{x : \mathbb{N} \mid x \neq 0\}$  in RUSSELL as well, in which case the generated obligation could certainly not be solved except if  $\Gamma$  is inconsistent. This phase distinction is essential: as soon as we want to do any proof we jump into an undecidable world because we have the full higher-order logic of COQ at hand, so we must not have any proving to do if we want our type-checking to be decidable. This is to contrast with the approaches of other dependently-typed programming languages, from DML (Xi and Pfenning 1999) to EPIGRAM. In these languages either the expressive power is limited so that an automatic prover can solve obligations during type-checking (e.g. Pressburger arithmetic in DML) or there is no limit but proofs and code are intermingled in the same language (for EPIGRAM, COQ and AGDA) or a mix of the two styles in ATS (Xi 2004) or  $\Omega$ MEGA (Sheard 2004) for example.

### 3.2 RUSSELL’s incarnation: PROGRAM

Using RUSSELL we can write arbitrarily complex code with arbitrarily complex specifications, it will typecheck only if it is “structurally” well-typed. We do not care about proofs at the type-checking step but later, when recombining obligations with the algorithmical skeleton we wrote. Let us define the strongly specified euclidean division of  $a$  by  $b$  where  $b$  must be non-zero. This

function is defined by well-founded recursion using the less-than order on the argument  $a$ :

```
Program Fixpoint div (a : nat) (b : nat | b ≠ 0) { wf lt } :
  { (q, r) : nat × nat | a = b × q + r ∧ r < b } :=
  if less_than a b then (0, a)
  else dest div (a - b) b as (q', r) in (S q', r).
```

The PROGRAM vernacular indicates that we use the RUSSELL type-checker followed by the interpretation into COQ instead of COQ’s original type-checker. The construct `dest t as p in b` is equivalent to a destructive `let` construct where  $p$  is an arbitrary pattern (allowing one to pull values from nested tuples for example) but it also reflects the equality between  $t$  and  $p$  in the context for later use in the obligations. We also use sugar for the binder  $(x : \text{nat} \mid P) \triangleq (x : \{x : \text{nat} \mid P\})$ . Finally, the `if` construct works on any inductive type with two constructors.

This code type-checks and PROGRAM generates obligations which are discharged automatically by COQ’s tactics except the one for the inductive case. We can discharge it interactively using COQ’s proof language. When all obligations are solved the PROGRAM system adds the completed definition of `div` in the environment. We can then use `div` like any other COQ object, reasoning and computing with it as desired.

We were able to discharge the obligations because hypotheses  $a < b$  or  $a \geq b$  were present in their contexts. However, these hypotheses are not automatically generated by COQ. They come directly from the term which uses a dependent function `less_than` of type:  $\forall x y, \{x < y\} + \{x \geq y\}$ . Had we simply used a boolean function instead, we would have no information about the comparison in the obligations. Hence we must always take care of *reflecting* the logic in our code if we want to reason on it later. In case of boolean conditionals, we can always use the combinator `dec` :  $\forall b, \{b = \text{true}\} + \{b = \text{false}\}$  to achieve this reflection.

### 3.3 Extension to inductive types

It is natural to extend the support for subset types in RUSSELL to inductive families. For example, we have seen we can have two encodings of vectors, one using subsets and another using an inductive family. The facility we added for handling subset types should then transpose easily to inductive types. Indeed we can obtain the RUSSELL derivation:

$$\frac{\Gamma \vdash x : \text{vector } A m \quad \Gamma \vdash \text{vector } A m \equiv \text{vector } A n : \text{Type}}{\Gamma \vdash x : \text{vector } A n}$$

The type conversion results from an instantiation of the rule given figure 6 which states that an inhabitant of an inductive family  $I$  with indexes  $\vec{x}_i$  also inhabits the type  $I \vec{y}_i$  (coercion does not happen on parameters).

$$\frac{I \text{ inductive family} \quad \Gamma \vdash I \vec{x}_i, I \vec{y}_i : \text{Type}}{\Gamma \vdash I \vec{x}_i \equiv I \vec{y}_i : \text{Type}}$$

Figure 6. Dependent inductive equivalence rule

When interpreting into COQ, proof obligations will be generated to show that indices are equal. For the example above, a proof of  $\Gamma \vdash m = n : \text{Prop}$  will be required. This extension was crucial in the development of Dependent Finger Trees where a good part of the obligations are generated due to applications of this rule.

Dually to this construction used for building an inductive object with arbitrary index, we have enriched the pattern-matching construct to get information about destructured values and their indexes (i.e. automatically reflecting pattern-matching at the logical level). Informally, in our setting each branch of a pattern-matching construct will be typed in a context extended with equalities between

the patterns and the matched values up to their indices. For example, in a context where  $v$  is of type `vector A (S n)`, when typing the term `match v with vnil => ... | vcons x n' v' => ... end` the first branch will be typed in a context containing hypotheses  $S\ n = 0$  and  $v \simeq \text{vnil}$ , hence an absurd case. The equality symbol  $\simeq$  will be presented in § 4.4.1. We will also be able to deduce that  $n = n'$  in the second branch because we will have an equality  $S\ n = S\ n'$  in the context (and constructors are injective). As a further refinement of pattern-matching, we add inequality hypotheses in the context of branches for which the pattern was intersecting with a previous one. Consider the following code:

```
Program Fixpoint div2 (n : nat) :
  { n' : nat | n = 2 * n' ∨ n = 2 * n' + 1 } :=
  match n with
  | S (S p) => S (div2 p)
  | x => 0
  end.
```

The second branch is typed in a context extended with the hypotheses  $x = n$  and  $\forall p, x \neq S\ (S\ p)$ , which is crucial to solve the obligation  $n = 2 * 0 \vee n = 2 * 0 + 1$ .

**A conclusion on RUSSELL** We think having an unrestricted power of expression in the logic is essential to having useful software development and certification tools and our method achieves it without losing the facilities we are accustomed to in a practical programming language. So we get decidability of type-checking and our programs look exactly like their simply-typed counterparts with no proof handling code appearing in the term but we also have a highly expressive logic and a complete system to prove and check statements in this logic.

## 4. Dependent Finger Trees

We will now illustrate RUSSELL's expressiveness by presenting a certified implementation of Finger Trees in COQ. Finger Trees (Hinze and Paterson 2006) are a functional, general purpose, efficient data structure based on an optimization of 2-3 trees. It is implemented in HASKELL as a *nested* datatype which is parameterized by a type of *measures*, which can be instantiated by various types to give specializations of the Finger Trees. The data structure builds upon simpler structures like digits and 2-3 nodes which we will present first.

### 4.1 Digits

A digit is simply a buffer of one to four values.

Section *Digit*.

```
Variable A : Type.
Inductive digit : Type :=
| One : A → digit
| Two : A → A → digit
| Three : A → A → A → digit
| Four : A → A → A → A → digit.
```

We build simple functional predicates on digits for use in specifications.

```
Definition full (x : digit) : Prop :=
  match x with Four _ _ _ _ => True | _ => False end.
```

We now define addition of an element to the left of a digit.

```
Program Definition add_digit_left (a : A)
  (d : digit | ¬ full d) : digit :=
  match d with
  | One x => Two a x
  | Two x y => Three a x y
```

```
a      : A
d      : digit
Hd     : ¬ full d
x, y, z, w : A
Heq_d  : d = Four x y z w
False
```

Figure 7. Obligation of `add_digit_left`

```
| Three x y z => Four a x y z
| Four _ _ _ => !
end.
```

It has become interesting here, as we define a *partial* function. We can add to a digit if and only if it is not already full. So, we require the argument `digit` to be accompanied by a proof that it is not full. We will use it to prove that the last branch is inaccessible. Note that we can pattern-match on  $d$  as if it was just a digit: properties have no influence on code, only on proofs.

As in COQ, there has to be a branch for each and every constructor in RUSSELL pattern-matchings, but we can use `!` to assert that a branch is inaccessible due to contextual information on the matched object. We must treat every constructor because exhaustiveness of pattern matching with dependent types is undecidable in general. The generated obligation (figure 7) is easily solved, as we have a contradiction in the context: both  $\neg \text{full } d$  and  $d = \text{Four } \_ \_ \_$  are present. We can define in a similar fashion addition on the right and the various accessors on digits (*head*, *tail*, *last* and *liat*). From now on we will omit the obligations and the proof scripts used to solve them; they can be found in the COQ contribution.

As an other example of the usefulness of PROGRAM, consider defining the monoid  $([], ++)$  on lists. To construct a monoid implementation (c.f. §2.2.3), we have to apply the `mkMonoid` constructor to `[], ++` and proofs of identity and associativity. Fortunately, those can be filled automatically by COQ using lemmas proved in the standard library. When PROGRAM processes the following definition it turns the holes  $(\_)$  in the term into obligations which are thus automatically discharged. We explicitly give the element and operation using the syntax  $(x:=t)$ ; they are considered implicit arguments by the system as they are derivable from the types of the proofs.

```
Program Definition listMonoid (A : Type) : monoid (list A) :=
  mkMonoid (empty:=nil) (mappend:=app) _ _ _.
```

We shall now define the Finger Trees over some monoid and measure. We will see in section 5 how particular instantiations of these will give different structures.

Section *DependentFingerTree*.

```
Variables (v : Type) (mono : monoid v).
```

The variable  $v$  is the monoid carrier type, *mono* is the monoid implementation. As before, we use  $\varepsilon$  and  $\cdot$  for the monoid empty element and operation respectively.

### 4.2 Nodes

Finger Trees are implemented using 2-3 nodes as in 2-3 trees, with the addition of a cached value that stores the combined measure of the subobjects.

Section *Nodes*.

```
Variables (A : Type) (measure : A → v).
Notation "'||' x '||'" := (measure x).
```

Intuitively, the measure represents an abstraction of the sequence of objects in the node (later, in the tree) considered from left to right. Taking the list monoid previously defined and the singleton measure  $[-]$  gives the most concrete abstraction of this sequence of elements: the sequence itself. The *measure* function is denoted by  $\| \_ \|$  in the following, when possible (notations local to a function are not supported). We now define 2-3 nodes with cached measures:

Inductive node : Type :=  
 | Node2 :  $\forall x y, \{ s : v \mid s = \| x \| \cdot \| y \| \} \rightarrow \text{node}$   
 | Node3 :  $\forall x y z, \{ s : v \mid s = \| x \| \cdot \| y \| \cdot \| z \| \} \rightarrow \text{node}$ .

We use a subset type here to specify the invariant on the cached value. This invariant could not be checked using simple types, because it is dependent on the values carried by the node. Moreover, it uses the *measure function* as a *datatype parameter* so it would require an additional abstraction mechanism like type classes or functors to achieve it in a regular ML-based functional language. A regular dependent product is used here instead.

We have smart constructors *node2* and *node3* which compute the measure, e.g:

Program Definition *node2* ( $x y : A$ ) : node :=  
 Node2  $x y (\| x \| \cdot \| y \|)$ .

The generated obligations are trivially true as they are of the form  $x = x$ . Correspondingly, *node\_measure* pulls the cached measure.

Program Definition *node\_measure* ( $n : \text{node}$ ) : v :=  
 match  $n$  with Node2 \_ \_  $s \Rightarrow s$  | Node3 \_ \_ \_  $s \Rightarrow s$  end.  
 End *Nodes*.

Although it may seem that the *node\_measure* function is independent of the *measure* function, it is not. Witness its type after we closed the section:

$\forall (A : \text{Type})(\text{measure} : A \rightarrow v), \text{node\_measure} \rightarrow v$

The node datatype itself is parameterized by the measure function, hence all operations on nodes take it as an implicit parameter. Had we not added the invariant, we would not need this parameter, but we could not prove much about node measures either, because they might be any element of type  $v$ . We could define an inductive invariant on nodes asserting that their measures were built using the measure function coherently, but we would need to show that this invariant is preserved in each of the functions, while in our case we get it for free by typing.

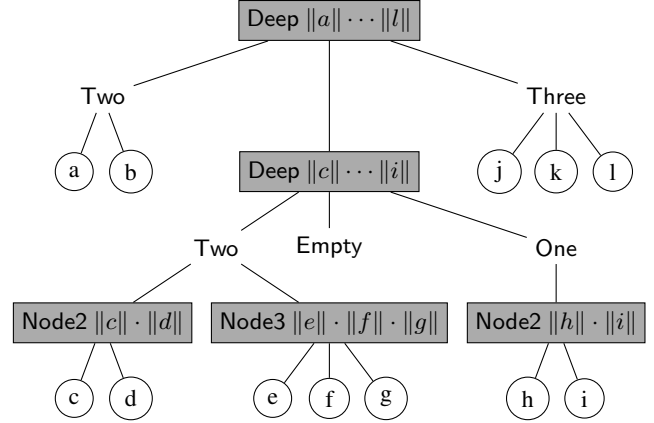
### 4.3 The fingertree datatype

Before presenting the definition of fingertree in COQ, we introduce the original HASKELL datatype and justify why the direct translation would be unsatisfactory. In HASKELL, the FingerTree algebraic datatype is defined as:

```
data FingerTree v a = Empty | Single a
  | Deep v (Digit a) (FingerTree v (Node v a)) (Digit a)
```

The Empty constructor represents the empty tree, the Single constructor builds a singleton tree from an element of  $a$  and the Deep constructor is the branching node. It builds a FingerTree  $v a$  from a cached measure of type  $v$ , two digits of  $a$  and a middle FingerTree  $v$  (Node  $v a$ ) of 2-3 nodes of  $a$  with measures of type  $v$ . This is where the nesting occurs; it is best described by a picture (figure 8).

We could directly define the Finger Tree data structure in COQ by translating the HASKELL definition. However doing so would



**Figure 8.** A sample FingerTree. At the first level we have a Deep node with a 2-digit of  $a$  at the left and a 3-digit at the right. The middle tree is a FingerTree  $v$  (Node  $v a$ ) consisting of a Deep node with an Empty middle tree. It has a 2-digit of Node  $v a$  at its left and a 1-digit at its right. The circled nodes represent elements of type  $a$  and the shaded ones indicate loci where cached measures are stored

cause a rather subtle problem: the cached measures in two Finger Trees could be computed by two *different* measure functions but these trees would still have the same type because the FingerTree datatype is parameterized only by the type of the computed measures and not by the measure function. Again, we could live with it and have a predicate formalizing the fact that the measures appearing in a FingerTree were built coherently with some measure function, but preservation of this invariant would have to be proved for each new definition on Finger Trees. We prefer to view the measure function as an additional parameter of the Finger Tree datatype. In this case coherence is ensured by the type system because two Finger Trees will have different types if they are built from different measure functions. Hence the following definition:

A fingertree inductive is parameterized by a type  $A$  and a measure function  $\text{measure} : A \rightarrow v$  on this type. Each fingertree object is also indexed by its corresponding measure:

- Empty builds the empty tree of measure  $\varepsilon$ ;
- Single  $x$  builds the tree with sole element  $x$  of measure  $\|x\|$ ;
- Deep  $pr\ ms\ m\ sf$  builds the tree of prefix  $pr$ , subtree  $m$  of measure  $ms$  and suffix  $sf$ . Its measure is given by combining the measures of its subtrees. The argument  $ms$  will be implicit when constructing Deep nodes, as it can be inferred from the type of  $m$ . We place this argument  $ms$  just before the middle Finger Tree contrary to the original datatype in which the cached measure is the first field and stores the measure of the *whole* tree whereas for us the latter is given by the index.

We present the definition using inference rules for enhanced readability, omitting  $A$  and  $\text{measure}$  in the premises:

$$\frac{}{\text{Empty} : \text{fingertree } A \text{ measure } \varepsilon}$$

$$\frac{x : A}{\text{Single } x : \text{fingertree } A \text{ measure } (\text{measure } x)}$$

$$\frac{pr, sf : \text{digit } A, ms : v}{m : \text{fingertree } (\text{node\_measure } A) (\text{node\_measure } \text{measure}) ms}$$

$$\frac{}{\text{Deep } pr\ ms\ m\ sf : \text{fingertree } A \text{ measure } (\text{digit\_measure } \text{measure } pr \cdot ms \cdot \text{digit\_measure } \text{measure } sf)}$$

This inductive family is indexed by values of type  $v$ . Why we need a dependent type follows from a simple observation. If we want the cached measure on the Deep node and the invariant that the measure is really the one of the middle tree, we must have a way to refer to the measure of the middle tree, but we are actually defining trees, so we cannot recurse on them at this time. Note also that we need *polymorphic* recursion to go into the middle subtrees of Deep nodes.

**Adding elements** We can add an element  $a$  to the left of a tree  $t$  of measure  $s$  to get a tree with measure  $\|a\| \cdot s$ . Due to polymorphic recursion, our functions will now always have  $A$  and *measure* arguments as they are *real* arguments which will change during recursive calls. If  $t$  is empty, a singleton or a tree with a left prefix which is not full, then we simply put  $a$  at the leftmost position of the tree. Otherwise, we need to reorganize the tree to make space at the left for  $a$ , and this requires recursing polymorphically to add a node  $a$  to the left of the middle subtree in the Deep node.

```

Program Fixpoint add_left (A : Type) (measure : A → v)
  (a : A) (s : v) (t : fingertree measure s) {struct t} :
  fingertree measure (measure a · s) :=
  match t in fingertree _ s
  return @fingertree A measure (measure a · s) with
  | Empty ⇒ Single a
  | Single b ⇒ Deep (One a) Empty (One b)
  | Deep pr st' t' sf ⇒
    match pr with
    | Four b c d e ⇒
      let sub := add_left (node3 measure c d e) t' in
      Deep (Two a b) sub sf
    | x ⇒ Deep (add_digit_left a pr) t' sf
  end
end.

```

The first match expression uses dependent elimination. Its meaning is that from a particular fingertree of measure  $s$  each branch will build a fingertree of measure  $measure\ a \cdot s$  where  $s$  has been substituted by the measure of the matched pattern. For example in the first branch we must build an object of type `fingertree measure (measure a · ε)`. However, the right hand side `Single a` has type `fingertree measure (measure a)`, hence we must use the inductive equivalence rule (figure 6), which generates an obligation  $\vdash measure\ a \cdot \varepsilon = measure\ a$ , easily solved using the right identity of the monoid. The `in` and `return` clauses are mandatory in COQ due to undecidability of higher-order unification, but they can be dropped in RUSSELL, in which case substitution is replaced by equational reasoning. Had we dropped the clauses, we would have the equation  $s = \varepsilon$  in the context of the first branch hence the generated obligation  $s = \varepsilon \vdash measure\ a \cdot s = measure\ a$ . This one would be solved by first substituting  $s$  by  $\varepsilon$  in the goal then applying the right identity ; we just postponed the substitution.

The nested match expression uses an alias  $x$  for capturing prefixes which are not `Four` and applies the partial function `add_digit_left` defined earlier. There is an application of the subset equivalence rule here, which generates an obligation to show that  $pr$  is not full. It can be solved because we have  $\forall b\ c\ d\ e, pr \neq Four\ b\ c\ d\ e$  in the context.

It is an essential property of this function that the measure is preserved. To see that, let us instantiate the Finger Tree by the list monoid  $([], ++)$  and measure  $[-]$  defined earlier. You can check here that adding to the left is prepending the measure of the element to the measure of the Finger Tree, hence consing to the sequence of elements of the tree with the list monoid interpretation. For each of the following operations, this correspondence with the list instantiation of the measure will hold.

#### 4.4 The view from the left... of a fingertree

We will now construct views on the Finger Trees which decompose a tree into its leftmost element and a remaining Finger Tree (`View_L`) (or dually, the rightmost one and the "preceding" Finger Tree). It serves to abstract from the implementation and give a list-like view of fingertree, which is used to write functions which recurse on fingertrees without having to deal with the intricacies of the structure (c.f. figure 5.1). The `View_L` inductive is a little less polymorphic, as it does not need carry the measure function which views ignore. However `View_L` still stores the measure  $s$  of the rest of the tree in the `cons_L` constructor ( $s$  will be implicit), hence we need to abstract by the sequence's type *seq* which is indexed by  $v$ . It will be instantiated by `fingertree measure`.

```

Inductive View_L (A : Type) (seq : v → Type) : Type :=
| nil_L : View_L A seq
| cons_L : A → ∀ s, seq s → View_L A seq.

```

Such a view will be constructed by the `view_L` function, by structural (polymorphic) recursion on the fingertree. We can seamlessly use the `digit_tail` function which is partial (it accepts only digits which are not `One`) and need not add any type annotations when calling `view_L` recursively on  $t'$ . Note that we use a partial type application (`fingertree measure`) in the return type, which is perfectly legal.

```

Program Fixpoint view_L A (measure : A → v)
  (s : v) (t : fingertree measure s) {struct t} :
  View_L A (fingertree measure) :=
  match t with
  | Empty ⇒ nil_L
  | Single x ⇒ cons_L x Empty
  | Deep pr st' t' sf ⇒
    match pr with
    | One x ⇒
      match view_L t' with
      | nil_L ⇒ cons_L x (digit_to_tree _ sf)
      | cons_L a st' t' ⇒
        cons_L x (Deep (node_to_digit a) t' sf)
      end
    | y ⇒ cons_L (digit_head y) (Deep (digit_tail y) t' sf)
    end
  end.

```

We can show that `view_L` preserves the measure of the input tree; had we indexed `View_L` by the measure of the input tree, these generation lemmas would have appeared as obligations.

```

Lemma view_L_nil : ∀ A (measure : A → v) s
  (t : fingertree measure s), view_L t = nil_L → s = ε.

```

```

Lemma view_L_cons : ∀ A (measure : A → v) s
  (t : fingertree measure s) x st' t',
  view_L t = cons_L x (s:=st') t' → s = measure x · st'.

```

##### 4.4.1 Dependence hell

Our views are useful to give a high-level interfaces to Finger Trees, but in their current state they are very limited as we can write only non-recursive definitions on views. That is, we will not be able to convince COQ that functions defined by recursion on the view of a tree are just as valid as those defined by recursion on the tree itself. To do that, we must have a measure on our Finger Trees, e.g. their number of elements (named *tree\_size*), from which we can trivially build a measure on the views `View_L_size`. Then it is only a matter of showing that for all  $t$ , `view_L t` returns a view of size `tree_size t` to prove that a recursive call on the tail of a view is correct (c.f. §5.1).

However, doing such a thing is not as easy as it looks because  $view\_L$  manipulates objects of dependent types and reasoning about them is somewhat tricky. An essential difficulty is that the usual Leibniz equality is not large enough to compare objects in dependent types as they may be comparable but in different types. A simple example is that the proposition  $vnil = vcons\ x\ n\ v$  is not well-typed because  $vnil$  is of type  $vector\ 0$  and  $vcons\ x\ n\ v$  of type  $vector\ (S\ n)$  which are not convertible.

In our case, we want to say that an arbitrary tree  $t$  of measure  $s$  with  $view\ nil\_L$  must be the Empty tree, but those two trees do not have the same type. We apply the usual trick of heterogeneous equality (McBride 1999) : prove they must be in the same type. The inductive  $JMeq$  defines an heterogeneous equality (previously denoted by  $\simeq$ ) in COQ. It is used to compare objects which are not of the same type. Its sole constructor is  $JMeq\_refl$  of type  $\forall A\ a, JMeq\ A\ a\ A\ a$ . The point of this admittedly strange notion of equality is to postpone the moment at which we need to prove that the types of the compared objects are equal. When we arrive at this point, we can apply the  $JMeq\_eq$  axiom of type  $\forall A\ x\ y, JMeq\ A\ x\ A\ y \rightarrow x = y$  to get back a regular equality between the objects.

In the first lemma, we compare  $t$  of measure  $s$  with Empty of measure  $\varepsilon$ ; clearly, replacing  $JMeq$  by regular equality would yield a type error here.

Lemma  $view\_L\_nil\_JMeq\_Empty : \forall A\ (measure : A \rightarrow v)\ s$   
 $(t : fingertree\ measure\ s), view\_L\ t = nil\_L \rightarrow$   
 $JMeq\ t\ (Empty\ (measure := measure))$ .

Now that we have shown the equality for a general  $s$  index, we can instantiate it with a particular one, i.e.  $\varepsilon$ . Given that  $t$  is now of measure  $\varepsilon$ , we can use the regular equality with Empty.

Lemma  $view\_L\_nil\_Empty : \forall A\ (measure : A \rightarrow v)$   
 $(t : fingertree\ measure\ \varepsilon), view\_L\ t = nil\_L \rightarrow t = Empty$ .

These technical lemmas about  $nil\_L$  and Empty will help us build our measure. Indeed, they are needed to prove the following:

Lemma  $view\_L\_size : \forall A\ (measure : A \rightarrow v)\ (s : v)$   
 $(t : fingertree\ measure\ s), View\_L\_size\ (view\_L\ t) = tree\_size\ t$ .

This gives us a decreasing measure on  $view\_L$  results. We will use it when building the instances.

Lemma  $view\_L\_size\_measure : \forall A\ (measure : A \rightarrow v)\ (s : v)$   
 $(t : fingertree\ measure\ s)\ x\ st' (t' : fingertree\ measure\ st')$   
 $view\_L\ t = cons\_L\ x\ t' \rightarrow tree\_size\ t' < tree\_size\ t$ .

We can also define  $deep\_L$ , the smart constructor which rearranges a tree given a potentially empty digit for the prefix. It is a dependent version of the internal function for the Deep case of  $view\_L$ , which is used when splitting Finger Trees. The code has not changed at all from the non-dependent version hence we do not show it. However, the specification would greatly benefit from an overloading mechanism to factor the different  $measure$  specializations.

Program Definition  $deep\_L\ (A : Type)\ (measure : A \rightarrow v)$   
 $(d : option\ (digit\ A))\ (s : v)$   
 $(mid : fingertree\ (node\_measure\ measure)\ s)$   
 $(sf : digit\ A) : fingertree\ measure\ (option\_digit\_measure\ measure\ d \cdot s \cdot digit\_measure\ measure\ sf) := \dots$

#### 4.4.2 Back to normal

We are now ready to implement regular deque viewing operations on our Finger Trees. We do not need recursion on the trees here, so we can define the operations for a fixed element type, measure function and measure index.

Section *View*.

Variables  $(A : Type)\ (measure : A \rightarrow v)\ (s : v)$ .

We define an  $isEmpty$  predicate for some operations will be partial, i.e. defined only for non-empty fingertree objects. We do not match directly on the object to maintain the abstraction from the implementation.

Definition  $isEmpty\ (t : fingertree\ measure\ s) :=$   
 $match\ view\_L\ t\ with\ nil\_L \Rightarrow True\ |\_ \Rightarrow False\ end$ .

Emptiness is decidable.

Definition  $isEmpty\_dec\ (t : fingertree\ measure\ s) :$   
 $\{isEmpty\ t\} + \{\neg isEmpty\ t\}$ .

The obvious operations are definable, we show the  $liat$  operation dual to  $tail$ . Here we return the index along with the fingertree in a dependent pair of type  $\{s : v \& fingertree\ measure\ s\}$ , which corresponds more closely to the view that the cached measure is really data, not only an index used to refine a datatype. The pairing of a tree  $t$  with its measure  $m$  is denoted  $m :| t$  and reads "m measures t".

Program Definition  $liat\ (t : fingertree\ measure\ s\ |$   
 $\neg isEmpty\ t) : \{s' : v \& fingertree\ measure\ s'\} :=$   
 $match\ view\_R\ t\ with$   
 $| nil\_R \Rightarrow !\ | cons\_R\ st'\ t'\ last \Rightarrow st' :| t'$   
 $end$ .

End *View*.

#### 4.5 Catenation & splitting, dependently

We can also define catenation using a dependent type which clearly states the function's specification. The implementation is the same as the one of Hinze & Paterson, except we proved the 100 obligations generated by PROGRAM concerning measures. The five mutually recursive functions hidden here which define  $app$  have the particularity of being quite big because they implement a specialization of concatenation for each combination of the digits at the right of the left tree and the left of the right tree. Hinze & Paterson even recommend to produce the 200 lines long definition automatically to prevent mistakes. Here the definition is checked directly: the type itself expresses that we did not mess up the obvious property of  $app$  on measures.

Definition  $app\ (A : Type)\ (measure : A \rightarrow v)$   
 $(xs : v)\ (x : fingertree\ measure\ xs)$   
 $(ys : v)\ (y : fingertree\ measure\ ys) :$   
 $fingertree\ measure\ (xs \cdot ys)$ .

Notation " $x ++ y$ " :=  $(app\ x\ y)$  (at level 20).

The last and most interesting operation from the specifier's point of view, is splitting a tree by a predicate on its measure. We begin by splitting a node.

Section *Nodes*.

Variables  $(A : Type)\ (measure : A \rightarrow v)$ .  
Variables  $(p : v \rightarrow bool)\ (i : v)$ .

We split a node  $n$  by finding where some predicate  $p$  on measures turns to true, starting with an initial measure  $i$  and accumulating the measures of subtrees from left to right. We simply copy the invariant given by Hinze and Paterson (2006) here, there is nothing to change, we only add the corresponding property on the measure which clearly generalizes the  $to\_list$  equation. The code is also a direct translation from HASKELL to RUSSELL except we use a regular tuple instead of a custom split datatype. We have also defined a  $split\_digit$  function with the same specification.



```

Program Definition split_node (n : node measure) :
{ (l, x, r) : option (digit A) × A × option (digit A) |
  let odm := option_digit_measure measure in
  node_measure measure n = odm l · || x || · odm r ∧
  node_to_list n = option_digit_to_list l ++ [x] ++
  option_digit_to_list r ∧
  (l = None ∨ p (i · ls) = false) ∧
  (r = None ∨ p (i · ls · || x ||) = true)} := ...

```

End Nodes.

The interesting case is for trees, as usual. Instead of returning a tuple, we use a dependent `tree_split` inductive which directly captures the invariant that a split is a decomposition of a `fingerTree` preserving measures. We also put the invariants on the left and right trees, which can be used by clients to prove properties about their code. This is one of the most distinctive parts of this work: we derive not only reusable code but also reusable proofs using rich types. Indeed, `split_tree` can be seen as a proof combinator, turning a property on a measure and monoid to a property on words on this monoid.

Section Trees.

Variable  $p : v \rightarrow \text{bool}$ .

```

Inductive tree_split (A : Type) (measure : A → v)
(i : v) : v → Type :=
mkTreeSplit : ∀ (xsm : v) (xs : fingerTree measure xsm |
  isEmpty xs ∨ p (i · xsm) = false)
(x : A) (ysm : v) (ys : fingerTree measure ysm |
  isEmpty ys ∨ p (i · xsm · measure x) = true),
tree_split measure i (xsm · measure x · ysm).

```

This inductive combines both subsets and dependent inductive types, but we can still write our code as usual. Approximately 100 lines of proof are necessary to discharge the generated obligations.

```

Program Fixpoint split_tree (A : Type) (measure : A → v)
(i s : v) (t : fingerTree measure s | ¬ isEmpty t)
: tree_split measure i s := ...

```

End Trees.

This concludes our implementation of (dependent) Finger Trees with PROGRAM. To sum up, we have proven that:

- All functions are total, once properly annotated with their pre-conditions.
- All functions respect the measure semantics and use them coherently.
- All functions respect the invariants given in the paper by Hinze & Paterson.

The complete file we walked through comprises 600 lines of specifications (declarations, types, programs, proof statements) and 450 lines of proof. In comparison the source file provided by Hinze & Paterson is 650 lines including comments. We argue that this is proof of the scalability of RUSSELL as a programming language.

## 5. Instances

We will now instantiate the `fingerTree` implementation by particular monoids and measures to get higher-level structures. Obviously, we can instantiate the Finger Trees just like in the original paper of Hinze & Paterson using weak, simply typed specifications and extract the code, but that would not be too exciting. Nevertheless, we will first present a way to recover the simply-typed interface and instantiate it to ordered sequences. We will then focus on proving a particular specialization using the dependent interface directly.

### 5.1 FingerTree: a non-dependent interface

As shown before, we can wrap the measure and the tree in a dependent pair to offer a simpler interface to Finger Trees.

Section FingerTree.

Variables  $(v A : \text{Type})$  ( $\text{mono} : \text{monoid } v$ ) ( $\text{measure} : A \rightarrow v$ ).

```

Definition FingerTree :=
{ m : v & fingerTree mono measure m }.

```

This gives us a type `FingerTree v mono A measure`, which can be compared with the original `FingerTree v a` datatype in HASKELL. It adds the monoid implementation and `measure` function as explicit parameters whereas in HASKELL they are passed implicitly using the typeclass mechanism. Morally, the parameterization we did using the section mechanism correspond to the systematic prefixing of functions on `FingerTree` objects by a `Measured v a` constraint in HASKELL (`Measured v a` is a class with a single function `measure` of type  $a \rightarrow v$  where  $v$  must have a `Monoid` instance). We think this equivalence indicates an interesting field of investigations on how to integrate overloading in a language like RUSSELL, or a section system in HASKELL.

We do not describe the wrapping in detail, it is just destructuring a `FingerTree` object, calling the appropriate function on `fingerTrees` and packing back the measure and the tree. However, an important point is that we can wrap views too and obviously recover the associated measure:

```

Inductive left_view : Type :=
| nil_left : left_view | cons_left : A → FingerTree → left_view.

```

```

Lemma view_left_size : ∀ t x t', view_left t = cons_left x t' →
tree_size t' < tree_size t.

```

We can finally wrap splitting on `FingerTrees` along a predicate. We offer the same simply-typed interface as HASKELL by defining the function `split_with p x`. It splits any tree  $x$  by first checking if  $x$  is empty and if not calls the split function on it using the predicate  $p$  with an empty accumulator.

```

Program Definition split_with (p : v → bool)
(x : FingerTree) : FingerTree × FingerTree :=
if is_empty_dec x then (empty, empty)
else let (l, x, r) := split p ε x in (l, cons x r).

```

End FingerTree.

As an example use of the simply-typed interface we develop the specialization of Finger Trees to Ordered Sequences. The main idea is to use as a measure on elements of  $A$  the element itself and have the monoid implement an operation which returns the rightmost element in the tree. Then if the operations maintain the invariant that the sequence of elements in the tree are ordered we have an ordered sequence where the maximal element can be retrieved in constant time.

```

Module OrdSequence(AO : OrderedType).

```

We use modules instead of the section mechanism as it gives compile-time instantiation which is exactly what we need here. The module system of COQ is a superset of OCAML's. We suppose given an implementation of an ordered type  $AO$ .

```

Definition A : Type := AO.t.
Definition measure (x : A) := Some x.

```

We measure one element by itself.

```

Module KM := KeyMonoid AO. Import KM.

```

The *KeyMonoid* module implements the monoid (*None*, *keyop*) on the type option *A*, where *keyop* is defined by:

```
keyop x None      = x
keyop _ (Some _ as y) = y
```

Clearly this will give the rightmost element of the structure as its measure if there is one. It also defines the function *key\_gt* which lift the operation of the ordered type to the monoid's type. We can then declare the ordered sequence type as *FingerTree* applied to the monoid and measure.

**Definition** *OrdSeq* := *FingerTree m measure*.

We can get the maximal element in constant time by taking the measure of the whole *FingerTree*.

**Definition** *max* (*x* : *OrdSeq*) := *tree\_measure x*.

We skip the other operations defined in the original paper which permit to partition an ordered sequence or insert or delete in it, they are no more difficult to code. We focus on the merge operation which is the only termination challenge in the whole development. Merging of ordered sequences simply does a merge sort of the list of elements of both sequences. It does so using the left view defined earlier and hence needs to recurse on results of the *view\_left* function. Note however that the termination argument is complicated by the swapping of arguments for the recursive call. We can use *pair\_size* measure defined below to define the function by well-founded recursion on the pair of sequences. The generated obligations can be solved using the *view\_left\_size* lemma defined earlier.

**Definition** *pair\_size* (*x* : *OrdSeq* × *OrdSeq*) :=  
let (*l*, *r*) := *x* in *tree\_size l* + *tree\_size r*.

**Program Fixpoint** *merge* (*x* : *OrdSeq* × *OrdSeq*)  
{measure *pair\_size*} : *OrdSeq* :=  
dest *x* as (*xs*, *ys*) in  
match *view\_left xs* with  
| *nil\_left* ⇒ *ys*  
| *cons\_left x xs'* ⇒  
 dest *split\_with* (fun *y* ⇒ *key\_gt y (measure x)*) *ys*  
 as (*l*, *r*) in *cat l (cons x (merge (r, xs')))*  
end.

## 5.2 Dependent Sequences.

We will now define random-access sequences as a specialization of *fingertree*. This structure provides insertion and deletion at both ends in constant amortized time, concatenation and splitting in logarithmic time and access (*get*) and mutation (*set*) operations in logarithmic time too. We will build a certified implementation of this structure by showing that the operations respect a functional specification. This method slightly differs from the usual separation of operation definition and invariant proving. Here, invariant preservation will be proved simultaneously with the definition of operations. We define sequences over a type of elements *A*.

**Section** *DependentSequence*.

**Variable** *A* : Type.

First, a useful definition: the type *below i* represents the naturals below some *i*, so *below 0* is not inhabited.

**Definition** *below i* := { *x* : nat | *x* < *i* }.  
**Lemma** *not\_below\_0* : *below 0* ⇒ False.

We use naturals to measure sequences by their length. Our measure is a bit special as it also carries a function giving the map realized

by the *FingerTree*. It is used solely for specification and could be removed in extracted code using a dead code analysis or even rejected at extraction using a finer type system.

**Definition** *v* := { *i* : nat & (*below i* ⇒ *A*) }.

We define a notation for our measure objects: an object *i* > *m* is a dependent pair composed by an index *i* and a map *m* from naturals below *i* to elements of *A*. The epsilon will represent empty sequences. As it contains no elements, no value is returned by the function. We have an obligation to show that in a context with an hypothesis of type *below 0*, we can prove False: we just use *not\_below\_0*.

**Program Definition** *ε* : *v* := 0 > (fun \_ ⇒ !).

Appending two maps requires some reindexing: the new map is formed by projecting the new index to either the first or the second map.

**Program Definition** *append* (*xs ys* : *v*) : *v* :=  
let (*n*, *fx*) := *xs* in let (*m*, *fy*) := *ys* in  
(*n* + *m*) > (fun *i* ⇒ if *less\_than i n* then *fx i* else *fy (i - n)*).

We can build a monoid *seqMonoid* from these operations, we skip the proofs which are relatively easy.

The measure of a singleton sequence built from *x* gives the constant map to *x*.

**Program Definition** *single* (*x* : *A*) : *v* := 1 > (fun \_ ⇒ *x*).

We define an abbreviation *seq* for our sequence type. Sequences are a specialization of *fingertree* with the *seqMonoid* monoid and *single* measure defined above. Hence an object of type *seq (i > m)* is a *FingerTree* representing a sequence of objects of length *i* given by the map *m*.

**Definition** *seq* (*x* : *v*) := *fingertree seqMonoid single x*.

We can get the length of the sequence in constant time as it is part of the cached measure.

**Program Definition** *length* (*s* : *v*) (*x* : *seq s*) : nat :=  
dest *s* as *size* > \_ in *size*.

The constructor *make n x* creates a sequence of length *n* with value *x* in each cell. Note that obligations are generated to prove that e.g., when *make n x* is of type *seq (n > fun \_ ⇒ x)* then *add\_left x (make n x)* is of type *seq (S n > fun \_ ⇒ x)*. We prove the preservation of semantics of our sequence's *make* operation directly here.

**Program Fixpoint** *make* (*i* : nat) (*x* : *A*) {struct *i*} :  
*seq (i > (fun \_ ⇒ x))* :=  
match *i* with  
| 0 ⇒ *empty*  
| S n ⇒ *add\_left x (make n x)*  
end.

We now define the strongly specified *get* function which gets the *j*th element of a sequence *x* of length *i*. We ensure that no out-of-bound access is made because *j* is of type *below i* and we assert that *get* indeed returns the right value of the map *m*. Note that *m* is never used in the code, only in the types. Again, some obligations are generated to prove that this specification is indeed a correct abstraction of the semantics of the code. It is crucial that the *split\_tree* function be dependent and return a *tree\_split* object carrying proofs to solve these obligations. As before, the code stays simple and compact. However, we skip around 50 lines of proof needed to discharge the obligations. The boolean predicate *lt\_idx i x* tests whether an index *i* is less than the first component of a measure *x*.

**Program Definition**  $get\ i\ m\ (x : seq\ (i > m))$   
 $(j : below\ i) : \{ value : A \mid value = m\ j \} :=$   
 $dest\ split\_tree\ (lt\_idx\ j) \varepsilon\ x$   
 $as\ mkTreeSplit\ ls\ l\ x\ rs\ r\ in\ x.$

The natural companion of  $get$  is  $set$  which given a sequence  $x$  of length  $i$  sets its  $j$ th cell to  $value$ . We have as precondition that  $j$  is below  $i$  and as postcondition that the new sequence has  $value$  at its  $j$ th index and the same elements as  $x$  otherwise. The function denoted by  $=n$  decides equality on naturals.

**Program Definition**  $set\ i\ (m : below\ i \rightarrow A)\ (x : seq\ (i > m))$   
 $(j : below\ i)\ (value : A) :$   
 $seq\ (i > (\text{fun}\ idx \Rightarrow \text{if}\ idx = n\ j\ \text{then}\ value\ \text{else}\ m\ idx)) :=$   
 $dest\ split\_tree\ (lt\_idx\ j) \varepsilon\ x\ as\ mkTreeSplit\ ls\ l\_rs\ r$   
 $in\ add\_right\ l\ value\ ++\ r.$

Now that we have defined our operations with built-in invariants, it is a simple matter of using the type information to relate operations. Here we formalize how  $get$  and  $set$  behave together. The proofs use solely the types of  $get$  and  $set$ , not their code. We show that our sequences respect the two axioms of functional arrays: if getting at an index  $j$  just set to  $value$  we get  $value$  ( $get\_set$ ), otherwise we get the value from the original map ( $get\_set\_diff$ ).

**Program Lemma**  $get\_set : \forall\ i\ m\ (x : seq\ (i > m))$   
 $(j : below\ i)\ (value : A),\ value = get\ (set\ x\ j\ value)\ j.$

We were able to state this lemma only because we are using the RUSSELL type-checker, indeed there is an automatically inserted coercion at the right hand side of the equality to go from a subset of  $A$  to an object of  $A$  (the implicit type of an equality is fixed by its first argument). This seamless integration of PROGRAM with COQ as a theorem prover is also an improvement over a previous attempt to provide a programming environment in COQ, namely the Program tactic by Parent (1995). However, we are not yet able to get the most natural statements in RUSSELL. In the next definition for example, we coerce objects  $j$  and  $k$  of type  $below\ i$  to their underlying  $nat$  components and the results of  $get$  to  $A$ . This is because the comparison of objects from subset types only makes sense on their first component, the witness, whereas comparing proofs is useless. Our pragmatic solution is to let the user insert a cast to get the expected behavior, but in a proof-irrelevant Type Theory the problem disappears because comparison of proofs become trivial. We discuss this solution in §7.1.

**Program Lemma**  $get\_set\_diff : \forall\ i\ m\ (x : seq\ (i > m))$   
 $(j : below\ i)\ (value : A)\ (k : below\ i),$   
 $(j : nat) \neq k \rightarrow (get\ x\ k : A) = get\ (set\ x\ j\ value)\ k.$

Finally, we define splitting of sequences. It requires creating two projections of the original map, hence the following:

**Program Definition**  $split\_l\ (i : nat)\ (j : below\ i)$   
 $(idx : below\ j) : below\ i := idx.$

**Program Definition**  $split\_r\ (i : nat)\ (j : below\ i)$   
 $(idx : below\ (i - j)) : below\ i := j + idx.$

We can finally split a sequence  $x$  at a particular index  $j$ , returning two sequences whose maps are just projections of the splitted  $x$ . Approximately a hundred lines of proof are necessary to discharge the three obligations generated by PROGRAM.

**Program Definition**  $split\ i\ m\ (x : seq\ (i > m))\ (j : below\ i) :$   
 $seq\ (j > (m \circ split\_l\ j)) \times seq\ ((i - j) > (m \circ split\_r\ j)) :=$   
 $dest\ split\_tree\ (lt\_idx\ j) \varepsilon\ x\ as\ mkTreeSplit\ ls\ l\ x\ rs\ r\ in$   
 $(l,\ add\_left\ x\ r).$

End *DependentSequence*.

This gives us a certified implementation of sequences in less than 500 lines whose dependent interface can be used directly to implement programs and prove properties at the same time.

## 6. Extraction

We can extract the code we just certified to both HASKELL and OCAML (bypassing the type checker for polymorphic recursion) to get certified implementations of Finger Trees and sequences. We are guaranteed to get the same algorithmic code than what we wrote using PROGRAM. This is an important aspect of developing with RUSSELL, because usually when developing programs with rich specifications in COQ one tends to use the proof language as much as the programming language and this can have devastating results on the extracted code. Our method enforces a distinction between code and proofs that is healthy in this respect.

Unsurprisingly, our extracted Finger Trees have indeed the same performance as the original implementation: they have the same code. The `Data.Sequence` implementation in HASKELL is in fact a specialization of the `FingerTree` implementation to a particular monoid and measure, avoiding dynamic lookups and unnecessary boxing. Had we used the module system of COQ to parameterize our development instead of the section mechanism, we would have had this instantiation at no cost. However the lack of a powerful module system in HASKELL prevented us from doing that, as we would not be able to extract to it then.

Unfortunately, extraction of our dependent sequences will not give excellent results because the measure contains a function giving the map which will be extracted even though it has no algorithmical role. We are in a case where only part of an index need be stored (unlike the setting of Brady et al. (2003)), so a finer distinction than `Prop/Type` has to be found to distinguish between algorithmical and non-algorithmical content. Current work by B. Barras and B. Bernardo on an adaptation of the Calculus of Implicit Constructions (Miquel 2001) as the core calculus of a dependently-typed language ought to give the expressivity we seek.

We conclude that RUSSELL is not overly verbose and that doing the proofs is actually not an insurmountable task, knowing that most of them were solved automatically using a normalization tactic for monoid expressions.

## 7. Discussion

### 7.1 Proof-Carrying Code

As displayed before, COQ ultimately melts proof and code in its terms, so as to be able to check the correctness of any part of a term at any time. On the surface, we can manage to separate the two activities of coding and proving using PROGRAM, but there are parts of the system where it still bites. The same notable problem appear during proofs and computations. When proving, we are faced with terms containing proofs we do not want to name or manipulate and when computing (in the call-by-value case only) the system itself is stuck with irreducible opaque proofs. However, in the two cases the proofs appear only in parts of the term that should not be reduced because they have no algorithmical content and need not be manipulated as their structure has no importance. This is formalized by proof-irrelevance (PI), the statement that two proofs of the same proposition are equal. The meta theory of the Calculus of Constructions extended with PI has been studied extensively by Werner and Miquel. Recently, Werner (2006) also gave a method to implement it effectively in COQ, which is currently pursued by the author. It has also been studied in the context of Observational Type Theory, the (would be) core calculus of EPIGRAM 2. Having PI in the system would make these two problems disappear and would probably improve performance of computations inside COQ too.

### 7.2 A note on the implementation

PROGRAM is composed of two distinct parts: the type-checker and the proof-obligation handling machinery. In fact only a small part of the COQ type-checker needed to be changed, so RUSSELL's

type-checker is just a different instantiation of the typing functor of COQ which is parameterized by the conversion algorithm. This is to contrast with the work by Catherine Parent on the previous PROGRAM tactic which was a completely different solution for program-synthesis in COQ. The system was not integrated with COQ and so went unmaintained.

### 7.3 Related Work

RUSSELL can be seen as one point in the design space of programming languages with dependent types. Its most distinctive feature is the separation of code from proof, à la PVS, and its treatment of subsets and indexed data types. It is also original in the sense that it *elaborates* into a well-studied, safe and mature language while most other solutions coming from the programming languages community have no such safety net or only for part of the system. RUSSELL could in fact be implemented in EPIGRAM or AGDA rather easily as they are based on roughly the same foundations as COQ. The main difference is that COQ already has a comprehensive standard library and a tactic system. Compared to programming languages like ATS, CONCOPTION (Fogarty et al. 2007) or HASKELL, RUSSELL has no effects and only well-founded recursion. We found that in practice the programs we write usually have relatively simple termination arguments and real non-termination could be accommodated by using a coinductive encoding. We think having a layered approach to treating computational effects using functional encodings like monads is a more disciplined approach to development and certification than trying to design a system with some computational effects built-in, if this is at all possible while keeping strong properties.

Other certifications of tree-based data structures include a certification of the AVL trees used in OCAML and also red-black trees in COQ (Filliâtre and Letouzey 2004). The development was done by first building a dependent interface and providing a simpler one on top of it. The second author reported great success adapting the dependently-typed implementation which was originally done using tactics to PROGRAM.

## 8. Conclusion

What we would like to assess is that writing programs in RUSSELL is not more difficult than in HASKELL, with the added benefit of a complete environment for proving properties about your programs in a highly expressive specification language. Clearly, there is room for improvement on the side of programming language technology, for example some kind of overloading mechanism would allow more conciseness, but RUSSELL is already very powerful in terms of specification possibilities and actual support for proving properties, as it (re-)lies on an already mature proof assistant. One thing we learned is that once we begin to use dependent types, embracing them completely become a sensible and desirable goal, carrying the strong properties bottom-up. PROGRAM allows to do this in a tractable and hopefully maintainable way.

**Acknowledgments** Thanks to Christine Paulin-Mohring, Jean-Christophe Filliâtre and the anonymous referees for their insightful comments on drafts of this paper, along with Adam Chlipala and my colleagues of the DÉMONS team for their lasting enthusiasm and criticism. Finally, I thank Yann Régis-Gianas for introducing me to Finger Trees.

## References

- Edwin Brady, Conor McBride, and James McKinna. Inductive families need not store their indices. In Stefano Berardi, Mario Coppo, and Ferruccio Damiani, editors, *TYPES*, volume 3085 of *LNCS*, pages 115–129. Springer, 2003.
- Adam Chlipala. Position paper: Thoughts on programming with proof assistants. In *PLPV'06: Proceedings of the Programming Languages meets Program Verification Workshop*, August 2006.
- Coq. The Coq proof assistant. [coq.inria.fr](http://coq.inria.fr).
- Thierry Coquand. Alfa/agda. In Freek Wiedijk, editor, *The Seventeen Provers of the World*, volume 3600 of *LNCS*, pages 50–54. Springer, 2006.
- Jean-Christophe Filliâtre and Pierre Letouzey. Functors for proofs and programs. In David A. Schmidt, editor, *ESOP*, volume 2986 of *LNCS*, pages 370–384. Springer, 2004.
- Seth Fogarty, Emir Pasalic, Jeremy Siek, and Walid Taha. Concoction: indexed types now! In G. Ramalingam and Eelco Visser, editors, *PEPM*, pages 112–121. ACM, 2007.
- Haskell. The Haskell programming language. [haskell.org](http://haskell.org).
- Ralf Hinze and Ross Paterson. Finger Trees: A Simple General-purpose Data Structure. *J. Funct. Program.*, 16(2):197–217, 2006.
- Pierre Letouzey. A new extraction for coq. In Herman Geuvers and Freek Wiedijk, editors, *TYPES'02*, volume 2646 of *LNCS*, pages 200–219. Springer, 2002.
- Conor McBride. *Dependently Typed Functional Programs and Their Proofs*. PhD thesis, University of Edinburgh, 1999.
- Conor McBride and James McKinna. The view from the left. *J. Funct. Program.*, 14(1):69–111, 2004.
- Alexandre Miquel. The implicit calculus of constructions. In *TLCA*, volume 2044 of *LNCS*, pages 344–359. Springer, 2001.
- OCaml. The Ocaml programming language. [caml.inria.fr](http://caml.inria.fr).
- Sam Owre and Natarajan Shankar. The formal semantics of PVS. Technical Report SRI-CSL-97-2, Computer Science Laboratory, SRI International, Menlo Park, CA, August 1997.
- Catherine Parent. Synthesizing proofs from programs in the Calculus of Inductive Constructions. In Bernhard Möller, editor, *MPC*, volume 947 of *LNCS*, pages 351–379. Springer, 1995.
- Christine Paulin-Mohring. Inductive definitions in the system COQ. In *Typed Lambda Calculi and Applications*, volume 664 of *LNCS*, pages 328–345. Springer, 1993.
- Tim Sheard. Languages of the future. *SIGPLAN Notices*, 39(12):119–132, 2004.
- Matthieu Sozeau. Subset coercions in Coq. In *TYPES'06*, volume 4502 of *LNCS*, pages 237–252. Springer, 2007.
- Benjamin Werner. On the strength of proof-irrelevant type theories. *3rd International Joint Conference on Automated Reasoning*, 2006.
- Hongwei Xi. Applied Type System (extended abstract). In *post-workshop Proceedings of TYPES 2003*, pages 394–408. Springer-Verlag LNCS 3085, 2004.
- Hongwei Xi and Frank Pfenning. Dependent types in practical programming. In *ACM SIGPLAN–SIGACT Symposium on Principles of Programming Languages (POPL)*, San Antonio, Texas, pages 214–227, January 1999.