

Generic Programming for Dependent Types

Constructing Strictly Positive Families

Peter Morris Thorsten Altenkirch

University of Nottingham, UK
 {pwm,txa}@cs.nott.ac.uk

Abstract

We begin by revisiting the idea of using a universe of types to write generic programs in a dependently typed setting by constructing a universe for Strictly Positive Types (SPTs). Here we extend this construction to cover dependent types, i.e. Strictly Positive Families (SPFs), thereby fixing a gap left open in previous work. Using the approach presented here we are able to represent all of Epigram’s datatypes within Epigram including the universe of datatypes itself.

Keywords Data type generic programming, dependently typed programming, Epigram

1. Introduction

In a dependently typed language like Epigram [8, 7, 3] generic programming is normal programming. This is achieved by defining a universe [6, 10] consisting of a type of names $U : \star$ and a family of elements $EI : U \rightarrow \star$ indexed by type names. We have exploited this opportunity in [9] by defining the universe of regular tree types and developing generic programs and proofs for this universe. However, there is an obvious asymmetry in our previous definitions where we exploit the power of dependent types to encode a universe of non-dependent types. In the present paper we show how to correct this imbalance and construct universes of dependent types. We define the universe of strictly positive families, which contains codes for all datatypes definable within Epigram, including the definition of this universe itself. We also consider a smaller universe of regular families, which is the dependent counterpart of the universe of regular tree types. This smaller universe which excludes infinitely branching trees is interesting because it allows more programs including a generic program to decide equality.

In related work, [5] presents an alternative approach to defining a universe of indexed strictly positive definitions. This has been the base for generic programming within the AGDA system [4].

1.1 Programming in Epigram

Epigram is an dependently typed functional language with an interactive environment for developing programs with the aid of the type checker.

All Epigram programs (currently) are total to ensure that type checking is decidable. We ensure this by only allowing structural recursion. Programs are presented as decision trees, representing the structure of the analysis of the problem being solved. Each node is presented with the information available as a pattern on the left hand side, right hand sides can either be:

- $\Rightarrow t$ the function *returns* t , an expression of the appropriate type, constructed over the pattern variables on the left;
- $\Leftarrow e$ the function’s analysis is refined by e , an *eliminator* expression, or ‘gadget’, characterizing some scheme of case analysis or recursion, giving rise to a bunch of sub nodes with more informative left-hand sides;
- $|| w$ the sub nodes’ left-hand sides are to be extended *with* the value of w , some intermediate computation, in an extra column: this may then be analysed in addition to the function’s original arguments.

In this paper we will need only two ‘by’ gadgets, `rec` which constructs the structural recursive calls available to the programmer, and `case` which applies the appropriate derived case analysis principle and introduces a set of more informative patterns in the sub-nodes. We will use the convention that we suppress the use of `case` when its presence is inferable from the presence of constructors in the patterns. We will always be explicit about which input we are being structurally recursive on.

Epigram’s data types are presented by declaring their formation rules and constructors in natural deduction style as are the types of functions. In these rules arguments whose types are inferable can be omitted for brevity.

As a warm up to Epigram’s syntax we’ll start with a simple datatype, the natural numbers and define addition:

```

data Nat :  $\star$  where 0 : Nat 1+ n :  $\star$ 
let plus m n : Nat
  plus m n  $\Leftarrow$  rec m
  plus 0 n  $\Rightarrow$  n
  plus (1+ m) n  $\Rightarrow$  1+ (plus m n)

```

We can then define types which are dependent on the natural numbers, the finite types and vectors (lists of a given length), we can then define safe projection from using the finite types to ensure there are only as many indexes as elements in the array.

[copyright notice will appear here]

$$\text{data } \frac{n : \text{Nat}}{\text{Fin } n : \star} \text{ where } \frac{}{0 : \text{Fin } (1+n)} \frac{i : \text{Fin } n}{1+i : \text{Fin } (1+n)}$$

$$\text{data } \frac{A : \star \quad n : \text{Nat}}{\text{Vec } n \ A : \star} \text{ where } \frac{}{\varepsilon : \text{Vec } 0 \ A} \frac{a : A \quad as : \text{Vec } n \ A}{a:as : \text{Vec } (1+n) \ A}$$

$$\text{let } \frac{as : \text{Vec } n \ A \quad i : \text{Fin } n}{\text{proj } as \ i : A} \\ \text{proj } as \ i \Leftarrow \text{rec } i \\ \text{proj } (a:as) \ 0 \Rightarrow a \\ \text{proj } (a:as) \ (1+i) \Rightarrow \text{proj } as \ i$$

Notice that the nil case doesn't appear since $\text{Fin } 0$ is uninhabited.

We will be returning to each of the above datatypes to encode them into our universe, along with another nice example of Epigram, the scoped lambda terms:

$$\text{data } \frac{n : \text{Nat}}{\text{Lam } n : \star} \text{ where } \frac{i : \text{Fin } n \quad t : \text{Lam } (1+n)}{\text{var } i : \text{Lam } n \quad \text{abs } t : \text{Lam } n} \\ \frac{f, a : \text{Lam } n}{\text{app } f \ a : \text{Lam } n}$$

2. Strictly Positive Types

Strictly positive types can be constructed using polynomial expressions ($0, 1, +, \times$), initial algebras (μ) and exponentiation by a constant $K \rightarrow -$. Examples are natural numbers $\text{Nat} = \mu X.1 + X$, lists $\text{List } A = \mu X.1 + A \times X$, rose trees $\text{RT } A = \mu Y.\text{List } (A \times Y) = \mu Y.\mu X.1 + A \times Y \times X$ and ordinal notations $\text{Ord} = \mu X.1 + X + \text{Nat} \rightarrow X$. The first three examples, which don't use exponentiation are regular tree types which are a proper subset of strictly positive types.

Strictly positive types with n free type variables are represented by the Epigram type $\text{SPT } n$, see figure 1. We use de Bruijn variables, where 'Z' stands for the last variable with index 0 and previous variables can be accessed using weakening 'wk'. Note that the fix-point constructor ' μ ' reduces the number of free variables by 1 because the last variable has been bound. We introduce the type of nested type substitutions, or telescopes, of length n , $\text{Tel } n$. Given a type $T : \text{SPT } n$ and a matching telescope $\vec{T} : \text{Tel } n$ we define the type of elements $\text{El } \vec{T} \ T$ as an inductive family.

We can now encode the examples mentioned above (Nat, List, RT, Ord):

$$\text{let } \text{'Nat'} : \text{SPT } 0 \\ \text{'Nat'} \Rightarrow \mu ('1' '+' 'Z')$$

$$\text{let } \text{'List'} : \text{SPT } 1 \\ \text{'List'} \Rightarrow \mu ('1' '+' (('wk' 'Z') '\times' 'Z'))$$

$$\text{let } \text{'RT'} : \text{SPT } 1 \\ \text{'RT'} \Rightarrow \mu (('wk' 'Z') '+' (\mu (('wk' 'Z') '\times' 'Z')))$$

$$\text{let } \text{'Ord'} : \text{SPT } 0 \\ \text{'Ord'} \Rightarrow \mu ('1' '+' ('Z' '+' (\text{Nat } \rightarrow 'Z')))$$

In the system we can build 'constructors' which construct elements of the interpretation of a code, for example with 'Nat':

$$\text{let } \text{'0'} : \text{El } \varepsilon \ \text{'Nat'} \\ \text{'0'} \Rightarrow \text{in } (\text{inl } \text{void}) \\ \text{let } \text{'1+'} : \text{El } \varepsilon \ \text{'Nat'} \rightarrow \text{El } \varepsilon \ \text{'Nat'} \\ \text{'1+'} \Rightarrow (\lambda n \Rightarrow \text{in } (\text{inr } (\text{top } n)))$$

2.1 Generic Map

As our first example of a generic program we shall present generic map. We shall define this by first considering morphisms between telescopes. Our base case is going to be the identity morphism between a telescope and itself, and we can obviously extend a morphism with a function. To ensure our map is structurally recursive (and total) we also need to be able to go under local bindings without extending the, morphism ϕ by the non structural recursive call $\text{gMap } \phi$. So we add a third constructor to morphisms mMap which we will use to denote an extension under a binder.

$$\text{data } \frac{\vec{S}, \vec{T} : \text{Tel } n}{\text{Morph } \vec{S} \ \vec{T} : \star} \text{ where} \\ \text{mld} : \text{Morph } \vec{S} \ \vec{S} \\ \phi : \text{Morph } \vec{S} \ \vec{T} \quad f : \text{El } \vec{S} \ S \rightarrow \text{El } \vec{T} \ T \\ \text{mFun } \phi \ f : \text{Morph } (\vec{S}:S) \ (\vec{T}:T) \\ \frac{\phi : \text{Morph } \vec{S} \ \vec{T}}{\text{mMap } \phi : \text{Morph } (\vec{S}:T) \ (\vec{T}:T)}$$

Most of the definition for gMap is structural, we extend uniformly under ' μ ' and when we see a mMap extension at the top variable we reconstruct the (now obviously structural) call we'd expect.

$$\text{let } \frac{\phi : \text{Morph } \vec{S} \ \vec{T} \quad x : \text{El } \vec{S} \ T}{\text{gMap } \phi \ x : \text{El } \vec{T} \ T} \\ \text{gMap } \phi \quad x \quad \Leftarrow \text{rec } x \\ \text{gMap } \text{mld} \quad (\text{top } x) \quad \Rightarrow \text{top } x \\ \text{gMap } (\text{mFun } \phi \ f) \quad (\text{top } x) \quad \Rightarrow \text{top } (f \ x) \\ \text{gMap } (\text{mMap } \phi) \quad (\text{top } x) \quad \Rightarrow \text{top } (\text{gMap } \phi \ x) \\ \text{gMap } \text{mld} \quad (\text{pop } x) \quad \Rightarrow \text{pop } x \\ \text{gMap } (\text{mFun } \phi \ f) \quad (\text{pop } x) \quad \Rightarrow \text{pop } (\text{gMap } \phi \ x) \\ \text{gMap } (\text{mMap } \phi) \quad (\text{pop } x) \quad \Rightarrow \text{pop } (\text{gMap } \phi \ x) \\ \text{gMap } \phi \quad \text{void} \quad \Rightarrow \text{void} \\ \text{gMap } \phi \quad (\text{inl } x) \quad \Rightarrow \text{inl } (\text{gMap } \phi \ x) \\ \text{gMap } \phi \quad (\text{inr } x) \quad \Rightarrow \text{inr } (\text{gMap } \phi \ x) \\ \text{gMap } \phi \quad (\text{pair } x \ y) \\ \Rightarrow \text{pair } (\text{gMap } \phi \ x) \ (\text{gMap } \phi \ y) \\ \text{gMap } \phi \quad (\text{fun } f) \\ \Rightarrow \text{fun } (\lambda k \Rightarrow \text{gMap } \phi \ (f \ k)) \\ \text{gMap } \phi \quad (\text{in } x) \\ \Rightarrow \text{in } (\text{gMap } (\text{mMap } \phi) \ x)$$

In our work on the regular tree types [9], we present a number of other algorithms in this style for the regular tree types, including a decidable equality. Types in this universe do not have such an equality, there is no such function for the ordinals for instance. It is clear that the larger the universe of types the fewer generic operations we may define. In a system of generic programming it is conceivable that we would need a number of successively large universes to cope with this trade off.

$$\begin{array}{c}
\text{data } \frac{n : \text{Nat}}{\text{SPT } n : \star} \text{ where} \\
\frac{}{\text{'Z'} : \text{SPT } (1+n)} \quad \frac{T : \text{SPT } n}{\text{'wk'} T : \text{SPT } (1+n)} \quad \frac{}{\text{'0'} : \text{SPT } n} \quad \frac{}{\text{'1'} : \text{SPT } n} \\
\frac{S, T : \text{SPT } n}{S \text{'+' } T : \text{SPT } n} \quad \frac{S, T : \text{SPT } n}{S \text{'\times'} T : \text{SPT } n} \quad \frac{K : \star \quad T : \text{SPT } n}{K \text{'\to'} T : \text{SPT } n} \quad \frac{F : \text{SPT } (1+n)}{\text{'}\mu\text{' } F : \text{SPT } n} \\
\\
\text{data } \frac{n : \text{Nat}}{\text{Tel } n : \star} \quad \frac{}{\varepsilon : \text{Tel } 0} \quad \frac{\vec{T} : \text{Tel } n \quad T : \text{SPT } n}{\vec{T} : T : \text{Tel } (1+n)} \\
\\
\text{data } \frac{\vec{T} : \text{Tel } n \quad T : \text{SPT } n}{\text{El } \vec{T} T : \star} \text{ where} \\
\frac{e : \text{El } \vec{T} T}{\text{top } e : \text{El } (\vec{T} : T) \text{'Z'}} \quad \frac{e : \text{El } \vec{T} T}{\text{pop } e : \text{El } (\vec{T} : S) \text{'wk'} T} \quad \frac{}{\text{void} : \text{El } \vec{T} \text{'1'}} \quad \frac{f : K \to \text{El } \vec{T} T}{\text{fun } f : \text{El } \vec{T} (K \text{'\to'} T)} \\
\frac{s : \text{El } \vec{T} S}{\text{inl } s : \text{El } \vec{T} (S \text{'+' } T)} \quad \frac{t : \text{El } \vec{T} T}{\text{inr } t : \text{El } \vec{T} (S \text{'+' } T)} \quad \frac{s : \text{El } \vec{T} S \quad t : \text{El } \vec{T} T}{\text{pair } s t : \text{El } \vec{T} (S \text{'\times'} T)} \quad \frac{e : \text{El } (\vec{T} : \mu F)}{\text{in } e : \text{El } \vec{T} (\mu F)}
\end{array}$$

Figure 1. The SPT Universe

3. Strictly Positive Families

There is an obvious imbalance in the above construction, we use the dependent types of Epigram to model a class of simple types for generic programming. If this technique is ever to prove useful then it has to be able to include as much of the power of the host language's type system as possible.

We have been using strictly positive families (presented in Epigram's natural deduction style) throughout this paper, but to illustrate the universe below we will return to the examples from the introduction, the finite types **Fin**, vectors **Vec** and scoped lambda terms **Lam**.

To tell this story we consider the positive ways to build a family of types in $O \rightarrow \star$, for the moment we'll call these codes **Fam**. We call O the output index type. In all of the above examples $O \equiv \text{Nat}$, though it can be *any* type. The simplest such families are constant and ignore the indexing information, we have two such types, **'0'** and **'1'** which are have zero or one elements (respectively) at all indices. We introduce the codes for this universe, and the interpretation $\llbracket - \rrbracket : \text{Fam } O \rightarrow O \rightarrow \star$ simultaneously.

$$\overline{\text{'0'}, \text{'1'} : \text{Fam } O} \quad \text{void} : \llbracket \text{'1'} \rrbracket_o$$

Clearly another possibility is to substitute for the given index, that is given a $O \rightarrow \star$ and a function in $O' \rightarrow O$ we can create a new family in $O' \rightarrow \star$. Categorically this relates to a pullback, and we will refer to it as re-indexing of a family.

$$\frac{f : O' \rightarrow O \quad T : \text{Fam } O}{\text{'}\Delta\text{' of } T : \text{Fam } O'} \quad \frac{v : \llbracket T \rrbracket (f o)}{\delta v : \llbracket \text{'}\Delta\text{' of } T \rrbracket_o}$$

It is interesting now to consider what we can do to construct values if we have a function on output types that goes the other way. Given a **Fam** O and a function $O \rightarrow O'$ we must construct values at a given index $o' : O'$, to do this we consider only the values $o : O$ for which $f o = o'$ but do we consider all such values, or just pick one? The first option gives us dependent sum, the second dependent product:

$$\begin{array}{c}
\frac{f : O \rightarrow O' \quad T : \text{Fam } O}{\text{'}\Sigma\text{' of } T, \text{'}\Pi\text{' of } T : \text{Fam } O'} \\
\frac{v : \exists o : O \Rightarrow (f o = o') \times \llbracket T \rrbracket_o}{\sigma_v : \llbracket \text{'}\Sigma\text{' of } T \rrbracket_{o'}} \\
\frac{\vec{v} : \forall o : O \Rightarrow (f o = o') \rightarrow \llbracket T \rrbracket_o}{\pi \vec{v} : \llbracket \text{'}\Pi\text{' of } T \rrbracket_{o'}}
\end{array}$$

Categorically, **'Σ'** and **'Π'** are respectively the left and right adjoints of **'Δ'**.

Although we called these sum and product there is only one type involved, in fact all of our constructs are linear in this way. **'Σ'** and **'Π'** work only on the index level so we must be able to pick between types based on the index we are given. There are a number of possibilities here but it seems sensible to restrict ourselves to finite collections of types so, using **Fin**:

$$\frac{f : \forall t : \text{Fin } n \Rightarrow \text{Fam } O}{\text{'}\text{Tag}\text{' } f : \text{Fam } (O \times \text{Fin } n)} \quad \frac{v : \llbracket f t \rrbracket_o}{\text{tag}_t v : \llbracket \text{'}\text{Tag}\text{' } f \rrbracket (o; t)}$$

Finally we add type variables and the fixed point constructor. Our input types will now be dependent so we have to index our codes not only by the input index type but the set of input index types, **Fam** above becomes:

$$\text{data } \frac{\vec{I} : \text{Vec } \star \quad n \quad O : \star}{\text{SPF } \vec{I} O : \star}$$

For the interpretation we follow the **SPT** construction and introduce telescopes as closing substitutions, this time with indexing information:

$$\begin{array}{c}
\text{data } \frac{\vec{I} : \text{Vec } \star \quad n}{\text{dTel } \vec{I} : \star} \text{ where} \\
\frac{}{\varepsilon : \text{dTel } \varepsilon} \quad \frac{\vec{T} : \text{dTel } \vec{I} \quad T : \text{SPF } \vec{I} I}{\vec{T} : T : \text{dTel } (\vec{I} : I)}
\end{array}$$

and our interpretation now has this type:

$$\text{data } \frac{T : \text{SPF } \vec{I} O \quad \vec{T} : \text{dTel } \vec{I} o : O}{\llbracket T \rrbracket \vec{T} o : \star}$$

the introduction rules for the codes given above can be uniformly modified to thread the context unchanged. The variable rules are as follows:

$$\frac{}{\text{'Z'} : \text{SPF } (\vec{I}:O) O} \quad \frac{v : \llbracket T \rrbracket \vec{T} o}{\text{top } v : \llbracket \text{'Z'} \rrbracket (\vec{T}:T) o}$$

$$\frac{T : \text{SPT } \vec{I} O}{\text{'wk'} T : \text{SPF } (\vec{I}:I) O} \quad \frac{v : \llbracket T \rrbracket \vec{T} o}{\text{pop } T : \llbracket \text{'wk'} T \rrbracket (\vec{T}:S) o}$$

Finally the fixed point construct:

$$\frac{T : \text{SPF } (\vec{I}:O) O}{\text{'}\mu\text{' } T : \text{SPF } \vec{I} O} \quad \frac{v : \llbracket T \rrbracket (\vec{T}:(\text{'}\mu\text{' } T)) o}{\text{in } v : \llbracket \text{'}\mu\text{' } T \rrbracket \vec{T} o}$$

The whole construction is given in figure 2.

We can now encode our examples into this universe, as an abbreviation we will first define disjoint union, Cartesian product and non-dependent arrow in a form similar to the **SPT** construction:

$$\text{let } \frac{A, B : \text{SPF } \vec{I} O}{A \text{'+' } B, A \text{'\times'} B : \text{SPF } \vec{I} O}$$

$$A \text{'+' } B \Rightarrow \text{'}\Sigma\text{' fst } \left(\text{'Tag'} \left(\begin{array}{l} 0 \Rightarrow A \\ 1 \Rightarrow B \end{array} \right) \right)$$

$$\text{let } \frac{a : \llbracket A \rrbracket \vec{T} o}{\text{inl } a : \llbracket A \text{'+' } B \rrbracket \vec{T} o}$$

$$\text{inl}_{\vec{T} o} a \Rightarrow \sigma_{(o;0)} (\text{tag } a)$$

$$\text{let } \frac{b : \llbracket B \rrbracket \vec{T} o}{\text{inr } b : \llbracket A \text{'+' } B \rrbracket \vec{T} o}$$

$$\text{inr}_{\vec{T} o} b \Rightarrow \sigma_{(o;1)} (\text{tag } b)$$

$$A \text{'\times'} B \Rightarrow \text{'}\Pi\text{' fst } \left(\text{'Tag'} \left(\begin{array}{l} 0 \Rightarrow A \\ 1 \Rightarrow B \end{array} \right) \right)$$

$$\text{let } \frac{a : \llbracket A \rrbracket \vec{T} o \quad b : \llbracket B \rrbracket \vec{T} o}{\text{pair } a b : \llbracket A \text{'\times'} B \rrbracket \vec{T} o}$$

$$\text{pair } a b \Rightarrow \pi \left(\begin{array}{l} (o;0) \text{ refl} \Rightarrow \text{tag } a \\ (o;1) \text{ refl} \Rightarrow \text{tag } b \end{array} \right)$$

$$\text{let } \frac{K : \star \quad T : \text{SPF } \vec{I} O}{K \text{'}\rightarrow\text{' } T : \text{SPF } \vec{I} O}$$

$$K \text{'}\rightarrow\text{' } T \Rightarrow \text{'}\Pi\text{'}_{(O \times K)} \text{fst } (\text{'}\Delta\text{' fst } T)$$

$$\text{let } \frac{\vec{v} : K \rightarrow \llbracket T \rrbracket \vec{T} T}{\text{fun } \vec{v} : \llbracket K \text{'}\rightarrow\text{' } T \rrbracket \vec{T} o}$$

$$\text{fun } \vec{v} \Rightarrow \pi (\delta \vec{v})$$

Notice how the **'Tag'** interacts with the **'Σ'** and **'Π'**. These auxiliary constructions allow us to construct codes in a sums of products style as before, except with added indexing information, for instance:

$$\text{let } \text{'Fin'} : \text{SPF } \square \text{Nat}$$

$$\text{'Fin'} \Rightarrow \text{'}\mu\text{' } ((\text{'}\Sigma\text{' } 1 + \text{'1'}) \text{'+' } (\text{'}\Sigma\text{' } 1 + \text{'Z'}))$$

Remember that both constructors for **Fin** target non zero numbers, the **'Σ' 1+** construct ensures this, it forces us to choose an index n' that satisfies $n = 1 + n'$ for the output index n since the equality type is uninhabited if $n = 0$ then so is the interpretation of this code.

In the **'Vec'** example we assume that the element type is fully applied, and so is indexed only by the unit type. When we refer to its type variable, we use delta to force the index to $()$. In the nil case we use **'Σ'** to ensure that the index is 0 .

$$\text{let } \text{'Vec'} : \text{SPF } [1] \text{Nat}$$

$$\text{'Vec'} \Rightarrow$$

$$\text{'}\mu\text{' } \left(\begin{array}{l} (\text{'}\Sigma\text{' } (\text{const } 0) \text{'1'}) \text{'+' } \\ (\text{'}\Sigma\text{' } 1 + (\text{'}\Delta\text{' } (\text{const } ())) (\text{'wk'} \text{'Z'})) \text{'\times'} \text{'Z'} \end{array} \right)$$

As before we can create **'constructors'** for these types to reflect the conventional ones. For instance for **'list'**:

$$\text{let } \text{'}\epsilon\text{' } : \llbracket \text{'Vec'} \rrbracket [A] 0$$

$$\text{'}\epsilon\text{' } \Rightarrow \text{in } (\text{inl } (\sigma_{\text{void}}))$$

$$\text{let } \frac{a : \llbracket A \rrbracket () \quad as : \llbracket \text{'Vec'} \rrbracket [A] n}{(a \text{'z'} as) : \llbracket \text{'Vec'} \rrbracket [A] (1 + n)}$$

$$(a \text{'z'} as) \Rightarrow \text{in } (\text{inr } (\text{pair } (\delta a) as))$$

As promised, we can also define the well scoped lambda terms:

$$\text{let } \text{'Lam'} : \text{SPF } \square \text{Nat}$$

$$\text{'Lam'} \Rightarrow$$

$$\text{'}\mu\text{' } ((\text{'wk'} \text{'Fin'}) \text{'+' } ((\text{'Z'} \text{'\times'} \text{'Z'}) \text{'+' } (\text{'}\Delta\text{' } 1 + \text{'Z'})))$$

3.1 Embedding SPT into SPF

It is clear that the **SPT** universe represents a subset of the the types in **SPF** indeed we have overloaded the names **'Z'**, **'+'**, \dots , **'μ'** to suggest the relation. Will will now be more explicit about this.

We first give a function that transforms **SPT** codes into **SPF** codes, since **SPFs** have to be indexed by some type, we have no choice but to index the embedded codes by the unit type:

$$\text{let } \frac{n : \text{Nat} \quad a : A}{\text{vec } a n : \text{Vec } A n}$$

$$\text{vec } a n \Leftarrow \text{rec } n$$

$$\text{vec } a 0 \Rightarrow \epsilon$$

$$\text{vec } a (1 + n) \Rightarrow a : (\text{vec } a n)$$

$$\text{let } \frac{T : \text{SPT } n}{\text{embspt } T : \text{SPF } (\text{vec } 1 n) 1}$$

$$\text{embspt } T \Leftarrow \text{rec } T$$

$$\text{embspt } \text{'Z'} \Rightarrow \text{'Z'}$$

$$\text{embspt } (\text{'wk'} T) \Rightarrow \text{'wk'} (\text{embspt } T)$$

$$\text{embspt } \text{'0'} \Rightarrow \text{'0'}$$

$$\text{embspt } \text{'1'} \Rightarrow \text{'1'}$$

$$\text{embspt } (S \text{'+' } T) \Rightarrow (\text{embspt } S) \text{'+' } (\text{embspt } T)$$

$$\text{embspt } (S \text{'\times'} T) \Rightarrow (\text{embspt } S) \text{'\times'} (\text{embspt } T)$$

$$\text{embspt } (K \text{'}\rightarrow\text{' } T) \Rightarrow K \text{'}\rightarrow\text{' } (\text{embspt } T)$$

$$\text{embspt } (\text{'}\mu\text{' } F) \Rightarrow \text{'}\mu\text{' } (\text{embspt } F)$$

We map this embedding across the telescope:

$$\text{let } \frac{\vec{T} : \text{Tel } n}{\text{embtel } \vec{T} : \text{dTel } (\text{vec } 1 n)}$$

$$\text{embtel}_n \vec{T} \Leftarrow \text{rec } n$$

$$\text{embtel}_0 \epsilon \Rightarrow \epsilon$$

$$\text{embtel}_{(1+n)} (\vec{T}:T) \Rightarrow (\text{embtel } \vec{T}) : (\text{embspt } T)$$

$$\begin{array}{c}
\text{data } \frac{\vec{I} : \text{Vec } \star \ n \ O : \star}{\text{SPF } \vec{I} \ O : \star} \text{ where } \frac{}{\text{'Z'} : \text{SPF } (\vec{I}:O) \ O} \frac{T : \text{SPF } \vec{I} \ O}{\text{'wk'} \ T : \text{SPF } (\vec{I}:I) \ O} \\
\frac{f : \forall t : \text{Fin } n \Rightarrow \text{SPF } \vec{I} \ O}{\text{'Tag'} \ f : \text{SPF } \vec{I} \ (O \times \text{Fin } n)} \frac{}{\text{'0'}, \text{'1'} : \text{SPF } \vec{I} \ O} \frac{T : \text{SPF } (\vec{I}:O) \ O}{\text{'}\mu\text{' } T : \text{SPF } \vec{I} \ O} \\
\frac{f : O \rightarrow O' \ T : \text{SPF } \vec{I} \ O}{\text{'}\Sigma\text{' of } T : \text{SPF } \vec{I} \ O'} \frac{f : O' \rightarrow O \ T : \text{SPF } \vec{I} \ O}{\text{'}\Delta\text{' of } T : \text{SPF } \vec{I} \ O'} \frac{f : O \rightarrow O' \ T : \text{SPF } \vec{I} \ O}{\text{'}\Pi\text{' of } T : \text{SPF } \vec{I} \ O'} \\
\\
\text{data } \frac{\vec{I} : \text{Vec } \star \ n}{\text{dTel } \vec{I} : \star} \text{ where } \frac{}{\varepsilon : \text{dTel } \varepsilon} \frac{\vec{T} : \text{dTel } \vec{I} \ T : \text{SPF } \vec{I} \ I}{(\vec{T}:T) : \text{dTel}(\vec{I}:I)} \\
\\
\text{data } \frac{T : \text{SPF } \vec{I} \ O \ \vec{T} : \text{dTel } \vec{I} \ o : O}{\llbracket T \rrbracket \vec{T} \ o : \star} \text{ where } \frac{v : \llbracket T \rrbracket \vec{T} \ o}{\text{top } v : \llbracket \text{'Z'} \rrbracket (\vec{T}:T) \ o} \frac{v : \llbracket T \rrbracket \vec{T} \ o}{\text{pop } v : \llbracket \text{'wk'} \rrbracket T \rrbracket (\vec{T}:S) \ o} \\
\frac{v : \llbracket f \ t \rrbracket \vec{T} \ o}{\text{tag}_t v : \llbracket \text{'Tag'} \ f \rrbracket \vec{T} \ (o; t)} \frac{}{\text{void} : \llbracket \text{'1'} \rrbracket \vec{T} \ o} \frac{v : \llbracket T \rrbracket (\vec{T}:(\text{'}\mu\text{' } T)) \ o}{\text{in } v : \llbracket \text{'}\mu\text{' } T \rrbracket \vec{T} \ o} \\
\frac{v : \llbracket T \rrbracket \vec{T} \ o}{\sigma_o v : \llbracket \text{'}\Sigma\text{' } f \ T \rrbracket \vec{T} \ (f \ o)} \frac{v : \llbracket T \rrbracket \vec{T} \ (f \ o)}{\delta v : \llbracket \text{'}\Delta\text{' } f \ T \rrbracket \vec{T} \ o} \frac{\vec{v} : \forall o : O \ p : (f \ o) = o' \Rightarrow \llbracket T \rrbracket \vec{T} \ o}{\pi \vec{v} : \llbracket \text{'}\Pi\text{' } f \ T \rrbracket \vec{T} \ o'}
\end{array}$$

Figure 2. The SPF Universe

Finally we embed elements in the simpler universe as elements of the embedded code interpreted by the embedding of the telescope:

$$\begin{array}{c}
\text{let } \frac{x : \text{El } \vec{T} \ T}{\text{embel } T : \llbracket \text{embspt } T \rrbracket (\text{embtel } \vec{T}) \ ()} \\
\text{embel } T \leftarrow \text{rec } T \\
\text{embel } (\text{top } t) \Rightarrow \text{top } (\text{embel } t) \\
\text{embel } (\text{pop } t) \Rightarrow \text{top } (\text{embel } t) \\
\text{embel } \text{void} \Rightarrow \text{void} \\
\text{embel } (\text{inl } t) \Rightarrow \text{inl } (\text{embel } t) \\
\text{embel } (\text{inr } t) \Rightarrow \text{inr } (\text{embel } t) \\
\text{embel } (\text{pair } s \ t) \Rightarrow \text{pair } (\text{embel } s) \ (\text{embel } t) \\
\text{embel } (\text{fun } f) \Rightarrow \text{fun } (\lambda k \Rightarrow \text{embel } (f \ k)) \\
\text{embel } (\text{in } t) \Rightarrow \text{in } (\text{embel } t)
\end{array}$$

4. Generic Map

To construct the generic map operation is a simple extension of the SPT map, indeed the morphisms are almost identical we only generalise to say that the functions must operate at any output index as we cannot before-hand anticipate which will turn up.

$$\begin{array}{c}
\text{data } \frac{\vec{S}, \vec{T} : \text{dTel } \vec{I}}{\text{Morph } \vec{S} \ \vec{T} : \star} \text{ where} \\
\text{mld} : \text{Morph } \vec{S} \ \vec{S} \\
\frac{\phi : \text{Morph } \vec{S} \ \vec{T} \ f : \forall i : I \Rightarrow \llbracket S \rrbracket \vec{S} \ i \rightarrow \llbracket T \rrbracket \vec{T} \ i}{\text{mFun } \phi \ f : \text{Morph } (\vec{S}:S) \ (\vec{T}:T)} \\
\frac{\phi : \text{Morph } \vec{S} \ \vec{T}}{\text{mMap } \phi : \text{Morph } (\vec{S}:T) \ (\vec{T}:T)}
\end{array}$$

The function itself is structural except at the variables where we do exactly as before:

$$\begin{array}{c}
\text{let } \frac{\phi : \text{Morph } \vec{S} \ \vec{T} \ x : \llbracket T \rrbracket \vec{S} \ o}{\text{gMap } \phi \ x : \llbracket T \rrbracket \vec{T} \ o} \\
\text{gMap } \phi \ x \leftarrow \text{rec } x \\
\text{gMap } \text{mld} \ (\text{top } x) \Rightarrow \text{top } x \\
\text{gMap } (\text{mFun } \phi \ f) \ (\text{top } x) \Rightarrow \text{top } (f \ x) \\
\text{gMap } (\text{mMap } \phi) \ (\text{top } x) \Rightarrow \text{top } (\text{gMap } \phi \ x) \\
\text{gMap } \text{mld} \ (\text{pop } x) \Rightarrow \text{pop } x \\
\text{gMap } (\text{mFun } \phi \ f) \ (\text{pop } x) \Rightarrow \text{pop } (\text{gMap } \phi \ x) \\
\text{gMap } (\text{mMap } \phi) \ (\text{pop } x) \Rightarrow \text{pop } (\text{gMap } \phi \ x) \\
\text{gMap } \phi \ (\text{tag } x) \Rightarrow \text{tag } (\text{gMap } \phi \ x) \\
\text{gMap } \phi \ \text{void} \Rightarrow \text{void} \\
\text{gMap } \phi \ (\text{in } x) \Rightarrow \text{in } (\text{gMap } (\text{mMap } \phi) \ x) \\
\text{gMap } \phi \ (\sigma_o x) \Rightarrow \sigma_o (\text{gMap } \phi \ x) \\
\text{gMap } \phi \ (\delta x) \Rightarrow \delta (\text{gMap } \phi \ x) \\
\text{gMap } \phi \ (\pi \vec{x}) \Rightarrow \pi (\lambda o \ p \Rightarrow \text{gMap } \phi \ (\vec{x} \ o \ p))
\end{array}$$

As we mentioned earlier the class of strictly positive types does not support all of the generic programs we might like to write in this way. The solution with SPTs is to loose the ‘→’ constructor and we arrive at the regular tee types from our previous work. We can play a similar trick here, although the source of the infinities that cause the trouble, ‘Π’, is rather more integral to the construction (giving us our ‘×’, for example) so, instead of removing it, we restrict it’s power to finite domains. We introduce the regular families (RF), whose construction is identical to the above except we replace the ‘Π’ rules with:

$$\begin{array}{c}
\frac{n : \text{Nat} \ T : \text{RF } \vec{I} \ (O \times \text{Fin } n)}{\text{'}\Pi^{<\omega}\text{' } n \ T : \text{RF } \vec{I} \ O} \\
\frac{\vec{v} : \forall i : \text{Fin } n \Rightarrow \llbracket T \rrbracket \vec{T} \ (o; i)}{\pi^{<\omega} \vec{v} : \llbracket \text{'}\Pi^{<\omega}\text{' } n \ T \rrbracket \vec{T} \ o}
\end{array}$$

This restriction is one that is easy to bear, only the construction of ‘ \rightarrow ’ and ‘**Ord**’ need the infinite pi in the examples given above and there is an obvious embedding of the finite into the infinite:

$$\begin{aligned} \cdot\Pi^{\leftarrow\omega},_n T : \mathbf{RF} \vec{I} O &\mapsto \cdot\Pi'_{(O \times \mathbf{Fin} \ n)} \mathbf{fst} T : \mathbf{SPF} \vec{I} O \\ \pi^{\leftarrow\omega} \vec{v} : \llbracket \cdot\Pi^{\leftarrow\omega},_n T \rrbracket \vec{T} o &\mapsto \\ \pi(\lambda(o; i); \mathbf{refl} \Rightarrow \vec{v} i) : \llbracket \cdot\Pi'_{(O \times \mathbf{Fin} \ n)} \mathbf{fst} T \rrbracket \vec{T} o & \end{aligned}$$

In the **RF** universe it is possible to define a generic equality which is structurally recursive on the data.

$$\begin{aligned} \text{let } & \frac{bs : \mathbf{Fin} \ n \rightarrow \mathbf{Bool}}{\mathbf{all}_n bs : \mathbf{Bool}} \\ & \mathbf{all}_n bs \leftarrow \mathbf{rec} \ n \\ & \mathbf{all}_0 \quad bs \Rightarrow \mathbf{true} \\ & \mathbf{all}_{(1+ \ n)} bs \Rightarrow (bs \ 0) \wedge (\mathbf{all} (bs \ . \ 1+)) \\ \\ \text{let } & \frac{T : \mathbf{RF} \vec{I} O \quad a : \llbracket T \rrbracket \vec{T} ob \quad b : \llbracket T \rrbracket \vec{T} ob}{\mathbf{gEq} \ a \ b : \mathbf{Bool}} \end{aligned}$$

$$\begin{aligned} \mathbf{gEq} \ a \ b &\leftarrow \mathbf{rec} \ a \\ \mathbf{gEq} \ (\mathbf{top} \ a) \ (\mathbf{top} \ b) &\Rightarrow \mathbf{gEq} \ a \ b \\ \mathbf{gEq} \ (\mathbf{pop} \ a) \ (\mathbf{pop} \ b) &\Rightarrow \mathbf{gEq} \ a \ b \\ \mathbf{gEq} \ (\mathbf{tag}_{ta} \ a) \ (\mathbf{tag}_{tb} \ b) &\left| \begin{array}{l} ta == tb \\ \mathbf{yes} \ (\mathbf{refl} \ t) \\ \mathbf{no} \ _ \end{array} \right. \begin{array}{l} \Rightarrow \mathbf{gEq} \ a \ b \\ \Rightarrow \mathbf{false} \\ \Rightarrow \mathbf{true} \end{array} \\ \mathbf{gEq} \ \mathbf{void} \ \mathbf{void} &\Rightarrow \mathbf{true} \\ \mathbf{gEq} \ (\sigma_{oa} \ a) \ (\sigma_{ob} \ b) &\Rightarrow \mathbf{gEq} \ a \ b \\ \mathbf{gEq} \ (\delta \ a) \ (\delta \ b) &\Rightarrow \mathbf{gEq} \ a \ b \\ \mathbf{gEq} \ (\pi^{\leftarrow\omega} \vec{a}) \ (\pi^{\leftarrow\omega} \vec{b}) &\Rightarrow \wedge (\forall i : \mathbf{Fin} \ n \Rightarrow \mathbf{gEq} \ \vec{a} \ i \ \vec{b} \ i) \\ \mathbf{gEq} \ (\mathbf{in} \ a) \ (\mathbf{in} \ b) &\Rightarrow \mathbf{gEq} \ a \ b \end{aligned}$$

Notice that we can decide the equality of values in a purely syntactic manner, in fact this test equates values at different output indexes as long as the syntax is the same (so for instance $0 : \mathbf{Fin} \ n = 0 : \mathbf{Fin} \ (1+ \ n)$). In practice it would be better to restrict ourselves only to comparing things for equality at the same index.

5. Conclusions and Further Work

We have here closed the gap left open in our previous work [9] and show how to construct a universe which captures strictly positive families, i.e. all of Epigram’s datatypes. This includes the universe itself, since **SPF**, **dTel**, $\llbracket _ \rrbracket$ are strictly positive families and can be laboriously encoded as instances of **SPF** — for reasons of space, and sanity, we refrain from spelling out the details here. One may wonder whether such a circular construction is necessarily paradoxical, as in Girard’s paradox which rules out a type of all types in a consistent theory. However, closer inspection shows that this is not the case but that the universe construction raises the level of the predicative hierarchy of types $\star_n : \star_{n+1}$, which is ignored in the current implementation of Epigram.

We have only touched the potential of the universe presented here. E.g. instead of generic simply typed programs like **gMap** or **gEq** we should consider their dependently typed counterparts, which integrate generic proofs and programs. In the case of generic map this involves defining the modality \square which, in logic (and here, if we appeal to the Curry-Howard isomorphism), lifts a predicate $P : A \rightarrow \star$ over a collection $F : A \Rightarrow \star$ which holds iff P holds for every A position in an $x : F \ A$. For instance $\square\mathbf{List}$:

$$\begin{aligned} \text{data } & \frac{P : A \rightarrow \star \quad as : \mathbf{List} \ A}{\square\mathbf{List} \ P \ as : \star} \text{ where} \\ & \frac{p : P \ a \quad ps : \square\mathbf{List} \ P \ as}{\varepsilon : \square\mathbf{List} \ P \ A \ \varepsilon} \quad \frac{p : P \ a \quad ps : \square\mathbf{List} \ P \ as}{p : ps : \square\mathbf{List} \ P \ (a : as)} \end{aligned}$$

Given a dependent function $f : \forall x : A \Rightarrow B \ a$, **mapList** f then has the type $\forall as : \mathbf{List} \ A \Rightarrow \square\mathbf{List} \ B \ as$. It turns out that we can write a function \square that calculates a code for this modality for any **SPF**, meaning that dependent map is also a generic program. In the case of generic equality between **RF**s we can not only give a function into **Bool** but give a function that returns the evidence of equality or inequality between the arguments much as we’ve done for the regular tree types.

In joint, so far unpublished work with Ghani, Hancock and McBride [2] we have shown how to extend the notion of container types [1] to cover strictly positive families, arriving at a semantic interpretation of our universe which gives an alternative access to generic programming with and for dependent types.

We plan to implement a reflection principle based on the universe construction presented here in the forthcoming new release of Epigram, thus giving the Epigram programmer full access to all datatypes definable in Epigram.

References

- [1] M. Abbott, T. Altenkirch, and N. Ghani. Containers - constructing strictly positive types. *Theoretical Computer Science*, 342:3–27, September 2005. Applied Semantics: Selected Topics.
- [2] T. Altenkirch, N. Ghani, P. Hancock, C. McBride, and P. Morris. Indexed containers. Manuscript, available online, February 2006.
- [3] T. Altenkirch, C. McBride, and J. McKinna. Why dependent types matter. Manuscript, available online, April 2005.
- [4] M. Benke, P. Dybjer, and P. Jansson. Universes for generic programs and proofs in dependent type theory. *Nordic Journal of Computing*, 10:265–269, 2003.
- [5] P. Dybjer and A. Setzer. Indexed induction-recursion. In R. Kahle, P. Schroeder-Heister, and R. F. Stärk, editors, *Proof Theory in Computer Science*, volume 2183 of *Lecture Notes in Computer Science*, pages 93–113. Springer, 2001.
- [6] P. Martin-Löf. *Intuitionistic Type Theory*. Bibliopolis-Napoli, 1984.
- [7] C. McBride. Epigram, 2004. <http://www.e-pig.org/>.
- [8] C. McBride and J. McKinna. The view from the left. *Journal of Functional Programming*, 14(1), 2004.
- [9] P. Morris, T. Altenkirch, and C. McBride. Exploring the regular tree types. In C. P.-M. Jean-Christophe Filliatre and B. Werner, editors, *Types for Proofs and Programs (TYPES 2004)*, Lecture Notes in Computer Science, 2006.
- [10] B. Nordström, K. Petersson, and J. Smith. *Programming in Martin-Löf’s Type Theory: an introduction*. Oxford University Press, 1990.