

# Computing with Generic Trees in Agda

Stephen Dolan  
stedolan@stedolan.net

## Abstract

Dependently-typed programming languages offer powerful new means of abstraction, allowing the programmer to work generically across data structures. However, using the standard generic encoding of tree-like data structures (the *W-types*), we soon notice a caveat: the computational behaviour of *W-types* does not quite match their first-order counterparts. Here, we show how a tweak to the definition of *W-types* avoids this caveat, making the generic definition work just as well as the direct one.

### ACM Reference Format:

Stephen Dolan. 2022. Computing with Generic Trees in Agda. In *Proceedings of ACM Conference (Conference'17)*. ACM, New York, NY, USA, 6 pages. <https://doi.org/10.1145/nnnnnnnn.nnnnnnnn>

## 1 Introduction

Part of the promise of dependent types is that the ability to abstract over types just as easily as one can abstract over values makes generic programming straightforward. One example is the definition of *W-types*, which in a single stroke defines a whole family of tree-like data structures:

```
data W (Sh : Set) (Pl : Sh → Set) : Set where
  sup : (s : Sh) → (Pl s → W Sh Pl) → W Sh Pl
```

This type represents tree-shaped data generically. A tree datatype is given by a set *Sh* of *shapes*, describing the possible kinds of node, and for each shape *s* a set *Pl s* of *places*, listing the subtrees of such nodes. A tree of such a type is then given as a node of a specified shape, with one subtree per place of that shape. (This is the least fixed point of a *container* [2], from where the “shapes and places” terminology arises)

This is the simplest form of *W-type*, generically representing a single recursive datatype with no parameters and no indices. While this is enough to illustrate the point of this paper, note that the idea has been generalised much further, covering nested types [1], indices [3, 5], and more)

The trouble with this definition, at least in standard intensional type theory, is that what would normally be a record of several values (“one subtree for each place”) is instead encoded as a function (“a function from places to subtrees”), and this causes difficulties with equality.

### 1.1 Encoding records as functions

A pair  $A \times A$  can be encoded as a function  $2 \rightarrow A$ , where 2 is the type containing two elements. (Indeed, both are often written  $A^2$ ).

We can try this in Agda, implementing construction and projection functions for pairs-as-functions:

```
Pair : Set → Set
```

```
Pair A = (Bool → A)
```

```
make : {A : Set} → A → A → Pair A
```

```
make a b = λ { false → a; true → b }
```

```
proj1 proj2 : {A : Set} → Pair A → A
```

```
proj1 p = p false
```

```
proj2 p = p true
```

The problem arises when we consider equality. The  $\eta$ -equality rule for pairs states:

$$p \equiv (\text{proj}_1 p, \text{proj}_2 p)$$

But in our functional encoding in Agda, we’d need to show

$$p \equiv \lambda \{ \text{false} \rightarrow p \text{ false}; \text{true} \rightarrow p \text{ true} \}$$

This does not hold definitionally. Agda compares the two functions by comparing them after applying an abstract argument  $x : \text{Bool}$ , but there is no  $\eta$  rule for **Bool** which would allow it to continue by case analysis on  $x$ .

If function extensionality is available propositionally (e.g. because it is postulated, or because we’re working in a system like HoTT where it is provable), then we can prove  $\eta$ -equality for functional pairs. However, this is less useful than the definitional equality of native pairs, since it is not used in computation and must be explicitly appealed to.

One could imagine adding special-case rules to Agda for definitional equality at type  $2 \rightarrow A$ , by comparing two functions at arguments **true** and **false**. This approach does not generalise, however, because of the following example due to McBride [7]:

Consider the functional encoding of the empty tuple, or the unit type. An tuple of no  $A$  is encoded as a function  $0 \rightarrow A$  (where 0 is the empty type), and by the  $\eta$  rule for empty records we expect any two such functions to be equal. In particular, this means that in an arbitrary context  $\Gamma$ :

$$\Gamma \vdash (\lambda x. \text{true}) \equiv (\lambda x. \text{false}) : 0 \rightarrow 2$$

If there is some  $e$  such that  $\Gamma \vdash e : 0$  (that is, if  $\Gamma$  is an inconsistent context), then we have:

$$\Gamma \vdash (\lambda x. \text{true})e \equiv (\lambda x. \text{false})e : 2$$

$$\Gamma \vdash \text{true} \equiv \text{false} : 2$$

---

Conference’17, July 2017, Washington, DC, USA  
2022. ACM ISBN 978-x-xxxx-xxxx-x/YY/MM...\$15.00  
<https://doi.org/10.1145/nnnnnnnn.nnnnnnnn>

So, deciding whether `true`  $\equiv$  `false` means first deciding whether  $\Gamma$  is inconsistent. Since the latter is undecidable, we have broken decidability of definitional equality.

## 1.2 Induction on Nat

These difficulties with functional encoding of records crop up when we try to write the induction principles for  $W$ -types. For `Nat`, for instance, we expect to end up with:

```
nat-ind :
  (R : Nat → Set)
  (Rzero : R zero)
  (Rsucc : ∀ n → R n → R (succ n)) →
  ∀ x → R x
```

But when we try to write the case for the zero shape, with an empty set of places, we find that our provided `Rzero` does not apply directly, because the two empty collections of subtrees are not definitionally equal. We can prove them equal propositionally, but then we lose the computation rule:

```
nat-ind R Rzero Rsucc zero  $\equiv$  Rzero
```

## 1.3 Contribution

The contribution of this paper is to write the induction principle for  $W$ -types generically, and have it compute, inside intensional type theory.

It is not a new result that this can be done by restricting to *finitary*  $W$ -types, where each tree has only finitely many subtrees. (These are fixed points of what Girard calls *normal functors* [6], rather than fixed points of general containers). In this case, one can encode the subtrees as a finite vector (cf. McBride [8]), with the expected computation rules.

While this approach is not novel, we present an alternative construction of it in section 3, based on a universe of finite types defined in section 2.

The new result here is to show that a small generalisation of the technique can work for infinitary  $W$ -types as well. In section 4, we introduce *partitioned sets*, which are disjoint unions of finitely many sets (which need not themselves be finite). By using partitioned sets rather than finite sets as the set of places, we can describe even infinitary  $W$ -types that still compute, as demonstrated in section 5.

The above takes place using the proof assistant Agda, and this paper is a literate Agda script.

## 2 Codes for Finite Types

Our goal for this section is to encode function types  $A \rightarrow B$  as records, for finite  $A$ , and we begin by writing a type of codes for finite types, following the universes approach explored by Benke et al. [4].

The most straightforward choice is to use  $\mathbb{N}$ , which has exactly one representative for each finite cardinality. However, we do not want our finite types to be unique up to cardinality, as a type of exactly two elements is not necessarily `Bool`.

So instead, we define finite codes as containing the empty type, singletons, and being closed under sum, and we allow singletons to be named:

```
infixr 20 _+_
data Fin : Set where
  none : Fin
  one : String → Fin
  _+_ : Fin → Fin → Fin
```

The finite types themselves are defined by interpreting `Fin` into `Set`:

```
record NamedUnit (Name : String) {l} : Set l where
  constructor tt

[[_]] : Fin → Set
[ none ] = ⊥
[ one name ] = NamedUnit name
[ A + B ] = [ A ] ∪ [ B ]
```

We use the `NamedUnit` type to ensure that distinct codes have distinct interpretations. This condition is not semantically important, but aids Agda's typechecking since in many cases it is then able to uniquely deduce the code from the interpretation, allowing us to mostly leave codes as implicit arguments to be filled in automatically.

Our named singletons mean we can write finite types with named members:

```
foobarbaz : Fin
foobarbaz = one "foo" + one "bar" + one "baz"
```

Sadly, its inhabitants have names like `inj2 (inj1 tt)` instead of "bar". To let us make use of the names, we add a convenience function for looking up elements by name:

```
lookup : (A : Fin) → String → Maybe [ A ]
lookup none s = nothing
lookup (one t) s with primStringEquality t s
... | false = nothing
... | true = just tt
lookup (A + B) s with lookup A s
... | just x = just (inj1 x)
... | nothing with lookup B s
... | just x = just (inj2 x)
... | nothing = nothing
```

as well as some syntactic trickery for making use of it:

```
data Found : Set where
  } : Found

inhab : ∀ {A : Set} → Maybe A → Set
inhab nothing = ⊥
inhab (just _) = Found

⟨ : ∀ {A : Fin} → (s : String) → inhab (lookup A s) → [ A ]
⟨ {A} s with lookup A s
```

```

221 ... | nothing = λ ()
222 ... | just x = λ _ → x

```

Now, we can refer to inhabitants of `foobarbaz` compactly:

```

225 bar : [[ foobarbaz ]]
226 bar = < "bar" >

```

The trick here is that the second argument to `<` is of type `inhab (lookup A s)`, which is uninhabited if `lookup` fails, but inhabited by `>` if it succeeds.

## 2.1 Universe polymorphism and generalization

We are going to use these finite codes to describe both values and types, and to allow the same definitions to be used for both we employ Agda's *universe polymorphism*. We are not making much use of this powerful feature: the only universe levels we actually use are 0 and 1, and we could get the same effect by duplicating most definitions.

From here on, the number of quantified variables in our types increases, so to remove some clutter we allow Agda to implicitly generalise “*l*” as a universe level and “*A*” as a finite code:

```

243 variable l : Level
244 variable A : Fin

```

## 2.2 Function types with finite domain

Next, we define functions with finite domain, by recursion on the code of their domain:

```

250 record One (Name : String) (S : Set l) : Set l where
251   constructor v
252   field contents : S

```

```

254 _→°_ : Fin → Set l → Set l

```

```

255 none →° S = T

```

```

256 one name →° S = One name S

```

```

257 (A + B) →° S = (A →° S) × (B →° S)

```

As before, using the `One` type ensures distinct codes have distinct interpretations, improving inference. We also write an alternative constructor for `One`, and a convenience function for proving equations on it.

```

263 _↦_ : (Name : String) → {S : Set l} → S → One Name S

```

```

264 _↦ x = v x

```

```

265 ≡/v : ∀ {n} {A : Set l} {a a' : A} →

```

```

266   (a ≡ a') → n ↦ a ≡ n ↦ a'

```

```

268 ≡/v refl = refl

```

The purpose of `↦` is to let us use explicit names when writing functions with finite domain:

```

272 is-bar : foobarbaz →° Bool

```

```

273 is-bar =

```

```

274   ("foo" ↦ false),

```

```

276   ("bar" ↦ true),

```

```

277   ("baz" ↦ false)

```

These names are redundant, since typechecking works by position rather than by name: in each occurrence of `↦`, there is only one string that can appear on the left and Agda already knows what it is. However, being able to write these names (and have them checked) makes the code readable.

Having defined `→°`, we now define generic introduction and elimination forms (lambda-abstraction and application):

```

286 λ° : {S : Set l} → ([[ A ]] → S) → (A →° S)

```

```

287 λ° {A = none} f = tt

```

```

288 λ° {A = one _} f = v (f tt)

```

```

289 λ° {A = A + B} f = λ° (f ∘ inj₁), λ° (f ∘ inj₂)

```

```

290 _<°_ : {S : Set l} → (A →° S) → ([[ A ]] → S)

```

```

291 _<°_ {A = one _} (v f) tt = f

```

```

292 _<°_ {A = A + B} (f, g) (inj₁ x) = f <° x

```

```

293 _<°_ {A = A + B} (f, g) (inj₂ x) = g <° x

```

These function types have slightly different computation rules than the usual, as their underlying implementation is as records rather than as functions. In particular, the  $\beta$  and  $\eta$  rules for functions no longer hold: we do not have  $(\lambda x.f)e \equiv f[x/e]$  nor  $f \equiv \lambda x.f x$  definitionally in general. However, these rules are provable (that is, they hold propositionally):

```

301 beta° : {S : Set l} (f : [[ A ]] → S) (x : [[ A ]]) →

```

```

302   (λ° f <° x) ≡ f x

```

```

303 beta° {A = one _} f tt = refl

```

```

304 beta° {A = A + B} f (inj₁ x) = beta° (f ∘ inj₁) x

```

```

305 beta° {A = A + B} f (inj₂ x) = beta° (f ∘ inj₂) x

```

```

306 eta° : {S : Set l} (f : A →° S) →

```

```

307   f ≡ λ° λ x → f <° x

```

```

308 eta° {A = none} tt = refl

```

```

309 eta° {A = one _} (v x) = refl

```

```

310 eta° {A = A + B} (f, g) = ≡/, (eta° f) (eta° g)

```

Additionally,  $\beta$ -equality holds definitionally as long as the code of the domain and the argument are both in canonical form, while  $\eta$ -equality holds definitionally when the code of the domain is canonical. So these functions do compute, but not when their types or arguments are stuck. That is, while  $(\lambda f \lambda x)$  does not reduce with  $x$  a variable, an application to a concrete argument like  $\lambda f \lambda x (\langle "bar" \rangle)$  reduces to  $f (\langle "bar" \rangle)$  for any  $f$  of type `[[ foobarbaz ]] → Set`.

A useful property that these functions also have is that extensionality is provable:

```

323 ext° : {S : Set l} (f g : [[ A ]] → S) →

```

```

324   (eq : ∀ x → f x ≡ g x) →

```

```

325   λ° f ≡ λ° g

```

```

326 ext° {A = none} f g eq = refl

```

```

327 ext° {A = one _} f g eq = ≡/v (eq tt)

```

```

328 ext° {A = A + B} f g eq =

```

331  $\equiv/, (\text{ext}^\circ (f \circ \text{inj}_1) (g \circ \text{inj}_1) (eq \circ \text{inj}_1))$   
 332  $(\text{ext}^\circ (f \circ \text{inj}_2) (g \circ \text{inj}_2) (eq \circ \text{inj}_2))$

333 Again, this holds definitionally for canonical domain codes.  
 334

### 335 2.3 Dependent functions with finite domain

336 Next, we generalise from simple function types  $\rightarrow^\circ$  to depen-  
 337 dent ones  $\Pi^\circ$ , where the type of the result may depend on  
 338 the argument. These are semantically tuples, consisting of  
 339 finitely many values of different types.  
 340

341 We need to define  $\rightarrow^\circ$  and  $\Pi^\circ$  separately, because the for-  
 342 mer is used in the definition of the latter: the result type of  
 343  $\Pi^\circ$  is given as a finitary function into **Set**:

344  $\Pi^\circ : (A : \mathbf{Fin}) (U : A \rightarrow^\circ \mathbf{Set} \ l) \rightarrow \mathbf{Set} \ l$

345  $\Pi^\circ \ \text{none} \ U = \top$

346  $\Pi^\circ \ (\text{one } \text{name}) \ (v \ U) = \mathbf{One} \ \text{name} \ U$

347  $\Pi^\circ \ (A + B) \ (U, V) = (\Pi^\circ \ A \ U) \times (\Pi^\circ \ B \ V)$

348 As above, we have abstraction and application:  
 349

350  $\Lambda^\circ : \{U : A \rightarrow^\circ \mathbf{Set} \ l\} \rightarrow ((x : \llbracket A \rrbracket) \rightarrow U \triangleleft^\circ x) \rightarrow (\Pi^\circ \ A \ U)$

351  $\Lambda^\circ \ \{A = \text{none}\} \ f = \text{tt}$

352  $\Lambda^\circ \ \{A = \text{one } \_ \} \ f = v \ (f \ \text{tt})$

353  $\Lambda^\circ \ \{A = A + B\} \ f = \Lambda^\circ \ (f \circ \text{inj}_1), \Lambda^\circ \ (f \circ \text{inj}_2)$

354  $\_ \triangleleft^\circ \_ : \{U : A \rightarrow^\circ \mathbf{Set} \ l\} \rightarrow (\Pi^\circ \ A \ U) \rightarrow (a : \llbracket A \rrbracket) \rightarrow U \triangleleft^\circ a$

355  $\_ \triangleleft^\circ \_ \ \{A = \text{one } \_ \} \ (v \ f) \ x = f$

356  $\_ \triangleleft^\circ \_ \ \{A = A + B\} \ (f, g) \ (\text{inj}_1 \ x) = f \triangleleft^\circ x$

357  $\_ \triangleleft^\circ \_ \ \{A = A + B\} \ (f, g) \ (\text{inj}_2 \ x) = g \triangleleft^\circ x$

358 The  $\beta$  and  $\eta$  rules hold (with the same caveats about compu-  
 359 tation as before), as does extensionality:  
 360

361  $\text{Beta}^\circ : \{U : A \rightarrow^\circ \mathbf{Set} \ l\} \rightarrow (f : ((x : \llbracket A \rrbracket) \rightarrow U \triangleleft^\circ x)) \rightarrow$

362  $(a : \llbracket A \rrbracket) \rightarrow \Lambda^\circ \ f \triangleleft^\circ a \equiv f \ a$

363  $\text{Eta}^\circ : \{U : A \rightarrow^\circ \mathbf{Set} \ l\} \rightarrow (f : \Pi^\circ \ A \ U) \rightarrow$

364  $f \equiv \Lambda^\circ \ \lambda x \rightarrow f \triangleleft^\circ x$

365  $\text{Ext}^\circ : \{U : A \rightarrow^\circ \mathbf{Set} \ l\} \rightarrow (f \ g : (x : \llbracket A \rrbracket) \rightarrow U \triangleleft^\circ x) \rightarrow$

366  $(eq : \forall x \rightarrow f \ x \equiv g \ x) \rightarrow \Lambda^\circ \ f \equiv \Lambda^\circ \ g$

367 The proofs are omitted, as they are identical to those for  
 368 simple function types.  
 369

## 371 3 Finitary W-types

372 Having function types with finite domains available, we are  
 373 now able to generically describe finite trees, by taking the  
 374 definition of W-types above and replacing  $\rightarrow$  with  $\rightarrow^\circ$ :  
 375

376  $\text{data } \mathbf{W}^\circ (Sh : \mathbf{Set}) (Pl : Sh \rightarrow \mathbf{Fin}) : \mathbf{Set} \ \text{where}$

377  $\ \ \ \ \ \text{sup} : (sh : Sh) \rightarrow (Pl \ sh \rightarrow \mathbf{W}^\circ \ Sh \ Pl) \rightarrow \mathbf{W}^\circ \ Sh \ Pl$   
 378

379 Note that this being accepted as a strictly positive inductive  
 380 type relies on the precise definition of  $\rightarrow^\circ$ .

381 The full eliminator for W-types is a bit of a mouthful. It  
 382 states that to compute a result  $R$  for all trees of a given W-  
 383 type, it suffices to compute  $R$  for every tree of the form  $\text{sup}$   
 384  $sh \ \text{sub}$ , given  $R$  is already computed for each subtree in  $\text{sub}$ .

The type is almost that of the standard eliminator, ex-  
 cept uses finitary function spaces  $\rightarrow^\circ$  and  $\Pi^\circ$  instead of the  
 usual ones. The implementation is slightly different, doing  
 an explicit recursion on the set of places to ensure that the  
 recursion is structural:

386  $\text{elim}^\circ : \forall \{Sh \ Pl\} (R : \mathbf{W}^\circ \ Sh \ Pl \rightarrow \mathbf{Set}) \rightarrow$

387  $(F : (sh : Sh) \rightarrow$

388  $(sub : Pl \ sh \rightarrow^\circ \mathbf{W}^\circ \ Sh \ Pl) \rightarrow$

389  $(subR : \Pi^\circ (Pl \ sh) (\lambda^\circ \lambda \ p \rightarrow R (sub \triangleleft^\circ p))) \rightarrow$

390  $R (\text{sup } sh \ sub)) \rightarrow$

391  $(x : \mathbf{W}^\circ \ Sh \ Pl) \rightarrow R \ x$

392  $\text{elim}^\circ \ \{Sh\} \ \{Pl\} \ R \ F \ (\text{sup } sh \ t) = F \ sh \ t \ (\text{IH } t)$

393 where

394  $\text{IH} : \forall \{Ps\} \rightarrow (t : Ps \rightarrow^\circ \mathbf{W}^\circ \ Sh \ Pl) \rightarrow$

395  $\Pi^\circ \ Ps \ (\lambda^\circ (\lambda \ p \rightarrow R (t \triangleleft^\circ p)))$

396  $\text{IH} \ \{\text{none}\} \ t = \text{tt}$

397  $\text{IH} \ \{\text{one } n\} \ (v \ t) = n \mapsto \text{elim}^\circ \ R \ F \ t$

398  $\text{IH} \ \{Ps_1 + Ps_2\} \ (t_1, t_2) = \text{IH } t_1, \text{IH } t_2$

399 As an example, we implement the Peano natural numbers,  
 400 which are written in Agda directly as:  
 401

402  $\text{data } \mathbf{Nat} : \mathbf{Set} \ \text{where}$

403  $\ \ \ \ \ \text{zero} : \mathbf{Nat}$

404  $\ \ \ \ \ \text{succ} : (x : \mathbf{Nat}) \rightarrow \mathbf{Nat}$

405 The set of shapes of a  $\mathbf{W}^\circ$ -type is an arbitrary **Set**, so we are  
 406 not obliged to use **Fin**. In this case it happens to be finite,  
 407 making it convenient to use **Fin** anyway. We use a helper  
 408 function to easily eliminate **Fin** in ordinary functions:  
 409

410  $\text{cases} : \{B : \llbracket A \rrbracket \rightarrow \mathbf{Set} \ l\} \rightarrow (\Pi^\circ \ A \ (\lambda^\circ B)) \rightarrow$

411  $(a : \llbracket A \rrbracket) \rightarrow B \ a$

412  $\text{cases } f \ a = \text{transp} \ (\text{beta}^\circ \_ \ a) \ (f \triangleleft^\circ a)$

413 Then, the definition of **Nat** is:  
 414

415  $\mathbf{Nat} = \mathbf{W}^\circ \ \llbracket \text{one } \text{"zero"} + \text{one } \text{"succ"} \rrbracket$

416  $(\text{cases} \ ($

417  $\ \ \ \ \ \text{"zero"} \mapsto \text{none} ,$

418  $\ \ \ \ \ \text{"succ"} \mapsto \text{one } \text{"x"})$

419 and the constructors are:  
 420

421  $\text{zero} : \mathbf{Nat}$

422  $\text{zero} = \text{sup} \ (\langle \text{"zero"} \rangle) \ \text{tt}$

423  $\text{succ} : \mathbf{Nat} \rightarrow \mathbf{Nat}$

424  $\text{succ } x = \text{sup} \ (\langle \text{"succ"} \rangle) \ (\text{"x"} \mapsto x)$

425 The usual induction principle for  $\mathbb{N}$  is now definable by  
 426 appeal to  $\text{elim}^\circ$ :  
 427

428  $\text{nat-ind} :$

429  $(R : \mathbf{Nat} \rightarrow \mathbf{Set})$

430  $(Pzero : R \ \text{zero})$

431  $(Psucc : \forall n \rightarrow R \ n \rightarrow R \ (\text{succ } n)) \rightarrow$

432  $\forall x \rightarrow R \ x$

```

441 nat-ind R Pzero Psucc =
442   elim° R (cases (
443     "zero" ↦ (λ _ _ → Pzero) ,
444     "succ" ↦ λ { (v x) (v Rx) → Psucc x Rx })
445

```

The point of this exercise is that unlike with plain W-types, the induction principle so defined has the right computation behaviour, in particular having the right behaviour on zero:

```

449 nat-ind-zero : ∀ { R Pzero Psucc } →
450   nat-ind R Pzero Psucc zero ≡ Pzero
451 nat-ind-zero = refl
452
453 nat-ind-succ : ∀ { R Pzero Psucc x } →
454   nat-ind R Pzero Psucc (succ x)
455   ≡ Psucc x (nat-ind R Pzero Psucc x)
456 nat-ind-succ = refl

```

The important thing is not that these are true, but that they are true by `refl`: our definition of `nat-ind` computes.

## 4 Partitioned Sets

Next, we generalise from finite sets to *partitioned sets*, which are disjoint unions of finitely many components, where the components themselves need not be finite:

```

465 record PSet : Set₁ where
466   constructor pset
467   field
468     parts : Fin
469     elems : parts →° Set

```

```

471 [ ]* : PSet → Set
472 [ pset none E ]* = ⊥
473 [ pset (one name) (v E) ]* = NamedUnit name × E
474 [ pset (P + Q) (E , F) ]* = [ pset P E ]* ∪ [ pset Q F ]*

```

An element of a partitioned set can be constructed by specifying a component and a member of that component:

```

478 el : ∀ { P E } → (p : [ P ]*) → (E ◁° p) → [ pset P E ]*
479 el {P = one x} p e = tt , e
480 el {P = P + Q} (inj₁ p) e = inj₁ (el p e)
481 el {P = P + Q} (inj₂ q) e = inj₂ (el q e)

```

### 4.1 Functions on Partitioned Sets

Mirroring the definitions of  $\rightarrow^\circ$  and  $\Pi^\circ$  earlier, we define  $\rightarrow^*$  and  $\Pi^*$  as functions with partitioned rather than finite domain. The definitions are essentially curried, where a function from a partitioned set  $P$  to a set  $S$  is a finitary function from the components of  $P$  to an ordinary function from the members of that component to  $S$ :

```

491 →* _ : ∀ { l } → PSet → Set l → Set l
492 pset none tt →* S = ⊤
493 pset (one name) (v E) →* S =
494   One name (E → S)

```

```

496 pset (P + Q) (E , F) →* S =
497   ((pset P E) →* S) × ((pset Q F) →* S)

```

```

498 Π* : (X : PSet) (M : X →* Set l) → Set l
499 Π* (pset none tt) M = ⊤
500 Π* (pset (one name) (v E)) (v M) =
501   One name ((x : E) → M x)
502 Π* (pset (P + Q) (E , F)) (M , N) =
503   Π* (pset P E) M × Π* (pset Q F) N

```

As before, these can be introduced and eliminated with abstraction and application operators:

```

504 variable X : PSet

```

```

505 λ* : {S : Set l} → ([ X ]* → S) → (X →* S)
506 λ* {X = pset none tt} f = tt
507 λ* {X = pset (one n) (v E)} f = n ↦ λ x → f (tt , x)
508 λ* {X = pset (P + Q) (E , F)} f = λ* (f ◦ inj₁) , λ* (f ◦ inj₂)
509
510 _◁* _ : {S : Set l} → (X →* S) → ([ X ]* → S)
511 _◁* _ {X = pset (one _) (v E)} (v f) (tt , e) = f e
512 _◁* _ {X = pset (P + Q) (E , F)} (f , g) (inj₁ x) = f ◁* x
513 _◁* _ {X = pset (P + Q) (E , F)} (f , g) (inj₂ x) = g ◁* x
514
515 Λ* : {M : X →* Set l} → ((x : [ X ]*) → M ◁* x) →
516   Π* X M
517
518 Λ* {X = pset none tt} f = tt
519 Λ* {X = pset (one n) E} f = n ↦ λ x → f (tt , x)
520 Λ* {X = pset (P + Q) (E , F)} f = Λ* (f ◦ inj₁) , Λ* (f ◦ inj₂)

```

```

521 _◁* _ : {M : X →* Set l} → (Π* X M) →
522   (x : [ X ]*) → M ◁* x
523
524 _◁* _ {X = pset (one n) (v E)} (v f) (tt , e) = f e
525 _◁* _ {X = pset (P + Q) (E , F)} (f , g) (inj₁ x) = f ◁* x
526 _◁* _ {X = pset (P + Q) (E , F)} (f , g) (inj₂ x) = g ◁* x

```

Agda's termination checker is relatively generous here, by accepting `pset P E` as smaller than `pset (P + Q) (E , F)` – with other typecheckers, we may have had to use a separate recursion on  $P$  to make the recursion count as structural.

We again have  $\beta$  and  $\eta$ , with essentially identical proofs to the finite case:

```

531 beta* : {S : Set l} (f : [ X ]* → S) →
532   (x : [ X ]*) → (λ* f ◁* x) ≡ f x
533
534 eta* : {S : Set l} (f : X →* S) →
535   f ≡ λ* λ x → f ◁* x
536
537 Beta* : {U : X →* Set l} → (f : ((x : [ X ]*) → U ◁* x)) →
538   (a : [ X ]*) → Λ* f ◁* a ≡ f a
539
540 Eta* : {U : X →* Set l} → (f : Π* X U) →
541   f ≡ Λ* λ x → f ◁* x

```

However, lacking function extensionality, we no longer have the `ext` and `Ext` rules, as showing equality of functions on partitioned sets requires both equality of their finite, record-based part and their possibly-infinite, functional part.

## 5 Infinitary W-types with Partitioned Sets

Finally, we are ready to implement W-types with a partitioned set of places:

```
data W* (Sh : Set) (Pl : Sh → PSet) : Set where
  sup : (sh : Sh) → (Pl sh →* W* Sh Pl) → W* Sh Pl
```

The eliminator is identical to that of  $W^\circ$ , with  $*$  replacing  $^\circ$ :

```
elim* : ∀ {Sh Pl} (R : W* Sh Pl → Set) →
  (F : (sh : Sh) →
    (sub : (Pl sh) →* W* Sh Pl) →
    (subR : Π* (Pl sh) (λ* λ p → R (sub <* p))) →
    R (sup sh sub)) →
  (x : W* Sh Pl) → R x
elim* {Sh} {Pl} R F (sup sh t) = F sh t (IH t)
where
  IH : ∀ {Ps} → (t : Ps →* W* Sh Pl) →
    Π* Ps (λ* (λ p → R (t <* p)))
  IH {pset none Es} t = tt
  IH {pset (one n) Es} (v t) = n ↦ λ e → elim* R F (t e)
  IH {pset (Ps1 + Ps2) Es} (t1, t2) = IH t1, IH t2
```

As an example, we code the Brouwer ordinal trees, which are defined by two finitary shapes and one infinitary one:

```
data Ord : Set where
  ozero : Ord
  osucc : Ord → Ord
  olim : (Nat → Ord) → Ord
```

Using  $W^*$ , this translates to:

```
Ord = W* [ [ one "zero" + one "succ" + one "lim" ] ]
  (cases (
    "zero" ↦ pset none tt,
    "succ" ↦ pset (one "x") ("x" ↦ ⊤),
    "lim" ↦ pset (one "f") ("f" ↦ Nat)))
```

with constructors:

```
ozero : Ord
ozero = sup (<"zero">) tt
osucc : Ord → Ord
osucc x = sup (<"succ">) ("x" ↦ λ _ → x)
olim : (Nat → Ord) → Ord
olim f = sup (<"lim">) ("f" ↦ f)
```

and an induction principle (defined using the general  $\text{elim}^*$ ):

```
ord-ind :
  (R : Ord → Set)
  (Pzero : R ozero)
  (Psucc : ∀ n → R n → R (osucc n))
  (Plim : ∀ f → (∀ n → R (f n)) → R (olim f)) →
  ∀ x → R x
ord-ind R Pzero Psucc Plim =
```

```
elim* R (cases (
  "zero" ↦ (λ _ _ → Pzero),
  "succ" ↦ (λ { (v x) (v Rx) → Psucc (x tt) (Rx tt) },
  "lim" ↦ (λ { (v f) (v Rf) → Plim f Rf })))
```

Like  $\text{nat-ind}$  previously, this eliminator computes, as shown by the following properties (where again, the point is not so much that they are true but that they are true by  $\text{refl}$ ):

```
ord-ind-zero : ∀ { R Pzero Psucc Plim } →
  ord-ind R Pzero Psucc Plim ozero ≡ Pzero
ord-ind-zero = refl
```

```
ord-ind-succ : ∀ { R Pzero Psucc Plim x } →
  ord-ind R Pzero Psucc Plim (osucc x)
  ≡ Psucc x (ord-ind R Pzero Psucc Plim x)
ord-ind-succ = refl
```

```
ord-ind-lim : ∀ { R Pzero Psucc Plim f } →
  ord-ind R Pzero Psucc Plim (olim f)
  ≡ Plim f (λ n → ord-ind R Pzero Psucc Plim (f n))
ord-ind-lim = refl
```

## 6 Conclusion

The powerful abstraction capabilities of dependently-typed programming languages make it possible to write extremely general definitions of families of data structures. Yet even when it is fairly straightforward to define a type with the desired inhabitants, it can be much trickier to ensure that the type has the right computation rules.

However, computation rules are fundamentally finite things, and as we've seen here careful attention to which parts of a definition are finitary can yield definitions that compute as expected.

## References

- [1] Michael Abbott, Thorsten Altenkirch, and Neil Ghani. 2004. Representing nested inductive types using W-types. In *International Colloquium on Automata, Languages, and Programming*. Springer, 59–71.
- [2] Michael Abbott, Thorsten Altenkirch, and Neil Ghani. 2005. Containers: Constructing strictly positive types. *Theoretical Computer Science* 342, 1 (2005), 3–27.
- [3] Thorsten Altenkirch, Neil Ghani, Peter Hancock, Conor McBride, and Peter Morris. 2015. Indexed containers. *Journal of Functional Programming* 25 (2015).
- [4] Marcin Benke, Peter Dybjer, and Patrik Jansson. 2003. Universes for generic programs and proofs in dependent type theory. *Nord. J. Comput.* 10, 4 (2003), 265–289.
- [5] James Chapman, Pierre-Evariste Dagand, Conor McBride, and Peter Morris. 2010. The gentle art of levitation. In *ICFP 2010*. 3–14.
- [6] Jean-Yves Girard. 1988. Normal functors, power series and  $\lambda$ -calculus. *Annals of pure and applied logic* 37, 2 (1988), 129–177.
- [7] Conor McBride. 2009. Grins from my Ripley Cupboard. <http://strictlypositive.org/Ripley.pdf>.
- [8] Conor McBride. 2015. Datatypes of datatypes. *Summer School on Generic and Effectful Programming, St Anne's College, Oxford* (2015).