

Multirec

Johan Jeuring

Utrecht University

9 June, 2015

Before the break ...

- Regular: a library in which you can use the recursive structure in a *single* datatype.

Now...

Multirec

- A DGP library (`multirec` on Hackage)
- Uses a type-indexed representation type
- Based on sum-of-products fixed-point view
- Supports recursion over a family of datatypes

Problem

Recall:

```
data Fix f = In { out :: f (Fix f) }
```

Question

How do we represent the following datatypes in a fixed-point view?

```
data Expr = Lit Int  
          | Add Expr Expr  
          | Mul Expr Expr  
          | Var Name  
          | Let Decl Expr
```

```
data Decl = Name := Expr  
          | Seq Decl Decl
```

```
type Name = String
```

Higher-Kinded `Fix`

- The datatypes `Expr` and `Decl` are **mutually recursive**.
- `Fix` takes a functor `f` with a single type of recursion.

```
data Fix f = In { out :: f (Fix f) }
```

- We need a fixed-point representation with **two recursive types**.

```
data Fix2 f g = In2 { out2 :: f (Fix2 f g) (Fix2 g f) }
```

Mutually Recursive Datatypes With `Fix2`

```
data ExprF2 rE rD = LitF Int
    | AddF rE rE
    | MulF rE rE
    | VarF Name
    | LetF rD rE

data DeclF2 rD rE = Name ::= rE
    | SeqF rD rD
```

```
type Expr' = Fix2 ExprF2 DeclF2
type Decl' = Fix2 DeclF2 ExprF2
```

And then we also need a `Functor2` type class...

Fix₁

Fix₁ can have one recursive type:

$$\text{Fix}_1 :: (* \rightarrow *) \rightarrow *$$

Fix₂

Fix₂ can have two recursive types:

$$\text{Fix}_1 :: (* \rightarrow *) \rightarrow *$$
$$\text{Fix}_2 :: (* \rightarrow * \rightarrow *) \rightarrow (* \rightarrow * \rightarrow *) \rightarrow *$$

Fix₃

Fix₃ can have three recursive types:

$$\text{Fix}_1 :: (* \rightarrow *) \rightarrow *$$

$$\text{Fix}_2 :: (* \rightarrow * \rightarrow *) \rightarrow (* \rightarrow * \rightarrow *) \rightarrow *$$

$$\text{Fix}_3 :: (* \rightarrow * \rightarrow * \rightarrow *) \rightarrow (* \rightarrow * \rightarrow * \rightarrow *) \rightarrow (* \rightarrow * \rightarrow * \rightarrow *) \rightarrow *$$

Fix_n

Question

How do we generalize to arbitrary arities?

Fix₁ :: ($* \rightarrow *$) $\rightarrow *$

Fix₂ :: ($* \rightarrow * \rightarrow *$) $\rightarrow (* \rightarrow * \rightarrow *)$ $\rightarrow *$

Fix₃ :: ($* \rightarrow * \rightarrow * \rightarrow *$) $\rightarrow (* \rightarrow * \rightarrow * \rightarrow *) \rightarrow (* \rightarrow * \rightarrow * \rightarrow *) \rightarrow *$

Fix_n :: ?

Uncurry

Let's uncurry the functional kinds.

$\text{Fix}_1 :: (* \rightarrow *)$		$\rightarrow *$
$:: (* \rightarrow *)$		$\rightarrow *$
$\text{Fix}_2 :: (* \rightarrow * \rightarrow *)$	$\rightarrow (* \rightarrow * \rightarrow *)$	$\rightarrow *$
$:: (* \times * \rightarrow *)$	$\times (* \times * \rightarrow *)$	$\rightarrow *$
$\text{Fix}_3 :: (* \rightarrow * \rightarrow * \rightarrow *)$	$\rightarrow (* \rightarrow * \rightarrow * \rightarrow *) \rightarrow (* \rightarrow * \rightarrow * \rightarrow *)$	$\rightarrow *$
$:: (* \times * \times * \rightarrow *)$	$\times (* \times * \times * \rightarrow *) \times (* \times * \times * \rightarrow *)$	$\rightarrow *$

Exponents

Then, we can reduce the product kinds to exponential kinds.

$$\text{Fix}_1 :: (*^1 \rightarrow *)^1 \rightarrow *$$

$$\text{Fix}_2 :: (*^2 \rightarrow *)^2 \rightarrow *$$

$$\text{Fix}_3 :: (*^3 \rightarrow *)^3 \rightarrow *$$

As a result, we can use `n` as the kind exponent.

$$\text{Fix}_n :: (*^n \rightarrow *)^n \rightarrow *$$

n Functors

Given:

$$\text{Fix}_n :: (*^n \rightarrow *)^n \rightarrow *$$

Fix_n is used for n functors:

$$\text{Fix}_n :: ((*^n \rightarrow *)^n \rightarrow *)^n$$

Aside: Laws of Exponents (1)

Function types and functional kinds are **exponential**.

$$t \rightarrow u \equiv u^t$$

Aside: Laws of Exponents (2)

Note:

- 1 means the unit type, isomorphic to $()$, a type with 1 constructor
- $2 \equiv 1 + 1$, isomorphic to `Bool`, a type with 2 constructors
- etc.

The function type $2 \rightarrow 2$ can be written as `Bool \rightarrow Bool`. For example:

```
not True = False
not False = True
```

The type is isomorphic to 2^2 . Note that for each argument constructor to `not` (`True` or `False`), you have 2 possible output constructors. Thus, you can define $2 * 2$ or 2^2 different functions.

Try other experiments to convince yourself.

Aside: Laws of Exponents (3)

The standard laws for exponents also hold on types.

$$1 \rightarrow u$$

$$\equiv$$

$$u^1$$

$$\equiv$$

$$u$$

$$t + u \rightarrow v$$

$$\equiv$$

$$v^{t+u}$$

$$\equiv$$

$$v^t * v^u$$

$$\equiv$$

$$(t \rightarrow v) * (u \rightarrow v)$$

$$t * u \rightarrow v$$

$$\equiv$$

$$v^{t*u}$$

$$\equiv$$

$$v^{u*t}$$

$$\equiv$$

$$(v^u)^t$$

$$\equiv$$

$$t \rightarrow v^u$$

$$\equiv$$

$$t \rightarrow (u \rightarrow v)$$

Rearranging

We can apply the laws of exponents to Fix_n :

$$\begin{aligned}\text{Fix}_n &:: ((*^n \rightarrow *)^n \rightarrow *)^n \\ &:: n \rightarrow (n \rightarrow ((n \rightarrow *) \rightarrow *)) \rightarrow *\end{aligned}$$

Rearranging a bit:

$$\text{Fix}_n :: ((n \rightarrow *) \rightarrow (n \rightarrow *)) \rightarrow (n \rightarrow *)$$

n as a Type or Kind

Given:

$$\text{Fix}_n :: ((n \rightarrow *) \rightarrow (n \rightarrow *)) \rightarrow (n \rightarrow *)$$

We can express it in Agda with a dependent type:

$$\text{Fix} : (n : \text{Nat}) \rightarrow ((\text{Fin } n \rightarrow *) \rightarrow (\text{Fin } n \rightarrow *)) \rightarrow \text{Fin } n \rightarrow *$$

But in Haskell, we are restricted to simple kinds:

$$\text{HFix} :: ((* \rightarrow *) \rightarrow (* \rightarrow *)) \rightarrow (* \rightarrow *)$$

We would like to say that `*` should only be instantiated by a suitable index type. Really, we use a bit of “dynamic kinding.”

```
data HFix (f :: (* -> *) -> (* -> *)) (ix :: *) = HIn { hout :: f (HFix f) ix }
```

Back to Expr and Decl

Recall:

```
data Expr = Lit Int
          | Add Expr Expr
          | Mul Expr Expr
          | Var Name
          | Let Decl Expr

data Decl = Name := Expr
          | Seq Decl Decl

type Name = String
```

Question

How do we represent this in a fixed-point view?

First: A GADT

Instead of three separate types, we could define it as a GADT.

data AST :: * \rightarrow * **where**

```

Lit    :: Int                 $\rightarrow$  AST Expr
Add    :: AST Expr  $\rightarrow$  AST Expr  $\rightarrow$  AST Expr
Mul    :: AST Expr  $\rightarrow$  AST Expr  $\rightarrow$  AST Expr
Var    :: AST Name           $\rightarrow$  AST Expr
Let    :: AST Decl  $\rightarrow$  AST Expr  $\rightarrow$  AST Expr
Bind   :: AST Name  $\rightarrow$  AST Expr  $\rightarrow$  AST Decl
Seq    :: AST Decl  $\rightarrow$  AST Decl  $\rightarrow$  AST Decl
Name   :: String             $\rightarrow$  AST Name

```

Expr, Decl, and Name are empty datatype declarations.

Next: A GADT Pattern Functor

Lift the recursion to a parameter.

data $\text{PF}_{\text{AST}} :: (* \rightarrow *) \rightarrow * \rightarrow *$ **where**

```

LitF    :: Int                → PFAST r Expr
AddF    :: r Expr → r Expr    → PFAST r Expr
MulF    :: r Expr → r Expr    → PFAST r Expr
VarF    :: r Name             → PFAST r Expr
LetF    :: r Decl → r Expr     → PFAST r Expr
BindF   :: r Name → r Expr     → PFAST r Decl
SeqF    :: r Decl → r Decl     → PFAST r Decl
NameF   :: String             → PFAST r Name

```

Tying the Knot with HFix

We can recover the recursion with HFix :

```
type AST' = HFix PFAST
```

Note how the index `ix` is used in `HFix` .

```
data HFix f ix = HIn { hout :: f (HFix f) ix }
```

```
data PFAST r ix where
```

```
  AddF :: r Expr → r Expr → PFAST r Expr
```

```
  BindF :: r Name → r Expr → PFAST r Decl
```

```
  ...
```

Pattern Functor in Regular

Question

Why can't we use Regular?

```

data U      r = U
data (f :+: g) r = L (f r) | R (g r)
data (f :×: g) r = f r :×: g r
newtype I    r = I r
newtype K a r = K a
  
```

Pattern Functor With Indexes

We need the index of the type in the family.

```

data U      (r :: * → *) ix = U
data (f :+: g) (r :: * → *) ix = L (f r ix) | R (g r ix)
data (f :×: g) (r :: * → *) ix = f r ix :×: g r ix
data l xi    (r :: * → *) ix = l (r xi)
data K a     (r :: * → *) ix = K a
  
```

Question

How do we represent the difference between an `Expr` constructor and a `Decl` constructor?

We **tag** the constructor with its index.

```

data f :>: ix :: (* → *) → * → * where
  Tag :: f r ix → (f :>: ix) r ix
  
```


Pattern Functor of AST

Here's the pattern functor type for AST :

```
type PFAST = K Int           :> Expr :+:  -- Lit
              (I Expr  :×: I Expr) :> Expr :+:  -- Add
              (I Expr  :×: I Expr) :> Expr :+:  -- Mul
              I Name           :> Expr :+:  -- Var
              (I Decl  :×: I Expr) :> Expr :+:  -- Let
              (I Name  :×: I Expr) :> Decl :+:  -- :=
              (I Decl  :×: I Decl) :> Decl :+:  -- Seq
              K String         :> Name      -- Name
```

Families of Datatypes

- As with Regular, we can use a type-indexed type for structure representation types such as `PFAST`.
- But we're missing one thing.

Question

What is the index of `PF` ?

type family `PF` ...

Describing a Family

```
type family PF ( $\phi :: * \rightarrow *$ ) :: ( $* \rightarrow *$ )  $\rightarrow * \rightarrow *$ 
```

- The index of the `PF` indicates the **family** or **system** of types.
- All functions in Multirec operate on a system of types. (“System” being a better word to distinguish it from the type families in GHC.)
- The system type ϕ :
 - ▶ Enumerates all types in the system
 - ▶ Reveals the index type (of kind `*`)
- The system datatype is a GADT:

```
data AST ::  $* \rightarrow *$  where
```

```
Expr  :: AST Expr
```

```
Decl  :: AST Decl
```

```
Name :: AST Name
```

Representation Class

Also, as with Regular, we use a class to define the embedding-projection pair for translating between the Haskell type and its pattern functor representation.

```
class Fam  $\phi$  where
  from ::  $\phi$  ix  $\rightarrow$  ix  $\rightarrow$  PF  $\phi$  l0 ix
  to   ::  $\phi$  ix  $\rightarrow$  PF  $\phi$  l0 ix  $\rightarrow$  ix
```

```
newtype l0 a = l0 a
```

Representation Instances

The representation instances are straightforward, if tedious.

type instance PF AST = PF_{AST}

$i_0 = I \circ I_0$

instance Fam AST **where**

from Expr (Lit i)	= L \$ Tag \$	K i
from Expr (Add e ₁ e ₂)	= R \$ L \$ Tag \$	$i_0\ e_1 \times i_0\ e_2$
...		
from Decl (v := e)	= R \$ R \$ R \$ R \$ R \$ L \$ Tag \$	$i_0\ v \times i_0\ e$
...		
from Name n	= R \$ R \$ R \$ R \$ R \$ R \$ R \$ Tag \$	K n
...		

Representation Via Template Haskell

Fortunately, you can generate these instances with Template Haskell:

```
$ (deriveConstructors [ ' ' Expr, ' ' Decl, ' ' Name])  
$ (deriveFamily ' ' AST [ ' ' Expr, ' ' Decl, ' ' Name] "PF_AST")  
type instance PF AST = PFAST
```

Generic Left

As an example of Multirec, we will define the generic function that produces the leftmost value of a datatype. It's called “empty” in other GP libraries.

```
left :: (El  $\phi$  ix, Fam  $\phi$ , Left  $\phi$  (PF  $\phi$ ))  $\Rightarrow$   $\phi$  ix  $\rightarrow$  ix
left w = to w $ fromJust $ gleft w (I0  $\circ$  left)
```

Defining a Generic Function: Left

```
class Left  $\phi$  f where
```

```
  gleft ::  $\phi$  ix  $\rightarrow$  r ix  $\rightarrow$  Maybe (f r ix)
```

```
instance Left  $\phi$  U where
```

```
  gleft _ _ = Just U
```

```
instance (Left  $\phi$  a, Left  $\phi$  b)  $\Rightarrow$  Left  $\phi$  (a :+: b) where
```

```
  gleft w r = fmap L (gleft w r)
```

```
instance (Left  $\phi$  a, Left  $\phi$  b)  $\Rightarrow$  Left  $\phi$  (a : $\times$ : b) where
```

```
  gleft w r = (: $\times$ :)  $\langle$ $ $\rangle$  gleft w r  $\langle$  $\star$  $\rangle$  gleft w r
```


Constant Types

For constant types, we use an additional type class:

```
instance (LeftK a)  $\Rightarrow$  Left  $\phi$  (K a) where  
  gleft _ _ = Just (K leftK)
```

```
class LeftK a where  
  leftK :: a  
  
instance LeftK Char where  
  leftK = minBound
```

Recursion

For recursion, we use the argument:

```
instance Left  $\phi$  (I xi) where
  gleft _ r = Just (I r)
```

Oops!

Couldn't match type 'ix' with 'xi'

'ix' is a rigid type variable bound by
the type signature for 'gleft'

'xi' is a rigid type variable bound by
the instance declaration

In the first argument of 'I', namely 'r'

We Need a Witness!

- Of course, `ix` and `xi` are not necessarily going to be the same.
- We support recursion on multiple types.
- We need to identify which type is this particular recursive reference.
- To produce a witness, we use a type class.

```
class El  $\phi$  ix where
  proof ::  $\phi$  ix
```

We need instances for all the possible proofs, namely all types in the family.

```
instance El AST Expr where proof = Expr
instance El AST Decl where proof = Decl
instance El AST Name where proof = Name
```

Recursion (Actually)

To handle recursion properly, we need to redefine the `Left` class.

- Since recursive indexes will be different, this indicates that we need a rank-2 type.
- Since we need to get a witness of the recursive index, this indicates that we need a `El` constraint.

```
class Left  $\phi$  f where
```

```
  gleft ::  $\phi$  ix  $\rightarrow$  ( $\forall$  xi. El  $\phi$  xi  $\Rightarrow$   $\phi$  xi  $\rightarrow$  r xi)  $\rightarrow$  Maybe (f r ix)
```

And so the recursive case can now be defined. (The rest is unchanged.)

```
instance (El  $\phi$  xi)  $\Rightarrow$  Left  $\phi$  (l xi) where
```

```
  gleft _ r = Just (l (r proof))
```

Tags

Recall:

```
data f :>: ix :: (* -> *) -> * -> * where
  Tag :: f r ix -> (f :>: ix) r ix
```

For tags, we need to prove that the tag index and the family index are equal.

```
instance (TEq  $\phi$ , El  $\phi$  ix, Left  $\phi$  f) => Left  $\phi$  (f :>: ix) where
  gleft w r = do
    Refl ← teq (proof ::  $\phi$  ix) w
    fmap Tag (gleft w r)
```

Type Equality

To determine type equality on the indexes, we use a GADT and type class.

```
data (:=:) :: * → * → * where  
  Refl :: a :=: a  
  
class TEq f where  
  teq :: f a → f b → Maybe (a :=: b)
```

HFunctor

Here is a generalization of `fmap`:

```
class HFunctor  $\phi$  f where
  hmapA :: (Applicative a)
     $\Rightarrow (\forall \mathbf{x}i. \phi \mathbf{x}i \rightarrow \mathbf{r} \mathbf{x}i \rightarrow a (\mathbf{s} \mathbf{x}i))$ 
     $\rightarrow \phi \mathbf{i}x \rightarrow f \mathbf{r} \mathbf{i}x \rightarrow a (f \mathbf{s} \mathbf{i}x)$ 
```

Note:

- The recursive type changes.
- The function parameter for recursion is polymorphic on the index.

We won't look at the definition of `HFunctor`.

hmap

With `hmapA`, we can define `hmap`:

```
hmap :: (HFunctor  $\phi$  f)
       $\Rightarrow (\forall xi. \phi\ xi \rightarrow r\ xi \rightarrow s\ xi)$ 
       $\rightarrow \phi\ ix \rightarrow f\ r\ ix \rightarrow f\ s\ ix$ 
hmap f p x = unI0 (hmapA ( $\lambda ix\ x \rightarrow I0\ (f\ ix\ x)$ ) p x)
```


fold

With `hmap`, we can define `fold`:

```
type Algebra  $\phi$  r =  $\forall ix. \phi\ ix \rightarrow (PF\ \phi)\ r\ ix \rightarrow r\ ix$ 
```

```
fold :: (Fam  $\phi$ , HFunctor  $\phi$  (PF  $\phi$ ))
```

```
     $\Rightarrow$  Algebra  $\phi$  r  $\rightarrow \phi\ ix \rightarrow ix \rightarrow r\ ix$ 
```

```
fold f p = f p  $\circ$  hmap ( $\lambda p\ (I0\ x) \rightarrow fold\ f\ p\ x$ ) p  $\circ$  from p
```

But this isn't the only sort of `fold` that can be used with `Multirec`. See the library for more.

Other Multirec Features

Some features that we did not discuss:

- Metadata on constructors
- Type composition

Using Multirec

Using Multirec requires certain prerequisites:

- Define a GADT for the system of datatypes
- Define the pattern functor for the system (i.e. the `PF` instance)
- Define instances of `Fam` and `El` for the datatypes in the system.
- Define other instances for functions as necessary

Applications

There are numerous applications for Multirec:

- General purpose functions: eq, enum, show, size, etc.
- Recursion schemes such as folds, paramorphisms, etc.
- Rewriting (including unification)
- Generation of values
- The zipper

Conclusions

- This library solved an open generic programming problem in Haskell: how can we write generic functions that use the fixed-point structure of datatypes on mutually recursive datatypes?
- The library has quite a number of useful applications.