

Generic Programming of All Kinds

Alejandro Serrano Mena, Victor Cacciari Miraldo

February 28, 2018



Context and Motivation

data $Exp :: * \rightarrow *$ **where**

$Val :: Int \rightarrow Exp$ Int

$Add :: Exp$ $Int \rightarrow Exp$ $Int \rightarrow Exp$ Int

$Eq :: Exp$ $Int \rightarrow Exp$ $Int \rightarrow Exp$ $Bool$

...

deriving instance $(Serialize\ a) \Rightarrow Serialize\ (Exp\ a)$



Context and Motivation

data $Exp :: * \rightarrow *$ **where**

$Val :: Int \rightarrow Exp$ Int

$Add :: Exp$ $Int \rightarrow Exp$ $Int \rightarrow Exp$ Int

$Eq :: Exp$ $Int \rightarrow Exp$ $Int \rightarrow Exp$ $Bool$

...

deriving instance $(Serialize\ a) \Rightarrow Serialize\ (Exp\ a)$

We would like this feature!



Context and Motivation

data $Exp :: * \rightarrow *$ **where**

$Val :: Int \rightarrow Exp$ Int

$Add :: Exp$ $Int \rightarrow Exp$ $Int \rightarrow Exp$ Int

$Eq :: Exp$ $Int \rightarrow Exp$ $Int \rightarrow Exp$ $Bool$

...

deriving instance $(Serialize\ a) \Rightarrow Serialize\ (Exp\ a)$

We would like this feature!

Implementing it in a general fashion requires some **generic programming over GADTs and arbitrarily kinded types**.



Context and Motivation

- ▶ GHC's modern extensions allow for more expressive generic programming.



Context and Motivation

- ▶ GHC's modern extensions allow for more expressive generic programming.
- ▶ Inability to currently handle arbitrarily kinded datatypes.



Context and Motivation

- ▶ GHC's modern extensions allow for more expressive generic programming.
- ▶ Inability to currently handle arbitrarily kinded datatypes.
- ▶ GADTs are becoming more common: **deriving** clauses would be handy.



Representing Datatypes (generics-sop)

Haskell datatypes come in **sums-of-products** shape:

```
data Tree a = Leaf | Bin a (Tree a) (Tree a)
```



Representing Datatypes (generics-sop)

Haskell datatypes come in **sums-of-products** shape:

```
data Tree a = Leaf | Bin a (Tree a) (Tree a)
```

Our **codes** will follow that structure:

```
type family Code (x :: *) :: '[[*]]
```

```
type instance Code (Tree a) = '[[], '[a, Tree a, Tree a]]
```



Representing Datatypes (generics-sop)

Haskell datatypes come in **sums-of-products** shape:

```
data Tree a = Leaf | Bin a (Tree a) (Tree a)
```

Our **codes** will follow that structure:

```
type family Code (x :: *) :: '['[*]]
```

```
type instance Code (Tree a) = '['[], '[a, Tree a, Tree a]]
```

Given a map from '['[*]] into *, call it *Rep*, package:

```
class Generic a where
```

```
  from :: a → Rep (Code a)
```

```
  to   :: Rep (Code a) → a
```



N-ary Sums and Products

$$NS \ p [x_1, \dots, x_n] \approx \textit{Either} (p \ x_1) (\textit{Either} \ \dots \ (p \ x_n))$$

$$NP \ p [x_1, \dots, x_n] \approx (p \ x_1, \dots, p \ x_n)$$



N-ary Sums and Products

$NS\ p\ [x_1, \dots, x_n] \approx \text{Either}\ (p\ x_1)\ (\text{Either}\ \dots\ (p\ x_n))$

$NP\ p\ [x_1, \dots, x_n] \approx (p\ x_1, \dots, p\ x_n)$

data $NS :: (k \rightarrow *) \rightarrow [k] \rightarrow *$ **where**

$\text{Here} :: f\ x \rightarrow NS\ f\ (x' : xs)$

$\text{There} :: NS\ f\ xs \rightarrow NS\ f\ (x' : xs)$

data $NP :: (k \rightarrow *) \rightarrow [k] \rightarrow *$ **where**

$\text{Nil} :: NP\ f\ []$

$\text{Cons} :: f\ x \rightarrow NP\ f\ xs \rightarrow NP\ f\ (x' : xs)$



Interpreting Codes (generics-sop)

data $I\ x = I\ x$

type $Rep\ (c :: '[[*]]) = NS\ (NP\ I)\ c$



Interpreting Codes (generics-sop)

data $I\ x = I\ x$

type $Rep\ (c :: '[[*]]) = NS\ (NP\ I)\ c$

Recall the *Tree* example:

type instance $Code\ (Tree\ a) = '[[], '[a, Tree\ a, Tree\ a]]$

leaf $:: Rep\ (Code\ (Tree\ a))$

leaf $= Here\ Nil$

bin $:: a \rightarrow Tree\ a \rightarrow Tree\ a \rightarrow Rep\ (Code\ (Tree\ a))$

bin\ e\ l\ r $= There\ (Here\ (Cons\ e\ (Cons\ l\ (Cons\ r\ Nil))))$



Writing Generic Functions

Package it in a class

```
class Size a where  
  size :: a → Int
```



Writing Generic Functions

Package it in a class

class *Size* *a* **where**

size :: *a* → *Int*

Then write the generic infrastructure:

gsize :: (*Generic* *x*, *All2* *Size* (*Code* *x*))

⇒ *x* → *Int*

gsize = *goS* ◦ *from*

where

goS (*Here* *x*) = *goP* *x*

goS (*There* *x*) = *goS* *x*

goP *Nil* = 0

goP (*Cons* *x* *xs*) = *size* *x* + *goP* *xs*



Generics of All Kinds

- ▶ So far, only handle types of kind `*` with no parameters.



Generics of All Kinds

- ▶ So far, only handle types of kind `*` with no parameters.
- ▶ Consequence of little structure on *Codes*.



Generics of All Kinds

- ▶ So far, only handle types of kind $*$ with no parameters.
- ▶ Consequence of little structure on *Codes*.
- ▶ **Solution:** Augment the language of codes!

type *DataType* ($\zeta :: \textit{Kind}$) = '['[*Atom* ζ (*)]]



Generics of All Kinds

- ▶ So far, only handle types of kind $*$ with no parameters.
- ▶ Consequence of little structure on *Codes*.
- ▶ **Solution:** Augment the language of codes!

type *DataType* ($\zeta :: \textit{Kind}$) = '['[*Atom* ζ (*)]]

- ▶ *Atom* is the applicative fragment of the λ -calculus with de Bruijn indices.



Generics of All Kinds

data *Atom* ($\zeta :: \textit{Kind}$) ($k :: \textit{Kind}$) :: (*) **where**

Var :: *TyVar* ζ k \rightarrow *Atom* ζ k

Kon :: k \rightarrow *Atom* ζ k

(:@) :: *Atom* ζ ($l \rightarrow k$) \rightarrow *Atom* ζ $l \rightarrow$ *Atom* ζ k

data *TyVar* ($\zeta :: \textit{Kind}$) ($k :: \textit{Kind}$) :: (*) **where**

VZ :: *TyVar* ($x \rightarrow xs$) x

VS :: *TyVar* xs $k \rightarrow$ *TyVar* ($x \rightarrow xs$) k



Generics of All Kinds

```
data Atom ( $\zeta :: \text{Kind}$ ) ( $k :: \text{Kind}$ ) :: (*) where  
  Var :: TyVar  $\zeta$  k → Atom  $\zeta$  k  
  Kon :: k → Atom  $\zeta$  k  
  (:@:) :: Atom  $\zeta$  ( $l \rightarrow k$ ) → Atom  $\zeta$  l → Atom  $\zeta$  k
```

Going back to our *Tree* example:

```
data Tree a = Leaf | Bin a (Tree a) (Tree a)
```

```
type V0 = Var VZ
```

```
type TreeCode
```

```
= '[[], '[V0, Kon Tree :@: V0, Kon Tree :@: V0]]  
:: '[@[Atom (* → *) *]]
```



Interpreting Atoms

Interpreting atoms needs environment.

data Γ ($\zeta :: Kind$) **where**

ϵ $::$ Γ (*)

$(:\&:)$ $:: k \rightarrow \Gamma ks \rightarrow \Gamma (k \rightarrow ks)$



Interpreting Atoms

Interpreting atoms needs environment.

data $\Gamma (\zeta :: Kind)$ **where**

$$\begin{aligned} \epsilon &:: \Gamma (*) \\ (:&:) &:: k \rightarrow \Gamma ks \rightarrow \Gamma (k \rightarrow ks) \end{aligned}$$

For example,

$$Int \ :&: \ Maybe \ :&: \ Char \ :&: \ \epsilon$$

Is a well-formed environment of kind

$$\Gamma (* \rightarrow (* \rightarrow *) \rightarrow * \rightarrow *)$$



Interpreting Atoms

Interpreting atoms needs environment.

data $\Gamma (\zeta :: Kind)$ **where**

$\epsilon \quad :: \quad \Gamma (*)$
 $(: \& :) :: k \rightarrow \Gamma ks \rightarrow \Gamma (k \rightarrow ks)$

type family $Ty \zeta (tys :: \Gamma \zeta) (t :: Atom \zeta k) :: k$ **where**

$Ty (k \rightarrow ks) (t : \& : ts) (Var VZ) = t$
 $Ty (k \rightarrow ks) (t : \& : ts) (Var (VS v)) = Ty ks ts (Var v)$
 $Ty \zeta ts (Kon t) = t$
 $Ty \zeta ts (f : @ : x) = (Ty \zeta ts f) (Ty \zeta ts x)$



Interpreting Codes

We are now ready to map a code, of kind *DataType* ζ , into $*$.

First, package *Ty* into a GADT:

```
data NA ( $\zeta :: Kind$ ) ::  $\Gamma \zeta \rightarrow Atom \zeta (*) \rightarrow *$  where  
  T ::  $\forall \zeta t a . Ty \zeta a t \rightarrow NA \zeta a t$ 
```



Interpreting Codes

We are now ready to map a code, of kind $\text{DataType } \zeta$, into $*$.

First, package Ty into a GADT:

```
data  $NA$  ( $\zeta :: Kind$ ) ::  $\Gamma \zeta \rightarrow Atom \zeta (*) \rightarrow *$  where  
   $T :: \forall \zeta t a . Ty \zeta a t \rightarrow NA \zeta a t$ 
```

Then, assemble NS , NP and NA :

```
type  $Rep$  ( $\zeta :: Kind$ ) ( $c :: DataType \zeta$ ) ( $a :: \Gamma \zeta$ )  
  =  $NS (NP (NA \zeta a)) c$ 
```



Approaching a Unified API

Usually, GP libraries provide a class:

```
class Generic f where  
  type Code f :: CodeKind  
  from :: f → Rep (Code f)  
  to    :: Rep (Code f)
```



Approaching a Unified API

Usually, GP libraries provide a class:

```
class Generic f where  
  type Code f :: CodeKind  
  from :: f → Rep (Code f)  
  to    :: Rep (Code f)
```

In our case, though, the number of arguments to f depend on it's kind!

```
from :: f      → Rep (*) (Code f) ε  
from :: f x    → Rep (*) (Code f) (x :&: ε)  
from :: f x y → Rep (*) (Code f) (x :&: y :&: ε)
```



Approaching a Unified API

Write a GADT:

data *ApplyT* ζ ($f :: k$) ($\alpha :: \Gamma \zeta$) :: * **where**

AO :: $f \rightarrow \text{ApplyT } (*) \quad f \in$

AS :: *ApplyT* ks (f t) $ts \rightarrow \text{ApplyT } (k \rightarrow ks) f (t \&: ts)$



Approaching a Unified API

Write a GADT:

```
data ApplyT ζ (f :: k) (α :: Γ ζ) :: * where  
  AO :: f → ApplyT (*) f ∈  
  AS :: ApplyT ks (f t) ts → ApplyT (k → ks) f (t &: ts)
```

Which allows us to unify the interface:

```
from :: ApplyT ζ f a → Rep ζ (Code f) a
```



Wait?! type-in-type?

- ▶ We require `-XTypeInType` to type check our code because we need to promote GADTs and work with kinds as types.



Wait?! type-in-type?

- ▶ We require `-XTypeInType` to type check our code because we need to promote GADTs and work with kinds as types.
- ▶ We do not require the `*:*` axiom



Wait?! type-in-type?

- ▶ We require `-XTypeInType` to type check our code because we need to promote GADTs and work with kinds as types.
- ▶ We do not require the `*:*` axiom
- ▶ We provide an Agda model of our approach to prove so. Basic types live in Set_0 , our codes inhabit Set_1 and the interpretations inhabit Set_2 .



Representing Constraints

With small modifications, we can handle constraints.



Representing Constraints

With small modifications, we can handle constraints.

Add one layer on top of *Atom*:

data *Field* ($\zeta :: \text{Kind}$) **where**

Explicit :: *Atom* ζ (*) \rightarrow *Field* ζ

Implicit :: *Atom* ζ *Constraint* \rightarrow *Field* ζ

type *DataType* $\zeta = '[[Field \zeta]]$



Representing Constraints

With small modifications, we can handle constraints.

Add one layer on top of *Atom*:

```
data Field ( $\zeta :: \textit{Kind}$ ) where  
  Explicit :: Atom  $\zeta$  (*)       $\rightarrow$  Field  $\zeta$   
  Implicit :: Atom  $\zeta$  Constraint  $\rightarrow$  Field  $\zeta$   
type DataType  $\zeta = '[[Field  $\zeta$ ]]$ 
```

Adapt the interpretation of *Atom* to work on top of *Field*:

```
data NA ( $\zeta :: \textit{Kind}$ ) ::  $\Gamma$   $\zeta \rightarrow$  Field  $\zeta \rightarrow *$  where  
  E ::  $\forall \zeta t a . \textit{Ty}$   $\zeta$   $a$   $t \rightarrow$  NA  $\zeta$   $a$  (Explicit  $t$ )  
  I ::  $\forall \zeta t a . \textit{Ty}$   $\zeta$   $a$   $t \Rightarrow$  NA  $\zeta$   $a$  (Implicit  $t$ )
```



Example: Representing a GADT

data *Expr* :: * → * **where**

Lit :: *a* → *Expr a*

IsZero :: *Expr Int* → *Expr Bool*

If :: *Expr Bool* → *Expr a* → *Expr a* → *Expr a*



Example: Representing a GADT

data *Expr* :: * → * **where**

Lit :: *a* → *Expr a*

IsZero :: (*a* ~ *Bool*) ⇒ *Expr Int* → *Expr a*

If :: *Expr Bool* → *Expr a* → *Expr a* → *Expr a*



Example: Representing a GADT

data *Expr* :: * → * **where**

Lit :: *a* → *Expr a*

IsZero :: (*a* ~ *Bool*) ⇒ *Expr Int* → *Expr a*

If :: *Expr Bool* → *Expr a* → *Expr a* → *Expr a*

type *CodeExpr*

= '['[*Explicit V0*]

, '[*Implicit (Kon (~) :@: V0 :@: Kon Bool)*

, *Explicit (Kon Expr :@: Kon Int)*]

, ...

]



Generic GADTs: Extensions Limitations

- ▶ On our paper we discuss how to handle existential types.



Generic GADTs: Extensions Limitations

- ▶ On our paper we discuss how to handle existential types. The resulting interface is not user-friendly and make the writing of generic combinators cumbersome.



Generic GADTs: Extensions Limitations

- ▶ On our paper we discuss how to handle existential types. The resulting interface is not user-friendly and make the writing of generic combinators cumbersome.
- ▶ Existential kinds pose a problem on the other hand. We can't represent telescopes like:

data *Problem* :: $k \rightarrow *$ **where**

Constructor :: $\forall k (a :: k) . X a \rightarrow Problem a$



Arity-generic fmap

We are able to generalize *Functor* and *BiFunctor* to *NFunctor*.



Arity-generic fmap

We are able to generalize *Functor* and *BiFunctor* to *NFunctor*.

That is, let f be of kind $* \rightarrow * \rightarrow \dots \rightarrow *$, we can then write:

$$\begin{aligned} fmapN &:: (a_1 \rightarrow b_1) \\ &\rightarrow \dots \\ &\rightarrow (a_n \rightarrow b_n) \\ &\rightarrow f a_1 \dots a_n \\ &\rightarrow f b_1 \dots b_n \end{aligned}$$



Discussion, Future and Related Work

- ▶ R. Scott “Generalized Abstract GHC.Generics” paper at HIW, last Sunday.



Discussion, Future and Related Work

- ▶ R. Scott “Generalized Abstract GHC.Generics” paper at HIW, last Sunday.
- ▶ We are able to represent a reasonable amount of GADTs generically.



Discussion, Future and Related Work

- ▶ R. Scott “Generalized Abstract GHC.Generics” paper at HIW, last Sunday.
- ▶ We are able to represent a reasonable amount of GADTs generically.
- ▶ Our approach also extend to mutually recursive types as long as we do not bring in explicit fixpoints.



Discussion, Future and Related Work

- ▶ R. Scott “Generalized Abstract GHC.Generics” paper at HIW, last Sunday.
- ▶ We are able to represent a reasonable amount of GADTs generically.
- ▶ Our approach also extend to mutually recursive types as long as we do not bring in explicit fixpoints.
- ▶ Fork generics-mrsop and package these ideas into a usable library.



Generic Programming of All Kinds

Alejandro Serrano Mena, Victor Cacciari Miraldo

February 28, 2018

