



Faculty of Science



# Parametric Compositional Data Types

Patrick Bahr Tom Hvitved

University of Copenhagen, Department of Computer Science  
{ paba , hvitved }@diku.dk

Mathematically Structured Functional Programming 2012,  
Tallinn, Estonia, March 25th, 2012



# Outline

- 1 Motivation
- 2 Compositional Data Types
- 3 Higher-Order Abstract Syntax



# The Issue

## Implementation/Prototyping of DSLs

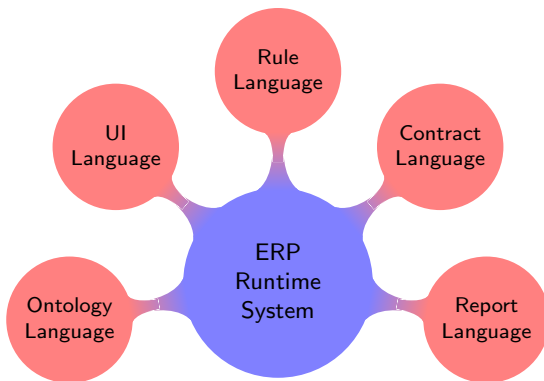


ERP  
Runtime  
System



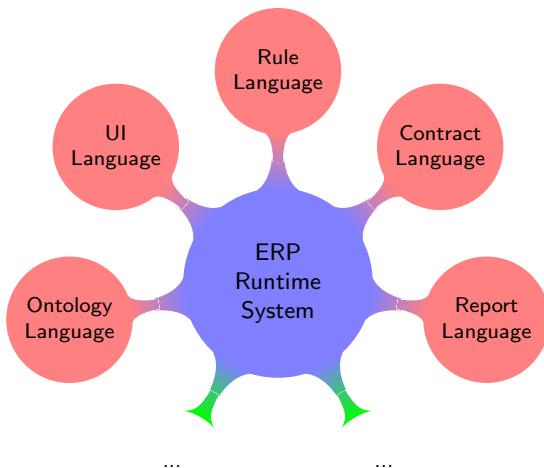
# The Issue

## Implementation/Prototyping of DSLs



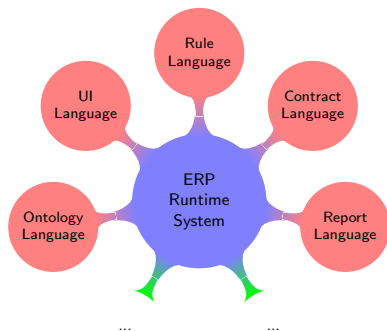
# The Issue

## Implementation/Prototyping of DSLs



# The Issue

## Implementation/Prototyping of DSLs



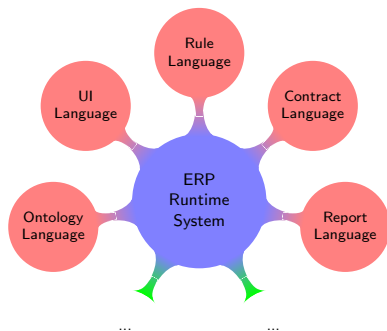
### The abstract picture

- We have a number of **domain-specific languages**.
- Each pair of DSLs shares some **common sublanguage**.
- All of them share a common language of values.
- We have the same situation on the type level!



# The Issue

## Implementation/Prototyping of DSLs



### The abstract picture

- We have a number of **domain-specific languages**.
- Each pair of DSLs shares some **common sublanguage**.
- All of them share a common language of values.
- We have the same situation on the type level!

How do we implement this system without duplicating code?!



## More General Application

Even with only one language to implement this issue appears!





## More General Application

Even with only one language to implement this issue appears!

Different stages of a compiler work on different languages.

- Desugaring:  $FullExp \rightarrow CoreExp$
- Evaluation:  $Exp \rightarrow Value$
- $\vdots$



## More General Application

Even with only one language to implement this issue appears!

Different stages of a compiler work on different languages.

- Desugaring:  $FullExp \rightarrow CoreExp$
- Evaluation:  $Exp \rightarrow Value$
- $\vdots$

Manipulating/extending syntax trees

- annotating syntax trees
- adding/removing type annotations

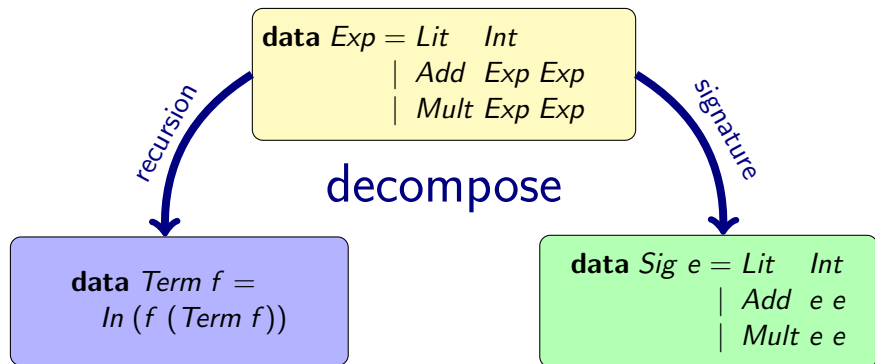


# Compositional Data Types

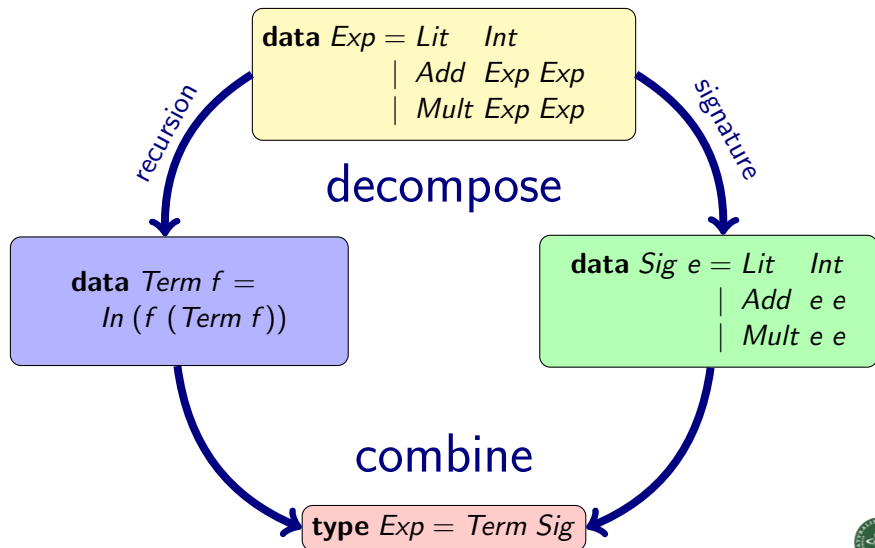
```
data Exp = Lit Int  
         | Add Exp Exp  
         | Mult Exp Exp
```



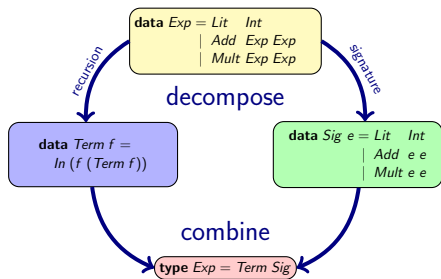
# Compositional Data Types



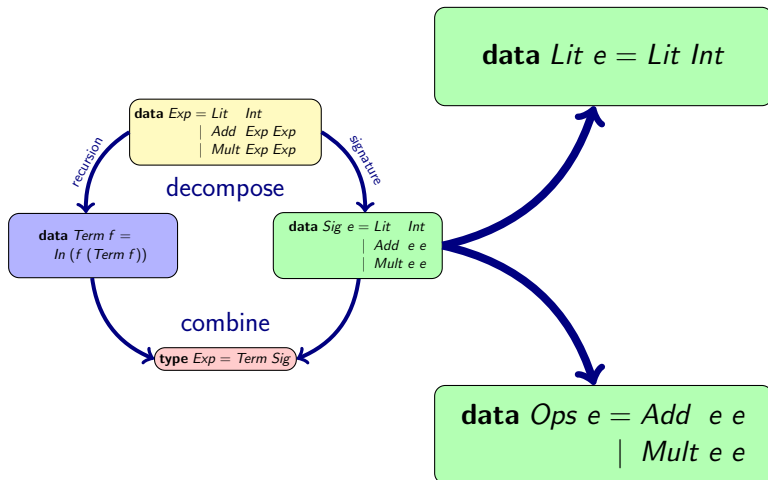
# Compositional Data Types



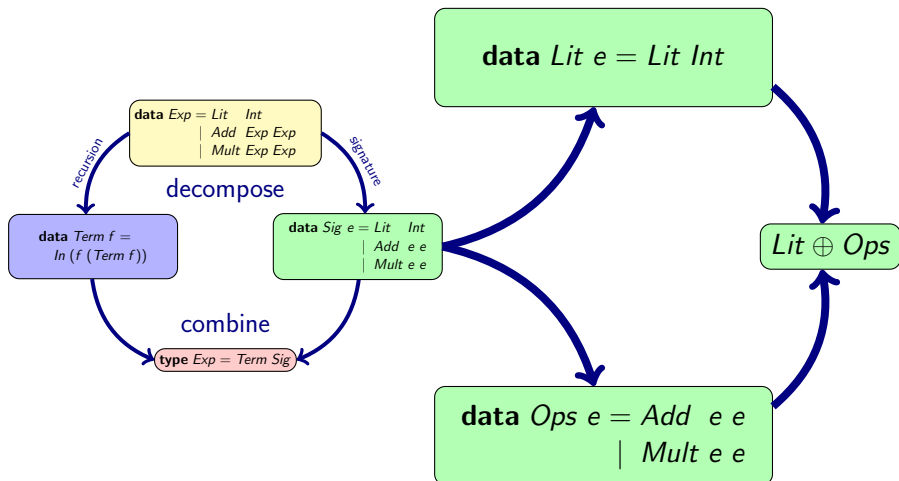
# Compositional Data Types



# Compositional Data Types

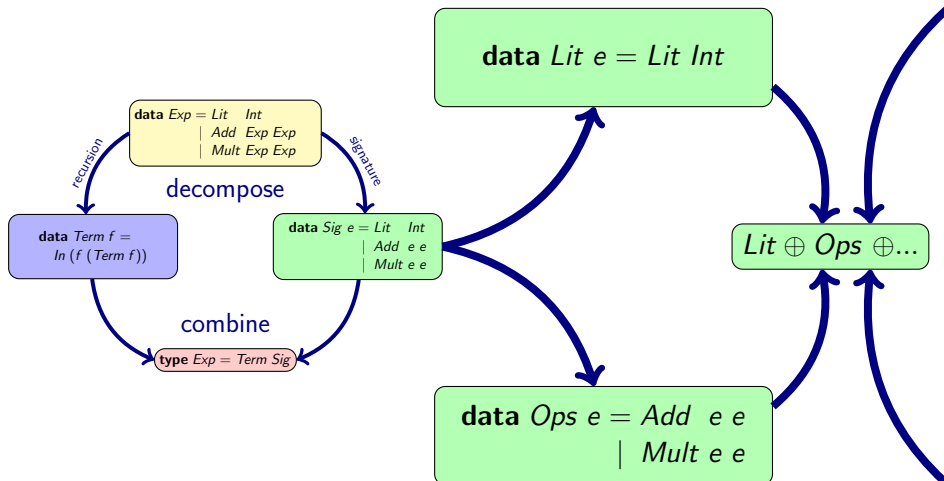


# Compositional Data Types

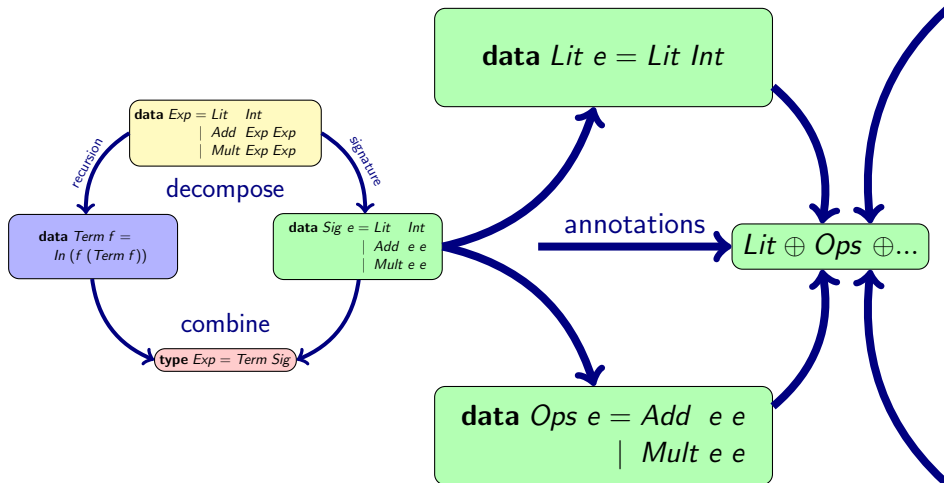




# Compositional Data Types



# Compositional Data Types



# Variable Binding

A straightforward solution

```
type Name = String
data Lam e = Lam Name e
data Var e = Var Name
data App e = App e e
```



# Variable Binding

A straightforward solution

**type** *Name* = *String*

**data** *Lam* *e* = *Lam* *Name* *e*

**data** *Var* *e* = *Var* *Name*

**data** *App* *e* = *App* *e* *e*

**type** *Sig* = *Lam*  $\oplus$  *Var*  $\oplus$  *App*

**type** *Lambda* = *Term* *Lam*



# Variable Binding

## A straightforward solution

**type** *Name* = *String*

**data** *Lam* *e* = *Lam* *Name* *e*

**data** *Var* *e* = *Var* *Name*

**data** *App* *e* = *App* *e* *e*

**type** *Sig* = *Lam*  $\oplus$  *Var*  $\oplus$  *App*

**type** *Lambda* = *Term* *Lam*

## Issues

- Definitions modulo  $\alpha$ -equivalence
- Capture-avoiding substitutions
- Implementing embedded languages



# Variable Binding

## A straightforward solution

**type** *Name* = *String*

**data** *Lam* *e* = *Lam* *Name* *e*

**data** *Var* *e* = *Var* *Name*

**data** *App* *e* = *App* *e* *e*

**type** *Sig* = *Lam*  $\oplus$  *Var*  $\oplus$  *App*

**type** *Lambda* = *Term* *Lam*

## Issues

- Definitions modulo  $\alpha$ -equivalence
- Capture-avoiding substitutions
- Implementing embedded languages

## Goal

Use higher-order abstract syntax to leverage the variable binding mechanism of the host language.

# Higher-Order Abstract Syntax

## Explicit Variables

**type** *Name* = *String*

**data** *Lam* *e* = *Lam Name e*

**data** *Var* *e* = *Var Name*

**data** *App* *e* = *App e e*



# Higher-Order Abstract Syntax

## Explicit Variables

**type** *Name* = *String*

**data** *Lam* *e* = *Lam* *Name* *e*

**data** *Var* *e* = *Var* *Name*

**data** *App* *e* = *App* *e* *e*

## Higher-Order Abstract Syntax

**data** *Lam* *e* = *Lam* ( $e \rightarrow e$ )

**data** *App* *e* = *App* *e* *e*





# Higher-Order Abstract Syntax

## Explicit Variables

**type** *Name* = *String*

**data** *Lam* *e* = *Lam* *Name e*

**data** *Var* *e* = *Var* *Name*

**data** *App* *e* = *App* *e e*

## Higher-Order Abstract Syntax

**data** *Lam* *e* = *Lam* (*e* → *e*)

**data** *App* *e* = *App* *e e*



# Higher-Order Abstract Syntax

Explicit Variables

$Lam\ "x"\ (...Var\ "x" \dots)$

**type**  $Name = String$

**data**  $Lam\ e = Lam\ Name\ e$

**data**  $Var\ e = Var\ Name$

**data**  $App\ e = App\ e\ e$

Higher-Order Abs

$Lam\ (\lambda x \rightarrow \dots x \dots)$

**data**  $Lam\ e = Lam\ (e \rightarrow e)$

**data**  $App\ e = App\ e\ e$



# Higher-Order Abstract Syntax

Explicit Variables

$Lam\ "x"\ (...Var\ "x" \dots)$

**type**  $Name = String$

**data**  $Lam\ e = Lam\ Name\ e$

**data**  $Var\ e = Var\ Name$

**data**  $App\ e = App\ e\ e$

Higher-Order Abs

$Lam\ (\lambda x \rightarrow \dots x \dots)$

**data**  $Lam\ e = Lam\ (e \rightarrow e)$

**data**  $App\ e = App\ e\ e$

## Issues

- inefficient and cumbersome **recursion schemes**  
(catamorphism needs an algebra and the **inverse coalgebra**)
- **Full function space** allows for **exotic terms**

# Higher-Order Abstract Syntax

Explicit Variables

$Lam\ "x"\ (...Var\ "x"\ ...)$

**type**  $Name = String$

**data**  $Lam\ e = Lam\ Name\ e$

**data**  $Var\ e = Var\ Name$

**data**  $App\ e = App\ e\ e$

Higher-Order Abs

$Lam\ (\lambda x \rightarrow \dots x \dots)$

**data**  $Lam\ e = Lam\ (e \rightarrow e)$

**data**  $App\ e = App\ e\ e$

## Issues

- inefficient and cumbersome **recursion schemes**  
(catamorphism needs an algebra and the **inverse coalgebra**)  
 $\rightsquigarrow$  Fegaras & Sheard (1996): parametric functions space
- Full function space** allows for **exotic terms**

# Higher-Order Abstract Syntax

Explicit Variables

$Lam\ "x"\ (...Var\ "x"\ ...)$

**type**  $Name = String$

**data**  $Lam\ e = Lam\ Name\ e$

**data**  $Var\ e = Var\ Name$

**data**  $App\ e = App\ e\ e$

Higher-Order Abs

$Lam\ (\lambda x \rightarrow ...x...)$

**data**  $Lam\ e = Lam\ (e \rightarrow e)$

**data**  $App\ e = App\ e\ e$

## Issues

- inefficient and cumbersome **recursion schemes**  
(catamorphism needs an algebra and the **inverse coalgebra**)  
 $\rightsquigarrow$  Fegaras & Sheard (1996): parametric functions space
- **Full function space** allows for **exotic terms**  
 $\rightsquigarrow$  Washburn & Weirich (2008): polymorphism & abstract type of terms

# Parametric Higher-Order Abstract Syntax

[Chlipala 2008]

## Idea

- Signature for lambda bindings:

**data**  $Lam\ a\ e = Lam\ (a \rightarrow e)$



# Parametric Higher-Order Abstract Syntax

[Chlipala 2008]

## Idea

- Signature for lambda bindings:

**data**  $Lam\ a\ e = Lam\ (a \rightarrow e)$

- Recursive construction of terms:

**data**  $Trm\ f\ a = In\ (f\ a\ (Trm\ f\ a))$



# Parametric Higher-Order Abstract Syntax

[Chlipala 2008]

## Idea

- Signature for lambda bindings:

**data**  $Lam\ a\ e = Lam\ (a \rightarrow e)$

- Recursive construction of terms:

**data**  $Trm\ f\ a = In\ (f\ a\ (Trm\ f\ a))$

**type**  $Term\ f = \forall a. Trm\ f\ a$





# Parametric Higher-Order Abstract Syntax

[Chlipala 2008]

## Idea

- Signature for lambda bindings:

**data**  $Lam\ a\ e = Lam\ (a \rightarrow e)$

- Recursive construction of terms:

**data**  $Trm\ f\ a = In\ (f\ a\ (Trm\ f\ a)) \mid Var\ a$

**type**  $Term\ f = \forall a. Trm\ f\ a$



# Parametric Higher-Order Abstract Syntax

[Chlipala 2008]

## Idea

- Signature for lambda bindings:

**data**  $Lam\ a\ e = Lam\ (a \rightarrow e)$

- Recursive construction of terms:

**data**  $Trm\ f\ a = In\ (f\ a\ (Trm\ f\ a)) \mid Var\ a$

**newtype**  $Term\ f = Term\ (\forall a. Trm\ f\ a)$



# Parametric Higher-Order Abstract Syntax

[Chlipala 2008]

## Idea

- Signature for lambda bindings:

**data**  $Lam\ a\ e = Lam\ (a \rightarrow e)$

- Recursive construction of terms:

**data**  $Trm\ f\ a = In\ (f\ a\ (Trm\ f\ a)) \mid Var\ a$

**newtype**  $Term\ f = Term\ (\forall a . Trm\ f\ a)$

## Example

**data**  $Sig\ a\ e = Lam\ (a \rightarrow e) \mid App\ e\ e$



# Parametric Higher-Order Abstract Syntax

[Chlipala 2008]

## Idea

- Signature for lambda bindings:

**data**  $Lam\ a\ e = Lam\ (a \rightarrow e)$

- Recursive construction of terms:

**data**  $Trm\ f\ a = In\ (f\ a\ (Trm\ f\ a)) \mid Var\ a$

**newtype**  $Term\ f = Term\ (\forall a. Trm\ f\ a)$

## Example

**data**  $Sig\ a\ e = Lam\ (a \rightarrow e) \mid App\ e\ e$

$e :: Term\ Sig$

$e = Term\ \$\ Lam\ (\lambda x \rightarrow Var\ x\ 'App'\ Var\ x)$



# Parametric Higher-Order Abstract Syntax

[Chlipala 2008]

## Idea

- Signature for lambda bindings:

**data**  $Lam\ a\ e = Lam\ (a \rightarrow e)$

- Recursive construction of terms:

**data**  $Trm\ f\ a = In\ (f\ a\ (Trm\ f\ a)) \mid Var\ a$

**newtype**  $Term\ f = Term\ (\forall a. Trm\ f\ a)$

## Example

**data**  $Sig\ a\ e = Lam\ (a \rightarrow e) \mid App\ e\ e$

$e :: Term\ Sig$

$e = Term\ \$\ Lam\ (\lambda x \rightarrow Var\ x\ 'App'\ Var\ x)$

$e = \lambda x. x\ x$



# Adding Compositionality

## Coproducts

**data**  $(f \oplus g) a e = \text{Inl } (f a e) \mid \text{Inr } (g a e)$



# Adding Compositionality

## Coproducts

**data**  $(f \oplus g) a e = \text{Inl } (f a e) \mid \text{Inr } (g a e)$

## Example

**data**  $\text{Lam } a e = \text{Lam } (a \rightarrow e)$

**data**  $\text{App } a e = \text{App } e e$

**type**  $\text{Sig} = \text{Lam} \oplus \text{App}$



# Recursion Schemes

## Generalising functors

```
class Functor f where  
  fmap :: (a → b) → f a → f b
```





# Recursion Schemes

## Generalising functors to difunctors

**class** *Difunctor* *f* **where**

*dimap* ::  $(a \rightarrow b) \rightarrow (c \rightarrow d) \rightarrow f\ b\ c \rightarrow f\ a\ d$



# Recursion Schemes

## Generalising functors to difunctors

**class** *Difunctor* *f* **where**

*dimap* ::  $(a \rightarrow b) \rightarrow (c \rightarrow d) \rightarrow f\ b\ c \rightarrow f\ a\ d$

**instance** *Difunctor*  $(\rightarrow)$  **where**

*dimap* *f* *g* *h* = *g* . *h* . *f*



# Recursion Schemes

## Generalising functors to difunctors

**class** *Difunctor* *f* **where**

$dimap :: (a \rightarrow b) \rightarrow (c \rightarrow d) \rightarrow f\ b\ c \rightarrow f\ a\ d$

**instance** *Difunctor*  $(\rightarrow)$  **where**

$dimap\ f\ g\ h = g . h . f$

## Algebras

**type** *Alg* *f* *c* =  $f\ c\ c \rightarrow c$



# Recursion Schemes

## Generalising functors to difunctors

**class** *Difunctor* *f* **where**

$dimap :: (a \rightarrow b) \rightarrow (c \rightarrow d) \rightarrow f\ b\ c \rightarrow f\ a\ d$

**instance** *Difunctor*  $(\rightarrow)$  **where**

$dimap\ f\ g\ h = g . h . f$

## Algebras

**type** *Alg* *f* *c* =  $f\ c\ c \rightarrow c$

## Catamorphisms

$cata :: Difunctor\ f \Rightarrow Alg\ f\ c \rightarrow Term\ f \rightarrow c$

$cata\ \phi\ (Term\ t) = cat\ t$

**where**  $cat :: Trm\ f\ c \rightarrow c$

$cat\ (In\ t) = \phi\ (dimap\ id\ cat\ t)$

$cat\ (Var\ x) = x$

## Example

### Declaring a catamorphism

**class** *Count* *f* **where**

$\phi_{\text{Count}} :: \text{Alg } f \text{ Int}$

*count* :: (*Difunctor* *f*, *Count* *f*)  $\Rightarrow$  *Term* *f*  $\rightarrow$  *Int*

*count* = *cata*  $\phi_{\text{Count}}$



## Example

### Declaring a catamorphism

**class** *Count f* **where**

$\phi_{\text{Count}} :: \text{Alg } f \text{ Int}$

$\text{count} :: (\text{Difunctor } f, \text{Count } f) \Rightarrow \text{Term } f \rightarrow \text{Int}$

$\text{count} = \text{cata } \phi_{\text{Count}}$

### Instantiation

**instance** *Count Lam* **where**

$\phi_{\text{Count}} (\text{Lam } f) = f \ 1$

**instance** *Count App* **where**

$\phi_{\text{Count}} (\text{App } e_1 \ e_2) = e_1 + e_2$



# Extending the Signature

## Let expressions

**data**  $Let\ a\ e = Let\ e\ (a \rightarrow e)$

**type**  $Sig' = Sig \oplus Let$



## Extending the Signature

### Let expressions

**data**  $Let\ a\ e = Let\ e\ (a \rightarrow e)$   
**type**  $Sig' = Sig \oplus Let$

Note:  $Sig \prec Sig'$





## Extending the Signature

### Let expressions

**data**  $Let\ a\ e = Let\ e\ (a \rightarrow e)$

**type**  $Sig' = Sig \oplus Let$

Note:  $Sig \prec Sig'$

### Example

$e \quad :: Term\ Sig$

$e = Term\ \$\ iLam\ (\lambda x \rightarrow x\ 'iApp'\ x)$



## Extending the Signature

### Let expressions

**data**  $Let\ a\ e = Let\ e\ (a \rightarrow e)$

**type**  $Sig' = Sig \oplus Let$

Note:  $Sig \prec Sig'$

### Example

$e \quad :: Term\ Sig'$

$e = Term\ \$\ iLam\ (\lambda x \rightarrow x\ 'iApp'\ x)$



## Extending the Signature

### Let expressions

**data**  $Let\ a\ e = Let\ e\ (a \rightarrow e)$

**type**  $Sig' = Sig \oplus Let$

Note:  $Sig \prec Sig'$

### Example

$e, e' :: Term\ Sig'$

$e = Term\ \$\ iLam\ (\lambda x \rightarrow x\ 'iApp'\ x)$

$e' = Term\ \$\ iLet\ (iLam\ (\lambda x \rightarrow x\ 'iApp'\ x))\ (\lambda y \rightarrow y\ 'iApp'\ y)$



## Extending the Signature

### Let expressions

**data**  $Let\ a\ e = Let\ e\ (a \rightarrow e)$

**type**  $Sig' = Sig \oplus Let$

Note:  $Sig \prec Sig'$

### Example

**let**  $y = \lambda x. x\ x$  **in**  $y\ y$

$e, e' :: \text{Term}$   $\sigma :: \text{Sig}$

$e = \text{Term } \$ \text{ iLam } (\lambda x \rightarrow x \text{ 'iApp' } x)$

$e' = \text{Term } \$ \text{ iLet } (\text{ iLam } (\lambda x \rightarrow x \text{ 'iApp' } x)) (\lambda y \rightarrow y \text{ 'iApp' } y)$



## Extending the Signature

### Let expressions

**data**  $Let\ a\ e = Let\ e\ (a \rightarrow e)$

**type**  $Sig' = Sig \oplus Let$

Note:  $Sig \prec Sig'$

### Example

$e, e' :: Term\ Sig'$

$e = Term\ \$\ iLam\ (\lambda x \rightarrow x\ 'iApp'\ x)$

$e' = Term\ \$\ iLet\ (iLam\ (\lambda x \rightarrow x\ 'iApp'\ x))\ (\lambda y \rightarrow y\ 'iApp'\ y)$

### Extending the variable counter

**instance**  $Count\ Let\ where$

$\phi_{Count}\ (Let\ e\ f) = e + f + 1$

# But Wait, There is More!

## Term transformations

- functions of type  $Term\ f \rightarrow Term\ g$
- e.g. desugaring, constant folding, type inference, annotations
- efficient recursion schemes derived from tree automata



## But Wait, There is More!

### Term transformations

- functions of type  $Term\ f \rightarrow Term\ g$
- e.g. desugaring, constant folding, type inference, annotations
- efficient recursion schemes derived from tree automata

### Monadic computations

- functions of type  $Term\ f \rightarrow m\ r$
- e.g. for pretty printing, evaluation with side effects



# But Wait, There is More!

## Term transformations

- functions of type  $Term\ f \rightarrow Term\ g$
- e.g. desugaring, constant folding, type inference, annotations
- efficient recursion schemes derived from tree automata

## Monadic computations

- functions of type  $Term\ f \rightarrow m\ r$
- e.g. for pretty printing, evaluation with side effects

## Generalised Algebraic Data Types

- difunctors  $\rightsquigarrow$  **indexed** difunctors
- algebras  $\rightsquigarrow$  **many-sorted** algebras



# But Wait, There is More!

## Term transformations

- functions of type  $Term\ f \rightarrow Term\ g$
- e.g. desugaring, constant folding, type inference, annotations
- efficient recursion schemes derived from tree automata

## Monadic computations

- functions of type  $Term\ f \rightarrow m\ r$
- e.g. for pretty printing, evaluation with side effects

## Generalised Algebraic Data Types

- difunctors  $\rightsquigarrow$  **indexed** difunctors
- algebras  $\rightsquigarrow$  **many-sorted** algebras
- mutually recursive data types (with binders)
- for simple type systems (e.g. simply typed lambda calculus)

## Current Work

We use our library constantly.  $\rightsquigarrow$  We extend it constantly.



# Current Work

We use our library constantly.  $\rightsquigarrow$  We extend it constantly.

## Other extensions

- algebras with **nested monadic effect**
- tree homomorphisms
- tree transducers
- attribute grammars



# Current Work

We use our library constantly.  $\rightsquigarrow$  We extend it constantly.

## Other extensions

- algebras with **nested monadic effect**
- tree homomorphisms
- tree transducers
- attribute grammars

## Try it yourself

- `http://hackage.haskell.org/package/compdata`
- `cabal install compdata`



# An example – Language Definition & Desugaring

```

data Lam a b = Lam (a → b)
data App a b = App b b
data Lit a b = Lit Int
data Plus a b = Plus b b
data Let a b = Let b (a → b)
data Err a b = Err

$(derive [smartConstructors, makeDifunctor, makeShowD, makeEqD, makeOrdD]
  [''Lam, ''App, ''Lit, ''Plus, ''Let, ''Err])

e :: Term (Lam :+: App :+: Lit :+: Plus :+: Let :+: Err)
e = Term (iLet (iLit 2) (λx → (iLam (λy → y 'iPlus' x) 'iApp' iLit 3)))

-- * Desugaring
class Desug f g where
  desugHom :: Hom f g

$(derive [liftSum] [''Desug]) -- lift Desug to coproducts

desug :: (Difunctor f, Difunctor g, Desug f g) ⇒ Term f → Term g
desug (Term t) = Term (appHom desugHom t)

instance (Difunctor f, Difunctor g, f <: g) ⇒ Desug f g where
  desugHom = In . fmap Hole . inj -- default instance for core signatures

instance (App <: f, Lam <: f) ⇒ Desug Let f where
  desugHom (Let e1 e2) = inject (Lam (Hole . e2)) 'iApp' Hole e1

```



# An example – Call-By-Value Evaluation

```

data Sem m = Fun (Sem m → m (Sem m)) | Int Int

class Monad m ⇒ Eval m f where
  evalAlg :: Alg f (m (Sem m))

$(derive [liftSum] ['Eval]) -- lift Eval to coproducts

eval :: (Difunctor f, Eval m f) ⇒ Term f → m (Sem m)
eval = cata evalAlg

instance Monad m ⇒ Eval m Lam where
  evalAlg (Lam f) = return (Fun (f . return))

instance MonadError String m ⇒ Eval m App where
  evalAlg (App mx my) = do x ← mx
                          case x of Fun f → my >>= f
                                   -      → throwError "stuck"

instance Monad m ⇒ Eval m Lit where
  evalAlg (Lit n) = return (Int n)

instance MonadError String m ⇒ Eval m Plus where
  evalAlg (Plus mx my) = do x ← mx
                          y ← my
                          case (x,y) of (Int n,Int m) → return (Int (n + m))
                                   -                → throwError "stuck"

instance MonadError String m ⇒ Eval m Err where
  evalAlg Err = throwError "error"

```

