

# Strongly Typed Programs and Proofs

Patrik Jansson, fp.st.cse.chalmers.se

Chalmers University of Technology and University of Gothenburg

2015-08-21



“It is one of the first duties of a professor, in any subject, to exaggerate a little both the importance of his subject and his own importance in it” [A Mathematician's Apology, G. H. Hardy]



This talk: <https://github.com/patrikja/ProfLect>

# My best results over the years (part 1)

Among my best results I count

- early work on Generic Programming [Backhouse et al., 1999, Jansson and Jeuring, 1997] (well cited)
- Polytypic Data Conversion Programs [Jansson and Jeuring, 2002]
- the Bologna structure (3y BSc + 2y MSc) at [cse.chalmers.se](http://cse.chalmers.se) in my role as Vice Head of Department
- self-evaluation reports for the CSE degrees (in my role as Head of the CSE programme). The BSc got “very high quality”.
- Global Systems Science work [Jaeger et al., 2013] leading to the FETPROACT1 call, the GRACeFUL project and the CoEGSS project.



# My best results over the years (part 2)

...continued

- my PhD graduates: Norell, Danielsson, and Bernardy
- Fast and Loose Reasoning [Danielsson et al., 2006]
- Parametricity and dependent types [Bernardy et al., 2010]
- Algebra of Programming in Agda [Mu et al., 2009]
- Feat: functional enumeration of algebraic types [Duregård et al., 2012]



# Functional Polytypic Programming (1995–2000)

What is a “polytypic function”?

Start from the normal *sum* function on lists:

```
sum :: Num a => [a] -> a
sum []      = 0
sum (x : xs) = x + sum xs
```

then generalise to other datatypes like these

```
data [a]      = []      | a : [a]
data Tree a   = Leaf a  | Bin (Tree a) (Tree a)
data Maybe a  = Nothing | Just a
data Rose a   = Fork a [Rose a]
```



# The Haskell language extension PolyP (POPL'97)

We obtain

$$\begin{aligned} psum &:: (Regular\ d, Num\ a) \Rightarrow d\ a \rightarrow a \\ psum &= cata\ fsum \end{aligned}$$

where  $fsum$  is defined by induction over the pattern functor  $f$  of the regular datatype  $d\ a$ .

$$\begin{aligned} \text{polytypic } fsum &:: Num\ a \Rightarrow f\ a\ a \rightarrow a \\ &= \text{case } f \text{ of} \end{aligned}$$

$$g + h \rightarrow \text{either } fsum\ fsum$$

$$g * h \rightarrow \lambda (x, y) \rightarrow fsum\ x + fsum\ y$$

$$Empty \rightarrow \backslash x \rightarrow 0$$

$$Par \rightarrow id$$

$$Rec \rightarrow id$$

$$d @ g \rightarrow psum \circ pmap\ fsum$$

$$Const\ t \rightarrow \backslash x \rightarrow 0$$



# Polytypic $\rightsquigarrow$ Generic Programming

Summer schools lecture notes ( $> 150$  citations each):

- Polytypic Programming [Jeuring & Jansson, 1996]
- Generic Programming - An Introduction [Backhouse, Jansson, Jeuring & Meertens, 1999]

$$\begin{array}{ccc} F \mu F & \xrightarrow{\text{inn}} & \mu F \\ \downarrow F \llbracket \alpha \rrbracket & & \downarrow \llbracket \alpha \rrbracket \\ F A & \xrightarrow{\alpha} & A \end{array}$$

Notation:

$$\begin{aligned} \llbracket \alpha \rrbracket &= \text{cata } \alpha \\ F h &= \text{fmap } h \end{aligned}$$



## Categories, functors and algebras

Category  $C$ , (endo-)functor  $F : C \rightarrow C$ ,  $F$ -algebra  $(A, \alpha : F A \rightarrow A)$ ,



## Categories, functors and algebras

Category  $C$ , (endo-)functor  $F : C \rightarrow C$ ,  $F$ -algebra  $(A, \alpha : F A \rightarrow A)$ ,

## Homomorphisms between algebras

$$h : (A, \alpha) \rightarrow (B, \beta) \quad \text{with} \quad \begin{array}{ccc} F A & \xrightarrow{\alpha} & A \\ \downarrow F h & & \downarrow h \\ F B & \xrightarrow{\beta} & B \end{array}$$



## Categories, functors and algebras

Category  $C$ , (endo-)functor  $F : C \rightarrow C$ ,  $F$ -algebra  $(A, \alpha : F A \rightarrow A)$ ,

## Homomorphisms between algebras

$$h : (A, \alpha) \rightarrow (B, \beta) \quad \text{with} \quad \begin{array}{ccc} F A & \xrightarrow{\alpha} & A \\ \downarrow F h & & \downarrow h \\ F B & \xrightarrow{\beta} & B \end{array}$$

## Catamorphisms

$$([\_]) : (F A \rightarrow A) \rightarrow (\mu F \rightarrow A) \quad \text{with} \quad \begin{array}{ccc} F \mu F & \xrightarrow{\text{inn}} & \mu F \\ \downarrow F([\alpha]) & & \downarrow([\alpha]) \\ F A & \xrightarrow{\alpha} & A \end{array}$$

# Implementing the theory ( $cata = ([\cdot])$ in Haskell)

## Catamorphisms towards implementation

$$\begin{array}{ccc} F \mu F & \xrightarrow{inn} & \mu F \\ \downarrow F([\alpha]) & & \downarrow([\alpha]) \\ F A & \xrightarrow{\alpha} & A \end{array}$$



## Catamorphisms towards implementation

$$\begin{array}{ccc} F \mu F & \xleftarrow{\text{out}} & \mu F \\ \downarrow F([\alpha]) & & \downarrow([\alpha]) \\ F A & \xrightarrow{\alpha} & A \end{array}$$

**data**  $Mu f$  **where** -- Notation:  $Mu f = \mu F$

$Inn :: f (Mu f) \rightarrow Mu f$

$out :: Mu f \rightarrow f (Mu f)$  -- The inverse of  $Inn$

$out (Inn x) = x$

$cata :: Functor f \Rightarrow (f a \rightarrow a) \rightarrow (Mu f \rightarrow a)$

$cata \alpha = \alpha \circ fmap (cata \alpha) \circ out$



## Catamorphisms towards implementation

$$\begin{array}{ccccc} F (F \mu F) & \xleftarrow{F \text{ out}} & F \mu F & \xleftarrow{\text{out}} & \mu F \\ \downarrow F (F ([\alpha])) & & \downarrow F ([\alpha]) & & \downarrow ([\alpha]) \\ F (F A) & \xrightarrow{F \alpha} & F A & \xrightarrow{\alpha} & A \end{array}$$

**data**  $Mu\ f$  **where** -- Notation:  $Mu\ f = \mu F$

$Inn :: f (Mu\ f) \rightarrow Mu\ f$

$out :: Mu\ f \rightarrow f (Mu\ f)$  -- The inverse of  $Inn$

$out (Inn\ x) = x$

$cata :: Functor\ f \Rightarrow (f\ a \rightarrow a) \rightarrow (Mu\ f \rightarrow a)$

$cata\ \alpha = \alpha \circ fmap\ (cata\ \alpha) \circ out$



# Implementing the theory ( $cata = ([\cdot])$ in Haskell)

**data** *Mu f* **where**

*Inn* ::  $f (Mu\ f) \rightarrow Mu\ f$

*out* ::  $Mu\ f \rightarrow f (Mu\ f)$

$out (Inn\ x) = x$

*cata* ::  $Functor\ f \Rightarrow (f\ a \rightarrow a) \rightarrow (Mu\ f \rightarrow A)$

$cata\ \alpha = \alpha \circ fmap\ (cata\ \alpha) \circ out$

Example: *Mu FTree* is the datatype of binary trees with *Int* leaves.

**data** *FTree subtree* **where**

*Leaf* ::  $Int \rightarrow FTree\ subtree$

*Bin* ::  $subtree \rightarrow subtree \rightarrow FTree\ subtree$



# Implementing the theory (arrows in Haskell)

```
class Category cat where    -- In the Haskell library Control.Category
  id  :: cat a a            -- the identity arrow
  (◦) :: cat b c → cat a b → cat a c -- arrow composition
  -- Identity laws:  $id \circ p = p = p \circ id = p$ 
  -- Associativity:  $(p \circ q) \circ r = p \circ (q \circ r)$ 
```



# Implementing the theory (arrows in Haskell)

```
class Category cat where -- In the Haskell library Control.Category
  id :: cat a a           -- the identity arrow
  (◦) :: cat b c → cat a b → cat a c -- arrow composition
  -- Identity laws:  $id \circ p = p = p \circ id = p$ 
  -- Associativity:  $(p \circ q) \circ r = p \circ (q \circ r)$ 
```

```
instance Category (→) where id x = x; (f ◦ g) x = f (g x)
```

```
instance Category (SA s) where -- ...
```

```
data SA s a b = SA ((a, s) → (b, s)) -- “Stateful functions”
```

and many other instances.



# Polytypic Data Conversion Programs

While John Hughes wrote “Generalising Monads to Arrows” [SCP’00] we used them for data conversion [SCP’02].

Motivation:

- save / load documents in editors should preserve the meaning
- but the source code for saving is not connected to that for loading
- proofs of pretty-print / parse round-trip properties are rare

Observations / contributions:

- we can describe both the saving and the loading using arrows
- we formalize the properties required
- we provide generic proofs of the round-trip properties



# Polytypic Data Conversion Programs (cont.)

The starting point was separation of a datastructure (of type  $d a$ ) into its shape ( $d ()$ ) and contents ( $[a]$ ).

$separate :: Regular\ d \Rightarrow SA\ [a]\ (d\ a)\ (d\ ())$

$separate = pmapAr\ put$

$combine :: Regular\ d \Rightarrow SA\ [a]\ (d\ ())\ (d\ a)$

$combine = pmapAl\ get$

$put :: SA\ [a]\ a\ ()$

$get :: SA\ [a]\ ()\ a$

$put = SA\ (\lambda\ (a,\ xs)\ \rightarrow\ ((),\ a : xs))$

$get = SA\ (\lambda\ ((),\ a : xs)\ \rightarrow\ (a,\ xs))$



- 2002: Director of studies
- 2005: Vice Head of Department for education
- 2008: Deputy project leader of “Pedagogical development of Master’s Programmes for the Bologna Structure at Chalmers”
- 2011: Head of the 5-year education programme in Computer Science and Engineering (Civilingenjör Datateknik, Chalmers).
- 2013: Head of the Division of Software Technology



# My PhD graduates: Norell, Danielsson, and Bernardy

I worked on

- generic programs and proofs with Norell



# My PhD graduates: Norell, Danielsson, and Bernardy

I worked on

- generic programs and proofs with Norell

⇒ Agda,



# My PhD graduates: Norell, Danielsson, and Bernardy

I worked on

- generic programs and proofs with Norell

⇒ Agda,

- on program correctness through types with Danielsson



# My PhD graduates: Norell, Danielsson, and Bernardy

I worked on

- generic programs and proofs with Norell

⇒ Agda,

- on program correctness through types with Danielsson  
⇒ Fast'n Loose Reasoning,



# My PhD graduates: Norell, Danielsson, and Bernardy

I worked on

- generic programs and proofs with Norell

⇒ Agda,

- on program correctness through types with Danielsson  
⇒ Fast'n Loose Reasoning, Chasing Bottoms, ...



# My PhD graduates: Norell, Danielsson, and Bernardy

I worked on

- generic programs and proofs with Norell

⇒ Agda,

- on program correctness through types with Danielsson  
⇒ Fast'n Loose Reasoning, Chasing Bottoms, ...
- parametricity for dependent types & testing with Bernardy



I worked on

- generic programs and proofs with Norell

⇒ Agda,

- on program correctness through types with Danielsson  
⇒ Fast'n Loose Reasoning, Chasing Bottoms, ...
- parametricity for dependent types & testing with Bernardy  
Proofs for free:

$$\llbracket \_ \rrbracket : PTS \rightarrow PTS$$
$$\Gamma \vdash A : B \Rightarrow \llbracket \Gamma \rrbracket \vdash \llbracket A \rrbracket : \llbracket B \rrbracket \bar{A}$$

**where**

$\llbracket A \rrbracket$  is the free proof and

$\llbracket B \rrbracket \bar{A}$  is the free theorem

and  $PTS =$  Pure Type System (Barendregt, et al.)



# Global Systems Science (GSS)

with the Potsdam institute for Climate Impact Research (PIK)

- Collaboration from 2007 onwards (main contact: Cezar Ionescu)
- Aim: *correct* software models for simulating global systems



# Global Systems Science (GSS)

with the Potsdam institute for Climate Impact Research (PIK)

- Collaboration from 2007 onwards (main contact: Cezar Ionescu)
- Aim: *correct* software models for simulating global systems
- Algebra of Programming [PhD course and two papers]
- Global Systems Dynamics and Policy (GSDP) [FET-Open 2010–13]
- Workshops including “Domain Specific Languages for Economical and Environmental Modelling”, 2011



# Global Systems Science (GSS)

with the Potsdam institute for Climate Impact Research (PIK)

- Collaboration from 2007 onwards (main contact: Cezar Ionescu)
- Aim: *correct* software models for simulating global systems
- Algebra of Programming [PhD course and two papers]
- Global Systems Dynamics and Policy (GSDP) [FET-Open 2010–13]
- Workshops including “Domain Specific Languages for Economical and Environmental Modelling”, 2011
- The call **FETPROACT1** (Future and Emerging Technology, Proactive support for GSS) in Horizon 2020 is concrete evidence on the success of this line of work.
- Project GRACeFUL: “Global systems Rapid Assessment tools through Constraint FUnctional Languages” granted 2015–2018 by the FETPROACT1 call.



# Global Systems Science (GSS)

with the Potsdam institute for Climate Impact Research (PIK)

- Collaboration from 2007 onwards (main contact: Cezar Ionescu)
- Aim: *correct* software models for simulating global systems
- Algebra of Programming [PhD course and two papers]
- Global Systems Dynamics and Policy (GSDP) [FET-Open 2010–13]
- Workshops including “Domain Specific Languages for Economical and Environmental Modelling”, 2011
- The call **FETPROACT1** (Future and Emerging Technology, Proactive support for GSS) in Horizon 2020 is concrete evidence on the success of this line of work.
- Project GRACeFUL: “Global systems Rapid Assessment tools through Constraint FUnctional Languages” granted 2015–2018 by the FETPROACT1 call.
- Upcoming project CoEGSS: “Center of Excellence for Global Systems Science”, start 2015-10-01, 3y.



# Algebra of Programming in Agda

While Agda was implemented by Norell, Danielsson et al. we used it for the Algebra of Programming.

One highlight is the notation for equality proofs

```
begin
  term1
≡ ⟨ step1 ⟩  -- step1 : term1 ≡ term2
  term2
≡ ⟨ step2 ⟩  -- step2 : term2 ≡ term3
  term3
```



Roughly equivalent to *trans step1 step2* but often more readable (at least in more complicated cases).



# An example proof in Agda, part 1

*expLemma* :  $(x : \mathbb{R}) \rightarrow (m\ n : \mathbb{N}) \rightarrow (x^m *_{\mathbb{R}} x^n \equiv x^{(m+n)})$

*baseCase* :  $(x : \mathbb{R}) \rightarrow (n : \mathbb{N}) \rightarrow (x^Z *_{\mathbb{R}} x^n \equiv x^{(Z+n)})$

*stepCase* :  $(x : \mathbb{R}) \rightarrow (m\ n : \mathbb{N}) \rightarrow$   
 $(ih : x^m *_{\mathbb{R}} x^n \equiv x^{(m+n)}) \rightarrow$   
 $(x^{(S\ m)} *_{\mathbb{R}} x^n \equiv x^{((S\ m)+n)})$

*expLemma* x Z n = *baseCase* x n

*expLemma* x (S m) n = *stepCase* x m n (*expLemma* x m n)



# An example proof in Agda, part 2

$baseCase : (x : \mathbb{R}) \rightarrow (n : \mathbb{N}) \rightarrow (x^Z *_{\mathbb{R}} x^n \equiv x^{(Z + n)})$

$baseCase\ x\ n =$

$begin$

$x^Z *_{\mathbb{R}} x^n$

$\equiv \langle refl \rangle$  -- By definition of  $\_ \hat{\_}$

$one *_{\mathbb{R}} x^n$

$\equiv \langle unitMult\ (x^n) \rangle$  -- Use  $one *_{\mathbb{R}} y = y$  for  $y = x^n$

$x^n$

$\equiv \langle refl \rangle$  -- By definition of  $\_ + \_$

$x^{(Z + n)}$



# Feat: functional enumeration of algebraic types

[with Duregård and Wang, Haskell Symposium 2012]

An efficiently computable bijective function  $index_a :: \mathbb{N} \rightarrow a$ , much like *toEnum* in the *Enum* class.

Example: enumerate “raw abstract syntax trees” for Haskell.

```
*Main> index (105) :: Exp
AppE (LitE (StringL ""))
      (CondE (ListE []) (ListE []) (LitE (IntegerL 1)))
```



# Feat: functional enumeration of algebraic types

[with Duregård and Wang, Haskell Symposium 2012]

An efficiently computable bijective function  $index_a :: \mathbb{N} \rightarrow a$ , much like *toEnum* in the *Enum* class.

Example: enumerate “raw abstract syntax trees” for Haskell.

```
*Main> index (10^5) :: Exp
AppE (LitE (StringL ""))
      (CondE (ListE []) (ListE []) (LitE (IntegerL 1)))
```

```
*Main> index (10^100) :: Exp
ArithSeqE (FromR (AppE (AppE (ArithSeqE (FromR (ListE []))))
... -- and 20 more lines!
```



## DSL: Presenting Math. Analysis Using Functional Programming

$$\forall \epsilon \in \mathbb{R}. (\epsilon > 0) \Rightarrow \exists a \in A. (|a - \sup A| < \epsilon)$$

## Sequential Decision Problems

“Sequential Decision Problems, dependent types and generic solutions”  
“A computational theory of policy advice and avoidability”

## AUTOSAR calculus

“A semantics of core AUTOSAR”  
(AUTOSAR = AUTomotive Open System ARchitecture)

## ValiantAgda

Certified Context-Free Parsing: A form. of Valiant's Algorithm in Agda  
Solve  $C = W + C * C$  for matrices of sets of non-terminals!

Solve  $C = W + C * C$  for strictly upper triangular matrices of something.



# ValiantAgda (the chocolate part;-)

Solve  $C = W + C * C$  for strictly upper triangular matrices of something.



 **Patrik Jansson**  
@patrikja

Applied algebra: an upper triangular matrix.  
Eat the diagonal to make it strictly upper  
triangular. #Agda #chocolate



**Patrik Jansson** @patrikja · 9 Dec 2013

@patrikja Strictly upper triangular matrix. Aim: use these #chocolate block matrices to explain Valiants alg. #Agda



# ValiantAgda (a part in the middle)

[Valiant, 1975] provides a rather awkward iterated def. for all bracketings:

$$C = W + W \cdot W + W \cdot (W \cdot W) + (W \cdot W) \cdot W + (W \cdot W) \cdot (W \cdot W) + \dots$$

We use the smallest solution to the following equation:

$$C \equiv W + C \cdot C$$

(for strictly upper triangular  $W$ ). Or more precisely

$$\text{Clo} : U \rightarrow U \rightarrow \text{Set}$$

$$\text{Clo } W \quad C = \quad C \equiv W + C \cdot C$$

$$\text{LowerBound} : \{A : \text{Set}\} \rightarrow (A \rightarrow \text{Set}) \rightarrow A \rightarrow \text{Set}$$

$$\text{LowerBound } P \ x = \forall z \rightarrow (P \ z \rightarrow x \leq z)$$

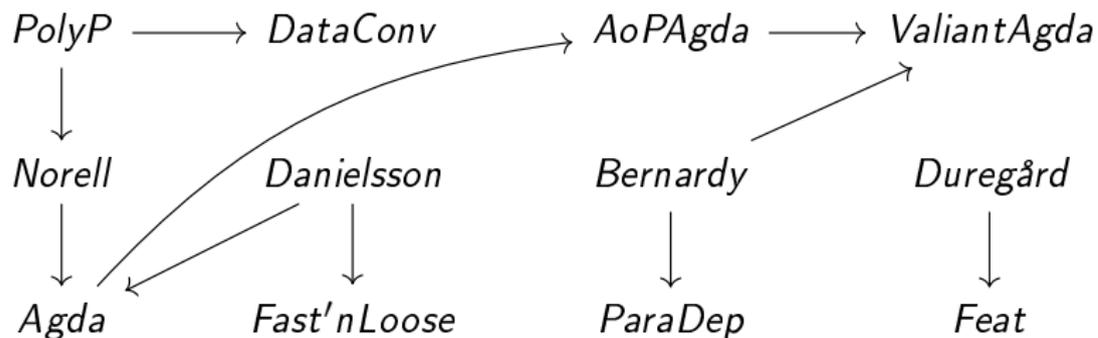
$$\text{Minimal} : \{A : \text{Set}\} \rightarrow (A \rightarrow \text{Set}) \rightarrow A \rightarrow \text{Set}$$

$$\text{Minimal } P \ x = P \ x \wedge \text{LowerBound } P \ x$$

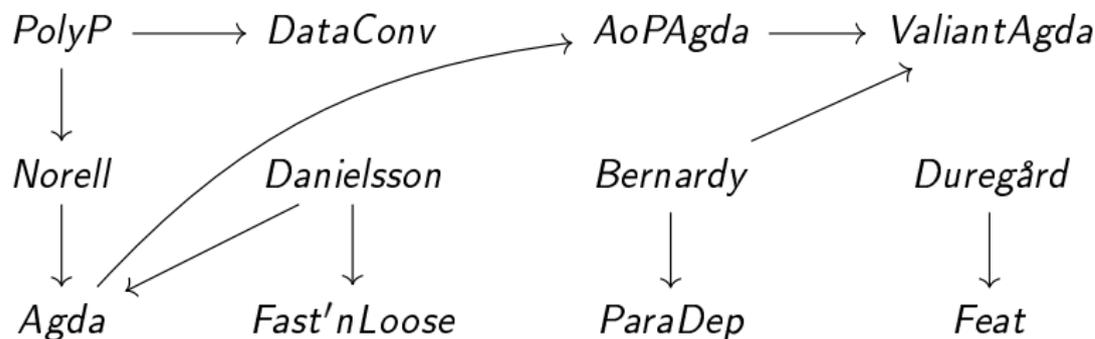
$$\text{Spec} = \forall (W : U) \rightarrow \exists \lambda (C : U) \rightarrow \text{Minimal } (\text{Clo } W) \ C$$



# Summary



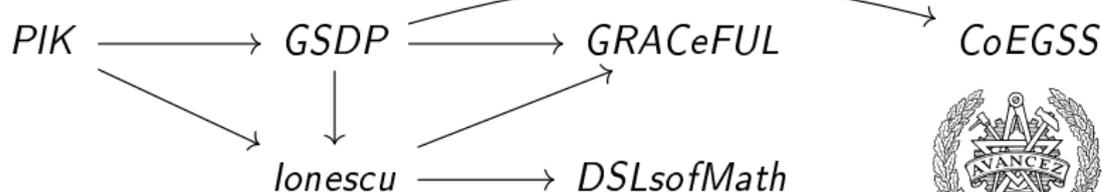
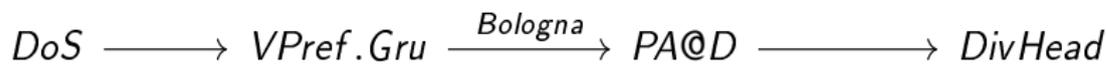
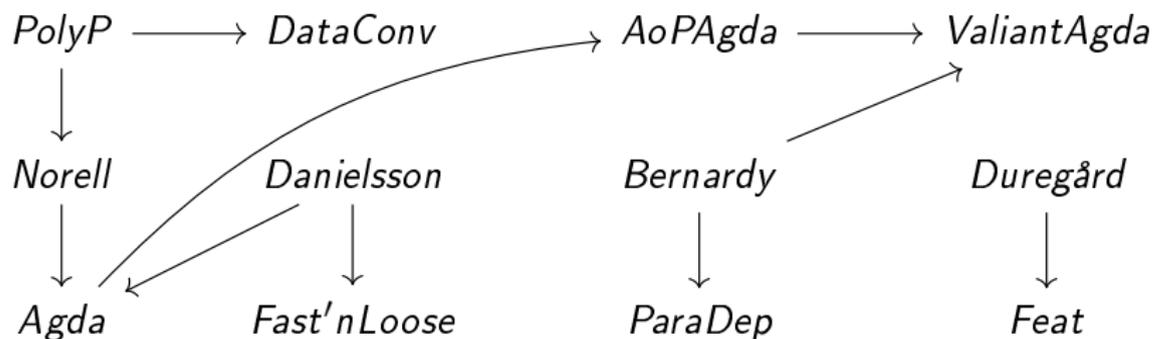
# Summary



*DoS*  $\longrightarrow$  *VPref.Gru*  $\xrightarrow{\text{Bologna}}$  *PA@D*  $\longrightarrow$  *DivHead*



# Summary



- R. Backhouse, P. Jansson, J. Jeuring, and L. Meertens. Generic programming: An introduction. In *Advanced Functional Programming*, volume 1608 of *LNCS*, pages 28–115. Springer, 1999. URL <http://www.cse.chalmers.se/~patrikj/poly/afp98/>.
- J.-P. Bernardy, P. Jansson, and R. Paterson. Parametricity and dependent types. In *Proc. of the 15th ACM SIGPLAN international conference on Funct. programming*, pages 345–356, Baltimore, Maryland, 2010. ACM. doi: 10.1145/1863543.1863592.
- N. A. Danielsson, J. Hughes, P. Jansson, and J. Gibbons. Fast and loose reasoning is morally correct. In *POPL'06*, pages 206–217. ACM Press, 2006. doi: 10.1145/1111037.1111056.
- J. Duregård, P. Jansson, and M. Wang. Feat: Functional enumeration of algebraic types. In *Haskell'12*, pages 61–72. ACM, 2012. doi: 10.1145/2364506.2364515.
- C. Jaeger, P. Jansson, S. van der Leeuw, M. Resch, and J. D. Tabara. GSS: Towards a research program for Global Systems Science. <http://blog.global-systems-science.eu/?p=1512>, 2013. ISBN 978.3.94.1663-12-1. Conference Version, prepared for the Second Open Global Systems Science Conference June 10-12, 2013, Brussels.
- P. Jansson and J. Jeuring. PolyP — a polytypic programming language extension. In *Proc. POPL'97: Principles of Programming Languages*, pages 470–482. ACM Press, 1997. doi: 10.1145/263699.263763.
- P. Jansson and J. Jeuring. Polytypic data conversion programs. *Science of Computer Programming*, 43(1): 35–75, 2002. doi: 10.1016/S0167-6423(01)00020-X.
- S.-C. Mu, H.-S. Ko, and P. Jansson. Algebra of programming in Agda: dependent types for relational program derivation. *J. Funct. Program.*, 19:545–579, 2009. doi: 10.1017/S0956796809007345.

