

# **Generic Programs, Generic Proofs, and Dependent Types**

Peter Dybjer  
Chalmers University of Technology  
Göteborg, Sweden

Cambridge  
24 October 2003

# A new result!

**Theorem.**

$$x = x$$

## Should reflexivity of equality be an axiom?

A. M. Turing (1936):

I shall also suppose that the number of symbols which may be printed is finite. If we were to allow an infinity of symbols, then there would be symbols differing to an arbitrary small extent. The effect of this restriction of the number of symbols is not very serious. It is always possible to use sequences of symbols in the place of single symbols. ... The difference from our point of view between the single and compound symbols is that the compound symbols, if they are too lengthy, cannot be observed at a glance. This is in accordance with experience. We cannot tell at a glance whether 999999999999999999 and 9999999999999999 are the same.

## Reflexivity of generic equality

**Theorem.**

$$(\Sigma : \text{Sig}) \rightarrow (x : T_\Sigma) \rightarrow x =_\Sigma x$$

where

$$(\text{=}_\Sigma) : T_\Sigma \rightarrow T_\Sigma \rightarrow \text{Bool}$$

is a generic equality test,

$$\text{Sig} : \text{Set}$$

is the set of signatures, ie codes for a suitable class of datatypes (cf ML's *equality types*), and

$$T_\Sigma : \text{Set}$$

is (the carrier of) the algebra of  $\Sigma$ -terms, ie an initial  $\Sigma$ -algebra.

## Generic map

Another motivating example of generic programming. We have

$$\text{map} : (X \rightarrow Y) \rightarrow [X] \rightarrow [Y]$$

An analogous function for binary trees:

$$\text{mapBT} : (X \rightarrow Y) \rightarrow \text{BT } X \rightarrow \text{BT } Y$$

In general

$$\text{map}_\Sigma : (X \rightarrow Y) \rightarrow T_\Sigma X \rightarrow T_\Sigma Y$$

where  $\Sigma : \text{Sig}$  is a code for a “regular” parameterized datatype

$$T_\Sigma : \text{Set} \rightarrow \text{Set}$$

## Generic programs and dependent types

Generic programs are written by induction on codes for datatypes (signatures for term algebras). Generic functional programming languages (eg PolyP, Generic Haskell) extend Haskell with special facilities for such definitions.

With dependent types we have natural “internal” types for generic programs, such as

$$(=) : (\Sigma : \text{Sig}) \rightarrow T_\Sigma \rightarrow T_\Sigma \rightarrow \text{Bool}$$

In dependent type theory you can also do generic proofs (Pfeiffer and Rueß 1999).

## Signatures and universes

We here make use of a universe, that is a set of codes

$$\text{Sig}$$

and a decoding function (family of sets)

$$T : \text{Sig} \rightarrow \text{Set}$$

which to each code (signature)  $\Sigma : \text{Sig}$  assigns the datatype (set, term algebra)  $T_\Sigma$  it denotes. (Usually, universes are denoted by the letter  $U$  in dependent type theory, but here we prefer the term signature since we borrow terminology from universal algebra.)

## **Why generic programs and generic proofs?**

Important for libraries of programs!

Libraries are important for proof assistants!

Generic proofs of generic programs help limit the size of libraries.



## Generic dependent type theory

The usual formulations of Martin-Löf Type Theory or the Calculus of Inductive Constructions do not allow definitions on the code of a “datatype” (= inductive definition).

# Martin-Löf type theory and inductive definitions - a brief history

**1970-79** Basic set formers:  $\Pi, \Sigma, +, I, N, N_n, W, U_n$

**1979-85** Adding new set formers with their rules when there is a need for them: lists, binary trees, the well-founded part of a relation, ....

**1985-91** Exactly what is a good inductive definition? Schemata for inductive definitions, indexed inductive definitions, inductive-recursive definitions

**1999-** Generic formulation: universes for inductive definitions, indexed inductive definitions, inductive-recursive definitions

## References

- A finite axiomatization of inductive-recursive definitions (with Anton Setzer). Pages 129 - 146 in Proceedings of TLCA 1999, LNCS 1581.
- Indexed induction-recursion (with Anton Setzer). Pages 93-113 in Proof Theory in Computer Science International Seminar, Dagstuhl Castle, Germany, October 7-12, 2001, LNCS 2183.
- Induction-recursion and initial algebras (with Anton Setzer), 2003. *Annals of Pure and Applied Logic*, in press.
- Universes for generic programs and proofs in dependent type theory (with Marcin Benke and Patrik Jansson), 2003. To appear in *Nordic Journal of Computing*.

# Plan

1. Generic dependent type theory
2. Generic programming - generic equality
3. Generic proofs - reflexivity of generic equality
4. More general classes of datatypes (if time permits)

# Generic dependent type theory

We work in Martin-Löf type theory with rules for one-sorted term algebras (axiomatizing “initial algebra semantics”):

- Logical framework. Rules for  $(x : A) \rightarrow B$ ,  $(x : A) \times B$ ,  $0$ ,  $1$ ,  $2$ ,  $\text{Set}$ .
- Rules for arities and signatures. For simplicity only one-sorted signatures, but much more powerful classes of datatypes can be represented.
- Auxiliary constructions for setting up “initial algebra semantics”
- Generic formation, introduction, elimination, and equality rules for  $\mathbb{T}_\Sigma$ .

## Notation

Dependent function types are written

$$(x : A) \rightarrow B$$

This is the type of functions  $f$  mapping  $a : A$  to  $f a : B[x := a]$ .  
 $B : A \rightarrow \text{Set}$  is an  $A$ -indexed family of sets.

Dependent product types are written

$$(x : A) \times B$$

This is the type of pairs  $(a, b)$  such that  $a : A$  and  $b : B[x := a]$ .

## Recall: rules for natural numbers

$\mathbf{N} : \text{Set}$

$0 : \mathbf{N}$

$\text{Succ} : \mathbf{N} \rightarrow \mathbf{N}$

$\text{natrec} : (C : \mathbf{N} \rightarrow \text{Set}) \rightarrow C\ 0 \rightarrow ((x : \mathbf{N}) \rightarrow C\ x \rightarrow C\ (\text{Succ}\ x))$   
 $\rightarrow (n : \mathbf{N}) \rightarrow C\ n$

$\text{natrec}\ C\ d\ e\ 0 = d$

$\text{natrec}\ C\ d\ e\ (\text{Succ}\ n) = e\ n\ (\text{natrec}\ C\ d\ e\ n)$

## Generic rules for $T_\Sigma$

Formation rule

$$T_\Sigma : \text{Set}$$

Introduction rule

$$\text{Intro}_\Sigma : ? \rightarrow T_\Sigma$$

Elimination rule

$$\text{rec}_\Sigma : (C : T_\Sigma \rightarrow \text{Set}) \rightarrow ? \rightarrow (x : T_\Sigma) \rightarrow C x$$

Equality rule

$$\text{rec}_\Sigma C d (\text{Intro}_\Sigma y) = d y \dots ? \dots$$



## One-sorted term algebras

A one-sorted signature is a finite list of natural numbers, representing the arities of the operations of the signature. Examples:

$\Sigma$	$T_\Sigma$
$[\ ]$	empty set
$[0, 1]$	natural numbers
$[0, 0]$	truth values
$[0, 1, 1]$	lists of truth values
$[0, 2]$	binary trees without information in the nodes

We could consider more general class  $\text{Sig}$  of signatures, eg giving rise to the ML-notion of equality class. Even more, capturing all kinds of inductive definitions in Martin-Löf type theory (indexed, inductive-recursive, ...).

## Signature = universe for one-sorted term algebras

We introduce the type of signatures

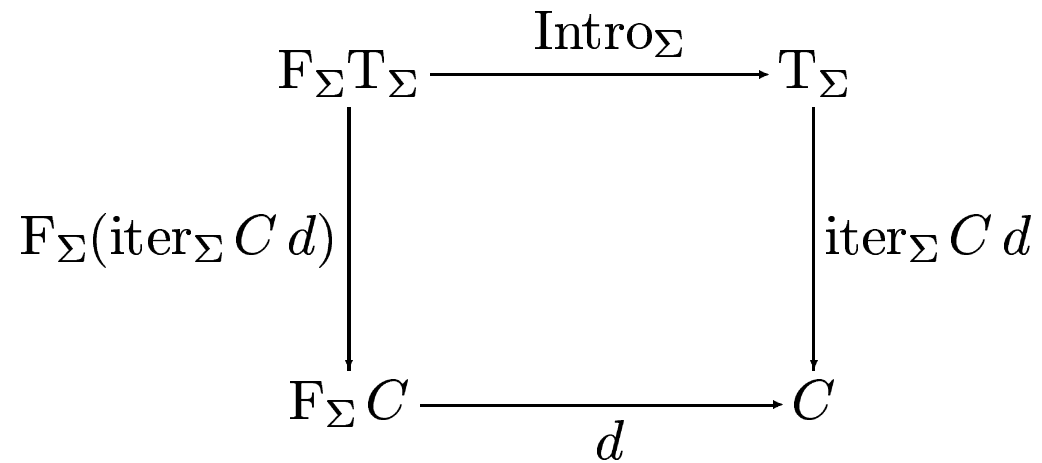
Sig Type

and the decoding function

$T : \text{Sig} \rightarrow \text{Set}$

which maps a signature  $\Sigma$  to (the carrier of) its term algebra  $T_\Sigma : \text{Set}$ .

## The initial $\Sigma$ -algebra diagram



Pattern functor:

$$F_{[n_1, \dots, n_m]} X = X^{n_1} + \dots + X^{n_m}$$

## A diagram for $T_\Sigma$ -elimination

$$\begin{array}{ccccc}
 & & F_\Sigma T_\Sigma & \xrightarrow{\text{Intro}_\Sigma} & T_\Sigma \\
 & & \downarrow \langle \text{id}, F_\Sigma^{\text{map}} T_\Sigma C (\text{rec}_\Sigma C d) \rangle & & \downarrow \langle \text{id}, \text{rec}_\Sigma C d \rangle \\
 F_\Sigma \langle \text{id}, \text{rec}_\Sigma C d \rangle & \swarrow & & & \\
 F_\Sigma((x : T_\Sigma) \times C x) & \xrightarrow{\cong_{\Sigma, T_\Sigma, C}} & (y : F_\Sigma T_\Sigma) \times F_\Sigma^{\text{IH}} T_\Sigma C y & \xrightarrow{e} & (x : T_\Sigma) \times C x
 \end{array}$$

$$e(y, z) = (\text{Intro}_\Sigma y, d y z)$$

## Auxiliary constructions

$$F_{\Sigma}^{\text{IH}} : (X : \text{Set}) \rightarrow (X \rightarrow \text{Set}) \rightarrow F_{\Sigma} X \rightarrow \text{Set}$$

$$F_{[n_1, \dots, n_m]}^{\text{IH}} X C (\text{In}_i (x_1, \dots, x_{n_i})) = C x_1 \times \dots \times C x_{n_i}$$

and

$$F_{\Sigma}^{\text{map}} : (X : \text{Set}) \rightarrow (C : X \rightarrow \text{Set})$$

$$\rightarrow ((x : X) \rightarrow C x) \rightarrow (y : F_{\Sigma} X) \rightarrow F_{\Sigma}^{\text{IH}} X C y$$

$$F_{[n_1, \dots, n_m]}^{\text{map}} X C h (\text{In}_i (x_1, \dots, x_{n_i})) = (h x_1, \dots, h x_{n_i})$$

## Generic rules for $T_\Sigma$

$T_\Sigma$  : Set

Intro $_\Sigma$  :  $F_\Sigma T_\Sigma \rightarrow T_\Sigma$

rec $_\Sigma$  :  $(C : T_\Sigma \rightarrow \text{Set}) \rightarrow ((y : F_\Sigma T_\Sigma) \rightarrow F_\Sigma^{\text{IH}} T_\Sigma C y \rightarrow C (\text{Intro}_\Sigma y))$   
 $\rightarrow (x : T_\Sigma) \rightarrow C x$

Equality rule

$$\text{rec}_\Sigma C d (\text{Intro}_\Sigma y) = d y (F_\Sigma^{\text{map}} C (\text{rec}_\Sigma C d) y)$$

## Large elimination

We may add a large version of this elimination too, where  $C$  can be an arbitrary family of types, that is,

$$C[x] \text{ Type} \quad (x : T_\Sigma)$$

not just a family of sets.

# Martin-Löf type theory with one-sorted term algebras

Start with logical framework (including at least dependent function and product types, and **0, 1, 2**). Add

**arities** formation, introduction, (and elimination and equality) rules for  $\mathbb{N}$

**signatures** formation, introduction (and elimination and equality) rules for lists of natural numbers

**pattern functors** defining rules for object part of  $F_{\Sigma}$ , and the auxiliary  $F_{\Sigma}^{\text{IH}}$  and  $F_{\Sigma}^{\text{map}}$

**term algebras** formation, introduction, elimination, and equality rules for  $T_{\Sigma}$  with constants  $\text{Intro}_{\Sigma}$  and  $\text{rec}_{\Sigma}$ .



# Generic programming

A generic size function.

A special case of the initial algebra diagram. Let  $\Sigma = [n_1, \dots, n_m]$ .

$$\begin{array}{ccc} T_{\Sigma}^{n_1} + \dots + T_{\Sigma}^{n_m} & \xrightarrow{\text{Intro}_{\Sigma}} & T_{\Sigma} \\ \downarrow \text{size}_{\Sigma}^{n_1} + \dots + \text{size}_{\Sigma}^{n_m} & & \downarrow \text{size}_{\Sigma} \\ N^{n_1} + \dots + N^{n_m} & \xrightarrow{\text{size}_{\text{step}}_{\Sigma}} & N \end{array}$$

## Generic recursion step

The recursion step is defined by induction on the signature:

$$\begin{aligned}\text{sizestep}_{n::\Sigma} (\text{Inl } xs) &= 1 + \text{sum}_n xs \\ \text{sizestep}_{n::\Sigma} (\text{Inr } y) &= \text{sizestep}_{\Sigma} y\end{aligned}$$

where

$$\text{sum} : (n : \mathbb{N}) \rightarrow \mathbb{N}^n \rightarrow \mathbb{N}$$

is a function summing the elements of a vector of natural numbers.

# Generic equality

$$\begin{array}{ccc} T_{\Sigma}^{n_1} + \dots + T_{\Sigma}^{n_m} & \xrightarrow{\text{Intro}_{\Sigma}} & T_{\Sigma} \\ \downarrow \scriptstyle{ (=_{\Sigma})^{n_1} + \dots + (=_{\Sigma})^{n_m} } & & \downarrow \scriptstyle{ (=_{\Sigma}) } \\ (T_{\Sigma} \rightarrow \text{Bool})^{n_1} + \dots + (T_{\Sigma} \rightarrow \text{Bool})^{n_m} & \xrightarrow{\text{eqstep}_{\Sigma}} & T_{\Sigma} \rightarrow \text{Bool} \end{array}$$

## Generic recursion step

$$\text{eqstep}_{\Sigma} (\text{In}_i (f_1, \dots, f_{n_i})) (\text{Intro}_{\Sigma} (\text{In}_i (y_1, \dots, y_{n_i}))) = f_1 y_1 \wedge \dots \wedge f_{n_i} y_{n_i}$$

and for  $i \neq j$

$$\text{eqstep}_{\Sigma} (\text{In}_i (f_1, \dots, f_{n_i})) (\text{Intro}_{\Sigma} (\text{In}_j (y_1, \dots, y_{n_j}))) = \text{False}$$

can be defined by induction on the signature.

## Code for the generic recursion step

$$\begin{aligned} (=_{\Sigma}) & : T_{\Sigma} \rightarrow T_{\Sigma} \rightarrow \text{Bool} \\ x =_{\Sigma} x' & = \text{iter}_{\Sigma} \text{eqstep}_{\Sigma} x x' \end{aligned}$$

$$\begin{aligned} \text{eqstep}_{\Sigma} & : F_{\Sigma} (T_{\Sigma} \rightarrow \text{Bool}) \rightarrow T_{\Sigma} \rightarrow \text{Bool} \\ \text{eqstep}_{\Sigma} y x & = \text{recogAll}_{\Sigma} T_{\Sigma} y (\text{out}_{\Sigma} x) \end{aligned}$$

$$\begin{aligned} \text{out}_{\Sigma} & : T_{\Sigma} \rightarrow F_{\Sigma} T_{\Sigma} \\ \text{out}_{\Sigma} x & = \text{rec}_{\Sigma} (\lambda x. F_{\Sigma} T_{\Sigma}) (\lambda yz. y) \end{aligned}$$

## More code for the generic recursion step

$\text{recogAll}_\Sigma : (X : \text{Set}) \rightarrow \text{F}_\Sigma (X \rightarrow \text{Bool}) \rightarrow \text{F}_\Sigma X \rightarrow \text{Bool}$

$\text{recogAll}_{n::\Sigma} \quad X \quad (\text{Inl } f s) \quad (\text{Inl } x s) \quad = \quad \text{andArgs}_n X f s x s$

$\text{recogAll}_{n::\Sigma} \quad X \quad (\text{Inr } x) \quad (\text{Inr } y) \quad = \quad \text{recogAll}_\Sigma X x y$

$\text{recogAll}_{n::\Sigma} \quad X \quad (\text{Inl } f s) \quad (\text{Inr } y) \quad = \quad \text{False}$

$\text{recogAll}_{n::\Sigma} \quad X \quad (\text{Inr } x) \quad (\text{Inl } x s) \quad = \quad \text{False}$

$\text{andArgs}_n : (X : \text{Set}) \rightarrow (X \rightarrow \text{Bool})^n \rightarrow X^n \rightarrow \text{Bool}$

$\text{andArgs}_0 \quad X \quad () \quad () \quad = \quad \text{True}$

$\text{andArgs}_{m+1} \quad X \quad (f, f s) \quad (x, x s) \quad = \quad f x \wedge \text{andArgs}_m X f s x s$

## Reflexivity of generic equality

Instantiate the generic elimination rule (structural recursion on  $T_\Sigma$ )!

$$\begin{array}{ccc}
 F_\Sigma T_\Sigma & \xrightarrow{\text{Intro}_\Sigma} & T_\Sigma \\
 \downarrow \langle \text{id}, F_\Sigma^{\text{map}} T_\Sigma C \text{ref}_\Sigma \rangle & & \downarrow \langle \text{id}, \text{ref}_\Sigma \rangle \\
 (y : F_\Sigma T_\Sigma) \times F_\Sigma^{\text{IH}} T_\Sigma C y & \xrightarrow{\lambda(y, z).(\text{Intro}_\Sigma y, \text{refstep}_\Sigma y z)} & (x : T_\Sigma) \times |x =_\Sigma x|
 \end{array}$$

$$\begin{array}{ccc}
F_{\Sigma} T_{\Sigma} & \xrightarrow{\text{Intro}_{\Sigma}} & T_{\Sigma} \\
\downarrow \langle \text{id}, F_{\Sigma}^{\text{map}} T_{\Sigma} C \text{ref}_{\Sigma} \rangle & & \downarrow \langle \text{id}, \text{ref}_{\Sigma} \rangle \\
(y : F_{\Sigma} T_{\Sigma}) \times F_{\Sigma}^{\text{IH}} T_{\Sigma} C y & \xrightarrow{\lambda(y, z).(\text{Intro}_{\Sigma} y, \text{refstep}_{\Sigma} y z)} & (x : T_{\Sigma}) \times C x
\end{array}$$

$$C x = |x =_{\Sigma} x|$$

$$F_{\Sigma}^{\text{IH}} T_{\Sigma} C (\text{In}_i(x_1, \dots, x_{n_i})) = |x_1 =_{\Sigma} x_1| \times \dots \times |x_{n_i} =_{\Sigma} x_{n_i}|$$

$$F_{\Sigma}^{\text{map}} T_{\Sigma} C \text{ref}_{\Sigma} (\text{In}_i(x_1, \dots, x_{n_i})) = (\text{ref}_{\Sigma} x_1, \dots, \text{ref}_{\Sigma} x_{n_i})$$



## Converting from a truth value to a set

Truth values, sets, and propositions:

$$|-| : \text{Bool} \rightarrow \text{Set}$$

$$|\text{False}| = \mathbf{0}$$

$$|\text{True}| = 1$$

$$\perp = \mathbf{0}$$

$$\top = 1$$

## How to write $\text{refstep}_\Sigma$

$$\text{refstep}_\Sigma \quad : \quad (y : F_\Sigma T_\Sigma) \rightarrow F_\Sigma^{\text{IH}} T_\Sigma C y \rightarrow C (\text{Intro}_\Sigma y)$$

Hence, informally

$$\begin{aligned} & \text{refstep}_\Sigma (\text{In}_i (x_1, \dots, x_{n_i})) \\ : & \quad |x_1 =_\Sigma x_1| \times \dots \times |x_{n_i} =_\Sigma x_{n_i}| \rightarrow |\text{Intro}_\Sigma (\text{In}_i (x_1, \dots, x_{n_i})) =_\Sigma \text{Intro}_\Sigma (\text{In}_i (x_1, \dots, x_{n_i}))| \end{aligned}$$

Simplify the result type:

$$|x_1 =_\Sigma x_1| \times \dots \times |x_{n_i} =_\Sigma x_{n_i}| \rightarrow |(x_1 =_\Sigma x_1) \wedge \dots \wedge (x_{n_i} =_\Sigma x_{n_i})|$$

## Two abbreviations

Rel : Set  $\rightarrow$  Set

Rel  $X = X \rightarrow X \rightarrow$  Bool

lref : (X : Set)  $\rightarrow$  Rel X  $\rightarrow$  X  $\rightarrow$  Set

lref  $X r x = |r x x|$

where the first argument to lref will be hidden for brevity.

## Code for generic reflexivity

$$\begin{aligned} \text{ref}_\Sigma & : (x : \mathbb{T}_\Sigma) \rightarrow |x =_\Sigma x| \\ \text{ref}_\Sigma & = \text{rec}_\Sigma (\text{lref } (=_\Sigma)) (\text{refmatch}_\Sigma \mathbb{T}_\Sigma (=_\Sigma)) \end{aligned}$$

ie  $\text{refstep}_\Sigma$  is a special case of  $\text{refmatch}$  defined by:

$$\begin{aligned} \text{refmatch}_\Sigma & : (X : \text{Set}) \rightarrow (r : \text{Rel } X) \rightarrow \\ & (y : \mathbb{F}_\Sigma X) \rightarrow \mathbb{F}_\Sigma^{\text{IH}} X (\text{lref } r) y \rightarrow |\text{recogAll}_\Sigma X (\mathbb{F}_\Sigma^1 X \text{ Bool } r y) y| \end{aligned}$$

$$\begin{aligned} \text{refmatch}_{n::\Sigma} X r (\text{Inl } xs) & = \text{refargs}_n X r xs \\ \text{refmatch}_{n::\Sigma} X r (\text{Inr } y) & = \text{refmatch}_\Sigma X r y \end{aligned}$$

## More code for generic reflexivity

refargs handles the arguments to the constructor:

$$\begin{aligned} \text{refargs}_n & : (X : \text{Set}) \rightarrow (r : \text{Rel } X) \rightarrow \\ & (xs : X^n) \rightarrow (\text{lref } r)^n xs \rightarrow |\text{andArgs}_n X (r^n xs) xs| \end{aligned}$$

$$\begin{array}{l} \text{refargs}_0 \quad X \quad r \quad () \quad () \quad = \quad () \\ \text{refargs}_{m+1} \quad X \quad r \quad (x, xs) \quad (ih, ihs) \quad = \quad ih (\text{refargs}_m X r xs ihs) \end{array}$$

**This concludes the proof ...**

This is the first, or perhaps more accurately the second ..., generic proof.

## Is this a tour de force?

- Equality is not the simplest generic function. Reflexivity of equality is not the simplest generic proof! Nevertheless, the proof is not so long. But it uses complex dependent types - comes from general framework.
- $(=) : \mathbb{N} \rightarrow \mathbb{N} \rightarrow \text{Bool}$  written with primitive recursion on higher type is superficially a bit more complex than written by 4 recursion equations.
- Would it be simpler to prove reflexivity in a framework with general recursion? YOU set up this logic for generic general recursive programs.
- Type-checking dependent types sometimes difficult - types may be *extensionally* but not *definitionally (computationally)* equal.

# Generalizing the notion of a signature

## iterated inductive definitions

$$\text{Cons} : \mathbf{N} \rightarrow \text{ListN} \rightarrow \text{ListN}$$

the first argument is a side-condition.

## generalized inductive definitions

$$\text{Sup} : (\mathbf{N} \rightarrow \mathcal{O}) \rightarrow \mathcal{O}$$

## constructors with dependent types

$$\text{Sup} : (x : A) \rightarrow (B x \rightarrow W) \rightarrow W$$



# Signatures for generalized inductive definitions

$$\epsilon : \text{Sig}$$

$$\sigma : (A : \text{Set}) \rightarrow (A \rightarrow \text{Sig}) \rightarrow \text{Sig}$$

$$\rho : \text{Set} \rightarrow \text{Sig} \rightarrow \text{Sig}$$

$$F_{\epsilon} X = \mathbf{1}$$

$$F_{\sigma A \Sigma} X = (x : A) \times F_{\Sigma x} X$$

$$F_{\rho A \Sigma} X = (A \rightarrow X) \times F_{\Sigma} X$$

# Parameterized inductive definitions

## parameterized inductive definitions

$$\text{Cons} : (A : \text{Set}) \rightarrow A \rightarrow \text{List } A \rightarrow \text{List } A$$

The first argument is a parameter - it is the same for all constructors.

We need universes for parameterized inductive definitions because much generic programming is about such definitions (generic map, zip, etc).

If we just want to formalize parameterized inductive definitions in Martin-Löf type theory, parameterization is taken care of by the logical framework.