# Mathematical specification of programming languages using monads and modules over them

#### Ambroise Lafont<sup>1</sup>

<sup>1</sup>University of New South Wales Sydney, Australia

November 26, 2020

## That is the question

What is a programming language, mathematically?

### Differential $\lambda$ -calculus [Ehrhard-Regnier '03]

 $\sim$ 10 pages (section 2  $\rightarrow$  beginning of section 3) describing the programming language and proving some properties.

#### Contributions presented in this talk

- a notion of programming language, transition monads [Hirschowitz-Hirschowitz-Lafont '20], and
- a discipline for automatically generating them.

#### Features of this approach

- monads and modules to take care of substitution.
- works with simple types (in this talk: untyped case)

### Related work

Syntax with variable binding

#### Two main notions of syntax:

- Nominal sets [Gabbay-Pitts '99].
  - Injective renamings of variables built-in.
- Substitution monoids [Fiore-Plotkin-Turi '99].
  - Simultaneous substitution of variables built-in.
  - Transition monads (syntax): a variant of this approach

### Related work

#### Notion of programming language and specification

- Reduction monads [Ahrens-Hirschowitz-Lafont-Maggesi '20]
  - cbv λ-calculus out of reach
  - + transition monads = a generalisation of reduction monads
- Mathematical Operational Semantics [Turi-Plotkin '97]
  - + Deeply developed
  - Higher-order languages (such as λ-calculus) only starting to be investigated [Peressotti '17]
- Rewriting with variable binding (categorical approach)
  - e.g. [Hamana '03, Hirschowitz '13, Ahrens '16]
  - only congruent transitions ⇒ weak reduction out of reach

# Examples of transition monads

- cbv/cbn λ-calculus (big/small-step)
- $\overline{\lambda}\mu$ -calculus
- π-calculus
- differential λ-calculus
- computational  $\lambda$ -calculus (variant of [Dal Lago-Gavazzo-Levy '17])
- GSOS systems

# What is still missing

### Metatheorems, e.g., congruence of bisimilarity

- [Borthelle-Hirschowitz-<u>Lafont</u> '20], Howe's method in a different setting.
- [Dal Lago-Gavazzo-Levy '17], Howe's method for particular cases of transition monads (computational λ-calculus).
- Can we generalize both approaches?

#### Limitations

- $\pi$ -calculus as a labelled transition system?
- signature for the computational  $\lambda$ -calculus?
- simple types ok
  - linear types?
  - polymorphic/dependent types?
  - subtyping?

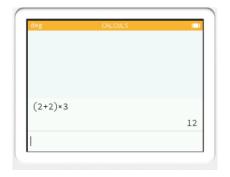
- Transition monads
  - Substitution
  - Graphs
  - Definition
- @ Generating transition monads
  - Three-level specification
  - Examples
- Generating compilations by initiality
- 4 Conclusion

- Transition monads
  - Substitution
  - Graphs
  - Definition
- 2 Generating transition monads
  - Three-level specification
  - Examples
- 3 Generating compilations by initiality
- 4 Conclusion

# What is a programming language?

#### 2 components:

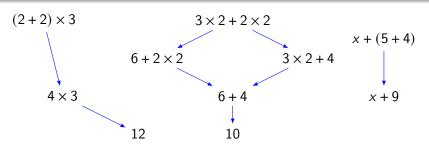
- Syntax: formal language for writing programs;
- Operational semantics: how do programs execute.



$$(2+2) \times 3 \xrightarrow{\qquad \qquad 1 \text{ execution step} \qquad \qquad 4 \times 3 \xrightarrow{\qquad \qquad \qquad 12}$$

# What is a programming language?

A graph whose vertices are programs.



#### Variables = placeholders for expressions

- Substitution: (x + (5 + 4))[x := 12] = 12 + (5 + 4)
- Reductions are stable under substitution

$$\frac{x + (5 + 4) \to x + 9}{12 + (5 + 4) \to 12 + 9}$$

→ Transition monads!

# Ingredients

- Programming languages (PLs) as graphs
  - (Syntax) vertices = terms
  - (Semantics) arrows = reductions between terms
- Simultaneous substitution: variables → terms
  - monads and modules over them

### Example

 $\lambda$ -calculus with  $\beta$ -reduction:

Syntax:

$$S, T ::= x \mid S T \mid \lambda x.S$$

• Modulo  $\alpha$ -equivalence, e.g.

$$\lambda x.x = \lambda y.y$$

• Reductions:

$$(\lambda x.t) u \xrightarrow{\beta} t[x := u] +$$

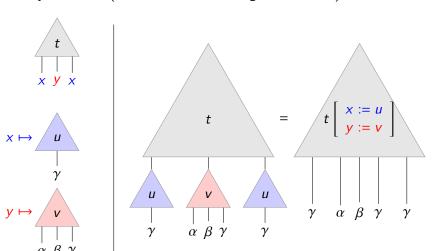
congruences

- Transition monads
  - Substitution
  - Graphs
  - Definition
- 2 Generating transition monads
  - Three-level specification
  - Examples
- Generating compilations by initiality
- 4 Conclusion

### Simultaneous substitution

Syntax comes with substitution

terms as syntax trees (free variables as distinguished leaves).



# Simultaneous substitution made formal

#### Free variables indexing

 $L(X) = \{\text{terms taking free variables in } X\}$ 

#### Example: $\lambda$ -calculus

### Simultaneous substitution (bind)

$$\forall f: X \to L(Y),$$

$$\begin{array}{ccc} L(X) & \to L(Y) \\ t & \mapsto t[x \mapsto f(x)] & \text{(or } t[f]) \end{array}$$

14/4/

### Simultaneous substitution

### Monads model simultaneous substitution

 $\lambda$ -calculus as a monad  $(L, \underline{\ }[\underline{\ }], \eta)$ 

- Simultaneous substitution (L, \_[\_])
- Variables are terms

$$\eta_X: X \to L(X) \\
x \mapsto \frac{x}{x}$$

Substitution laws:

$$x[f] = f(x)$$
  $t[x \mapsto x] = t$ 

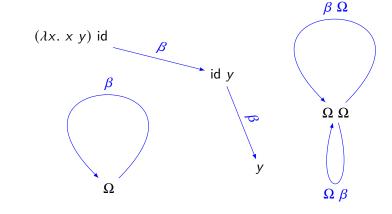
+ associativity:

$$t[f][g] = t[x \mapsto f(x)[g]]$$

- Transition monads
  - Substitution
  - Graphs
  - Definition
- Quantification Generating transition monads
  - Three-level specification
  - Examples
- 3 Generating compilations by initiality
- 4 Conclusion

### PLs as graphs

Example:  $\lambda$ -calculus with  $\beta$ -reduction



- (Syntax) vertices = terms e.g.  $\Omega = (\lambda x.xx)(\lambda x.xx)$
- (Semantics) arrows = reductions

# Graphs

Definition

$$\begin{aligned} \mathsf{Graph} &= \mathsf{a} \; \mathsf{quadruple} \; (A,V,\sigma,\tau) \; \mathsf{where} \\ &\quad A = \{\mathsf{arrows}\} \\ &\quad V = \{\mathsf{vertices}\} \\ &\quad A \xrightarrow{source} V \\ &\quad \sigma : \quad A \quad \to V \qquad \tau : \quad A \quad \to V \\ &\quad t \xrightarrow{r} u \quad \mapsto t \qquad \qquad t \xrightarrow{r} u \quad \mapsto u \\ &\quad \sigma(r) \xrightarrow{r} \tau(r) \end{aligned}$$

### Substitution for semantics

Syntax supports substitution. This is also true of semantics.

#### Our notion of PL:

- **Syntax**: a monad (*L*, \_[\_], η)
- Semantics:
  - graphs  $R(X) \xrightarrow{source_X} L(X)$  for each X

$$R(X) =$$
total set of reductions between terms taking free variables in  $X$ 

• substitution of reduction: variables  $\mapsto$  *L*-terms.

$$\frac{t \xrightarrow{r} u}{t[f] \xrightarrow{r[f]} u[f]} \qquad f: X \to L(Y)$$

⇒ R is a L-module, and source, target are module morphisms (see next slide)

### Substitution for semantics made formal

#### R as a **module** over L

R supports L-monadic substitution:

$$\forall f: X \to \mathbf{L}(Y),$$

$$R(X) \to R(Y)$$

$$r \mapsto r[x \mapsto f(x)] \text{ (or } r[f])$$
+ substitution laws

Other examples of *L*-modules:  $L, L \times L, 1, \dots$ 

### source and target as L-module morphisms

if 
$$source(r) \xrightarrow{r} target(r)$$
 then  $source(r[f]) \xrightarrow{r[f]} target(r[f])$ .

We want 
$$source(r)[f] \xrightarrow{r[f]} target(r)[f], i.e.,$$

$$source(r)[f] = source(r[f])$$
  $target(r)[f] = target(r[f])$ 

Commutation with substitution  $\Leftrightarrow$  Module morphisms  $R \xrightarrow[target]{source} L$ .

- Transition monads
  - Substitution
  - Graphs
  - Definition
- Quantification Generating transition monads
  - Three-level specification
  - Examples
- Generating compilations by initiality
- 4 Conclusion

# Transition monads (first attempt)

Summary: graphs + substitution.

#### Definition

A **transition monad**  $R \xrightarrow[target]{source} T$  consists of

- T = monad (= module over itself)
- R = module over T
- source, target :  $R \rightarrow T$  are T-module morphisms.

#### Example

 $\lambda$ -calculus with  $\beta$ -reduction.

What about cbv  $\lambda$ -calculus?

 Reductions are stable under substitution with values, not with terms!

### Transition monads

cbv λ-calculus (big-step)	Values (monad)	Transitions  source  Values
transition monads	a monad T	$T$ -module morphisms $M_1 \stackrel{source}{\longleftarrow} Trans \stackrel{target}{\longrightarrow} M_2$ (bipartite graph)
reduction monads <sup>1</sup>	a monad T	$T \stackrel{source}{\longleftarrow} Trans \stackrel{target}{\longrightarrow} T$

- Untyped case: base category = Set
- Simply-typed case: base category =  $Set^{Types}$

<sup>&</sup>lt;sup>1</sup>[Ahrens-Hirschowitz-<u>Lafont</u>-Maggesi '20]

# Morphisms of transition monads

Simple case  $M_i = T$ 

Morphism 
$$(T \leftarrow Trans \rightarrow T) \rightarrow (T' \leftarrow Trans' \rightarrow T') =$$

$$(Syntax) \text{ A monad morphism}^1 T \xrightarrow{c} T'$$

$$(Semantics) \text{ Forward simulation}^2: \text{ if } t_1 \xrightarrow{r} t_2, \text{ then } c(t_1) \xrightarrow{\llbracket r \rrbracket} c(t_2)$$

### Examples (detailed later)

- $\lambda$ -calculus + fixpoint op.  $\longrightarrow \lambda$ -calculus
- $\lambda$ -calculus + explicit substitution  $t[x/u] \longrightarrow \lambda$ -calculus

<sup>&</sup>lt;sup>1</sup>mapping preserving substitution and variables

<sup>&</sup>lt;sup>2</sup>backward simulations are often considered as a correctness criteria

- Transition monads
  - Substitution
  - Graphs
  - Definition
- 2 Generating transition monads
  - Three-level specification
  - Examples
- Generating compilations by initiality
- 4 Conclusion

- Transition monads
  - Substitution
  - Graphs
  - Definition
- 2 Generating transition monads
  - Three-level specification
  - Examples
- Generating compilations by initiality
- 4 Conclusion

# Constructing transition monads

programming language = transition monad.

Can we construct them from simple specifications?

#### Overview

- *simple specification* = **signature** for transition monads
- existence (unique up to iso) of a transition monad matching a spec

Transition monads Generating transition monads Generating compilations by initiality Conclusion

# Specification through initial semantics

Notion of signature

### Example (syntax)

A list of operation symbols with associated arities

- To each signature is associated
  - a notion of model

#### Example

- a monad equipped with the operations of the signature
  - a notion of morphism of models

#### Example

- a monad morphism preserving operations
  - a proof that the category of models has an initial object
  - ullet object specified by the signature  $\stackrel{def}{=}$  initial model
    - Initiality  $\Rightarrow$  recursion principle.

# Three-level specification

Transition monad = 
$$(T, M_1 \stackrel{source}{\longleftarrow} Trans \stackrel{target}{\longrightarrow} M_2)$$

#### Three spec steps:

Step	Component	Nature	Specification
1	T	monad	Operations + Equations
2	$M_1, M_2$	T-modules	Operations + Equations
	Trans,		Transition rules as
3	source,	"transition structure"	$t_1 \rightarrow u_1 \dots t_n \rightarrow u_n$
	target		$t \rightarrow u$

⇒ signature for transition monads = signature for each component

- Transition monads
  - Substitution
  - Graphs
  - Definition
- 2 Generating transition monads
  - Three-level specification
  - Examples
- Generating compilations by initiality
- 4 Conclusion

# Examples

Transition monad =  $(T, M_1 \stackrel{source}{\longleftarrow} Trans \stackrel{target}{\longrightarrow} M_2)$ 

#### Upcoming examples

1.	cbn λ-calculus	full signature (sketched)
2.	cbn λ-calculus	signature for <i>T</i>
3.	cbn λ-calculus	left congruence rule for application
4.	cbn λ-calculus	congruence rule for abstraction (involves a binding variable)
5.	cbv λ-calculus	signature for $M_i$
6.	differential $\lambda$ -calculus	signature for $M_i$
7.	differential $\lambda$ -calculus	signature for <i>T</i>

# Example 1/7: small-step cbn $\lambda$ -calculus

Transition monad =  $(T, M_1 \stackrel{source}{\longleftarrow} Trans \stackrel{target}{\longrightarrow} M_2)$ 

### Signature for cbn $\lambda$ -calculus

Step	Component	Nature	Specification
1	T	monad	Operations = app, abs
2	$M_1, M_2$	T-modules	$M_1 = M_2 = T$
3	Trans, source, target	"transition structure"	eta-rule $+$ congruences

# Example 2/7: Specify the monad of $\lambda$ -terms

(untyped) cbn  $\lambda$ -calculus:  $(T, T \xleftarrow{source} Trans \xrightarrow{target} T)$ 

Syntax "generated" by

application	$T \times T \to T$	
$\lambda$ -abstraction	$T' \to T$	T' = module of terms depending
λx.t	$I \rightarrow I$	on an extra variable
(variables)	Var  o T	

#### Signature for T

2 operations (application/abstraction)

- Monads always have variables: no need to specify them
- "operation" = module morphism, i.e., compatible with substitution:

$$(t_1 t_2)[y \mapsto u_y] = t_1[y \mapsto u_y] t_2[y \mapsto u_y]$$

References "Second-order equational logic", Fiore-Hur '10, [Ahrens-Hirschowitz-Lafont-Maggesi. '19]

# Disgression on T'

• M' = derivative of a module M:

$$X$$
 extended with a fresh variable  $x$   
 $M'(X) = M(X \coprod \{x\})$ 

used to model an operation binding a variable.

$$\mathsf{abs}: \ L' \to L \qquad \left\{ \begin{array}{c} \mathsf{abs}_X : L(X \amalg \{x\}) \to L(X) \\ t \mapsto \lambda x. t \end{array} \right.$$

# Example 3/7: Left congruence for application

cbn 
$$\lambda$$
-calculus:  $(T, T \xleftarrow{source} Trans \xrightarrow{target} T)$ 

#### Left congruence rule for application

$$\frac{t_1 \to t_2}{app(t_1, u) \to app(t_2, u)}$$

Easy interpretation of transition rules:

Components of the rule	Interpreted as
3 "metavariables":	a "metavariable" $T$ -module
$t_1, t_2, u$	$V = T \times T \times T$
1 "premise":	$V \rightarrow M_1 \times M_2$ ( <i>T</i> -module
$t_1 \rightarrow t_2$	$(t_1,t_2,u)\mapsto (t_1,t_2)$ morphism)
"conclusion":	$V \rightarrow M_1 \times M_2$
$app(t_1, u) \rightarrow app(t_2, u)$	$(t_1, t_2, u) \mapsto (app(t_1, u), app(t_2, u))$

# Example 4/7: Binding variables in rules

cbn  $\lambda$ -calculus:  $(T, T \xleftarrow{source} Trans \xrightarrow{target} T)$ 

#### Congruence rule for abstraction

$$\frac{t_1 \to t_2}{\lambda x. t_1 \to \lambda x. t_2}$$

- "metavariables"  $t_1$  and  $t_2$ : terms that may depend on x.
- T' = T-module of terms depending on an additional variable

Components of the rule	Interpreted as	
2 "metavariables": $t_1, t_2$	a "metavariable" $T$ -module $V = T' \times T'$	
1 "premise":	$V \to T' \times T'$ ( <i>T</i> -module	
$t_1 \rightarrow t_2$	$(t_1, t_2) \mapsto (t_1, t_2)$ morphism)	
"conclusion":	$V \to T \times T$	
$\lambda x.t_1 \rightarrow \lambda x.t_2$	$(t_1, t_2) \mapsto (\lambda x. t_1, \lambda x. t_2)$	

# Example 5/7: Specify $M_i$ for cbv

Transition monad 
$$= (T, M_1 \xleftarrow{source} Trans \xrightarrow{target} M_2)$$
  
big-step cbv  $\lambda$ -calculus  $= (Vals, Tms \xleftarrow{source} Trans \xrightarrow{target} Vals)$ 

#### Syntax of values and terms

$$Vals: v, w ::= x | \lambda x.t$$

$$Tms: t, u ::= \underbrace{x | \lambda x.t}_{v} | t u \qquad \Rightarrow \qquad terms = binary trees of values$$

$$Tms = BinTree \quad o \quad Vals$$

In fact, by definition of a transition monad,

•  $M_i$  is always of the shape  $S_i \circ T$ . Here,

$$T = Vals$$
  $M_1 = BinTree \circ T$   $M_2 = Id \circ T (= T)$ 

• Signature for  $M_i$  = Signature for  $S_i$ 

#### Signature for BinTree

variables (= labelled leaves) + 1 binary operation (building nodes)

# Example 6/7: Specify $M_i$ for DLC

Transition monad =  $(T, M_1 \xleftarrow{source} Trans \xrightarrow{target} M_2)$ 

#### Differential $\lambda$ -calculus (DLC)

Syntax monad 
$$T$$
 of terms (a variant of  $\lambda$ -calculus)  
Semantics a term  $t$  reduces to a multiterm  $t_1 + \cdots + t_n$   
 $M_1 = Id \circ T \ (=T)$  multiterms = formal sum of terms  
 $M_2 = Formal Sum \circ T$ 

#### Signature for FormalSum

Operations	a constant 0, a binary operation $+$ , variables
Equations	commutativity, associativity, unitality

# Example 7/7: the monad of DLC

differential  $\lambda$ -calculus:  $(T, M_1 \xleftarrow{source} Trans \xrightarrow{target} M_2)$ 

• Syntax of DLC = variant of  $\lambda$ -calculus

### Application of DLC

$$app:(t,U)\mapsto tU$$

input of app = a term t and a multi-term  $U = u_1 + \cdots + u_n$ = a term and a formal sum of terms

input module of  $app = T \times (FormalSum \circ T)$ 

## Signature for T

3 operations (no equation):

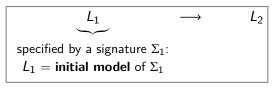
application t U	$T \times (FormalSum \circ T) \rightarrow T$
differential application $Dt \cdot u$	$T \times T \to T$
$\lambda$ -abstraction	(as before)

## Outline

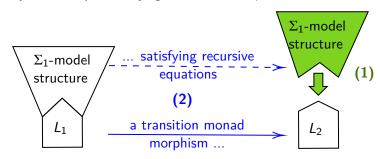
- Transition monads
  - Substitution
  - Graphs
  - Definition
- 2 Generating transition monads
  - Three-level specification
  - Examples
- Generating compilations by initiality
- 4 Conclusion

# Generating compilations by initiality

Initiality ≈ recursion principle



**Data generating a compilation**: a  $\Sigma_1$ -model structure for  $L_2$   $\Rightarrow$  By recursion/**initiality**, get a model morphism  $L_1 \rightarrow L_2$ 



# Examples

$$L_1 \longrightarrow L_2$$

specified by a signature  $\Sigma_1$ :

#### Recipe:

- **1** provide a  $\Sigma_1$ -model structure for  $L_2$
- 2 as a model morphism, the induced compilation satisfies recursive equations.

#### Upcoming examples

- $\lambda$ -calculus + formal fixpoint op.  $\longrightarrow \lambda$ -calculus
  - **①** construct a fixpoint operator in  $\lambda$ -calculus
  - ② formal fixpoint operator → constructed fixpoint operator
- $\lambda$ -calculus + explicit substitution  $t[x/u] \longrightarrow \lambda$ -calculus
  - **1** consider  $\lambda$ -calculus with its unary substitution operation
  - ② explicit substitution → real substitution

<sup>1&</sup>quot;A Theory of Explicit Substitutions with Safe and Full Composition", Kesner '09

# Example 1/2: compiling $\lambda$ -calculus + formal fixpoint op.

$$\underbrace{L_{\text{fix}}}_{\text{specified by "}\Sigma_{L} + \Sigma_{\text{fix}"}} \longrightarrow \underbrace{L}_{\text{specified by }\Sigma_{L}} \left( \lambda\text{-calculus} \right)$$

#### Signature $\Sigma_{fix}$ specifying a fixpoint operator

- an operation  $T' \xrightarrow{\text{fix}} T$
- reductions  $fix(t) \rightarrow t[x := fix(t)]$

the fresh variable

## Model structure on L for $\Sigma_{fix}$ ( $\Rightarrow$ compilation $L_{fix} \rightarrow L$ )

- choose a fixpoint combinator: a term Y s.t. Y  $u \rightarrow_{R}^{*} u$  (Y u)
- define fix(t) :=  $Y(\lambda x.t)$

$$\underbrace{Y(\lambda x.t)}_{\mathsf{fix}(t)} \to_{\beta}^{*} (\lambda x.t)(Y(\lambda x.t)) \to_{\beta} \underbrace{t[x := Y(\lambda x.t)]}_{t[x := \mathsf{fix}(t)]}$$

44/47

# Example 2/2: compiling $\lambda$ -calculus + explicit substitution

$$\underbrace{L_{\rm ex}}_{\rm specified \ by \ "\Sigma_L \setminus \{\beta\} + \Sigma_{\rm ex}"} \longrightarrow \underbrace{L}_{\rm specified \ by \ \Sigma_L} (\lambda \text{-calculus})$$

### Signature $\Sigma_{ex}$ for the explicit substitution

• an operation  $T' \times T \xrightarrow{(t,u) \mapsto t[x/u]} T$  s.t.

$$\boxed{t[x/u][y/v] = t[y/v][x/u]} \quad \text{if } x \notin fv(v), y \notin fv(u)$$

•  $\beta$ -reduction  $(\lambda x.t)u \to t[x/u] + \text{congruences} +$ 

$$t[x/u][y/v] \to t[y/v][x/u[y/v]] \quad x \notin f_V(v), \ y \in f_V(u) \quad (1)$$

## Model structure on L for $\Sigma_{ex}$ ( $\Rightarrow$ compilation $L_{ex} \rightarrow L$ )

- ullet use the real susbtitution  $T' \times T \xrightarrow{(t,u) \mapsto t[x:=u]} T$
- $\beta$ -reduction + congruences + reflexive reduction (1)

45/47

## Outline

- Transition monads
  - Substitution
  - Graphs
  - Definition
- 2 Generating transition monads
  - Three-level specification
  - Examples
- Generating compilations by initiality
- 4 Conclusion

## Conclusion

#### Summary

- PLs as transition monads
- Compilation as transition monad morphisms
- Signatures for transition monads

#### Perspectives

- Develop the metatheory (e.g., congruence of bisimilarity)
- Extensions to linear types, subtyping, polymorphism, ...
- $\pi$ -calculus as a transition system?
- signature for computational  $\lambda$ -calculus?

# Example: computational λ-calculus<sup>1</sup>

#### Parameterized by:

- ullet a set  $\Sigma$  of operation symbols  $\sigma$  with specified arities
- a monad T with operations  $T \times \cdots \times T \xrightarrow{\sigma^T} T$ .

$$\begin{array}{lll} M,N & ::= & \operatorname{return} \ V \mid VW \mid M \ \operatorname{to} \ x.N \mid \sigma(M,\ldots,M); \\ V,W & ::= & x \mid \lambda x.M. \end{array}$$

 $\Rightarrow$  a monad  $L_v$  of **values** + a  $L_v$ -module of terms

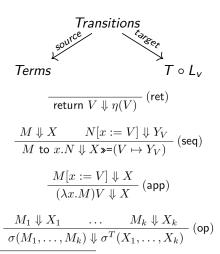


<sup>&</sup>lt;sup>1</sup> Effectful Applicative Bisimilarity: Monads, Relators, and Howe's Method, Dal Lago-Gavazzo-Levy LICS 2017

Transition monads Generating transition monads Generating compilations by initiality Conclusion

# Example: computational λ-calculus<sup>1</sup>

**Semantics** 



<sup>&</sup>lt;sup>1</sup> Effectful Applicative Bisimilarity: Monads, Relators, and Howe's Method, Lago-Gavazzo-Levy LICS 2017