Mathematical specification of programming languages using monads and modules over them

Ambroise Lafont¹

¹University of New South Wales Sydney, Australia

October 27, 2020

That is the question

What is a programming language, mathematically?

• In the literature, no well-established consensus.

Differential λ-calculus [Ehrhard-Regnier 2003]

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

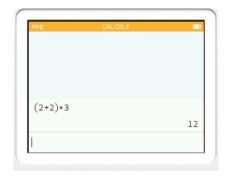
In this talk:

- a tentative notion of programming language, transition monads (FSCD 2020, with Tom and Andre Hirschowitz), and
- a discipline for automatically generating well-behaved transition monads.
- in the untyped case for ease of presentation (simply-typed case works as well)

What is a programming language?

2 components:

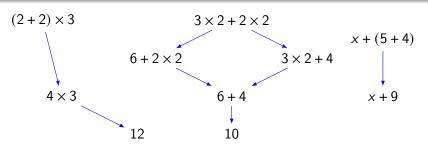
- Syntax: formal language for writing programs;
- Operational semantics: how do programs execute.



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

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!

Related work

- Mathematical Operational Semantics (Turi-Plotkin '97)
 - Deeply developed
 - \times Higher-order languages (such as λ -calculus) only starting to be investigated (Peressoti '17)
- Rewriting with variable binding (categorical approach)
 - e.g. Hamana 2003, T. Hirschowitz 2013, Ahrens 2016
 - x only congruent transitions ⇒ weak reduction out of reach

In this talk

- Mathematical definition of programming languages as transition monads.
- Signatures for specifying them
- Systematic use of monads and modules for taking care of substitution.

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

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

Ingredients

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

Example

 λ -calculus with β -reduction:

Syntax:

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

• Modulo α -equivalence, e.g.

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

• Reductions:

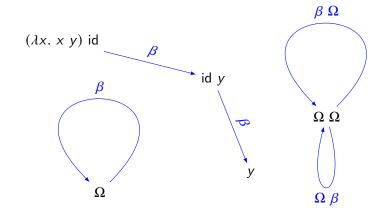
$$(\lambda x.t) u \xrightarrow{\beta} t[x := u] +$$

congruences

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

PLs as graphs

Example: λ -calculus with β -reduction



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

Graphs

Definition

Graph = a quadruple
$$(A, V, \sigma, \tau)$$
 where
$$A = \{\text{arrows}\} \qquad \sigma = \text{source of an arrow}$$

$$V = \{\text{vertices}\} \qquad \tau = \text{target of an arrow}$$

$$A \xrightarrow{\sigma} V$$

$$\sigma : \qquad A \qquad \rightarrow V \qquad \tau : \qquad A \qquad \rightarrow V$$

$$t \xrightarrow{r} u \qquad \mapsto t \qquad \qquad t \xrightarrow{r} u \qquad \mapsto u$$

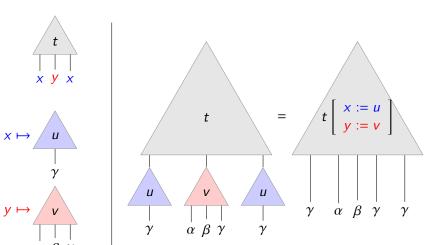
$$\sigma(r) \xrightarrow{r} \tau(r)$$

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

Simultaneous substitution

Syntax comes with substitution

terms (e.g. λ -terms) = trees with free variables as (distinguished) leaves.



Simultaneous substitution made formal

Free variables indexing

 $X \mapsto \{\text{terms taking free variables in } X\}$

Example: λ -calculus

Simultaneous substitution (bind)

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

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

15/4/

Simultaneous substitution

Monads model simultaneous substitution

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

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

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

Substitution laws:

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

+ associativity:

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

Substitution for semantics

Syntax supports substitution. This is also true of semantics.

Our notion of PL:

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

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

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

$$\frac{t \xrightarrow{r} u}{t[f] \xrightarrow{r[f]} u[f]}$$

• \Rightarrow R is a L-module, and σ , τ are module morphisms (see next slide)

Substitution for semantics made formal

R as a **module** over L

R supports L-monadic substitution:

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

$$R(X) \rightarrow R(Y)$$

 $r \mapsto r[x \mapsto f(x)]$ (or $r[f]$)

+ substitution laws

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

σ and τ as *L*-module morphisms

$$t \xrightarrow{r} u \ \, \sim \ \, t' \xrightarrow{r[f]} u' \quad \text{with} \quad \begin{cases} t' = t[f] \\ u' = u[f] \end{cases} \text{ i.e., } \begin{cases} \sigma(r[f]) = \sigma(r)[f] \\ \tau(r[f]) = \tau(r)[f] \end{cases}$$

Commutation with substitution \Leftrightarrow Module morphisms $\sigma, \tau : R \to L$.

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

Transition monads (first attempt)

Summary: graphs + substitution.

Definition

A transition monad $R \xrightarrow{\sigma} T$ consists of

- T = monad (= module over itself)
- R = module over T
- $\sigma, \tau : R \to T$ are T-module morphisms.

Example

 λ -calculus with β -reduction.

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

What about big-step cbv λ -calculus? Terms reduce to values, not terms!

Transition monads

Generalising cbv λ -calculus, and reduction monads

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

Examples: $\dfrac{\overline{\lambda}\mu\text{-calculus}}{\text{cbv }\lambda\text{-calculus}}$ $\dfrac{\pi\text{-calculus}}{\text{differential }\lambda\text{-calculus}}$ GSOS specs

¹POPL'20 with B.Ahrens, A. Hirschowitz, M. Maggesi.

Example: computational λ-calculus¹

Parameterized by:

- ullet a set Σ of operations σ 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

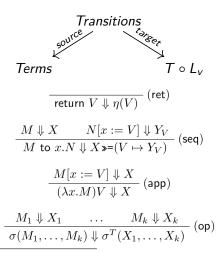


¹Effectful Applicative Bisimilarity: Monads, Relators, and Howe's Method, Lago-Gavazzo-Levy LICS 2017

Transition monads Generating transition monads Generating compilations by initiality Conclusion

Example: computational λ-calculus¹

Semantics



¹ Effectful Applicative Bisimilarity: Monads, Relators, and Howe's Method, Lago-Gavazzo-Levy LICS 2017

Morphisms of transition monads

Simple case $M_i = T$

Morphism
$$(T \leftarrow Trans \rightarrow T) \rightarrow (T' \leftarrow Trans' \rightarrow T') =$$
(Syntax) A monad morphism¹ $T \stackrel{c}{\rightarrow} T'$
(Semantics) Forward simulation²: if $t_1 \stackrel{r}{\rightarrow} t_2$, then $c(t_1) \stackrel{\llbracket r \rrbracket}{\longrightarrow} c(t_2)$

Examples (POPL'20, detailed later)

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

¹mapping preserving substitution and variables

²backward simulations are often considered as a correctness criteria

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

- Transition monads
 - Graphs
 - Substitution
 - Definition
- 2 Generating transition monads
 - Three-level specification
 - Examples
- 3 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

Specification through initial semantics

- Constructing syntax and reductions of a given PL may be complex (cf. differential λ-calculus).
- Often easier to describe the models.

Model \approx transition monad + interpretation of the operations and reductions

Initial Semantics

To each signature is associated

- ullet a notion of model = transition monad + additional structure
- a notion of morphism of models = compilation preserving additionnal structure
- a proof that the category of models has an initial object = object specified by the signature

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
3	Trans, source, target	"transition structure"	Transition rules as $\frac{t_1 \to u_1 \dots t_n \to u_n}{t \to u}$

⇒ Three notions of signatures.

- Transition monads
 - Graphs
 - Substitution
 - 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 λ -calculus	signature for M_i
7.	differential λ -calculus	signature for <i>T</i>

Example 1/7: small-step cbn λ -calculus

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

Signature for cbn λ -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 λ -terms

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

Syntax "generated" by

application	$T \times T \to T$	
λ -abstraction	$T' \rightarrow T$	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: 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, "Modular specification of monads" Ahrens et al. '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 λ -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":	$V \to T' \times T'$ (T-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 =
$$(T, M_1 \xleftarrow{source} Trans \xrightarrow{target} M_2)$$

cbv λ -calculus = $(Vals, Tms \xleftarrow{source} Trans \xrightarrow{target} Vals)$

Syntax of values and terms

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

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

$$\vdots = \underbrace{v | t u}_{t} | t u \qquad \Rightarrow \qquad Tms = BinTree \quad o \quad Vals$$

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 λ -calculus (DLC)

Syntax monad
$$T$$
 of terms (a variant of λ -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 λ -calculus: $(T, M_1 \xleftarrow{source} Trans \xrightarrow{target} M_2)$

• Syntax of DLC = variant of λ -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

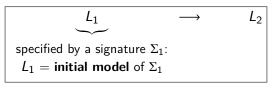
3 operations (no equation):

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

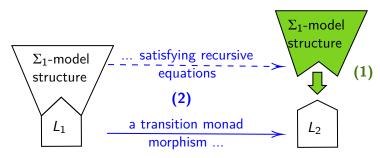
- Transition monads
 - Graphs
 - Substitution
 - Definition
- 2 Generating transition monads
 - Three-level specification
 - Examples
- Generating compilations by initiality
- 4 Conclusion

Generating compilations by initiality

Initiality ≈ recursion principle



Data generating a compilation: a Σ_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 Σ_1 :

Recipe:

- **1** provide a Σ_1 -model structure for L_2
- as a model morphism, the induced compilation satisfies recursive equations.

Upcoming examples (POPL'20)

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

¹A Theory of Explicit Substitutions with Safe and Full Composition, [Kesner 2009]

Example 1/2: compiling λ -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 Σ_{fix} specifying a fixpoint operator

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

the fresh variable

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

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

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

44/47

Example 2/2: compiling λ -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 Σ_{ex} for the explicit substitution

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

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

• β -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 Σ_{ex} (\Rightarrow compilation $L_{ex} \rightarrow L$)

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

45/47

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

Perspectives

- Generalise well-known theorems, e.g. Howe's method:
 - "A cellular Howe's theorem", LICS'20 with T. Hirschowitz and P. Borthelle, in a different setting.
 - "Effectful Applicative Bisimilarity: Monads, Relators, and Howe's Method", Lago-Gavazzo-Levy LICS 2017, for particular cases of transition monads (computational λ -calculus).
- Morphisms of transition monads = compilations
 - explore different variants (different correctness criteria).
 - try "academic" examples, e.g, Plotkin's CPS translations of \(\delta\)-calculus
 - "effective" Coq formalization (theory already formalized using UniMath for the syntax)