High-level signatures and initial semantics

Ambroise Lafont

joint work with Benedikt Ahrens, André Hirschowitz, Marco Maggesi
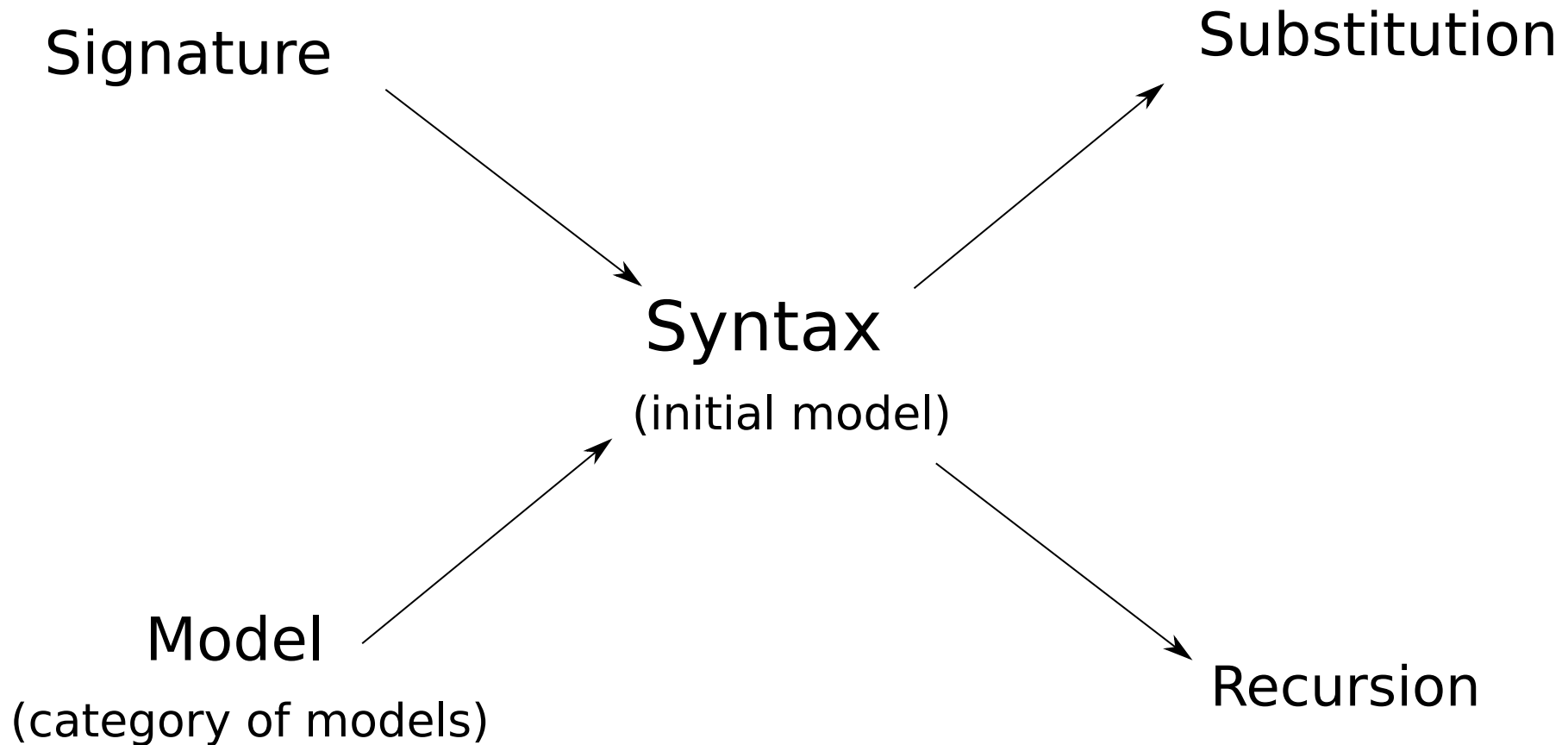
CSL 2018

Introduction

Purpose of our work: specify and construct untyped syntaxes with variables and a well-behaved substitution (e.g. lambda-calculus).

(We expect that our work straightforwardly generalizes to simply-typed syntaxes)

What is a syntax?

Signature → Syntax (initial model)

Syntax (initial model) → Substitution

Model (category of models) → Syntax (initial model)

Syntax (initial model) → Recursion

Signatures which we care about: those whose category of models have an *initial object*.

Our work

We present an alternative notion of signature (and associated models) based on the notion of module over a monad.

Goal of our work: Identify a large class of these signatures whose category of models have an initial object.

Table of contents

1. Standard signatures and their models

2. Languages, monads and modules

3. Recursion

4. Presentables signatures

Example: 0, ★

Consider the syntax generated by a binary operation ★ and a constant **0** (and variables):

$$
\begin{array}{lll}
\text{expr} ::= & \text{x} & \textit{(variable)} \\
& |\ t_1 ★ t_2 & \textit{(binary operation)} \\
& |\ 0 & \textit{(constant)}
\end{array}
$$

The syntax induces an endofunctor **B** (on Set) mapping a set of variables to the set of expressions built out of them.

$$B(\emptyset) = \{0, 0 \star 0, \dots\}$$
$$B(\{x, y\}) = \{0, 0 \star 0, \dots, x, y, x \star y, \dots\}$$

Example: 0, ★

The binary operation construction induces a natural transformation:

$$\mathbf{B} \times \mathbf{B} \to \mathbf{B}$$

The constant **0** induces a natural transformation:

$$\mathbf{1} \to \mathbf{B}$$

Variables induce a natural transformation

$$\mathbf{Id}_{Set} \to \mathbf{B}$$

Using disjoint union, they gather into a single natural transformation:

$$\mathbf{B} \times \mathbf{B} + \mathbf{1} + \mathbf{Id}_{Set} \to \mathbf{B}$$

i.e. **B** is an algebra for the endofunctor $\mathbf{F} \mapsto \mathbf{F} \times \mathbf{F} + \mathbf{1} + \mathbf{Id}_{Set}$ on the category $\mathbf{End}_{Set}$ of endofunctors on Set.

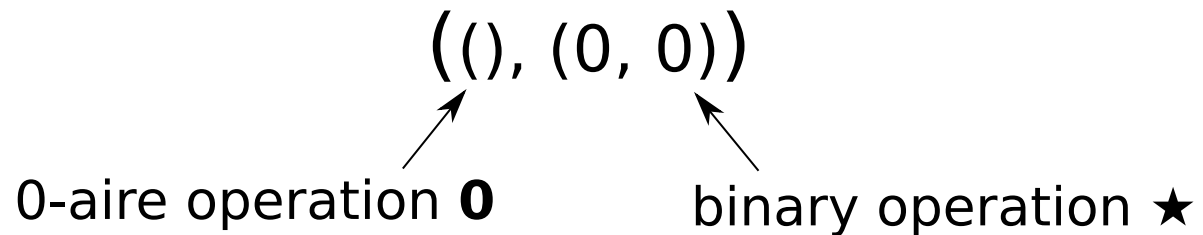Actually, **B** is the initial algebra.

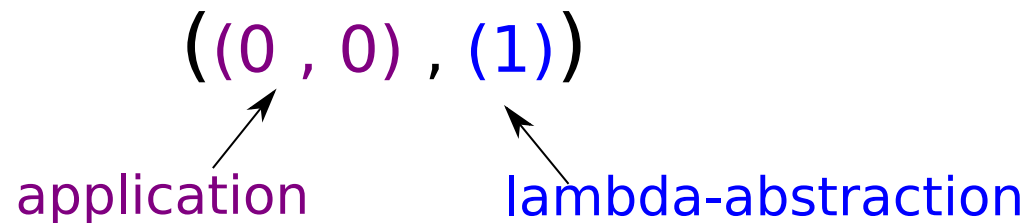**Binding signature** = a family of lists of natural numbers.

Each list specifies an operation in the syntax:

- length of the list = number of arguments of the operation

- natural number in the list = number of bound variables in the

  corresponding argument

Syntax with 0, ⋆:
$$\bigl((),\ (0,\ 0)\bigr)$$

0-aire operation **0**   binary operation ⋆

Lambda calculus:
$$\bigl((0\ ,\ 0)\ ,\ (1)\bigr)$$

application   lambda-abstraction

Models and signatures following [FPT]

In the same spirit as in the first example $(0,\star)$, any binding signature can be turned into an endofunctor Σ on the category $\mathbf{End}_{\mathrm{Set}}$.

A natural notion of model: $\Sigma + \mathbf{Id}_{\mathrm{Set}}$ -algebra

Theorem [FPT]: The initial $\Sigma + \mathbf{Id}_{\mathrm{Set}}$ -algebra of a binding signature Σ exists and comes with a *well-behaved substitution*.

Models and signatures following [FPT]

The endofunctor Σ induced by a binding signature comes with a *strength.*
It allows to define a new notion of model:

Σ-monoids = $\Sigma + \mathrm{Id}_{Set}$-algebras equipped with a well-behaved substitution.

Σ-monoid morphisms = algebra morphisms commuting with substitution.

Theorem [FPT]: Initial **Σ + Id**$_{Set}$-algebra morphisms commute with substitution (when the target algebra is a Σ-monoid).

In other words, the initial **Σ + Id**$_{Set}$-algebra is also initial in this category of models.

Any endofunctor on End$_{Set}$ *with strength* can be viewed as a signature (as in [Matthes-Uustalu 2004]) together with this notion of model.

Table of contents

Our signatures

We give an overview of our definition of signatures and models.

A **signature** Σ assigns functorially:

- to any endofunctor R with substitution: a "**monad**"

- an endofunctor **Σ**(**R**) with substitution: a "**module** over the input monad".

A **model** is a monad **R** with a natural transformation from **Σ**(**R**) to **R** compatible with the substitution (a "**module morphism**").

We will show how binding signatures yield signatures in our sense. Their category of models are (essentially) equivalent to the [FPT] ones.

Monads

A monad **R** corresponds to a language with variables as placeholders for any expression of **R**.

R(**X**) denotes the set of expressions taking variables in **X**.

Given any family $(\mathbf{t_x})_{\mathbf{x} \in \mathbf{X}}$ of elements of **R**(**Y**), variables of any expression **e** in **R**(**X**) can be substituted to yield an expression $\mathbf{e}[\mathbf{x} \mapsto \mathbf{t_x}]$ in **R**(**Y**).

The substitution is required to satisfy some intuitive equations.

A **monad morphism** between two monads **R** and **S** is a family of maps $(\mathbf{f}_X : \mathbf{R}(\mathbf{X}) \to \mathbf{S}(\mathbf{X}))_X$ preserving variables and substitution.

Operations as module morphisms

In the lambda-calculus,

$$\mathrm{app}(t, u)[x \mapsto v_x] = \mathrm{app}(t[x \mapsto v_x], u[x \mapsto v_x])$$

application commutes with substitution?

Yes: rewrite the right hand side as:

$$\mathrm{app}(t, u)[x \mapsto v_x] = \mathrm{app}((t, u)[x \mapsto v_x])$$

considering the obvious substitution on pairs of lambda terms.

We abstract this situation as follows:

- pairs of lambda-terms form a **module** L x L over the lambda-calculus monad L,

- application is a **module morphism** from L x L to L

Module over a monad

A module $\mathbf{M}$ over a monad $\mathbf{R}$ corresponds to expressions with variables as placeholders for any expression in the language $\mathbf{R}$.

Given a module $\mathbf{M}$, the set $\mathbf{M(X)}$ is the set of expressions taking variables in $\mathbf{X}$.

Given any family $(\mathbf{t_x})_{\mathbf{x} \in \mathbf{X}}$ of elements in $\mathbf{R(Y)}$, variables of any expression $\mathbf{e}$ in $\mathbf{M}(\mathbf{X})$ can be substituted to yield an expression $\mathbf{e}[\mathbf{x} \mapsto \mathbf{t_X}]$ in $\mathbf{M}(\mathbf{Y})$.

As for monads, the substitution is required to statisfy some intuitive equations.

Examples of modules

Modules over a monad:

Some examples of modules over a monad **R**:

- **R** itself

- **R** x **R** (i.e. pairs of expressions of **R**)

- **M** x **N** for any modules **M** and **N**

- **M'** is the module defined by $\mathbf{M'}(\mathbf{X}) = \mathbf{M}(\mathbf{X} + \{\mathbf{x}\})$ for any set **X** of variables given a module **M**. We call it the **derivative of M**.

  The new variable **x** is used to model an operation binding a variable (e.g. the lambda-abstraction).

Examples of module morphisms

A module morphism between two modules **M** and **N** on the same monad **R** is a family of maps $(\mathbf{f}_X : \mathbf{M}(\mathbf{X}) \rightarrow \mathbf{N}(\mathbf{X}))_X$ commuting with substitution.

Examples:

$id_M : M \rightarrow M$

the family of identity maps $(id_{M(X)} : M(X) \rightarrow M(X))_X$ for any module **M**

$app : L \times L \rightarrow L$

the application operation of the lambda calculus monad **L**.

$abs : L' \rightarrow L$

Indeed, in $\lambda x.t,$ the term t depends on an additional free variable x:

If $\lambda x.t \in L(Y)$, then $t \in L(Y + \{x\})$ **= L'(Y)**

Signatures

A **signature** Σ assigns (functorially) to each monad R a module Σ_R over it.

A **model** of a signature Σ is a monad R together with a morphism of modules $\sigma : \Sigma_R \to R$.

Models form a category (morphisms are monad morphisms compatible with σ).

Notion of signature too general: existence of the syntax (initial model) ?

- R ↦ R x R + 1

    By universal property of the disjoint sum, models are monads R

    equipped with module morphisms R x R → R and 1 → R. The syntax

    corresponds to our example with **0** and ★.

- R ↦ R x R + R'

    Models are monads R equipped with two modules morphisms R x R →

    R and R' → R. The syntax corresponds to lambda calculus.

Algebraic signatures

More generally, the syntax always exists for any signature induced by a disjoint sum of products of finite derivatives of the monad ($R \mapsto R' \times R'' \times R''' + R \times R'' \times R''' \times R + \ldots$).

We call such a signature an **algebraic signature**. They correspond to binding signatures.

Our main result: quotients of algebraic signatures also generate a syntax

Table of contents

Copie de Recursion

Question: cette section est-elle vraiment éncessaire, puisque on n'a pas vraiment de contribution là-dedans ? (ca fait partie du background initial semantics, non ?).

Benedikt suggère de ne pas faire plus d'une slide là dessus

Recursion

Initiality of the syntax allows recursion.

Example: computing the set of free variables of a lambda-term

Let **LC** be the monad of lambda-calculus.

Let **t** $\in$ **LC**(**X**) be a term (whose free variables are in **X**). We want to

compute its set of free variables **fv**(**t**) $\subset$ **X** (i.e. **fv**(**t**) $\in$ $\mathscr{P}$(**X**)).

Strategy:

The only thing to do is to give the assignement **X** $\mapsto$ $\mathscr{P}$(**X**) the

adequate structure of a monad, then of a model.

define **fv** : **LC** $\to$ $\mathscr{P}$ by initiality of LC in the category of

models of its signature.

Computing free variables

The assignement which to any set **X** associates its power set $\mathscr{P}(\mathbf{X})$ can be given the structure of a monad (variables are singletons, substitution is union).

$\mathbf{app}_{\mathscr{P}} : \mathscr{P}(\mathbf{X}) \times \mathscr{P}(\mathbf{X}) \rightarrow \mathscr{P}(\mathbf{X})$ and $\mathbf{abs}_{\mathscr{P}} : \mathscr{P}(\mathbf{X}+\{\mathbf{x}\}) \rightarrow \mathscr{P}(\mathbf{X})$ should be given to yield a model. Let us study the case of **app**:

expected equation of *fv*

$$fv(\mathrm{app}(t, u)) \quad = \quad \textcolor{blue}{fv(t)} \cup \textcolor{red}{fv(u)}$$

$$\|$$

$$\mathrm{app}_{\mathcal{P}}(\textcolor{blue}{fv(t)}, \textcolor{red}{fv(u)})$$

fv should be a model morphism

Computing free variables

The assignement which to any set **X** associates its power set $\mathscr{P}(\mathbf{X})$ can be given the structure of a monad (variables are singletons, substitution is union).

app$_{\mathscr{P}}$: $\mathscr{P}(\mathbf{X})$ x $\mathscr{P}(\mathbf{X}) \rightarrow \mathscr{P}(\mathbf{X})$ and **abs**$_{\mathscr{P}}$: $\mathscr{P}(\mathbf{X}+\{\mathbf{x}\}) \rightarrow \mathscr{P}(\mathbf{X})$ should be given to yield a model. Let us study the case of **app**:

expected equation of *fv*

$$fv(\mathrm{app}(t, u)) \quad = \quad \textcolor{blue}{fv(t)} \cup \textcolor{red}{fv(u)}$$

$$\| \quad$$

$$\mathrm{app}_{\mathcal{P}}(\textcolor{blue}{fv(t)}, \textcolor{red}{fv(u)})$$

fv should be a model morphism

Thus, we pose:

$$\mathrm{app}_{\mathcal{P}}(\textcolor{blue}{A}, \textcolor{red}{B}) := \textcolor{blue}{A} \cup \textcolor{red}{B}$$

Computing free variables

The case of **abs**$_{\mathscr{P}}$ is similar.

It can be shown that **app**$_{\mathscr{P}}$ and **abs**$_{\mathscr{P}}$ are module morphisms, hence give the monad $\mathscr{P}$ the structure of a model for the signature of the lambda-calculus.

By initiality of the syntax **LC**, we get a (unique) model morphism from **LC** to $\mathscr{P}$ which satisfies:

$$fv(t\,u) = fv(t) \cup fv(u)$$
$$fv(\lambda x.t) = fv(t)\backslash\{x\}$$
$$fv(x) = \{x\}$$

Table of contents

Quotient of a signature

Quotient of a set:

A quotient of a set X is a set Y together with a surjection p : X → Y.

$$x \sim x' \qquad \Longleftrightarrow \qquad p(x) = p(x')$$

Quotient of a signature:

A quotient of a signature **Σ** is a signature **Ψ** together with a (natural) family of module morphisms $(\mathbf{f_R} : \mathbf{\Sigma_R} \to \mathbf{\Psi_R})_\mathbf{R}$ that is pointwise surjective.

A **presentable signature** is a quotient of a binding signature.

Main Theorem: Any presentable signature generates a syntax.

Examples of presentable signatures

Presentable signatures allow to extend a syntax generated by an algebraic (or combinatorial) signature with new kinds of operations.

A binary commutative operation:

as a quotient of the signature of a binary operation $R \mapsto R \times R$ by the the action of the symmetry.

A syntactic closure operator:

Such an operator allows to bind a given set of variables in an expression (thus invariant under permutation of these variables).

The signature is obtained as a quotient of the algebraic signature specifying a sequence of increasingly sequential binding operators.

Examples of presentable signatures

Explicit substitution:

It is possible to specify an operation $\_\langle \mathbf{x_i} \mapsto \mathbf{t_i} \rangle$ that mimics the behavior of the true substitution $\_[\mathbf{x_i} \mapsto \mathbf{t_i}]$ in the sense that it enjoys some of its coherences, for example:

- if **u** does not depend on **y**,

$$u\langle x \mapsto v, y \mapsto w \rangle = u\langle x \mapsto v \rangle$$

- let **u'** be **u** where the variables **x** and **y** have been swapped,

$$u'\langle x \mapsto v, y \mapsto w \rangle = u\langle x \mapsto w, y \mapsto v \rangle$$

Examples of presentable signatures

A coherent fixedpoint operator:

A language with (mutual) fixedpoints comes with a construction

```
let rec f₁ = t₁
    and f₂ = t₂            where each fⱼ may appear as a
         ...               variable in each expression tᵢ.
    and fₙ = tₙ
in fᵢ
```

where each $\mathbf{f}_j$ may appear as a variable in each expression $\mathbf{t}_i$.

Thus, it takes **n** expressions $\mathbf{t}_1,..,\mathbf{t}_n$ depending on **n** new variables $\mathbf{f}_1,..,\mathbf{f}_n$ and produces an expression which does not depend on these variables.

As such, it can be specified by an algebraic signature.

Coherent fixedpoint operator

But we would like to encode some of the expected behaviour of such a fixed point. For instance:

$$\begin{array}{l}\texttt{let rec } \mathbf{f}_1 = \mathbf{t}_1 \\ \quad \texttt{and } \mathbf{f}_2 = \mathbf{t}_2 \\ \texttt{in } \mathbf{f}_1 \end{array} \quad = \quad \begin{array}{l}\texttt{let rec } \mathbf{f}_1 = \mathbf{t}_2' \\ \quad \texttt{and } \mathbf{f}_2 = \mathbf{t}_1' \\ \texttt{in } \mathbf{f}_1 \end{array}$$

($\mathbf{t}_i'$ is $\mathbf{t}_i$ where $\mathbf{f}_1$ and $\mathbf{f}_2$ have been swapped)

or, if $\mathbf{t}_1$ does not depend on $\mathbf{f}_2$,

$$\begin{array}{l}\texttt{let rec } \mathbf{f}_1 = \mathbf{t}_1 \\ \quad \texttt{and } \mathbf{f}_2 = \mathbf{t}_2 \\ \texttt{in } \mathbf{f}_1 \end{array} \quad = \quad \begin{array}{l}\texttt{let rec } \mathbf{f}_1 = \mathbf{t}_1 \\ \texttt{in } \mathbf{f}_1 \end{array}$$

A construction satisfying these equations can be specified by quotienting the naive algebraic signature.

Conclusion

We found a criterion for high-level 'monadic' signatures to specify a syntax. This criterion encompasses the classical binding signatures, and allows fancier operations at the level of the syntax.

Our main theorem have been formalized using the Coq library UniMath.

Future work:

We plan to take into account more sophisticated equations in the syntax than just quotients, extend our framework to simply typed syntaxes.

FIN PROVISOIRE

Ne pas lire les slides qui suivent (ce sont des anciennes slides que je garde au cas où).

FIN PROVISOIRE

Ne pas lire les slides qui suivent (ce sont des anciennes slides que je garde au cas où).

Copie de Models and signatures following [FTP]

For endofunctors induced by binding signatures, they define the
category of algebras equipped with a well-behaved substitution.

The syntax (as endofunctor) is still the same in this category of models:
the initial morphism from the syntax to such a model commutes with
substitution.

This notion of model works with any endofunctor with a strength, seen
as a general notion of signature.

By definition, the initial model (if it exists) comes with a well-behaved
substitution.

Copie de Models following [FPT]

In the same spirit as in the first example (0,1,+), any binding signature can be turned into an endofunctor **Σ** on the category **End**$_{Set}$.

A natural notion of model: **Σ + Id**$_{Set}$ -algebra

Theorem [FPT]: The initial model of a binding signature exists and comes with a *well-behaved substitution*.

Example : binary trees (leaves labelled by variables)

binary node (2 arguments)   unary node (1 argument)

Example : Binary Trees (leaves are labelled by variables)

binary node (2 arguments)   unary node (1 argument)

$$\mathbf{B} \times \mathbf{B} + \mathbf{B} + \mathbf{1}$$

application (2 arguments)   lambda-abstraction (1 argument binding 1 variable)

$$\mathbf{F} \mapsto \mathbf{F} \times \mathbf{F} + \mathbf{1} + \mathbf{1}$$

binary node (2 arguments)   unary node (1 argument)

$$(2, 1)$$

Copie de Examples of presentable signatures

Copie de Introduction

Copie de Purpose of our work

Copie de High-level signatures

Signatures with strength

Copie de First-order signatures

Copie de Example

First-order signatures

Signatures suggested by [FTP]

First-order signatures

Models following [FTP]

Examples of monads (à siupprimer ?)

- the syntax of arithmetic expressions

- the (untyped) syntax of lambda-calculus *L (modulo alpha equivalence)*

$$
\begin{aligned}
expr ::= \ &x && \textit{(variable)}\\
| \ &t \ u && \textit{(application)}\\
| \ &\lambda x.t && \textit{(abstraction)}
\end{aligned}
$$

- the (untyped) syntax of lambda-calculus modulo beta-equivalence and eta-equivalence

'High-level' VS classical signatures

+   Our 'high-level' signatures are more abstract and contrast with 'low-level' signatures which seem quite ad-hoc.

—   Our signatures, are too general: **we don't expect that all of them specify a language** (i.e. that the initial object always exist in the category of models associated to a signature)**.**

Goal of our work:

Identify a large class of (high-level) signatures which actually specify a language.

Copie de Languages as monads

A monad *A* as a language with variables:

- for each set X, a set A(X) of expressions taking free variables in X.

- any variable $x \in X$ is a valid expression that we note var$_X$(x) = $\underline{x} \in A(X)$

- given a family $(t_x)_{x \in X}$ of expressions in *A(Y)*, we can perform for

any expression **e** in **A(X)** the substitution e[x ↦ t$_x$] lying in A(Y)

Three monadic laws:

COMPOSITION OF SUBSTITUTIONS   *e[x ↦ t$_x$][y ↦ u$_y$] = e[x ↦ t$_x$[y ↦ u$_y$]]*

IDENTITY SUBSTITUTION $$e[x \mapsto x] = e$$

VARIABLE SUBSTITUTION $$\forall x \in X \quad x[y \mapsto t_y] = t_x$$

Copie de Overview of the methodology

1. Introduce a notion of signature.

2. Construct an associated notion of model (suitable as domain of interpretation of the syntax generated by the signature). Such models form a category.

3. Define the syntax generated by a signature as its initial model, when it exists.

4. Identify a class of signatures that generate a syntax: **presentable signatures**

:

For each set X, the sum of two expressions *e,e'* ∈ *A(X)* take free variables in *X*:

$$\forall X,\ add_X : A(X) \times A(X) \rightarrow A(X)$$
$$(e, e') \mapsto e + e'$$

Note that *(commutation with substitution)*:

$$(e + e')[x \mapsto t_x] = e[x \mapsto t_x] + e'[x \mapsto t_x]$$

We characterize this situation as follows:

A(X) x A(X) expressions are "*substitutable*" ⟿ A x A is a **module** on A

    add commutes with substitution ⟿ *add* is a **module morphism**

Examples of monads

- the assignement X $\mapsto$ $\mathscr{P}$(X) = { U | U $\subset$ X } yields a monad $\mathscr{P}$:

$$\forall X, \; var_X : X \to \mathcal{P}(X)$$
$$x \mapsto \{x\}$$

Let U $\subset$ X (i.e. U $\in$ $\mathscr{P}$(X)) and $(V_x)_{x \in x}$ a family of subsets of Y.

Substitution is defined as union:

$$U[x \mapsto V_x] = \bigcup_{x \in U} V_x \quad \in \mathcal{P}(Y)$$

Induction

Example: computing the free variables of a lambda-term

We compute it by induction on the syntax:

$$fv(x) = \{x\} \qquad\qquad\qquad \text{(variable)}$$
$$fv(tu) = fv(t) \cup fv(u) \qquad\qquad \text{(application)}$$
$$fv(\lambda x.t) = fv(t) \backslash \{x\} \qquad\qquad \text{(abstraction)}$$

This is formalized in our setting as a family of maps $(fv_X : L(X) \to \mathscr{P}(X))_X$ which *commutes with variable and substitution:*

$$fv(var_L(x)) = \{x\} \qquad\qquad fv(u[x \mapsto t_x]_L) = \bigcup_{y \in fv(u)} t_y$$
$$= var_{\mathcal{P}}(x) \qquad\qquad\qquad\qquad = fv(u)[x \mapsto fv(t_x)]_{\mathcal{P}}$$

(This is a definition of a monad morphism)

Induction

Example: computing the free variables of a lambda-term

fv also commutes with 'application' and 'abstraction'

$$app_{\mathcal{P}} : \mathcal{P} \times \mathcal{P} \to \mathcal{P}$$
$$(V, V') \mapsto V \cup V'$$

$$abs_{\mathcal{P}, X} : \overbrace{\mathcal{P}'(X)}^{\mathcal{P}(X + \{n\})} \to \mathcal{P}$$
$$V \mapsto V \backslash \{n\}$$

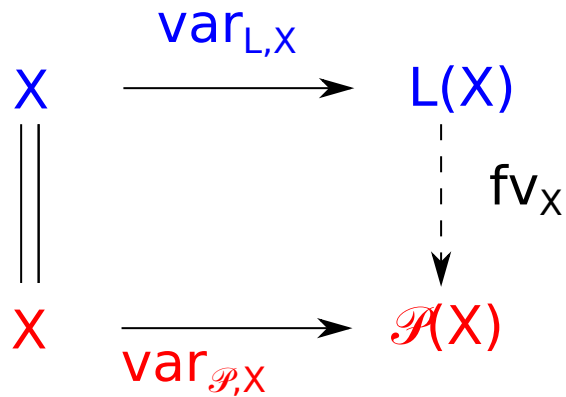Actually, these commutations **define** *fv* uniquely by induction:

$$fv(x) = \{x\} \qquad \text{(commutation with variable)}$$
$$fv(tu) = fv(t) \cup fv(u) \qquad \text{(commutation with application)}$$
$$fv(\lambda x.t) = fv(t) \backslash \{x\} \qquad \text{(commutation with abstraction)}$$

fv is the unique family of maps that makes the following diagrams commute:
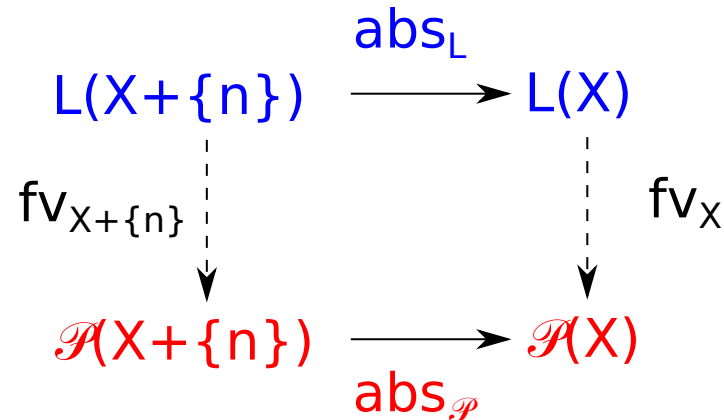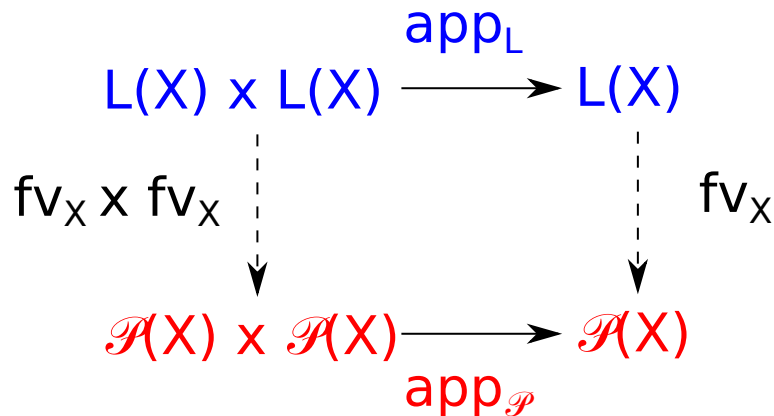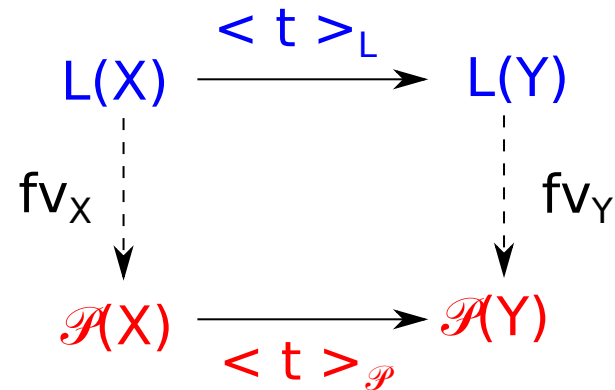
$$X \xrightarrow{\text{var}_{L,X}} L(X)$$

$$X \xrightarrow{\text{var}_{\mathscr{P},X}} \mathscr{P}(X)$$

with fv_X on the right vertical map.

$$L(X) \xrightarrow{< t >_L} L(Y)$$

$$\mathscr{P}(X) \xrightarrow{< t >_{\mathscr{P}}} \mathscr{P}(Y)$$

with fv_X and fv_Y vertical maps.

$$L(X) \times L(X) \xrightarrow{\text{app}_L} L(X)$$

$$\mathscr{P}(X) \times \mathscr{P}(X) \xrightarrow{\text{app}_{\mathscr{P}}} \mathscr{P}(X)$$

with $fv_X \times fv_X$ and fv_X vertical maps.

$$L(X+\{n\}) \xrightarrow{\text{abs}_L} L(X)$$

$$\mathscr{P}(X+\{n\}) \xrightarrow{\text{abs}_{\mathscr{P}}} \mathscr{P}(X)$$

with $fv_{X+\{n\}}$ and fv_X vertical maps.

Induction and initiality

More generally, let R be a monad with application and abstraction.

$$X \xrightarrow{\text{var}_{R,X}} R(X)$$

$$R(X) \xrightarrow{<t>_R} R(Y)$$

$$R(X) \times R(X) \xrightarrow{\text{app}_R} R(X)$$

$$R(X+\{n\}) \xrightarrow{\text{abs}_R} R(X)$$

Induction and initiality

More generally, let R be a monad with application and abstraction.

Then there is a unique family $(f_X)_X$ of maps (defined by induction) that

makes the following diagrams commute:

More generally, let R be a monad with application and abstraction. Then there is a unique family $(\mathbf{f_X})_\mathbf{X}$ of maps (defined by induction) that makes the following diagrams commute:

In other words, the monad L is initial in the category of monads with application and abstraction
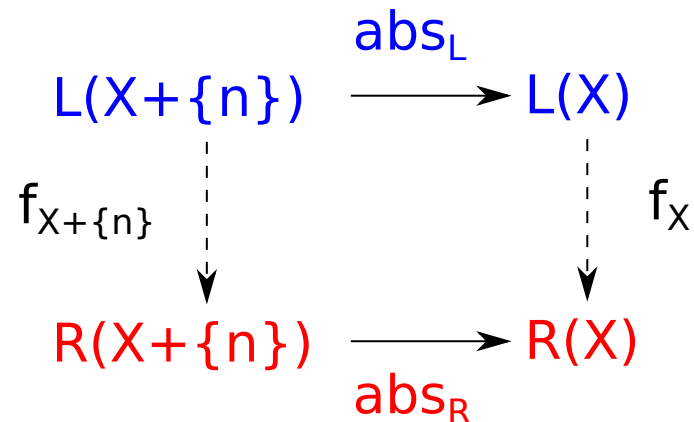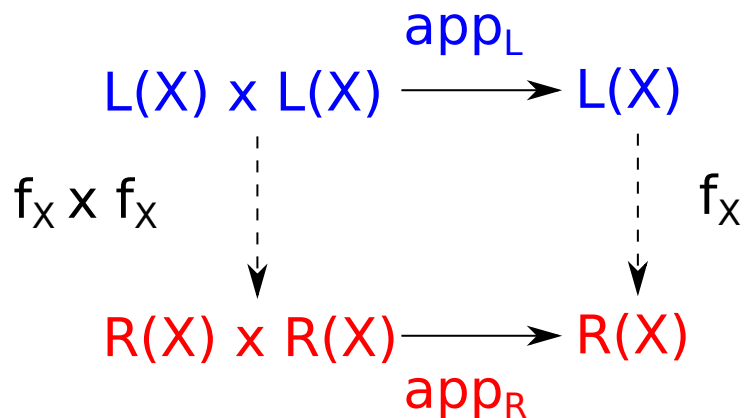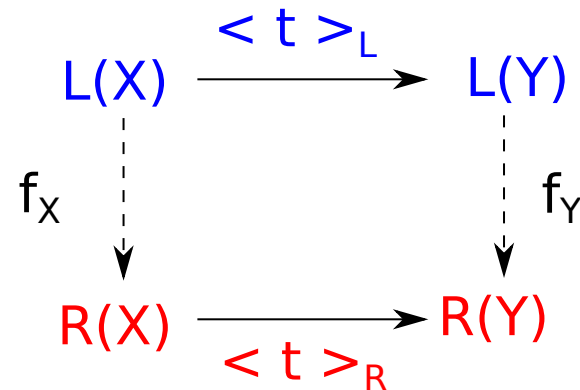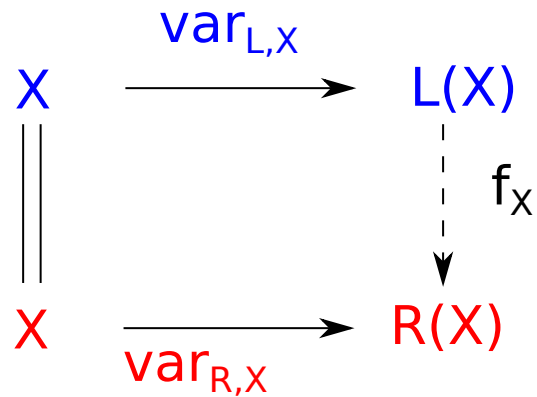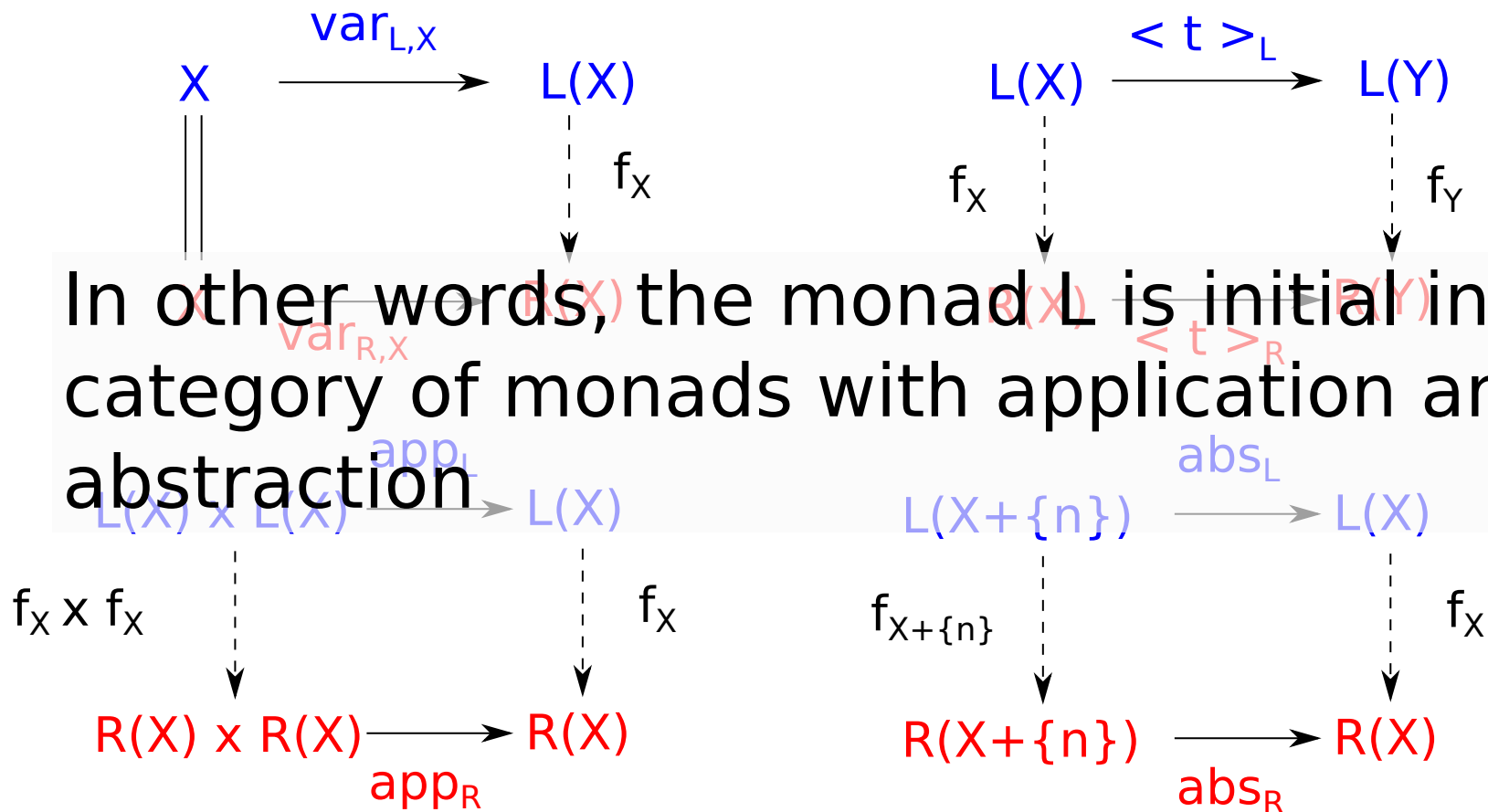
Syntax and initiality

A definition of a syntax:

A **syntax** is a monad that comes with an *induction principle, i.e.* which is initial in a suitable category of *monads + operations that it implements.*

Example:

The monad L of lambda calculus is initial in the category of *monads + application and abstraction*.

We say that L is the **syntax *generated* by the *signature* of application and abstraction.**

We will now present a general definition of **signatures**.

Signatures

What a signature should be:

L is initial among the monads R that model the signature ΣL of application and abstraction, i.e. monads R that come with module morphisms:

$$app_R : R \times R \to R$$
$$abs_R : R' \to R$$

or

$$[app_R, abs_R] : \underbrace{R \times R + R'}_{} \to R$$

ΣL(R)

A syntax *S* is initial among the monads R that model its associated signature Σ, i.e. monads R that come with a module morphism:

$$\sigma_R : \Sigma_R \to R$$

Thus, a signature Σ should assign to any monad R a module Σ_R over it.

Signatures

Let **R** be a monad that models the signature of application and abstraction. Then there exists a unique monad morphism **f** : **L** → **R** which commutes with abstraction and application:

$$\begin{array}{ccc} L(X) \times L(X) & \xrightarrow{\ app_L\ } & L(X) \\ {\scriptstyle f_X \times f_X}\Big\downarrow & & \Big\downarrow{\scriptstyle f_X} \\ R(X) \times R(X) & \xrightarrow{\ app_R\ } & R(X) \end{array}$$

$$f_X(app_L(t,u)) = app_R(f_X(t), f_X(u))$$

(and similarly for abs)

$$f_X(abs_L(t)) = abs_R(f_{X+\{n\}}(t))$$

Let **R** be a monad that models a signature **Σ** (there is a module morphism **σ**R : **Σ_R** → **R**). Then there exists a unique monad morphism **f** : **S** → **R** which commutes with **σ**:

$$\begin{array}{ccc} \Sigma_L(X) & \xrightarrow{\ \sigma_L\ } & L(X) \\ {\scriptstyle ??}\Big\downarrow & & \Big\downarrow{\scriptstyle f_X} \\ \Sigma_R(X) & \xrightarrow{\ \sigma_R\ } & R(X) \end{array}$$

Signatures

Let **R** be a monad that models the signature of application and abstraction. Then there exists a unique monad morphism **f** : **L** → **R** which commutes with abstraction and application:

$$L(X) \times L(X) \xrightarrow{\text{app}_L} L(X)$$

$$\mathbf{f_X \times f_X} \downarrow \qquad \qquad \downarrow f_X$$

$$R(X) \times R(X) \xrightarrow{\text{app}_R} R(X)$$

$$f_X(app_L(t, u)) = app_R(f_X(t), f_X(u))$$

(and similarly for abs)

$$f_X(abs_L(t)) = abs_R(f_{X+\{n\}}(t))$$

Let **R** be a monad that models a signature **Σ** (there is a module morphism **σ**R : **Σ_R** → **R**). Then there exists a unique monad morphism **f** : **S** → **R** which commutes with **σ**:

$$\Sigma_L(X) \xrightarrow{\sigma_L} L(X)$$

$$\mathbf{\Sigma(f)_X} \downarrow \qquad \qquad \downarrow f_X$$

$$\Sigma_R(X) \xrightarrow{\sigma_R} R(X)$$

Signatures

Thus, a signature Σ assigns to any monad morphism f : R → R'
a family of maps $(\Sigma(f)_X : \Sigma_R(X) \to \Sigma_{R'}(X))_X$.

As for module morphisms, we require that this family
commutes with substitution:

$$\Sigma(f)_Y(e[x \mapsto t_x]_{\Sigma_R}) = \Sigma(f)_X(e)[x \mapsto f_X(t_x)]_{\Sigma'_R}$$

PLAN

1. Languages, monads and modules

2. Induction and Initiality

3. Signatures

Definition of signatures

A **signature** Σ is given by:

- for each monad R, a module Σ_R over it

- for each monad morphism f : R → S, a family Σ(f) : Σ_R → Σ_S of morphisms which commutes with substitution:

$$\Sigma(f)_Y(e[x \mapsto t_x]_{\Sigma_R}) = \Sigma(f)_X(e)[x \mapsto f_X(t_x)]_{\Sigma'_R}$$

- such that (functoriality)

$$\Sigma(f \circ g) = \Sigma(f) \circ \Sigma(g) \qquad \text{and} \qquad \Sigma(id_R) = id_{\Sigma R}$$

A **model** of a signature Σ is a monad R together with a morphism of modules σ_R : Σ_R → R

A **model morphism** of a signature Σ between two models R and R' is a monad morphism f : R → S which commutes with σ: $\quad \sigma_R \circ f = \Sigma_f \circ \sigma_{R'}$

The **syntax generated by** a signature Σ is its initial model.

Syntax generated by a signature

This notion of signature is very general so that we do not expect that all of them generate a syntax.

Examples of syntax generating signatures:

- $R \mapsto R \times R$:

    models are monads R that comes with a module morphism $R \times R \to R$.

    The syntax corresponds to a language with variables and a binary

    operator b: expr ::= x   *(variable)*

                           | b(t, u)   *where t and u are any expressions*

- $R \mapsto R \times R + R'$:

    By universal property of the disjoint sum +, models are monads R

    equipped with two modules morphisms $R \times R \to R$ and $R' \to R$.

    The syntax corresponds to lambda calculus

Syntax generated by a signature

This notion of signature is very general so that we do not expect that all of them generate a syntax.

Examples of syntax generating signatures:

- $R \mapsto R \times R$:

    models are monads R that comes with a module morphism $R \times R \to R$.

    The syntax corresponds to a language with variables and a binary

    operator b: expr ::= x   *(variable)*

                                    | b(t, u)   *where t and u are any expressions*

- $R \mapsto R \times R + R'$:

    By universal property of the disjoint sum +, models are monads R

    equipped with two modules morphisms $R \times R \to R$ and $R' \to R$.

    The syntax corresponds to lambda calculus