High-level signatures and initial semantics

Ambroise Lafont

joint work with Benedikt Ahrens, André Hirschowitz, Marco Maggesi

CSL 2018

Overview

Topic: specification and construction of untyped syntaxes with variables and a well-behaved substitution (e.g. lambda calculus).

Our work:

1. general notion of ***signature*** based on ***monads*** and ***modules***.
   - *Caveat:* Not all of them do ***generate a syntax***
   - special case: classical **algebraic signatures**
2. our main result: any ***quotient*** of algebraic signatures generates a syntax.

This talk: explain the words in bold

Operations covered by our result

Some examples:

- Symmetric operations

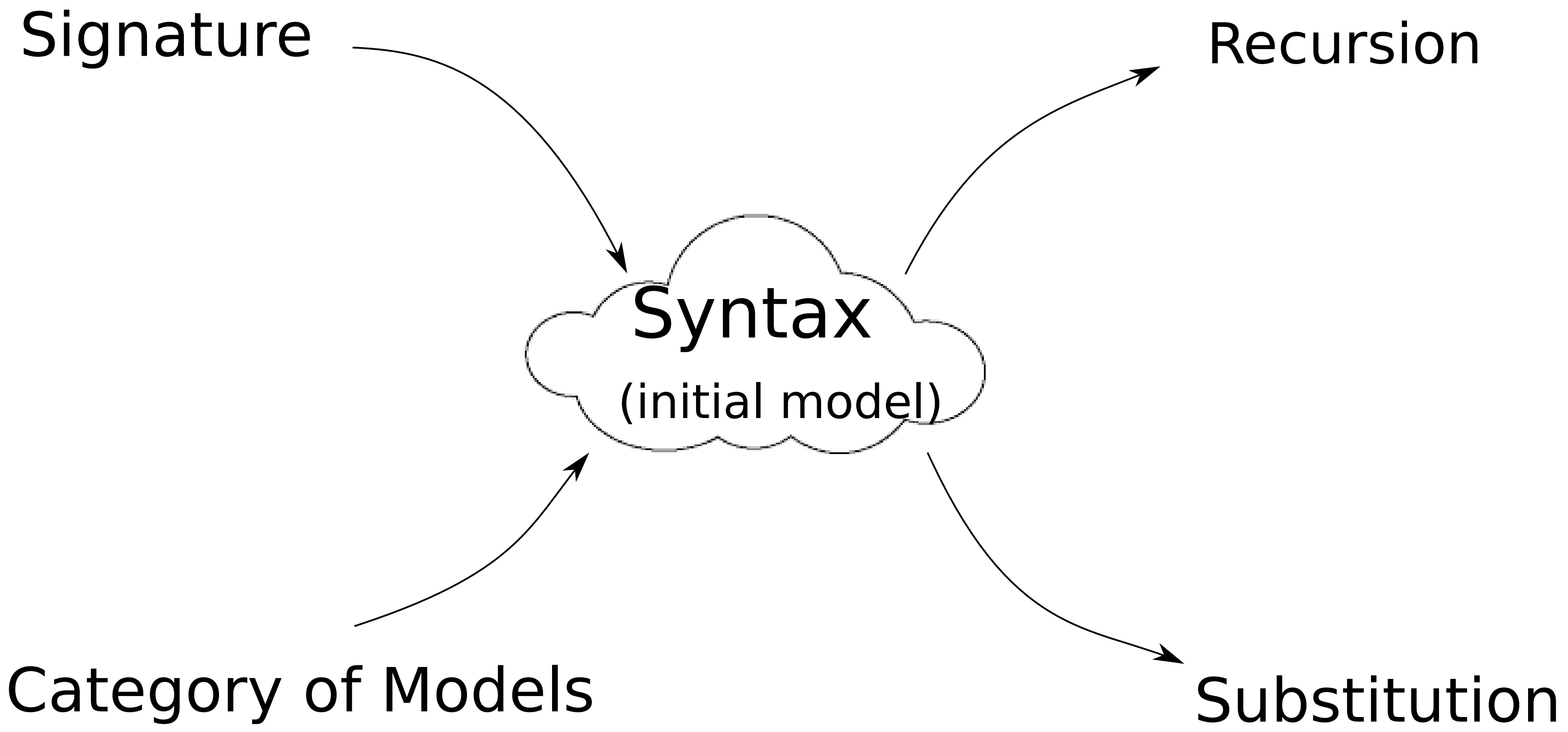
$$m : T \times T \to T \qquad \text{s.t.} \qquad m(t, u) = m(u, t)$$

- Explicit substitution with coherences

- Fixed point operation with coherences

- Syntactic closure operator with coherences

What is a syntax?



generates a syntax = existence of the initial model

Table of contents

1. **Binding signatures and their models**
    - Categorical formulation of term languages
    - Initial semantics for binding signatures
    - Categorical formulation of substitution
2. Signatures and models based on monads and modules
3. Our main result

Categorical formulation of a term language

Example: syntax with a binary operation, a constant, and variables

$$\begin{aligned} \mathrm{expr} ::= {} & x && \textit{(variable)} \\ & \mid t_1 \bigstar t_2 && \textit{(binary operation)} \\ & \mid 0 && \textit{(constant)} \end{aligned}$$

The syntax can be considered as the endofunctor B (on Set):

$$B : X \mapsto \{\text{expressions over } X\}$$

For example:

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

Categorical formulation of a term language

Then we have:

$$\begin{aligned} \star : \; & B \times B \xrightarrow{\bullet} B \\ 0 : \; & \quad 1 \quad \xrightarrow{\bullet} B \\ \mathrm{var} : \; & \quad \mathrm{Id}_{\mathrm{Set}} \xrightarrow{\bullet} B \end{aligned}$$

Putting all together:

$$B \times B + 1 + \mathrm{Id}_{\mathrm{Set}} \xrightarrow{\bullet} B$$

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

Actually, B can be **defined** to be the initial algebra.

Binding Signatures

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

Each list specifies one 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, ★:

$$\Big(\underset{0}{()}, \underset{\star}{(0, 0)}\Big)$$

Lambda calculus:

$$\Big(\underset{app}{(0\ ,\ 0)}\ ,\ \underset{abs}{(1)}\Big)$$

Initial semantics for binding signatures

Reminder

The syntax $(0,\star)$ is the initial algebra for the endofunctor:

$$F \mapsto F \times F + 1 + \mathrm{Id}_{\mathrm{Set}}$$

More generally, any binding signature gives rise to an endofunctor Σ.

Definition

Model = $(\Sigma + \mathrm{Id}_{\mathrm{Set}})$-algebra

Classical Theorem

The initial $(\Sigma + \mathrm{Id}_{\mathrm{Set}})$-algebra of a binding signature Σ always exists.

Does this initial algebra come with a well-behaved substitution?

Classical results on initial semantics

[Fiore-Plotkin-Turi 1999]: Initial semantics for binding signatures

[Ghani-Uustalu 2003]: Syntax for endufonctors with strength

Table of contents

1. Binding signatures and their models

2. Signatures and models based on monads and modules

- Our categorical formulation of substitution
- Our take on signatures, models and syntax
- Our take on binding signatures

3. Our main result

The Big Picture of signatures and models

Binding signatures $\hookrightarrow$ Endofunctors with strength $\hookrightarrow$ Our signatures

A **signature** Σ is a functorial assignment:

$$R \mapsto \Sigma(R)$$

A **model of $\boldsymbol{\Sigma}$** is a pair:

$$(R, \quad \rho : \Sigma(R) \to R)$$

monad	:=	endofunctor with substitution
module over a monad	:=	endofunctor with substitution
module morphism	:=	natural transformation preserving substitution

The Big Picture of signatures and models

Binding signatures $\hookrightarrow$ Endofunctors with strength $\hookrightarrow$ Our signatures

A **signature** Σ is a functorial assignment:

$$R \mapsto \Sigma(R)$$

monad

A **model of Σ** is a pair:

$$(R, \quad \rho : \Sigma(R) \to R)$$

monad	:=	endofunctor with substitution
module over a monad	:=	endofunctor with substitution
module morphism	:=	natural transformation preserving substitution

The Big Picture of signatures and models

Binding signatures $\hookrightarrow$ Endofunctors with strength $\hookrightarrow$ Our signatures

A **signature** Σ is a functorial assignment:

$$R \mapsto \Sigma(R)$$

monad

module over R

A **model of Σ** is a pair:

$$(R, \quad \rho : \Sigma(R) \to R)$$

monad	:=	endofunctor with substitution
module over a monad	:=	endofunctor with substitution
module morphism	:=	natural transformation preserving substitution

The Big Picture of signatures and models

Binding signatures $\hookrightarrow$ Endofunctors with strength $\hookrightarrow$ Our signatures

A **signature** Σ is a functorial assignment:

$$R \mapsto \Sigma(R)$$

monad

module over R

A **model of Σ** is a pair:

$$(R, \quad \rho : \Sigma(R) \to R)$$

monad

monad	:=	endofunctor with substitution
module over a monad	:=	endofunctor with substitution
module morphism	:=	natural transformation preserving substitution

The Big Picture of signatures and models

Binding signatures $\hookrightarrow$ Endofunctors with strength $\hookrightarrow$ Our signatures

A **signature** Σ is a functorial assignment:

$$R \mapsto \Sigma(R)$$

monad module over R

A **model of** $\boldsymbol{\Sigma}$ is a pair:

$$(R, \quad \rho : \Sigma(R) \to R)$$

monad module morphism

monad	:=	endofunctor with substitution
module over a monad	:=	endofunctor with substitution
module morphism	:=	natural transformation preserving substitution

Substitution and monads

Reminder:

- $B(X)$ = expressions built out of 0, ⋆ and variables taken in X
- Variables induce a natural transformation $\mathrm{var} : \mathrm{Id}_{\mathrm{Set}} \to B$

substitution:

$$\mathrm{bind} : B(X) \to (X \to B(Y)) \to B(Y)$$

+ laws

A triple (B, var, bind) is called a **monad**.

monad morphism = mapping preserving var and bind.

Preview: Operations are module morphisms

$\star$ commutes with substitution

$$(t \star u)[x \mapsto v_x] = t[x \mapsto v_x] \star u[x \mapsto v_x]$$

Categorical formulation

$B \times B$ supports B-substitution $\rightsquigarrow$ $B \times B$ is a **module over $\boldsymbol{B}$**

$\star$ commutes with substitution $\rightsquigarrow$ $\star : B \times B \to B$ is a **module morphism**

Building blocks for binding signatures

Essential constructions of **modules over a monad R**:

- R itself
- M x N for any modules M and N (in particular, R x R)
- The **derivative of a module M** is the module M' defined by $M'(X) = M(X + \{\bullet\})$.

  The derivative is used to model an operation binding a variable (Cf next slide).

Syntactic operations are module morphisms

module morphism = maps commuting with substitution.

$$id_M : M \to M$$

$$0 : 1 \to B$$

$$\star : B \times B \to B$$

$$app : \Lambda \times \Lambda \to \Lambda$$

$$abs : \Lambda' \to \Lambda$$

The Big Picture again

A **signature** Σ is a functorial assignment:

monad module over R

$$R \mapsto \Sigma(R)$$

A **model of Σ** is a pair:

$$(R, \quad \rho : \Sigma(R) \to R)$$

monad module morphism

A **model morphism** $m : (R, \rho) \to (S, \sigma)$ is a monad morphism commuting with the module morphism:

$$\begin{tikzcd}
\Sigma(R) \arrow[r, "\rho"] \arrow[d, "\Sigma(m)"'] & R \arrow[d, "m"] \\
\Sigma(S) \arrow[r, "\sigma"'] & S
\end{tikzcd}$$

Syntax

Definition

Given a signature Σ, its **syntax** is an initial object in its category of models.

Question: Does the syntax exist for every signature?

Answer: No.

Counter-example: the signature $R \mapsto \mathscr{P} \circ R$

powerset endofunctor on $\mathbb{S}\mathrm{et}$

Examples of signatures generating syntax

- **(0,★) language**:

  Signature: $R \mapsto 1 + R \times R$

  Model: $(R\ ,\quad 0 : 1 \to R\ ,\quad \bigstar : R \times R \to R)$

  Syntax: $(B\ ,\quad 0 : 1 \to B\ ,\quad \bigstar : B \times B \to B)$

- **lambda calculus**:

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

  Model: $(R\ ,\quad abs : R' \to R\ ,\quad app : R \times R \to R)$

  Syntax: $(\Lambda\ ,\quad abs : \Lambda' \to \Lambda\ ,\quad app : \Lambda \times \Lambda \to \Lambda)$

Can we generalize this pattern?

Initial semantics for algebraic signatures

Theorem

Syntax exists for any **algebraic signature**, i.e. signature built out of derivatives, products, and the trivial signature $R \mapsto R$.

Algebraic signatures correspond to binding signatures through the embedding:

$$\text{Binding signatures} \hookrightarrow \text{Our signatures}$$

Question: Can we identify a larger class of signatures generating a syntax?

Table of contents

1. Binding signatures and their models

2. Signatures and models based on monads and modules

3. Our main result

- Definition of presentable signatures
- Generated syntax for presentable signatures
- Examples of presentable signatures

Quotient of a signature

Quotient of a set:

A quotient of a set X consists of:

- a set Y
- a surjection function $f : X \to Y$

$$\begin{array}{c} X \\ \Big\downarrow f \\ Y \end{array}$$

Quotient of a signature:

A quotient of a signature Σ consists of:

- a signature Ψ
- a (natural) family of surjective module morphisms $(f_R : \Sigma(R) \to \Psi(R))_R$

$$R \mapsto \begin{array}{c} \Sigma(R) \\ \Big\downarrow f_R \\ \Psi(R) \end{array}$$

Syntax for presentable signatures

Definition

A **presentable signature** Ψ is a quotient of an algebraic signature Σ:

$$\begin{tikzcd} \Sigma \arrow[d, two heads] \\ \boldsymbol{\Psi} \end{tikzcd}$$

Theorem

Any presentable signature generates a syntax.

Question: Are there interesting examples of presentable signatures?

Answer:

- Symmetric operations
- Explicit substitution
- Coherent fixed point operation
- ...

Example 1: Symmetric operations

Binary commutative operation +:

$$t + u = u + t$$

As a quotient of an algebraic signature:

$$R \mapsto \begin{array}{c} R \times R \\ \big\downarrow\!\!\downarrow \\ R \times R \,/\, \{(x,y) \sim (y,x)\} \end{array}$$

This generalizes to **n-ary permutation invariant operations**.

Example 2: Explicit substitution

- an operation $\_\langle x_i \mapsto t_i \rangle$
- satisfying coherence equations:
  - invariance under **permutation**

    $$F(x, y)\langle x \mapsto t, y \mapsto u\rangle = F(y, x)\langle x \mapsto u, y \mapsto t\rangle$$

  - invariance under **weakening**

    $$F(x)\langle x \mapsto t, y \mapsto u\rangle = F(x)\langle x \mapsto u\rangle$$

  - invariance under **contraction**

    $$F(x, y)\langle x, y \mapsto t\rangle = F(x, x)\langle x \mapsto t\rangle$$

Example 2: Explicit substitution

$$\Sigma(R) := R' \times R \;+\; R'' \times R \times R \;+\; R''' \times R \times R \times R \;+\; \ldots$$

$t\langle x \mapsto u\rangle$ $\qquad$ $t\langle x \mapsto u, y \mapsto v\rangle$ $\qquad$ $t\langle x \mapsto u, y \mapsto v, z \mapsto w\rangle$

Signature of explicit substitution as a quotient of an algebraic signature:

$$R \mapsto \begin{tikzcd} \Sigma(R) \arrow[d, two heads] \\ \Sigma(R)\,/\!\sim \end{tikzcd}$$

- **permutation**: $t\langle x \mapsto u, y \mapsto v\rangle \sim t[x \rightleftarrows y]\langle x \mapsto v, y \mapsto u\rangle$
- **weakening**: $t\langle x \mapsto u\rangle \sim t\langle x \mapsto u, y \mapsto v\rangle$
- **contraction**: $t\langle x \mapsto u, y \mapsto u\rangle \sim t[y := x]\langle x \mapsto u\rangle$

Conclusion

Summary of the talk:

- presented a notion of signature and models
- identified a class of signatures that generate a syntax
  - encompasses the classical binding signatures
  - encompasses operations satisfying some equations

Future work:

- add equations (e.g. lambda calculus modulo beta/eta equivalence);
- extend our framework to simply typed syntaxes.

Conclusion

Summary of the talk:

- presented a notion of signature and models
- identified a class of signatures that generate a syntax
  - encompasses the classical binding signatures
  - encompasses operations satisfying some equations

Future work:

- add equations (e.g. lambda calculus modulo beta/eta equivalence);
- extend our framework to simply typed syntaxes.

Thank you!

Copie de Classical results on initial semantics

The endofunctor Σ induced by a binding signature comes with a *strength* which allows [FPT] to refine the notion of model:

***Σ*-monoid**:

$\Sigma+\mathrm{Id}_{\mathrm{Set}}$-algebra **equipped with a well-behaved substitution**.

***Σ*-monoid morphisms**:

algebra morphisms commuting with substitution.

Theorem [FPT]:

The initial $\Sigma + \mathrm{Id}_{\mathrm{Set}}$-algebra of a binding signature comes with a well-behaved substitution that makes it initial in the category of ***Σ*-monoids**.

This suggests defining signatures to be endofunctors on $\mathrm{End}_{\mathrm{Set}}$ *with strength* (as in [Matthes-Uustalu 2004]).