Initial semantics

B. Ahrens

Initial semantics

Methodology for defining/characterizing a language:

1. Introduce a notion of signature.
2. Construct an associated notion of model. Such models should form a category.
3. Define the **syntax generated by a signature** to be its initial model, when it exists.
4. Find a satisfactory sufficient condition for a signature to generate a syntax.

In this talk

- notions of signature for untyped languages
- notions of model of a signature
- some sufficient conditions for signatures to generate a syntax

Outline

How to specify the lambda calculus

1.

$$t, u \quad ::= \quad x \quad | \quad \mathsf{app}(t, u) \quad | \quad \lambda x.t$$

plus information about binding of variables

2.

$$\frac{x \in X}{X \vdash \mathsf{var}(x)} \qquad \frac{X \vdash t \qquad X \vdash u}{X \vdash \mathsf{app}(t, u)} \qquad \frac{X + 1 \vdash t}{X \vdash \mathsf{abs}(t)}$$

What kind of mathematical object is the lambda calculus?

Lambda calculus as a monad

1.
$$\begin{aligned} \mathsf{LC} : \mathsf{Set} &\to \mathsf{Set} \\ X &\mapsto \mathsf{LC}(X) \end{aligned}$$

2.
$$\mathsf{var} : \mathsf{Id} \to \mathsf{LC}$$

3.
$$\mathsf{subst}_{X,Y} : \mathsf{LC}(X) \times (X \to \mathsf{LC}(Y)) \to \mathsf{LC}(Y)$$

4. monad laws satisfied

Application and abstraction

$$\mathsf{app} : \mathsf{LC} \times \mathsf{LC} \to \mathsf{LC}$$
$$\mathsf{abs} : \mathsf{LC} \circ \mathsf{option} \to \mathsf{LC}$$

- ✓ natural transformations
- ✗ monad morphisms
- ✓ morphism of modules over monad LC

Definition

Given monad R, a module M over R is a

1. functor $M : \mathsf{Set} \to \mathsf{Set}$
2. substitution $\mathrm{subst}_{X,Y} : M(X) \times (X \to RY) \to M(Y)$

Module morphisms

Definition

Given modules M, N over monad R, a morphism from M to N is

- a nat. transformation $\tau : M \to N$
- for any $f : X \to R(Y)$,

$$\begin{array}{ccc} M(X) & \xrightarrow{\tau} & N(X) \\ {\scriptstyle \mathsf{subst}^M(f)} \downarrow & & \downarrow {\scriptstyle \mathsf{subst}^N(f)} \\ M(Y) & \xrightarrow{\tau} & N(Y) \end{array}$$

$\rightsquigarrow$ category $\mathrm{Mod}(R)$ of modules over monad R

$$\mathsf{subst}(f)(\mathsf{app}(t,u)) = \mathsf{app}\big(\mathsf{subst}(f)(t), \mathsf{subst}(f)(u)\big)$$
$$\mathsf{subst}(f)(\mathsf{abs}(t)) = \mathsf{abs}\big(\mathsf{subst}(\uparrow f)(t)\big)$$

Summary of the lambda calculus

The lambda calculus is a triple $(\mathsf{LC}, \mathsf{app}, \mathsf{abs})$,

1. monad $\mathsf{LC} : \mathsf{Set} \to \mathsf{Set}$
2. module morphisms over monad LC,

$$\begin{aligned} \mathsf{app} &: \mathsf{LC} \times \mathsf{LC} \to \mathsf{LC} \\ \mathsf{abs} &: \mathsf{LC} \circ \mathsf{option} \to \mathsf{LC} \end{aligned}$$

Definition

A **model of** LC is given by a triple $(R, \mathsf{app}, \mathsf{abs})$

1. monad $R : \mathsf{Set} \to \mathsf{Set}$
2. module morphisms over monad R,

$$\begin{aligned} \mathsf{app} &: R \times R \to R \\ \mathsf{abs} &: R \circ \mathsf{option} \to R \end{aligned}$$

Category of models of LC

Definition

Given two models $(R, \mathsf{app}, \mathsf{abs})$ and $(S, \mathsf{app}, \mathsf{abs})$ of LC, a morphism of models is a monad morphism $f : R \to S$ commuting with app and abs.

Theorem (Hirschowitz & Maggesi)

$(\mathsf{LC}, \mathsf{app}, \mathsf{abs})$ *is initial in the category of models.*

Outline

What specifies the lambda calculus

Model of LC:

1. monad R
2. morphism of modules $(R \times R) + (R \circ \mathsf{option}) \longrightarrow R$

The data specific to the lambda calculus:

$$R \mapsto (R \times R) + (R \circ \mathsf{option})$$

Definition

Signature Σ is a section to π:

$$\begin{array}{c} \int_R \mathsf{Mod}(R) \\ \pi \Big\downarrow \Big\uparrow \Sigma \\ \mathsf{Mon} \end{array}$$

Models of a signature

Definition

A **model of** Σ is a pair of

1. a monad R
2. a morphism $\Sigma(R) \longrightarrow R$ of R-modules

$\rightsquigarrow$ category Mon^{Σ} of models of Σ

Definition

Call Σ **representable** if Mon^{Σ} has an initial object.

Examples of signatures

Hypotheses	On objects	Name
	$R \mapsto R$	Θ
F functor, Σ signature	$R \mapsto F \cdot \Sigma(R)$	$F \cdot \Sigma$
	$R \mapsto 1_R$	1
Σ, Ψ signatures	$R \mapsto \Sigma(R) \times \Psi(R)$	$\Sigma \times \Psi$
Σ, Ψ signatures	$R \mapsto \Sigma(R) + \Psi(R)$	$\Sigma + \Psi$
	$R \mapsto R' := R \circ \mathsf{option}$	Θ'
$n \in \mathbb{N}$	$R \mapsto R^{(n)}$	$\Theta^{(n)}$
$(a) = (a_1, \ldots, a_n) \in \mathbb{N}^n$	$R \mapsto R^{(a)} = R^{(a_1)} \times \ldots \times R^{(a_n)}$	$\Theta^{(a)}$

Definition (Signatures are called)

elementary of the form $\Theta^{(a)}$

algebraic coproduct of elementary, e.g., $\Sigma_{\mathsf{LC}} := \Theta \times \Theta + \Theta'$

Not all signatures are representable

Non-example

Let $\mathscr{P} : \mathsf{Set} \to \mathsf{Set}$ denote the powerset functor. The signature $\mathscr{P} \cdot \Theta$ associates, to any monad R, the module $\mathscr{P} \cdot R$ that sends a set X to the powerset $\mathscr{P}(RX)$ of RX.
This signature is not representable.

Goal

Identify sufficient conditions for signatures to be representable.

Algebraic signatures

Theorem (Hirschowitz & Maggesi)

Algebraic signatures are representable.

Earlier variants of this theorem with essentially the same notion of signature by Fiore, Plotkin & Turi, by Gabbay & Pitts, by Hofmann, each using a different notion of model.

Category of signatures

Definition

Given signatures Σ and Ψ, a **morphism** $\Sigma \to \Psi$ **of signatures** is a natural transformation that is the identity when postcomposed with $\int \mathsf{Mod} \to \mathsf{Mon}$.

$\rightsquigarrow$ category Sig of signatures

Proposition

Sig *is cocomplete.*

Modularity

Given a pushout diagram of representable signatures

$$\begin{array}{ccc} \Sigma_0 & \longrightarrow & \Sigma_1 \\ \downarrow & & \downarrow \\ \Sigma_2 & \longrightarrow & \Sigma \end{array}$$

Modularity

Given a pushout diagram of representable signatures

$$\begin{array}{ccc} \Sigma_0 & \longrightarrow & \Sigma_1 \\ \big\downarrow & & \big\downarrow \\ \Sigma_2 & \longrightarrow & \Sigma \end{array} \qquad\qquad \begin{array}{ccc} \hat{\Sigma}_0 & \longrightarrow & \hat{\Sigma}_1 \\ \big\downarrow & & \big\downarrow \\ \hat{\Sigma}_2 & \longrightarrow & \hat{\Sigma} \end{array}$$

the corresponding diagram of representations is again a pushout, in the category $\int_\Sigma \mathsf{Mon}^\Sigma$:

Definition

object is a triple (Σ, R, r) where Σ is a signature, R is a monad, and r is an action of Σ in R.

morphism from (Σ_1, R_1, r_1) to (Σ_2, R_2, r_2) is a pair (i, m) of a signature morphism $i : \Sigma_1 \longrightarrow \Sigma_2$ and a morphism m of Σ_1-models from (R_1, r_1) to $(R_2, i^*(r_2))$.

Signatures and their models

Modularity follows from the projection being a Grothendieck fibration

$$\begin{array}{c} \int_{\Sigma} \mathsf{Mon}^{\Sigma} \\ \big\downarrow \\ \mathsf{Sig} \end{array}$$

Outline

Goal

Goal

Integrate some “semantic” equalities into the syntax.

- For instance, a binary operator is semantically symmetric (e.g., addition, parallel-or,...).
- Instead of defining that a posteriori, integrate this symmetry in the syntax.

Can be expressed as quotients of signatures

Presentations of a signature

Definition

A **presentation of** Σ is an algebraic signature Ψ and an epimorphism

$$\begin{array}{c}\Psi \\ \downarrow\!\!\downarrow \\ \Sigma\end{array}$$

Theorem (Ahrens, Hirschowitz, Lafont, Maggesi)

Presentable signatures are representable.

Signature for a commutative binary operator

$$\begin{array}{c} \Theta \times \Theta \\ \big\downarrow\!\!\!\downarrow \\ \mathscr{S}_2 \cdot \Theta \end{array}$$

Model of $\mathscr{S}_2 \cdot \Theta$ is a pair $(R, m : R \times R \to R)$ such that $m_X(a, b) = m_X(b, a)$.

Example: explicit substitution

Consider p-ary substitution

$$\mathsf{subst}_p : R^{(p)} \times R^p \longrightarrow R$$

Calculi with **explicit substitution** allow delaying substitutions. If $u : I_p \longrightarrow I_q$ a function, we expect

$$\begin{array}{ccc}
R^{(p)} \times R^q & \xrightarrow{R^{(p)} \times R^u} & R^{(p)} \times R^p \\
{\scriptstyle R^{(u)} \times R^q}\big\downarrow & & \big\downarrow{\scriptstyle \mathsf{subst}_p} \\
R^{(q)} \times R^q & \xrightarrow[\mathsf{subst}_q]{} & R
\end{array}$$

Signature for a coherent family of explicit substitutions

$$\int^{p:\mathbb{N}} \Theta^{(\underline{p})} \times \Theta^{\overline{p}}$$

Outline

Goal

1. Develop an explicit notion of **equation on a signature**
2. Define a **2-signature** to be a pair of a (1-)signature and a family of equations over it
3. Define **models** and **representability** for a 2-signature
4. Identify sufficient criteria for a 2-signature to be representable

η and β for lambda calculus

$$\mathsf{abs}(\mathsf{app}(f,*)) = f$$
$$\mathsf{app}(\mathsf{abs}(f),a) = f[* := a]$$

What are f and a?
Consider

$$\begin{aligned} &\mathsf{LC} \to \mathsf{LC} \\ &\quad f \mapsto \mathsf{abs}(\mathsf{app}(f,*)) && \text{(lhs)} \\ &\quad f \mapsto f && \text{(rhs)} \end{aligned}$$

and

$$\begin{aligned} &\mathsf{LC}' \times \mathsf{LC} \to \mathsf{LC} \\ &\quad (f,a) \mapsto \mathsf{app}(\mathsf{abs}(f),a) && \text{(lhs)} \\ &\quad (f,a) \mapsto f[* := a] && \text{(rhs)} \end{aligned}$$

Equations over Σ_{LC}

We can abstract from LC:

$$(R, \mathsf{app}, \mathsf{abs}) \quad \mapsto \quad \begin{array}{l} R \to R \\ f \mapsto \mathsf{abs}(\mathsf{app}(f, *)) \\ f \mapsto f \end{array}$$

and

$$(R, \mathsf{app}, \mathsf{abs}) \quad \mapsto \quad \begin{array}{l} R \times R \to R \\ (f, a) \mapsto \mathsf{app}(\mathsf{abs}(f), a) \\ (f, a) \mapsto f[* := a] \end{array}$$

2-signatures

Definition (Equation over Σ)

A Σ-equation is a pair (e_1, e_2) of parallel morphisms in a suitable category.

Definition (2-signature)

A **2-signature** is a pair (Σ, E) where Σ is a (1-)signature and E is a family of equations over Σ.

Definition

A **model of** (Σ, E) is a model M of Σ such that, for any equation (e_1, e_2) of E, we have $e_1(M) = e_2(M)$.

$$\mathsf{Mon}^{(\Sigma,E)} \subset \mathsf{Mon}^{\Sigma}$$

An unsatisfiable equation

Let $\Sigma := \Theta$. The equation

$$(R, r) \quad \mapsto \quad \begin{array}{l} R \to R + R \\ x \mapsto \mathsf{inl}(x) \\ x \mapsto \mathsf{inr}(x) \end{array}$$

is never satsified.

Elementary equations

Definition

An equation is **elementary** if

1. the source is of the form $(R, r) \mapsto R^{(a_1)} \times \ldots \times R^{(a_n)}$
2. the target is of the form $(R, r) \mapsto R^{(a)}$

Theorem (A-H-L-M)

If Σ is representable, and E is a family of elementary Σ-equations, then (Σ, E) is representable.

Theorem (with axiom of choice)

Reflection

$$\mathsf{Mon}^{(\Sigma,E)} \underset{L}{\overset{R}{\rightleftarrows}} \mathsf{Mon}^{\Sigma} \quad (L \dashv R)$$

Example: fixpoint operator

Definition

A **unary fixpoint operator for a monad** R is a module morphism $f : R' \to R$ such that

$$\begin{array}{ccc} R' & \xrightarrow{\langle 1_{R'}, f \rangle} & R' \times R \\ & {\scriptstyle f} \searrow \quad \swarrow {\scriptstyle \mathsf{subst}_R} & \\ & R & \end{array}$$

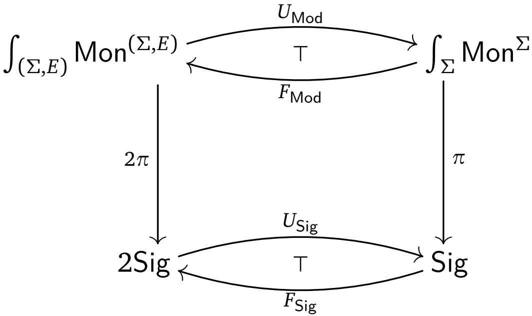
Signature Θ'

Model of signature $(R, \mathsf{fix} : R' \to R)$

Equation

$$(R, \mathsf{fix} : R' \to R) \quad \mapsto \quad \begin{array}{l} R' \xrightarrow{\langle 1, \mathsf{fix} \rangle} R' \times R \xrightarrow{\mathsf{subst}_R} R \\ R' \xrightarrow[\mathsf{fix}]{\qquad\qquad\qquad\qquad} R \end{array}$$

Modularity



A pushout diagram of representable 2-signatures yields a pushout diagram of representations “above”.

Outline

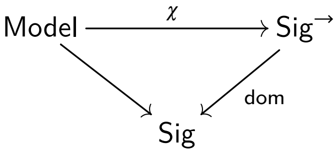
Goal I

Devise notions of signature and model for type theories with a Grothendieck fibration

$$\begin{array}{c} \mathsf{Model} \\ \big\downarrow \\ \mathsf{Sig} \end{array}$$

1. Pullback corresponds to restricting a model to a smaller signature (forgetting information)
2. **Pushforward** corresponds to extending a model to a larger signature
3. In particular, pushforward of empty model along $0 \to \Sigma$ exists iff initial model of Σ exists

Goal II



“Co-comprehension” category structure corresponds to “reflecting features of a model back into a signature”.