Categorical Semantics and Topos Theory
Homotopy type theory Seminar
University of Oxford, Michaelis 2011

November 16, 2011

References

- Johnstone, P.T.: *Sketches of an Elephant. A Topos-Theory Compendium.* Oxford University Press, 2002. (Main source)
- Caramello, O.: *Lectures on Topos Theory.* Cambridge, 2010. (Additional source)
- Bell, J.L.: *Notes on Toposes and Local Set Theories.* 2009. (Additional source)
- Moerdijk, I. and Mac Lane, S.: *Sheaves in Geometry and Logic: A First Introduction to Topos Theory.* Springer, 1992.
- Abramsky, S. and Tzevelekos, N.: *Introduction to Categories and Categorical Logic*. Oxford.

Introduction

Let's start with a simple example. The definition of a **group** can be realized in many different categories:

Sets	G any group
FinSets	G finite group
Top	G topological group
$\mathbf{Sh}(\mathbf{X})$, X topological space	G sheaf of groups.

We call the purely axiomatic definition of a group the **syntax** and the different realizations in categories **semantics**. Properties of groups can be proved on the syntactic level, and then also hold on the semantic level. **Categorical semantics** makes this relationship, which we often use, precise.

Note: All the categories **Sets**, **FinSets**, **Top**, $\mathbf{Sh}(\mathbf{X})$ are examples of *toposes*.

Aim: To give an idea the relation between syntax and semantics, and to advertise toposes as mathematical universes.

Logic and Lattices

Definition: A **preorder** is a category with at most one morphism for each pair of objects. A **partially ordered set (poset)** is a small preorder where the only isomorphisms are identity morphism. A **lattice** is a poset L where each pair $x, y \in L$ has a supremum **(join)** $x \vee y$ and an infimum **(meet)** $x \wedge y$.
Remark: Note that $\wedge$ and $\vee$ are product, respectively coproducts in L seen as a category. We say a lattice is **complete** if is is complete (and hence cocomplete) as a category, i.e. all small limits (and colimits) exist.
Here, this just means that each subset of L has an infimum and supremum. In particular,

$$\bigvee \emptyset = \bigwedge L =: 0, \text{ bottom (initial)}, \ \bigwedge \emptyset = \bigvee L =: 1, \text{ top (terminal)}$$

exist. Such a lattice is called **bounded**.

Logic and Lattices

Definition: A **Heyting algebra** is a bounded lattice H s.t. for all $x, y \in H$, there exists an element $x \Rightarrow y \in H$ s.t.

$$z \leq (x \Rightarrow y) \quad \Leftrightarrow \quad z \wedge x \leq y, \quad \forall z \in H.$$

This is the largest element z s.t. $z \wedge x \leq y$.
We write $\neg x := x \Rightarrow 0$, the **pseudocomplement** of x.
A **Boolean algebra** B is a complete latte s.t.

$$\begin{aligned} &\neg(\neg x) = x, \forall x \in B, \\ \Leftrightarrow\ &\neg x \vee x = 1, \forall x \in B. \end{aligned}$$

Proposition: H is a complete Heyting algebra if and only if H is a complete lattice and

$$x \wedge \bigvee_{i \in I} y_i = \bigvee_{i \in I} (x \wedge y_i), \quad \forall x, y_i \in H,$$

for any set I.

Logic and Lattices

As the notations suggest, these structures are closely related to logic, in particular **propositional logic**.

Classical propositional logic	$\leftrightarrow$	Boolean algebras
Intuitionistic propositional logic	$\leftrightarrow$	Heyting algebras
Free intuitionistic logic (with $\forall, \exists$)	$\leftrightarrow$	Complete Heyting algebras.

Classical logic allows the Law of Excluded Middle, while intuitionistic logic does not. By these correspondences, I mean certain *completeness*, more about this later.

Signatures and Theories

Brief introduction to how Logic formalizes mathematics. Back to our example, **groups**

Signature: One type: G; one constant (function) symbol (0-ary): 1; two function symbols $inv \colon G \to G$ (1-ary), and $\cdot \colon G \times G \to G$ (2-ary).

Axioms:

$$\top \vdash_x (x \cdot 1 = x \wedge 1 \cdot x = 1)$$
$$\top \vdash_x (x \cdot x^{-1} = 1 \wedge x^{-1} \cdot x = 1)$$
$$\top \vdash_{x,y,z} (x \cdot y) \cdot z = x \cdot (y \cdot z).$$

Signatures and Theories

More generally:

Definition: A **signature** Σ consists of

- **Types** (or **sorts**).
- **Function symbols**, each such f has a configuration $f : A_1 \times \ldots A_n \to B$, where $A_1, \ldots, A_n, B$ are types (f is n-ary). 0-ary function symbols are called **constants**.
- **Relation symbols**, each such R has a configuration $R \rightarrowtail A_1 \times \ldots \times A_n$, where $A_1, \ldots, A_n, B$ are types (R is n-ary). 0-ary relation symbols are called **propositions**.

Signatures and Theories

Given a signature Σ, we can recursively define **terms** over Σ

(I) $x : A$ means x is a variable of type A.

(II) $f(t_1, \ldots, t_n) : B$ if $f : A_1 \times \ldots \times A_n \to B$ is a function symbol, and $t_i : A_i$ are terms.

Signatures and Theories

From terms, we can build up **formulae** recursively by applying

- **Relations:** $R(t_1, \ldots, t_n)$ is a formula if $R \rightarrowtail A_1 \times \ldots \times A_n$ is a relation symbol, and $t_i \colon A_i$ are types. We assume that there is always the relation symbol **equality** $= \rightarrowtail A \times A$.
- **Truth and falsity:** Truth $\top$ and falsity $\bot$ are formulae without variables.
- **Binary constructs:** conjunction $\phi \wedge \psi$, disjunction $\phi \vee \psi$, and implication $\phi \Rightarrow \psi$ are formulae provided ϕ and ψ are.
- **Quantification:** $(\exists x)\phi$ and $(\forall x)\phi$ are formulae for x a variable and ϕ a formulae.
- Sometimes infinitary disjunction and conjunction are included.

First order logic does not allow infinitary conjunction and disjunction. If these occur, we are in **higher order logic**.

Definition: A **sequent** over Σ is an expression

$$\phi \vdash_{x_1,\dots,x_n} \psi,$$

where ϕ, ψ are formulae, and $x_1, \dots, x_n$ are variables containing at least all **free** variables in ϕ and ψ, i.e. those not being bound by a quantifier $\exists$ or $\forall$ (such a context is suitable for ϕ and ψ).
The interpretation of such a sequent is that for all $x_1, \dots, x_n$, $\phi(x_1, \dots, x_n)$ implies $\psi(x_1, \dots, x_n)$.
We can now define what a **theory** $\mathbb{T}$ over Σ is: it is just a collection of sequents, which we call the **axioms** of $\mathbb{T}$.
Examples: See above for groups. The theory of categories can be formulated using two sorts O and M, 2-categories with three sorts O, M_1, and M_2, etc.

Signatures and Theories

In logic, we are interested in proof rather than validity of certain statements. We want to know *what follows from what?*.
This is formulated using **formal** or **natural derivations**

$$\frac{\Gamma}{\sigma}$$

where Γ is a collection of sequents, and σ is a sequent. This is interpreted as Γ collectively implies σ.
There are, depending on what kind of logic we are working with, certain structural rules and axioms which are assumed to be given. Many of them may seem like trivialities, but have to be specified on such an elementary level. For example,

$$\frac{\phi \vdash \psi \qquad \chi \vdash \sigma}{\phi \vdash \psi} \text{ Cut rule,} \quad \text{or}$$

Signatures and Theories

$$\frac{\phi \vdash \psi \qquad \phi \vdash \chi}{\phi \vdash \psi \wedge \chi} \text{ Rule for } \wedge.$$

An important example of an axiom is the **Law of Excluded Middle**

$$\top \vdash_{x_1,\ldots,x_n} (\phi \vee \neg\phi),$$

which is not derivable in intuitionistic logic. It gives the basis for proofs by contradiction in classical logic.
Example for natural deduction: Implication is associative:

$$\cfrac{\cfrac{A \Rightarrow B, B \Rightarrow C, A \vdash B \Rightarrow C \qquad \cfrac{A \Rightarrow B, B \Rightarrow C, A \vdash A \Rightarrow B \qquad A \Rightarrow B, B \Rightarrow C, A \vdash A}{A \Rightarrow B, B \Rightarrow C, A \vdash B}\text{ Elimination}}{A \Rightarrow B, B \Rightarrow C, A \vdash C}\text{ Elimination}}{A \Rightarrow B, B \Rightarrow C \vdash A \Rightarrow C}\text{ Rule for } \Rightarrow$$

Toposes

Toposes are "almost as nice as sets". Let's see what makes **Set** special. For sets, we have the well-known adjunction

$$\mathrm{Hom}(Z \times X, Y) \cong \mathrm{Hom}(Z, Y^X),$$

where $Y^X = \{f\colon X \to Y\}$. That is, $(-) \times X \vdash (-)^X$ as functors. We want to generalize this idea.

Definition: Let $\mathcal{C}$ be a category with finite products. We say that $A \in \mathrm{Ob}\,\mathcal{C}$ is **exponentiable** if the functor $(-) \times A$ has a right adjoint, denoted by $(-)^A$.

A category $\mathcal{C}$ is called **cartesian closed** if it has all finite products and exponentials. **Cartesian functors** are those preserving these structures.

Remark: The counit $\epsilon_B \colon B^A \times A \to B$ gives an **evaluation map**, i.e.

$$\begin{tikzcd} C \times A \arrow[r, "h"] \arrow[d, "\exists! g \times \mathrm{Id}_A"'] & B \\ B^A \times A \arrow[ur, "\epsilon_B"'] & \end{tikzcd}$$

commutes for any such morphism h. In **Set**, $\epsilon_B(f, a) = f(a)$, and

$$g(c) = (a \mapsto h(c, a)).$$

Example: We have exponentials in the functor category $[\mathcal{O}(X)^{\mathrm{op}}, \mathbf{Set}] := \mathrm{Fun}(\mathcal{O}(X)^{\mathrm{op}}, \mathbf{Set})$ for a topological space X (this is called a **presheaf topos**). They can be found using the Yoneda lemma:

$$
\begin{aligned}
R^Q(C) &\overset{\text{Yoneda}}{\cong} \mathrm{Hom}_{[\mathcal{O}(X)^{\mathrm{op}}, \mathbf{Set}]}(yC, R^Q) \\
&\overset{\text{assuming exp. exist}}{\cong} \underbrace{\mathrm{Hom}_{[\mathcal{O}(X)^{\mathrm{op}}, \mathbf{Set}]}(yC \times Q, R)}_{\text{use as definition}}.
\end{aligned}
$$

for presheaves Q, R and an open set C (yC is the image of C under the Yoneda embedding).
If Q, R are sheaves, then so is R^Q, i.e. $\mathbf{Sh}(X)$ has the same exponentials as $[\mathcal{O}(X)^{\mathrm{op}}, \mathbf{Set}]$.

In **Set**, we can describe subset by *classifying arrows*, i.e for $S \subset X$ we have

$$\chi_S \colon X \to \Omega := \{0,1\}, \quad x \mapsto \begin{cases} 1 & \text{if } x \in S \\ 0 & \text{if } x \notin S. \end{cases}$$

This map contains all the information about S. Denote by

$$\text{true} \colon \{*\} = 1_{\mathbf{Set}} \to \Omega$$

the function $* \mapsto 1$. Then

$$\begin{tikzcd} S \arrow[r, "1"] \arrow[d, tail, "\iota"'] & \{*\} \arrow[d, tail, "\text{true}"] \\ X \arrow[r, "\chi_S"] & \Omega \end{tikzcd}$$

is a **pullback** square, $\Omega = \{0,1\}$ is the set of **truth values**.

Toposes

We can categorify this:
Definition: Let $\mathcal{C}$ be a category with all finite limits. A **subobject classifier** is an object Ω with a map $\mathrm{true}\colon 1_{\mathcal{C}} \to \Omega$ s.t. for each monomorphism $m\colon A' \rightarrowtail A$ there is a *unique* morphism $\chi_m\colon A \to \Omega$ (called **classifying arrow** of m) s.t.

$$\begin{tikzcd} A' \arrow[r, "1"] \arrow[d, tail, "m"'] & 1_{\mathcal{C}} \arrow[d, tail, "\mathrm{true}"] \\ A \arrow[r, "\chi_m"] & \Omega \end{tikzcd}$$

is a pullback square.

Remark: We denote by $\mathrm{Sub}_{\mathcal{C}}$ the lattice of subobjects of A. For a subobject, the classifying arrow is unique only up to isomorphism of morphisms. For a morphism $f : A \to B$, pullback induces a mapping $\mathrm{Sub}_{\mathcal{C}}\, B \to \mathrm{Sub}_{\mathcal{C}}\, A$. Thus we have a functor

$$\mathrm{Sub}_{\mathcal{C}} : \mathcal{C}^{\mathrm{op}} \to \mathbf{Set}$$

which is, provided $\mathcal{C}$ is locally small, represented by the pair (Ω, true).

Toposes

Definition (Lawvere-Tierney): A **topos** is category $\mathcal{C}$ with all finite limits, exponentials, and a subobject classifier.
Remark: This implies that $\mathcal{C}$ also has all finite colimits.
Examples: For any topological space X, the category of sheaves on X, $\mathbf{Sh}(X)$, is a topos. More generally, one can define $\mathbf{Sh}(\mathcal{C}, J)$ for any small category $\mathcal{C}$ and a **Grothendieck topology** J, this is a generalization of the notion of covering used to define sheaves. Such a topos is called a **Grothendieck topos**. It has the advantage that it is cocomplete and all colimits commute with finite limits, hence it is categorically "as good as" **Set**, which is also a topos.

Toposes

Toposes may be seen as **mathematical universes** (i.e. we can perform logic in them) because of the following fact:
Theorem: For each topos, the subobject sets $\mathrm{Sub}_{\mathcal{C}}\, A$ have the structure of a Heyting algebra (complete for Grothendieck toposes).
We call the logical operations in these subobject lattices the **external Heyting algebra** of the topos. The Yoneda lemma gives that these operations are described by a model structure on Ω, the so-called **internal Heyting algebra** of the topos. By this, we mean here that for each logical operation, we have maps from powers of Ω to Ω, e.g.

$$\wedge \colon \Omega \times \Omega \to \Omega,$$

induced by conjunction on the subobject lattices.

Toposes

Thus toposes give a way of "doing logic in them", more specifically

Toposes	$\longleftrightarrow$	First order logic
Grothendieck toposes	$\longleftrightarrow$	*Geometric* logic, having infinitary conjunction and disjunction.

Categorical Semantics

We want to interpret theories in categories. First, we look at realizations of signatures:

Definition: Let $\mathcal{C}$ be a category with finite limits and Σ a signature. A Σ-structure M in $\mathcal{C}$ consists of

- An object MA for each type A of Σ, s.t. $M(A_1 \times \ldots \times A_n) = MA_1 \times \ldots \times MA_n$.
- A morphism $Mf : MA_1 \times \ldots \times MA_n \to MB$ for each function symbol $f : A_1 \times \ldots \times A_n \to B$ of Σ.
- A subobject $MR \rightarrowtail MA_1 \times \ldots \times MA_n$ for each relation symbol $R \rightarrowtail A_1 \times \ldots \times A_n$ of Σ.

Categorical Semantics

A Σ-**structure homomorphism** $h\colon N \to M$ between two Σ-structures M and N is a collection of morphisms

$$h_A\colon MA \to NA, \quad \text{for each type } A \text{ in } \Sigma, \text{ s.t.}$$

- For any $f\colon A_1 \times \ldots \times A_n \to B$ in Σ,

$$\begin{tikzcd} MA_1 \times \ldots \times MA_n \arrow[r, "Mf"] \arrow[d, "h_{A_1} \times \ldots \times h_{A_n}"'] & MB \arrow[d, "h_B"] \\ NA_1 \times \ldots \times NA_n \arrow[r, "Nf"] & NB \end{tikzcd}$$

  commutes.

- For each relation symbol $R \rightarrowtail A_1 \times \ldots \times A_n$ in Σ,

$$\begin{tikzcd} MR \arrow[r, tail] \arrow[d, "\exists"'] & MA_1 \times \ldots \times MA_n \arrow[d, "h_{A_1} \times \ldots \times h_{A_n}"] \\ NR \arrow[r, tail] & NA_1 \times \ldots \times NA_n \end{tikzcd}$$

  commutes.

Any finite limit preserving functor $F\colon \mathcal{C} \to \mathcal{D}$ transfers Σ-structures in $\mathcal{C}$ into Σ-structures in $\mathcal{D}$.

★

Categorical Semantics

We can recursively define how to interpret **terms** in context $\{\vec{x}.t\}$, where $\vec{x} = x_1, \ldots, x_n$ is a list of variables containing those occurring in t, as morphisms

$$[[\vec{x}.t]]_M \colon MA_1 \times \ldots \times MA_n \to MB.$$

(I) If t is a variable of type A_i, then

$$[[\vec{x}.t]]_M = \pi_i \colon MA_1 \times \ldots \times MA_n \to MA_i$$

is the projection on the i-th component.

(II) If $t = f(t_1, \ldots, t_n)$, $t_i \colon C_i$, then

$$[[\vec{x}.t]]_M = Mf([[\vec{x}.t_1]]_M, \ldots, [[\vec{x}.t_1]]_M).$$

Categorical Semantics

To interpret **formulae** in $\mathcal{C}$, we need to assume that $\mathcal{C}$ has enough structure. Formulae in context $\{\vec{x}.\phi\}$ will be interpreted as subobjects

$$[[\vec{x}.\phi]]_M \rightarrowtail MA_1 \times \ldots \times MA_n \in \mathrm{Sub}_{\mathcal{C}}\, MA_1 \times \ldots \times MA_n.$$

These subobject lattice are all Heyting algebras in a topos. Thus, for a topos, we have enough structure to interpret $\wedge, \vee, \top, \bot, \exists, \forall$ of formulae in it. One just uses the corresponding operations in the subobject lattices.

Note that formulae build up using **relation symbols** are interpreted using the pullback, i.e. if $R \rightarrowtail B_1 \times \ldots \times B_m$ and $t_i \colon A_i$ are terms with a list of variable $\vec{x} = x_1, \ldots, x_n$, $x_i \colon A_i$, suitable for all of them, define $[[\vec{x}.R(t_1, \ldots, t_m)]]_M$ by the subobject in the pullback square

Categorical Semantics

$$\begin{tikzcd}
{[[\vec{x}.R(t_1,\ldots,t_m)]]_M} \arrow[r] \arrow[d, tail] & MR \arrow[d, tail] \\
MA_1 \times \ldots \times MA_n \arrow[r, "{([[\vec{x}.t_1]]_M,\ldots,[[\vec{x}.t_m]]_M)}"] & MB_1 \times \ldots \times MB_m.
\end{tikzcd}$$

We say that a sequent $\phi \vdash_{\vec{x}} \psi$ is **valid** in M if $[[\vec{x}.\phi]]_M \leq [[\vec{x}.\psi]]_M$ as subobjects. A **model** of a theory $\mathbb{T}$ is a Σ-structure M in which all axioms of $\mathbb{T}$ are satisfied.

★

Let X, Y be sets. For $S \subset X \times Y$, consider $p \colon X \times Y \to Y$ which induces a functor

$$p^* \colon \mathcal{P}(Y) \to \mathcal{P}(X \times Y), \quad T \mapsto p^{-1}(T) = X \times T.$$

Now $\forall_p S := \{y \in Y \mid \forall x \in X, (x, s) \in S\}$, and $\exists_p S := \{y \in Y \mid \exists x \in X, (x, y) \in S\}$ define functors $\exists_p, \forall_p \colon \mathcal{P}(X \times Y) \to \mathcal{P}(Y)$ satisfying the chain of adjunctions

$$\exists_p \vdash p^* \vdash \forall_p, \text{ i.e. } \begin{aligned} \exists_p S \subseteq T &\Leftrightarrow S \subseteq X \times T \\ X \times T \subseteq S &\Leftrightarrow T \subseteq \forall_p S. \end{aligned}$$

Categorical Semantics

More generally, in a topos $\mathcal{C}$ $f\colon A \to B$ induces

$$f^*\colon \operatorname{Sub}_{\mathcal{C}} B \to \operatorname{Sub}_{\mathcal{C}} A$$

preserving the lattice structures. This functor has adjoints

$$\exists_f \vdash f^* \vdash \forall_f.$$

These adjunctions give a way of interpreting the **quantifiers** $\exists$ and $\forall$ in $\mathcal{C}$.

Categorical Semantics

An important feature of categorical semantics is that it gives a way of arguing in a pointless category as if there where elements. In a signature, variables $x : A$ may be interpreted as "$x \in A$". Hence we can still interpret formulae using variables in categories in which the objects are no sets.

One can also proceed conversely. Given a category $\mathcal{C}$ one can define the **canonical signature** $\Sigma_{\mathcal{C}}$ of $\mathcal{C}$ as having a type $\ulcorner A \urcorner$ for each object, a function symbol $\ulcorner f \urcorner \colon \ulcorner A \urcorner \to \ulcorner B \urcorner$ for each morphism $f \colon A \to B$ in $\mathcal{C}$, and a relation symbol $\ulcorner R \urcorner \rightarrowtail \ulcorner A \urcorner$ for each subobject $R \rightarrowtail A$ in $\mathcal{C}$. In addition, all the compositional identities in $\mathcal{C}$ are formulated as axioms in $\Sigma_{\mathcal{C}}$. If $\mathcal{C}$ is a topos one can for example proof

$$f \text{ is mono in C} \quad \Leftrightarrow \quad f(x) = f(y) \vdash_{x,y} x = y \text{ is derivable in } \Sigma_{\mathcal{C}}.$$

Thus, $\Sigma_{\mathcal{C}}$ gives a way of arguing in a topos using elements.

Soundness and Completeness

Soundness is the property one naturally expects from a well-defined way of interpreting logic in a category.

Theorem: Let $\mathbb{T}$ be a first-order theory over Σ. If a sequent σ is derivable in $\mathbb{T}$, then σ is valid in any $\mathbb{T}$-model in any topos.

Proof (Sketch): One needs to verify that all the structural rules $\frac{\Gamma}{\omega}$ hold in any model M satisfying the axioms of $\mathbb{T}$ in the following way: If all the sequents of Γ hold in M, then so does ω. For most rules, this is trivial, e.g. for the *Cut rule*

$$\frac{\phi \vdash_{\vec{x}} \psi \qquad \chi \vdash_{\vec{x}} \lambda}{\phi \vdash_{\vec{x}} \psi}$$

this just means that $[[\vec{x}.\phi]]_M \leq [[\vec{x}.\psi]]_M$, and $[[\vec{x}.\chi]]_M \leq [[\vec{x}.\lambda]]_M$ implies $[[\vec{x}.\phi]]_M \leq [[\vec{x}.\psi]]_M$, i.e. $\phi \vdash_{\vec{x}} \psi$ holds in M.

Soundness and Completeness

One is also interested in the converse of the Soundness Theorem, the so-called **Completeness Theorem**. Such a theorem was first found for classical logic in **Set**:

Classical Completeness (Gödel, 1929): Let $\mathbb{T}$ be a coherent theory (this is basically a first order theory only using $\wedge, \vee, \top, \bot, \exists$ with some extra axioms) then a sequent σ is derivable in $\mathbb{T}$ using classical logic if and only if it is valid in each model of $\mathbb{T}$ in **Set**.

To prove completeness, one constructs a model $M_{\mathbb{T}}$ with the property that a sequent is valid in $M_{\mathbb{T}}$ if and only if it is derivable in $\mathbb{T}$.

Soundness and Completeness

Toposes satisfy completeness with respect to all first order theories: **Theorem (Completeness):** Let Σ be a signature. Then a first order formula $\{\vec{x}.\phi\}$ over Σ is provable in intuitionistic first order logic if it is valid in every Σ-structure in any elementary topos. The proof of this uses **Kripke-Joyal semantics**. Starting with a propositional theory $\mathbb{P}$, one can find a Kripke frame which is a Heyting algebra and contains all the information about validity of propositions in $\mathbb{P}$. For a more general intuitionistic theory $\mathbb{T}$, one obtains a collection of Kripke frames $\{U_p\}_{p \in P}$ related by a poset P. Then these Kripke models $\{U_p\}$ correspond to a model U^* in $[P, \mathbf{Set}]$ which is also a Kripke frame, and sequents are valid in all U_p if and only if they are valid in U^*.

This completeness justifies the view of toposes as *mathematical universes*.