November 16, 2011
References

- Bell, J.L.: *Notes on Toposes and Local Set Theories*. 2009. (Additional source)
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
- **Sh($X$)**, $X$ 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**, **Sh($X$)** 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) $x \lor y$ and an infimum (meet) $x \land y$.

**Remark:** Note that $\land$ and $\lor$ 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,

$$\bigvee \emptyset = \bigwedge L =: 0,$$

bottom (initial),

$$\bigwedge \emptyset = \bigvee L =: 1,$$

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.

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

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 lattice s.t.

$$\neg(\neg x) = x, \quad \forall x \in B,$$

$$\iff \neg x \lor x = 1, \quad \forall x \in B.$$

**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 $I$.  

Categorical Semantics and Topos Theory Homotopy type theory Seminar University of Oxford, Michaelis 2011
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.
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 $\text{inv}: G \to G$ (1-ary), and $\cdot: G \times G \to G$ (2-ary).

**Axioms:**

\[
\begin{align*}
\top & \vdash_x (x \cdot 1 = x \land 1 \cdot x = 1) \\
\top & \vdash_x (x \cdot x^{-1} = 1 \land x^{-1} \cdot x = 1) \\
\top & \vdash_{x,y,z} (x \cdot y) \cdot z = x \cdot (y \cdot z).
\end{align*}
\]
More generally:

**Definition:** A signature $\Sigma$ consists of

- **Types** (or **sorts**).
- **Function symbols**, each such $f$ has a configuration $f : A_1 \times \ldots \times A_n \to B$, 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 $R \ni A_1 \times \ldots \times A_n$, 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 $\Sigma$, we can recursively define terms over $\Sigma$:

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

- **Relations:** $R(t_1, \ldots, t_n)$ is a formula if $R \to A_1 \times \ldots \times A_n$ is a relation symbol, and $t_i : A_i$ are types. We assume that there is always the relation symbol equality $\equiv : A \times A$.

- **Truth and falsity:** Truth $\top$ and falsity $\bot$ are formulae without variables.

- **Binary constructs:** conjunction $\phi \land \psi$, disjunction $\phi \lor \psi$, and implication $\phi \Rightarrow \psi$ are formulae provided $\phi$ and $\psi$ are.

- **Quantification:** $(\exists x)\phi$ and $(\forall x)\phi$ are formulae for $x$ a variable and $\phi$ 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 $\Sigma$ is an expression

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

where $\phi, \psi$ are formulae, and $x_1, \ldots, x_n$ are variables containing at least all free variables in $\phi$ and $\psi$, i.e. those not being bound by a quantifier $\exists$ or $\forall$ (such a context is suitable for $\phi$ and $\psi$).

The interpretation of such a sequent is that for all $x_1, \ldots, x_n$, $\phi(x_1, \ldots, x_n)$ implies $\psi(x_1, \ldots, x_n)$.

We can now define what a theory $T$ over $\Sigma$ is: it is just a collection of sequents, which we call the axioms of $T$.

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.
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**

\[
\Gamma \\ \frac{}{\sigma}
\]

where \(\Gamma\) is a collection of sequents, and \(\sigma\) is a sequent. This is interpreted as \(\Gamma\) collectively implies \(\sigma\).

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,

\[
\phi \vdash \psi \\ \chi \vdash \sigma \\ \frac{}{\phi \vdash \psi} \text{ Cut rule, or}
\]
Signatures and Theories

\[
\frac{\phi \vdash \psi \quad \phi \vdash \chi}{\phi \vdash \psi \land \chi}
\]
Rule for \(\land\).

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

\[
\top \vdash x_1, \ldots, x_n (\phi \lor \neg \phi),
\]

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:

\[
\begin{align*}
\frac{A \Rightarrow B, B \Rightarrow C, A \vdash B \Rightarrow C}{A \Rightarrow B, B \Rightarrow C, A \vdash C} \quad & \text{Elimination} \\
\frac{A \Rightarrow B, B \Rightarrow C, A \vdash B}{A \Rightarrow B, B \Rightarrow C, A \vdash B \Rightarrow C} \quad & \text{Elimination} \\
\frac{A \Rightarrow B, B \Rightarrow C, A \vdash C}{A \Rightarrow B, B \Rightarrow C \vdash A \Rightarrow C} \quad & \text{Rule for \(\Rightarrow\)}
\end{align*}
\]
Toposes are “almost as nice as sets”. Let’s see what makes \( \textbf{Set} \) special. For sets, we have the well-known adjunction

\[
\text{Hom}(Z \times X, Y) \cong \text{Hom}(Z, Y^X),
\]

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

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

A category \( C \) is called **cartesian closed** if it has all finite products and exponentials. **Cartesian functors** are those preserving these structures.
Remark: The counit $\epsilon_B : B^A \times A \to B$ gives an evaluation map, i.e.

$$
\begin{array}{ccc}
C \times A & \xrightarrow{h} & B \\
\downarrow & & \downarrow \\
B^A \times A & \xrightarrow{\epsilon_B} & B
\end{array}
$$

exists $g \times \text{Id}_A$ commutes for any such morphism $h$. In $\textbf{Set}$, $\epsilon_B(f, a) = f(a)$, and

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

$$R^Q(C) \overset{\text{Yoneda}}{\cong} \text{Hom}_{[\mathcal{O}(X)^{\text{op}}, \text{Set}]}(yC, R^Q)$$

assuming exp. exist

$$\cong \underbrace{\text{Hom}_{[\mathcal{O}(X)^{\text{op}}, \text{Set}]}(yC \times Q, R)}_{\text{use as definition}}$$

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. $\text{Sh}(X)$ has the same exponentials as $[\mathcal{O}(X)^{\text{op}}, \text{Set}]$. 

---

**Toposes**

Categorical Semantics and Topos Theory Homotopy type theory Seminar University of Oxford, Michaelis 2011
In **Set**, we can describe subset by *classifying arrows*, i.e. for $S \subset X$ we have

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

This map contains all the information about $S$. Denote by

$$\text{true}: \{\ast\} = 1_{\text{Set}} \to \Omega$$

the function $\ast \mapsto 1$. Then

$$\begin{array}{ccc}
S & \rightarrow^1 & \{\ast\} \\
\downarrow & & \downarrow \\
X & \rightarrow^{\chi_S} & \Omega \\
\end{array}$$

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

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

\[
\begin{array}{ccc}
A' & \xrightarrow{1} & 1_{\mathcal{C}} \\
\downarrow{m} & & \downarrow{\text{true}} \\
A & \xrightarrow{\chi_m} & \Omega
\end{array}
\]

is a pullback square.
Remark: We denote by $\text{Sub}_C$ 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 $\text{Sub}_C B \to \text{Sub}_C A$. Thus we have a functor

$$\text{Sub}_C : C^{\text{op}} \to \text{Set}$$

which is, provided $C$ is locally small, represented by the pair $(\Omega, \text{true})$. 
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 $X$, $\text{Sh}(X)$, is a topos. More generally, one can define $\text{Sh}(\mathcal{C}, J)$ 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” $\text{Set}$, which is also a topos.
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 $\text{Sub}_C A$ 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 $\Omega$, 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 $\Omega$ to $\Omega$, e.g.

$$\land: \Omega \times \Omega \to \Omega,$$

induced by conjunction on the subobject lattices.
Thus toposes give a way of “doing logic in them”, more specifically

\[
\text{Toposes} \iff \text{First order logic}
\]

\[
\text{Grothendieck toposes} \iff \text{Geometric logic, having infinitary conjunction and disjunction.}
\]
We want to interpret theories in categories. First, we look at realizations of signatures:

**Definition:** Let $C$ be a category with finite limits and $\Sigma$ a signature. A $\Sigma$-structure $M$ in $C$ consists of

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

Categorical Semantics and Topos Theory
Homotopy type theory
Seminar
University of Oxford, Michaelis 2011
A \( \Sigma \)-structure homomorphism \( h : N \to M \) between two \( \Sigma \)-structures \( M \) and \( N \) is a collection of morphisms

\[
h_A : MA \to NA, \quad \text{for each type } A \text{ in } \Sigma, \text{ s.t.}
\]

\[
\begin{align*}
&M_{A_1} \times \ldots \times MA_n \\
\xrightarrow{Mf} & MB
\end{align*}
\]

\[
\begin{align*}
h_{A_1} \times \ldots \times h_{A_n} \\
\xrightarrow{Nf} & NB
\end{align*}
\]

commutes.
For each relation symbol \( R \leftrightarrow A_1 \times \ldots \times A_n \) in \( \Sigma \),

\[
\begin{array}{ccc}
MR & \rightarrow & MA_1 \times \ldots \times MA_n \\
\exists & & h_{A_1} \times \ldots \times h_{A_n} \\
NR & \rightarrow & NA_1 \times \ldots \times NA_n
\end{array}
\]

commutes.

Any finite limit preserving functor \( F : C \rightarrow D \) transfers \( \Sigma \)-structures in \( C \) into \( \Sigma \)-structures in \( D \).
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

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

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

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

is the projection on the \( i \)-th component.

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

\[
[[\vec{x}.t]]_M = Mf([[\vec{x}.t_1]]_M, \ldots, [[\vec{x}.t_1]]_M).
\]
To interpret formulae in $\mathcal{C}$, we need to assume that $\mathcal{C}$ has enough structure. Formulae in context $\{\overrightarrow{x}.\phi\}$ will be interpreted as subobjects

$$[[\overrightarrow{x}.\phi]]_M \hookrightarrow MA_1 \times \ldots \times MA_n \in \text{Sub}_C MA_1 \times \ldots \times MA_n.$$ 

These subobject lattice are all Heyting algebras in a topos. Thus, for a topos, we have enough structure to interpret $\land, \lor, \top, \bot, \exists, \forall$ 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 $R \hookrightarrow B_1 \times \ldots \times B_m$ and $t_i: A_i$ are terms with a list of variable $\overrightarrow{x} = x_1, \ldots, x_n$, $x_i: A_i$, suitable for all of them, define $[[\overrightarrow{x}.R(t_1, \ldots, t_m)]]_M$ by the subobject in the pullback square
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 $T$ is a $\Sigma$-structure $M$ in which all axioms of $T$ are satisfied.

★
Let $X, Y$ be sets. For $S \subseteq X \times Y$, consider $p: X \times Y \to Y$ which induces a functor

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

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 $\exists_p, \forall_p : \mathcal{P}(X \times Y) \to \mathcal{P}(Y)$ satisfying the chain of adjunctions

$$\exists_p \dashv p^* \dashv \forall_p,$$

i.e. $\exists_p S \subseteq T \iff S \subseteq X \times T$

$$X \times T \subseteq S \iff T \subseteq \forall_p S.$$
More generally, in a topos $\mathcal{C}$, $f : A \to B$ induces

$$f^* : \text{Sub}_\mathcal{C} B \to \text{Sub}_\mathcal{C} A$$

preserving the lattice structures. This functor has adjoints

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

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 were 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 \( C \) one can define the **canonical signature** \( \Sigma_C \) of \( C \) as having a type \( \Gamma_A \) for each object, a function symbol \( \Gamma f : \Gamma A \to \Gamma B \) for each morphism \( f : A \to B \) in \( C \), and a relation symbol \( \Gamma R \rightharpoonup \Gamma A \) for each subobject \( R \rightharpoonup A \) in \( C \). In addition, all the compositional identities in \( C \) are formulated as axioms in \( \Sigma_C \). If \( C \) is a topos one can for example prove

\[
f \text{ is mono in } C \quad \iff \quad f(x) = f(y) \vdash_{x,y} x = y \text{ is derivable in } \Sigma_C.
\]

Thus, \( \Sigma_C \) gives a way of arguing in a topos using elements.
**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 $\Sigma$. If a sequent $\sigma$ is derivable in $\mathbb{T}$, then $\sigma$ is valid in any $\mathbb{T}$-model in any topos.

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

$$
\frac{
\phi \vdash \overline{x} \psi \
\chi \vdash \overline{x} \lambda
}{
\phi \vdash \overline{x} \psi
}
$$

this just means that $\llbracket \overline{x}.\phi \rrbracket_M \leq \llbracket \overline{x}.\psi \rrbracket_M$, and $\llbracket \overline{x}.\chi \rrbracket_M \leq \llbracket \overline{x}.\lambda \rrbracket_M$ implies $\llbracket \overline{x}.\phi \rrbracket_M \leq \llbracket \overline{x}.\psi \rrbracket_M$, i.e. $\phi \vdash \overline{x} \psi$ 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 $\land, \lor, \top, \bot, \exists$ with some extra axioms) then a sequent $\sigma$ is derivable in $\mathbb{T}$ using classical logic if and only if it is valid in each model of $\mathbb{T}$ in $\text{Set}$. 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 $\mathbb{T}$. 

---

Categorical Semantics and Topos Theory Seminar University of Oxford, Michaelis 2011
Soundness and Completeness

Toposes satisfy completeness with respect to all first order theories:

**Theorem (Completeness):** Let $\Sigma$ be a signature. Then a first order formula $\vec{x}.\phi$ over $\Sigma$ is provable in intuitionistic first order logic if it is valid in every $\Sigma$-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 $[P, \text{Set}]$ 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.