Sketches of an Elephant
—An Introduction to Topos Theory

Jinyi Wang
Qiuzhen College, Tsinghua University

Autumn 2024

Topos Theory is an elephant that reveals itself in a completely different way as one touches different parts of it.

- To algebraic geometers like Grothendieck and Deligne, a topos is a space supporting exotic cohomologies.
- To categorical logicians W. Lawvere, M. Tierney, A. Joyal etc., a topos is the geometrical incarnation of a first-order theory, or semantics for intuitionistic higher-order logic.
- To set theorists, a topos is a non-standard model of set theory (e.g. Clausen-Scholze condensed sets).
- To homotopy theorists, a topos in the ∞-categorical setting provids a place to do homotopy theory.

This talk is an audacious attempt to sketch all the aspects of topos theory within two hours, providing illustrative examples. In order not to make the presentation too pedantic, some definitions in formal logic are only informally stated.

* * *

1 Toposes as generalized spaces

1.1 Categories of sheaves

Historically, the concept of toposes first appeared as a generalization of topological spaces. The category

$$\mathrm{Sh}(X)$$

of sheaves on a topological space X satisfies some important properties, but not all categories with these properties arise from topological spaces. Therefore it is natural to *define* a new kind of space *by* the "category of sheaves" on it. This kind of spaces is called *toposes*.

What good properties does the category of sheaves have? In one word,

The category of sheaves behaves like the category of sets.

The intuition behind this fact is that a sheaf is a "parametrized family" of sets over a space, so that constructions on sets can be performed "fiberwise" on sheaves.

The category of sets has the following properties, making it the foundation of mathematics.

- It has *limits* and *colimits*; for example
  - a *terminal object* $\{*\}$ and an *initial object* $\varnothing$,

  – *products* $X \times Y$ and *sums* (i.e. disjoint unions) $X + Y$,
  – *equalizers* $\mathrm{eq}(f, g\colon X \rightrightarrows Y)$ and *coequalizers* $\mathrm{coeq}(f, g\colon X \rightrightarrows Y)$.

- It has *exponential objects* (i.e. sets of functions) Y^X, such that there is a natural isomorphism

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

- It has a *subobject classifier* $\Omega = \{\bot, \top\}$, such that there is a natural isomorphism

$$\{\text{subobjects of } X\} \simeq \mathrm{Hom}(X, \Omega).$$

A morphism $f\colon X \to Y$ of topological spaces induces an adjunction

$$\mathrm{Sh}(X) \underset{f_*}{\overset{f^*}{\underset{\bot}{\leftrightarrows}}} \mathrm{Sh}(Y)$$

between the categories of sheaves, where f^* preserves finite limits. This is generalized to morphisms of toposes.

Definition 1.1.1 (topos)**.** A *topos* (plural: *topoi* or *toposes*) is a category with finite limits, exponential objects and a subobject classifier. A morphism $f\colon \mathcal{C} \to \mathcal{D}$ of toposes is an adjunction

$$\mathcal{C} \underset{f_*}{\overset{f^*}{\underset{\bot}{\leftrightarrows}}} \mathcal{D}$$

where f^* preserves finite limits.

Topological spaces are not the only objects on which we can define sheaves. The most general place where the notion of sheaves makes sense is a *site*, which is a category with a notion of *covering*. They are an important way to construct toposes.

Definition 1.1.2 (site)**.** A *site* $(\mathcal{C}, J)$ consists of a category $\mathcal{C}$, and for every object c, a collection $J(c)$ of families of morphisms $\{f_i\colon c_i \to c\}$ called *coverings* of c, satisfying the following condition:

- If $\{f_i\colon c_i \to c\}$ is a covering of c, then for any morphism $g\colon d \to c$ there exists a covering $\{h_j\colon d_j \to d\}$, such that each $g \circ h_j$ factors through some f_i.

Definition 1.1.3 (sheaves on a site)**.** A *sheaf* on a site $(\mathcal{C}, J)$ is a presheaf

$$F\colon \mathcal{C}^{\mathrm{op}} \to \mathsf{Set}$$

satisfying the following *sheaf condition*:

- For every covering $\{f_i\colon c_i \to c\}$ and every compatible family $(s_i \in F(c_i))$ (being compatible means that whenever $g\colon d \to c_i$ and $h\colon d \to c_j$ are such that $f_i g = f_j h$, we have $F(g)(s_i) = F(h)(s_j)$), there exists a unique $s \in F(c)$ such that $F(f_i)(s) = s_i$.

The category $\mathrm{Sh}(\mathcal{C}, J)$ of sheaves on $(\mathcal{C}, J)$ is the full subcategory of $\mathrm{Psh}(\mathcal{C})$ spanned by sheaves on $(\mathcal{C}, J)$.

Example 1.1.4. For X a topological space, the category $\mathrm{Open}(X)$ of open subsets of X form a site, where a covering $\{f_i\colon U_i \to U\}$ is a collection of subset inclusions with $\bigcup_i U_i = U$. Traditional sheaves on X are sheaves on this site.

Constructing a topos from a site $(\mathcal{C}, J)$ is like constructing an algebra from generators and relations: objects of the category $\mathcal{C}$ are like generators, and coverings are like relations. The category $\mathrm{Psh}(\mathcal{C}) = \mathrm{Sh}(\mathcal{C}, \varnothing)$ of presheaves on $\mathcal{C}$ is like the *free algebra* on the generators.

Definition 1.1.5 (Grothendieck topos)**.** A *Grothendieck topos* is a category equivalent to the category of sheaves on some site.

A classical theorem states that the definition is equivalent to the following.

Definition 1.1.6 (Grothendieck topos, alternative definition)**.** A *Grothendieck topos* is a left exact localization[1] of a presheaf category.

Grothendieck toposes are the most important class of toposes, but are not all of them[2]. It is a remarkable fact, discovered by Grothendieck's student J. Giraud, that Grothendieck toposes can also be characterized by a small set of axioms.

Definition 1.1.7 (Giraud axioms)**.** A category $\mathcal{C}$ is said to satisfy the *Giraud axioms* if

(0) $\mathcal{C}$ is *presentable* (i.e. cocomplete and generated under colimits by a small set of small objects);

(1) sums in $\mathcal{C}$ are disjoint;

(2) colimits in $\mathcal{C}$ are stable under pullback;

(3) epimorphisms are effective (i.e. arise as coequalizers).

Example 1.1.8. The archetypical topos is Set, the category of sets, which is also the category $\mathrm{Sh}(\mathrm{pt})$ of sheaves on the point. Every Grothendieck topos $\mathcal{C}$ admits a unique morphism to Set

$$\mathcal{C} \underset{\pi_*}{\overset{\pi^*}{\underset{\longrightarrow}{\overset{\longleftarrow}{\perp}}}} \mathsf{Set}$$

where

- π^* is called the *constant sheaf* functor, and
- π_* is called the *global sections* functor.

This is an analog of the fact that every topological space admits a unique map to the point.

Definition 1.1.9 (point of a topos)**.** A *point* of a topos $\mathcal{C}$ is a morphism $p\colon \mathsf{Set} \to \mathcal{E}$. For such a point p,

- p^* is called the *stalk* functor, and
- p_* is called the *skyscraper sheaf* functor.

In a topological space the points play a dominant role; in toposes they don't. For instance, a map $f\colon X \to Y$ of topological spaces is completely determined by the value of f on the points of X; however this is not true in toposes. In fact, there exists nontrivial toposes with *no point at all*!

Example 1.1.10 (A nontrivial topos without points)**.** P. Deligne gave the following example of a nontrivial topos without points. Let $\mathcal{P}$ be the partially ordered set of measurable subsets $U \subset [0,1]$ modulo difference by a null set (i.e. $U \sim V$ iff $U \setminus V$ and $V \setminus U$ are both of measure zero). Define a covering to be a countable collection of subsets $\{f_i\colon U_i \to U\}$ whose union is equivalent to U. The topos defined by this site is a subtopos of $[0,1]$, but it does not contain any point $x \in [0,1]$ because the measure of any point is zero.

There is a notion called *existence of enough points*, which means isomorphy can be tested stalkwise. However, the assumption is artificial and according to Deligne, in many cases we can do without it.

1.2 Locales

A *locale* is a space defined by the collection of its open subspaces, which behaves like that of a topological space; the difference is that the definition of a locale does not involve "points".

Definition 1.2.1 (locale)**.** A *locale* X is given by a set $\mathcal{O}(X)$, the elements of which are called *open subspaces* of X, with a partial order relation $\leq$ closed under *finite intersections* $\wedge$ and *arbitrary unions*[3] $\bigvee$. A morphism $f\colon X \to Y$ of locales is a mapping $f^*\colon \mathcal{O}(Y) \to \mathcal{O}(X)$ that preserves finite intersections and arbitrary unions.

[1] A left exact localization is a fully faithful functor with a left adjoint that preserves finite limits.
[2] For example, the category Fin of finite sets is a topos, but not a Grothendieck topos.
[3] The *intersection* of a family of elements $\{x_i \mid i \in I\}$ is the largest element $y \in \mathcal{O}(X)$ such that $y \leq x_i$ for all $i \in I$; the union is defined dually.

By adjoint functor theorem a morphism $f\colon X \to Y$ of locales is equivalently an adjunction

$$\mathcal{O}(X) \underset{f_*}{\overset{f^*}{\underset{\perp}{\rightleftarrows}}} \mathcal{O}(Y)$$

where f^* preserves finite limits. Note the similarity to morphisms of toposes.

Also by adjoint functor theorem the functor $(- \wedge x)$ has a right adjoint; this shows that

- $\mathcal{O}(X)$ has an *implication* operation $\Rightarrow$ such that

$$z \leq (x \Rightarrow y) \quad \text{if and only if} \quad (z \wedge x) \leq y.$$

Notice the similarity to the exponential objects in a topos; yet we have used a logical operation to denote it. This is to suggest that the collection of open subspaces behaves like the collection of *propositions*, or *truth values*.

In some sense, locales can also be seen as a generalization of topological spaces, a "lower" analogue which does not go as far as toposes. An open subspace of X is a "sheaf of truth values" over X. In the view of homotopy theory, sets are 0-truncated types and propositions are (-1)-truncated types. Therefore locales are also called 0-*toposes*. A locale can be turned into a topos by taking the category of sheaves.

Definition 1.2.2 (point of a locale)**.** Let pt be the locale with $\mathcal{O}(\mathrm{pt}) = \{\perp, \top\}$. A *point* of a locale X is a morphism $p\colon \mathrm{pt} \to X$. Equivalently a point is a map $p^*\colon \mathcal{O}(X) \to \{\perp, \top\}$ preserving finite intersections and arbitrary unions, and thus corresponds to a *completely prime filter* of $\mathcal{O}(X)$.

2 Toposes and logics

Logic has long been treated by mathematicians as a "foundational" area of study far away from other branches of mathematics. For example, when encountering the term "axiom of choice", they take it for granted as a matter of fact, and think only the most unconventional logicians would bother to investigate it. However, with the work of categorical logicians from late-20th century, in which toposes has played a very important role, it became clear that

Logic is something that can be studied just like algebra.

2.1 Locales and propositional logic

Definition 2.1.1 (propositional theory, informal definition)**.** A *propositional theory* $\mathbb{T}$ consists of

- a collection of *atomic propositional symbols* $p, q, r, \cdots$,
- a set of logical formulas called *axioms*.

A *standard model* (or just *model*) of a propositional theory $\mathbb{T}$ is an assignment of truth value to all the atomic propositional symbols satisfying the axioms.

For X a locale, an X-*model* of a propositional theory $\mathbb{T}$ is an assignment of value in $\mathcal{O}(X)$ to all the atomic propositional symbols, such that the axioms have value $\top \in \mathcal{O}(X)$. A standard model is thus a pt-model.

Example 2.1.2 (theory of real numbers)**.** A theory $\mathbb{T}_{\mathbb{R}}$, called the *theory of real numbers*, has

- for every pair of rational numbers $a < b$, an atomic propositional symbol $p_{a,b}$;
- for every sequence of rational numbers $a \leq b < c \leq d$, an axiom $(p_{a,c} \wedge p_{b,d}) \Leftrightarrow p_{b,c}$;
- for every rational number $\varepsilon > 0$, an axiom $\bigvee_{a \in \mathbb{Q}} p_{a-\varepsilon, a+\varepsilon}$.

A model of $\mathbb{T}_{\mathbb{R}}$ is a real number.

Example 2.1.3 (theory of random numbers)**.** The topos in Example 1.1.10 is presented by a propositional theory called the *theory of random numbers* in $[0, 1]$. It has

- for every equivalence class of measurable subsets (i.e. events) $U \subset [0,1]$ an atomic propositional symbol p_U;
- for every inclusion $U \subset V$ an axiom $p_U \Rightarrow p_V$.

If this theory had a model x, then x would be a point in $[0,1]$ contained in any subset of measure 1, and would be called a "random number" in $[0,1]$. But no number in $[0,1]$ is a random number; thus the theory has no models.

Locales are in close relation to *geometric propositional theories*, where we allow finite $\wedge$ and infinite $\bigvee$ in logical formulas. Such a theory $\mathbb{T}$ can be regarded as a *presentation* of a locale $\mathcal{L}_{\mathbb{T}}$ in terms of generators and relations; the atomic symbols are the generators, and the axioms are the relations. In this presentation, a *proposition* (logical formula) in $\mathbb{T}$ is an *open subspace* of $\mathcal{L}_{\mathbb{T}}$, and logical operations $\wedge, \bigvee$ correspond to finite intersections and arbitrary unions of open subspaces.

A crucial observation is that a *model* of $\mathbb{T}$ corresponds to a *point* of the locale $\mathcal{L}_{\mathbb{T}}$; more generally, an X-model of the theory is a morphism of locales $X \to \mathcal{L}_{\mathbb{T}}$. For example the theory of real numbers present the locale $\mathbb{R}$ of real numbers, and a model is a point of $\mathbb{R}$, while an X-model is a morphism of locales $X \to \mathbb{R}$.

An X-model of a theory $\mathbb{T}$ can be interpreted as a family of models *parametrized* by X, or a model of $\mathbb{T}$ *internal* to X. The identity of the locale $\mathcal{L}_{\mathbb{T}}$ corresponds to the *generic model* of $\mathbb{T}$, and every other model is a pullback of it.

Definition 2.1.4 (deduction system of geometric propositional logic)**.** Any type of logic has an associated *deduction system*, which is a set of rules defining the ways we can derive new propositions from existing ones. The deduction system of geometric propositional logic consists of the following rules.

- (truth) We have $\phi \Rightarrow \top$.
- (falsehood) We have $\bot \Rightarrow \phi$.
- (finite conjunction introduction) Given $\phi \Rightarrow \psi$ and $\phi \Rightarrow \chi$, we can derive $\phi \Rightarrow (\psi \wedge \chi)$.
- (finite conjunction elimination) We have $(\phi \wedge \psi) \Rightarrow \phi$ and $(\phi \wedge \psi) \Rightarrow \psi$.
- (infinite disjunction introduction) For any set $\{\phi_i\}_{i \in I}$ of formulas we have $\phi_{i_0} \Rightarrow \bigvee_{i \in I} \phi_i$.
- (infinite disjunction elimination) For any set $\{\phi_i\}_{i \in I}$ of formulas, if for each i is given $\phi_i \Rightarrow \chi$, we have $\bigvee_{i \in I} \phi_i \Rightarrow \chi$.
- (distributivity) For any set $\{\phi_i\}_{i \in I}$ of formulas we have $\psi \wedge (\bigvee_{i \in I} \phi_i) \Leftrightarrow \bigvee_{i \in I}(\psi \wedge \phi_i)$.

Note that the *law of excludded middle* is not present; there is no rule saying that we can always derive $\phi \vee \neg\phi$ (where $\neg\phi$ means $\phi \Rightarrow \bot$). Moreover, we cannot derive it from the other rules, because there exists a theory in which this rule fails.

Example 2.1.5 (theory where the LEM fails)**.** In a topological space X viewed as a theory in geometric propositional logic, an open subset U is a proposition, and the negation $\neg U$ is the largest open subset V such that $U \wedge V = \bot$; in other words, $\neg U$ is the interior of the complement of U. Therefore $U \vee \neg U = \top$ holds if and only if U is clopen.

A propositional theory can be consistent (i.e. cannot derive the contradiction $\bot$) while having no models, just like a locale can be nontrivial while having no points.

So far we have seen a clear ongoing analogy between logic and algebra:

Logic	Ring theory
propositional theory	ring (presented as quotient of polynomial ring)
atomic symbol	generator of ring
logical formula	element of ring
logical operations $\wedge, \vee$	algebraic operations $\times, +$
deduction system	algebraic laws
axioms	relations (system of polynomial equations)
locale presented by propositional theory	spectrum of ring
point of locale (prime filter)	point of spectrum (prime ideal)
model of theory	solution of system of polynomial equations
standard model of theory	closed point (field-valued point)
inconsistent theory ($\top = \bot$)	zero ring ($1 = 0$)
making assumption	taking localization
...	...

2.2 First-order logic and categorical semantics

Mathematical language is the *tool* of our mathematical reasoning, but seldom the *object* of our study. Ordinary mathematical language is the reasoning in the category Set, and it is believed that nothing can be said unless they are converted into something in this category[4]. However, an important insight of categorical logic is that,

A language can be interpreted in any category with sufficient structures.

Moreover, all theorems we may derive using a language is automatically true in any category with sufficient structures to interprete it. This is called *categorical semantics.*

Example 2.2.1 (interpretation of logical formula)**.** This example shows how logical formulas are interpreted in categories.

- For maps $f, g\colon X \to Y$, the formula

$$\{x \in X \mid f(x) = g(x)\}$$

  is interpreted as the equalizer $\mathrm{eq}(f, g)$ in a category having equilizers.

- For a binary relation $R \hookrightarrow X \times Y$ the formula

$$\{x \in X \mid \exists y\, R(x, y)\}$$

  is interpreted as the *image* of R under the projection $\pi_1 : X \times Y \to X$, which is the largest subobject of X through which $R \hookrightarrow X \times Y \to X$ factors. Interpretation of this formula requires the existence and well-behavedness of images in the category, and a topos is a good enough category for this.

- For a binary relation $R \hookrightarrow X \times Y$ the formula

$$\{x \in X \mid \forall y\, R(x, y)\}$$

  is interpreted as the largest subobject $U \hookrightarrow X$ such that $\pi_1^* U \leq R$. Again this structure exists in any topos.

- The *axiom of choice*

$$\forall f \in Y^X \big((\forall y \in Y\, \exists x \in X\, f(x) = y) \Rightarrow \exists s \in X^Y\, f \circ s = \mathrm{id}_Y \big)$$

  can be interpreted in any topos. Even if we take it for granted in the topos Set, it is easy to construct a topos that does not satisfy this axiom[5].

Definition 2.2.2 (first-order theory, informal definition)**.** A *first-order theory* $\mathbb{T}$ consists of

[4] For example, even in formulations of ∞-category theory people use concepts like simplicial sets and model structures which ultimately reduce to set-theoretical notions.

[5] Consider the topos $\mathrm{Sh}(S^1)$ of sheaves on the circle. The nontrivial two-fold covering of S^1 is a surjection to the terminal object that does not have a section.

- a collection of *types* $A, B, C, \cdots$;
- a collection of *function symbols* $f, g, h, \cdots$, each of the form $f\colon A_1 \times \cdots \times A_n \to B$ (when $n = 0$ it is also called a *constant symbol* and denoted $f \in B$);
- a collection of *relation symbols* $R, S, \cdots$, each of the form $R \hookrightarrow A_1 \times \cdots \times A_n$ (when $n = 0$ it is also called an *atomic proposition*);
- a set of logical formulas called *axioms*.

A *model* of $\mathbb{T}$ in a category $\mathcal{C}$ is an assignment of

- an object $[\![A]\!]$ to every type A,
- a morphism $[\![f]\!]\colon [\![A_1]\!] \times \cdots \times [\![A_n]\!] \to [\![B]\!]$ to every function symbol $f\colon A_1 \times \cdots \times A_n \to B$, and
- a subobject $[\![R]\!] \hookrightarrow [\![A_1]\!] \times \cdots \times [\![A_n]\!]$ to every relation symbol $R \hookrightarrow A_1 \times \cdots \times A_n$,

satisfying the axioms.

Example 2.2.3 (theory of groups)**.** The *theory of groups* $\mathbb{T}_{\mathrm{Grp}}$ can be given by

- a type G;
- three function symbols $e \in G$ (identity), $i\colon G \to G$ (inverse) and $m\colon G^2 \to G$ (multiplication);
- the axioms
  - $m(x, e) = m(e, x) = x$ (identity),
  - $m(x, m(y, z)) = m(m(x, y), z)$ (associativity),
  - $m(x, i(x)) = m(i(x), x) = e$ (inverse).

A model of $\mathbb{T}_{\mathrm{Grp}}$ in Set is an ordinary group. A model of $\mathbb{T}_{\mathrm{Grp}}$ in the category Mfd of manifolds is a Lie group.

Just like a propositional theory is a presentation of a locale, a first-order theory can be regarded as a presentation of a topos called the *classifying topos*. This presentation is achieved by first constructing the *syntactic site*, a site encoding the syntax of the first-order theory.

Example 2.2.4 (syntactic site of groups)**.** As a rudimentary example, we describe the *syntactic site* of groups $\mathcal{C}_{\mathrm{Grp}}$. An object of this category has the form

$$\langle x_1, \cdots, x_n \mid \varphi(x_1, \cdots, x_n) \rangle$$

where φ is a finite conjunction $(s_1 = t_1) \wedge \cdots \wedge (s_n = t_n)$ of formulas in the theory of groups. Note that a *formula* in this case is just an equation $s = t$ of two terms, and a *term* is inductively defined as follows,

- the constant e is a term;
- a single variable x, y, z, etc. is a term;
- if s, t are terms, then $m(s, t)$, $i(t)$ are terms.

A morphism of the syntactic site is a *substitution of variables* that respects the formulas. For example, there is a morphism

$$\langle x \rangle \to \langle y, z \mid y^2 = z^3 \rangle$$

given by the substitution

$$y = x^3, z = x^2.$$

In general, a morphism

$$f\colon \langle x_1, \cdots, x_n \mid \varphi(x_1, \cdots, x_n) \rangle \to \langle y_1, \cdots, y_m \mid \psi(y_1, \cdots, y_m) \rangle$$

is a substitution of variables

$$y_i = f_i(x_1, \cdots, x_n)\,(i = 1, \cdots, m)$$

such that, from $\varphi(x_1, \cdots, x_n)$, one may deduce $\psi(f_1, \cdots, f_m)$ in the theory of groups. Intuitively, a morphism is a map between abstract "sets of solutions" of polynomial equations, although we are not actually solving equations in any specific set. Another description of the category $\mathcal{C}_{\mathrm{Grp}}$ is that it is dual to the category $\mathsf{Grp}_{\mathrm{fp}}$ of finitely presented groups, and the morphism in the above example is dual to the morphism of finitely presented groups

$$\langle y, z \mid y^2 = z^3 \rangle \to \langle x \rangle, \quad y \mapsto x^3, z \mapsto x^2.$$

There are no nontrivial coverings in this site, but for theories with more logical structures like $\vee$ or $\exists$, we will need a notion of covering. (For example in the theory of *fields* we have an axiom $x = 0 \vee \exists y\, xy = 1$, and we should demand that $\langle x \mid x = 0 \rangle$ and $\langle x \mid \exists y\, xy = 1 \rangle$ together cover the object $\langle x \rangle$ in the syntactic site of fields.)

The syntactic site of groups has the universal property that for any category $\mathcal{D}$ with finite limits, a group in $\mathcal{D}$ (i.e. a model of $\mathbb{T}_{\mathrm{Grp}}$ in $\mathcal{D}$) is equivalent to a finite-limit-preserving functor

$$\mathcal{C}_{\mathrm{Grp}} \to \mathcal{D}.$$

The idea is that, given a group G in $\mathcal{D}$ we may construct objects like

$$\{(y, z) \in G^2 \mid y^2 = z^3\}$$

using finite limits in the category $\mathcal{D}$; this gives a finite-limit-preserving functor $\mathcal{C}_{\mathrm{Grp}} \to \mathcal{D}$. The multiplication $G \times G \to G$ is recovered from the morphism

$$\langle x, y \rangle \to \langle z \rangle, \quad z = xy.$$

The data of such a functor contains, in a coordinate-free manner, exactly what is required for a group.

Example 2.2.5 (classifying topos of groups)**.** The *classifying topos of groups* can be constructed as the presheaf topos $\mathrm{Psh}(\mathcal{C}_{\mathrm{Grp}})$ over the syntactic site of groups. It has the universal property that for any cocomplete topos $\mathcal{E}$, a group in $\mathcal{E}$ is equivalent to a morphism of toposes $f \colon \mathcal{E} \to \mathrm{Psh}(\mathcal{C}_{\mathrm{Grp}})$, where $f^* \colon \mathrm{Psh}(\mathcal{C}_{\mathrm{Grp}}) \to \mathcal{E}$ is the extension by colimits[6] of the finite-limit-preserving functor $\mathcal{C}_{\mathrm{Grp}} \to \mathcal{E}$.

The classifying topos of groups may be thought of as the *moduli space* $\mathcal{M}$ of groups. To illustrate this idea we can consider the sheaf topos $\mathrm{Sh}(X)$ on a space X. A group G in $\mathrm{Sh}(X)$ is a sheaf of groups on X, a "continuously varying family of groups" over X (indeed, for every point $p \colon \mathrm{pt} \to X$ there is an ordinary group p^*G which is the value of the family at the point p). The property of the classifying topos says that the group G is classified by some morphism $X \to \mathcal{M}$. Therefore it is intuitive to think of $\mathcal{M}$ as the "space of all groups" and the map $X \to \mathcal{M}$ as an assignment of a group to each point of X. As with all moduli spaces, there is a *generic model*, the *universal group* over $\mathcal{M}$ whose pullbacks give all groups in all toposes.

Example 2.2.6 (classifying topos of a discrete group)**.** This example given in [4] relates the concept of classifying toposes to the classical notion of classifying spaces of groups. Let G be a discrete group (in Set). In topology we know that, at least in good cases, isomorphism classes of principal G-bundles on X are in bijection with homotopy classes of maps $X \to BG$, where BG is the *classifying space* of G. In category theory, the symbol BG is used to denote something different but related: we use BG to denote the category with only one object $*$ and hom-set $\mathrm{Hom}(*, *) \simeq G$. Presheaves on BG are equivalently G-sets, i.e. sets with a right G-action.

It can be shown that for any topos $\mathcal{X}$, isomorphism classes of "principal G-bundles" on $\mathcal{X}$ are in bijection with morphisms of toposes $\mathcal{X} \to \mathrm{Psh}(BG)$. By definition, a "principal G-bundle" (also called a G-torsor) on $\mathcal{X}$ is a model for the theory with

- A type T,
- a function symbol $\rho_g \colon T \to T$ for every $g \in G$,
- (T is inhabited) an axiom $\exists x \in T$,
- (the action is free) an axiom $\rho_g x \neq \rho_h x$ for every pair of distinct elements $g, h \in G$, and
- (the action is transitive) an axiom $\bigvee_{g \in G} \rho_g x = y$.

A G-torsor in the sheaf topos $\mathrm{Sh}(X)$ over a space X is just a principal G-bundle over X.

[6]The presheaf category $\mathrm{Psh}(\mathcal{C})$ is the *free cocompletion* of $\mathcal{C}$, meaning that for any cocomplete category $\mathcal{D}$, a functor $\mathcal{C} \to \mathcal{D}$ extends uniquely to a colimit-preserving functor $\mathrm{Psh}(\mathcal{C}) \to \mathcal{D}$.

2.3 Internal language of topos

Definition 2.3.1 (internal language of category, informal definition)**.** To any category $\mathcal{C}$ with suitable structures we can associate a first-order language called the *internal language* of $\mathcal{C}$, it has

- objects of $\mathcal{C}$ as types,
- morphisms $f\colon X \to Y$ in $\mathcal{C}$ as function symbols,
- subobjects $R \hookrightarrow X$ in $\mathcal{C}$ as relation symbols.

The fact that toposes behave like the category of sets means that a large part of the ordinary theory of sets, which we use throughout mathematics to reason about mathematical structures, can actually be interpreted internal to any topos. We may talk about objects and morphisms, definitions and constructions, propositions and proofs, almost everything in mathematics as if we are working in the category of sets.

Example 2.3.2. The use of internal language clarifies many constructions in algebraic geometry. Recall that the category of schemes is a subcategory of the topos

$$\mathrm{Psh}(\mathsf{Aff}) \simeq \mathsf{Fun}(\mathsf{Ring}, \mathsf{Set}),$$

and an important object is the *affine line* $\mathbb{A}^1 = \operatorname{spec} \mathbb{Z}[x]$, the forgetful functor $\mathsf{Ring} \to \mathsf{Set}$.

Using the internal language of this topos, we can define some objects like

$$\{(x, y, z) \in (A^1)^3 \mid x^n + y^n = z^n\}$$

which is the functor $\mathsf{Ring} \to \mathsf{Set}$ given by

$$R \mapsto \{(x, y, z) \in R^3 \mid x^n + y^n = z^n\};$$

this turns out to be the Fermat scheme $\operatorname{spec} \mathbb{Z}[x, y, z]/(x^n + y^n - z^n)$.

The *multiplicative group* $\mathbb{G}_m$ can be defined by

$$\{x \in \mathbb{A}^1 \mid \exists y\, xy = 1\}.$$

The 1-dimensional *projective space* $\mathbb{P}^1$ can be defined by

$$\{(x, y) \in (\mathbb{A}^1)^2 \mid x \neq 0 \vee y \neq 0\}/\sim,$$

where $\sim$ is the equivalence relation where $(x, y) \sim (x', y')$ if and only if $\exists \lambda \in \mathbb{A}^1\, (x', y') = \lambda(x, y)$.

Roughly speaking, there are two flavors of toposes (although there is no logical distinction between them):

- a *petit topos* (small topos) is a topos regarded as a generalized space (whose objects are sheaves on this generalized space), while
- a *gros topos* (big topos) is a topos whose objects are regarded as generalized spaces.

The previous example is a gros topos, while the next example is a petit topos.

Example 2.3.3. For a scheme X, the internal language of the topos $\mathrm{Sh}(X)$ also simplifies various notions in algebraic geometry. For example the structure sheaf $\mathcal{O}_X$ is just a *local ring*. A sheaf of $\mathcal{O}_X$-module is finite type if and only if, from the internal perspective, it is a *finitely generated $\mathcal{O}_X$-module*. If in a short exact sequence of sheaves of $\mathcal{O}_X$-modules the two outer ones are of finite type, then the middle one is too. In this manner ([1]),

Any intuitionistically valid theorem about modules yields a theorem about sheaves of modules.

Example 2.3.4. The internal language of certain gros toposes help formulating theories involving "infinite-dimensional objects" like the space of fields in physics ([2]). Physical field configurations Φ are maps from a spacetime manifold X to some coefficient space F, so that the space of fields is not an ordinary manifold. However it can be seen as an object of the gros topos $\mathrm{Sh}(\mathsf{Mfd})$ where the "mapping space" from any object to any other object is again an object in the topos. We can then study differential forms etc. on this mapping space as if it were an ordinary manifold.

Mfd is the site given by open covering of manifolds. Define the *moduli space of differential k-forms* Ω^k by the presheaf

$$M \mapsto \Omega^k(M),$$

and call a map $X \to \Omega^k$ in $\mathrm{Sh}(\mathsf{Mfd})$ a *differential k-form* on X.

Example 2.3.5 (Continuum Hypothesis)**.** One of the most famous works done using ideas from topos theory is the proof of the independence of the Continuum Hypothesis by P. Cohen in the 1960s. The hypothesis states that

- There is no cardinality between $\mathbb{N}$ and $2^{\mathbb{N}}$.

This statement can be interpreted in the internal logic of any Boolean topos, that is, one where the law of excludded middle holds true. Cohen did the following: begin with a very large cardinal $\kappa > 2^{\mathbb{N}}$, define a Boolean topos which could serve as a model of set theory, and in which

$$\underline{\mathbb{N}} < \underline{2^{\mathbb{N}}} < \kappa \leq 2^{\underline{\mathbb{N}}},$$

thus violating the Continuum Hypothesis, showing that it cannot be derived from the axioms of set theory.

3 ∞-toposes and homotopy theory

In the same way that a topos is a category in which we can do *set theory*, an ∞-topos is an ∞-category in which we can do *homotopy theory*; we can perform in an ∞-topos everything we do on spaces, taking homotopy groups and cohomology, computing connectivity and truncatedness, forming Postnikov towers, looking for delooping and stabilization, etc. The archetypical ∞-topos is the category Grpd_∞ of ∞-groupoids. The full subcategory of $(n-1)$-truncated objects in an ∞-topos form an *n-topos.*

Analogous to Definition 1.1.6, a Grothendieck ∞-topos can be defined as a left exact localization of a presheaf category

$$\mathcal{E} \hookrightarrow \mathrm{Psh}(\mathcal{C}) := \mathsf{Fun}(\mathcal{C}, \mathsf{Grpd}_\infty).$$

In the context of ∞-category theory we can also state the Giraud axioms, with the only change that notion of *effective epimorphisms* is upgraded to that of *effective groupoids.*

3.1 Third Giraud axiom and principal bundles

Groupoids are generalizations of *equivalence relations* in the context of ∞-category theory. A *groupoid object* in an ∞-category $\mathcal{C}$ is a simplicial object $\mathcal{G}\colon \Delta^{\mathrm{op}} \to \mathcal{C}$ such that the natural map

$$\mathcal{G}_n \to \mathcal{G}_1 \times_{\mathcal{G}_0} \times \cdots \times_{\mathcal{G}_0} \mathcal{G}_1$$

is an equivalence for every $n \geq 1$. The object $\mathcal{G}_0$ is the "set of vertices" of the groupoid.

In accordance with the classical observation that a group is a one-object groupoid, a *group object* in an ∞-category $\mathcal{C}$ is a groupoid object $\mathcal{G}$ with an equivalence $\mathcal{G}_0 \simeq 1$. We call $\mathcal{G}_1$ the *underlying object* of $\mathcal{G}$, and denote it by G.

For $\mathcal{G}$ a groupoid object, the object $\operatorname{colim}\mathcal{G}$ may be thought of as a quotient of $\mathcal{G}_0$ obtained by gluing along the morphisms of $\mathcal{G}$, a generalization of the "quotient by equivalence relation" operation.

To any morphism $P \to X$ in $\mathcal{C}$ is associated a groupoid object

$$\check{\mathrm{C}}(P \to X) = \;\cdots \substack{\to\\\to\\\to\\\to} P \times_X P \times_X P \substack{\to\\\to\\\to} P \times_X P \rightrightarrows P$$

called the *Čech nerve.* A morphism $P \to X$ is called an *effective epimorphism* if it is the quotient projection of its own Čech nerve:

$$\operatorname{colim} \check{\mathrm{C}}(P \to X) \xrightarrow{\simeq} X.$$

The third Giraud axiom for ∞-toposes says

Groupoids are the same as effective epimorphisms.

Specifically, we require every groupoid object $\mathcal{G}$ in $\mathcal{C}$ to be equivalent to the Čech nerve of its quotient projection $\mathcal{G}_0 \to \operatorname{colim}\mathcal{G}$.

A remarkable consequence of the third Giraud axiom is the classification of principal G-bundles by the classifying space $\mathbf{B}G$, shown in [7]. The idea is the following. A group object G is by definition the *action groupoid* $*/\!/G$ of G acting on a point. The *delooping* $\mathbf{B}G$ is defined to be the quotient object of $*/\!/G$, which satisfies $\Omega\mathbf{B}G \simeq G$ by the third Giraud axiom. A *principal G-bundle* $P \to X$ over X is an action groupoid $P/\!/G$ whose quotient projection is $P \to X$, giving a morphism of groupoids $P/\!/G \to */\!/G$ and thus a morphism $X \to \mathbf{B}G$.

3.2 Internal language of ∞-topos

It is believed ([3]) that the internal language of ∞-toposes can be given by *Homotopy type theory* (HoTT, described in the amazing textbook [6]).

Identity types. A crucial feature of HoTT[7] is the internal expression of identities using *identity types* allowing equality to be not a *mere proposition*[8]. For every two elements a, b of a given type A there is an identity type $a =_A b$, whose elements are thought of as

- proofs that a and b are equal;
- paths from point a to point b in a space A.

Groupoids (and later ∞-groupoids) give the first model of a type theory with nontrivial identity types, as there can be more than one paths in a groupoid from one point to another, even up to homotopy.

The elimination rule of identity types is called *path induction*. It says

- Given a type family $B(x, y, z)$ depending on $x, y : A$ and $z : x =_A y$, to construct an element of $B(x, y, z)$ it suffices to construct for each $x : A$ an element $d(x) : B(x, x, \mathsf{refl}_x)$.

In particular,

- To prove something about an identity $z : x =_A y$, we may assume WLOG that y is *definitionally* equal to x and z is just refl_x.

The path induction is used throughout HoTT to define, for instance, the concatenation of paths

$$\prod_{x,y,z:A} (x = y) \times (y = z) \to (x = z).$$

Simply put, given a path $p : x = y$, to construct from $q : y = z$ an element $p \cdot q : x = z$, we may assume z is definitionally equal to y and q is just refl_y.

In model category[9] models of HoTT, identity types are modelled by *path objects*

$$(s, t) \colon A^I \to A \times A,$$

which is a replacement of the diagonal

$$(\mathrm{id}, \mathrm{id}) \colon A \to A \times A$$

by a fibration. Fibrations are the interpretation of type families in model category models ([5]).

In model category (or ∞-category) models, path induction is justified by the following fact.

TODO

Proposition as types. Unlike first-order logic, HoTT carries its own deduction system by interpreting *propositions as types*. Propositional calculus is thus encoded by type operations:

Type theory	Logic
type A	proposition
element $a : A$	proof
dependent type $B(x)$	predicate
0	false
1	true
$A + B$	A or B
$A \times B$	A and B
$A \to B$ (function type)	A implies B
$\sum_{x:A} B(x)$	there exists $x : A$ such that $B(x)$
$\prod_{x:A} B(x)$	for all $x : A$ we have $B(x)$
$a =_A b$	$a = b$

[7] To be more specific, the so-called *intensional type theories*

[8] A *mere proposition* is a (-1)-truncated type.

[9] A model category is a 1-category with some extra information (such as a class of weak equivalences, a class of fibrations etc.) that presents an ∞-category.

Proving a proposition is the same thing as *constructing* an element of the corresponding type, for instance

- proving "A or B" is equivalent to constructing an element of $A + B$, which is either an element of A or an element of B (note the conformity to *constructivism*);
- proving "A and B" is equivalent to constructing an element of $A \times B$, which is a pair (a, b) consisting of an element $a : A$ and an element $b : B$;
- proving "A implies B" is equivalent to constructing a function $f \colon A \to B$ which, given any proof of A, returns a proof of B.

It is argued in [6] that *untruncated logic* (i.e. representation of propositions by general types instead of (-1)-truncated ones) is sometimes closer to our intuition when we reason informally about mathematics. For instance, in this kind of logic,

Whenever you prove something exists, the proof already contains an algorithm to find such a thing.

This is impossible in classical logic. However in practice, when we prove the existence of some object x and later use x in further deduction, often times we are not using the fact that it *merely* exists; we are using the *construction* of x.

Σ-types and Π-types. The HoTT expressions of existential and universal quantifiers are respectively called Σ-types and Π-types. For $B(x)$ a family of types parametrized by $x : A$ (interpreted as a "bundle" $B \to A$), the Σ-type $\sum_{x:A} B(x)$ is

- the collection of pairs (a, b) where $a : A$ and $b : B(a)$,
- the internal expression of the statement "there exists x in A such that $B(x)$."

while the Π-type $\prod_{x:A} B(x)$ is

- the collection of sections of the bundle $B \to A$ (i.e. "functions" f assigning an element $f(a) = b : B(a)$ to every $a : A$),
- the internal expression of the statement "for all x in A we have $B(x)$."

Any ∞-topos (in particular, any topos) $\mathcal{E}$ interprets Σ-types and Π-types via the adjoint triple

$$\mathcal{E}_{/A} \; \underset{A_* = \prod_A}{\overset{A_! = \sum_A}{\underset{\perp}{\overset{\perp}{\longleftarrow A^* = A\times - \longrightarrow}}}} \; \mathcal{E},$$

a consequence of being a *cartesian closed category.* The functor $\sum_A \colon \mathcal{E}_{/A} \to \mathcal{E}$ (the "dependent sum") simply sends an object $f \colon B \to A$ in $\mathcal{E}_{/A}$ to the object B in $\mathcal{E}$. The functor $\prod_A \colon \mathcal{E}_{/A} \to \mathcal{E}$ (the "dependent product") can be defined in the internal language as the "set of sections"

$$\prod_A (f \colon B \to A) = \{ s \colon A \to B \mid fs = \mathrm{id}_A \}$$

which is externally the pullback

$$\begin{array}{ccc} \prod_A f & \longrightarrow & B^A \\ \big\downarrow & & \big\downarrow{\scriptstyle f^A} \\ 1 & \underset{\mathrm{id}_A}{\longrightarrow} & A^A. \end{array}$$

Example 3.2.1. The internal statement saying that there is a *unique* element in a type A is

$$\sum_{a:A} \prod_{x:A} x = a,$$

which literally means "there exists some a in A such that for all x in A we have $x = a$." Externally this statement means only that A is path-connected, that any point x of A is connected to some fixed point a. But internally this means A is *contractible* to the point a! One possible explanation is that a proof of $\prod_{x:A} x = a$ requires a *continuous* choice of path from x to a as x varies along A, resulting in a deformation retract of A to a point. In an ∞-topos model of HoTT, a proof that there is a *unique* element in A is a section of the left vertical map in the pullback

$$\begin{array}{ccc} \prod_{x:A} x = a & \longrightarrow & A \\ \big\downarrow & & \big\downarrow {\scriptstyle (\mathrm{id},\mathrm{id})} \\ A & \xrightarrow[x \mapsto (a,x)]{} & A \times A, \end{array}$$

and equivalently a homotopy from id_A to the constant map a.

3.3 Cohomologies

Sheaf cohomology can produce good invariant on "bad" spaces where other approaches to cohomology fail. Homotopy theory, which studies ∞-*groupoids* (also called *homotopy types*), gives an clean theory of cohomology. Therefore it is natural to believe that sheaves of ∞-groupoids, and more generally, objects of an ∞-topos, are good objects to define cohomologies.

Definition 3.3.1 (nonabelian cohomology, general abstract definition)**.** In an ∞-category $\mathcal{C}$, the *cohomology* of an object X with coefficients in A is the homotopy class of maps from X to A:

$$H(X, A) := \pi_0 \operatorname{Hom}_{\mathcal{C}}(X, A).$$

This is called *nonabelian cohomology*, which is conceptually simpler than it sounds; in the following we will see familiar cohomologies are all special cases of this concept. If the object A admits an n-fold delooping[10] $B^n A$, then we may define the n-th cohomology of X with coefficients in A to be

$$H^n(X, A) := \pi_0 \operatorname{Hom}_{\mathcal{C}}(X, B^n A).$$

Denote by Grpd_∞ the ∞-category of ∞-groupoids.

Example 3.3.2 (ordinary cohomology)**.** For good spaces X (e.g. CW complexes), the ordinary cohomology of X with coefficients in an abelian group A is

$$H^n(X, A) \simeq \pi_0 \operatorname{Hom}_{\mathsf{Grpd}_\infty}(X, K(A, n)).$$

Note that $K(A, n)$ is an n-fold delooping $B^n A$ of the abelian group A.

Example 3.3.3 (sheaf cohomology)**.** For F a sheaf of abelian groups on a space X, the *sheaf cohomology* of X with coefficients in F is

$$H^n(X, F) \simeq \pi_0 \operatorname{Hom}_{\mathrm{Sh}_{\mathsf{Grpd}_\infty}(X)}(X, B^n F),$$

where X denotes the terminal object of the ∞-category $\mathrm{Sh}_{\mathsf{Grpd}_\infty}(X)$ of sheaves of ∞-groupoids on X. By the Dold–Kan correspondence we can regard $B^n F$ as a complex of sheaves of abelian groups, and the above set becomes

$$\pi_0 \operatorname{Hom}_{\mathrm{Sh}_{\mathsf{Ch(Ab)}}(X)}(\underline{\mathbb{Z}}, B^n F)$$

($\underline{\mathbb{Z}}$ being the constant sheaf of integers) which is the traditional definition of sheaf cohomology in the derived category of abelian sheaves on X.

References

[1] Ingo Blechschmidt. *Using the internal language of toposes in algebraic geometry.* 2021. arXiv: 2111.03685.

[2] Urs Schreiber. *Higher Topos Theory in Physics.* 2023.

[10] A *delooping* BA of an object A satisfies $\Omega BA \simeq A$.

[3] Emily Riehl. *On the ∞-topos semantics of homotopy type theory.* arXiv: 2212.06937.

[4] Saunders Mac Lane and Ieke Moerdijk. *Sheaves in Geometry and Logic: A First Introduction to Topos Theory*. Springer New York, 1994.

[5] Steve Awodey and Michael A. Warren. *Homotopy theoretic models of identity types.* arXiv: 0709.0248

[6] The Univalent Foundations Program, *Homotopy Type Theory: Univalent Foundations of Mathematics.* Institute for Advanced Study, 2013. `https://homotopytypetheory.org/book/`

[7] Thomas Nikolaus, Urs Schreiber, Danny Stevenson, *Principal infinity-bundles – General theory.* arXiv: 1207.0248