Topos Theory, Toposes as Bridges and Applications

by Laurent Lafforgue

(Lagrange Center and Huawei Research Center in Boulogne)

Presentation at the Paris CSTT Workshop

September 18-20, 2024

Historical reminder:

A. Grothendieck first worked in Functional Analysis
before moving to Topology, Geometry and, in particular, discovering Toposes.

Principles of Functional Analysis:

- The main objects of study are no more geometric objects X (such as $\mathbb{R}^n$, open subsets of $\mathbb{R}^n$, manifolds, $\cdots$) but numerical functions f on X.
- Numerical functions f are no more considered in isolation but as elements of "topological vector spaces" $\widehat{\mathcal{C}}_J$ constructed as completions of some linear spaces $\mathcal{C}$ of "nice functions" on X endowed with translation-invariant topologies J (usually defined by a norm or a sequence of semi-norms).

Remarks:

- The possibility to choose different topologies J on the same space $\mathcal{C}$ is a key feature of Functional Analysis.
- Replacing geometric spaces X with function spaces $\widehat{\mathcal{C}}_J$ allows to generalize the notion of space.

Structures on spaces of numerical functions:

- **Point-wise operations:**

$$\left.\begin{array}{c}\text{addition}\\ \text{multiplication}\end{array}\right\} \Rightarrow \left\{\begin{array}{l}\underline{\text{linear structure}}\text{ of spaces of functions}\\ \text{(commutative) }\underline{\text{algebra structure}}\end{array}\right.$$

- **Changes of parameters:**
  - evaluations of functions f on X at points x of X
  $$f \longmapsto f(x)\,,$$
  - restrictions of functions f on X to open subspaces $U \subseteq X$ (localization),
  - more generally, any "geometric map" $x : X' \to X$ between "geometric spaces" induces a transform
  $$\{\text{numerical functions on } X\} \to \{\text{numerical functions on } X'\},$$
  $$f \longmapsto f \circ x\,.$$
- **Differentiation:**
  - On differential manifolds X, "smooth" functions can be differentiated.
  - This operation is linear and verifies the "Leibniz rule".
  - This operation is local and compatible with diffeomorphic changes of parameters.
- **Integration:**
  - Integration of differential forms or of functions w.r.t. a measure μ.
  - Integration is linear and additive w.r.t. its domain.
  - Integration and differentiation are related by the Stokes formula.

From numerical functions to linear sheaves:

The main drawback of numerical functions:

- Numerical functions are meaningless just as numbers are meaningless. (Numbers recover meaning e.g. when they count something.)

$\Rightarrow$ For building a theory of "semantic information", numerical functions are insufficient.

Leray's first idea of linear sheaf:

- A "linear sheaf" on a base space X is defined as a "continuous family of vector spaces" $X \ni x \mapsto V_x$ typically arising from geometry or topology, such as

$$X \ni x \longmapsto V_x = \text{cohomology space } H^i(F_x)$$

where F_x = fiber at $x \in X$ of some fibered space $F \xrightarrow{p} X$.

Gothendieck's "sheaf-function" dictionary:

- A sheaf lift of a numerical function f on a space X is a linear sheaf $X \ni x \longmapsto V_x$ endowed with a geometrically defined continuous family of linear maps

$$\sigma_x : V_x \to V_x\,, \quad x \in X\,,$$

such that $\qquad f(x) = \mathrm{Tr}(V_x \xrightarrow{\sigma_x} V_x)\,, \quad \forall\, x \in X.$

- A function can be considered "meaningfull" when it is lifted to a sheaf.
- Grothendieck has constructed liftings for all operations on functions, including differentiation and integration.

From linear sheaves to set-valued sheaves:

Building on Leray's idea of "sheaf",
Grothendieck introduced three ground-breaking innovations:

- **Replacing "linear sheaves" with "set-valued sheaves":**
  - A "set-valued sheaf" on a base space X can be seen as a "continuous family of sets" $X \ni x \longmapsto F_x$, such as the family of fibers F_x of a fibered space $F \xrightarrow{p} X$.
  - These "set-valued sheaves" may be endowed with any type of (linear or non-linear) structure.
- **Identifying the wider context where the notion of sheaf makes sense:**

This is the context of a "geometrical country" (or category) $\mathcal{C}$ consisting in

$$\mathcal{C} = \begin{cases} \text{“geometric cities” (e.g. the open subspaces of a topological space } X\text{)}, \\ \text{“geometric itineraries” relating these cities (e.g. embeddings of open subspaces)}, \\ \text{a rule for composing itineraries}, \end{cases}$$

and endowed with the choice of a "topology" J on $\mathcal{C}$.

- **Considering sheaves on** $(\mathcal{C}, J)$ **not in isolation:**

They make up a "completed geometric country" $\widehat{\mathcal{C}}_J$.

$\rightarrow$ Any such $\widehat{\mathcal{C}}_J$ can be seen as a "completion" $\mathcal{C} \rightarrow \widehat{\mathcal{C}}_J$.

$\rightarrow$ The chosen topology J on $\mathcal{C}$ can be seen as an extrapolation device, allowing to "complete" $\mathcal{C}$.

Toposes as geometric incarnations of semantics:

Definition. –
A "topos" is a "geometric country" $\mathcal{E}$
which can be presented as the completion $\mathcal{E} \simeq \widehat{\mathcal{C}}_J$
of a "small geometric country" $\mathcal{C}$
endowed with a topology J.

Theorem. –
There is a notion of geometric itinerary of toposes $\mathcal{E}' \to \mathcal{E}$
such that, for any presentation $\mathcal{E} \simeq \widehat{\mathcal{C}}_J$,
geometric itineraries $\mathcal{E}' \to \mathcal{E} \simeq \widehat{\mathcal{C}}_J$
interpret as the "structures" (or "models") of type $\mathbb{T}_{\mathcal{C},J}$ in $\mathcal{E}'$,
for some "first-order geometric theory" $\mathbb{T}_{\mathcal{C},J}$ associated with $(\mathcal{C}, J)$.

Theorem. –
Conversely, for any "first-order geometric theory" $\mathbb{T}$,
there is an associated topos $\mathcal{E}_{\mathbb{T}}$
such that all geometric itineraries $\mathcal{E}' \to \mathcal{E}_{\mathbb{T}}$
interpret as the structures of type $\mathbb{T}$ in $\mathcal{E}'$.

Toposes as universal invariants:

- All classical notions of space X naturally define associated "small geometric countries" $\mathcal{C}_X$ endowed with topologies J_X and, as a consequence, toposes $\mathcal{E}_X = \widehat{(\mathcal{C}_X)}_{J_X}$.
- All classical geometric notions (such as: point, subspace, $\cdots$) on spaces X can be defined directly in terms of the associated toposes $\mathcal{E}_X$, as well as all classical topological invariants, most importantly cohomology and homotopy.
- Furthermore, these classical geometric notions and classical topological invariants remain well-defined for arbitrary toposes, which means they make sense for
  - arbitrary "small geometric countries" $\mathcal{C}$ endowed with a topology J,
  - arbitrary "first-order geometric theories" $\mathbb{T}$.

Definition. – *A "topos invariant" can be*

(i) *a property of toposes $\mathcal{E}$ which is respected by topos equivalences $\mathcal{E}' \cong \mathcal{E}$,*

(ii) *a construction from the geometric country of toposes and their itineraries to another (usually more algebraic) mathematical country which respects or reverses the orientation of itineraries*

$$\begin{array}{ccl} \mathcal{E} & \longmapsto & H(\mathcal{E})\,, \\ (\mathcal{E}' \to \mathcal{E}) & \longmapsto & \begin{cases} \quad H(\mathcal{E}') \quad \to \quad H(\mathcal{E})\,, \\ \text{or } H(\mathcal{E}) \quad \to \quad H(\mathcal{E}')\,. \end{cases} \end{array}$$

Toposes as bridges:

O. Caramello has proposed an extremely general process for deriving mathematical results, based on the following facts:

Facts. –

(i) *Any topos $\mathcal{E}$ has an infinite diversity of presentations*

- *as the "completed geometric country" $\mathcal{E} \xrightarrow{\sim} \widehat{\mathcal{C}}_J$ of "sites" (= geometric sketches) $(\mathcal{C}, J)$,*
- *as the "geometric incarnation" $\mathcal{E} \xrightarrow{\sim} \mathcal{E}_{\mathbb{T}}$ of the semantics of "first-order geometric theories" (= formal languages) $\mathbb{T}$.*

(ii) *There is an ever expanding list of invariants (corresponding to the extraction of different types of partial knowledge) which are generally derived from practice in particular domains and prove to be well-defined for toposes in general.*

(iii) *Examples show that*

- *computations often allow to express invariants of toposes $\mathcal{E}$ in terms of presentations of $\mathcal{E}$ by sites $(\mathcal{C}, J)$ or theories $\mathbb{T}$,*
- *these expressions can be very different, so that topos equivalences*

$$\widehat{\mathcal{C}}_J \cong \widehat{\mathcal{C}}'_{J'}, \qquad \widehat{\mathcal{C}}_J \cong \mathcal{E}_{\mathbb{T}} \qquad \mathcal{E}_{\mathbb{T}'} \cong \mathcal{E}_{\mathbb{T}}$$

often generate unexpected highly non trivial results by the straightforward computation of chosen invariants.

(iv) *In the reverse direction, some classical math results can be lifted to topos equivalences which, through the computation of other topos invariants, generate unexpected "sister results" of the starting classical results.*

The project of implementing the "bridge technique" on machines:

Basic fact. – *In many cases,*
an expression of topos invariants
in terms of presenting "geometric sketches" $(\mathcal{C}, J)$ *or "formal languages"* $\mathbb{T}$
can be found through straightforward computation.

Remark: A research paper of O. Caramello has shown that
an explicit wide class of invariants are theoretically computable.

Elements of a "bridge" automatic system to be engineered:

(1) *An ever expanding library of topos invariants.*

(2) *A computing system for automatically expressing*
an expanding list of topos invariants
in an expanding list of topos presentations
by "geometric sketches" $(\mathcal{C}, J)$ *or "formal languages"* $\mathbb{T}$*.*

(3) *A library of topos equivalences*
relating families of "geometric sketches" or "formal languages".

(4) *A library of lifts of classical mathematical results*
at the higher level of topos equivalences,
through the computation of particular invariants.

Lattices of subtoposes as topos invariants, and the induced bridges:

There is a well-defined notion of "subtopos" of any topos $\mathcal{E}$,
verifying the following properties:

Theorem. –

(i) *For any topos $\mathcal{E}$, its subtoposes form a set, ordered by $\supseteq$, with*

$$\begin{cases} - & \text{arbitrary unions,} \\ - & \text{arbitrary intersections,} \\ - & \text{a "difference" operation } (\mathcal{E}_1, \mathcal{E}_2) \mapsto \mathcal{E}_1 \backslash \mathcal{E}_2 \text{ characterized by} \\ & \qquad \mathcal{E}_3 \supseteq \mathcal{E}_1 \backslash \mathcal{E}_2 \Leftrightarrow \mathcal{E}_3 \cup \mathcal{E}_2 \supseteq \mathcal{E}_1 \,. \end{cases}$$

(ii) *Any topos itinerary $\mathcal{E}' \xrightarrow{e} \mathcal{E}$ induces*

- *a push-forward map $\{\text{subtoposes of } \mathcal{E}'\} \xrightarrow{e_*} \{\text{subtoposes of } \mathcal{E}\}$ which respects arbitrary unions,*
- *a pull-back map $\{\text{subtoposes of } \mathcal{E}\} \xrightarrow{e^{-1}} \{\text{subtoposes of } \mathcal{E}'\}$ which respects arbitrary intersections and finite unions.*

Expression theorem on the geometric side (Grothendieck). –
For any equivalence $\mathcal{E} \cong \widehat{\mathcal{C}}_J$, subtoposes of $\mathcal{E}$ correspond to topologies J' on $\mathcal{C}$ which refine J.

Expression theorem on the logic side (O. Caramello). –
For any equivalence $\mathcal{E} \cong \mathcal{E}_{\mathbb{T}}$, subtoposes of $\mathcal{E}$ correspond to "quotient theories" $\mathbb{T}'$ of $\mathbb{T}$ deduced from $\mathbb{T}$ by retaining the same vocabulary and adding extra axioms.

The project of topology computing for provability problems:

The double expression of the notion of subtopos
in terms of topologies and in terms of theories implies:

Corollary. – *For any equivalence* $\mathcal{E}_{\mathbb{T}} \cong \widehat{\mathcal{C}}_J$,
any implication formulated in the vocabulary of $\mathbb{T}$

$$\varphi \vdash \psi$$

constructively defines families of itineraries of $\mathcal{C}$

$$\chi_{\varphi,\psi}$$

in such a way that:

- *an implication* $\varphi \vdash \psi$ *is provable from the axioms of* $\mathbb{T}$
and an extra family of implications $(\varphi_i \vdash \psi_i)_{i \in I}$
if and only if the families $\chi_{\varphi,\psi}$
belong to the topology J' *on* $\mathcal{C}$ *generated by* J *and the families* χ_{φ_i,ψ_i}, $i \in I$.

Remarks:

(i) There are closed formulas (of O.C. and L.L.)
which allow to express the topology
generated by any family of generators, in the context of any underlying $\mathcal{C}$.

(ii) The project of implementing "topology computing" on machines
has begun at the Lagrange Center with Anthony Bordg.

AI as a problem of language transformations and language elaboration:

Illustrations. –

(1) *Some of the most commonly used AI systems are those which realize automatic translations.*

(2) *LLM's mainly consist in transforming texts (written in some human languages) into vectors belonging to some high-dimensional linear spaces. Their success (which is not absolute but is real and impressive) means that such an a priori unlikely language transformation can be performed in ways which are not arbitrary.*

(3) *Image recognition consists in extracting information from numerical images (represented as vectors whose coordinates are numbers indexed by pixels) and expressing the extracted information in human language (e.g. answering the question: "Does the image contain a cat?").*

The two core problems:

- Given a type of elements of reality (e.g. images, or texts), how to represent them in some formal language which would be well-adapted to them and amenable to computer processing?
- Given a chosen good description formal language, how to process data expressed in this language, moving to another possibly completely different language, hopefully through a succession of intermediate languages?

The question of the type of math representations for data:

Starting remarks:

- *Most often, data are represented in the form of vectors in high-dimensional linear spaces, which are usually understood as discrete approximations of numerical functions defined on subdomains of some* $\mathbb{R}^n$*.*
- *In other words, data are usually understood as elements of some functional spaces.*

Reminder: Toposes can be considered as geometric liftings of functions spaces which, unlike these spaces, incarnate meaning.

The different topossic possibilities for representing data:

(1) *Linear sheaves (conceived as geometric meaningfull liftings of functions) subjected to the extremely elaborate operations introduced by Grothendieck as liftings of usual operations on functions.*

(2) *Ordinary (non-linear and unstructured) sheaves, subjected to the simpler operations of sum, product, exponentiation, push-forward and pull-back by topos itineraries* $\mathcal{E}' \xrightarrow{e} \mathcal{E}$*.*

(3) *Points of toposes, subjected to the unique operation of image*

$$x \longmapsto e(x) \quad \text{by topos itineraries} \quad \mathcal{E}' \xrightarrow{e} \mathcal{E}.$$

(4) *Subtoposes, with their double interpretation in terms of topologies or theories, subjected to the operations of union, intersection, difference, push-forward and pull-back by topos itineraries* $\mathcal{E}' \xrightarrow{e} \mathcal{E}$*.*

Syntactic learning for data represented as subtoposes:

Starting problem:

Given a natural family of data (e.g. images), one should look for a formal language $\mathbb{T}$ in a vocabulary Σ such that each datum can be described in the vocabulary Σ, and the axioms of $\mathbb{T}$ are derived from singular data through a mathematical modelling of inductive reasoning.

Principles:

- The fact that all data under consideration make up a natural family should translate in the possibility to choose a joint vocabulary Σ for them.
- This vocabulary should be large enough to allow to express differences but not too large, so as to express what the data have in common.

Proposition of a topos model for inductive reasoning:

- Each datum considered without interpretation could be represented by some "category" $\mathcal{C}_i$, $i \in I$, endowed with a "naming functor" $N_i : \mathcal{C}_i \to \mathcal{C}_\Sigma$ (= "sentences" in Σ).
- Elements of interpretation of each datum could be represented by topologies J_i on each $\mathcal{C}_i$, defining subtoposes $\widehat{(\mathcal{C}_i)}_{J_i} \hookrightarrow \widehat{\mathcal{C}}_i$.
- The axioms defining $\mathbb{T}$ from the vocabulary Σ would correspond to a subtopos $\mathcal{E}_\mathbb{T} \hookrightarrow \widehat{\mathcal{C}}_\Sigma$ such that $N_i^{-1}(\mathcal{E}_\mathbb{T}) = \widehat{(\mathcal{C}_i)}_{J_i}$, $\forall\, i \in I$, or, equivalently, $(N_i)_*(\widehat{\mathcal{C}}_i)_{J_i} \subseteq \mathcal{E}_\mathbb{T} \subseteq (N_i)_!\widehat{(\mathcal{C}_i)}_{J_i}$.

Looking for some topossic deep learning:

Proposition of a starting representation for data:

• For any natural family of data,
we should elaborate a starting description formal language $\mathbb{T}$
such that data (together with their interpretation coming from language)
would be related to $\mathbb{T}$ through geometric "naming" arrows $\quad N_i : \widehat{(\mathcal{C}_i)}_{J_i} \longrightarrow \mathcal{E}_{\mathbb{T}}$.

• Their images $\quad \mathrm{Im}(N_i) \hookrightarrow \mathcal{E}_{\mathbb{T}} \quad$ would be suptoposes
corresponding to "quotient theories" $\mathbb{T}_i$ of $\mathbb{T}$ $\qquad \mathrm{Im}(N_i) = \mathcal{E}_{\mathbb{T}_i} \hookrightarrow \mathcal{E}_{\mathbb{T}}$.
The theories $\mathbb{T}_i$ would be "descriptions" of the different data.

Proposition of a general form of transformation process for data:

• Such math representations of data as "subtoposes"
could be processed by chains of toposes:

$$\begin{tikzcd}[column sep=small]
 & \mathcal{E}_1 \arrow[dl, "p_1"'] \arrow[dr, "q_1"] & & \cdots \arrow[dl] \arrow[dr] & & \mathcal{E}_n \arrow[dl, "p_n"'] \arrow[dr, "q_n"] & \\
\mathcal{E}_{\mathbb{T}} = \mathcal{E}_{\mathbb{T}_0} & & \mathcal{E}_{\mathbb{T}_1} & & \mathcal{E}_{\mathbb{T}_{n-1}} & & \mathcal{E}_{\mathbb{T}_n}
\end{tikzcd}$$

• There would be a sequence of formal languages $\mathbb{T} = \mathbb{T}_0, \mathbb{T}_1, \cdots, \mathbb{T}_n$
related by "correspondences" consisting in toposes
$\mathcal{E}_i$, $1 \leq i \leq n$, carrying a double model structure of $\mathbb{T}_{i-1}$ and $\mathbb{T}_i$.

• Subtoposes would be processed by the composites of $\qquad (q_i)_* \circ p_i^{-1}\,, \quad 1 \leq i \leq n$.

An already existing topos-inspired system for processing time series:

History and success of the "top down forecast" system:

• In the last 15 years, O. Caramello's father Luigi (who is a computer scientist and has decades of experience) designed AI systems for stock market analysis on the basis of ideas borrowed from the theory of "toposes as bridges".

• These systems are very efficient (`https://www.eossrl.it/udf_rendim_en.html`): on average, they outperform every year 20% above the stock market index.

• Luigi Caramello made them more and more stable
by adding layers representing higher levels of abstraction.

The basic scheme as I understand it:

$$\left\{\begin{array}{c}\text{space of}\\ \text{time series}\\ \text{of stock prices}\end{array}\right\} \xrightarrow[\begin{array}{c}\text{identification of}\\ \text{invariants}\end{array}]{} \left\{\begin{array}{c}\text{space of}\\ \text{time series}\\ \text{of evaluations}\end{array}\right\} \xrightarrow[\begin{array}{c}\text{identification of}\\ \text{higher invariants}\end{array}]{} \left\{\begin{array}{c}\text{space of}\\ \text{time series}\\ \text{of evaluations}\\ \text{on evaluations}\end{array}\right\} \rightarrow \cdots$$

Large interest of these topos-inspired systems:

• Their conception is very abstract and could certainly be adapted
to other kinds of time series analysis problems.

• Presently, they are the most advanced AI systems
inspired from the theory of "toposes as bridges".