

Topoi of automata (SLACS 2025): Local state classifier for automata theory



Ryuya Hora

University of Tokyo

31 October, 2025

Slides (Local
state classifier
for automata
theory)

- ▶ *Topoi of automata I* [Hora 2024b]
- ▶ *Topoi of automata II* [Hora n.d.]
- ▶ *Internal Parameterization of Hyperconnected Quotients* [Hora 2024a]

Table of Contents

- 1 Introduction
- 2 LSCs: Definition
- 3 LSCs: More Examples
- 4 Atmt: LSC of Σ -Set
- 5 Atmt: Covariety of Σ -Set

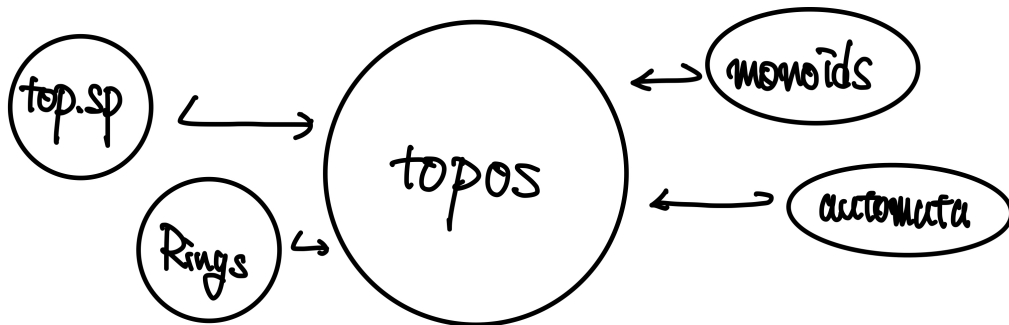
Table of Contents

- 1 Introduction
- 2 LSCs: Definition
- 3 LSCs: More Examples
- 4 Atmt: LSC of Σ -Set
- 5 Atmt: Covariety of Σ -Set

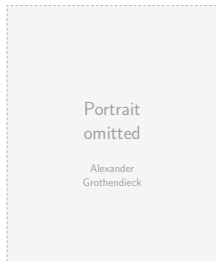
Motivation: Not Generalization, but Geometrization!

My motivation is neither generalisation nor abstraction.

I want to provide a new geometric method/intuition, just like Grothendieck did to number theory!

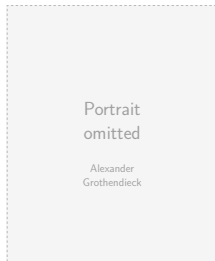


Naive question: Can we imitate Grothendieck?



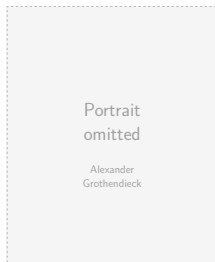
- ▶ Grothendieck defined **topoi** (sites, and schemes) to introduce infinite/continuous geometric methods into number theory, which is traditionally thought of as a finite/discrete subject.

Naive question: Can we imitate Grothendieck?



- ▶ Grothendieck defined **topoi** (sites, and schemes) to introduce infinite/continuous geometric methods into number theory, which is traditionally thought of as a finite/discrete subject.
- ▶ The theory of automata and regular languages is also (pro)finite/discrete combinatorial theory. Is it possible to introduce infinite/continuous geometric methods into this setting?

Naive question: Can we imitate Grothendieck?



- ▶ Grothendieck defined **topoi** (sites, and schemes) to introduce infinite/continuous geometric methods into number theory, which is traditionally thought of as a finite/discrete subject.
- ▶ The theory of automata and regular languages is also (pro)finite/discrete combinatorial theory. Is it possible to introduce infinite/continuous geometric methods into this setting?
- ▶ Primitive word conjecture $\leftarrow$ Topos-Galois theory
- ▶ Generalized star-height $\leftarrow$ Topos-cohomology

Original Motivation: Lawvere's open problems in topos theory

In 2009, W.Lawvere posed 7 open problems in topos theory [Lawvere 2009]:

- 1 **Quotient Toposes** [Kamio and Hora 2024]
- 2 Subquotients and Idempotents
- 3 Boundaries of classes of models
- 4 The jump operator on levels within a topos [Hora, Kamio, and Maehara 2025]
- 5 Coverings that admit averaging and microlinearity
- 6 How strong is the adjointness of fractional exponents?
- 7 The algebra of time

I defined **local state classifiers** as a partial solution to the first problem.

Today, I will focus on another point of view.

Today's motivation

I will try to convey the key idea, and I decided **not to talk about topos theory**.

- ▶ The colimit of all monomorphisms Ξ exists in surprisingly many categories!
- ▶ Ξ is the key to connect category theory and word-congruences.
- ▶ This is a theoretical background of the reason why I started to think about topos-theoretic automata theory.

Question

How can we deal with word-congruences in categorical settings?

Definition of LSC

Definition (local state classifier [Hora 2024a])

A **local state classifier** Ξ of a category $\mathcal{C}$ is a colimit of all monomorphisms:

$$\Xi = \operatorname{colim}(\mathcal{C}_{\text{mono}} \rightarrow \mathcal{C}).$$

The associated cocone is referred to as $\{\xi_X: X \rightarrow \Xi\}_{X \in \operatorname{ob} \mathcal{C}}$.

$$\begin{array}{ccc} U & \xrightarrow{\iota} & X \\ & \searrow \xi_U & \swarrow \xi_X \\ & \Xi & \end{array}$$

Toy Example (1/3): sets

Example (**Set**)

The local state classifier of **Set** is the terminal object $\Xi = \{*\}$.

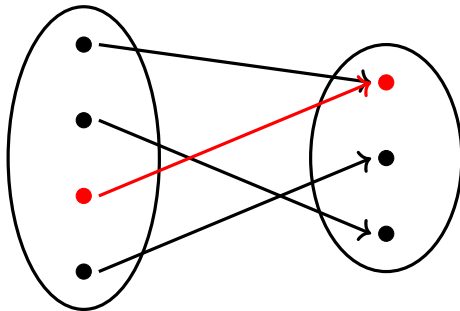
Consider an arbitrary cocone $\{\alpha_X : X \rightarrow L\}$ of $\mathbf{Set}_{\text{inj}} \rightarrow \mathbf{Set}$. Then, for any set X and its element $x \in X$, we have

$$\begin{array}{ccc} \{*\} & \xrightarrow{\text{pick}_x} & X \\ & \searrow \alpha_{\{*\}} & \swarrow \alpha_X \\ & L & \end{array}$$

and therefore $\alpha_X(x) = \alpha_{\{*\}}(*)$.

Toy Example (2/3): pointed sets

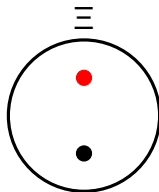
What is the local state classifier of the category of pointed sets **Set**_{*}?



Toy Example (2/3): pointed sets

What is the local state classifier of the category of pointed sets **Set**_{*}?

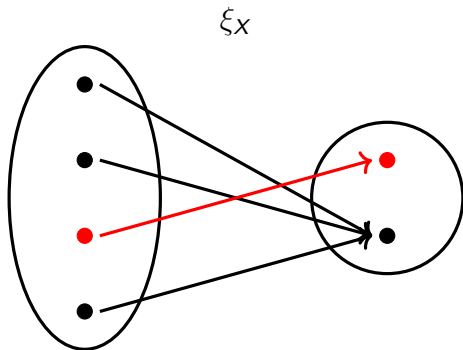
It is



equipped with ...

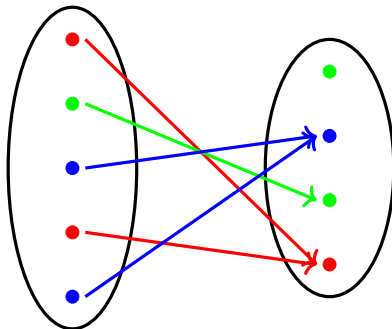
Toy Example (2/3): pointed sets

cocone maps



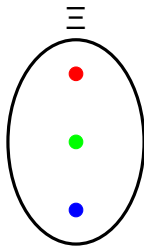
Toy Example (3/3): colored sets

What is the local state classifier of the category of 3-colored sets **Set**/ $\{R,G,B\}$?



Toy Example (3/3): colored sets

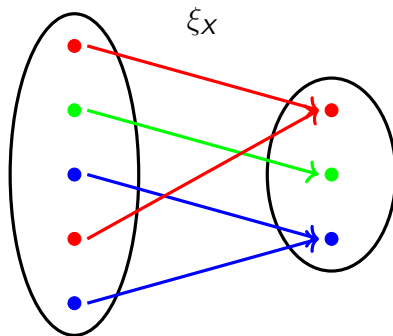
What is the local state classifier of the category of 3-colored sets **Set**/ $\{R,G,B\}$?
It is



equipped with ...

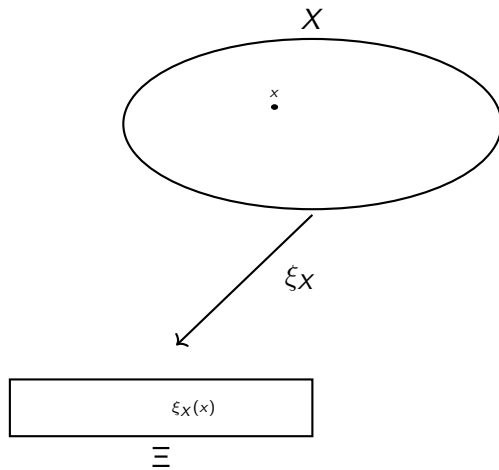
Toy Example (3/3): colored sets

morphisms that classify elements by their colors.



Intuition: why local?

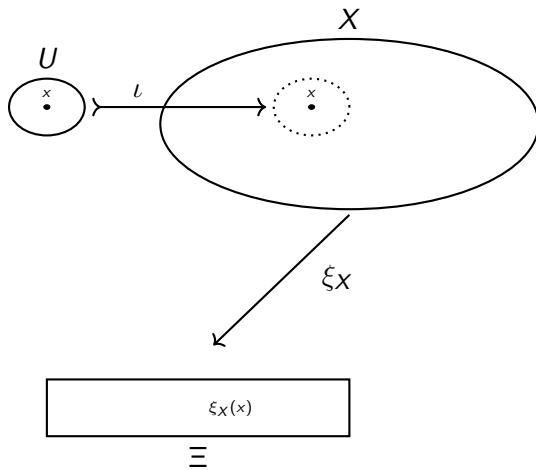
Let $\{\xi_X: X \rightarrow \Xi\}_{X \in \text{ob}\mathcal{C}}$ be a local state classifier.



The value of $\xi_X(x)$ can be calculated in arbitrarily small neighborhood of x .

Intuition: why local?

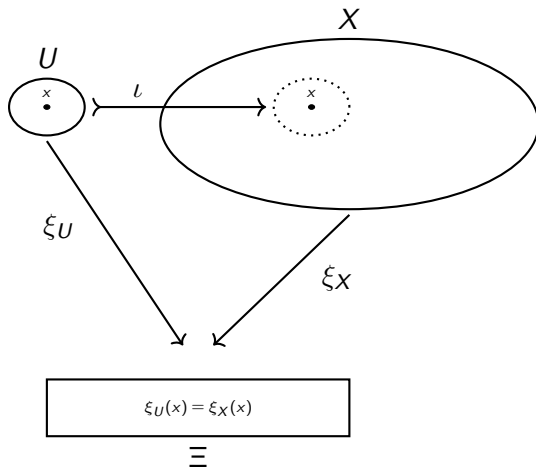
Let $\{\xi_X: X \rightarrow \Xi\}_{X \in \text{ob}\mathcal{C}}$ be a local state classifier.



The value of $\xi_X(x)$ can be calculated in arbitrarily small neighborhood of x .

Intuition: why local?

Let $\{\xi_X: X \rightarrow \Xi\}_{X \in \text{ob}\mathcal{C}}$ be a local state classifier.



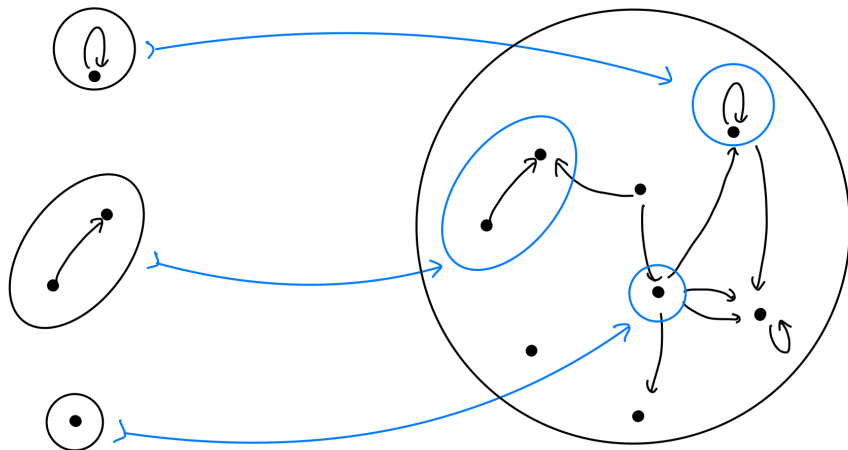
The value of $\xi_X(x)$ can be calculated in arbitrarily small neighborhood of x .

Table of Contents

- 1 Introduction
- 2 LSCs: Definition
- 3 LSCs: More Examples
- 4 Atmt: LSC of Σ -Set
- 5 Atmt: Covariety of Σ -Set

More Examples (1/4): directed graphs

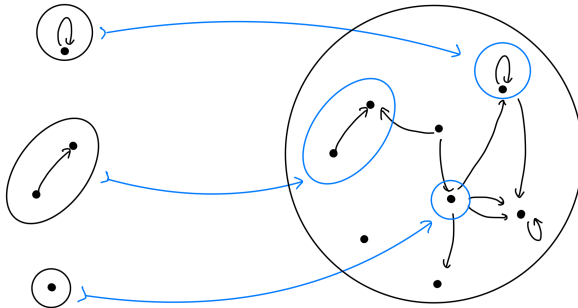
What is the local state classifier of the category of directed graphs $\mathbf{PSh}(\Rightarrow)$?



More Examples (1/4): directed graphs

The local state classifier of the topos of directed graphs $\mathbf{PSh}(\Rightarrow)$ is given by the "2-bouquet."

1. Ξ is [being a loop]  $\bullet$  [not being a loop].
2. ξ_X detects whether a given edge is a loop or not.



More Examples (2/4): group actions for a group G

What is the local state classifier of **PSh**(G)?

= What is the "local data" for each element of a G -set $X \curvearrowright G$?

More Examples (2/4): group actions for a group G

What is the local state classifier of **PSh**(G)?

= What is the "local data" for each element of a G -set $X \curvearrowright G$?

For the category of right G -actions **PSh**(G) for a group G ,

1. Ξ is the set of all subgroups $\text{Sub}_{\text{Grp}}(G)$ (equipped with the conjugate action).
2. ξ_X sends each element $x \in X$ to its stabilizer subgroup.

More Examples (3/4): presheaves

Let $\mathcal{C}$ be a small category. The local state classifier of the presheaf topos $\mathbf{PSh}(\mathcal{C})$ is the presheaf of quotient objects of representable presheaves:

$$\Xi: \mathcal{C}^{\text{op}} \rightarrow \mathbf{Set}: c \mapsto \{y(c) \twoheadrightarrow C\}$$

More Examples (3/4): presheaves

Let $\mathcal{C}$ be a small category. The local state classifier of the presheaf topos $\mathbf{PSh}(\mathcal{C})$ is the presheaf of quotient objects of representable presheaves:

$$\Xi: \mathcal{C}^{\text{op}} \rightarrow \mathbf{Set}: c \mapsto \{y(c) \twoheadrightarrow C\}$$

Remark (Dual similarity with the Subobject classifier)

It is a kind of dual to the subobject classifier, which is the presheaf of subobjects of representable presheaves:

$$\Omega: \mathcal{C}^{\text{op}} \rightarrow \mathbf{Set}: c \mapsto \{y(c) \hookleftarrow C\}$$

More Examples (4/4): Other categories

The following categories have LSCs.

- ▶ The category of simplicial sets **sSet**
- ▶ The topos of Joyal's species **FinSet**^{FinSet₀} (cf. [Fiore, Galal, and Paquet 2024])
- ▶ The category of coalgebraic automata **Coalg**_{2x Σ} = **Atmt** _{Σ} .
- ▶ The Schanuel topos **Cont**(Aut_{Set}($\mathbb{N}$)) $\cong$ **Sh**(**FinSet**_{mono}^{op}, $J_{\text{at.}}$)
- ▶ **Ab**, **SemiLattice**, and any other categories of algebras.
- ▶ Any Grothendieck topoi including **Sh**(X) for a topological space X .
- ▶ Any locally presentable categories, including **Poset**, **Cat**, ...

Table of Contents

- 1 Introduction
- 2 LSCs: Definition
- 3 LSCs: More Examples
- 4 Atmt: LSC of Σ -Set
- 5 Atmt: Covariety of Σ -Set

Σ -sets

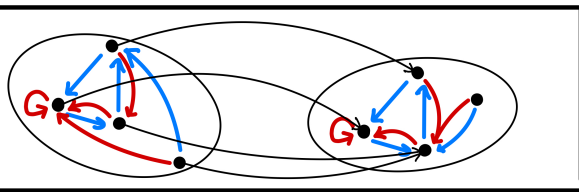
We fix a finite set Σ . A Σ -**set** is a pair (Q, δ) of a set Q and a function $\delta: Q \times \Sigma \rightarrow Q$.

Definition

Σ -**Set** is defined by $\Sigma\text{-Set} := \mathbf{PSh}(\Sigma^*)$.

$$\Sigma = \{a, b\}$$

Σ -Set



The Σ -set of right congruences

A **right congruence** is an equivalence relation $\sim$ on Σ^* such that

$$\forall u, v, w \in \Sigma^*, u \sim v \implies uw \sim vw.$$

Definition (The Σ -set of right congruences Ξ)

Element Right congruence (= quotient object of the representable Σ^* .)

Σ -action $u (\sim * w) v \iff wu \sim wv.$

Right congruences form LSC

For any Σ -set (Q, δ) , we define a morphism

$$\xi_{(Q, \delta)}: (Q, \delta) \rightarrow \Xi$$

by sending $q \in Q$ to the right congruence $\xi_{(Q, \delta)}(q) \in \Xi$:

$$u \xi_{(Q, \delta)}(q) v \iff q * u = q * v.$$

Proposition (LSC of $\Sigma\text{-Set}$, Universality of right congruences)

The Σ -set of right congruences Ξ , together with the cocone $\{\xi_{(Q, \delta)}: (Q, \delta) \rightarrow \Xi\}_{(Q, \delta) \in \text{ob}(\Sigma\text{-Set})}$ is the local state classifier of $\Sigma\text{-Set}$.

Toy example of LSC calculation

Example (The Σ -set of languages **Lan**)

Element Language

Σ -action $L * w := \{v \in \Sigma^* \mid wv \in L\}$ (Brzozowski derivative)

Universality $\Sigma\text{-Set}((Q, \delta), \mathbf{Lan}) \cong \mathcal{P}(Q)$

What's the cocone component $\xi_{\mathbf{Lan}}: \mathbf{Lan} \rightarrow \Xi$?

Toy example of LSC calculation

Example (The Σ -set of languages **Lan**)

Element Language

Σ -action $L * w := \{v \in \Sigma^* \mid wwv \in L\}$ (Brzozowski derivative)

Universality Σ -**Set**((Q, δ), **Lan**) $\cong \mathcal{P}(Q)$

What's the cocone component $\xi_{\mathbf{Lan}}: \mathbf{Lan} \rightarrow \Xi$?

Proposition (A categorical derivation of Nerode congruence)

For any language L , $\xi_{\mathbf{Lan}}(L)$ coincides with the **Nerode congruence** of L .

$$u \xi_{\mathbf{Lan}}(L) v \iff L * u = L * v \iff (\forall w \in \Sigma^* \quad uw \in L \iff vw \in L)$$

Table of Contents

- 1 Introduction
- 2 LSCs: Definition
- 3 LSCs: More Examples
- 4 Atmt: LSC of Σ -Set
- 5 Atmt: Covariety of Σ -Set

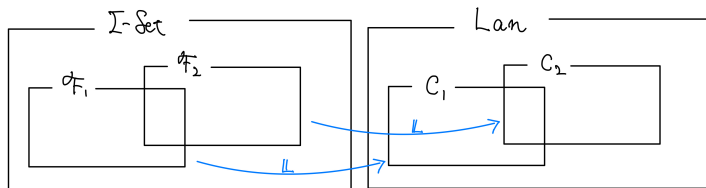
Idea: Σ -set classes $\rightarrow$ language classes

From a full subcategory $\mathcal{F} \subset \Sigma\text{-Set}$, we can consider the **induced language class** $\mathbb{L}_{\mathcal{F}} \subset \mathbf{Lan}$.

$$\mathbb{L}_{\mathcal{F}} := \{L \in \mathbf{Lan} \mid \text{some } \Sigma\text{-set in } \mathcal{F} \text{ recognizes } L\}.$$

Example (regular languages)

$\Sigma\text{-FinSet}$ is the category of finite Σ -sets. $\mathbb{L}_{\Sigma\text{-FinSet}}$ is the set of regular languages.



Q: What kind of full subcategories should we consider?

It suffices to consider **nice full subcategories**. but which?

In category theory, we tend to consider reflective subcategories:

- ▶ Subtopoi? (= lex reflective subcat)
- ▶ Birkhoff's HSP varieties? (= subquotient-closed reflective sub)
- ▶ Orthogonality classes?

¹This is the reason why it's related to the open problem of **quotient topoi**.

Q: What kind of full subcategories should we consider?

It suffices to consider **nice full subcategories**. but which?

In category theory, we tend to consider reflective subcategories:

- ▶ Subtopoi? (= lex reflective subcat)
- ▶ Birkhoff's HSP varieties? (= subquotient-closed reflective sub)
- ▶ Orthogonality classes?

We need to go the "opposite" direction¹, coreflective ones!

¹This is the reason why it's related to the open problem of **quotient topoi**.

A: Covariety of $\Sigma\text{-Set}$

Definition (Covariety)

A **covariety**² of $\Sigma\text{-Set}$ is a full subcategory of $\Sigma\text{-Set}$ that is closed under

- ▶ subobjects,
- ▶ quotient objects, and
- ▶ coproducts.

Every covariety is **coreflective**.

Lemma (Covarieties are enough.)

For any full subcategory $\mathcal{F} \subset \Sigma\text{-Set}$, there exists the minimum covariety $\overline{\mathcal{F}} \subset \Sigma\text{-Set}$ that contains $\mathcal{F}$. Furthermore, we have $\mathbb{L}_{\mathcal{F}} = \mathbb{L}_{\overline{\mathcal{F}}}$.

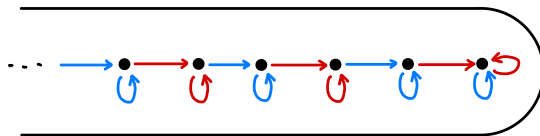
²This terminology is due to Peter Johnstone. This is defined for any Grothendieck topos.

Example: What is $\overline{\Sigma\text{-FinSet}}$? Orbit-finiteness

Definition ($\Sigma\text{-Set}_{\text{o.f.}}$)

A Σ -set $(Q, \delta: Q \times \Sigma \rightarrow Q)$ is said to be **orbit-finite** if, for any state $q \in Q$, its orbit $q\Sigma^* := \{qw \mid w \in \Sigma^*\}$ is a finite set.

Let $\Sigma\text{-Set}_{\text{o.f.}}$ denote the full subcategory of $\Sigma\text{-Set}$ that consists of all orbit-finite Σ -sets.



Example

$\overline{\Sigma\text{-FinSet}} = \Sigma\text{-Set}_{\text{o.f.}} \hookrightarrow \Sigma\text{-Set}$ is a covariety such that $\mathbb{L}_{\Sigma\text{-Set}_{\text{o.f.}}} = \mathbf{Reg}$.

covariety classification (1/2)

But how can we calculate covarieties of Σ -**Set**? (Even the smallness of the number of them is non-trivial a priori.)

covariety classification (1/2)

But how can we calculate covarieties of Σ -**Set**? (Even the smallness of the number of them is non-trivial a priori.)

Theorem (Classification of covarieties)

*For any Grothendieck topos $\mathcal{E}$, (in particular $\mathcal{E} = \Sigma$ -**Set**,) there is a bijective correspondence between*

- ▶ *covarieties of $\mathcal{E}$, and*
- ▶ *upward closed^a subobjects $U \mapsto \Xi$.*

This theorem reduces the study of covarieties of Σ -**Set** to the word combinatorics of congruences Ξ .

^a Ξ admits a unique order structure such that $\xi_X(x) \wedge \xi_Y(y) = \xi_{X \times Y}(x, y)$.

covariety classification (2/2)

For a given upward closed subobject $U \rhd \Xi$, the corresponding covariety $\mathcal{F}_U$ is given by

$$\text{ob}(\mathcal{F}_U) := \left\{ X \in \text{ob}(\mathcal{E}) \left| \begin{array}{ccc} & & U \\ & \nearrow \Xi & \downarrow \\ X & \xrightarrow{\xi_X} & \Xi \end{array} \right. \right\}.$$

Example $(\Sigma\text{-}\mathbf{Set}_{\text{o.f.}} \xleftrightarrow{\text{corresponding}} U_{\text{o.f.}})$

The covariety $\Sigma\text{-}\mathbf{Set}_{\text{o.f.}}$ corresponds to the subobject $U_{\text{o.f.}} := \{\sim \in \Xi \mid \Sigma^*/\sim \text{ is finite}\}.$

What I cannot talk: Other examples of calculation with LSC

≡ translates abstract definitions into concrete word combinatorics:

What I cannot talk: Other examples of calculation with LSC

Ξ translates abstract definitions into concrete word combinatorics:

- ▶ **Myhill-Nerode theorem** $\mathbb{L}_{\mathcal{F}} = \xi_{\mathbf{Lan}}^{-1}(U)$ for each covariety $\mathcal{F}$.
 - ▶ $\mathbb{L}_{\Sigma\text{-Set}_{\text{o.f.}}} = \xi_{\mathbf{Lan}}^{-1}(U_{\text{o.f.}})$ (cf. $\xi_{\mathbf{Lan}}: \mathbf{Lan} \rightarrow \Xi$ is the nerode congruence.)

What I cannot talk: Other examples of calculation with LSC

Ξ translates abstract definitions into concrete word combinatorics:

- ▶ **Myhill-Nerode theorem** $\mathbb{L}_{\mathcal{F}} = \xi_{\mathbf{Lan}}^{-1}(U)$ for each covariety $\mathcal{F}$.
 - ▶ $\mathbb{L}_{\Sigma\text{-Set}_{\text{o.f.}}} = \xi_{\mathbf{Lan}}^{-1}(U_{\text{o.f.}})$ (cf. $\xi_{\mathbf{Lan}}: \mathbf{Lan} \rightarrow \Xi$ is the nerode congruence.)
- ▶ **Syntactic congruence** $\cong_L = \bigwedge_{w \in \Sigma^*} (\xi_{\mathbf{Lan}}(L) * w)$ in Ξ .
 - ▶ We need $\xi_{\Xi}: \Xi \rightarrow \Xi$ to capture the behavior of $\xi_{\mathbf{Lan}}(L)$.

What I cannot talk: Other examples of calculation with LSC

Ξ translates abstract definitions into concrete word combinatorics:

- ▶ **Myhill-Nerode theorem** $\mathbb{L}_{\mathcal{F}} = \xi_{\mathbf{Lan}}^{-1}(U)$ for each covariety $\mathcal{F}$.
 - ▶ $\mathbb{L}_{\Sigma\text{-Set}_{\text{o.f.}}} = \xi_{\mathbf{Lan}}^{-1}(U_{\text{o.f.}})$ (cf. $\xi_{\mathbf{Lan}}: \mathbf{Lan} \rightarrow \Xi$ is the nerode congruence.)
- ▶ **Syntactic congruence** $\cong_L = \bigwedge_{w \in \Sigma^*} (\xi_{\mathbf{Lan}}(L) * w)$ in Ξ .
 - ▶ We need $\xi_{\Xi}: \Xi \rightarrow \Xi$ to capture the behavior of $\xi_{\mathbf{Lan}}(L)$.
- ▶ **Syntactic monoids** $\pi_1(C)$.
 - ▶ Prodiscrete for arbitrary C : **work in progress; not established**.
 - ▶ $\pi_1(\mathbf{Reg}) = \widehat{\Sigma^*}$: the profinite monoid of profinite words
 - ▶ $\pi_1(\{L\}) = \Sigma^*/\cong_L \iff \xi_{\Xi}(\xi_{\mathbf{Lan}}(L)) \in U_{\text{o.f.}} \iff \xi_{\mathbf{Lan}}(L) \in U_{\text{o.f.}}$
- ▶ **Dyck example** (**work in progress; not established**):
 $\pi_1(\{\text{well-parenthesized}\}) = \langle a, b \mid ab = 1 \rangle \sqcup \{*\}$ with a non-trivial topology.

What I cannot talk: Other examples of calculation with LSC

Ξ translates abstract definitions into concrete word combinatorics:

- ▶ **Myhill-Nerode theorem** $\mathbb{L}_{\mathcal{F}} = \xi_{\mathbf{Lan}}^{-1}(U)$ for each covariety $\mathcal{F}$.
 - ▶ $\mathbb{L}_{\Sigma\text{-Set}_{\text{o.f.}}} = \xi_{\mathbf{Lan}}^{-1}(U_{\text{o.f.}})$ (cf. $\xi_{\mathbf{Lan}}: \mathbf{Lan} \rightarrow \Xi$ is the nerode congruence.)
- ▶ **Syntactic congruence** $\cong_L = \bigwedge_{w \in \Sigma^*} (\xi_{\mathbf{Lan}}(L) * w)$ in Ξ .
 - ▶ We need $\xi_{\Xi}: \Xi \rightarrow \Xi$ to capture the behavior of $\xi_{\mathbf{Lan}}(L)$.
- ▶ **Syntactic monoids** $\pi_1(C)$.
 - ▶ Prodiscrete for arbitrary C : **work in progress; not established**.
 - ▶ $\pi_1(\mathbf{Reg}) = \widehat{\Sigma^*}$: the profinite monoid of profinite words
 - ▶ $\pi_1(\{L\}) = \Sigma^*/\cong_L \iff \xi_{\Xi}(\xi_{\mathbf{Lan}}(L)) \in U_{\text{o.f.}} \iff \xi_{\mathbf{Lan}}(L) \in U_{\text{o.f.}}$
- ▶ **Dyck example** (**work in progress; not established**):
 $\pi_1(\{\text{well-parenthesized}\}) = \langle a, b \mid ab = 1 \rangle \sqcup \{*\}$ with a non-trivial topology.
- ▶ **Syntactic topos** $\mathcal{T}_C$ for each $C \subset \mathbf{Lan}$.
 - ▶ Classification of hyperconnected quotient topoi (= covarieties closed under finite products)[Hora 2024a]
 - ▶ **Work in progress; not established**: $\mathcal{T}_C \simeq \mathbf{Cont}(\pi_1(C))$
 (semi-Galois theory; cf. [Uramoto 2025]).

References

-  Fiore, Marcelo, Zeinab Galal, and Hugo Paquet (2024). “Stabilized profunctors and stable species of structures”. In: *Logical Methods in Computer Science* 20.
-  Hora, Ryuya (2024a). “Internal parameterization of hyperconnected quotients”. In: *Theory and Applications of Categories* 42.11, pp. 263–313.
-  — (2024b). “Topoi of automata I: Four topoi of automata and regular languages”. In: *arXiv:2411.06358*.
-  — (n.d.). *Topoi of automata II: Hyperconnected geometric morphisms, syntactic monoids, and language classes (ongoing)*.
-  Hora, Ryuya, Yuhi Kamio, and Yuki Maehara (2025). *Lawvere’s fourth open problem: Levels in the topos of symmetric simplicial sets*. *arXiv: 2503.03439 [math.CT]*. URL: <https://arxiv.org/abs/2503.03439>.
-  Kamio, Yuhi and Ryuya Hora (2024). “A solution to the first Lawvere’s problem A Grothendieck topos that has a proper class many quotient topoi”. In: *arXiv:2407.17105*.
-  Lawvere, F. William (2009). *Open problems in topos theory*. Can be found via “<https://lawverearchives.com>”.
-  Uramoto, Takeo (2025). “Semi-galois Categories I: The classical Eilenberg variety theory”. In: *Journal of Pure and Applied Algebra* 229.2, p. 107863.