Overview of the Preliminary Results

Philip Kaluđerčić

philip.kaludercic@fau.de

26Jul24, typeset on August 8, 2024

In this document I intend to briefly summarise the work related to my *work-in-progress* master's thesis ("Semantics of Categorical Nondeterministic Automata in a Topos") up until this point.

1 Problem Statement

The semantics of a (non-deterministic) automaton is the language, i.e. the words it accepts as input. Besides the conventional set-theoretical formalisation, we want to consider the following:

1. In the context of a topos $\mathscr E,$ a coalgebra for a non-deterministic automaton is

 $FQ = \Omega \times \mathbf{P} \left(Q \right)^{\Sigma},$

where Ω is the subobject classifier and $\mathbf{P}(-)$ is the covariant power-object functor.

The accepted words of this construction are describable using " \mathcal{EM} -style semantics" [JSS12], that require the functor F to be the composition of a functor and a monad. This is the case, for the $GQ = 2 \times Q^{\Sigma}$ and the $\mathbf{P}(-)$ (where unit and multiplication are analogous to the power-sets monad). A morphism $[\![-]\!]: Q \longrightarrow \Sigma^*$ then describes the accepted words starting from some state.

2. A "categorical nondeterministic automata" [FMU23] requires a category \mathscr{C} to have subobjects, epi-mono splits and finite limits (all of these properties are granted by toposes). The general idea is that we have an object $Q \in Ob(\mathscr{C})$ of states and $\Sigma \in Ob(\mathscr{C})$ of letters in our input alphabet. For each word-length, we consider subobjects of $(Q \times \Sigma \times Q)^n$, one designating the legal transitions, the other transitions that begin in an initial, and end in an accepting state. The "accepted runs" results from the pullback of the two aforementioned subobjects. By taking the image of the projection that extracts letters out from the runs, we have a description of the accepted words.

The question is if these two formulations are equivalent, specifically when assuming that the base category is a topos \mathscr{E} . This structure allows us to argue within the internal logic of \mathscr{E} , simplifying the proofs.

2 Necessary Background and Noteworthy Properties of Toposes

Definition 1 (Topos). A category \mathscr{E} is elementary topos [MM12, p. 161], has all finite limits and an object Ω (subobject), with a function that assigns each object A: \mathscr{E} to an object $\mathbf{P}A$: E, such that for any object B: \mathscr{E}

$$\operatorname{Sub}_{\mathscr{E}}(A) \cong \operatorname{Hom}_{\mathscr{E}}(A, \Omega)$$
 (1)

$$\operatorname{Hom}_{\mathscr{E}}(A \times B, \Omega) \cong \operatorname{Hom}_{\mathscr{E}}(B, \mathbf{P}A) \tag{2}$$

which is natural in A.

Definition 2 (Internal Logic). Admitted.

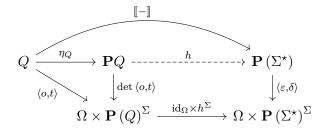


Figure 1: Overview of the structure required to define [-], given a determination det $\langle o, t \rangle$ of a NDA in \mathscr{E}

3 Coalgebraic Trace Semantics in a Topos

The semantics of a coalgebra is a morphism in the base category that maps a state to a representation of the accepted words. In a topos \mathscr{E} this means a coalgebra $FQ = \Omega \times Q^{\Sigma}$ has a morphism

$$\llbracket - \rrbracket : Q \longrightarrow \Sigma^{\star}. \tag{3}$$

Recalling that in **Sets** the carrier of the final coalgebra of $FQ = 2 \times Q^{\Sigma}$ is the powerset of words $\wp(\Sigma^*)$, we notice that the terminal arrow from an arbitrary coalgebra to the final coalgebra bears similarities with [-] from above.

3.1 Necessary Structure of \mathscr{E}

It remains open what Σ^* means. In **Sets** we would expect X^* to be the free monoid over a set X.

Definition 3 (Countably Extensive). If in a category $\mathscr{C} X^* \cong \coprod_{n \in \mathbb{N}} X^n$, then we call \mathscr{C} "countably extensive". [FMU23, p. 10]

For this to be possible, a category requires *countable* coproducts. This is something an elementary topos does not have by default, as it is only required to have *finite* limits. Therefore it appears we have to strengthen our assumptions and imbue \mathscr{E} with countable coproducts before proceeding.

Furthermore, it is necessary to define the extension of a function $f: A \longrightarrow X$ along a free monoid.

3.2 Definition of [-]

Definition 4 (Eilenberg-Moore category). Admitted.

Definition 5 (Eilenberg-Moore Law). Admitted.

How can we ensure the existence of an final coalgebra in \mathscr{E} , analogous to the canonical one in **Sets**? One approach is to translate the results of Jacobs, et. al. [JSS12], from **Sets** to \mathscr{E} . This states, that for an endofunctor G and a monad (T, η, μ) , for which a definition 5 $\rho: TG \Rightarrow GT$ holds¹, the coalgebra on G is the same as of \hat{G} , i.e. the lifted functor from \mathscr{E} to the Eilenberg-Moore category (c.f. definition 4) on T.

¹Note the meaning of this natural transformation in **Sets**: It states that given a set of deterministic automata, we can either check if any of these accept a word, or transform these into a single non-deterministic automata and run that.

In \mathscr{E} , we take $GQ = \Omega \times Q^{\Sigma}$ and $TQ = \mathbf{P}Q$. It is possible to define ρ component-wise,

$$\rho_Q \colon \mathbf{P}\left(\Omega \times Q^{\Sigma}\right) \longrightarrow \Omega \times \mathbf{P}Q^{\Sigma} \qquad \rho_Q(A) \coloneqq \left\langle \rho_Q^1, \rho_Q^2 \right\rangle \tag{4}$$

$$\rho_Q^1 \colon \mathbf{P}\left(\Omega \times Q^{\Sigma}\right) \longrightarrow \Omega \qquad \qquad \rho_Q^1(A) \coloneqq \exists \langle \varepsilon, \delta \rangle \in A. \, \varepsilon = \text{true} \tag{5}$$

$$\rho_Q^2 \colon \mathbf{P}\left(\Omega \times Q^{\Sigma}\right) \longrightarrow \mathbf{P}Q^{\Sigma} \qquad \qquad \rho_Q^2(A) \coloneqq \sigma \mapsto \{q \colon \mathbf{P}Q \mid \exists \langle \varepsilon, \delta \rangle \in A. \, \delta(\sigma) = q \} \tag{6}$$

Theorem 1 (\mathcal{EM} -law). There exists a natural transformation $\rho: TG \Rightarrow GT$, for which distributively of the unit

$$\rho_X \circ \eta_{FX} = F(\eta_X) \tag{7}$$

and of multiplication,

$$\rho_X \circ \mu_{FX} = F(\mu_X) \circ \rho_{\mathbf{P}X} \circ \mathbf{P}\rho_X \tag{8}$$

hold.

Proof. We can express the statement in the internal logic of \mathscr{E} .

Admitted.

Due to a bijective correspondence between the existence of ρ^2 and a lifting from \mathscr{E} to $\mathcal{EM}(\mathbf{P})$ [JSS12], we can now lift the coalgebra $Q \longrightarrow \Omega \times \mathbf{P}Q^{\Sigma}$ in \mathscr{E} to $\mathbf{P}Q \longrightarrow \Omega \times \mathbf{P}Q^{\Sigma}$ in $\mathscr{EM}(\mathbf{P})$. The terminal map in $\mathscr{EM}(\mathbf{P})$ turns out to be the intuitive definition,

$$h(\mathbb{Q}) = \left\{ \vec{\sigma} \in \Sigma^{\star} \mid \exists q \in \mathbb{Q}. \, o(\overline{t(q)}(\vec{\sigma})) = \text{true} \right\},\tag{9}$$

as also seen in fig. 1. The jump to an internal definition of

$$\llbracket q \rrbracket = \left\{ \vec{\sigma} \in \Sigma^{\star} \mid o(\overline{t(q)}(\vec{\sigma})) = \text{true} \right\}$$
(10)

is easy, just as arguing that the internal logic of ${\mathscr E}$ grants us

$$\llbracket - \rrbracket = h \circ \eta_Q. \tag{11}$$

4 Internal Description of a Categorical Automaton

The second description of a non-deterministic automaton requires a category $\mathcal E$ with certain structure:

- All finite limits,
- Arbitrary subobjects $m: S \rightarrow A$, i.e. equivalence classes of monos [MM12, p. 11],
- Epi-mono factorisation [MM12, p. 185],

but not properties like being countably extensive (definition 3). Any elementary topos satisfies these conditions.

We define a NDA, as a tuple $A = (Q, \Sigma, \delta, m_I, m_F)$, where $Q, \Sigma \in \mathsf{Ob}(\mathscr{E})$, and $\delta \colon Q \times \Sigma \times Q \longrightarrow$ (transitions), $m_I \colon I \longrightarrow Q$ (initial states) and $m_F \colon F \rightarrowtail Q$ (final states) are subobjects.

As \mathscr{E} is not necessarily countably extensive, meaning we cannot construct $\Sigma^* \cong \coprod_i \Sigma^i$, we define a language as a family of subobjects

$$L \coloneqq \left(m_n^{(L)} \colon L^{(n)} \rightarrowtail \Sigma^n \right)_{n \in \mathbb{N}}$$
(12)

instead of Σ^* . To define a language, we have to define the subobjects individually for each n.

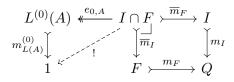


Figure 2: Commutative diagrams describing a categorical automaton fur n = 0

Figure 3: Commutative diagrams describing a categorical automaton for $n \ge 1$

4.1 Accepting the Empty Word

On a high-level, we know that a NDA accepts the empty word ϵ , if an initial state is also final. Intuitively, this is a subobject as well (the pullback of m_I and m_F), describing the "accepted runs" of length n = 0. Recall that in a topos, each subobject $m: S \rightarrow A$ corresponds to a predicate $\varphi = \text{char } S$ [MM12, p. 165]. This gives us a convenient, internal description of the accepting runs:

$$I \cap F = \{ q \in Q \mid (\operatorname{char} I)(q) \land (\operatorname{char} F)(q) \}$$
(13)

A map from $I \cap F$ to the accepting words $\Sigma^0 \cong 1$, as seen in fig. 2 has an image, with the internal description

$$L^{(0)}(A) \coloneqq \begin{cases} \{\epsilon\} & \text{if } I \cap F \not\cong \emptyset \\ \{ \} & \text{otherwise} \end{cases}$$
(14)

4.2 Accepting Non-Empty Words

For a non-empty word of length n, an accepting run is a sequence of n transitions, beginning in an initial state, connected by Σ s in δ and ending in a final state. An external description is given by fig. 3, where $\mathsf{AccRun}_{A}^{(n)}$ is the pullback of two subobjects into $(Q \times \Sigma \times Q)^{n}$,

- m_{δ}^{n} : an *n*-times product of δ , and
- $d_{n,A}$: an "injection" from $I \times (\Sigma \times Q)^{n-1} \times \Sigma \times F$. The domain of this map is by associativity

$$\underbrace{I \times \Sigma \times Q \times \cdots \times Q \times \Sigma \times F}_{n\text{-many }\Sigma},$$

i.e. a "chain" of transitions that are reordered into the intended form.

²The direction that interests us, is that we can define $\hat{G} = \rho_X \circ G$, such that it behaves as intended.

4.2.1 Internal Description of Accepting Runs

We can describe the accepted runs in the internal language of $\mathcal E,$ by formalising the above description

$$\mathfrak{A}_{n,A} = \left\{ a \in \delta^n \left| \underbrace{\pi_1(\pi_1(a)) \in I}_{\substack{\text{begins in an}\\\text{initial state}}} \wedge \underbrace{\pi_3(\pi_n(a)) \in F}_{\text{ends in final state}} \wedge \underbrace{\forall 1 \le i < n. \, \pi_3(\pi_i(a)) = \pi_1(\pi_{i+1}(a))}_{\text{all transitions are legal}} \right\}.$$
(15)

Theorem 2. $\mathfrak{A}_{n,A}$ is a pullback of $d_{n,A}$ and m_{δ}^n .

The statement involves the following subclaims:

- 1. There exist two morphisms $\overline{d}_{n,A} : \mathfrak{A}_{n,A} \longrightarrow \delta^n$ and $\overline{m}_{\delta}^n : \mathfrak{A}_{n,A} \longrightarrow I \times (\Sigma \times Q)^{n-1} \times \Sigma \times F$, for which $m_{\delta}^n \circ \overline{d}_{n,A} = \overline{m}_{\delta}^n \circ d_{n,A}$ holds, and
- 2. for any other object P that satisfies the UMP of a pullback for $d_{n,A}$ and m_{δ}^n , there is a unique morphism from P to $\mathfrak{A}_{n,A}$.

Proof. We consider the subclaims separately,

1. The morphisms are

$$d_{n,A}(a) = a \tag{16}$$

$$\overline{m}^{n}_{\delta}(a) = \left\langle \pi_{1} \circ \pi_{1}, \left\langle \pi_{2}, \pi_{3} \right\rangle^{n-1}, \pi_{2} \circ \pi_{n}, \pi_{3} \circ \pi_{n} \right\rangle(a).$$
(17)

It is easy to see that $\overline{d}_{n,A}$ composed with m_{δ}^n is just an injection into $(Q \times \Sigma \times Q)^n$, seeing as both are respectively just injections. As for $\overline{m}_{\delta}^n(a)$, we have to verify that

$$\vdash d_{n,A} \circ \overline{m}^n_{\delta} = m^n_{\delta} \circ d_{n,A} = \iota_{(Q \times \Sigma \times Q)^n},\tag{18}$$

which can be done in the internal logic of \mathscr{E}^3 :

$$\begin{aligned} &d_{n,A} \circ \overline{m}_{\delta}^{n} \\ &= \left(m_{I} \times (\mathrm{id}_{\Sigma} \times \Delta_{Q})^{n-1} \times \mathrm{id}_{\Sigma} \times m_{F} \right) \circ \left\langle \pi_{1} \circ \pi_{1}, \langle \pi_{2}, \pi_{3} \rangle^{n-1}, \pi_{2} \circ \pi_{n}, \pi_{3} \circ \pi_{n} \right\rangle \\ &= \left\langle m_{I} \circ \pi_{1} \circ \pi_{1}, (\mathrm{id}_{\Sigma} \times \Delta_{Q})^{n-1} \circ \langle \pi_{2}, \pi_{3} \rangle^{n-1}, \mathrm{id}_{\Sigma} \circ \pi_{2} \circ \pi_{n}, m_{F} \circ \pi_{3} \circ \pi_{n} \right\rangle \\ &= \left\langle m_{I} \circ \pi_{1} \circ \pi_{1}, \langle \pi_{2}, \Delta_{Q} \circ \pi_{3} \rangle^{n-1}, \pi_{2} \circ \pi_{n}, m_{F} \circ \pi_{3} \circ \pi_{n} \right\rangle \end{aligned}$$

We can drop the injections, knowing that for any $a \in \mathfrak{A}_{n,A}$, $\pi_1(\pi_1(a)) \in I$ and $\pi_3(\pi_n(a)) \in F$,

$$= \left\langle \pi_{1} \circ \pi_{1}, \langle \pi_{2}, \Delta_{Q} \circ \pi_{3} \rangle^{n-1}, \pi_{2} \circ \pi_{n}, \pi_{3} \circ \pi_{n} \right\rangle$$
$$\cong \left\langle \pi_{1} \circ \pi_{1}, \underset{i=1}{\overset{n-1}{\times}} \langle \pi_{2}, \pi_{3}, \pi_{3} \rangle \circ \pi_{i}, \pi_{2} \circ \pi_{n}, \pi_{3} \circ \pi_{n} \right\rangle$$
$$= \left\langle \pi_{1} \circ \pi_{1}, \underset{i=1}{\overset{n-1}{\times}} \langle \pi_{2} \circ \pi_{i}, \pi_{3} \circ \pi_{i}, \pi_{3} \circ \pi_{i} \rangle, \pi_{2} \circ \pi_{n}, \pi_{3} \circ \pi_{n} \right\rangle$$

As for any $a: \mathfrak{A}_{n,A}$, we know that $\forall 1 \leq i < n. \pi_3(\pi_i(a)) = \pi_1(\pi_{i+1}(a))$ is satisfied, we can use the associativity of products to infer

$$\cong \sum_{i=1}^{n} \langle \pi_2 \circ \pi_i, \pi_3 \circ \pi_i, \pi_3 \circ \pi_i \rangle = \sum_{i=1}^{n} \langle \pi_2, \pi_3, \pi_3 \rangle \circ \pi_i = \sum_{i=1}^{n} \pi_i$$
$$= \iota_{(Q \times \Sigma \times Q)^n} = m_{\delta}^n \circ \overline{d}_{n,A}$$

The choice of notation is a cludge for now, somethine else would be preferable 2. As it is known that toposes have all finite limits, hence

 $Pb(d_{n,A}, m_{\delta}^{n}) = \{ r \in (Q \times \Sigma \times Q)^{n} \mid r \in Im(d_{n,A}) \land r \in Im(m_{\delta}^{n}) \}$

is a valid, general description of a pullback in \mathscr{E} .

We can argue the case that $\mathfrak{A}_{n,A} \cong \mathrm{Pb}\left(d_{n,A}, m^n_{\delta}\right)$ in the internal logic of \mathscr{E} by extensionality over $r: (Q \times \Sigma \times Q)^n \vdash$

$$r \in \mathfrak{A}_{n,A}$$

$$\iff r \in \{a \in \delta^n \mid \ldots\}$$
(defn.)
$$\iff \pi_{1,1}(r) \in I \land \pi_{3,n}(r) \in F \land \forall 1 \leq i < n. \pi_{3,i}(r) = \pi_{1,i+1}(r) \land r \in \delta^n$$

$$\iff (\exists a: I \times \cdots \times F. d_{n,A}(a) = r) \land (\exists l: \delta^n. m_{\delta}^n(l) = r)$$
(†)
$$\iff r \in \{r \mid (\exists a. d_{n,A}(a) = r) \land (\exists l. m_{\delta}^n(l) = r)\}$$

$$\iff r \in \{r \mid \exists a. d_{n,A}(a) = r\} \land r \in \{r \mid \exists l. m_{\delta}^n(l) = r\}$$

$$\iff r \in \mathrm{Im}(d_{n,A}) \land r \in \mathrm{Im}(m_{\delta}^n)$$

$$\iff r \in \{r \in (Q \times \Sigma \times Q)^n \mid r \in \mathrm{Im}(d_{n,A}) \land r \in \mathrm{Im}(m_{\delta}^n)\}$$

$$\iff r \in \mathrm{Pb}(d_{n,A}, m_{\delta}^n)$$
(undefn.)

The (†) inference constitutes the intuitive yet critical step in this chain, specifically

$$\pi_{1,1}(r) \in I \land \pi_{3,n}(r) \in F \land \forall 1 \le i < n. \pi_{3,i}(r) = \pi_{1,i+1}(r)$$
$$\iff \exists a \colon I \times (\Sigma \times Q)^{n-1} \times \Sigma \times F. d_{n,A}(a) = r,$$
(20)

as $r \in \delta^n \iff \exists l \colon \delta^n . m_{\delta}^n(l) = r$ ought to be clear.

By considering $d_{n,A}$ component-wise $(m_I \times (\mathrm{id}_{\Sigma} \times \Delta_Q)^{n-1} \times \mathrm{id}_{\Sigma} \times m_F)$:

- m_I : By construction, this matches $\pi_{1,1}(r)$
- $(\mathrm{id}_{\Sigma} \times \Delta_{O})^{n-1}$: Recall that to the \cong -step in fig. 3 uses associativity of products to reparenthesise the product. Due to Δ_Q , the string of unparentheses objects will have $\pi_{3k} = \pi_{3k+1}$, for $1 \le k < n$, matching the internal description.
- id_{Σ} : The formula does not describe this, as there are no restrictions on what input constitutes a *legal* run.
- m_F : By construction, this matches $\pi_{3,n}(r)$

4.2.2**Internal Description of Accepted Words**

The image of a map $\pi_{n,A}$ from AccRun⁽ⁿ⁾ to Σ^n , that projects the Σ -components out may be defined as the compositions

$$p_{n,A} \circ \overline{m}_{\delta}^{(n)}, \text{ or } \pi_2^n \circ m_{\delta}^n \circ \overline{d}_{n,A},$$

$$(21)$$

as given in fig. 3, or simply given $\mathfrak{A}_{n,A}$ from eq. (15)

$$\pi_{n,A}(a) = \pi_2^n(a) \tag{22}$$

For a n > 0, the image of $\pi_{n,A}$ denotes the accepted words. In the internal language of \mathscr{E} , we can describe this on a high-level by

$$\operatorname{Im}(\pi_{n,A}) = \left\{ \vec{\sigma} \in \Sigma^n \mid \exists a \in \operatorname{AccRun}_A^{(n)} . \, \pi_{n,A}(a) = \vec{\sigma} \right\}$$
(23)

or by expanding definitions, $L^{(n)}(A) =$

$$\frac{\{\vec{\sigma} \in \Sigma^n \mid \exists a \in \delta^n. \pi_{1,1}(a) \in I \land \pi_{3,n}(a) \in F \land (\forall i < n. \pi_{3,i}(a) = \pi_{1,i+1}(a)) \land \pi_2^n(a) = \vec{\sigma} \}. (24)$$
ity its a statistic for $\langle f_1, \dots, f_n \rangle.$

This ar gument is not rigorous enough

Is there

point?

Cite a source or demonstrate why this is a pullback (for two

monos)

(19)

5 Comments & Considerations on the Future Work

Given the above, we want to relate eq. (10) for some $q \in Q$ to eq. (15). Both representations can share the state space Q and the input alphabet Σ . Yet note that the accepted words cannot be directly compared, as $\text{Im}(\pi_{n,A})$ describes only the accepted words of length n. Therefore, it is necessary to consider each n > 0 separately

$$\operatorname{Im}(\pi_{n,A}) \stackrel{?}{=} \left\{ \sigma^{n} \in \Sigma^{n} \mid o(t^{n}(q)(\sigma^{n})) = \operatorname{true} \right\},$$
(25)

whereas for n = 0

$$I \cap F \ncong \varnothing \iff o(q) = \text{true}.$$
 (26)

Point of Note: Im $(\pi_{n,A})$ depends on A (specifically the subobjects δ , I and F), while the coalgebraic representation involve o and t. Any further progress in proving the above equivalences depends on a reliable translation of the one into the other.

5.1 Categorical Automata into Coalgebras

Given δ , I and F we can define a coalgebra over the same state space Q, with the initial state I. The structure morphism $\langle o, t \rangle : Q \longrightarrow \Omega \times \mathbf{P} (Q)^{\Sigma}$ is easily construct the individual morphisms element-wise:

$$o(q) = q \in F$$

$$t(q) = \sigma \mapsto \{ q' \mid (q, \sigma, q') \in \delta \}$$
(27)
(28)

5.2 Coalgebras into Categorical Automata

Given a Coalgebra $\langle o, t \rangle$ in a state $q \in Q$, be can describe an equivalent automaton constituting

$$I = \{q\}\tag{29}$$

$$F = \left\{ q' \in Q \mid \exists \vec{\sigma} \in \Sigma^{\star}. \, \overline{t(q)}(\vec{\sigma}) = q' \implies o(q') = \text{true} \right\}$$
(30)

$$\delta = \left\{ \left(q', \sigma, q'' \right) \mid \exists \vec{\sigma} \in \Sigma^{\star} . \, \overline{t(q)}(\vec{\sigma}) = q' \implies t(q')(\sigma) = q'' \right\}$$
(31)

5.3 Equivalence of Descriptions

An intuitive condition for equivalence would be that converting a categorical automaton into a coalgebra and back (or vice versa) results in the same automaton. Note that this fails at least if there exist any states that are not accessible from the initial state, as eq. (31) will reconstruct only the "reachable" parts of δ . It is therefore at the very least necessary to ease the conditions and not require a full isomorphism.

References

- [FMU23] Florian Frank, Stefan Milius, and Henning Urbat. *Positive Data Languages*. 2023. arXiv: 2304.12947 [cs.FL].
- [JSS12] Bart Jacobs, Alexandra Silva, and Ana Sokolova. "Trace semantics via determinization". In: International Workshop on Coalgebraic Methods in Computer Science. Springer. 2012, pp. 109–129.
- [MM12] Saunders Mac Lane and Ieke Moerdijk. *Sheaves in geometry and logic: A first introduction to topos theory.* Springer Science & Business Media, 2012.

Would it be possible to have some third representation, e.g. in **Sets** and map that to the two other?