Higher-Dimensional Automata Theory

Uli Fahrenberg

LRE & EPITA Rennes, France

RP 2023

Languages of higher-dimensional automata

- Languages of Higher-Dimensional Automata. MSCS 2021
- Posets with Interfaces as a Model for Concurrency. I&C 2022
- A Kleene Theorem for Higher-Dimensional Automata. CONCUR 2022
- A Myhill-Nerode Theorem for Higher-Dimensional Automata. Petri Nets 2023
- Closure and Decision Properties for Higher-Dimensional Automata. ICTAC 2023

Today:

1. What are HDAs (and why should I be interested)?
2. What are languages of HDAs (and why should I be interested)?
3. What can I do with languages of HDAs (that I cannot do with other models)?

Nice people

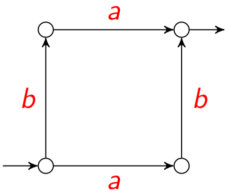
- Christian Johansen, NTNU
- Georg Struth, Sheffield
- Krzysztof Ziemiański, Warsaw

- Amazigh Amrane, Hugo Bazille, EPITA
- Emily Clement, Marie Fortin, Roman Kniazev, Jérémy Ledent, IRIF
- Loïc Hélouët, IRISA
- Safa Zouari, NTNU
- Eric Goubault, LIX

- See `https://ulifahrenberg.github.io/pomsetproject/` for more

Higher-dimensional automata

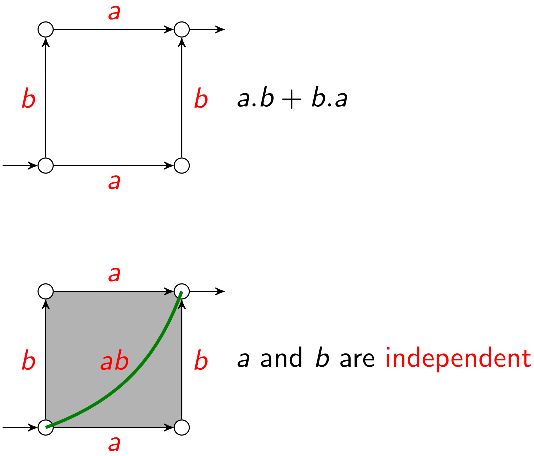
semantics of "*a* parallel *b*":



$a.b + b.a$

Higher-dimensional automata

semantics of "a parallel b":



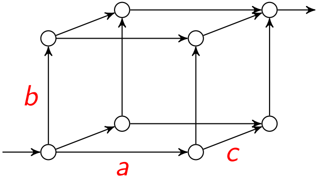
Higher-dimensional automata & concurrency

HDAs as a model for concurrency:

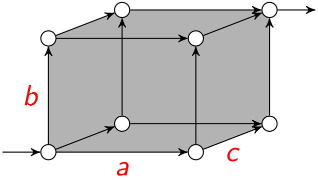
- points: states
- edges: transitions
- squares, cubes etc.: independency relations (concurrently executing events)
- two-dimensional automata $\cong$ asynchronous transition systems [Bednarczyk]
- [Pratt 1991, POPL], [van Glabbeek 1991, email message]

[van Glabbeek 2006, TCS]: Up to history-preserving bisimilarity, HDAs "generalize the main models of concurrency proposed in the literature" (notably, event structures and Petri nets)

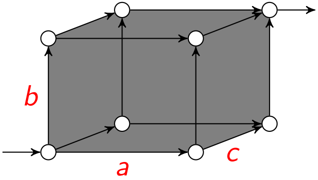
Examples



no concurrency



two out of three



full concurrency

Precubical sets and higher dimensional automata

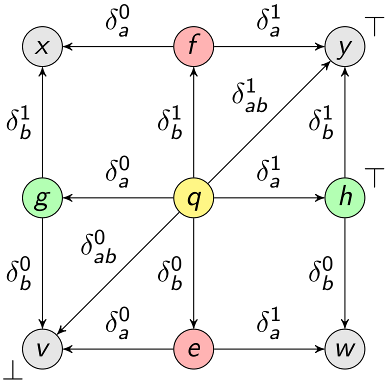
An loset is a finite, ordered and Σ-labelled set. (a list of events)

A precubical set X consists of:

- A set of cells X
- Every cell $x \in X$ has an loset $\mathrm{ev}(x)$ (list of events active in x)
- We write $X[U] = \{x \in X \mid \mathrm{ev}(x) = U\}$ for an loset U (cells of type U)
- For every loset U and $A \subseteq U$ there are:
  upper face map $\delta_A^1 : X[U] \to X[U - A]$ (terminating events A)
  lower face map $\delta_A^0 : X[U] \to X[U - A]$ ("unstarting" events A)
- Precubical identities: $\delta_A^\mu \delta_B^\nu = \delta_B^\nu \delta_A^\mu$ for $A \cap B = \emptyset$ and $\mu, \nu \in \{0, 1\}$

A higher dimensional automaton (HDA) is a finite precubical set X with start cells $\bot \subseteq X$ and accept cells $\top \subseteq X$ (not necessarily vertices)

Example



$X[\emptyset] = \{v, w, x, y\}$

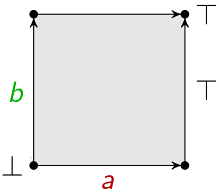
$X[a] = \{e, f\}$

$X[b] = \{g, h\}$

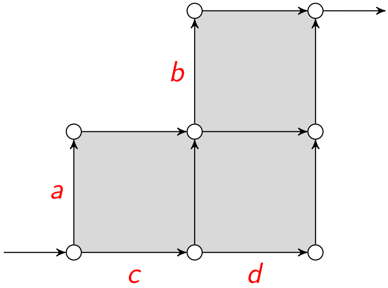
$X[ab] = \{q\}$

$\bot_X = \{v\}$

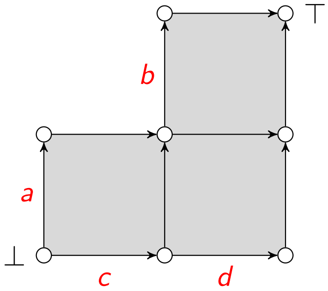
$\top_X = \{h, y\}$



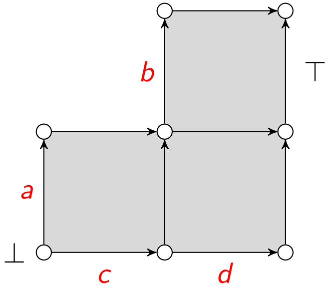
More interesting



More interesting



More interesting



Precubical sets as presheaves

A presheaf over a category $\mathcal{C}$ is a functor $\mathcal{C}^{\mathsf{op}} \to \mathsf{Set}$ (contravariant functor on $\mathcal{C}$)

The precube category $\square$ has (iso classes of) losets as objects.
Morphisms are coface maps $d_{A,B} : U \to V$, where

- $A, B \subseteq V$ are disjoint subsets,
- $U \simeq V - (A \cup B)$ are isomorphic losets,
- $d_{A,B} : U \to V$ is a unique order and label preserving map with image $V - (A \cup B)$.

Composition of coface maps $d_{A,B} : U \to V$ and $d_{C,D} : V \to W$ is

$$d_{\partial(A) \cup C, \partial(B) \cup D} : U \to W,$$

where $\partial : V \to W - (C \cup D)$ is the loset isomorphism.

Intuitively, $d_{A,B}$ terminates events B and "unstarts" events A.

- precubical sets: presheaves over $\square$

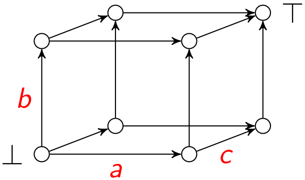
Languages of HDAs

- automata have languages
- HDAs don't (hitherto)
- (focus has been on geometric and topological aspects)

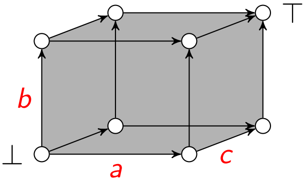
- automata and language theory is the very basis of computer science
- happy mix of operational and algebraic theory
- glue provided by Kleene and Myhill-Nerode theorems (among others)

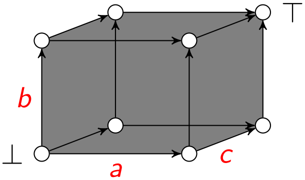
- Let's go!

Examples



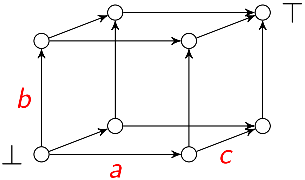
$$L_1 = \{abc, acb, bac, bca, cab, cba\}$$



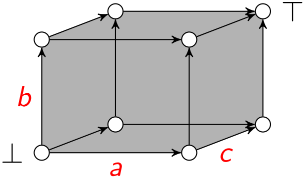


$$L_3 = \left\{ \begin{pmatrix} a \\ b \\ c \end{pmatrix}, \dots \right\}$$

Examples

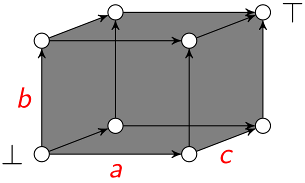


$$L_1 = \{abc, acb, bac, bca, cab, cba\}$$



$$L_2 = \left\{ \begin{pmatrix} a \\ b \to c \end{pmatrix}, \begin{pmatrix} a \\ c \to b \end{pmatrix}, \begin{pmatrix} b \\ a \to c \end{pmatrix}, \begin{pmatrix} b \\ c \to a \end{pmatrix}, \begin{pmatrix} c \\ a \to b \end{pmatrix}, \begin{pmatrix} c \\ b \to a \end{pmatrix} \right\} \cup L_1$$

sets of pomsets



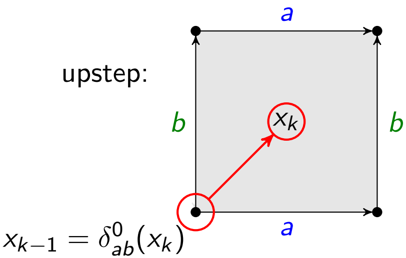
$$L_3 = \left\{ \begin{pmatrix} a \\ b \\ c \end{pmatrix} \right\} \cup L_2$$

Computations of HDAs

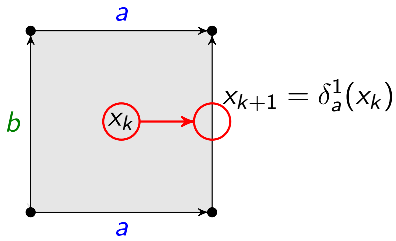
A path on an HDA X is a sequence $(x_0, \phi_1, x_1, \dots, x_{n-1}, \phi_n, x_n)$ such that for every k, (x_{k-1}, ϕ_k, x_k) is either

- $(\delta^0_A(x_k), \nearrow^A, x_k)$ for $A \subseteq \mathrm{ev}(x_k)$ or (upstep: start A)
- $(x_{k-1}, \searrow_B, \delta^1_B(x_{k-1}))$ for $B \subseteq \mathrm{ev}(x_{k-1})$ (downstep: terminate B)

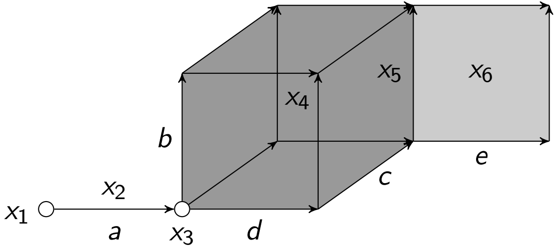
upstep:



downstep:

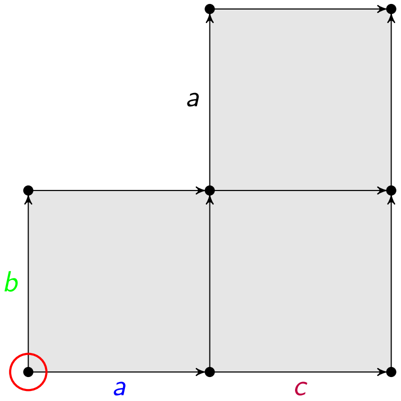


Example

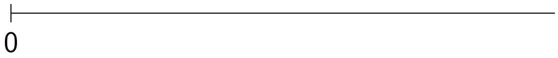


$$\big(x_1 \nearrow^{a} x_2 \searrow_{a} x_3 \nearrow^{\{b,c,d\}} x_4 \searrow_{\{c,d\}} x_5 \nearrow^{e} x_6\big)$$

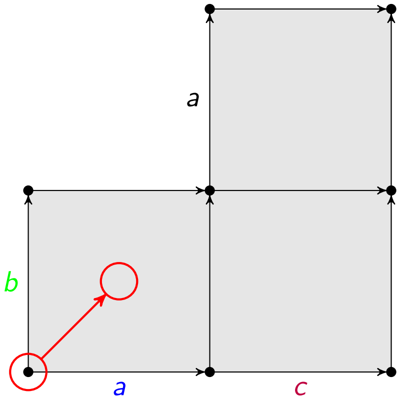
Event ipomset of a path



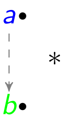
Lifetimes of events



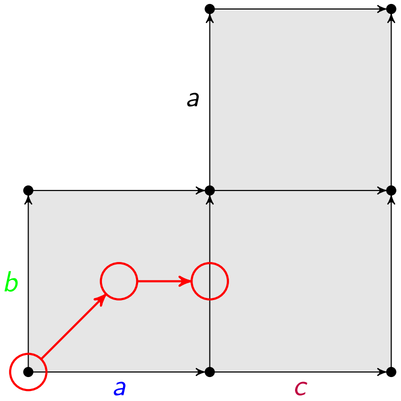
Event ipomset of a path



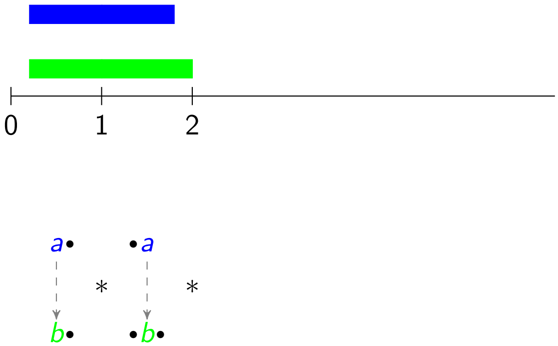
Lifetimes of events



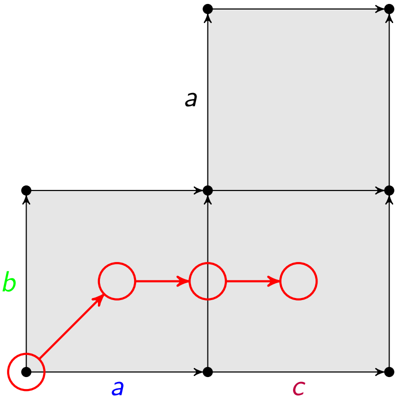
Event ipomset of a path



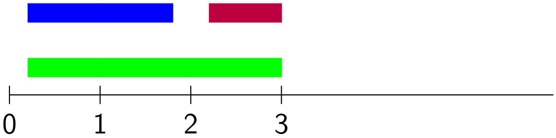
Lifetimes of events

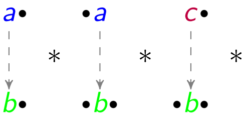


Event ipomset of a path

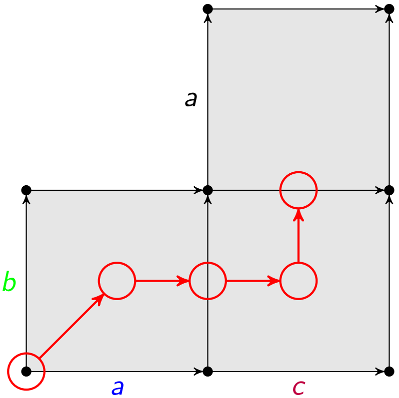


Lifetimes of events

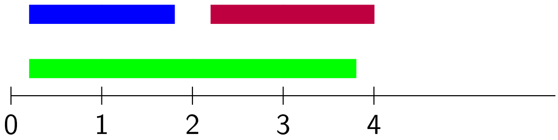


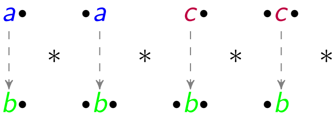


Event ipomset of a path

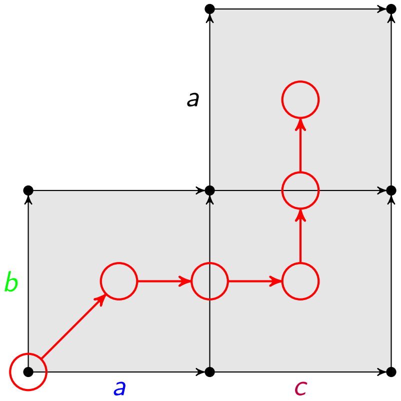


Lifetimes of events

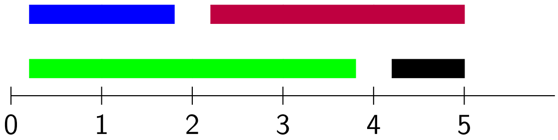


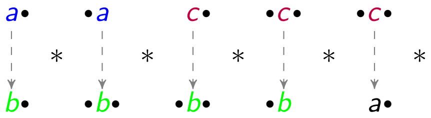


Event ipomset of a path

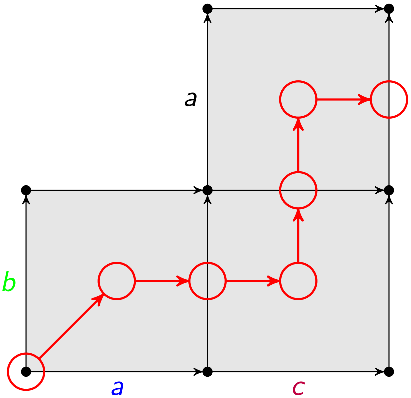


Lifetimes of events

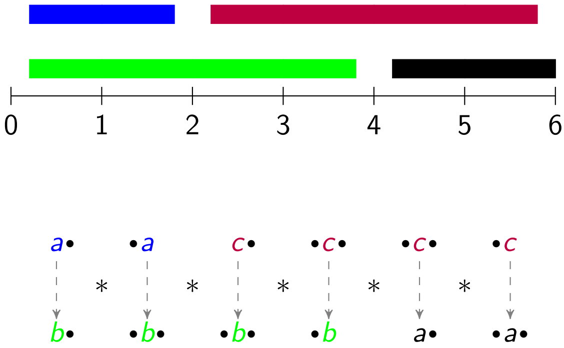




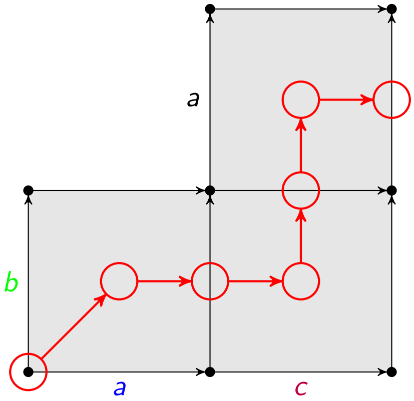
Event ipomset of a path



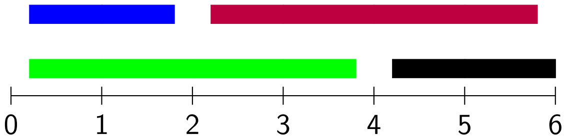
Lifetimes of events



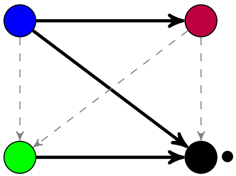
Event ipomset of a path



Lifetimes of events



Event ipomset

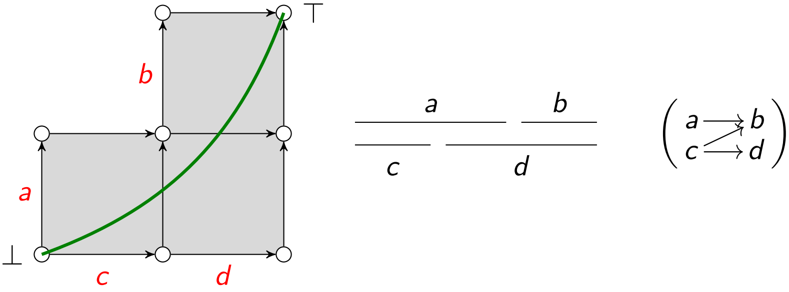


not series-parallel!

Are all pomsets generated by HDAs?

No, only (labeled) interval orders

- Poset $(P, \leq)$ is an interval order iff it has an interval representation:
  - a set $I = \{[l_i, r_i]\}$ of real intervals
  - with order $[l_i, r_i] \preceq [l_j, r_j]$ iff $r_i \leq l_j$
  - and an order isomorphism $(P, \leq) \leftrightarrow (I, \preceq)$
- [Fishburn 1970]



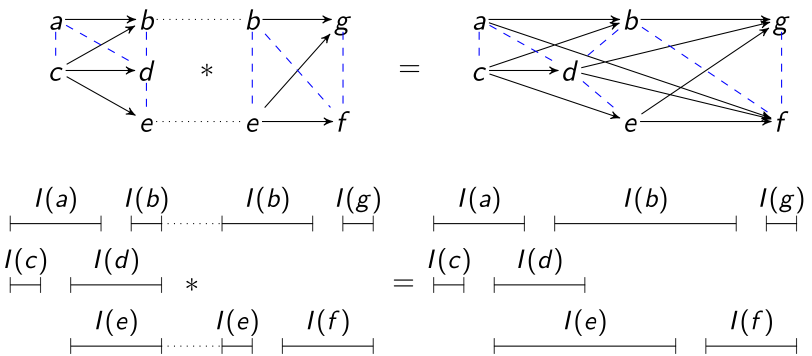
Pomsets with interfaces

Definition (Ipomset)

A pomset with interfaces (and event order): $(P, <, \dashrightarrow, S, T, \lambda)$:

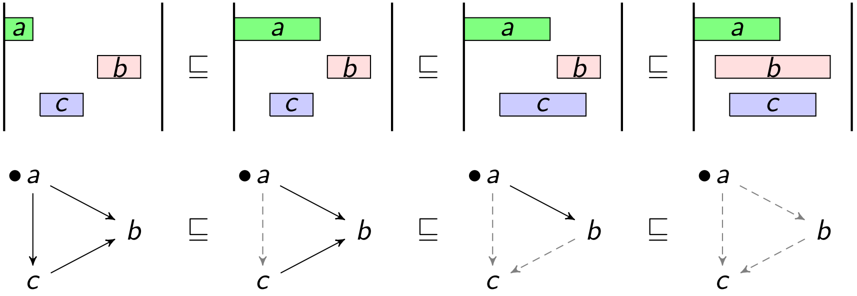
- finite set P;
- two partial orders $<$ (precedence order), $\dashrightarrow$ (event order)
  - s.t. $< \cup \dashrightarrow$ is a *total relation*;
- $S, T \subseteq P$ source and target interfaces
  - s.t. S is $<$-minimal, T is $<$-maximal.

Composition of ipomsets



- Gluing $P * Q$: P before Q, except for interfaces (which are identified)
- Parallel composition $P \parallel Q$: P above Q (disjoint union)

Subsumption



P refines Q / Q subsumes P / $P \sqsubseteq Q$ iff

- P and Q have same interfaces
- P has more $<$ than Q
- Q has more $\dashrightarrow$ than P

Languages of HDAs

Definition

The language of an HDA X is the set of event ipomsets of all accepting paths:

$$L(X) = \{\mathrm{ev}(\pi) \mid \pi \in \mathrm{Paths}(X), src(\pi) \in \bot_X, tgt(\pi) \in \top_X\}$$

- $L(X)$ contains only interval-order ipomsets
- and is closed under subsumption

Path objects

Important tool:

Proposition

For any interval-order ipomset P there exists an HDA $\square^P$ for which $L(\square^P) = \{P\}{\downarrow}$.

Lemma

For any HDA X and ipomset P, $P \in L(X)$ iff $\exists f : \square^P \to X$.

From Petri nets to automata

Definition (to fix notation)

A Petri net $N = (S, T, F)$: sets S of places and T of transitions; $S \cap T = \emptyset$;
$F : S \times T \cup T \times S \to \mathbb{N}$ weighted flow relation.

- for $t \in T$ let ${}^\bullet t, t^\bullet : S \to \mathbb{N}$ be ${}^\bullet t(s) = F(s, t)$, $t^\bullet(s) = F(t, s)$

Definition (standard)

The reachability graph of Petri net $N = (S, T, F)$ is the weighted graph $[\![N]\!]_1 = (V, E)$ given by $V = \mathbb{N}^S$ and $E = \{(m, t, m') \in V \times T \times V \mid {}^\bullet t \leq m, m' = m - {}^\bullet t + t^\bullet\}$.

From Petri nets to HDAs

- This goes back to [van Glabbeek 2006, TCS]
- Let $N = (S, T, F)$ be a Petri net
- Define $X = \mathbb{N}^S \times T^*$
- and $\text{ev} : X \to \square$ by $\text{ev}(m, \tau) = \tau$
- For $x = (m, \tau) \in X[\tau]$ with $\tau = (t_1, \dots, t_n)$ non-empty and $i \in \{1, \dots, n\}$, define

$$\delta^0_{t_i}(x) = (m + {}^\bullet t_i, (t_1, \dots, t_{i-1}, t_{i+1}, \dots, t_n))$$
$$\delta^1_{t_i}(x) = (m + t_i^\bullet, (t_1, \dots, t_{i-1}, t_{i+1}, \dots, t_n))$$

$\Rightarrow$ precubical set $[\![N]\!] = X$

- 0-cells are markings
- in an n-cell, n transitions are active concurrently
- auto-concurrency; collective token interpretation

Properties

Lemma

The reachability graph of N is isomorphic to the 1-skeleton of $[\![N]\!]$*:* $[\![N]\!]_1 \cong [\![N]\!]^{\leq 1}$.

Lemma

If N is bounded, then the essential part $\mathrm{ess}([\![N]\!])$ *is finite.*

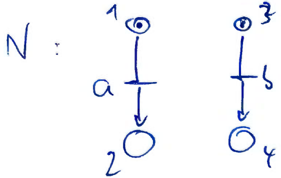
Conjecture

The concurrent step reachability graph of N is "essentially isomorphic" to the symmetrization of $[\![N]\!]$.

Conjecture

The language of $[\![N]\!]$ is "essentially the same" as the interval language of N.

Examples



$S = \{p_1, \dots, p_4\}$ $T = \{a, b\}$

$F = \{(p_1, a), (a, p_2), (p_3, b), (b, p_4)\} \mapsto 1$

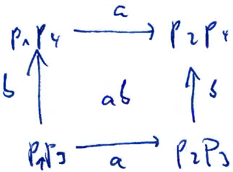
$i = \{p_1, p_3\} \mapsto 1$

$X :$ $X[\emptyset] = \{p_1p_3, p_2p_3, p_1p_4, p_2p_4\}$

$\uparrow$ $X[a] = \{(p_3, a), (p_4, a)\}$

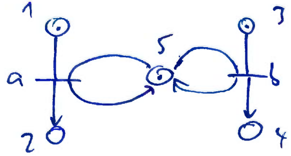
$\mathrm{ess}(X)$ $X[b] = \{(p_1, b), (p_2, b)\}$

$X[ab] = \{(\varepsilon, ab)\}$



Examples

N:



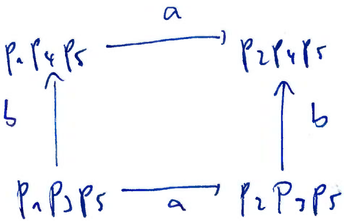
$S = \{p_1, \dots, p_5\}$ $T = \{a, b\}$

F

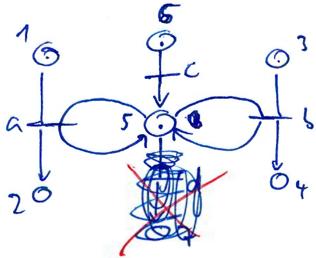
$ess(X)$: $X[\emptyset] = \{p_1 p_3 p_5, \dots\}$

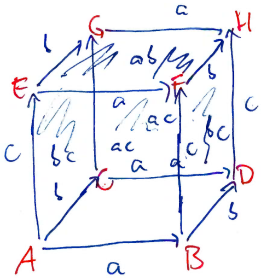
$X[a] = \{(p_3, a), (p_4, a)\}$

$X[b] = \{(p_1, b), (p_2, b)\}$



N:





ess(X): $X[\emptyset] = \{p_1+p_3+p_5+p_6,\ p_2+p_3+p_5+p_6,\ p_1+p_4+p_5+p_6,\ p_2+p_4+p_5+p_6,\ p_1+p_3+2p_5,\ p_2+p_3+2p_5,\ p_1+p_4+2p_5,\ p_2+p_4+2p_5\}$

$X[a] = \{(p_3+p_6, a), (p_4+p_6, a), (p_3+p_5, a), (p_4+p_5, a)\}$

$X[b] = \{(p_1+p_6, b), (p_2+p_6, b), (p_1+p_5, b), (p_2+p_5, b)\}$

$X[c] = \{(p_1+p_3+p_5, c), (p_2+p_3+p_5, c), (p_1+p_4+p_5, c), (p_2+p_4+p_5, c)\}$

$X[d] = \{(\ldots, d)\}$

Theorems

Definition (Rational Languages over Σ)

- Generated by $\emptyset$, $\{\epsilon\}$, and all $\{[a]\}$, $\{[\bullet\, a]\}$, $\{[a\,\bullet]\}$, $\{[\bullet\, a\,\bullet]\}$ for $a \in \Sigma$
- under operations $\cup$, $*$, $\|$ and (Kleene plus) $^+$

- $L^+ = \bigcup_{n\ge 1} L^n$
- no Kleene star; no parallel star

Theorem (à la Kleene)

A language is rational iff it is recognized by an HDA. CONCUR'22

Theorem (à la Myhill-Nerode)

A language is rational iff it has finite prefix quotient. Petri Nets'23

Recent Results

- regular languages are closed under ($\cup$, $*$, $\|$, $^+$, and) $\cap$
- but not under complement
  - L regular $\Rightarrow$ L has finite width $\Rightarrow$ $(\text{iiPoms} - L)\downarrow$ has infinite width
- width-bounded complement: $\overline{L}^k = \{P \in \text{iiPoms} - L \mid \text{wid}(P) \le k\}\downarrow$
- regular languages are closed under $\overline{}^k$ (for all k)
- not all HDAs are determinizable
- in fact, there are languages of unbounded ambiguity
  - for example $L = ([\,{}^{a}_{b}\,]\, cd + ab\, [\,{}^{c}_{d}\,])^+$

Lemma (Pumping Lemma)

(just like for finite automata!) ICTAC'23

Theorem

Inclusion of regular languages is decidable. ICTAC'23

Conclusion & Further Work

Higher-Dimensional Automata Theory for Fun and Profit!

- Kleene and Myhill-Nerode: a good start
- are HDAs learnable?
- trouble with determinization and non-ambiguity: residual automata?

- logical characterization: Büchi-Elgot theorem
- higher-dimensional timed automata

- relation to trace theory?
- languages vs homotopy?
- presheaf automata?
- coalgebra?
- higher-dimensional omega-automata