Discrete and Continuous Models for Concurrent Systems

4. Higher-Dimensional Automata Theory

Uli Fahrenberg

LMF, Université Paris-Saclay, France

ISIMM, April 2026

1. Languages of Higher-Dimensional Automata: Properties
2. Kleene Theorem
3. Büchi-Elgot-Trakhtenbrot Theorem
4. Exercises

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) $^+$
- (these need to take subsumption closure into account)

Definition (Monadic Second-Order Logics over Ipomsets)

$$\psi ::= a(x) \mid \mathsf{s}(x) \mid \mathsf{t}(x) \mid x < y \mid x \dashrightarrow y \mid x \in X \mid$$
$$\exists x.\, \psi \mid \forall x.\, \psi \mid \exists X.\, \psi \mid \forall X.\, \psi \mid \psi_1 \wedge \psi_2 \mid \psi_1 \vee \psi_2 \mid \neg \psi$$

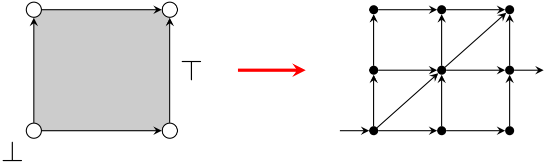
Theorem (à la Kleene): regular $\iff$ rational

Theorem (à la Myhill-Nerode): regular $\iff$ finite prefix quotient

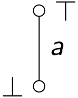
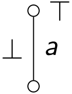
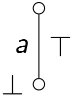
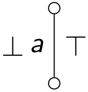
Theorem (à la Büchi-Elgot-Trakhtenbrot):
regular $\iff$ MSO-definable, of finite width, and subsumption-closed

Kleene Theorem: Easy Parts

- regular $\Longrightarrow$ rational: by reduction to ST-automata



- rational $\Longrightarrow$ regular: generators:

$L(X)$	$\emptyset$	$\{\epsilon\}$	$\{[a]\}$	$\{[\bullet\, a]\}$	$\{[a\, \bullet]\}$	$\{[\bullet\, a\, \bullet]\}$
X	$\emptyset$	$\bot \circ \top$				

- rational $\Longrightarrow$ regular: $\cup$ and $\parallel$

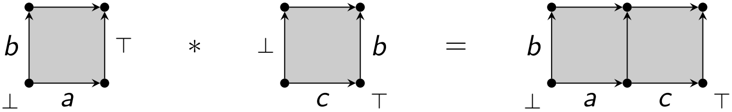
$$L(X) \cup L(Y) = L(X \sqcup Y) \qquad L(X) \parallel L(Y) = L(X \otimes Y)$$

Kleene Theorem: Difficult Parts

- miss to see: gluings and iterations of regular languages are regular:

$$L(X) * L(Y) = L(X * Y) \qquad L(X)^+ = L(X^+)$$

- much more difficult: higher-dimensional gluings identify too much
- for example:



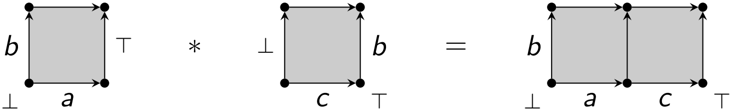
$L_2 = \emptyset$ $\qquad$ $L_1 * L_2 \ni ac$

Kleene Theorem: Difficult Parts

- miss to see: gluings and iterations of regular languages are regular:

$$L(X) * L(Y) = L(X * Y) \qquad L(X)^+ = L(X^+)$$

- much more difficult: higher-dimensional gluings identify too much
- for example:



$L_2 = \emptyset$ $\qquad$ $L_1 * L_2 \ni ac$

- use HDAs with interfaces and cylinder objects

HDAs with Interfaces

A conclist with interfaces (iconclist) is a conclist U with subsets $S \subseteq U \supseteq T$, denoted ${}_S U_T$

(events in T cannot be terminated; events in S cannot be "unstarted")

A precubical set with interfaces (ipc-set) X consists of a set of cells X such that:

- Every cell $x \in X$ has an iconclist $\mathrm{ev}(x)$
- We write $X[{}_S U_T] = \{x \in X \mid \mathrm{ev}(x) = {}_S U_T\}$.
- For every $A \subseteq U - S$ there is a lower face map $\delta_A^0 : X[U] \to X[{}_S U_T - A]$.
- For every $B \subseteq U - T$ there is an upper face map $\delta_B^1 : X[U] \to X[{}_S U_T - b]$.
- Precubical identities: $\delta_A^\mu \delta_B^\nu = \delta_B^\nu \delta_A^\mu$ for $A \cap B = \emptyset$ and $\mu, \nu \in \{0, 1\}$
- (presheaves over a category $\mathrm{I}\square$)

An HDA with interfaces (iHDA) is a finite ipc-set with start and accept cells.

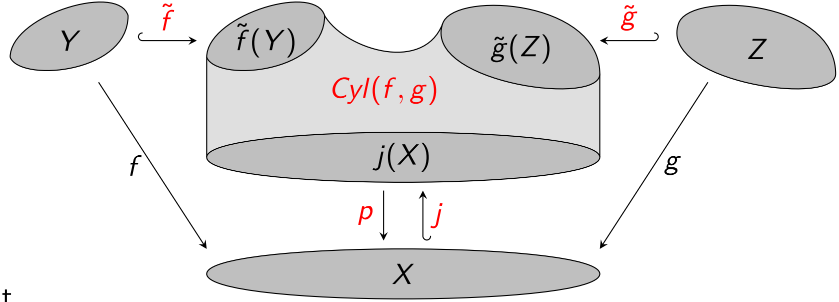
Extra conditions:

If $x \in X[{}_S U_T]$ is a start cell, then $S = U$.

If $x \in X[{}_S U_T]$ is an accept cell, then $T = U$.

Cylinders

Let X, Y, Z be ipc-sets and $f : Y \to X$, $g : Z \to X$ ipc-maps with $f(Y) \cap g(Z) = \emptyset$
There is a diagram of ipc-sets



such that

- $\tilde{f}$ is an initial inclusion;
- $\tilde{g}$ is a final inclusion;
- all paths in X from $f(Y)$ to $g(Z)$ lift to paths in $Cyl(f,g)$.

Cylinders: Construction

X, Y, Z: ipc-sets, $f : Y \to X$, $g : Z \to X$: ipc-maps with $f(Y) \cap g(Z) = \emptyset$.
For ${}_S U_T \in \mathrm{I}\square$ let

$$Cyl(f, g)[{}_S U_T] = \{(x, K, L, \varphi, \psi)\}$$

such that

- $x \in X[{}_S U_T]$;
- $K \subseteq \mathrm{I}\square^U$ is an initial subset;
- $L \subseteq \mathrm{I}\square^U$ is a final subset;
- $\varphi : K \to Y$, $\psi : L \to Z$ are ipc-maps satisfying $f \circ \varphi = \iota_x|_K$ and $g \circ \psi = \iota_x|_L$:

$$\begin{tikzcd}
K \arrow[r, hook] \arrow[d, "\varphi"'] & \mathrm{I}\square^U \arrow[d, "\iota_x"] & L \arrow[l, hook'] \arrow[d, "\psi"] \\
Y \arrow[r, "f"] & X & Z \arrow[l, "g"']
\end{tikzcd}$$

Gluing Composition of Regular Languages Is Regular

Proposition

Gluing composition of regular languages is regular.

Proof sketch: Let L and M be regular languages.

1. We may ssume that L, M are simple, i.e., $L = L(X)$, $M = L(Y)$ for iHDAs X, Y having one initial and one accepting cell each.
2. Now replace X by $X' = \mathit{Cyl}(X \leftarrow \top_X : j)$ and Y by $Y' = \mathit{Cyl}(i : \bot_Y \to Y)$, then $L(X') = L(X)$ and $L(Y') = L(Y)$.
3. Go back to HDA and glue:

$$L(\mathbf{Cl}(X') * \mathbf{Cl}(Y')) = L(X') * L(Y') = L * M.$$

   (closure $\mathbf{Cl}$: iHDA $\to$ HDA “adds missing cells”)

4. So $L * M$ is recognized by a finite HDA, hence regular.

Monadic Second-Order Logics over Ipomsets

Definition

MSO-iiPoms is built using the grammar

$$\begin{aligned}\psi ::= {} & a(x) \mid \mathsf{s}(x) \mid \mathsf{t}(x) \mid x < y \mid x \dashrightarrow y \mid x \in X \mid \\ & \exists x.\, \psi \mid \forall x.\, \psi \mid \exists X.\, \psi \mid \forall X.\, \psi \mid \psi_1 \wedge \psi_2 \mid \psi_1 \vee \psi_2 \mid \neg\psi\end{aligned}$$

- signature $\{<, \dashrightarrow, (a)_{a\in\Sigma}, \mathsf{s}, \mathsf{t}\}$
- $x \prec y := x < y \wedge \neg(\exists z. x < z < y)$
- for $\psi \in$ MSO with free variables $x_1, \dots, x_n, X_1, \dots, X_m$ and $P \in$ iiPoms, a valuation is $\nu = (\nu_1, \nu_2)$ with $\nu_1 : \{x_1, \dots, x_n\} \to P$ and $\nu_2 : \{X_1, \dots, X_m\} \to 2^P$
- $P \models_\nu \psi$ if ψ holds with x_i and X_j interpreted as $\nu(x_i)$ and $\nu(X_j)$
- $P \models \psi$ if $P \models_{(\emptyset,\emptyset)} \psi$ (no free variables: a sentence)
- $L(\psi) = \{P \in \text{iiPoms} \mid P \models \psi\}$
- L is MSO-definable if $\exists \psi \in \text{MSO} : L = L(\psi)$

Example

$$\varphi = \exists x\, \exists y.\, a(x) \wedge b(y) \wedge \neg(x < y) \wedge \neg(y < x)$$

- satisfied by all P which contain an a in parallel with a b
- in particular any conclists $\begin{bmatrix} \vdots \\ a \\ \vdots \\ b \\ \vdots \end{bmatrix}$ or $\begin{bmatrix} \vdots \\ b \\ \vdots \\ a \\ \vdots \end{bmatrix}$
- but not by ab nor ba

$\Longrightarrow$ $L(\varphi)$ has infinite width and is not closed under subsumption

Results

Theorem

For all $L \subseteq \text{iiPoms}$,

- *if L is MSO-definable, then* $L_{\le k}{\downarrow}$ *is regular for all* $k \in \mathbb{N}$;
  (subsumption closure of width restriction of L)
- *if L is regular, then L is MSO-definable.*

The constructions are effective in both directions.

Corollary

For all $k \in \mathbb{N}$, *a language* $L \subseteq \text{iiPoms}_{\le k}$ *with* $L = L{\downarrow}$ *is regular iff it is MSO-definable.*

Results

Corollary

For all $k \in \mathbb{N}$ *and* $\varphi \in \mathrm{MSO}$ *with* $L(\varphi) = L(\varphi)_{\leq k}{\downarrow}$, *satisfiability and model checking are decidable.*

- also implied by Courcelle's theorem, even more generally for $L(\varphi) = L(\varphi)_{\leq k}$, because $\mathrm{iiPoms}_{\leq k}$ has bounded treewidth

Corollary

For all $k \in \mathbb{N}$ *and* $L \subseteq \mathrm{iiPoms}_{\leq k}$, *if* L *is MSO-definable, then so is* $L{\downarrow}$.

- does not hold for general (non-interval) ipomsets [Fanchon-Morin 2009]

Proofs: Special ipomsets

Definition

An ipomset $(P, <, \dashrightarrow, S, T, \lambda)$ is

- discrete if $<$ is empty (hence $\dashrightarrow$ is total) $\begin{bmatrix} \bullet a \\ \bullet b \bullet \\ a \bullet \end{bmatrix}$
- a starter if it is discrete and $T = P$ $\begin{bmatrix} \bullet a \bullet \\ \bullet b \bullet \\ a \bullet \end{bmatrix}$
- a terminator if it is discrete and $S = P$ $\begin{bmatrix} \bullet a \\ \bullet b \bullet \\ \bullet a \bullet \end{bmatrix}$
- an identity if it is both a starter and a terminator $\begin{bmatrix} \bullet a \bullet \\ \bullet b \bullet \\ \bullet a \bullet \end{bmatrix}$

Lemma

Any interval ipomset has a decomposition as a sequence of starters and terminators.

$$\left[\begin{tikzcd} a \arrow[r] & b \\ c \arrow[r] \arrow[ur] & a \end{tikzcd}\right] = \begin{bmatrix} a \bullet \\ c \end{bmatrix} * \begin{bmatrix} \bullet a \\ a \bullet \end{bmatrix} * \begin{bmatrix} b \\ \bullet a \end{bmatrix} = \begin{bmatrix} a \bullet \\ c \bullet \end{bmatrix} * \begin{bmatrix} \bullet a \bullet \\ \bullet c \end{bmatrix} * \begin{bmatrix} \bullet a \bullet \\ a \bullet \end{bmatrix} * \begin{bmatrix} \bullet a \\ \bullet a \bullet \end{bmatrix} * \begin{bmatrix} b \bullet \\ \bullet a \bullet \end{bmatrix} * \begin{bmatrix} \bullet b \\ \bullet a \end{bmatrix}$$

Unique decompositions

Notation: St: set of starters ${}_S U_U$
Te: set of terminators ${}_U U_T$
$\mathsf{Id} = \mathsf{St} \cap \mathsf{Te}$: set of identities ${}_U U_U$
$\Omega = \mathsf{St} \cup \mathsf{Te}$

Definition

A word $w = (S_1, U_1, T_1) \dots (S_n, U_n, T_n) \in \Omega^+$ is coherent if $T_i = S_{i+1}$ for all i. The set of coherent words is denoted Coh.

Definition

A coherent word is sparse if proper starters and proper terminators are alternating.

- additionally, all $w \in \mathsf{Id} \subseteq \Omega^+$ are sparse
- so that's $\mathsf{Id} \cup (\mathsf{St} \setminus \mathsf{Id})\big((\mathsf{Te} \setminus \mathsf{Id})(\mathsf{St} \setminus \mathsf{Id})\big)^* \cup (\mathsf{Te} \setminus \mathsf{Id})\big((\mathsf{St} \setminus \mathsf{Id})(\mathsf{Te} \setminus \mathsf{Id})\big)^*$

Lemma

Any interval ipomset P has a unique decomposition $P = P_1 * \cdots * P_n$ such that $P_1 \dots P_n \in \Omega^+$ is sparse.

From MSO to HDAs

Idea: translate ipomsets to decompositions in Ω^+ and MSO-iiPoms to MSO over Ω

- problem: Ω is infinite $\implies$ need to restrict width

Lemma

For every $k \in \mathbb{N}$ and $\varphi \in \text{MSO}_{\text{iiPoms}}$ there exists $\hat\varphi \in \text{MSO}_{\Omega_{\leq k}}$ such that $P_1 \dots P_n \models \hat\varphi$ iff $P_1 \dots P_n$ is coherent and $P_1 * \dots * P_n \models \varphi$.

- $\text{coherent}_{\leq k} := \forall x\, \forall y.\, x \prec y \implies \bigvee_{P_1 P_2 \in \text{Coh}_{\leq k} \cap \Omega^2_{\leq k}} P_1(x) \wedge P_2(y)$
- ...
- Let $K = \{P \in \text{iiPoms}_{\leq k} \mid P \models \varphi\}$

$\implies$ $L'' = \{P_1 \dots P_n \in \Omega^+_{\leq k} \mid P_1 * \dots * P_n \in K\}$ is $\text{MSO}_{\Omega_{\leq k}}$-definable (lemma)

- standard Büchi & Kleene $\implies$ L'' is $\Omega_{\leq k}$-rational
- replace concatenation by gluing $\implies$ $L = \{P \in \text{iiPoms}_{\leq k} \mid P \models \varphi\}{\downarrow}$ is rational

From HDAs to MSO

Idea: encode accepting paths into MSO

- (similar to classical construction)
- for $P \in$ iiPoms, use sparse decomposition $P = P_1 * \cdots * P_n$ to define $st : P \to \{-\infty\} \cup \mathbb{N}$, $te : P \to \mathbb{N} \cup \{+\infty\}$ to denote where events start and terminate in $P_1 \dots P_n$

$\Longrightarrow$ $e \in P_i$ iff $st(e) \leq i \leq te(e)$

- and then map st and te to upsteps and downsteps in an accepting path

Lemma

For $f, g \in \{st, te\}$ and $\bowtie \in \{=, <, >\}$, the relations $f(x) \bowtie g(y)$, $\min(f)$ and $\max(f)$ are MSO-definable.

From HDAs to MSO

Lemma

Let X *be an HDA and* $P \in \text{iiPoms}$ *with sparse decomposition* $P = P_1 * \cdots * P_n$. *Then* $P \in L(X)$ *iff there exist* $\rho_{\nearrow} : P \setminus S_P \to \text{ups}(X)$ *and* $\rho_{\searrow} : P \setminus T_P \to \text{downs}(X)$ *such that for all* $e_1, e_2 \in P$,

- $st(e_1) = st(e_2) \implies \rho_{\nearrow}(e_1) = \rho_{\nearrow}(e_2)$ *and* $te(e_1) = te(e_2) \implies \rho_{\searrow}(e_1) = \rho_{\searrow}(e_2)$
- $st(e_2) = te(e_1) + 1 \implies \text{src}(\rho_{\nearrow}(e_2)) = \text{tgt}(\rho_{\searrow}(e_1))$
- $te(e_2) = st(e_1) + 1 \implies \text{src}(\rho_{\searrow}(e_2)) = \text{tgt}(\rho_{\nearrow}(e_1))$
- $\rho_{\nearrow}(e_1) = (p, \nearrow^A, q) \implies A = \{e \mid st(e) = st(e_1)\} \,\&\, \text{ev}(q) = \{e \mid st(e) \le st(e_1) < te(e)\}$
- $\rho_{\searrow}(e_1) = (p, \searrow_A, q) \implies A = \{e \mid te(e) = te(e_1)\} \,\&\, \text{ev}(q) = \{e \mid st(e) < te(e_1) \le te(e)\}$
- $st(e_1) = 1 \implies \text{src}(\rho_{\nearrow}(e_1)) \in \bot_X$ *and* $te(e_1) = 1 \implies \text{src}(\rho_{\searrow}(e_1)) \in \bot_X$
- $st(e_1) = n \implies \text{tgt}(\rho_{\nearrow}(e_1)) \in \top_X$ *and* $te(e_1) = n \implies \text{tgt}(\rho_{\searrow}(e_1)) \in \top_X$

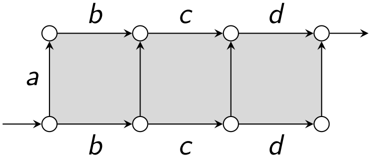
. . . and now encode these into MSO

Selected Bibliography

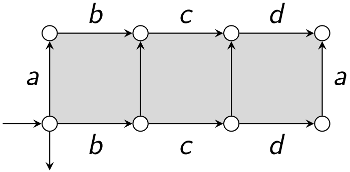
- U.F., C.Johansen, G.Struth, K.Ziemiański: *Kleene theorem for higher-dimensional automata*. Logical Meth.Comput.Sci. 2024
- U.F., K.Ziemiański: *Myhill-Nerode theorem for higher-dimensional automata*. Fund.Inf. 2024
- A.Amrane, H.Bazille, U.F., K.Ziemiański: *Closure and decision properties for higher-dimensional automata*. Theor.Comput.Sci. 2025
- A.Amrane, H.Bazille, U.F., M.Fortin: *Logic and languages of higher-dimensional automata*. DLT 2024

Exercises

Exercise 4: What is the language of the following HDA?

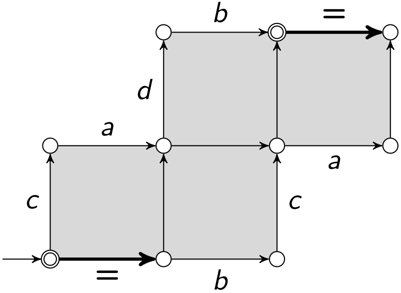


Exercise 5: What is the language of the HDA below? Now, horizontal cells with the same label are identified, as are the left and right vertical cells. (Geometrically, this is a torus.)



Exercises

Exercise 6: What is the language of the following HDA, with the cells marked "=" identified?



Exercises

Exercise 7: Construct HDAs which correspond to the following rational expressions.

① $\left[\begin{smallmatrix} b\bullet \\ a \end{smallmatrix}\right]\left[\begin{smallmatrix} \bullet b \\ c \end{smallmatrix}\right]$

② $\left(\left[\begin{smallmatrix} b\bullet \\ a \end{smallmatrix}\right]\left[\begin{smallmatrix} \bullet b \\ c \end{smallmatrix}\right]\right)^+$

③ $\left(\left[\begin{smallmatrix} \bullet b\bullet \\ a \end{smallmatrix}\right]\left[\begin{smallmatrix} \bullet b\bullet \\ c \end{smallmatrix}\right]\right)^+$

④ $\left(a\left[\begin{smallmatrix} a \\ a \end{smallmatrix}\right]+\left[\begin{smallmatrix} a \\ a \end{smallmatrix}\right]a\right)^+$

Exercise 8: Let's look a bit closer at the HDA of the last part of the last exercise, for $\left(a\left[\begin{smallmatrix} a \\ a \end{smallmatrix}\right]+\left[\begin{smallmatrix} a \\ a \end{smallmatrix}\right]a\right)^+$. Call it X.

① Let $n \geq 1$. How many paths in X recognize the word $(aaa)^n$?

② Does there exist another HDA Y with $L(Y)=\left(a\left[\begin{smallmatrix} a \\ a \end{smallmatrix}\right]+\left[\begin{smallmatrix} a \\ a \end{smallmatrix}\right]a\right)^+$ and in which fewer paths recognize the words $(aaa)^n$?

③ Conclude that the language $\left(a\left[\begin{smallmatrix} a \\ a \end{smallmatrix}\right]+\left[\begin{smallmatrix} a \\ a \end{smallmatrix}\right]a\right)^+$ is inherently infinitely ambiguous.