Regular Monoidal Languages

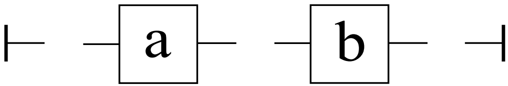
Matt Earnshaw[1]

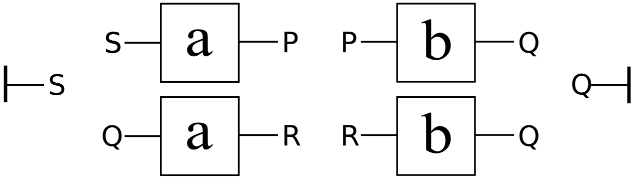
j.w.w. Paweł Sobociński[1]

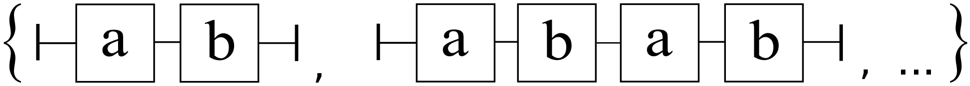
[1]Tallinn University of Technology, Estonia

SYCO 9, Como
September 2022

Regular languages with pictures



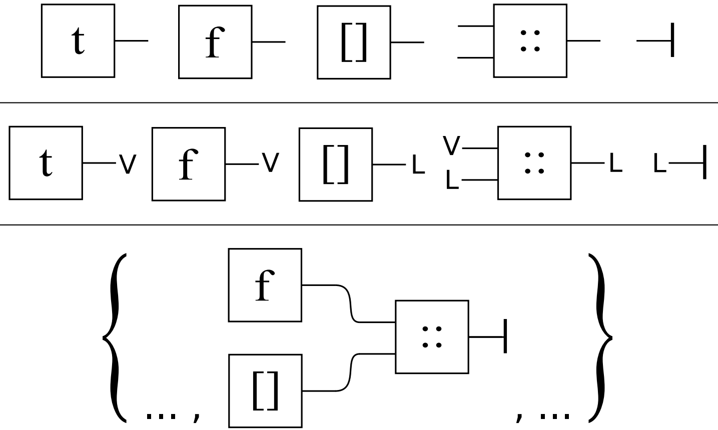




Fact: any regular language can be pictured in this way.

Bottom-up tree languages with pictures

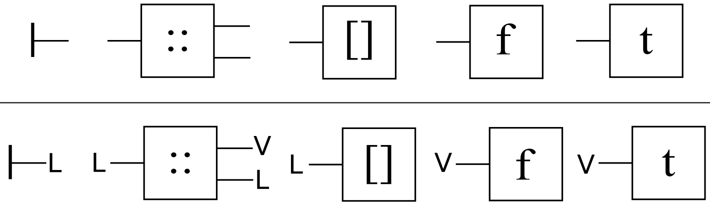
What about multiple wires on the left?

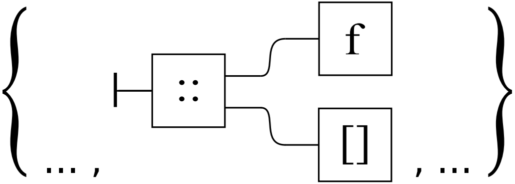


Fact: any bottom-up regular tree language can be pictured in this way.

Top-down tree languages with pictures

What about multiple wires on the right?

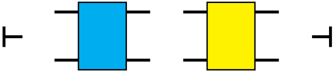


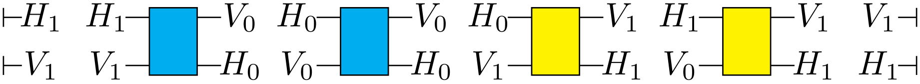


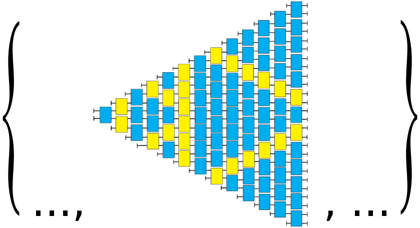
Fact: any top-down regular tree language can be pictured in this way.

Regular monoidal languages

What about multiple wires on the left and right?



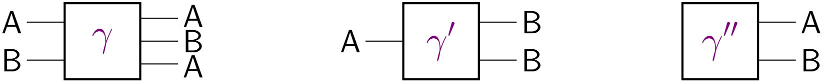




How to define these pictures formally?

Monoidal graphs

A monoidal graph is a pair of functions $s, t : E \rightrightarrows V^*$.



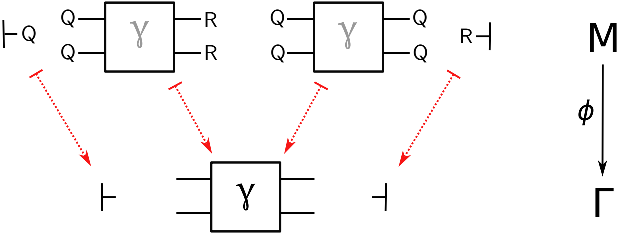
A monoidal graph generates a free pro.

When *single-sorted*, s and t give natural numbers: *arity* and *coarity*.

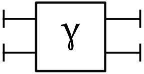
Morphism of monoidal graphs is a pair of functions
$E \to E', V \to V'$ commuting with s and t.

All of our monoidal graphs will be finite.

Regular monoidal grammars and languages



Gives a subset of $\mathscr{F}\Gamma(0,0)$, the $0 \to 0$ diagrams that can be built.



$\in L(\phi)$  $\notin L(\phi)$

Definition

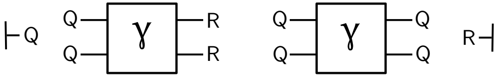
The subsets so definable are *regular monoidal languages*.

Non-deterministic monoidal automata

Definition $\Delta = (V, \Delta_\Gamma)$

- V, finite set
- Γ, monoidal alphabet
- $\Delta_\Gamma = \{V^{\mathrm{ar}(\gamma)} \xrightarrow{\Delta_\gamma} \mathscr{P}(V^{\mathrm{coar}(\gamma)})\}_{\gamma \in E_\Gamma}$,

set of transition relations



String diagrams $0 \to 0$ map to a $V^0 \to \mathscr{P}(V^0)$ (accept/reject).

By restricting Γ we recover:

- Ordinary non-deterministic automata
- Top-down n.d. tree automata
- Bottom-up n.d. tree automata

The problem of determinization

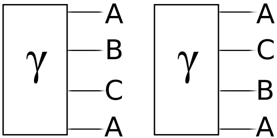
Challenge

Characterize the deterministically recognizable RMLs.

Partial answers:

- convex automata
- necessary property of deterministic language
- categorical invariant of language

Partial answer I: Convex automata



$\gamma : V^0 \to \mathscr{P}(V^4)$ is not convex

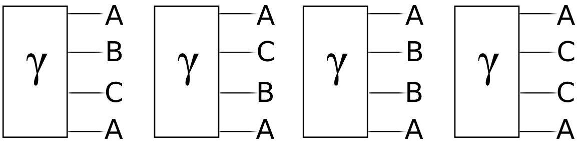
$$\begin{tikzcd}
(\mathscr{P}Q)^n \arrow[r, "\gamma^*"] \arrow[d, "\nabla"'] & (\mathscr{P}Q)^m \arrow[d, "\nabla"] \\
\mathscr{P}(Q^n) \arrow[r, "\gamma^\#"] & \mathscr{P}(Q^m)
\end{tikzcd}$$

A monoidal automaton is convex if its transition relations are convex.

Theorem

Convex automata can be determinized, by an analogue of the powerset construction. E.g. word automata and bottom-up tree automata.

Partial answer I: Convex automata



$\gamma : V^0 \to \mathscr{P}(V^4)$ is convex

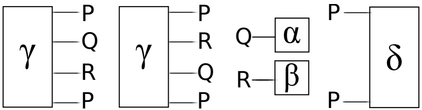
$$\begin{tikzcd} (\mathscr{P}Q)^n \arrow[r, "\gamma^*"] \arrow[d, "\nabla"'] & (\mathscr{P}Q)^m \arrow[d, "\nabla"] \\ \mathscr{P}(Q^n) \arrow[r, "\gamma^\#"] & \mathscr{P}(Q^m) \end{tikzcd}$$

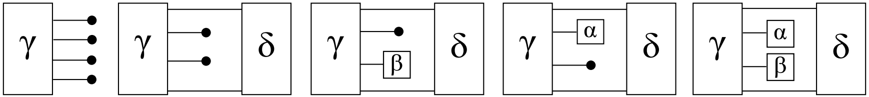
A monoidal automaton is convex if its transition relations are convex.

Theorem

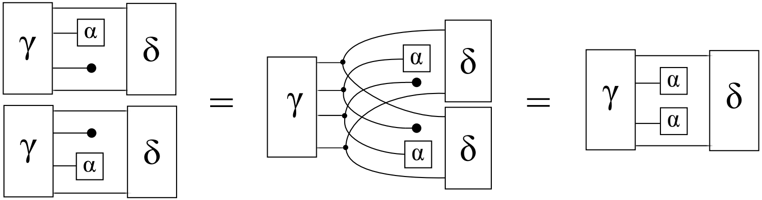
Convex automata can be determinized, by an analogue of the powerset construction. E.g. word automata and bottom-up tree automata.

Partial answer II: Causal closure



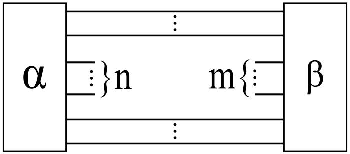


Causal histories recombinable via equations in cartesian restriction categories



Theorem: Deterministically recognizable RMLs are causally closed.

Partial answer III: Syntactic pro



$$\gamma \equiv_L \delta \text{ if } C[\gamma] \in L \iff C[\delta] \in L, \text{ for all contexts } C$$

Theorem

If L is an RML then its syntactic pro has finite homsets.

Theorem

If the syntactic pro of an RML has cartesian restriction category structure, then the language is deterministically recognizable.

Future work

- Completely characterize deterministic recognizability
- Embeddings of word languages
- Diagrammatics for pushdown and Zielonka automata, transducers, etc.
- Context-free monoidal languages via a monoidal multicategory of contexts

Thanks for your attention.