A Family of Effective Methods for Decompiling Canonical Acceptors, Instantiated for Languages of Dot-Depth One and Tier-Based Extensions

Dakotah Lambert, Department of Mathematics and Computer Science, Lake Forest College
dakotahlambert@acm.org

Varieties

A $*$-**variety** is a family of language classes, defining for each alphabet Σ a class $\Sigma^*\mathscr{V}$, based on structural properties.

Each class is closed under Boolean operations and both left and right quotients. The family as a whole is closed under inverse images of homomorphisms.

Lifting to a tier is taking the inverse image of a homomorphism, so this operation provides no additional power to a $*$-variety.

Quantifier-Free Logics

Base languages combined using some chosen connectives, usually Boolean operations. Many satisfy the other closure properties and correspond to varieties or families of varieties.

Examples (x and x_i in Σ)

- $\mathscr{J}_1$: x
  "x occurs"
  content (set of letters)
- $\mathscr{A}com_t$: x^k
  "x occurs k or more times"
  threshold multisets (count to t)
- $\mathscr{J}_k$: $x_1 .. \cdots .. x_k$
  "x_i precedes x_{i+1} for $1 \leq i < k$"
  piecewise k-testable

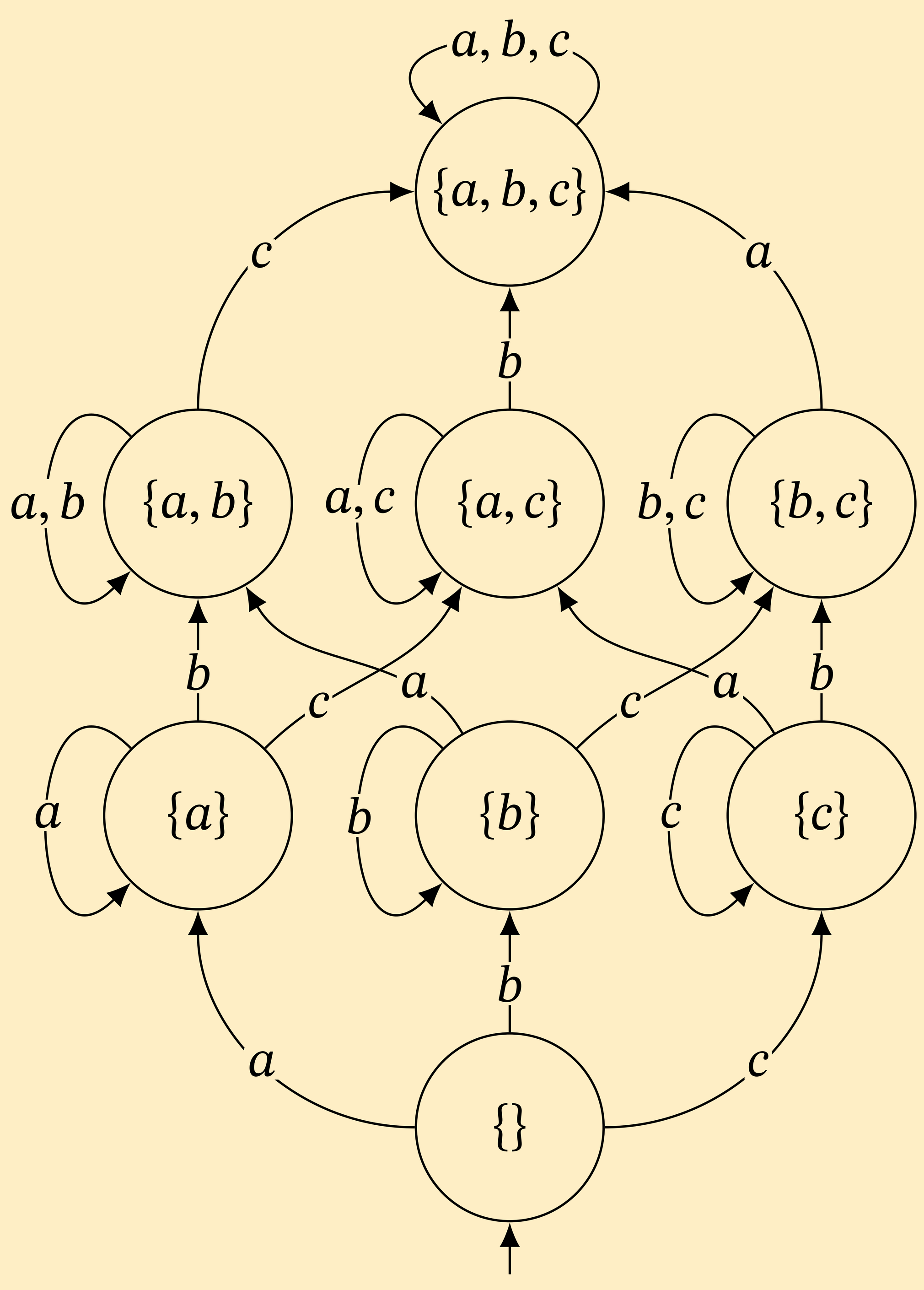
How can we pull logical descriptions out of automata?

Logic → Automata

Construct base languages, combine them with known operations.

Automata → Logic

For $*$-varieties. Build a **free object**: a general structure that encodes all and only the distinctions that $\Sigma^*\mathscr{V}$ can make. E.g. for $\{a,b,c\}\mathscr{J}_1$:



Unfold source automaton onto free object and mark states that coincide with an accepting state.

Interpret marked states as formulae and take their disjunction.

Some Non-$*$-Varieties

- Strictly local (SL)
- Locally testable (LT)
- Locally threshold testable (LTT)
- Dot-depth one
- …

Treat Windows as Symbols

Run a sliding window over the word with boundaries, and emit each view as its own symbol, using an input strictly local transduction:

$$\#x_1 \dots x_k \to \boxed{\#x_1 \dots x_k}$$
$$x_1 \dots x_{k+1} \to \boxed{x_1 \dots x_{k+1}}$$
$$x_1 \dots x_k\# \to \boxed{x_1 \dots x_k\#}$$
$$\#x_1 \dots x_j\# \to \boxed{\#x_1 \dots x_j\#}$$

where $x_i \in \Sigma$ for all i, and $j < k$.

Enhanced Extraction

Read $C \mapsto D$ as "given a way to extract formulae for class C, this transformation gives you a way to do it for class D".

$$\mathscr{J}_1 \mapsto \mathrm{LT}_k$$
$$\mathscr{A}com_t \mapsto \mathrm{LTT}_{k,t}$$
$$\mathscr{J}_\ell \mapsto \text{dot-depth one}$$
$$\mathscr{V} \mapsto \mathscr{V} * \mathscr{D}$$

Selected References

- Robert Knast. 1983. A semigroup characterization of dot-depth one languages. *RAIRO – Informatique théorique*, 17(4):321–330.
- Dakotah Lambert. 2026. Multitier phonotactics with logic and algebra. *Phonology*, 43:e9 1–31.
- Howard Straubing. 1985. Finite semigroup varieties of the form $V * D$. *Journal of Pure and Applied Algebra*, 36:53–94.