Machines, Models, Monoids, and Modal logic

Sam van Gool

University of Amsterdam and City College of New York

September 2017

Tbilisi Symposium on Language, Logic and Computation

Lagodekhi, Georgia

Outline

1. Part I: Formal Languages, Automata, and Algebra
2. Part II: Duality and Varieties of Monoids
3. Part III: Profiniteness, Pointlikes, and the Future

Recap

- Any regular language L comes with a finite closed Boolean subalgebra, $B(L)$, of $\mathrm{Reg}(\Sigma^*)$ and a finite monoid quotient, $M(L)$, of Σ^*.
- The syntactic monoid $M(L)$ is the Stone dual (= atom set) of the closed Boolean subalgebra $B(L)$.
- Properties of the syntactic monoid reflect properties of the language.
- The syntactic monoid $M(L)$ is aperiodic iff L is first-order definable.
- In general, varieties (HSP^{fin}-classes) of monoids correspond to varieties of languages.
- Full (non-finite) Stone duality can be used to explain and analyze this correspondence but how?

Outline Part III

1. Full Duality and Varieties
   - Varieties
   - Full Stone duality
   - Intuitionistic intermezzo
   - Profinite equations
2. Aperiodic pointlikes
   - Separation problem
   - Pointlike sets
   - Henckell's Theorem
3. The Future

1. Full Duality and Varieties
   - Varieties
   - Full Stone duality
   - Intuitionistic intermezzo
   - Profinite equations

Varieties of monoids and languages

- A class of finite monoids $\mathbf{V}$ is a (pseudo)*variety* if it is closed under homomorphic images (H), submonoids (S) and finite products ($\mathrm{P}^{\mathrm{fin}}$).
- For a variety of monoids $\mathbf{V}$, define $\mathcal{V}(\Sigma^*)$ to be the class of Σ-languages L such that $M(L) \in \mathbf{V}$.
- Then $\{\mathcal{V}(\Sigma^*)\}_\Sigma$ is a *variety of regular languages*: a collection of Boolean subalgebras of $\mathrm{Reg}(\Sigma^*)$ which is closed under inverse images of homomorphisms $\Sigma_1^* \to \Sigma_2^*$.

Theorem (Eilenberg)

The map $\mathbf{V} \mapsto \mathcal{V}$ is an order-bijection between varieties of finite monoids and varieties of regular languages.

Equations?

- Birkhoff's theorem: varieties of (arbitrary) algebras can be defined by (finite) equations.
- What about (pseudo)varieties of *finite* algebras?
- We need *pro*finite equations.
- To explain what these are, and why we need them: full Stone duality.

Stone duality: general case

Proposition

Every Boolean algebra B can be embedded into a Boolean algebra of the form $\mathcal{P}(X)$, and there is a unique such embedding for which the topology generated by the sets in the image of B is compact and Hausdorff (and zero-dimensional).

'Construction' of the embedding.

Take X to be the set of ultrafilters of B.
Identify $b \in B$ with the set, $\hat{b}$, of ultrafilters containing it. □

- A *Boolean space* is a compact Hausdorff zero-dimensional space.
- Equivalently, a *Boolean space* is a profinite object in the category of topological spaces.

Stone duality: example

Example

The dual space of the Lindenbaum algebra of CPL on a countable set $V = \{p_1, p_2, p_3, \dots\}$ is the *Cantor space* $\{0,1\}^V$.

Exercises

1. What is the dual space of the Boolean algebra of finite subsets of the natural numbers and their complements?
2. Use what you know about classical propositional logic to prove that the Lindenbaum algebra of CPL on a countable set $V = \{p_1, p_2, p_3, \dots\}$ can be embedded into $\mathcal{P}(\{0,1\}^V)$.
3. (*) Show that the topology generated by the image of the embedding in (2) is compact and Hausdorff.
4. (*) Show that the topology generated by the image of the embedding in (2) coincides with the topology of the Cantor space.

Duality: categorical level

- As in the finite case, all homomorphisms between Boolean algebras are of the form f^{-1}, for f a *continuous* function between the dual spaces.
- The *categories* of Boolean algebras and Boolean spaces are *dually equivalent*.

Algebras	dual to	Spaces
subalgebras	$\leftrightarrow$	quotient objects
quotient algebras	$\leftrightarrow$	subobjects
homomorphisms	$\leftrightarrow$	continuous functions
algebraic operations	$\leftrightarrow$	co-algebraic operations
unions (directed colimits)	$\leftrightarrow$	projective limits

Stone duality: summary

- Finite Boolean algebras are power sets.
- Boolean algebras are subalgebras of power sets.
- Boolean algebra homomorphisms are inverse images.
- Boolean algebras are algebras of clopen sets of a compact Hausdorff topological space, called the dual space.
- Subalgebras of the Boolean algebra correspond to quotient spaces of the dual space.
- Quotients of the Boolean algebra correspond to closed subspaces of the dual space.

Intuitionistic Intermezzo

An open mapping theorem for Esakia spaces

- Stone duality generalizes to *Heyting algebras*, the structures for *intuitionistic* propositional logic analogous to Boolean algebras.
- The Boolean space is equipped with a *partial order* (= the Kripke accessibility relation in semantics for intuitionistic logic).
- The spaces dual to Heyting algebras were characterized by L. Esakia and are now called *Esakia spaces*.
- Heyting algebra homomorphisms also require special attention: their duals are *continuous p-morphisms*.
- Esakia duality is useful, for example, for proving *interpolation* properties of intermediate logics.
- In recent joint work with L. Reggio, we prove an *open mapping theorem* for Esakia spaces dual to finitely presented Heyting algebras.
- Our result in particular implies Pitts' Uniform Interpolation Thoerem for IPC.

Another hint of duality for interpolation?

- The use of duality for analyzing quantifiers is not limited to the context of regular languages.
- A classical example is Rasiowa & Sikorski's proof of the completeness of classical predicate logic via Stone duality and the Baire category theorem.
- A recent example (I claim) is the counterexample to interpolation for constant domain intuitionistic predicate logic (Mints, Olkhovikov, Urquhart JSL 2013).
- In the latter, and in Olkhovikov's work on van-Benthem-style characterizations (2012-2015), the use of duality is not (yet) explicit.

End of Intuitionistic Intermezzo

Stone duality: crucial example for language varieties

Example

The dual space of the Boolean algebra $\mathrm{Reg}(\Sigma^*)$ of regular Σ-languages is the projective limit of the diagram $(q\colon \Sigma^* \to M)$ of *finite* quotients of Σ^*.
This is the space underlying the free profinite monoid, $\widehat{\Sigma^*}$, on Σ.
Thus, the free profinite monoid over Σ is the 'canonical Kripke model' for MSO on finite words (since the Lindenbaum algebra is $\mathrm{Reg}(\Sigma^*)$).

On the free profinite monoid

"In the mid-1970s when I was at Oxford, it occurred to me that, using regular events in the free monoid on a finite alphabet as neighborhoods, one could make a completion to 'infinite words'. I even suggested to one of my students to consider this idea for a thesis. Neither he nor I could make much progress in analyzing this algebra or applying this idea, however, and he went on to write a different thesis (fortunately)."
[D. Scott, via e-mail, Nov. 9, 2016]

Describing varieties: profinite equations

- Let $\mathbf{V}$ be a variety of finite monoids.
- Let $\mathcal{V}$ be the corresponding variety of regular languages.
- For every alphabet Σ, $\mathcal{V}(\Sigma^*)$ is a closed subalgebra of $\mathrm{Reg}(\Sigma^*)$.
- The dual of this closed subalgebra is a continuous monoid quotient $\widehat{\Sigma^*} \to \widehat{F}_{\mathbf{V}}(\Sigma)$, the free pro-$\mathbf{V}$-monoid on Σ.
- The fact that $\mathcal{V}$ is closed under inverse images of homomorphisms means that the quotients $\widehat{\Sigma^*} \to \widehat{F}_{\mathbf{V}}(\Sigma)$ are substitution-invariant.
- Thus, varieties can be described by profinite equations.

Profinite equations: example

Example

The free pro-aperiodic monoid, $\widehat{F}_{\mathbf{A}}(\Sigma^*)$, is the quotient of $\widehat{\Sigma^*}$ by the equivalence relation defined by the substitution-invariant equation $x^\omega = x^\omega x$.

Here, $()^\omega : \widehat{\Sigma^*} \to \widehat{\Sigma^*}$ is the operation which sends any x to the idempotent x^ω in $\overline{\{x^n \mid n \geq 1\}}$.

Duality beyond the profinite

- Let $\mathcal{V}$ be any closed subalgebra of $\mathcal{P}(\Sigma^*)$.
- It corresponds again to a topological quotient, but now of $\beta\Sigma^*$, the Stone-Cech compactification of the discrete free monoid.
- The story for the monoid operation is a bit more complicated (internal monoid action).
- This idea, combined with methods from circuit complexity theory, leads to 'ultrafilter equations' for characterizing classes of non-regular languages.
- M. Gehrke, A. Krebs, and J.-É. Pin. "From ultrafilters on words to the expressive power of a fragment of logic". In: *DCFS 2014*. Vol. 8614. Lect. Notes Comput. Sci. Springer, 2014, pp. 138–149
- M. Gehrke, D. Petrisan, and L. Reggio. "Quantifiers on languages and codensity monads". In: *TACL 2017*. https://arxiv.org/abs/1702.08841. 2017

2. Aperiodic pointlikes
   - Separation problem
   - Pointlike sets
   - Henckell's Theorem

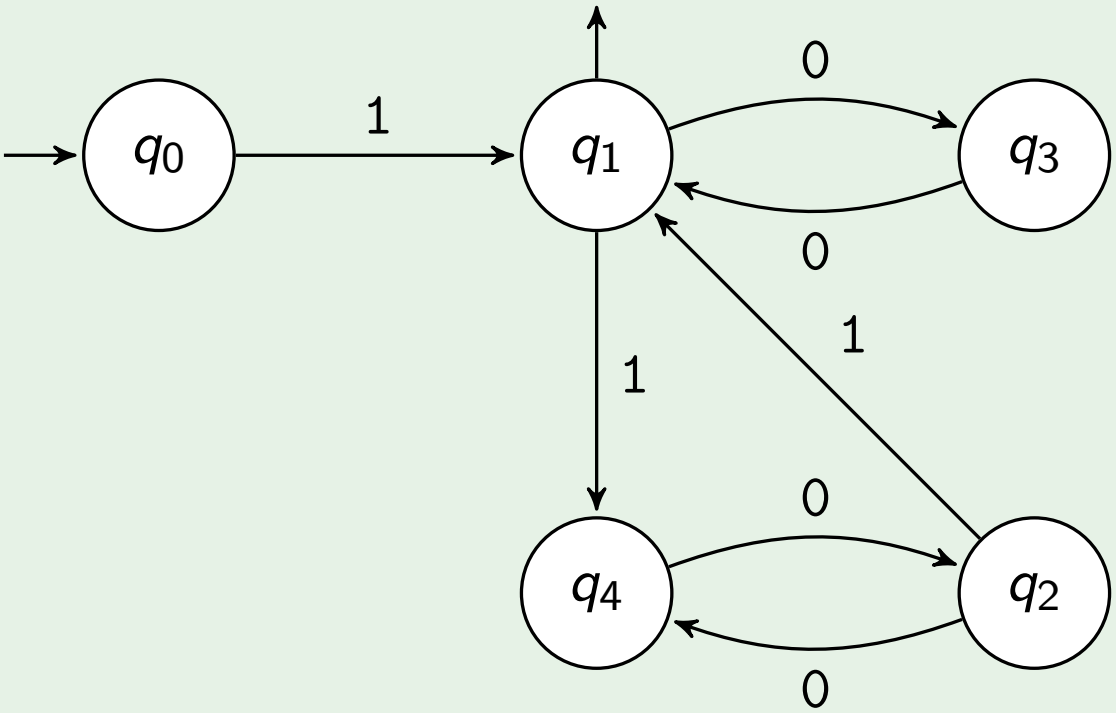
Separation problem: language version

- From here on, we work with semigroups, and 'Σ-language' means subset of Σ^+.
- Let $\mathbf{V}$ be a variety of finite semigroups with corresponding variety of languages $\mathcal{V}$.
- Separation Problem: Given two disjoint regular Σ-languages L_1, L_2, is it possible to find a language, K, in $\mathcal{V}(\Sigma)$ which *separates* L_1 from L_2?
- Here, K separates L_1 from L_2 if $L_1 \subseteq K$ and $L_2 \cap K = \emptyset$.
- If φ_1 and φ_2 are MSO sentences defining L_1 and L_2, respectively, then disjointness means $\varphi_1 \vdash \neg\varphi_2$.
- The *logic formulation* of the separation problem is: does there exist ψ such that $\varphi_1 \vdash \psi \vdash \neg\varphi_2$, with the language $K = L_\psi$ in $\mathcal{V}(\Sigma)$.
- In general, this problem can fail to be decidable, even when membership in $\mathcal{V}$ is decidable.

Example of non-separable languages

Example (Place & Zeitoun 2016)

Let $\Sigma = \{0, 1\}$. Consider the automaton

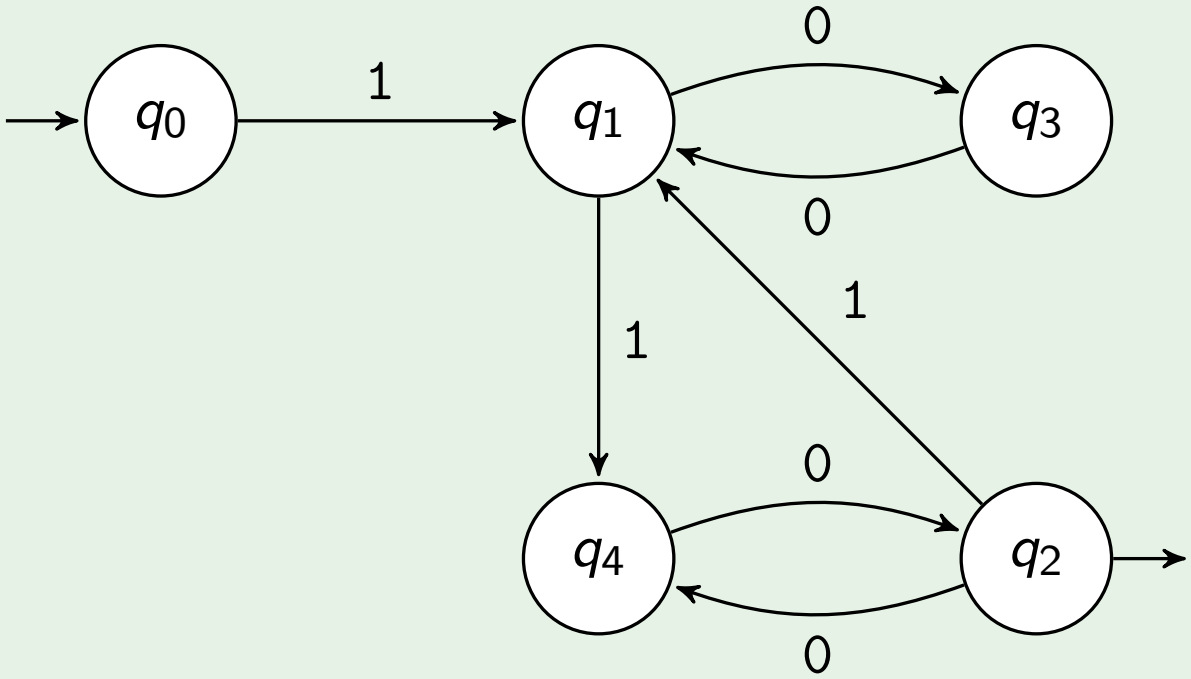


The language recognized with q_1 final is $L_1 = (1(00)^*10(00)^*)^*1(00)^*$.

Example of non-separable languages

Example (Place & Zeitoun 2016)

Let $\Sigma = \{0, 1\}$. Consider the automaton

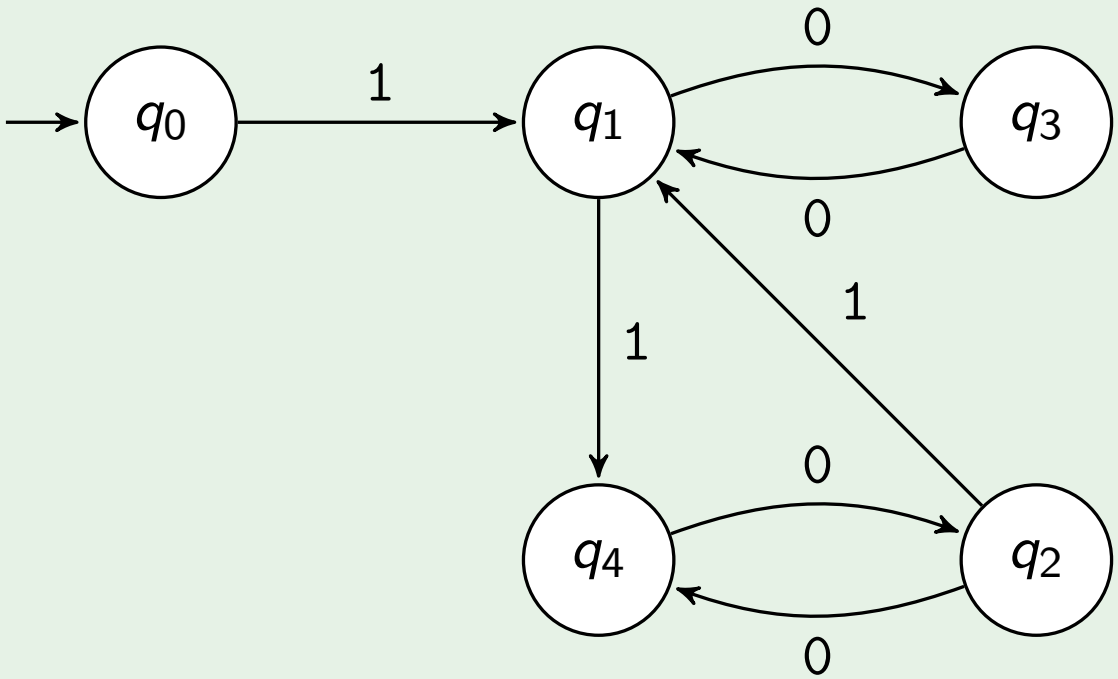


The language recognized with q_1 final is $L_1 = (1(00)^*10(00)^*)^*1(00)^*$.

The language recognized with q_2 final is $L_2 = (1(00)^*10(00)^*)^+$.

Example (Place & Zeitoun 2016)

Let $\Sigma = \{0, 1\}$. Consider the automaton



The language recognized with q_1 final is $L_1 = (1(00)^*10(00)^*)^*1(00)^*$.

The language recognized with q_2 final is $L_2 = (1(00)^*10(00)^*)^+$.

The languages L_1 and L_2 are *disjoint*, but not FO-separable.

Exercise: Use first-order logic games to prove this.

Separation problem: semigroup version

- The separation problem can be formulated as a problem about semigroups.
- We may assume L_1 and L_2 are recognized by the same semigroup homomorphism $\eta\colon \Sigma^+ \to S$.
- So $L_1 = \eta^{-1}(R_1)$ and $L_2 = \eta^{-1}(R_2)$, with R_1 and R_2 disjoint.
- Is there a semigroup homomorphism $\theta\colon \Sigma^* \to T \in \mathbf{V}$, and $P \subseteq T$, such that $\eta^{-1}(R_1)$ is contained in $\theta^{-1}(P)$, and $\eta^{-1}(R_2)$ is disjoint from $\theta^{-1}(P)$?
- Fact. The answer is 'no' if, and only if, for every $r_1 \in R_1$ and $r_2 \in R_2$, the subset $\{r_1, r_2\}$ of S is pointlike.

Pointlike sets

- A relational morphism from a semigroup S to a semigroup T is a relation $\varphi \subseteq S \times T$ such that $s\varphi \cdot s'\varphi \subseteq ss'\varphi$ and $s\varphi \neq \emptyset$ for all $s, s' \in S$.
- Equivalently, it is a relation of the form $\beta\alpha^{-1}$, where $\alpha\colon U \twoheadrightarrow S$ and $\beta\colon U \to T$ are homomorphisms from a semigroup U.
- A subset $X \subseteq S$ is **V**-pointlike if, for every relational morphism $\varphi\colon S \to T$ such that $T \in \mathbf{V}$, there exists a point $x \in T$ such that $X \subseteq \varphi^{-1}(x)$.
- If we can compute the (two-element) **V**-pointlike sets of S, then we can decide the **V**-separation problem:
- Given $L_1 = \eta^{-1}(R_1)$, $L_2 = \eta^{-1}(R_2)$ for $\eta\colon \Sigma^* \to M$, check if $\{r_1, r_2\}$ is pointlike for all $r_1 \in R_1$, $r_2 \in R_2$. If so, L_1 and L_2 are non-separable.
- In particular, to decide FO-separation, we will compute the **A**-pointlike sets, where **A** is the variety of aperiodic semigroups.

The monad of **V**-pointlikes

- The collection of **V**-pointlike sets, $\mathrm{PL}_{\mathbf{V}}(S)$, of a semigroup S is a subset of the *power semigroup*, 2^S, of S.
- Elements of 2^S are subsets of S, i.e., as a set, $2^S = \mathcal{P}(S)$.
- Multiplication on 2^S is given by: $X \cdot Y = \{xy \mid x \in X, y \in Y\}$.

Fact

*The collection, $\mathrm{PL}_{\mathbf{V}}(S)$, of **V**-pointlike subsets of a finite semigroup S, is a downward closed subsemigroup of 2^S which contains all the singletons.*

Fact

*The union of a **V**-pointlike subset of the semigroup $\mathrm{PL}_{\mathbf{V}}(S)$ is **V**-pointlike. That is, $\bigcup : \mathrm{PL}_{\mathbf{V}}(\mathrm{PL}_{\mathbf{V}}(S)) \to \mathrm{PL}_{\mathbf{V}}(S)$ is well-defined.*

Generating aperiodic-pointlike sets

- For $X \in 2^S$, define $X^{\omega+*} = \bigcup_{n \geq 0} X^\omega X^n$.
- Fact. If X is **A**-pointlike, then so is $X^{\omega+*}$.
- Singletons are **A**-pointlike.
- Products of **A**-pointlike sets are **A**-pointlike.
- Subsets of **A**-pointlike sets are **A**-pointlike.

Theorem (Henckell)

*For any finite semigroup S, the set of **A**-pointlikes of S is the smallest downward closed subsemigroup of 2^S which contains the singletons and is closed under the operation $X \mapsto X^{\omega+*}$.*

In particular, the **A**-pointlikes of any finite semigroup are computable.

Recent progress

- Henckell, Rhodes and Steinberg (2010) improved on Henckell's original proof and extended his methods to varieties of semigroups that avoid specific subgroups.
- Place and Zeitoun (2016) gave a logic proof of Henckell's Thm.
- Place and Zeitoun (2014-17) computed FO_2-pointlikes.
- Steinberg and I (2017) gave a semigroup proof of Henckell's Thm.
- To do so, we construct a 'merge decomposition' of homomorphisms.
- This is an algebraic version of 'quantifying over first and last occurrences'.
- In addition to a short elementary proof of Henckell's Theorem, we also give a short proof of the two-sided Krohn-Rhodes theorem.
- The latter, in a slogan, says:
  'semigroup theory = semilattice theory + group theory'.

References for Part III

- Basics on duality theory for Boolean algebras: Chapter 11 in
  B. A. Davey and H. A. Priestley. *Introduction to Lattices and Order*. 2nd. Cambridge University Press, May 6, 2002
- The duality-theoretic view on varieties:
  M. Gehrke. "Stone duality, topological algebra, and recognition". In: *J. Pure Appl. Algebra* 220.7 (2016), pp. 2711–2747
- On the relation between pointlikes and separation:
  J. Almeida. "Some algorithmic problems for pseudovarieties". In: *Publ. Math. Debrecen* 54.1 (1999), pp. 531–552
- Our recent preprint on Henckell's and Krohn-Rhodes theorems:
  S. J. v. Gool and B. Steinberg. "Merge decompositions, two-sided Krohn-Rhodes, and aperiodic pointlikes". arXiv:1708.08118, submitted. Aug. 2017

3 The Future

Seven questions

- How far can the decidability of pointlikes be stretched?
- How far do duality-theoretic methods reach beyond the regular?
- How does our semigroup-theoretic work fit with the category/duality approach?
- Is there a topos-theoretic interpretation of 'logic on words'?
- What can be said about regular languages with model theory? (Partial answers in joint work with Ghilardi)
- Can the relationship with modal logic be made more tight?
- Is anything I've said relevant for (formal) linguistics?

გმადლობ

XKCD 208: Regular Expressions (`https://xkcd.com/208/`)