Duality

and

Equational theory of regular languages

Yann Pequignot

Supervisor: Jacques Duparc,*Université de Lausanne, Switzerland.*

Co-supervisor: Jean-Éric Pin, *LIAFA, University Paris-Diderot and CNRS, France.*

Réunion FREC — May 2011
Île de Ré

Background

Formal languages, automata and monoids

- Finite set A, called an **alphabet**.
- Finite sequences on A, called **words**.
- Binary operation on words, called **concatenation**.

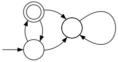
$$a_1 \cdots a_n \cdot b_1 \cdots b_n = a_1 \cdots a_n b_1 \cdots b_n$$

- **Free monoid** on A: A^* all words on A.
- Subsets of A^*, called **languages**.
- A machine can specify a language.
- **Finite state automaton**: simplest model.

Background

Formal languages, automata and monoids

For any $L \subseteq A^*$, TFAE:

- L is recognised by a finite automaton 
- L is recognised by a finite monoid $(M, \cdot, e)$

  there exists a surjective monoid morphism $\varphi : A^* \to M$ onto a finite monoid M and some $P \subseteq M$ such that for all $w \in A^*$

  $$w \in L \quad \textit{if and only if} \quad \varphi(w) \in P.$$

- The **syntactic congruence** $\sim_L$ defined by

  $$u \sim_L v \quad \text{iff} \quad \forall x, y \in A^* \; xuy \in L \leftrightarrow xvy \in L$$

  is of finite index.

$$\text{Rec}(A^*) = \{L \subseteq A^* \mid L \text{ is recognisable}\}$$

Eilenberg Theorem

Varieties of languages and finite monoids

A **variety of languages** is the association to each finite alphabet A of a

- Boolean subalgebra $\mathcal{V}(A^*)$ of $\mathrm{Rec}(A^*)$,
- **closed under quotienting**, $\forall L \in \mathcal{V}(A^*)$ and $\forall u \in A^*$

  $$u^{-1}L = \{w \in A^* \mid uw \in L\} \in \mathcal{V}(A^*),$$
  $$Lu^{-1} = \{w \in A^* \mid wu \in L\} \in \mathcal{V}(A^*),$$

- closed under inverse image by morphisms.

A **variety of finite monoids** is a class of finite monoids closed under

- submonoid,
- quotient monoid,
- finite direct products.

Theorem (Eilenberg, 1976)

There is a bijective and order preserving correspondence between varieties of languages and varieties of finite monoids.

Birkhoff Theorem

A **variety of monoids** is a class of (not necessarily finite) monoids closed under

- submonoid,
- quotient monoid,
- (not necessarily finite) direct products.

A monoid is commutative if $\forall x \forall y (xy = yx)$.

The commutative monoids form a variety, characterised by the equation $xy = yx$.

We can see xy and yx as words on the alphabet $A = \{x, y\}$.

A monoid M is commutative iff for all morphism $\varphi : \{x, y\}^* \to M$ we have $\varphi(xy) = \varphi(yx)$.

A monoid M **satisfies** the equation (u, v) if for all morphism $\varphi : A^* \to M$ we have $\varphi(u) = \varphi(v)$.

Birkhoff Theorem

A **variety of monoids** is a class of (not necessarily finite) monoids closed under

- submonoid,
- quotient monoid,
- (not necessarily finite) direct products.

An **equation** is a couple (u, v) of words on a finite alphabet A.

A monoid M **satisfies** the equation (u, v) if for all morphism $\varphi : A^* \to M$ we have $\varphi(u) = \varphi(v)$.

Theorem (Birkhoff, 1935)

A class of monoids is a variety if and only if it is definable by a set of equations.

Is there a similar characterisation for varieties of finite monoids?

Reiterman theorem

A profinite metric on A^*, $u, v \in A^*$:

$$d(u,v) = 2^{-\min\{|\mathcal{A}| \mid \mathcal{A} \text{ distinguishes } u \text{ and } v\}}$$

where an automaton $\mathcal{A}$ distinguishes u and v if it accepts one of them and rejects the other.

Theorem

The completion $(\widehat{A^*}, \mathcal{T}, \cdot)$ is a compact zero dimensional monoid called the **free profinite monoid** *on A.*

Points in $\widehat{A^*}$ are Cauchy sequences of finite words, called **profinite words**. Example: $x^\omega = \lim_{n \to \infty} x^{n!}$.

Reiterman theorem

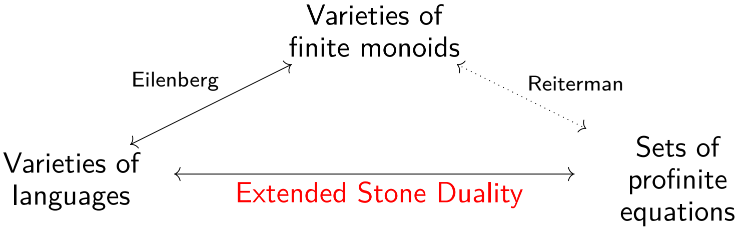
Define a **profinite equation** as a couple (u, v) of profinite words on a finite alphabet A.

A finite monoid M satisfies a profinite equation (u, v) if for all continuous monoid morphism $\widehat{\varphi} : \widehat{A^*} \to M$ we have $\widehat{\varphi}(u) = \widehat{\varphi}(v)$.

Theorem (Reiterman, 1982)

A class of finite monoids is a variety of finite monoids if and only if it is definable by a set of profinite equations.

Eilenberg-Reiterman theorem



Stone duality

A **Boolean algebra** is a structure $(B, \wedge, \vee, \{\}^{\complement}, 0, 1)$ s.t.

- $\wedge$, $\vee$ associative
- $\wedge$, $\vee$ commutative
- $\wedge$, $\vee$ distributive
- absorption $x \vee (x \wedge y) = x$
- complementation $x \wedge x^{\complement} = 0$, $x \vee x^{\complement} = 1$

Examples:

- field of sets, $\mathcal{P}(E)$
- $\mathrm{Rec}(A^*)$
- Boolean algebras of recognisable languages

A **Boolean space** is a topological space $(X, \mathcal{T})$ s.t.

- **Hausdorff**: distinct points are separated by neighbourhoods
- **Compact**: open covers contains finite subcovers
- **0-dimensional**: there is a clopen basis.

Examples:

- finite discrete spaces
- Cantor space 2^{ω}
- closed subspace of 2^{ω}

Stone duality

Let $(X, \mathcal{T})$ be a Boolean space:

$$\text{Clop}(X) = \{c \subseteq X \mid c \text{ closed and open for } \mathcal{T}\}$$

with the Boolean structure inherited from $\mathcal{P}(X)$.

Let $(B, \wedge, \vee, \{\}^{\complement}, 0, 1)$ be a Boolean algebra:

$$\text{Ult}(B) = \{u \subseteq B \mid u \text{ ultrafilter of } B\}$$

is a Boolean space for the topology generated by the sets of the form

$$\{u \in \text{Ult}(B) \mid b \in u\} \quad b \in B.$$

Theorem (Marshall H. Stone, 1936)

Every Boolean algebra is isomorphic to the algebra of clopen sets of a Boolean space. Furthermore

$$\text{Clop}(\text{Ult}(B)) = B \quad \textit{and} \quad \text{Ult}(\text{Clop}(X)) = X.$$

Stone duality and recognition

Examples

Boolean algebras (← clopen subsets)	Boolean spaces (ultrafilters →)
finite Boolean algebra	finite discrete space
finite or cofinite subsets of ω	one point compactification of ω
$\mathrm{Rec}(A^*)$	$(\widehat{A^*}, \mathcal{T})$
free BA on ω	Cantor space 2^ω
powerset of ω	Stone-Čech compactification $\beta\omega$

Theorem (Almeida, Pippenger, 1997)

The underlying topological space of the free profinite monoid on a finite set A is dual to the Boolean algebra of recognisable languages on A.

Stone duality

$$\text{For}\quad f : X \longrightarrow Y \quad \text{continuous}$$
$$\text{define}\quad \mathrm{Clop}(f) : \mathrm{Clop}(Y) \longrightarrow \mathrm{Clop}(X) \quad \text{by}$$
$$c \longmapsto f^{-1}(c)$$

$$\text{For}\quad h : A \longrightarrow B \quad \text{Boolean morphism}$$
$$\text{define}\quad \mathrm{Ult}(h) : \mathrm{Ult}(B) \longrightarrow \mathrm{Ult}(A) \quad \text{by}$$
$$u \longmapsto h^{-1}(u)$$

Theorem

$\mathrm{Clop}(f)$ *is a Boolean morphism and* $\mathrm{Ult}(h)$ *is continuous. Furthermore*

h *surjective* *iff* $\mathrm{Ult}(h)$ *injective*

h *injective* *iff* $\mathrm{Ult}(h)$ *surjective*

Stone duality and recognition

Boolean algebras	← clopen subsets / ultrafilters →	Boolean spaces
$(\mathrm{Rec}(A^*), \cap, \cup, \{\}^{\complement}, \emptyset, A^*)$		$(\widehat{A^*}, \mathcal{T})$
↑		↓
$(B, \cap, \cup, \{\}^{\complement}, \emptyset, A^*)$		$(\widehat{A^*}/E, \mathcal{T}/E)$

$B = \{L \in \mathrm{Rec}(A^*) \mid \forall (x, y) \in E \ (L \in x \leftrightarrow L \in y)\}$ $\overset{\text{Galois}}{\longleftrightarrow}$ connection $E = \{(x, y) \in \widehat{A^*} \times \widehat{A^*} \mid \forall L \in B \ (L \in x \leftrightarrow L \in y)\}$

A recognisable $L \subseteq A^*$ satisfies $x \leftrightarrow y$ if $L \in x \leftrightarrow y \in L$

Theorem (Gehrke,Grigorieff, Pin, 2008)

A set of recognisable languages on A is a Boolean algebra iff it can be defined by profinite equations of the form $x \leftrightarrow y$.

Extended Stone duality

Can duality account for the product on the free profinite monoid $(\widehat{A^*}, \mathcal{T}, \cdot)$?

Consider supplementary operations on the BA $\mathrm{Rec}(A^*)$.

The relevant operations are the **left** and **right** **residuals** by $M \in \mathrm{Rec}(A^*)$ are defined by

$$N \longmapsto M \backslash N = \{u \in A^* \mid \text{for all } v \in M,\ vu \in N\}$$
$$N \longmapsto N / M = \{u \in A^* \mid \text{for all } v \in M,\ uv \in N\}$$

These operations are characterised by the property that for all $L, M, N \in \mathrm{Rec}(A^*)$

$$L \subseteq N/M \quad \text{iff} \quad L \cdot M \subseteq N \quad \text{iff} \quad M \subseteq L \backslash N$$

Extended Stone duality

A **Boolean residuation algebra** is a Boolean algebra B with two binary operations $\backslash, / : B \times B \to B$ s.t.

- $(\bigvee_{\text{finite}} a_i)\backslash b = \bigwedge_{\text{finite}}(a_i\backslash b)$ and $b\backslash(\bigwedge_{\text{finite}} a_i) = \bigwedge_{\text{finite}}(b\backslash a_i)$
- $(\bigwedge_{\text{finite}} a_i)/b = \bigwedge_{\text{finite}}(a_i/b)$ and $b/(\bigwedge_{\text{finite}} a_i) = \bigvee_{\text{finite}}(b/a_i)$
- Galois property: $\forall a, b, c \in B \quad b \leq a\backslash c \leftrightarrow a \leq c/b$

Boolean algebras + operations $\longleftarrow$ Extended duality $\longrightarrow$ Boolean spaces + relations

For $x, y, z \in \mathrm{Ult}(B)$ define the relation

$$\begin{aligned} (x, y)Rz \quad &\text{iff} \quad \forall a, b \in B \ (b \in x \text{ and } a \notin z) \to b\backslash a \notin y \\ &\text{iff} \quad \forall a, b \in B \ (b \in y \text{ and } a \notin z) \to a/b \notin x \end{aligned}$$

Not functional from $\mathrm{Ult}(B) \times \mathrm{Ult}(B) \to \mathrm{Ult}(B)$ in general!

Extended Stone duality

Boolean algebras
+ operations

$\xleftarrow{\text{Extended}}$
$\xrightarrow[\text{Stone duality}]{}$

Boolean spaces
+ relations

$(\mathrm{Rec}(A^*), \cap, \cup, \{\}^{\complement}, \emptyset, A^*, \backslash, /)$

$(\widehat{A^*}, \mathcal{T}, R)$

$\backslash, /$

$(x, y)Rz \quad \text{iff} \quad x \cdot y = z$

functional !

Theorem (Gehrke, Grigorieff, Pin, 2008)

The dual space of the Boolean algebra of recognisable languages on A with residuals under extended Stone duality is the free profinite monoid on A.

Extended Stone duality and recognition

A **variety of languages** is the association to each finite alphabet A of a

- Boolean subalgebra $\mathcal{V}(A^*)$ of $\mathrm{Rec}(A^*)$,
- **closed under quotienting**, $\forall L \in \mathcal{V}(A^*)$ and $\forall u \in A^*$

  $$u^{-1}L = \{w \in A^* \mid uw \in L\} \in \mathcal{V}(A^*),$$
  $$Lu^{-1} = \{w \in A^* \mid wu \in L\} \in \mathcal{V}(A^*),$$

- closed under inverse image by morphisms.

A **Boolean residuation ideal** of $\mathrm{Rec}(A^*)$ is a Boolean subalgebra B of $\mathrm{Rec}(A^*)$ such that for all $L \in B$ and all $K \in \mathrm{Rec}(A^*)$

$$K \backslash L \in B,$$
$$L/K \in B.$$

Proposition

A Boolean algebra of recognisable languages is closed under quotienting iff it is a Boolean residuation ideal of $\mathrm{Rec}(A^*)$.

Extended Stone duality and recognition

$$(\mathrm{Rec}(A^*), \cap, \cup, \{\}^{\complement}, \emptyset, A^*, \backslash, /) \qquad \overset{\text{Extended duality}}{\longleftarrow\!\!\longrightarrow} \qquad (\widehat{A^*}, \mathcal{T}, \cdot)$$

residuation ideal $\uparrow$ $\qquad\qquad$ $\downarrow$ Stone monoid quotient

$$(B, \cap, \cup, \{\}^{\complement}, \emptyset, A^*, \backslash, /) \qquad\qquad\qquad (\widehat{A^*}/E, \cdot)$$

Theorem (Dekkers, 2008)

The Boolean residuation ideals of recognisable languages correspond dually to the profinite monoid quotients of $\widehat{A^*}$.

Say $L \in \mathrm{Rec}(A^*)$ satifies the equation of the form $x = y$, $(x, y) \in \widehat{A^*}$ if forall $u, v \in \widehat{A^*}$ it satisfies $uxv \leftrightarrow uyv$.

Theorem (Gehrke, Grigorieff, Pin, 2008)

A set of recognisable languages on A is a Boolean algebra closed under quotienting iff it can be defined by a set of profinite equations of the form $x = y$.