Part I: Duality and recognition

Dual spaces as completions of Pervin uniformities and their application to recognition of formal languages

Mai Gehrke

Radboud University Nijmegen

Joint work with Serge Grigorieff and Jean-Eric Pin

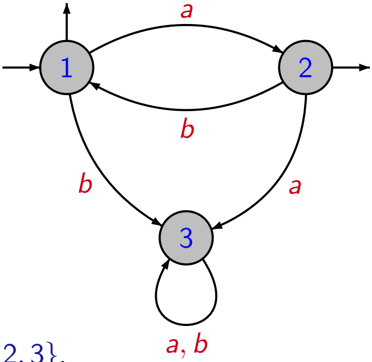
The syntactic monoid of a recognizable language and duality

Recognizable subsets, profinite completions, and duality

Reiterman's equational theory

Representation theory and Pervin uniform spaces

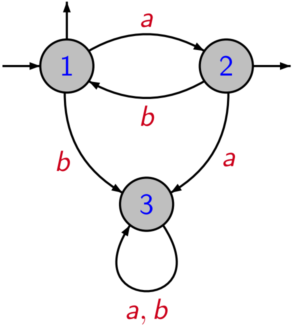
A finite state automaton



The states are $\{1, 2, 3\}$.
The initial state is 1, the final states are 1 and 2.
The alphabet is $A = \{a, b\}$ The transitions are

$$1 \cdot a = 2 \qquad 2 \cdot a = 3 \qquad 3 \cdot a = 3$$
$$1 \cdot b = 3 \qquad 2 \cdot b = 1 \qquad 3 \cdot b = 3$$

Recognizability



Transitions extend to words: $1 \cdot aba = 2$, $1 \cdot abb = 3$.
The language recognized by the automaton is the set of words u such that $1 \cdot u$ is a final state. Here:

$$L(\mathcal{A}) = (ab)^* \cup (ab)^* a$$

where $*$ means arbitrary iteration of the product.

Algebraic theory of automata

Given a language L, the syntactic monoid of L is given by

$$M(L) = A^*/\sim_L$$

where $\sim_L$ is the syntactic congruence of L, which is defined by

$$u \sim_L v \quad \text{if and only if} \quad \forall x, y \in A^* \ (xuy \in L \iff xvy \in L)$$

NB! It is not hard to see that $\varphi_L : A^* \to A^*/\sim_L$ is the furthest monoid quotient of A^* with $\varphi^{-1}(\varphi(L)) = L$.

<u>Theorem:</u> (Myhill '53, Rabin-Scott '59)
The syntactic monoid of a recognizable language is finite and there is an effective way of computing it.

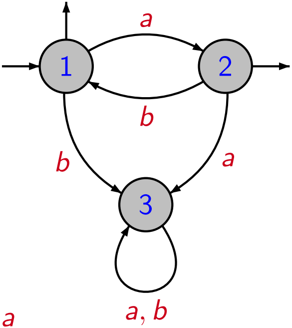
The syntactic monoid

Fact: Syntactic monoids provide a powerful tool in automata theory and yield decidability results for various classes of automata. They are definable for arbitrary languages but have mainly been successful for recognizable ones.

(Possibly a hint why it works well in the recognizable case:)

Theorem: [GGP2008] For a recognizable language L, the syntactic monoid of L is the dual space of a certain Boolean algebra with additional operations generated by L.

Quotient operations



$L(\mathcal{A}) = (ab)^* \cup (ab)^* a$

$$a^{-1}L = \{u \in A^* \mid au \in L\} = (ba)^* b \cup (ba)^*$$
$$La^{-1} = \{u \in A^* \mid ua \in L\} = (ab)^*$$
$$b^{-1}L = \{u \in A^* \mid bu \in L\} = \emptyset$$

Capturing the underlying machine

Given a recognizable language L the underlying machine is captured by the Boolean algebra $\mathcal{B}(L)$ of languages generated by

$$\{ x^{-1}Ly^{-1} \mid x, y \in A^* \}$$

NB! This generating set is finite since all the languages are recognized by the same machine with varying sets of initial and final states.

NB! $\mathcal{B}(L)$ is closed under quotients since the quotient operations commute will all the Boolean operations.

The residuation ideal generated by a language

Since $\mathcal{B}(L)$ is finite it is also closed under residuation with respect to arbitrary denominators.

For any $K \in \mathcal{B}(L)$ and any $S \in A^*$

$$S \backslash K = \bigcap_{u \in S} u^{-1}K \in \mathcal{B}(L)$$

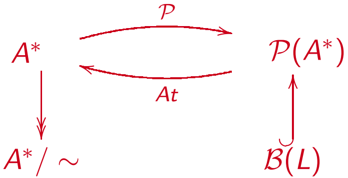
$$K/S = \bigcap_{u \in S} Ku^{-1} \in \mathcal{B}(L)$$

Theorem: [GGP2008] For a recognizable language L, the dual space of the algebra $(\mathcal{B}(L), \cap, \cup, (\ \ )^c, 0, 1, \backslash, /)$ is the syntactic monoid of L.

– *including the multiplication and all!*

The dual of the Boolean algebra $\mathcal{B}(L)$

Recall that $\mathcal{B}(L) = \langle x^{-1}Ly^{-1} \mid x, y \in A^* \rangle$.



where $u \sim v \iff \forall x, y \in A^* \; (u \in x^{-1}Ly^{-1} \iff v \in x^{-1}Ly^{-1})$
$\iff \forall x, y \in A^* \; (xuy \in L \iff xvy \in L)$

That is, $\sim = \sim_L$ and $M(L)$ is indeed the set underlying the dual of $\mathcal{B}(L)$.

Frobenius' complex algebras

Let A be an algebra with an n-ary operation $f : A^n \to A$
The operation lifts to the powerset

$$\begin{array}{rcl} f[\ ]: \ \mathcal{P}(A)^n & \to & \mathcal{P}(A) \\ (S_1, \ldots, S_n) & \mapsto & f[S_1 \times \ldots \times S_n] \end{array}$$

The complex algebra of A is

$$\mathbb{C}(A) = (\mathcal{P}(A), \cap, \cup, (\ )^c, 0, 1, f[\ ])$$

NB! The operation $f[\ ]$ is $\bigcup$-preserving in each coordinate

Residuation

A binary operation $\cdot : C \times C \to C$ is residuated provided there are operations

$$\backslash, / \ : C \times C \to C$$

satisfying

$$\forall\, a, b, c \in C \qquad (\ a \cdot b \leqslant c \quad \iff \quad b \leqslant a \backslash c \quad \iff \quad a \leqslant c/b \ )$$

Residuation

A binary operation $\cdot : C \times C \to C$ is residuated provided there are operations

$$\backslash, / \; : C \times C \to C$$

satisfying

$$\forall\, a, b, c \in C \qquad (\, a \cdot b \leqslant c \iff b \leqslant a \backslash c \iff a \leqslant c/b \,)$$

$$f_a(b) \leqslant c \iff b \leqslant g_a(c)$$

NB! $\cdot$ is residuated iff it is $\bigvee$-preserving in each coordinate

Residuated complex algebra

Given an abstract algebra A, the complex algebra $\mathbb{C}(A)$ is residuated, yielding (in the binary case)

$$
\begin{aligned}
K \cdot L &= \{uv \mid u \in K, v \in L\} \\
K \backslash M &= \{v \mid \forall u \in K\ (uv \in M)\} = \bigcap_{u \in K} u^{-1}M \\
M/L &= \{v \mid \forall w \in L\ (vw \in M)\} = \bigcap_{w \in L} Mw^{-1} \\
\{u\} \backslash M &= \{v \mid uv \in M\} = u^{-1}M \\
M/\{w\} &= \{v \mid vw \in M\} = Mw^{-1}
\end{aligned}
$$

Duals of operations – the finite distributive lattice case

A $\vee$-preserving operation $f : D \to D$ yields a binary relation R_f on $J(D)$ given by

$$R_f = \{(x, y) \mid x \leqslant f(y)\}.$$

It satisfies $\leqslant \circ R_f \circ \leqslant \; = R_f$.

We get a duality which, on the object level, is given by:

$$\begin{array}{ccc} (D, f) & \mapsto & (J(D), \leqslant, R_f) \\ (\mathcal{D}(X, \leqslant), R^{-1}[\;\;]) & \leftarrow\!\!\shortmid & (X, \leqslant, R) \end{array}$$

Here $R^{-1}[S] = \{x \mid \exists y \in S \; xRy\}$

NB! A UNARY operation corresponds to a BINARY relation

The dual of residuation operations

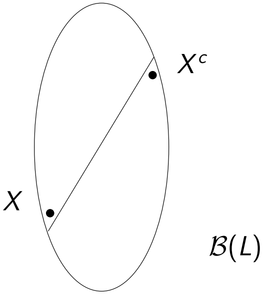
The residuation operation $\backslash : \mathcal{B} \times \mathcal{B} \to \mathcal{B}$

sends $\bigvee \mapsto \bigwedge$ in the first coordinate
sends $\bigwedge \mapsto \bigwedge$ in the second coordinate

$$R(X, Y, Z) \iff X\backslash(Z^c) \subseteq Y^c$$

$$\iff Y \not\subseteq X\backslash Z^c$$

$$\iff XY \not\subseteq Z^c$$



The dual of residuation operations

Now $XY \not\subseteq Z^c \iff XY \subseteq Z$ because $\sim = \sim_L$ is a congruence relation:

$$x \in X,\ y \in Y \text{ with } xy \in Z \implies XY \subseteq Z$$

That is, $(\mathcal{B}(L), \cap, \cup, (\ \ )^c, 0, 1, \backslash, /) = M(L)$ as required.

The dual of residuation operations

Now $XY \not\subseteq Z^c \iff XY \subseteq Z$ because $\sim = \sim_L$ is a congruence relation:

$$x \in X,\ y \in Y \text{ with } xy \in Z \quad \Longrightarrow \quad XY \subseteq Z$$

That is, $(\mathcal{B}(L), \cap, \cup, (\ \ )^c, 0, 1, \backslash, /) = M(L)$ as required.

NB! A language is recognized by an automaton if and only if it is recognized by a finite monoid in the sense that

$$L = \varphi^{-1}(P) \text{ where } \varphi : A^* \to M \supseteq P$$

as $(M, A, \{(m, a, m\varphi(a)) \mid m \in M, a \in A\}, \{1\}, P)$ is an automaton recognizing L.

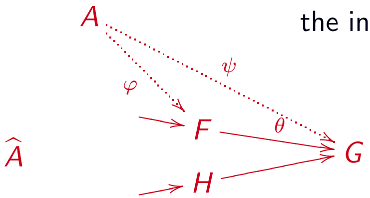
The recognizable subsets of an abstract algebra

$$\mathrm{Rec}(\mathrm{A}) \;=\; \{\varphi^{-1}(\mathrm{P}) \mid \varphi : \mathrm{A} \to \mathrm{F} \text{ hom, } \mathrm{F} \text{ finite, } \mathrm{P} \subseteq \mathrm{F}\}$$

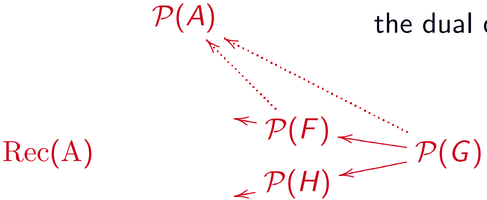
- $\emptyset\,,\ A\ \in\ \mathrm{Rec}(\mathrm{A})$
- $K, L \in \mathrm{Rec}(\mathrm{A}) \implies \mathrm{K} \cap \mathrm{L} \in \mathrm{Rec}(\mathrm{A})$
(recognized by the product of the two homomorphisms)
- $L \in \mathrm{Rec}(\mathrm{A}) \implies \mathrm{L}^{\mathrm{c}} \in \mathrm{Rec}(\mathrm{A})$
(recognized by the same hom with complementary subset)

$(\mathrm{Rec}(\mathrm{A}), \cap, \cup, (\ \ )^{\mathrm{c}}, \emptyset, \mathrm{A})$ is a Boolean algebra

Profinite completions and recognizable subsets



the inverse limit system $\mathcal{F}_A$



the dual direct limit system $\mathcal{G}_A$

$\mathrm{Rec}(A)$ as a subalgebra of $\mathbb{C}(A)$

The residuated complex algebra of (A, f)

$$\mathbb{C}(A) \;=\; (\mathcal{P}(A), f[\;\;], \{Res(f,i)\}_{i=1}^{n})$$

NB! $\mathrm{Rec}(A) \subseteq \mathbb{C}(A)$ MOSTLY NOT closed under the lifted operation, BUT

Proposition: The Boolean subalgebra $\mathrm{Rec}(A)$ is closed under $(S\backslash(\;\;),\ (\;\;)/S)_{S\in\mathcal{P}(A)}$.

Proof: For $L = \varphi^{-1}(P)$ we have $S\backslash L = \varphi^{-1}(\varphi(S)\backslash L)$

The dual of $(\mathrm{Rec}(A), /, \backslash)$

Theorem: [GGP2008] The dual space of

$$\mathrm{Rec}(A) + \text{residuals of liftings of operations}$$

is the profinite completion $\widehat{A}$ with its operations.

The dual of $(\mathrm{Rec(A)}, /, \backslash)$

Theorem: [GGP2008] The dual space of

$$\mathrm{Rec(A)}+\text{residuals of liftings of operations}$$

is the profinite completion $\widehat{A}$ with its operations.

In particular, the duals of the residual operations are functional and continuous. In binary case:

$$R_{(\backslash,/)} = \cdot : \widehat{A} \times \widehat{A} \to \widehat{A}$$

It is an open mapping iff $\mathrm{Rec(A)}$ is closed under the lifted multiplication.

Functional duals

Question: For which Boolean residuation ideals of

$$\mathcal{P}(A) + \text{ the residuals of the lifted operations}$$

is the dual of the residual operations functional?

Theorem: [GGP2010] For algebras A such as monoids, Boolean subalgebras B of $\mathcal{P}(A)$ closed under $(\{u\}\backslash(\ \ ), (\ \ )/\{u\})_{u \in A}$ have a functional dual if and only if B is contained in $\mathrm{Rec(A)}$.

Categorical dualities

subalgebras	$\longleftrightarrow$	quotient structures
quotient algebras	$\longleftrightarrow$	(generated) substructures
products	$\longleftrightarrow$	sums
sums	$\longleftrightarrow$	products

The mechanism behind Reiterman's theorem

Let A be an abstract algebra.

$\mathcal{L}$ a Boolean subalgebra (sublattice) of $\mathrm{Rec(A)}$
corresponds to
$E \subseteq \widehat{A} \times \widehat{A}$ a set of (in)equations in profinite terms

$\mathcal{L}$ a Boolean subalgebra (sublattice) of $\mathcal{P}(A)$
corresponds to
$E \subseteq \beta(A) \times \beta(A)$ a set of (in)equations in "β-terms"

A Galois connection for subsets of an algebra

Let B be a Boolean algebra, X the dual space of B.

The maps $\mathcal{P}(B) \leftrightarrows \mathcal{P}(X \times X)$ given by

$$S \mapsto \preceq_S = \{(x, y) \in X \mid \forall b \in S \quad (b \in y \Rightarrow b \in x)\}$$

and

$$E \mapsto B_E = \{b \in B \mid \forall (x, y) \in E \quad (b \in y \Rightarrow b \in x)\}$$

establish a Galois connection whose Galois closed sets are the compatible quasiorders and the bounded sublattices, respectively.

A Galois connection for subsets of an algebra

Let B be a Boolean algebra, X the dual space of B.

The maps $\mathcal{P}(B) \leftrightarrows \mathcal{P}(X \times X)$ given by

$$S \mapsto \approx_S = \{(x, y) \in X \mid \forall b \in S \quad (b \in y \iff b \in x)\}$$

and

$$E \mapsto B_E = \{b \in B \mid \forall (x, y) \in E \quad (b \in y \iff b \in x)\}$$

establish a Galois connection whose Galois closed sets are the compatible equivalence relations and the Boolean subalgebras, respectively.

Varying interpretations of equations

Consider a language $L \in \mathcal{P}(A^*)$ and $\mu, \nu \in \beta(A^*)$. Then L satisfies $\mu \leftrightarrow \nu$ provided

$$L \in \mu \iff L \in \nu$$

Varying interpretations of equations

Consider a language $L \in \mathcal{P}(A^*)$ and $\mu, \nu \in \beta(A^*)$. Then L satisfies $\mu \leftrightarrow \nu$ provided

$$L \in \mu \iff L \in \nu$$

If we think of $\mu \approx \nu$ as an equation of residuation ideals then the interpretation is

$$\forall x, y \in A^* \qquad (\ L \in x\mu y \iff L \in x\nu y\ )$$

or equivalently

$$\forall x, y \in A^* \qquad (\ x^{-1}Ly^{-1} \in \mu \iff x^{-1}Ly^{-1} \in \nu\ )$$

Varying interpretations of equations

If we think of $\mu = \nu$ as an equation of residuation ideals that is also invariant under substitution then

$$\varphi : A^* \to A^*$$
$$\varphi^{-1} : \mathcal{P}(A^*) \to \mathcal{P}(A^*)$$
$$S(\varphi^{-1}) : \beta(A^*) \to \beta(A^*)$$

and the interpretation is

$$\forall \varphi \, \forall x, y \in A^* \; (\, x^{-1} L y^{-1} \in S(\varphi^{-1})(\mu) \iff x^{-1} L y^{-1} \in S(\varphi^{-1})(\nu) \,)$$

e.g., if L is a commutative language it satisfies the substitution invariant equation $ab = ba$ (i.e., $\mu\nu = \nu\mu$ for all $\mu, \nu \in \beta(A^*)$)

The case of recognizable languages

In this case we may work at the level of $\widehat{A^*}$-equations. A recognizable language L satisfies $x = y$ corresponds to its syntactic monoid satisfying it.

$$f : A \to M$$
$$\varphi : A^* \to M$$
$$\varphi^{-1} : \mathcal{P}(M) \to \mathcal{P}(A^*)$$
$$S(\varphi^{-1}) : \beta(A^*) \to M$$

The case of recognizable languages

In this case we may work at the level of $\widehat{A^*}$-equations. A recognizable language L satisfies $x = y$ corresponds to its syntactic monoid satisfying it.

$$f : A \to M$$
$$\varphi : A^* \to M$$
$$\varphi^{-1} : \mathcal{P}(M) \to \mathcal{P}(A^*)$$
$$S(\varphi^{-1}) : \beta(A^*) \to M$$

E.g., there is an operation $(\quad)^\omega$ on $\widehat{A^*}$ which interprets in each finite monoid as the idempotent in the cyclic monoid generated by the element. The equation $x^\omega = x^{\omega+1}$ describes the star-free languages.

A fully modular Eilenberg-Reiterman theorem

Using the fact that sublattices of $\mathrm{Rec}(\mathrm{A}^*)$ correspond to Stone quotients of $\widehat{A^*}$ we get a vast generalization of the Eilenberg-Reiterman theory for recognizable languages

Closed under	**Equations**	**Definition**
$\cup, \cap$	$u \to v$	$\hat{\eta}(v) \in \hat{\eta}(L) \Rightarrow \hat{\eta}(u) \in \hat{\eta}(L)$
quotienting	$u \leqslant v$	for all $x, y,\ \ xuy \to xvy$
complement	$u \leftrightarrow v$	$u \to v$ and $v \to u$
quotienting and complement	$u = v$	for all $x, y,\ \ xuy \leftrightarrow xvy$
Closed under inverses of morphisms		**Interpretation of variables**
all morphisms		words
nonerasing morphisms		nonempty words
length multiplying morphisms		words of equal length
length preserving morphisms		letters

Eilenberg-Reiterman theory for arbitrary languages

The dual space of $\mathcal{P}(A^*)$ is the Stone-Čech compactification $\beta(A^*)$ of A^* as a discrete space.

Thus the sublattices of $\mathcal{P}(A^*)$ correspond to the Stone quotients of $\beta(A^*)$. We get theorem as for recognizable languages:

Closed under	**Equations**	**Definition**
$\cup, \cap$	$u \rightarrow v$	$v \in \widehat{L} \Rightarrow u \in \widehat{L}$
quotienting	$u \leqslant v$	for all x, y, $xuy \rightarrow xvy$
complement	$u \leftrightarrow v$	$u \rightarrow v$ and $v \rightarrow u$
quotienting and complement	$u = v$	for all x, y, $xuy \leftrightarrow xvy$
Closed under inverses of morphisms		**Interpretation of variables**
all morphisms		words
nonerasing morphisms		nonempty words
length multiplying morphisms		words of equal length
length preserving morphisms		letters

Concrete representations

An algebra of languages is more than just an abstract algebra. It is a concrete representation of an abstract algebra:

$$e : B \hookrightarrow \mathcal{P}(A)$$

This information is equivalent to

$$(A, \mathcal{B}) \quad \text{where} \quad \mathcal{B} = Im(e)$$

and dually it is equivalent to

$$A \to X_B \quad \text{where} \quad a \mapsto \mathcal{F}_a = \{L \in B \mid a \in e(L)\}$$

It may very well be that A and B (and thus $\mathcal{B}$) are countable while X_B is much bigger.

Concrete representations and Pervin uniformities

From a concrete representation $(A, \mathcal{B})$ we can make blocks

$$\mathcal{B} \ni L \quad \mapsto \quad (L \times L) \cup (L^c \times L^c)$$

and obtain a Pervin uniform space

$$(A, \mathcal{U}_{\mathcal{B}}) \text{ where } \mathcal{U}_{\mathcal{B}} = \langle (L \times L) \cup (L^c \times L^c) \mid L \in \mathcal{B} \rangle$$

Proposition: Generating a uniformity does NOT add blocks in the sense that $L \subseteq A$ is a block of $\mathcal{U}_{\mathcal{B}}$ iff $L \in \mathcal{B}$.

Pervin uniform spaces and Stone duals

Given a Pervin uniform space $(A, \mathcal{U}_\mathcal{B})$ its Hausdorff completion

$$(A, \mathcal{U}_\mathcal{B}) \to (X, \mathcal{U}_\mathcal{B})$$

yields a compact topological space.

Thus uniformity and topology carry the same information and in fact $(X, \mathcal{U}_\mathcal{B})$ is the Stone dual space of $\mathcal{B}$.

That is, in a natural way, we recover

$$A \to X_\mathcal{B} \quad \text{where} \quad a \mapsto \mathcal{F}_a = \{L \in \mathcal{B} \mid a \in L\}$$

Conclusions

- Stone duality yields canonical representations/recognizing objects
- The dual of binary residuation on regular languages is FUNCTIONAL
- (Interesting) functional duals is closely linked to (finite) recognition and is a new phenomenon for duality theory
- Equations à la Reiterman may be seen as a special case of the duality

$$\text{subalgebras} \quad \leftrightarrow \quad \text{quotient spaces}$$

- The theory of Pervin uniform spaces provides an ideal setting for the study of concrete representation/recognition and associated dual spaces