Duality and recognition

Mai Gehrke

Radboud University Nijmegen, The Netherlands

Abstract. The fact that one can associate a finite monoid with universal properties to each language recognised by an automaton is central to the solution of many practical and theoretical problems in automata theory. It is particularly useful, via the advanced theory initiated by Eilenberg and Reiterman, in separating various complexity classes and, in some cases it leads to decidability of such classes. In joint work with Jean-Éric Pin and Serge Grigorieff we have shown that this theory may be seen as a special case of Stone duality for Boolean algebras extended to a duality between Boolean algebras with additional operations and Stone spaces equipped with Kripke style relations. This is a duality which also plays a fundamental role in other parts of the foundations of computer science, including in modal logic and in domain theory. In this talk I will give a general introduction to Stone duality and explain what this has to do with the connection between regular languages and monoids.

1 Stone duality

Stone type dualities is *the* fundamental tool for moving between linguistic specification and spatial dynamics or transitional unfolding. As such, it should come as no surprise that it is a theory of central importance in the foundations of computer science where one necessarily is dealing with syntactic specifications and their effect on physical computing systems.

In 1936, M. H. Stone initiated duality theory by presenting what, in modern terms, is a dual equivalence between the category of Boolean algebras and the category of compact Hausdorff spaces having a basis of clopen sets, so-called Boolean spaces [13]. The points of the space corresponding to a given Boolean algebra are not in general elements of the algebra – just like states of a system are not in general available as entities in a specification language but are of an entirely different sort. In models of computation these two different sorts, specification expressions and states, are given a priori but in unrelated settings. Via Stone duality, the points of the space may be obtained from the algebra as homomorphisms into the two-element Boolean algebra or equivalently as ultrafilters of the algebra. In logical terms these are valuations or models of the Boolean algebra. In computational terms they are possible states of the system. Each element of the Boolean algebra corresponds to the set of all models in which it is true, or all states in which it holds, and the topology of the space is generated by these sets. A main insight of Stone is that one may recover the original algebra as the Boolean algebra of clopen subsets of the resulting space.

In Boole's original conception, Boolean algebras were meant to capture the arithmetic of propositions and he thought of propositions as 'classes' or sets of entities modelling them. In this sense Stone's theorem closes the circle by showing that every Boolean algebra is indeed isomorphic to a field of sets with the set theoretic operations of intersection, union, and complement as the Boolean operations. Stone duality is thus, in part, a representation theorem showing that the axioms of Boolean algebras exactly capture the fields of sets just like Cayley's theorem shows that the axioms of groups exactly capture the groups of permutations.

However, the fact that, with the topology in play, we obtain mathematical objects with their own and very separate theory and intuitions which fully capture the original Boolean algebras as well as their morphisms is the real power of Stone duality. The *duality* (as opposed to equivalence) aspect turns more complicated constructions such as quotients into simpler ones such as subobjects, it turns additional connectives on the algebras into transition structure on state spaces. This ability to translate faithfully between algebraic specification and spacial dynamics has often proved itself to be a powerful theoretical tool as well as a handle for making practical problems decidable. This principle was applied first by Stone himself in functional analysis, followed by Grothendieck in algebraic geometry who represented rings in terms of sheaves over the dual spaces of distributive lattices (i.e., 'positive' Boolean algebras) and has since, over and over again, proved itself central in logic and its applications in computer science. One may specifically mention Scott's model of the λ-calculus, which is a dual space, Esakia's duality [4] for Heyting algebras and the corresponding frame semantics for intuitionist logics, Goldblatt's paper [8] identifying extended Stone duality as the theory for completeness issues for Kripke semantics in modal logic, and Abramsky's path-breaking paper [1] linking program logic and domain theory. Our work with Grigorieff and Pin [7, 9, 6], with Pippenger [10] as a precursor, shows that the connection between regular languages and monoids also is a case of Stone duality.

1.1 Duality for finite distributive lattices

Lattices are partial orders with infima (meets) and suprema (joins) of finite sets, but may also be seen as algebras $(L, \wedge, \vee, 0, 1)$ satisfying certain equations, see [2] for the basics of lattice theory. A lattice is distributive provided the binary meet ($\wedge$) and the binary join ($\vee$) distribute over each other. Distributive lattices corresponds to the negation-free reduct of classical propositional logic, and if in a distributive lattice every element a has a complement (that is, an element b so that $a \wedge b = 0$ and $a \vee b = 1$) then the lattice is a Boolean algebra.

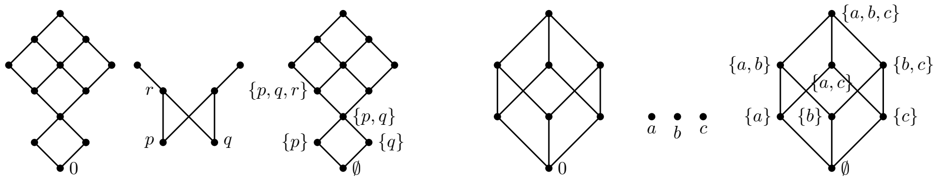
The restriction of Stone's duality for distributive lattices to finite objects yields a duality between finite posets and finite distributive lattices which restricts further to a duality between finite sets and finite Boolean algebras. This duality was a precursor to Stone's duality and is due to Birkhoff. We begin with a description of Birkhoff duality as the essential features are easiest to understand in this setting. This duality is based on the fact that each element in a finite lattice is the join of all join irreducible elements below it and that all down-sets of join irreducible elements yield distinct elements if the lattice is distributive. The component facts are proved in [2, Lemma 5.11, page 117].

Definition 1. *An element p in a bounded lattice D is join irreducible provided $p \neq 0$ and $p = x \vee y$ in D implies $p = x$ or $p = y$. An element p in a bounded lattice D is join prime provided $p \neq 0$ and $p \leqslant x \vee y$ in D implies $p \leqslant x$ or $p \leqslant y$.*

We denote by $\mathcal{D}_{fin}$ the category of finite bounded distributive lattices with bounded lattice homomorphisms and by $\mathcal{P}_{fin}$ the category of finite posets with order preserving maps. Birkhoff duality is given by two functors

$$J : \mathcal{D}_{fin} \to \mathcal{P}_{fin} \qquad \text{and} \qquad H : \mathcal{P}_{fin} \to \mathcal{D}_{fin}$$

that establish the dual equivalence of the two categories. The functor J sends a finite bounded distributive lattice D, to the poset $J(D)$ of join irreducible elements of D with the order induced from D. For a finite poset P, the dual lattice $H(P)$ is the lattice of all down-sets of P with meet and join given by intersection and union. On the object level the dual equivalence of the categories $\mathcal{D}_{fin}$ and $\mathcal{P}_{fin}$ is given by the isomorphisms: $D \cong H(J(D)), a \mapsto {\downarrow} a \cap J(D)$ and $P \cong J(H(P)), p \mapsto {\downarrow} p$, see [2, Chapter 5]. The following figure provides two examples. Note that an element of a Boolean algebra is join irreducible if and only if it is an atom, i.e., an element right above 0, and thus the dual of a Boolean algebra is just a set.



D $J(D)$ $H(J(D))$ B $J(B)$ $H(J(B))$

The fact that the above correspondence extends to a categorical duality is what makes it so powerful. In order to specify the categorical duality we have also to give the correspondence between the morphisms in the two categories. This correspondence is essentially based on the notion of adjoint maps.

Definition 2. *Let D and E be finite lattices. Let $f : D \to E$ and $g : E \to D$ be functions satisfying for all $d \in D$ and for all $e \in E$:*

$$f(d) \leqslant e \qquad \iff \qquad d \leqslant g(e).$$

Then g is called an upper adjoint of f and f is called a lower adjoint of g.

It is easy to see that adjoints are unique when they exist and that a map between complete lattices has an upper adjoint if and only if it preserves arbitrary joins and order dually for lower adjoints. If f has an upper adjoint, we will denote it by $f^\sharp$ and if g has a lower adjoint, we will denote it by $g^\flat$. Note that a bounded lattice homomorphism between finite lattices $h : D \to E$ preserves arbitrary joins and meets. So it has both an upper adjoint and a lower adjoint. The duality for maps is based on the fact that a map $f : E \to D$ (such as $h^\flat$) has an upper adjoint which has an upper adjoint if and only if it sends join irreducible elements to join irreducible elements.

Definition 3. *Let D and E be finite distributive lattices, $h : D \to E$ a bounded lattice homomorphism. The dual of h is*

$$J(h) = h^\flat \restriction J(E),$$

that is, the restriction of the lower adjoint $h^\flat$ of h viewed as a map from $J(E)$ to $J(D)$. For finite posets P and Q and $f : P \to Q$ an order preserving map, we define $H(f) = (f^\to)^\sharp$ where $f^\to : H(P) \to H(Q)$ is the forward image map, $S \mapsto f[S]$. Note that $H(f) = (f^\to)^\sharp$ is then actually the inverse image map $T \mapsto f^{-1}(T)$ because the inverse image map is the upper adjoint of the forward image map.

Using the uniqueness of upper and lower adjoints, it is easy to show that J and H on morphisms in the two categories establish one-to-one correspondences as needed for the duality.

In closing, we note that the functors J and H can be extended to a duality between the category $\mathcal{DL}^+$ of down-set lattices with complete lattice homomorphisms and the category $\mathcal{P}$ of posets with order preserving maps by replacing binary meets and joins by arbitrary ones in the definitions above. However, this duality does not encompass distributive lattices in general (as can be seen, e.g. from the example at the end of the next subsection).

1.2 Duality for bounded distributive lattices

The basic idea of the dualities is to represent a distributive lattice by its set of join irreducible elements. However, for infinite lattices, there may not be enough of these, and idealised elements, in the form of ideals or filters, must be considered. Let D be a bounded distributive lattice. A subset I of D is an *ideal* provided it is a down-set closed under finite joins. We denote by $Idl(D)$ the set of all ideals of D partially ordered by inclusion. The embedding $D \to Idl(D), a \mapsto {\downarrow}a$ is the free $\bigvee$-completion of the lattice D. In this sense one should think of an ideal as standing for the element which would be the supremum of the ideal if the supremum existed. A subset F of D is a *filter* provided it is an up-set closed under finite meets. We denote by $Filt(D)$ the partially ordered set of all filters of D. Filters represent (possibly non-existing) infima and thus the order on filters is given by *reverse* inclusion. The embedding $D \to Filt(D), a \mapsto {\uparrow}a$ is the free $\bigwedge$-completion of the lattice D. An ideal or filter is *proper* provided it isn't the entire lattice. A proper ideal I is *prime* provided $a \wedge b \in I$ implies $a \in I$ or $b \in I$. A proper filter F is *prime* provided $a \vee b \in F$ implies $a \in F$ or $b \in F$.

Note that a filter is prime if and only if its complement is a (prime) ideal so that prime filters and prime ideals come in complementary pairs. In particular this means that the set of prime ideals with the inclusion order is isomorphic to the set of prime filters with the reverse inclusion order. For a bounded distributive lattice D we will denote this partially ordered set by X_D or just X. Since there are so many set theoretic levels, we will revert to lower case letters $x, y, z \ldots$ for elements of X and to make clear when we talk about the corresponding prime filter or the complementary ideal we will denote these by F_x and I_x, respectively. In the case of a finite distributive lattice, filters and ideals are all principal generated by their meet and their join, respectively. In this case, the meets of prime filters are exactly the join prime elements of the lattice while the joins of the prime ideals are exactly the meet prime elements of the lattice. Thus these come in pairs $p, \kappa(p) = \bigwedge\{a \in D \mid p \nleqslant a\}$ which split the lattice in two disjoint pieces, that is,

$$\forall a \in D \qquad (\ p \nleqslant a \qquad \iff \qquad a \leqslant \kappa(p)\ )$$

In a finite Boolean algebra, the meet of a prime filter is necessarily an atom while a meet irreducible is a co-atom and $\kappa(p) = \neg p$ in finite Boolean algebras.

In the infinite case prime filters play the role of the join irreducible elements, and it is not hard to verify that the following map is a bounded lattice homomorphism

$$\begin{aligned} \eta_D : D &\to \mathcal{P}(X_D) \\ a &\mapsto \eta_D(a) = \{x \in X_D \mid a \in F_x\} \end{aligned}$$

Using the Axiom of Choice one may in addition show that any distributive lattice has enough prime filters in the sense that this map also is injective.

One may also show that the sets in the image of η_D are down-sets in the reverse order of inclusion. However, for an infinite distributive lattice, it is never the set of all such down-sets. Stone's insight was to generate a topology with the sets in the image of η_D. This works but yields a non-Hausdorff space in the non-Boolean case. A slight variant of Stone duality was later developed by Priestley and this is what we will use here. The (Priestley) dual space of bounded distributive lattice D is the ordered topological space $(X_D, \leqslant, \pi)$ where X_D is the set of prime filters of D under reverse inclusion order and π is the topology on X_D generated by the subbasis

$$\{\eta_D(a), (\eta_D(a))^c \mid a \in D\}.$$

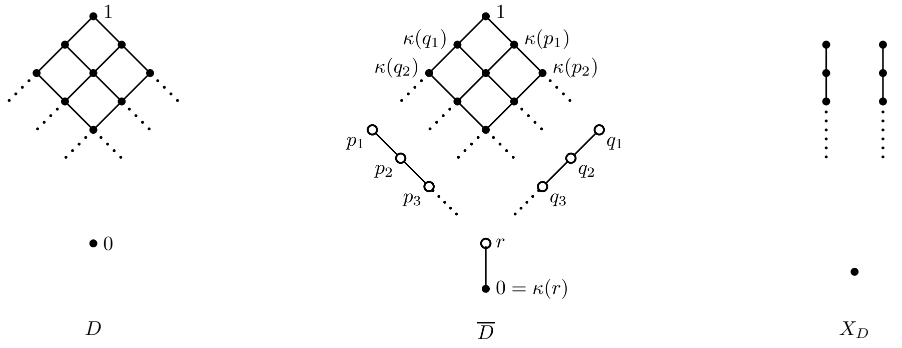
The space $(X_D, \leqslant, \pi)$ is then compact and *totally order disconnected*, that is, for $x, y \in X_D$ with $x \nleqslant y$ there is a clopen down-set U with $y \in U$ and $x \notin U$. The dual of a homomorphism $h : D \to E$ is the restriction of the inverse image map to prime filters, $h^{-1} : X_E \to X_D$, and, for any homomorphism $h : D \to E$, the map $h^{-1} : X_E \to X_D$ is continuous and order preserving.

A *Priestley space* is an ordered topological space that is compact and totally order disconnected and the morphisms of Priestley spaces are the order preserving continuous maps. The dual of a Priestley space $(X, \leqslant, \pi)$ is the bounded distributive lattice $ClopD(X, \leqslant, \pi)$ of all subsets of X that are simultaneously clopen and are down-sets. For $f : X \to Y$ a morphism of Priestley spaces, the restriction of the inverse image map to clopen down-sets, $f^{-1} : ClopD(Y) \to ClopD(X)$, is a bounded lattice homomorphism and is the dual of f under Priestley duality.

This accounts for Priestley duality. The point is that, for each distributive lattice D, the lattice $ClopD(X_D, \leqslant, \pi)$ is isomorphic to D via the map η_D as described above and these isomorphisms transform homomorphisms between lattices into their double dual homomorphisms. Similarly, any Priestley space is order isomorphic and homeomorphic to its double dual via the map which assigns to any point its neighbourhood filter and the double duals of order preserving continuous functions are naturally isomorphic to the original maps. This very tight relationship between the two categories allows one to translate essentially all structure, concepts, and problems back and forth between the two sides of the duality.

Note that in the case where the lattice D is a Boolean algebra, that is, each element a has a complement $\neg a$, then the order on prime filters (which are in this case the same as the ultrafilters or the maximal proper filters) is trivial, and since $(\eta_D(a))^c = \eta_D(\neg a)$, the image of η is already closed under complementation. In this case, the Priestley duality agrees with the original Stone duality and we may refer to it as Stone duality rather than as Priestley duality.

We close this subsection with an example. Let $D = \mathbf{0} \oplus (\mathbb{N}^{op} \times \mathbb{N}^{op})$ be the first lattice in the figure below. Note that D has *no* join irreducible elements whatsoever. The prime filters of D correspond to the hollow points in the lattice $\overline{D}$ (by taking the intersection with D of their individual up-sets) and the prime ideals of D are all principal down-sets given by the points as marked with κ's. The dual space X_D consists of two chains with a common lower bound, the image of η_D consists of the cofinite down-sets, and the topology is that of the one point compactification of a discrete space where the limit point is the least element. We recover D as the clopen down-sets.



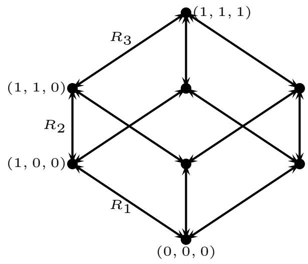
1.3 Duality for additional operations

Any further structure on either side of the duality may be translated as corresponding structure on the other side. The translation of additional operations on lattices and Boolean algebras to their dual spaces is particularly important. We give the simplest but prototypical example in the finite setting. Let B be a finite Boolean algebra and $f : B \to B$ a unary join and 0 preserving operation on B. This is usually called a normal modal (possibility) operator and it comes about in many applications. In order to dualise such an operation in case B is finite, we just need to know where the elements of $J(B)$ get sent. That is,

$$R_f = \{(x, y) \in X_B \times X_B \mid x \leqslant f(y)\}$$

encodes f as a binary relation on the dual of B. This relation is not a function unless f actually was a homomorphism on B. We illustrate this with an example.

Example 1. Consider the situation of the muddy children puzzle: there are n children each of whom may or may not have a muddy forehead thus giving rise to the Boolean algebra B whose atoms are the complete conjunctions over these n statements. Each child can see the foreheads of the others but not his own and we want to consider modal operators $<i>$, for each i from 1 to n, where $<i>\phi$ means ϕ is possible according to child i.



X_B

The dual space of B is the set of the 2^n atoms of B which may be thought of as n tuples each specifying which children have muddy foreheads and which don't. As explained above, the modal operators, $<i>$, are given dually by relations R_i where xR_iy for $x, y \in 2^n$ if and only if $x \leqslant <i>y$. Since the order in a Boolean algebra is the order of implication and x implies $<i>y$ precisely when x and y differ at most in the i^{th} coordinate, the relational image of each point in 2^n consists of precisely two points. In the case of three children for example, the dual space has 8 elements and the three relations each partition the points in two element sets along each of the three dimensions. Thus R_2 identifies points vertically above/below each other. This dual structure is quite simple and it is indeed also what one usually works with when analysing the associated dynamic epistemic puzzle [5].

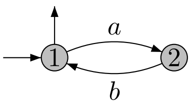
For infinite lattices one first has to extend the operation to be dualised to the filters or the ideals (depending on whether it preserves join or meet) in order to have the operation defined on prime filters or ideals and thus on points of the dual space. Despite this slight complication, we get here too, for an n-ary operation, an $(n + 1)$-ary relation on the dual space. The dual relations will have appropriate topological properties. For a unary modality these amount to R being point-closed ($R[x] = \{y \mid xRy\}$ is closed for each x) and pre-images of clopens are clopen ($R^{-1}[U] = \{x \mid \exists y \in U \text{ with } xRy\}$ is clopen for each clopen U). One can also describe the duals of morphisms of lattices with additional operations. These are are often called bounded morphisms or p-morphisms. Altogether this yields what is known as *extended Stone or Priestley duality.* The details for a fairly large class of additional operations may be found in the first section of [8].

2 Monoids and recognition by automata

The starting point of the algebraic approach to automata theory is the classical result that one can effectively assign a finite monoid to each finite state automaton. We give a derivation of this basic result using extended duality.

An *automaton* is a structure $\mathcal{A} = (Q, A, \delta, I, F)$ where Q is a finite set whose elements are called *states*, A is a finite alphabet for the automaton, and $\delta \subseteq Q \times A \times Q$ is the *transition relation* which specifies the transition behaviour of the machine when it is fed a letter in a given state. The set I of *initial states* is a subset of Q as is the set F of *final states*.
We denote by A^* the free monoid over the alphabet A and any subset $L \subseteq A^*$ is called a *language* over A^*. The *language recognised by* $\mathcal{A}$, denoted $L(\mathcal{A})$, is the subset of A^* of all words $a_1 \dots a_n$ over the alphabet A such that there are states $q_0, \dots, q_n$ in Q with $(q_{i-1}, a_i, q_i) \in \delta$ for each i with $1 \leqslant i \leqslant n$ and $q_0 \in I$ and $q_n \in F$.

Example 2. Let $\mathcal{A} = (Q, A, \delta, I, F)$ where $Q = \{1, 2\}, A = \{a, b\}$, and δ is as specified in the picture. That is, $(q, x, q') \in \delta$ if and only if there is an arrow from q to q' labelled by x. The initial and final states are $I = \{1\} = F$.



The language recognised by $\mathcal{A}$ consists of all those words that may be read by starting in 1 and ending in state 1. That is, $L(\mathcal{A}) = (ab)^*$ where S^* denotes the submonoid generated by $S \subseteq A^*$ and $u^* = \{u\}^*$ for a word $u \in A^*$.

There may be many different automata that produce a given language but some languages recognised by automata require inherently more complex machines than others. A fundamental insight is that we can get at the essential features of the machines recognising a given language in a purely algebraic way from the language. As we shall see, we may think of the underlying transition system of an automaton as a kind of state space, and the languages recognised by it with various choices of initial and final states as a dual algebra of sets. Then, given what we know about duality, it should come as no surprise that the operations on languages dual to concatenation are given by adjunction.

Let A be a finite alphabet. The concatenation operation on A^* gives rise to a *residuated* or *adjoint family* of operations on the set of all languages over A^* as follows. *Complex* or *lifted concatenation* on $\mathcal{P}(A^*)$ is given by

$$KL = \{uv \mid u \in K \text{ and } v \in L\}. \tag{1}$$

The *residuals* of this operation are uniquely determined by the *residuation* or *adjunction laws*:

$$\forall K, L, M \in \mathcal{P}(A^*) \qquad \begin{aligned} KM \subseteq L \quad &\iff \quad M \subseteq K\backslash L \\ &\iff \quad K \subseteq L/M. \end{aligned} \tag{2}$$

One easily sees from this that $K\backslash L = \{u \in A^* \mid \forall v \in K \; vu \in L\}$. In particular, for $K = \{x\}$ a singleton language $x\backslash L = \{u \in A^* \mid xu \in L\}$. The operations

$L \mapsto x\backslash L$ are widely used in language theory and usually $x\backslash L$ is denoted by $x^{-1}L$ and these operations are referred to as *quotients.*

One may now easily verify that the quotient operations on the left and the right correspond to moving the initial and final states along words respectively.

Proposition 1. *Let $L = L(\mathcal{A})$ be a language recognised by an automaton $\mathcal{A} = (Q, A, \delta, I, F)$. Then the languages $x^{-1}Ly^{-1}$ for $x, y \in A^*$ are recognised by automata $\mathcal{A}' = (Q, A, \delta, I', F')$ obtained from $\mathcal{A}$ by altering only the sets of initial and final states. Consequently, the set*

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

is finite.

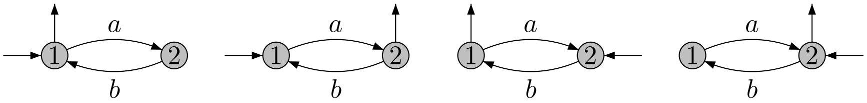
Definition 4. *Let A be a finite alphabet and $L \subseteq A^*$ a language over A. Let $\mathcal{B}(L)$ be the Boolean subalgebra of $\mathcal{P}(A^*)$ generated by the set $\{x^{-1}Ly^{-1} \mid x, y \in A^*\}$. We will call $\mathcal{B}(L)$ the* quotienting ideal *generated by L. More generally a quotienting ideal of $\mathcal{P}(A^*)$ is a Boolean subalgebra which is closed under the quotienting operations $x^{-1}(\ )$ and $(\ )y^{-1}$ for all $x, y \in A^*$.*

Using the fact that the quotienting operations $x^{-1}(\ )$ preserve all the Boolean operations and that $S\backslash(\ ) = \bigcap_{x \in S} x^{-1}(\ )$ (and the same on the right), we can then prove the following proposition.

Proposition 2. *Let L be a language recognised by some automaton. Then $\mathcal{B}(L)$ is closed under the operations $S\backslash(\ )$ and $(\ )/S$ for all $S \subseteq A^*$. In particular, $\mathcal{B}(L)$ is closed under the binary operations $\backslash$ and $/$.*

The stronger property that the Boolean subalgebra $\mathcal{B}(L)$ of $\mathcal{P}(A^*)$ has of being closed under residuation with arbitrary denominators, we call being a *residuation ideal.*

Example 3. For the language L of Example 2, it is clear that moving the final and initial states around along transitions yields the four automata given below corresponding to L, Lb^{-1}, $a^{-1}L$, and $a^{-1}Lb^{-1}$, respectively.



Thus $\mathcal{B}(L)$ is the Boolean subalgebra of $\mathcal{P}(A^*)$ generated by these four languages. It is not hard to see that this is the Boolean algebra generated by the atoms 1, $(ab)^+$, $a(ba)^*$, $b(ab)^*$, $(ba)^+$, and 0, where $1 = \{\varepsilon\} = L \cap a^{-1}Lb^{-1}$, and 0 is the complement of the union of the four generating languages. Note that $\mathcal{B}(L)$ is *not* closed under the lifted multiplication.

Theorem 1. *Let L be a language recognised by an automaton, then the extended dual of the Boolean algebra with additional operations $(\mathcal{B}(L), \backslash, /)$ is the syntactic monoid of L. In particular, it follows that the syntactic monoid of L is finite and is effectively computable.*

Proof. It is not hard to see that the atoms of the Boolean algebra generated by the finite collection $\mathcal{C} = \{x^{-1}Ly^{-1} \mid x, y \in A^*\}$ are the equivalence classes of the finite indexed equivalence relation

$$\begin{aligned} u \approx_L v \quad & \text{if and only if} \quad \forall x, y \in A^* \ (u \in x^{-1}Ly^{-1} \iff v \in x^{-1}Ly^{-1}) \\ & \text{if and only if} \quad \forall x, y \in A^* \ (xuy \in L \iff xvy \in L) \end{aligned}$$

and the set $A^*/\approx_L = S(L)$ is in fact the set underlying the syntactic monoid of L. It is a general fact that all the operations of a residuated family have the same dual relation up to the order of the coordinates. So we focus on the operation $\backslash$. It turns joins in the first coordinate into meets and meets in the second coordinate into meets. For this reason some swapping between join irreducible and meet irreducible elements using κ is needed. For $X, Y, Z \in A^*/\approx_L$ we have

$$\begin{aligned} R_\backslash(X, Y, Z) & \iff X\backslash\kappa(Y) \subseteq \kappa(Z) \\ & \iff X\backslash(Y^c) \subseteq Z^c \\ & \iff Z \not\subseteq X\backslash Y^c \\ & \iff XZ \not\subseteq Y^c \\ & \iff \exists x \in X, z \in Z \text{ with } xz \in Y \\ & \iff \exists x, z \text{ with } X = [x]_{\approx_L}, Z = [z]_{\approx_L}, Y = [xz]_{\approx_L} \end{aligned}$$

so that $R_\backslash$ is the graph of the operation on the quotient. Thus the dual space $(X_{\mathcal{B}(L)}, R_\backslash)$ is the quotient monoid $(A^*/\approx_L, \cdot/\approx_L)$. □

Example 4. Continuing with our running example $L = (ab)^*$, we have seen that the Boolean algebra $\mathcal{B}(L)$ has six atoms, namely $1 = \{\varepsilon\}$, $(ab)^+$, $a(ba)^*$, $b(ab)^*$, $(ba)^+$, and 0 (the latter consisting of all words in A^* containing two consecutive identical letters). Note that the product of two languages in $\mathcal{B}(L)$ may intersect several languages in $\mathcal{B}(L)$ (e.g., $(ab)^*(ba)^*$ intersects 1, $(ab)^+$, $(ba)^+$, and 0). However, the product of any two of the atoms is entirely contained in a unique other atom (while we didn't quite prove that above, it is a consequence of what we proved). It should be clear that in this example, the element 1 will be the neutral element, 0 will be absorbing, and the multiplication of the remaining four elements will be as given in the adjoining table.

	$(ab)^+$	$a(ba)^*$	$b(ab)^*$	$(ba)^+$
$(ab)^+$	$(ab)^+$	$a(ba)^*$	0	0
$a(ba)^*$	0	0	$(ab)^+$	$a(ba)^*$
$b(ab)^*$	$b(ab)^*$	$(ba)^+$	0	0
$(ba)^+$	0	0	$b(ab)^*$	$(ba)^+$

Duality theory is not just about objects but also about maps, and it is straight forward to check that the dual of the inclusion map $\iota : \mathcal{B}(L) \to \mathcal{P}(A^*)$ is the quotient map $\varphi_L : A^* \to A^*/\approx_L$, where $A^*/\approx_L = S(L)$ is the syntactic monoid of L. The content of this fact is that the dual of φ_L, which is $\varphi_L^{-1} : \mathcal{P}(S(L)) \to \mathcal{P}(A^*)$, is naturally isomorphic to $\iota : \mathcal{B}(L) \to \mathcal{P}(A^*)$:

$$\begin{tikzcd}
\mathcal{B}(L) \arrow[r, "\iota"] \arrow[d, leftrightarrow, "\cong"] & \mathcal{P}(A^*) \arrow[d, leftrightarrow, "="] \\
\mathcal{P}(S(L)) \arrow[r, "\varphi_L^{-1}"] & \mathcal{P}(A^*).
\end{tikzcd}$$

This in turn precisely means that $\mathcal{B}(L) = \{\varphi_L^{-1}(P) \mid P \subseteq S(L)\}$.

3 Recognisable subsets of an algebra and profinite completions

Let A be any algebra, $\varphi : A \to B$ a homomorphism into a finite algebra. A subset $L \subseteq A$ is said to be *recognised by* φ provided there is a subset $P \subseteq B$ with $L = \varphi^{-1}(P)$. A subset $L \subseteq A$ is said to be *recognised by* B provided there is a homomorphism $\varphi : A \to B$ which recognises L. The last observation of the previous section can then be phrased as saying that the residuation ideal generated by a regular language L consists precisely of those languages that are recognised by the quotient map φ_L onto the syntactic monoid of L and, in particular, that every language recognised by an automaton also is recognised by a finite monoid. It is not hard to see that the converse of the latter statement also holds so that the languages recognised by finite automata precisely are the languages recognised by finite monoids. This is then the starting point of the algebraic theory of automata. The next, and most crucial step, is the link to profinite completions.

In this section we describe the duality theoretic link between recognition and profinite completion. The technical result is Theorem 2 at the end of the section. This result is crucial for the applications in Section 4. For the reader who finds the proof given here too abstract, a more pedestrian proof may be found in [6] where this result in the case of monoids occurs as Theorem 3 and a different proof is given.

Let A be an algebra. A subset $L \subseteq A$ is said to be *recognisable* provided there is a finite algebra B such that L is recognised by B. We denote the Boolean algebra of all recognisable subsets of A by $\mathrm{Rec}(A)$. We have:

$$\begin{aligned}
\mathrm{Rec}(A) &= \{\varphi^{-1}(P) \mid \varphi : A \to B \supseteq P \text{ with } \varphi \text{ an onto morphism and } B \text{ finite}\} \\
&= \bigcup \{\varphi^{-1}(\mathcal{P}(B)) \mid \varphi : A \to B \text{ is an onto morphism and } B \text{ is finite}\}
\end{aligned}$$

By placing this definition in a more category-theoretic context we will be able to apply the dualities of Section 1. First note that the finite quotients of A are in one-to-one correspondence with the set $Con_\omega(A)$ of all congruences θ of A of finite index (i.e. for which the quotient algebra A/θ is finite). Also, $Con_\omega(A)$ is ordered by reverse set inclusion and is directed in this order since the intersection of two congruences of finite index is again a congruence of finite index. Thus we obtain an inverse limit system, $\mathcal{F}_A$, indexed by $Con_\omega(A)$ as follows:

1. For each $\theta \in Con_\omega(A)$ we have the finite algebra A/θ;

2. Whenever $\theta \subseteq \psi$ we have a (unique) homomorphism $A/\theta \to A/\psi$ which commutes with the quotient maps $q_\theta : A \to A/\theta$ and $q_\psi : A \to A/\psi$.

$$\begin{tikzcd}
A \arrow[drr, dashed, bend right=10] \arrow[ddddrr, dashed, bend right=15] \arrow[ddrrr, dashed, bend left=10] & & & \\
& & & \\
& {} \arrow[r] & B \arrow[dr] & \\
& & & D \\
& {} \arrow[r] & C \arrow[ur] &
\end{tikzcd}$$

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

The *profinite completion* of an algebra A, denoted $\widehat{A}$, is by definition the inverse limit, $\varprojlim \mathcal{F}_A$, of the system $\mathcal{F}_A$ viewed *as a system of topological algebras.* Note that A is not a topological algebra but the finite quotients are trivially so with the discrete topology.

Applying the discrete duality of Section 1.1, we get the direct limit system dual to $\mathcal{F}_A$:

$$\begin{tikzcd}
\mathcal{P}(A) & & & \\
& & & \\
& {} & \mathcal{P}(B) \arrow[l] \arrow[ull, dashed, bend left=10] & \\
& & & \mathcal{P}(D) \arrow[ul] \arrow[dl] \arrow[uuulll, dashed, bend right=10] \\
& {} & \mathcal{P}(C) \arrow[l] \arrow[uuuull, dashed, bend left=15] &
\end{tikzcd}$$

The direct limit system $\mathcal{G}_A$

In Section 2, we saw that for a regular language L, the dual of the residuation algebra $(\mathcal{B}(L), \backslash, /)$ is the syntactic monoid of L. One can actually show that the extended Stone dual of any finite algebra is the Boolean algebra with residuation operations obtained by taking the powerset of the algebra with the residuals of the liftings of the operations of the algebra (as illustrated for a binary operation in (1) and (2) in Section 2). Further, the quotient maps in the system $\mathcal{F}_A$ are dual to the inverse image maps which provide embeddings between the finite powerset algebras, and one can show that the fact that the maps in $\mathcal{F}_A$ are algebra homomorphisms corresponds to the fact that the maps in $\mathcal{G}_A$ embed the residuation algebras as quotienting ideals (in the case of a single binary operation, see Definition 4). All in all, we obtain a direct limit system $\mathcal{G}_A$ of finite residuation ideals of $\mathcal{P}(A)$. It is well-known in algebra that a direct limit of subalgebras of an algebra simply is the union of these subalgebras. Thus we get

$$\begin{aligned}
\varinjlim \mathcal{G}_A &= \bigcup \{\varphi^{-1}(\mathcal{P}(B)) \mid \varphi : A \to B \text{ is an onto morphism and } B \text{ is finite}\} \\
&= \mathrm{Rec}(A).
\end{aligned}$$

We have outlined the proof of the following theorem.

Theorem 2. *Let A be an abstract algebra. Then* Rec(A) *is residuation ideal in* $\mathcal{P}(A)$ *and the profinite completion, $\widehat{A}$, of the algebra A is homeomorphic as a topological algebra to the extended Stone dual of* Rec(A) *viewed as a Boolean algebra with residuation operations.*

4 Eilenberg-Reiterman: sub vs. quotient duality

In automata theory, deciding membership in a class of recognisable languages and separating two such classes are central problems. Classes of interest arise by restricting the class of automata allowed for recognition, or they arise via the description of languages by regular expressions by putting some restriction on the form of the expressions that are allowed. There is also, via Büchi's logic on words, a correspondence between recognisable languages and monadic second order sentences of this logic. Thus natural subclasses arise as the classes of languages corresponding to fragments of monadic second order logic.

The classical example is that of the star-free languages. These are the languages obtainable from the singleton languages by closure under the Boolean connectives and the lifted concatenation product (but *without* using the Kleene star). In terms of Büchi's logic, these are precisely the languages that are models of sentences of the first order fragment. While both of these descriptions of the star-free languages are nice, neither allows one readily to decide whether or not the language recognised by a given automaton is star-free or not. Schützenberger [12] made a breakthrough in the mid-sixties by using syntactic monoids to give an algebraic characterisation of the star-free languages which also provides a decision procedure for the class: a regular languages is star-free if and only if its syntactic monoid is aperiodic (for any element x in a finite monoid, there are m and n so that $x^{m+n} = x^m$; aperiodicity means the n uniformly can be taken to be equal to 1). This is clearly a decidable property of a finite monoid.

Eilenberg [3] greatly contributed to the success of the algebraic theory by isolating an essential feature of the above example: the finite aperiodic monoids form a *variety* of finite algebras. That is, a class of finite algebras closed under subalgebras, quotients, and finite Cartesian products. Further he showed that such varieties are in one-to-one correspondence with certain classes of regular languages, which he called *varieties of languages.* Later, Reiterman proved that profinite words (that is, elements of the profinite completion of A^* for A finite) can be viewed as $|A|$-ary term functions on any finite monoid, and that each variety of finite monoids is given by a set of identities in profinite words [11]. In conjunction we have: A class of regular languages is a variety of languages if and only if it can be defined by a set of profinite identities. If a variety of languages has a finite basis for its identities and the identities can be checked effectively, then it follows that the class is decidable. This has become a standard route to decidability, both for varieties of languages and for various generalisations for which Eilenberg-Reiterman theorems have subsequently been proved.

With Grigorieff and Pin, we gave a general and modular Eilenberg-Reiterman theorem based on duality theory [7]. The idea is the following. Let $\mathcal{C}$ be a sublattice of $\mathrm{Rec}(A^*)$, that is

$$\mathcal{C} \quad \hookrightarrow \quad \mathrm{Rec}(A^*).$$

Then, the Priestley dual $X_{\mathcal{C}}$ is a quotient space of the dual space of $\mathrm{Rec}(A^*)$, which we know to be $\widehat{A^*}$:

$$\widehat{A^*} \quad \twoheadrightarrow \quad X_{\mathcal{C}}.$$

That is, $\mathcal{C}$ is fully described by describing $X_{\mathcal{C}}$, and $X_{\mathcal{C}}$ is fully described by describing, in the case of a Boolean subalgebra of $\mathrm{Rec}(A^*)$, the equivalence relation on $\widehat{A^*}$ that yields the quotient $X_{\mathcal{C}}$. In the sublattice case, not only are some points identified going from $\widehat{A^*}$ to $X_{\mathcal{C}}$, but the order may also be strengthened. Thus sublattices correspond to certain quasiorders, called Priestley quasiorders, on $\widehat{A^*}$. This may be seen as the underlying source of profinite identities.

Fundamental to this relationship between sublattices and quotients is the following binary *satisfaction* relation between pairs $(u,v) \in \widehat{A^*} \times \widehat{A^*}$ and elements $L \in \mathrm{Rec}(A^*)$:

$$L \ \text{ satisfies } \ (u,v) \qquad \Longleftrightarrow \qquad \big(\eta_{\mathrm{Rec}(A^*)}(L) \in v \ \Rightarrow \ \eta_{\mathrm{Rec}(A^*)}(L) \in u\big).$$

A language L satisfies (u,v) provided L lies in a sublattice of $\mathrm{Rec}(A^*)$ corresponding to a Priestley quotient of $\widehat{A^*}$ in which u ends up being below v. For this reason we write $u \to v$ for these profinite inequations instead of just (u,v).

Theorem 3. *The assignments*

$$\Sigma \ \mapsto \ \mathcal{C}_\Sigma = \{L \in \mathrm{Rec}(A^*) \mid \forall\, u \to v \in \Sigma \ (L \quad \textit{satisfies} \quad u \to v)\}$$

for $\Sigma \subseteq \widehat{A^*} \times \widehat{A^*}$ *and*

$$\mathcal{K} \ \mapsto \ \Sigma_{\mathcal{K}} = \{u \to v \in \widehat{A^*} \times \widehat{A^*} \mid \forall\, L \in \mathcal{K} \ (L \quad \textit{satisfies} \quad u \to v)\}$$

for $\mathcal{K} \subseteq \mathrm{Rec}(A^*)$ *establish a Galois connection whose Galois closed sets are the Priestley quasiorders on* $\widehat{A^*}$ *and the bounded sublattices of* $\mathrm{Rec}(A^*)$*, respectively.*

Thus, for any sublattice $\mathcal{C}$ of $\mathrm{Rec}(A^*)$, we have $\mathcal{C}_{\Sigma_{\mathcal{C}}} = \mathcal{C}$ so that $\mathcal{C}$ is determined by a set of inequations. Also, we may look for bases $\Sigma \subseteq \Sigma_{\mathcal{C}}$ with $\mathcal{C}_\Sigma = \mathcal{C}$. In addition, if $\mathcal{C}$ is closed under the quotienting operations $a^{-1}(\ )$ and $(\ )a^{-1}$ on languages then we know that the corresponding Priestley quotient is also a monoid quotient, or equivalently, the corresponding Priestley quasiorder is a monoid congruence. In this case, we know that, with each $u \to v$ in $\Sigma_{\mathcal{C}}$, the inequation $xuy \to xvy$ is also in $\Sigma_{\mathcal{C}}$ and we can abbreviate this whole family of inequalities as $u \leqslant v$. Similarly, for Boolean sublattices $u \to v$ in $\Sigma_{\mathcal{C}}$ implies $v \to u$ in $\Sigma_{\mathcal{C}}$, and it can be shown that $\mathcal{C}$ being closed under inverse images for various kinds of homomorphisms of A^* corresponds to the set of inequations for $\mathcal{C}$ being closed with respect to various kinds of substitutions. The ensuing family of Eilenberg-Reiterman theorems thus obtained is summed up in the following table.

Closed under	**Equations**	**Definition**
$\cup, \cap$	$u \to v$	$\hat{\varphi}_L(v) \in P_L \Rightarrow \hat{\varphi}_L(u) \in P_L$
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

In order to understand the interpretation of profinite words in finite monoid quotients of A^*, it is important to realise that, by duality, any map $\varphi : A^* \to F$ has a unique extension $\widehat{\varphi} : \widehat{A^*} \to F$ obtained as the Stone dual of the Boolean algebra homomorphism $\varphi^{-1} : \mathcal{P}(F) \to \mathrm{Rec}(A^*)$.

Example 5. The class of star-free languages is axiomatised by $x^\omega = x^{\omega+1}$ with the interpretation of x ranging over all profinite words. The fact that the class is closed under the quotient operations and the Boolean operations means that L is star-free if and only if $\varphi_L^{-1}(P)$ is star-free for each $P \subseteq S(L)$, not just for P_L. Now, it can be shown that for $u \in \widehat{A^*}$ we have $\widehat{\varphi}_L(u^\omega) = e(\widehat{\varphi}_L(u))$ where $e : S(L) \to S(L)$ is the map that sends any element m of $S(L)$ to the unique idempotent in the cyclic monoid generated by m. Also, since each element of $S(L)$ is $\widehat{\varphi}_L(u)$ for some $u \in \widehat{A^*}$, we have L is star-free if and only if

$$\forall\ P \subseteq S(L)\ \forall m \in S(L) \qquad\qquad (\ e(m) \in P \quad \iff \quad e(m)\, m \in P\ ).$$

Since this has to hold in particular for singleton subsets P of $S(L)$, a language L is star-free if and only if $S(L)$ satisfies the identity $e(x) = e(x)x$. Here we get a genuine identity, that is, an equation scheme closed under substitution because the class of star-free languages is closed under inverse images of arbitrary morphisms between free finitely generated monoids. Finally we note that our language $L = (ab)^*$ is star-free since the elements 1, 0, $(ab)^+$, and $(ba)^+$ are idempotent and $e(a(ba)^*) = e(b(ab)^*) = 0$ is absorbent.

A *regular language with zero* is a regular language whose syntactic monoid has a 0. It is not hard to see that the class of regular languages with zero is closed under the quotient operations and the Boolean operations, but not under inverse images of arbitrary morphisms. Regular languages with 0 are given by the A-specific identities $x\rho_A = \rho_A = \rho_A x$ where ρ_A is an idempotent in the minimal (closed) ideal of $\widehat{A^*}$ and x can range over all elements of $\widehat{A^*}$. As in the case of star-freeness, the closure under the Boolean and the quotient operations allows us to quantify over all the subsets P of $S(L)$ and thus we must have

$$\forall m \in S(L) \qquad\qquad \big(\ m\, \widehat{\varphi}_L(\rho_A) = \widehat{\varphi}_L(\rho_A) = \widehat{\varphi}_L(\rho_A)\, m\ \big).$$

Since $\widehat{\varphi}_L(\rho_A)$ will necessarily belong to the minimum ideal of $S(L)$, it is easy to see that it will evaluate to 0 if and only if $S(L)$ has a zero so that these profinite equations precisely say that $S(L)$ has a zero.

For more details, see [9] where the various declensions of our theorem are illustrated with various examples.

References

1. S. Abramsky, Domain Theory in Logical Form, *Ann. Pure Appl. Logic* **51** (1991), 1–77.
2. B. A. Davey and H. A. Priestley, *Introduction to Lattices and Order, 2nd edition*, Cambridge University Press, 2002.
3. S. Eilenberg, *Automata, languages, and machines. Vol. B*, Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1976.
4. L. L. Esakia, Topological Kripke models, *Soviet Math. Dokl.* **15** (1974), 147–151.
5. R. Fagin, J. Y. Halpern, Y. Moses and M. Y. Vardi, *Reasoning About Knowledge*, MIT Press, 1995.
6. M. Gehrke, Stone duality and the recognisable languages over an algebra, in *CALCO 2009*, A. Kurz and al. (eds.), Berlin, 2009, pp. 236–250, *Lect. Notes Comp. Sci.* vol. 5728, Springer.
7. M. Gehrke, S. Grigorieff and J.-E. Pin, Duality and equational theory of regular languages, in *ICALP 2008, Part II*, L. Aceto and al. (eds.), Berlin, 2008, pp. 246–257, *Lect. Notes Comp. Sci.* vol. 5126, Springer.
8. R. Goldblatt, Varieties of complex algebras, *Ann. Pure App. Logic* **44** (1989), 173–242.
9. J.-E. Pin, Profinite methods in automata theory, in *26th International Symposium on Theoretical Aspects of Computer Science (STACS 2009)*, S. Albers and J.-Y. Marion (eds.), pp. 31–50, Internationales Begegnungs- und Forschungszentrum für Informatik (IBFI), Schloss Dagstuhl, Germany, 2009.
10. N. Pippenger, Regular languages and Stone duality, *Theory Comput. Syst.* **30**,2 (1997), 121–134.
11. J. Reiterman, The Birkhoff theorem for finite algebras, *Algebra Universalis* **14**,1 (1982), 1–10.
12. M. P. Schützenberger, On finite monoids having only trivial subgroups, *Inform. Control* **8** (1965), 190–194.
13. M. Stone, The theory of representations for Boolean algebras, *Trans. Amer. Math. Soc.* **40** (1936), 37–111.