

Equational theories of profinite structures

MICHAŁ SKRZYPCZAK*

Institute of Informatics
University of Warsaw

July 18, 2011

Abstract

In this paper we consider a general way of constructing profinite structures based on a given framework — a countable family of objects and a countable family of recognisers (e.g. formulas). The main theorem states:

A subset of a family of recognisable sets is a lattice if and only if it is definable by a family of profinite equations.

This result extends Theorem 5.2 from [GGEP08] expressed only for finite words and morphisms to finite monoids.

One of the applications of our theorem is the situation where objects are finite relational structures and recognisers are first order sentences. In that setting a simple characterisation of lattices of first order formulas arise.

1 Introduction

The following situation is very popular in computer science: the expressive power of a countable set of syntactical objects Φ is studied over a countable family of structures \mathbb{W} .

The following examples illustrate this situation:

1. $\mathbb{W} = A^*$ and $\Phi = \{A^* \rightarrow M : M \text{ is a finite monoid}\}$,
2. \mathbb{W} are finite trees and $\Phi = \{\varphi : \varphi \text{ is a first order sentence}\}$,
3. \mathbb{W} is a set of all finite trees and Φ is a family of all deterministic bottom-up tree automata,
4. \mathbb{W} are all finite graphs and Φ is the family of all FO or MSO formulas.

*Author supported by ERC Starting Grant "Sosna"

One of the natural ideas how to represent *recognition* is to treat a syntactical object $\varphi \in \Phi$ as a function $\varphi: \mathbb{W} \rightarrow K_\varphi$ to a finite set K_φ . Such a function recognises sets of objects of the form

$$\varphi^{-1}(V) \subseteq \mathbb{W},$$

for $V \subseteq K_\varphi$. Such sets are usually called regular or recognisable.

All the examples presented above fall into this schema: formula is a function to $\{\perp, \top\}$, homomorphism is a function to a finite monoid, deterministic automaton maps a tree into the state reached at the root.

In this paper we work with a very general setting of families of objects \mathbb{W} and of recognisers Φ . Using adequate topology on \mathbb{W} we show how to define a profinite structure $\widehat{\mathbb{W}}$ extending \mathbb{W} . Moreover we prove that it's possible to extend recognisers $\varphi \in \Phi$ to all profinite objects $w \in \widehat{\mathbb{W}}$.

The following theorem is the main result.

Theorem 1.1. *A family \mathcal{M} of recognisable sets of objects is a lattice if and only if it is defined by a set of profinite equations.*

In paper [GGEP08] authors show analogous theorem in the context of profinite monoids. It is a special case of our result where $\mathbb{W} = \Sigma^*$ and Φ are homomorphisms to finite monoids.

Paper [GGEP10] is devoted to the idea of recognisers — particular functions defined on a given topological (or more generally uniform) space. Authors show that each Boolean algebra of subsets of a space has a *minimal recogniser*. Additionally various additional structures of the space (e.g. the structure of a monoid) provides additional properties of a recogniser. The key tool is the Stone-Priestley duality.

The approach presented in this paper is different. We start with a fixed family of recognisers and using them we provide adequate topology on \mathbb{W} . Using this topology we extend the space to a profinite structure and study the particular algebra of recognisable sets.

2 Profinite structures

Definition 2.1. *A framework consists of*

1. *a countable set \mathbb{W} of objects,*
2. *a countable set Φ of recognisers, i.e. functions $\varphi: \mathbb{W} \rightarrow K_\varphi$ where K_φ is a finite set.*

Additionally two properties must hold:

- a) *Each object $w \in \mathbb{W}$ is totally described by some recogniser. That is for every object $w \in \mathbb{W}$ there is some recogniser $\varphi \in \Phi$ such that $\varphi(w) \neq \varphi(w')$ for $w' \neq w$.*

b) *Recognisers are closed under intersections. That is for every recognisers $\varphi_1, \varphi_2 \in \Phi$ and every sets of values $V_1 \in K_{\varphi_1}, V_2 \in K_{\varphi_2}$ there exists a recogniser $\varphi \in \Phi$ and a set $V \subseteq K_\varphi$ such that*

$$\forall w \in \mathbb{W} \quad (\varphi_1(w) \in V_1 \wedge \varphi_2(w) \in V_2) \Leftrightarrow \varphi(w) \in V.$$

It is easy to check that all the examples from the introduction satisfy both axioms so they form *frameworks*.

Fix a framework (\mathbb{W}, Φ) .

Definition 2.2. *A language $L \subseteq \mathbb{W}$ is called regular or recognisable if there exists $\varphi \in \Phi$ and $V \subseteq K_\varphi$ such that*

$$L = \varphi^{-1}(V).$$

The family of all regular languages is denoted as $\text{reg}(\mathbb{W})$.

Observation 2.3. *$\text{reg}(\mathbb{W})$ is a Boolean algebra.*

Proof. By Property b) regular languages are closed under intersection. Of course they are also closed under complementation. So they form a Boolean algebra. ■

List all recognisers in Φ in a sequence $\Phi = \{\varphi_0, \varphi_1, \dots\}$. It is good to think that recognisers appearing further in the sequence are *more complicated*. Let $K_i = K_{\varphi_i}$ and let $X = \prod_{i \in \mathbb{N}} K_i$. Because K_i 's are finite, X is a homeomorphic copy of Cantors discontinuum. Let $\mu: \mathbb{W} \rightarrow X$ be defined as follows

$$\mu(w) = (\varphi_0(w), \varphi_1(w), \varphi_2(w), \dots).$$

In other words μ maps an object $w \in \mathbb{W}$ to a sequence of values of all recognisers on that object.

It is easy to see that μ is 1-1 because of Property a).

Definition 2.4. *Let $\widehat{\mathbb{W}} = \overline{\mu(\mathbb{W})} \subseteq X$. The elements of the set $\widehat{\mathbb{W}}$ are called profinite objects of the framework (\mathbb{W}, Φ) . The topology on $\widehat{\mathbb{W}}$ is defined as a topology induced from X .*

Since the order of coordinates in Cantors discontinuum does not affect its topology, we obtain the following fact.

Fact 2.5. *The construction described above does not depend on the order of recognisers in the sequence $\varphi_0, \varphi_1, \dots$*

The following proposition summarises the properties of $\widehat{\mathbb{W}}$.

Proposition 2.6.

1. \mathbb{W} naturally embeds into $\widehat{\mathbb{W}}$,
2. $\widehat{\mathbb{W}}$ is a compact topological space,

3. for each $w \in \mathbb{W}$ the point $\mu(w) \in \widehat{\mathbb{W}}$ is isolated,
4. $\mu(\mathbb{W})$ is a countable dense subset of $\widehat{\mathbb{W}}$.

Proof.

1. The natural embedding is μ .
2. $\widehat{\mathbb{W}}$ is a closed subset of a compact space, so it is compact.
3. This is an easy consequence of property a) of the framework.
4. $\widehat{\mathbb{W}}$ is a closure of $\mu(\mathbb{W})$, so $\mu(\mathbb{W})$ is dense in $\widehat{\mathbb{W}}$.

■

Note that recognisers naturally extend to $\widehat{\mathbb{W}}$.

Proposition 2.7. *For each recogniser $\varphi_i \in \Phi$ we can extend it to all profinite objects $w \in \widehat{\mathbb{W}} \subseteq \prod K_i$ by an equation*

$$\varphi_i(w) = w_i \in K_i.$$

For $w \in \mathbb{W}$ this definition is consistent with the original one.

The following lemma enables us to define regular languages from the topological point of view.

Lemma 2.8. *A language $L \subseteq \mathbb{W}$ is regular if and only if \overline{L} is an clopen subset of $\widehat{\mathbb{W}}$.*

Moreover $L \mapsto \overline{L}$ is an isomorphism of the Boolean algebra of regular languages in \mathbb{W} and the Boolean algebra of clopen subsets of $\widehat{\mathbb{W}}$.

Proof. Regular languages form a clopen base of the topology and they are closed under finite Boolean operations. In a compact space each clopen subset is a Boolean combination of the clopen base sets. ■

3 Topology

In this section we provide a general characterisation of sublattices and Boolean subalgebras of the algebra of clopen subsets of a compact space. This whole theory is based only on topological properties and holds for any compact space.

Fix X to be a compact space and let \mathcal{B} denote the Boolean algebra of clopen subsets of X .

Definition 3.1. *A sublattice of \mathcal{B} is any subset $\mathcal{M} \subseteq \mathcal{B}$ closed under union and intersection and containing \emptyset and X .*

Definition 3.2. An equation is a formula $u \rightarrow v$ for $u, v \in X$. We say that a subset $A \subseteq X$ satisfies an equation $u \rightarrow v$ if the following property holds

$$u \in A \Rightarrow v \in A.$$

For a given set of equations $\mathcal{E} \subseteq X^2$, the family of subsets that satisfy all equations in \mathcal{E} will be called a family of subsets defined by \mathcal{E} .

Theorem 3.3. For a given compact space X a family $\mathcal{M} \subseteq \mathcal{B}(X)$ is a lattice if and only if it is defined by some set of equations.

Proof. Of course for a given set of equations, a family of clopen subsets satisfying them is a lattice.

Take any lattice $\mathcal{M} \subseteq \mathcal{B}$. Let \mathcal{E} be a set of all equations satisfied by every set in \mathcal{M} . Take any subset $A \in \mathcal{B}$ satisfying \mathcal{E} . We will show that $A \in \mathcal{M}$.

Let

$$\mathcal{U} = \{M \in \mathcal{M} : M \subseteq A\}.$$

If $\bigcup \mathcal{U} = A$, then \mathcal{U} is a covering by open sets of a compact A , so there is a finite family $M_1, M_2, \dots, M_n \in \mathcal{U}$ such that $\bigcup M_i = A$. So $A \in \mathcal{M}$. Assume by contradiction that there exists $x \in A \setminus \bigcup \mathcal{U}$.

Consider

$$\mathcal{V} = \{M \in \mathcal{M} : x \in M\}.$$

If $\bigcap \mathcal{V} \subseteq A$, then complements of elements in \mathcal{V} cover $X \setminus A$ so there is a finite family $M_1, M_2, \dots, M_n \in \mathcal{V}$ such that $\bigcap M_i \subseteq A$. But then $x \in \bigcap M_i \subseteq A$ and $\bigcap M_i \in \mathcal{M}$, so a contradiction to the fact that $x \notin \bigcup \mathcal{U}$. So there exists $y \in \bigcap \mathcal{V} \setminus A$.

Consider equation $x \rightarrow y$ and any $N \in \mathcal{M}$ such that $x \in N$. Then $N \in \mathcal{V}$ so by the definition $y \in N$ so N satisfies $x \rightarrow y$. Therefore $(x \rightarrow y) \in \mathcal{E}$. But A does not satisfy $x \rightarrow y$. A contradiction. ■

Corollary 3.4. A family $\mathcal{M} \subseteq \mathcal{B}$ is a Boolean subalgebra if and only if it is defined by a symmetric set of equations $u \leftrightarrow v \equiv u \rightarrow v, v \rightarrow u$.

4 Main theorem

Using the topological result from the previous section we can prove the main theorem from the introduction.

Theorem 4.1. A family \mathcal{M} of recognisable sets of objects is a lattice if and only if it is defined by a set of profinite equations.

Firstly we adopt the definition of equation to the case of regular languages. This definition follows the one from [GGEP08].

Definition 4.2. We say that a regular language $L \subseteq \mathbb{W}$ satisfies an equation $u \rightarrow v$ for $u, v \in \widehat{\mathbb{W}}$ if

$$u \in \bar{L} \Rightarrow v \in \bar{L}.$$

In other words language L recognised by a recogniser φ and a set of values $K \subseteq K_\varphi$ satisfies $u \rightarrow v$ iff

$$\varphi(u) \in K \Rightarrow \varphi(v) \in K.$$

Proof of Theorem 1.1. Take any subfamily of regular languages $\mathcal{M} \subseteq \text{reg}(\mathbb{W})$ and consider

$$\overline{\mathcal{M}} = \{\bar{L} \subseteq \widehat{\mathbb{W}} : L \in \mathcal{M}\}.$$

By Lemma 2.8 $\overline{\mathcal{M}} \subseteq \mathcal{B}(\widehat{\mathbb{W}})$ and \mathcal{M} is a lattice if and only if $\overline{\mathcal{M}}$ is a lattice. Additionally \mathcal{M} is defined by a set of equations in the meaning of Definition 4.2 if and only if $\overline{\mathcal{M}}$ is defined by a set of equations in the meaning of Definition 3.2.

But $\overline{\mathcal{M}}$ is defined by a set of equations of and only if it is a lattice, because of Theorem 3.3 and the fact that $\widehat{\mathbb{W}}$ is a compact topological space. ■

Corollary 4.3. A family \mathcal{M} of recognisable languages of objects is a Boolean algebra if and only if it is defined by a set of profinite symmetric equations.

5 Conclusions

In this section we propose various applications of the main theorem.

Firstly consider a case when $\mathbb{W} = \Sigma^*$ and Φ consists of all homomorphisms of Σ^* to finite monoids. In that case $\widehat{\mathbb{W}}$ is just a space of profinite words $\widehat{\Sigma}^*$ and Theorem 1.1 coincides with Theorem 5.2 from [GGEP08].

It turns out that Theorem 1.1 gives some insight into the structure of lattices of first order formulas. Fix a relational signature Σ and consider \mathbb{W} as a family of all finite structures over Σ . Let Φ be a set of all FO sentences over Σ . Of course \mathbb{W}, Φ is a framework. We denote this framework as *first order framework* over signature Σ .

Lemma 5.1. If \mathbb{W}, Φ is a first order framework then for every profinite object $w \in \widehat{\mathbb{W}}$ there exists a structure $S(w)$ over Σ such that for all FO sentences $\varphi \in \Phi$

$$\varphi(w) = \top \Leftrightarrow S(w) \models \varphi.$$

Proof. An easy application of the compactness theorem. ■

Of course there are infinite structures that don't represent any element of $\widehat{\mathbb{W}}$. For example take $\Sigma = \{\leq\}$ and let φ express that \leq is a linear order and that $\forall x \exists y x < y$. Then (ω, \leq) is a model for φ but no finite structure satisfies it. So (ω, \leq) does not represent any element of $\widehat{\mathbb{W}}$.

Theorem 5.2. *Every lattice of first order formulas $\mathcal{M} \subseteq FO(\Sigma)$ is defined by some family of implications $(u \rightarrow v)$, where u, v are (potentially infinite) Σ -structures.*

We say that a formula φ satisfies an implication $(u \rightarrow v)$ iff

$$(u \models \varphi) \Rightarrow (v \models \varphi).$$

Proof. By 1.1 a lattice of formulas is defined by some family of profinite equations. But using 5.1 each profinite object $w \in \widehat{\mathbb{W}}$ can be interpreted as a structure $S(w)$. Therefore, we can put structures instead of profinite objects in all equations. ■

5.1 Acknowledgements

The author would like to thank Mikołaj Bojańczyk, Damian Niwiński and Henryk Michalewski for their helpful comments.

References

- [GGEP08] Mai Gehrke, Serge Grigorieff, and Jean Éric Pin. Duality and equational theory of regular languages. In *ICALP*, pages 246–257, 2008.
- [GGEP10] Mai Gehrke, Serge Grigorieff, and Jean Éric Pin. A topological approach to recognition. In *Automata, Languages and Programming*, volume 6199 of *LNCS*, pages 151–162. Springer Berlin / Heidelberg, 2010.