Recognisable sets, profinite topologies and weak arithmetic

Jean-Éric Pin

IRIF, CNRS and University Paris 7

May 2018, Firenze

Partially funded by the ERC (grant agreement No 670624) and by the DeLTA project (ANR-16-CE40-0007)

Outline

The starting point

If A is a one-letter alphabet, the free monoid A^* is isomorphic to the additive monoid $\mathbb{N}$.

It seems natural to extend results on $\mathbb{N}$ to A^*. However, one may expect any result on A^* to become trivial on a one-letter alphabet.

Surprisingly enough, this is not always the case...

An example

Given a language $L \subseteq A^*$ and a word $u \in A^*$, let

$$u^{-1}L = \{x \in A^* \mid ux \in L\}$$
$$Lu^{-1} = \{x \in A^* \mid xu \in L\}$$

Theorem (Almeida, Esik, Pin 2017)

A class of regular languages closed under finite intersection, finite union, quotients and inverse of length-decreasing morphisms is also closed under inverse of morphisms.

For one-letter alphabets

For $L \subseteq \mathbb{N}$ and $k > 0$, let

$$L - 1 = \{n \in \mathbb{N} \mid n + 1 \in L\}$$
$$L \div k = \{n \in \mathbb{N} \mid kn \in L\}$$

Corollary (Cegielski, Grigorieff, Guessarian 2014)

Let $\mathcal{L}$ be a lattice of regular subsets of $\mathbb{N}$ such that if $L \in \mathcal{L}$, then $L - 1 \in \mathcal{L}$. Then for each positive integer k, $L \in \mathcal{L}$ implies $L \div k \in \mathcal{L}$.

Zoltán Ésik's original statement (January 07, 2010)

Corollary

The mapping

$$\mathcal{V} \mapsto \{X_L : L \in \mathcal{V}(a^*)\}$$

is an isomorphism from the lattice of commutative positive (ld-)varieties to the sublattice of $\mathcal{P}(\mathcal{P}(\mathbb{N}))$ consisting of those sets $\mathcal{X}$ of finite or ultimately periodic subsets of $\mathbb{N}$ that contain $\emptyset$ and $\mathbb{N}$ which are closed under union and intersection, moreover, the decrement operation defined by

$$X \to X - 1 = \{n - 1 \mid n \in X,\ n > 0\}.$$

Any such set $\mathcal{X}$ is closed under the division operations defined by:

$$X \to X/d = \{n \mid nd \in X\}, \quad d > 1.$$

When restricted to commutative (ld-)varieties, the same mapping creates an order isomorphism from the lattice of commutative (ld-)varieties to the sublattice of $\mathcal{P}(\mathcal{P}(\mathbb{N}))$ consisting of all those sets $\mathcal{X}$ of finite or ultimately periodic subsets of $\mathbb{N}$ which are additionally closed under complementation.

Original motivation

A function $f : A^* \to B^*$ is regularity-preserving if, for each regular language L of B^*, $f^{-1}(L)$ is also regular.

More generally, let $\mathcal{C}$ be a class of regular languages. A function $f : A^* \to B^*$ is $\mathcal{C}$-preserving if, for each $L \in \mathcal{C}$, $f^{-1}(L)$ is also in $\mathcal{C}$.

Goal. Find a complete description of regularity-preserving [$\mathcal{C}$-preserving] functions.

Same questions for transductions, that is, relations from A^* to B^*.

Part I

Recognisable sets

Monoids

A monoid is a set M equipped with an associative binary operation (the product) and an identity 1 for this operation.

A monoid M is finitely generated if there exists a finite subset F of M which generates M.

Examples. The free monoid A^*, with A finite.

Given a monoid M, the set $\mathcal{P}(M)$ of subsets of M is a monoid under the product defined, for $X, Y \subseteq M$, by $XY = \{xy \mid x \in X, y \in Y\}$.

Recognisable subsets of a monoid

A subset P of a monoid M is recognizable if there exists a finite monoid F, a monoid morphism $\varphi : M \to F$ and a subset Q of F such that $P = \varphi^{-1}(Q)$.

$\mathrm{Rec}(M)$ $=$ set of recognizable subsets of M .

Theorem (Kleene)

If $M = A^*$, then recognizable $=$ rational $=$ regular (that is, recognised by a finite automaton).

Recognisable subsets of $\mathbb{N}$

An arithmetic progression is a subset of $\mathbb{N}$ of the form $a + r\mathbb{N}$, with $r \geqslant 0$.

A subset of $\mathbb{N}$ is recognizable iff it is a finite union of arithmetic progressions.

$\{1, 3, 4, 7, 8, 9, 11, 12, 13, 17, 18, 22, 23, 27, 28, \ldots\} = \{1, 3, 4, 9, 11\} \cup \{7 + 5n \mid n \geqslant 0\} \cup \{8 + 5n \mid n \geqslant 0\}$ is a finite union of arithmetic progressions.

Recognisable subsets of a product of monoids

Theorem (Mezei)

Let $M = M_1 \times \cdots \times M_n$ be a product of monoids. A subset of M is recognisable iff it is a finite union of subsets of the form $R_1 \times \cdots \times R_n$, where each R_i is a recognisable subset of M_i.

Exercise: find the recognisable subsets of $\mathbb{N}^k$.

Transductions

Given two monoids M and N, a transduction from M into N is a relation on M and N.

If $\tau : M \to N$ is a transduction, then the inverse relation $\tau^{-1} : N \to M$ is also a transduction. If $R \subseteq N$, then

$$\tau^{-1}(R) = \{x \in M \mid \tau(x) \cap R \neq \emptyset\}$$

A function $f : M \to N$ is recognizability-preserving if, for each $R \in \mathrm{Rec}(N)$, $f^{-1}(R) \in \mathrm{Rec}(M)$.

Similarly, $\tau : M \to N$ is recognizability-preserving if, for each $R \in \mathrm{Rec}(N)$, $\tau^{-1}(R) \in \mathrm{Rec}(M)$.

Part II

Topological characterizations

Residually finite monoids

Let M be a monoid. A monoid F separates two elements $x, y \in M$ if there exists a morphism $\varphi : M \to F$ such that $\varphi(x) \neq \varphi(y)$.

A monoid is residually finite if any pair of distinct elements of M can be separated by a finite monoid.

Finite monoids, free monoids, free groups are residually finite. A product of residually finite monoids is residually finite.

Profinite metric

Let M be a residually finite monoid. The profinite metric d is defined by setting, for $u, v \in M$:

$$r(u,v) = \min\{|F| \mid F \text{ is a monoid separating } u \text{ and } v\}$$
$$d(u,v) = 2^{-r(u,v)}$$

with $\min \emptyset = +\infty$ and $2^{-\infty} = 0$. Then

$$d(u,w) \leqslant \max(d(u,v), d(v,w)) \quad \text{(ultrametric)}$$
$$d(uw,vw) \leqslant d(u,v)$$
$$d(wu,wv) \leqslant d(u,v)$$

Recognizability-preserving functions

Let M and N be two finitely generated, residually finite monoids. (For instance $M = A^*$ and $N = B^*$).

Theorem (Pin-Silva 2005)

A function $M \to N$ is recognizability-preserving iff it is uniformly continuous.

Another result

Proposition (Pin-Silva 2005)

The function $\tau : M \times \mathbb{N} \to M$ *defined by* $\tau(x, n) = x^n$ *is recognizability-preserving.*

Corollary. The function $u \to u^{|u|}$ (from A^* to A^*) is recognizability-preserving. Indeed it can be decomposed as

$$\begin{aligned} A^* &\to A^* \times \mathbb{N} \\ u &\to (u, |u|) \end{aligned} \qquad \begin{aligned} A^* \times \mathbb{N} &\to A^* \\ (u, n) &\to u^n \end{aligned}$$

Some examples of regularity preserving functions

$$u \to u^2 \qquad u \to \tilde{u}u$$

$$u \to u^{|u|} \qquad u \to a^{|u|_a} b^{|u|_b}$$

$$a^m c b^n \to a^n b^{mn}$$

$$u_0 \# u_1 \# u_2 \to u_2 \# u_1 \# u_0 \# u_1 \# u_2$$

Part III

Functions from $\mathbb{N}$ to $\mathbb{N}$

Ultimately periodic functions

A function $f : \mathbb{N} \to \mathbb{N}$ is ultimately periodic if there exists $t \geqslant 0$ and $p > 0$ such that, for all $n \geqslant t$, $f(n+p) = f(n)$.

A function $f : \mathbb{N} \to \mathbb{N}$ is ultimately periodic modulo n if the function $f \bmod n$ is ultimately periodic.

A function $f : \mathbb{N} \to \mathbb{N}$ is cyclically ultimately periodic if it is ultimately periodic modulo n for all $n > 0$.

Regularity-preserving functions from $\mathbb{N}$ to $\mathbb{N}$

Theorem (Siefkes 1970, SeiferasMcNaughton 1976)

A function $f : \mathbb{N} \to \mathbb{N}$ is ultimately periodic modulo n iff for $0 \leqslant k < n$, the set $f^{-1}(k + n\mathbb{N})$ is regular.

Theorem (Siefkes 1970, SeiferasMcNaughton 1976)

A function $f : \mathbb{N} \to \mathbb{N}$ is regularity-preserving iff it is cyclically ultimately periodic and, for every $k \in \mathbb{N}$, the set $f^{-1}(k)$ is regular.

Ultimately periodic functions

A function $f : \mathbb{N} \to \mathbb{N}$ is ultimately periodic modulo k if the function $f \bmod k$ is ultimately periodic.

It is cyclically ultimately periodic (cup) if it is ultimately periodic modulo n for all $n > 0$.

Proposition (Siefkes 70, SeiferasMcNaughton 76)

A function $f : \mathbb{N} \to \mathbb{N}$ is ultimately periodic modulo n iff for $0 \leqslant k < n$, the set $f^{-1}(k + n\mathbb{N})$ is regular. It is regularity-preserving iff it is cyclically ultimately periodic and $f^{-1}(k)$ is regular for every $k \in \mathbb{N}$.

Two examples

Theorem (Siefkes 1970)

The functions $n \to 2^n$ *and* $n \to 2^{2^{2^{\cdot^{\cdot^{\cdot^{2}}}}}}$ *(exponential stack of* 2*'s of height* n*) are cyclically ultimately periodic and hence regularity-preserving.*

Closure properties

Theorem (Siefkes 70, Zhang 98, Carton-Thomas 02)

Let $f, g : \mathbb{N} \to \mathbb{N}$ be cyclically ultimately periodic functions. Then so are the following functions:

(1) $g \circ f$, $f + g$, fg, f^g, *and* $f - g$ *provided that* $f \geqslant g$ *and* $\lim_{n\to\infty} (f - g)(n) = +\infty$,

(2) *(generalised sum)* $n \to \sum_{0 \leqslant i \leqslant g(n)} f(i)$,

(3) *(generalised product)* $n \to \prod_{0 \leqslant i \leqslant g(n)} f(i)$.

Two counterexamples

[Siefkes 1970] The function $n \to \lfloor \sqrt{n} \rfloor$ **is not** cyclically ultimately periodic and hence **not** regularity-preserving.

The function $n \to \binom{2n}{n}$ **is not** ultimately periodic modulo 4 and hence **not** regularity-preserving. Indeed

$$\binom{2n}{n} \bmod 4 = \begin{cases} 2 & \text{if } n \text{ is a power of } 2, \\ 0 & \text{otherwise.} \end{cases}$$

Recursivity

Let $f : \mathbb{N} \to \{0, 1\}$ be a non-recursive function. Then the function $n \to (\sum_{0 \leqslant i \leqslant n} f(i))!$ is regularity-preserving but non-recursive.

Open problem. Is it possible to describe all recursive regularity-preserving functions, respectively all recursive cyclically ultimately periodic functions?

One could try to use Siefkes' recursion scheme (1970).

Siefkes' recursion scheme

Theorem

Let $g : \mathbb{N}^k \to \mathbb{N}$ and $h : \mathbb{N}^{k+2} \to \mathbb{N}$ be cyclically ultimately periodic functions satisfying three technical conditions. Then the function f defined from g and h by primitive recursion, i.e.

$$\begin{aligned} f(0, x_1, \ldots, x_k) &= g(x_1, \ldots, x_k), \\ f(n+1, x_1, \ldots, x_k) &= h(n, x_1, \ldots, x_k, f(n, x_1, \ldots, x_k)) \end{aligned}$$

is cyclically ultimately periodic.

The three technical conditions

(1) h is cyclically ultimately periodic in x_{k+2} of decreasing period,

(2) g is essentially increasing in x_k,

(3) for all $x \in \mathbb{N}^{k+2}$, $x_{k+2} < h(x_1, \ldots, x_{k+2})$.

A function f is essentially increasing in x_j iff, for all $z \in \mathbb{N}$, there exists $y \in \mathbb{N}$ such that for all $x \in \mathbb{N}^n$, $y \leqslant x_j$ implies $z \leqslant f(x_1, \ldots, x_n)$.

A function f is c.u.p. of decreasing period in x_j iff, for all p, the period of the function $f \bmod p$ in x_j is $\leqslant p$.

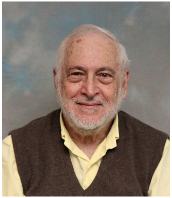
Part IV

An extension

Lattice of subsets

Let X be a set. A lattice of subsets of X is a set $\mathcal{L}$ of subsets of X containing $\emptyset$ and X and closed under finite union and finite intersection.

A Boolean algebra of subsets of X is a lattice of subsets of X closed under complement.



A Pervin space is a pair $(X, \mathcal{L})$ where $\mathcal{L}$ is a lattice of subsets of X.

Lattice-preserving functions

Let $f : X \to Y$ be a map, $\mathcal{K}$ be a lattice of subsets of X and $\mathcal{L}$ a lattice of subsets of Y.

Theorem

The following conditions are equivalent:

(1) *for each* $L \in \mathcal{L}$, $f^{-1}(L) \in \mathcal{K}$,

(2) f *is a uniformly continuous map from* $(X, \mathcal{K})$ *to* $(Y, \mathcal{L})$.

Lattice-preserving functions

Let $f : X \to Y$ be a map, $\mathcal{K}$ be a lattice of subsets of X and $\mathcal{L}$ a lattice of subsets of Y.

Theorem

The following conditions are equivalent:

(1) *for each* $L \in \mathcal{L}$, $f^{-1}(L) \in \mathcal{K}$,

(2) f *is a uniformly continuous map from* $(X, \mathcal{K})$ *to* $(Y, \mathcal{L})$.

Wait a second, what does uniformly continuous mean in this setting?

Uniform spaces

A uniformity on a set X is a nonempty set $\mathcal{U}$ of reflexive relations (entourages) on X such that:

(1) if a relation U on X contains an element of $\mathcal{U}$, then $U \in \mathcal{U}$, (extension property),

(2) the intersection of any two elements of $\mathcal{U}$ is in $\mathcal{U}$, (intersection),

(3) for each $U \in \mathcal{U}$, there exists $V \in \mathcal{U}$ such that $VV \subseteq U$ (sort of transitivity).

(4) for each $U \in \mathcal{U}$, ${}^tU \in \mathcal{U}$ (symmetry).

Quasi-uniform spaces

A quasi-uniformity on a set X is a nonempty set $\mathcal{U}$ of reflexive relations (entourages) on X such that:

(1) if a relation U on X contains an element of $\mathcal{U}$, then $U \in \mathcal{U}$ (extension property),

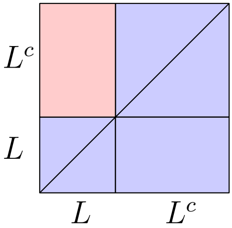
(2) the intersection of any two elements of $\mathcal{U}$ is in $\mathcal{U}$ (intersection),

(3) for each $U \in \mathcal{U}$, there exists $V \in \mathcal{U}$ such that $VV \subseteq U$ (sort of transitivity).

Pervin spaces as quasi-uniform spaces

Let $(X, \mathcal{L})$ be a Pervin space. For each $L \in \mathcal{L}$, let

$$\begin{aligned} U_L &= (X \times L) \cup (L^c \times X) \\ &= \{(x, y) \in X \times X \mid x \in L \Rightarrow y \in L\} \end{aligned}$$



The sets U_L form the subbasis of a quasi-uniformity.

Uniform continuity

Let X and Y be quasi-uniform spaces. A function $f : X \to Y$ is uniformly continuous if, for each entourage V of Y, $(f \times f)^{-1}(V)$ is an entourage of X.

Proposition

Let $(X, \mathcal{K})$ and $(Y, \mathcal{L})$ be two Pervin spaces. A function $f : X \to Y$ is uniformly continuous iff for each $L \in \mathcal{L}$, $f^{-1}(L) \in \mathcal{K}$.

Generalized ultrametric

A generalized ultrametric on a set X is a mapping $d : X \times X \to \mathbb{R}^+$ satisfying the following conditions:

(1) for all $x \in X$, $d(x, x) = 0$.

(2) for all $x, y, z \in X$,
$d(x, z) \leqslant \max(d(x, y), d(y, z))$.

Let $(X, \mathcal{L})$ be a Pervin space. Are equivalent:

(1) The associated quasi-uniformity can be defined by a generalized ultrametric,

(2) The quasi-uniformity has a countable basis,

(3) The lattice $\mathcal{L}$ is countable.

Boolean algebras

If $\mathcal{L}$ is a Boolean algebras, then one has a uniformity. Moreover if $\mathcal{L}$ is countable, this uniformity can be defined by an ultrametric.

If $\mathcal{L}$ is the set of recognizable subsets of a residually finite monoid M, then this ultrametric is the profinite ultrametric.

Part V

Transductions

Recognizability-preserving transductions

Let M and N be two finitely generated, residually finite monoids.

Theorem

A function $M \to N$ is recognizability-preserving iff it is uniformly continuous.

What about transductions from M to N?

Completion

Let M be a finitely generated, residually finite monoid. Let $\widehat{M}$ be the completion of the metric space (M, d).

Proposition

$\widehat{M}$ *is a compact monoid.*

Moreover, the set $\mathcal{K}(\widehat{M})$ of compact subsets of $\widehat{M}$ is also a compact monoid for the Hausdorff metric.

Back to transductions

Let M and N be two finitely generated, residually finite monoids and let $\tau : M \to N$ be a transduction.

Define a map $\widehat{\tau} : M \to \mathcal{K}(\widehat{N})$ by setting, for each $x \in M$, $\widehat{\tau}(x) = \overline{\tau(x)}$.

Theorem (Pin-Silva 2005)

The transduction τ is recognizability-preserving iff $\widehat{\tau}$ is uniformly continuous.

Exercises

Let L be a subset of A^*. Let

$$\frac{1}{2n+1}L = \{u \in A^* \mid \text{there exist } x, y \in A^*, \\ |x| = |y| = n \text{ and } xuy \in L\}$$

If L is regular, then so is the language

$$\bigcup_{p \text{ odd prime}} \frac{1}{p}L$$

The transduction $u \to u^*$ is regularity-preserving.

Part VI

p-group languages

Target class $\mathcal{G}_p$: the class of languages recognized by a finite p-group.

Goal. Characterization of $\mathcal{G}_p$-preserving functions.

Fonctions from $\mathbb{N}$ to $\mathbb{Z}$

The difference operator Δ associates to each function $f : \mathbb{N} \to \mathbb{Z}$, the function $\Delta f : \mathbb{N} \to \mathbb{Z}$ defined by $(\Delta f)(n) = f(n+1) - f(n)$.

A Newton polynomial is a function f such that $\Delta^k f = 0$ for almost all k.

Mahler's theorem

Let $\delta^k f = (\Delta^k f)(0)$.

Theorem (Mahler 58)

Let $f : \mathbb{N} \to \mathbb{Z}$ be a function. Are equivalent:

(1) *f is uniformly continuous for the p-adic metric,*

(2) *the functions $\Delta^n f$ tend uniformly to 0,*

(3) *the p-adic norm of $\delta^n f$ tends to 0,*

(4) *f is the uniform limit of a sequence of Newton polynomials.*