Automata, semigroups and groups: 60 years of synergy

Jean-Éric Pin[1]

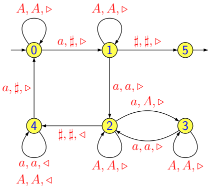
[1]LIAFA, CNRS and University Paris Diderot

60th birthday of Stuart W. Margolis

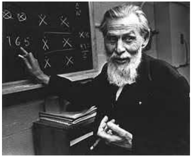
June 2013, Bar Ilan

Don't forget to turn your mobile phone back on AFTER this lecture

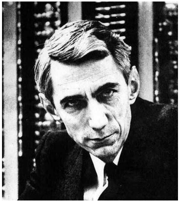
The precursors

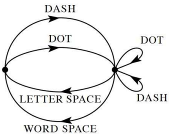


Turing (1936)
Turing Machine



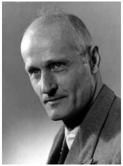
McCulloch and Pitts (1943)
Neural networks



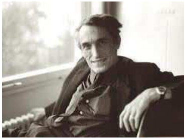


Shannon (1948)

The founders



Kleene (1951, 1956)
Equivalence between automata and regular expressions.



Schützenberger (1956):
Ordered syntactic monoid
Codes, unambiguous expressions.

Chomsky (1956):
Chomsky hierarchy

1956: Schützenberger's paper

M. P. SCHÜTZENBERGER

Une théorie algébrique du codage

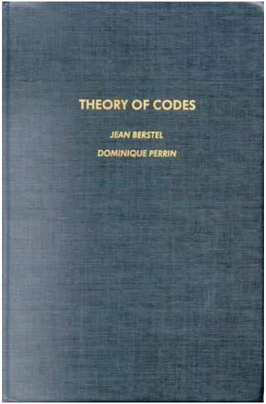
Séminaire Dubreil. Algèbre et théorie des nombres, tome 9 (1955-1956), exp. n° 15, p. 1-24.

<http://www.numdam.org/item?id=SD_1955-1956__9__A10_0>

- Birth of the theory of variable length codes
- Free submonoids of the free monoid
- Link with probabilities
- Definition of the syntactic preorder

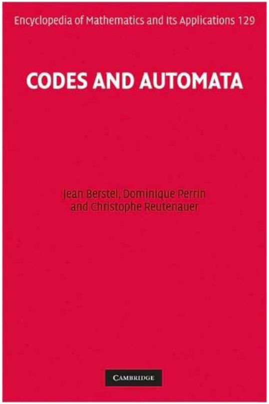
Theory of codes: lead to the following books

1985



Berstel, Perrin

2011



Berstel, Perrin, Reutenauer

Definition of the syntactic preorder

2.- Equivalence syntaxiques.

Soient A un demi-groupe contenant un élément neutre, K une partie quelconque de A .

Définition. On dira que a est syntaxiquement plus fort que b dans A , par rapport à K (a $\geqslant$ b (A , K)) si pour tout x , y $\in$ A :

(II) xby $\in$ K entraine xay $\in$ K

Si a $\geqslant$ b(A , K) et b $\geqslant$ (A , K) , a et b seront "syntaxiquement équivalents" (a $\equiv$ b(A , K).

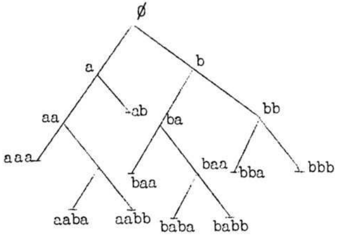
The syntactic preorder of a language K of A^* is the relation $\leqslant_K$ defined on A^* by $u \leqslant_K v$ iff, for every $x, y \in A^*$, $xuy \in K \Rightarrow xvy \in K$.

The syntactic congruence $\sim_K$ is the associated equivalence relation: $u \sim_K v$ iff $u \leqslant_K v$ and $v \leqslant_K u$.

Codes and syntactic monoids

Exemple.

Le cas le plus simple (K = 2 ; ℓ = 3) est décrit par l'arbre suivant. Il contient 9 mots et son GSF (privé de l'élément neutre bilatère) a 24 éléments et ne possède pas d'idéaux propres. Cette dernière particularité n'est pas une nécessité pour les GSF des codes de ce type.



(u = ab)

* a^3	a	a^2ba	aba	a^2	aba^2	* $abab$	a^2b	a^2bab	ab^2	aba^2b	ab
* $baba$	ba^2	ba^3	ba	ba^2ba	$baba^2$	* b^3	bab	ba^2b	b	b^2	bab^2

Early results

- Automata theory: Medvedev (1956), Myhill (1957), Nerode (1958), Rabin and Scott (1958), Brzozowski (1964).
- Logic and automata: Trahtenbrot (1958), Büchi (1960), McNaughton (1960), Elgot (1961).
- Semigroup theory: Clifford and Preston (1961) [Vol 2 in 1967]

Krohn-Rhodes theorem (1962-1965)

Every automaton divides a cascade product of permutation automata and flip-flops.

Every semigroup divides a wreath product of simple groups and copies of U_2. Every aperiodic semigroup divides a wreath product of copies of U_2.

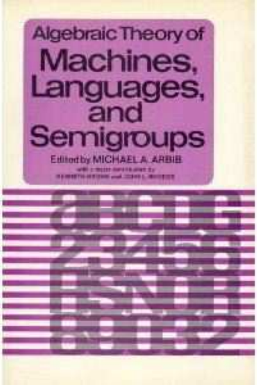
U_2: $aa = ab = b$, $ba = bb = b$

^{*1}

$^*a$	$^*b$

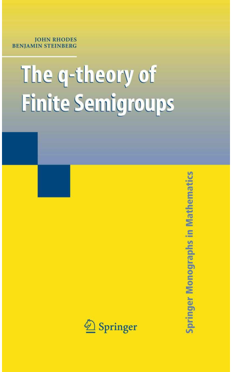
... lead to the following books

1968



Arbib (ed.)

2011



Rhodes and Steinberg

Schützenberger's theorem on star-free languages

[1] Sur les monoïdes finis n'ayant que des sous-groupes triviaux. In Séminaire Dubreil-Pisot, année 1964-65, Exposé 10, 6 pages. Inst. H. Poincaré, Paris, 1965.

[2] On finite monoids having only trivial subgroups. *Information and Control* **8** 190–194, 1965.

[3] Sur certaines variétés de monoïdes finis. In Automata Theory, Ravello 1964, 314–319. Academic Press, New York, 1966.

[4] On a family of sets related to McNaughton's L-language. In Automata Theory, Ravello 1964, 320–324. Academic Press, New York, 1966.

Schützenberger's theorem on star-free languages

Star-free languages = smallest class of languages containing the finite languages and closed under Boolean operations and concatenation product.

Theorem

A language is star-free iff its syntactic monoid is aperiodic.

Schützenberger product of two monoids. If $\mathbf{H}$ is a variety of groups, the variety $\overline{\mathbf{H}}$ of all monoids whose groups belong to $\mathbf{H}$ is closed under Schützenberger product.

The Asilomar conference (September 1966)

Conference on the Algebraic Theory of Machines, Languages and Semigroups (Asilomar, California).

The proceedings (Arbib 1968) contain several chapters on the Krohn-Rhodes theory and on the local structure of finite semigroups (by Arbib, Rhodes, Tilson, Zeiger, etc.), a chapter "The syntactic monoid of a regular event" by McNaughton-Papert and other material.

The Asilomar conference (September 1966)

Conference on the Algebraic Theory of Machines, Languages and Semigroups (Asilomar, California).

The proceedings (Arbib 1968) contain several chapters on the Krohn-Rhodes theory and on the local structure of finite semigroups (by Arbib, Rhodes, Tilson, Zeiger, etc.), a chapter “The syntactic monoid of a regular event” by McNaughton-Papert and other material.

Schützenberger's famous provocative sentence: *The garbage truck driven by Arbib*. Didn't help the synergy between Paris and Berkeley...

1968-1969

Cohen and Brzozowski, On star-free events, *Proc. Hawaii Int. Conf. on System Science* (1968).

Definition of the dot-depth. Proof of Schützenberger's theorem using Krohn-Rhodes decomposition.

Meyer, A note on star-free events, *J. Assoc. Comput. Mach.* **16**, (1969)

Proof of Schützenberger's theorem using Krohn-Rhodes decomposition.

1972: Imre Simon's thesis

Hierarchies of events of dot-depth one.

Piecewise testable languages = Boolean combination of languages of the form $A^*a_1A^*a_2\cdots A^*a_kA^*$, where $a_1, \ldots, a_k$ are letters:

Theorem (Simon 1972)

A regular language is piecewise testable iff its syntactic monoid is $\mathcal{J}$-trivial.

Consequences in semigroup theory

Easily proved to be equivalent to

Theorem (Straubing-Thérien 1985)

Every $\mathcal{J}$-trivial monoid is a quotient of an ordered monoid satisfying the identity $x \leqslant 1$.

Many known proofs of one of the two results: Combinatorics on words [Simon], induction on $|M|$ [Straubing-Thérien], profinite techniques [Almeida], direct construction [Henckell-Pin], simplified combinatorics [Straubing, Klima], etc.

Locally testable languages

Brzozowski and Simon (1973), McNaughton (1974)

Locally testable languages = Boolean combination of languages of the form uA^*, A^*u and A^*uA^* where $u \in A^+$.

Theorem

A regular language is locally testable iff its syntactic semigroup is locally idempotent and commutative.

For each idempotent $e \in S$, eSe is idempotent and commutative.

Follow up

Brzozowski-Simon proved the first theorem on graph congruences. Motivated by problems on languages, further results were proved by Knast, Thérien, Weiss, etc.

Lead to the study of global varieties of categories (Almeida, Jones, Rhodes, Szendrei, Steinberg, Tilson, Trotter, etc.).

Straubing (1985) studied the decidability of varieties of the form $\mathbf{V} * \mathbf{D}$, where $\mathbf{D} = [\![yx^\omega = x^\omega]\!]$ and described the corresponding languages.

Lead to Tilson's delay theorem.

Eilenberg's book

To prepare his book Automata, languages, and machines (Vol A, 1974, Vol B 1976), S. Eilenberg worked with Schützenberger and Tilson.

Theorem (Eilenberg 1976)

There is bijection between varieties of monoids and varieties of languages.

Finite monoids	Regular languages
Aperiodic monoids	Star-free languages
$\mathcal{J}$-trivial monoids	Piecewise testable languages

Perrot's conjectures (September 1977)

Open problems in the theory of

Syntactic monoids for rational languages

J.F. Perrot

Sept 1977

<u>Notations</u> A Variety of rational languages $\mathcal{V}$ is understood in the sense of Eilenberg. Given an alphabet X, $X^*\mathcal{V}$ denotes the family of languages over X that belong to the variety $\mathcal{V}$. $\mathcal{Ab}(\mathcal{V})$ is the variety of finite

Perrot's conjectures

(1) Is there a variety of languages closed under concatenation whose corresponding variety of monoids is not of the form $\overline{\mathbf{H}}$?

(2) Is there a nontrivial variety of languages closed under shuffle product whose corresponding variety of monoids contains a noncommutative monoid?

(3) Is there a nontrivial variety of languages closed under the star operation?

Nontrivial means not equal to the variety of all regular languages.

Status of Perrot's conjectures

(3) was solved in [Pin TCS (1978)]

(2) was solved in [Esik and Simon, Semigroup Forum (1998)]

Status of Perrot's conjectures

(3) was solved in [Pin TCS (1978)]

(2) was solved in [Esik and Simon, Semigroup Forum (1998)]

What about conjecture (1)?

When Perrot presented this conjecture, someone in the back of the audience raised his hand...

S.W. Margolis H. Straubing

Two former students of Rhodes

Straubing's results

[JPAA 1979]: Closure under concatenation product on varieties of languages corresponds to the Mal'cev product $\mathbf{V} \to \mathbf{A} \textcircled{m} \mathbf{V}$ on varieties of monoids. ($\mathbf{A}$ is the variety of aperiodic monoids)

[JPAA 1979]: Description of the languages whose syntactic monoid is a solvable group [respectively, a monoid in which all subgroups are solvable].

REED COLLEGE Portland, Oregon 97202

Oct 12, 1978

Dear Drs. Perrot and Pin,

Thank you for your letter and the interesting papers you sent. I am quite impressed with the results on the shuffle product and the star.

I am sending you 3 papers. The one on aperiodic homomorphisms gives a characterization of ~~[illegible]~~ $\mathcal{M}$-varieties whose corresponding $*$-varieties are closed under concatenation. (Incidentally, you don't need the full strength of this theorem to show that there are $*$-varieties closed under concatenation that are not "variétés à groupes" – see the attached sheet.)

The paper on the Schützenberger product was just finished and has not yet been submitted to a journal. My dissertation is very close in content to the paper "Families of Recognizable Sets ..." but contains some additional results. I will send you a copy shortly.

Again, thanks very much for your interest. I hope you will keep me informed of any new developments ~~[illegible]~~

S.W. MARGOLIS

The University of Vermont

COLLEGE OF ENGINEERING, MATHEMATICS AND BUSINESS ADMINISTRATION
DEPARTMENT OF MATHEMATICS, NICHOLSON BUILDING
BURLINGTON, VERMONT 05405
(802) 656-2940

(1)

2/14/80

Dear Professor Pin:

Thank you for your recent letter and papers. I have just solved one of your conjectures and thought you would like to know. I'll give you an outline of the proof.

Theorem 1 If V is a variety then $P(V) = M \Leftrightarrow B_2 \notin V$.

Lemma The above theorem is true $\Leftrightarrow P(V) \neq M$ where V is the variety of monoids all of whose regular $\mathcal{J}$ classes are subsemigroups.

Pf. This follows from the observation that B_2 is in

I met Howard Straubing for the first time in Mai 1979. Our collaboration led to 9 joint articles.

In 1981, I visited Stuart in Vermont. Then Stuart spent 8 months in my place near Paris (Sept. 82 – June 83). This collaboration considerably improved my English and led to 16 joint articles and an uncountable number of coffee stamps on my papers.

So my Margolis number is $\frac{1}{16}$...

Semigroup Forum Vol. 22 (1981) 339-353

RESEARCH ARTICLE

ON M-VARIETIES GENERATED BY POWER MONOIDS

by

Stuart W. Margolis

Communicated by G. Lallement

ACKNOWLEDGEMENTS

I would like to thank Howard Straubing for bringing this problem to my attention and Jean-Eric Pin for sending me his work before it was published. Conversations with Garance Pin were amusing.

A selection of successful topics

(1) Languages and power semigroups
(2) Languages and inverse semigroups
(3) Profinite groups and the Rhodes conjecture
(4) Concatenation hierarchies

Languages and power semigroups

Given a variety of monoids $\mathbf{V}$, let $\mathbf{PV}$ be the variety generated by all monoids of the form $\mathcal{P}(M)$, where $M \in \mathbf{V}$.

[Reutenauer TCS 1979, Straubing SF 1979]: Closure under projections on varieties of languages corresponds to the operation $\mathbf{V} \to \mathbf{PV}$ on varieties of monoids.

Also closely related to the shuffle product.

Varieties of the form $\mathbf{PV}$

Objective: full classification of the varieties of the form $\mathbf{PV}$. Many results by Almeida, Margolis, Perrot, Pin, Putcha, Straubing, etc.

$\mathbf{PV} = \mathbf{M}$ iff $B_2 \in \mathbf{V}$ [Margolis 81]

$\mathbf{P}^3\mathbf{V} = \mathbf{M}$ [Margolis-Pin 84].

$\mathbf{PG}$ Powergroups (see later).

Full classification for $\mathbf{V} \subseteq \mathbf{A}$ (Almeida 05).

Major open problems: $\mathbf{PB}$, $\mathbf{PJ}$.

Extension to ordered monoids (ongoing work)

Languages and inverse semigroups

Let $\mathbf{Inv}$ be the variety of monoids generated by inverse monoids. What is $\mathbf{Inv}$? What is the corresponding variety of languages?

Let $\mathbf{J}_1$ = idempotent and commutative monoids.
Let $\mathbf{Ecom}$ = monoids with commuting idempotents.

$$\Diamond_2\mathbf{G} = \mathbf{Inv} = \mathbf{J}_1 * \mathbf{G} = \mathbf{J}_1 \circledM \mathbf{G} \overset{\text{Ash}}{=} \mathbf{Ecom}$$

[Margolis-Pin 84, Ash 87]

Extensions of a group by a semilattice

A monoid M is an extension of a group by a semilattice if there is a surjective morphism π from M onto a group G such that $\pi^{-1}(1)$ is a semilattice.

- How to characterize the extensions of a group by a semilattice?
- Is there a synthesis theorem in this case?
- In the finite case, what is the variety generated by the extensions of a group by a semilattice?

Covers, categories and inverse semigroups

Margolis-Pin, Marquette Conf. (1984) + 3 articles in J. of Algebra 110 (1987)

Extensions of groups by semilattices are exactly the E-unitary dense semigroups.

Representation of E-unitary dense semigroups by groups acting on categories. (First example of a derived category, due to Stuart).

Extension of McAlister's P-theorem (representation theorem on inverse semigroups) to nonregular semigroups.

The 1984 conjecture and its follow up

Conjecture. Every [finite] E-dense semigroup is covered by a [finite] E-unitary dense semigroup and the covering is one-to-one on idempotents.

The conjecture was solved by Ash (finite case, 1987) and Fountain (infinite case, 1990).

Subsequent research: Birget, Margolis, Rhodes, (1990), Almeida, Pin, Weil (1992), Fountain, Pin and Weil (2004): General extensions of a monoid by a group.

Rhodes conjecture (1972, Chico 1986)

The group radical of a monoid M is the set

$$K(M) = \bigcap_{\tau : M \to G} \tau^{-1}(1)$$

where the intersection runs over all relational morphisms from M into a group.

Fact. M belongs to $\mathbf{V} \circledcirc \mathbf{G}$ iff $K(M) \in \mathbf{V}$.

Let $D(M)$ be the least submonoid T of M closed under weak conjugation: if $t \in T$ and $a\bar{a}a = a$, then $at\bar{a} \in T$ and $\bar{a}ta \in T$.

Rhodes conjecture: $K(M) = D(M)$. Proved by Ash [1991].

Connection with pro-group topologies

Hall (1950): A topology for free groups and related groups.

Reutenauer (1979): Une topologie du monoïde libre.

Pin, Szeged (1987) + Topologies for the free monoid (1991).

Topological conjecture. A regular language is closed in the pro-group topology iff its ordered syntactic monoid satisfies $e \leqslant 1$ for every idempotent e.

Thm: The topological conjecture implies the Rhodes conjecture and gives a simple algorithm to compute the closure of L.

Finitely generated subgroups of the free group

Pin and Reutenauer, A conjecture on the Hall topology for the free group, (1991).

Thm: The topological conjecture is equivalent to the following statement: Let $H_1, \ldots, H_n$ be finitely generated subgroups of the free group. Then $H_1 H_2 \cdots H_n$ is closed.

This later property was proved by Ribes and Zalesskii (1993).

Many further developments and open problems (solvable groups).

Powergroups

[Margolis-Pin 84, Ash 87]

$$\Diamond_2 \mathbf{G} = \mathbf{Inv} = \mathbf{J}_1 * \mathbf{G} = \mathbf{J}_1 \circledM \mathbf{G} \overset{\text{Ash}}{=} \mathbf{Ecom}$$

[Margolis-Pin 85, Ash 91, Henckell-Rhodes 91]

$$\Diamond \mathbf{G} = \mathbf{PG} = \mathbf{J} * \mathbf{G} \overset{\text{Ash} + \text{HR}}{=} \mathbf{J} \circledM \mathbf{G} = \mathbf{BG} = \mathbf{EJ}$$

Also needs [Knast 83].
$\mathbf{BG}$ = Block groups = At most one idempotent in each $\mathcal{R}$-class and each $\mathcal{L}$-class.

$\mathbf{EJ}$ = Idempotents generate a $\mathcal{J}$-trivial monoid.

Concatenation hierarchy of star-free languages

Level 0: $\emptyset$ and A^*.

Level $n + 1/2$: Union of products of languages of level n.

Level $n + 1$: Boolean combination of languages of level n.

The hierarchy is infinite (Brzozowski-Knast 1978).

Level 1 = Piecewise testable languages = languages corresponding to $\mathbf{J}$. (Simon 72).

Level $3/2$ is also decidable (Pin-Weil 2001, using varieties of ordered monoids).

Level 2

Thm. Let L be a rational language and let M be its syntactic monoid. Are equivalent [Pin-Straubing 81]

(1) L has concatenation level 2.

(2) M divides a finite monoid of upper triangular Boolean matrices.

(3) M divides $\mathcal{P}(M)$, for some $\mathcal{J}$-trivial monoid M.

(4) L is expressible by a Boolean combination of Σ_2-formulas of Büchi's logic. [Thomas 82]

Many partial results, but decidability is still an open problem!

Conclusion and new directions

A lot of exciting collaboration between automata, semigroups and group theory over the past 60 years. Keep going! Here are some new trends:

Cost functions (Colcombet).

Words over ordinals and linear orders (Carton). Tree languages (Bojanczyk, Straubing).

Profinite equational theory for lattices of regular languages (Gehrke, Grigorieff, Pin).

Conclusion and new directions

A lot of exciting collaboration between automata, semigroups and group theory over the past 60 years. Keep going! Here are some new trends:

Cost functions (Colcombet).

Words over ordinals and linear orders (Carton). Tree languages (Bojanczyk, Straubing).

Profinite equational theory for lattices of regular languages (Gehrke, Grigorieff, Pin).

Convince Stuart to have a birthday conference every year during the next 40 years.