When does partial commutative closure preserve regularity?

Antonio Cano Gómez[1] Giovanna Guaiana[2]
Jean-Éric Pin[3]

[1]Universidad Politécnica de Valencia

[2]LITIS, Université de Rouen

[3]LIAFA, CNRS and Université Paris Diderot

Cremona, October 2008 (based on ICALP 2008)

supported by the ESF (European Science Foundation)
network AutoMathA

Summary

(1) Algebraic tools
(2) Two questions, known results
(3) Commutative closure
(4) Closure under partial commutation
(5) Summary and open problems

Partial commutation

Let A be a finite alphabet and let I be an independence relation. We denote by $\sim_I$ the congruence on A^* generated by the set

$$\{ab = ba \mid (a, b) \in I\}$$

If L is a language on A^*, we denote by $[L]_I$ the closure of L under $\sim_I$ (I-commutation). If I is the total commutation, we simplify the notation to $\sim$ and $[L]$.

The dependence relation is $D = (A \times A) \setminus I$.

Syntactic ordered monoid

- The syntactic preorder of L is the relation $\leqslant_L$ defined by : $u \leqslant_L v$ iff, for every $x, y \in A^*$,

$$xvy \in L \Rightarrow xuy \in L$$

- The syntactic congruence of L is the relation $\sim_L$ defined by : $u \sim_L v$ iff, for every $x, y \in A^*$,

$$xvy \in L \Leftrightarrow xuy \in L$$

- The syntactic ordered monoid of L is $(A^*/{\sim_L}, \leqslant_L /{\sim_L})$

An equivalent definition

If $\mathcal{A}$ is a minimal deterministic automaton, the relation $\leqslant$ defined on Q by $p \leqslant q$ iff

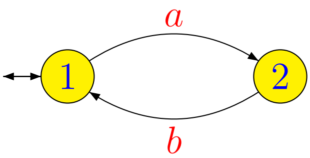
for all $u \in A^*$, $q{\cdot}u \in F \Rightarrow p{\cdot}u \in F$

is the syntactic order.

The syntactic ordered monoid of a language is the transition monoid of its ordered minimal automaton. The order is defined by

$$u \leqslant v \text{ iff, for all } q \in Q,\ q{\cdot}u \leqslant q{\cdot}v$$

The syntactic ordered monoid of $(ab)^*$



Order on the states :
$1 < 0$ and $2 < 0$

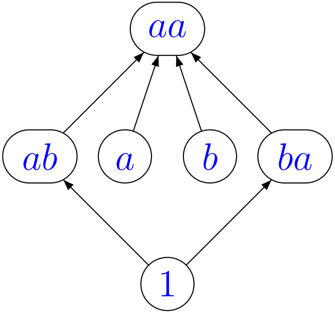
Elements

	1	2
1	1	2
a	2	0
b	0	1
aa	0	0
ab	1	0
ba	0	2

Relations

$bb = aa$
$aba = a$
$bab = b$

Order

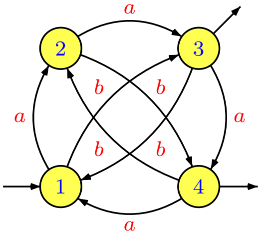


Group languages

A group language is a regular language whose syntactic monoid is a group, or, equivalently, is recognized by a deterministic automaton in which each letter defines a permutation of the set of states.

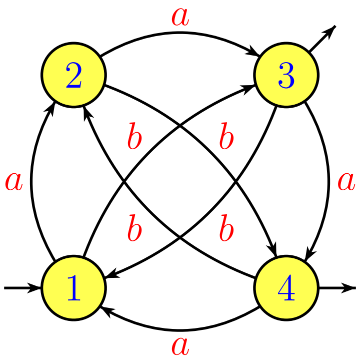
Examples on the alphabet $A = \{a, b\}$.

- {the words of odd length}
- $\{u \mid |u|_a + 2|u|_b \equiv 3 \bmod 5\}$
- $\{u \mid \binom{u}{ab} \equiv 1 \bmod 8\}$
- $(a(ab + bb)^*Aa)^*$



Commercial. . .

Use GasTeX !



Exponent and period of a regular language L

The exponent of L is the least integer ω such that, for all $u \in A^*$, $u^\omega \sim_L u^{2\omega}$.

The period of L is the least integer p for which, for all $u \in A^*$, $u^\omega \sim_L u^{\omega+p}$.

A language has period 1 iff it is star-fre (Schützenberger 65).

The period of a group language recognized by a group G divides $|G|$.

Two problems

- When is the closure of a regular language under [partial] commutation still regular?

- Are there any robust classes of languages closed under [partial] commutation?

Robust class = closed under some of the usual operations on languages, such as Boolean operations, product, star, shuffle, morphisms, inverses of morphisms, quotients, etc.

Known results on the first problem

Theorem (Ginsburg-Spanier 66, Gohon 85)

One can effectively decide whether the commutative closure of a regular language is regular or not.

Theorem (Sakarovitch 1992)

One can decide whether the closure $[L]_I$ of a regular language L is regular iff I is a transitive relation.

Star-free languages

Theorem (Muscholl, Petersen 1996)

One can decide whether the closure $[L]_I$ of a star-free language L is star-free iff I is a transitive relation.

Theorem (Muscholl, Petersen 1996)

Let L be a star-free language. If D is transitive, then $[L]_I$ is either star-free or non regular. If D is not transitive, then there exist star-free languages such that $[L]_I$ is regular but not star-free.

A convenient definition

Let $\mathcal{L}$ be a class of regular languages. We denote by $\mathrm{Pol}(\mathcal{L})$ the polynomial closure of $\mathcal{L}$, consisting of the finite unions of languages of the form

$$L_0 a_1 L_1 \cdots a_k L_k$$

where $a_1, \ldots, a_k \in A$ and $L_0, \ldots, L_k \in \mathcal{L}$.

Five classes of languages related to Problem 2

(1) $\mathsf{Pol}(\mathcal{I})$ = finite unions of languages of the form $A^*a_1A^* \cdots a_kA^*$,

(2) $\mathcal{J}$ = Boolean closure of $\mathsf{Pol}(\mathcal{I})$ (piecewise testable languages),

(3) $\mathsf{Pol}(\mathcal{J})$ = finite unions of languages of the form $A_0^*a_1A_1^* \cdots a_kA_k^*$ with $A_i \subseteq A$. Also called APC (Alphabetic Pattern Constraints) by Bouajjani, Muscholl and Touili.

(4) $\mathsf{Pol}(\mathcal{C}om)$ = polynomials of commutative languages.

Polynomial of group languages

Theorem (Pin, ICALP 1994)

Let L be a regular language. The following conditions are equivalent:

(1) *L is a polynomial of group languages,*

(2) *the ordered syntactic monoid of L satisfies the identity $x^\omega \leqslant 1$,*

(3) *L is open in the profinite topology.*

Known results on the second problem

We now summarise the results of Guaiana, Restivo and Salemi (00/04), Bouajjani, Muscholl and Touili (01/07) and Cécé, Héam and Mainier (07).

Theorem (Four classes Theorem)

Let I be any independence relation. Then

(1) *Pol*$(\mathcal{I})$ *is closed under commutation,*

(2) $\mathcal{J}$ *is closed under commutation,*

(3) *Pol*$(\mathcal{J})$ *is closed under I-commutation,*

(4) *Pol*$(\mathcal{C}om)$ *is closed under I-commutation.*

Our first result

Theorem

The commutative closure of a group language is regular.

Main argument: Insertion systems of Bucher, Ehrenfeucht and Haussler [1985].

Let π be a monoid morphism from A^* onto a finite group G and let $R = \pi^{-1}(1)$. Let us write $u \longrightarrow_R v$ if $u = u'u''$ and $v = u'ru''$ for some $r \in R$.

Let $\xrightarrow{*}_R$ be the reflexive transitive closure of the relation $\longrightarrow_R$. Then $\xrightarrow{*}_R$ is a well preorder.

A double improvement

Theorem

The commutative closure of a polynomial of group languages is also a polynomial of group languages.

Main argument: A refined description of the polynomials of group languages [Pin 1996] as a finite union of languages of the form $Ra_1R \cdots Ra_nR$, where R is a group language.
Then show that $[L]$ is closed under $\xrightarrow{*}_R$.

The positive variety $\mathcal{W}$

A positive variety of languages is a class of regular languages closed under union, intersection, quotients ($L \to u^{-1}L, Lu^{-1}$) and inverses of morphisms.

Theorem (Cano Gómez-Pin, 2002)

There exists a unique maximal positive variety $\mathcal{W}$ which does not contain the language $(ab)^*$, for all letters $a \neq b$.

Further, $\mathcal{W}$ is decidable, but this requires an algebraic characterization.

Algebraic characterization of $\mathcal{W}$

By a general result, a language belongs to $\mathcal{W}$ iff its syntactic ordered monoid belongs to a certain variety of finite ordered monoids **W**.

Theorem (Cano Gómez-Pin, 2002)

A finite ordered monoid M belongs to **W** *iff, for all pair (a, b) of elements of M such that $aba = a$ and $bab = b$, for all z of the minimal ideal of the submonoid of M generated by a and b, one has $(abzab)^\omega \leqslant ab$.*

The positive variety $\mathcal{W}$ is very robust!

- $\mathcal{W}$ is the unique maximal proper variety closed under shuffle,
- $\mathcal{W}$ is closed under product,
- $\mathcal{W}$ is closed under length decreasing morphism,
- $\mathcal{W}$ contains the group languages and their polynomial closure.
- $\mathcal{W}$ contains our "Four classes" and the languages whose syntactic monoid belongs to the variety **DS**.

Example. By definition, $(ab)^* \notin \mathcal{W}$. However $(ab)^* + A^*aaA^* \in \mathcal{W}$.

Main result

Theorem

If L is a language of $\mathcal{W}$, then $[L]$ is also in $\mathcal{W}$ (and hence is regular) and its period divides that of L.

Proof. Relies on the algebraic characterisation, which makes it difficult.

Corollary

The class of star-free languages of $\mathcal{W}$ is closed under commutation.

The case where I is transitive

Theorem

Let I be a transitive relation. If L is a language of $\mathcal{W}$, then $[L]_I$ is regular.

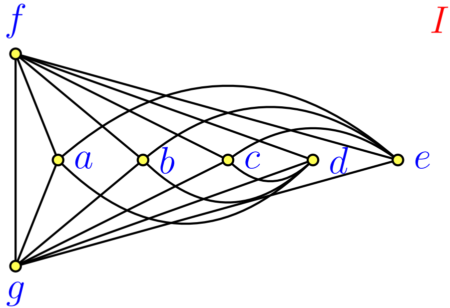
However, we don't know whether $[L]_I$ also belongs to $\mathcal{W}$.

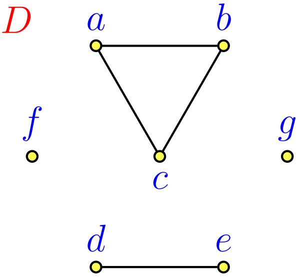
Proof. Makes use of ordered generalised automata (transitions are labeled by regular languages).

The case where D is transitive

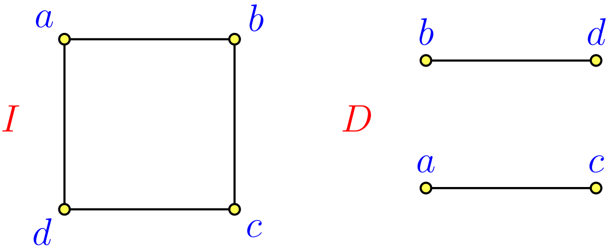
Theorem

Suppose that D is transitive. If L is a polynomial of group languages, then $[L]_I$ is also a polynomial of group languages.





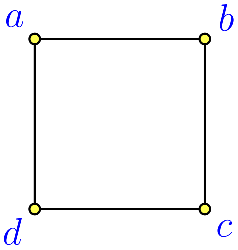
This result does not extend to $\mathcal{W}$.



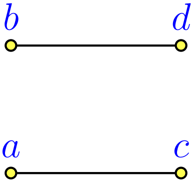
Here D is transitive. The language
$L = (abcd)^* + A^*aaA^* + A^*bbA^* + A^*ccA^* + A^*ddA^*$
$+ A^*ababA^* + A^*bcbcA^* + A^*cdcdA^* + A^*dadaA^*$
is in $\mathcal{W}$, but $[L]_I$ is not regular.

This result does not extend to $\mathcal{W}$.

I



D



Here D is transitive. The language
$L = (abcd)^* + A^*aaA^* + A^*bbA^* + A^*ccA^* + A^*ddA^*$
$+ A^*ababA^* + A^*bcbcA^* + A^*cdcdA^* + A^*dadaA^*$
is in $\mathcal{W}$, but $[L]_I$ is not regular.
It is not so easy to show that L is in $\mathcal{W}$, since its syntactic monoid has 170 elements...

Another extension

Let (A_1, I_1), ..., (A_k, I_k) be the connected components of the graph (A, I) and let (A_j, D_j) be the complement graph of (A_j, I_j).

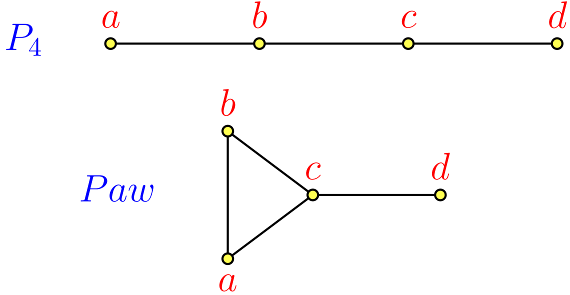
Theorem

If each graph (A_j, D_j) is transitive and L is a polynomial of group languages, then $[L]_I$ is regular.

A graph theoretic characterization

Theorem

The graphs (A_j, D_j) *are all* *transitive* *iff the graph* (A, I) *is* (P_4, Paw)*-free.*



Summary of our results

- Total commutation :
  if L is in $\mathcal{W}$, then so is $[L]$.
- I transitive :
  if L is in $\mathcal{W}$, then $[L]_I$ is regular.
- D transitive :
  if L is a polynomial of group languages, then so is $[L]_I$.
- (A, I) is (P_4, Paw)-free :
  if L is a polynomial of group languages, then $[L]_I$ is regular.

Conclusion and open problems

- The robust class $\mathcal{W}$ might offer an appropriate framework for modeling certain problems arising in the verification of concurrent systems.

- Open problems
  - If I is transitive and if L is in $\mathcal{W}$, does $[L]_I$ belong to $\mathcal{W}$?
  - If L is a group language, is $[L]_I$ always regular?
  - Is $\mathrm{Pol}(\mathcal{G})$ closed under partial commutations?
  - What about the polynomial closure of the family commutative languages $+$ group languages?