Theoretical Computer Science 103 (1992) 39–49
Elsevier

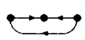
Semicommutations and algebraic languages*

M. Clerbout and Y. Roos

CNRS URA 369, L.I.F.L., Université de Lille 1, U.F.R. I.E.E.A. Informatique, 59655 Villeneuve d'Ascq Cedex, France

Abstract

Clerbout, M. and Y. Roos, Semicommutations and algebraic languages, Theoretical Computer Science 103 (1992) 39–49.

We give a decidable necessary and sufficient condition for a semicommutation function f_θ to transform any rational language into an algebraic language. This condition is the following: The semicommutation graph of f_θ has no subgraph isomorphic to or 

1. Introduction

The free partially commutative monoids study was initiated by Cartier and Foata [1], whose aim was to solve some combinatory problems. Trace languages, which are subsets of a free partially commutative monoid, were proposed by Mazurkiewicz [11] as tools for the description of concurrent program behaviour. Important results have been found and several syntheses have been written about this subject (see [5, 6, 9, 12, 13, 16, 17, 19]).

A partially commutative alphabet is a couple (A, θ), in which A is an alphabet and θ, the independence relation, is a symmetric and irreflexive binary relation over A. Associated with the commutation relation θ, an application $f_\theta : 2^{A^*} \to 2^{A^*}$ can be defined by: *For every language L over the alphabet A, $f_\theta(L)$ is the set of words which are equivalent to some word of L for the congruence generated by θ.* Thus, f_θ is a unary operation over languages, which is named partial commutation function associated with θ.

More recently, we introduced the notion of semicommutation, which generalizes the notion of partial commutation: a semicommutation is an irreflexive independence

* This work is supported by the P.R.C. Mathématiques et Informatique, and by the EBRA project Algebraic and Syntactic Methods in Computer Science.

Elsevier Science Publishers B.V.

relation over A (see [2, 3]). Since then, a lot of papers have been dealing with their properties (see [8, 10, 15, 18]).

When new operators, as semicommutation functions, are defined, a natural question is: *Do these operators preserve regularity*? Clearly, the answer is *no*: for any semicommutation function f_θ such that θ is not empty, there exist regular languages such that their image by f_θ is not regular. A more interesting question is the following: *If R is a regular language and f_θ a semicommutation function, is it decidable to know whether the image of R by f_θ is a regular language*? As a matter of fact, only decidable sufficient conditions have been found to ensure that the image by a semicommutation function of a regular language remains regular (see [4, 14, 15]).

In this paper, we answer the following question: *If f_θ is a semicommutation function, is it decidable to know whether the image of any regular language by f_θ is algebraic*? We name such functions *algebrico rational functions* and we give a decidable characterization of semicommutation functions which are algebrico rational.

2. Preliminaries

2.1. Notations

In the following text X is the used alphabet; u, v and w are words in X^*; Y is a subset of X.

$|w|$ is the length of the word w;

$|w|_x$ is the number of occurrences of the letter x that appear in the word w;

$|w|_Y$ is the number of occurrences of the letters of Y that appear in the word w;

$\mathrm{alph}(w)=\{x\in X \mid |w|_x\neq 0\}$ is the *alphabet* of the word w;

$\mathrm{com}(w)=\{u\in X^* \mid \forall x\in X, |w|_x=|u|_x\}$ is the *commutative closure* of the word w, and if $L\subseteq X^*$,

$$\mathrm{com}(L)=\bigcup_{w\in L}\mathrm{com}(w).$$

$F(w)$ is the set of *factors* of the word w, that is,

$$F(w)=\{u\in X^* \mid \exists v, v'\in X^*, \text{ with } w=vuv'\},$$

and if $L\subseteq X^*$, we extend the definition by

$$F(L)=\bigcup_{w\in L}F(w).$$

$\mathrm{LF}(w)$ is the set of *left factors* of the word w, that is,

$$\mathrm{LF}(w)=\{u\in X^* \mid \exists v\in X^*, \text{ with } w=uv\},$$

and if $L\subseteq X^*$, we extend the definition by

$$\mathrm{LF}(L)=\bigcup_{w\in L}\mathrm{LF}(w).$$

$\Pi_Y(w)$ is the *projection* of the word w over the *subalphabet* Y, i.e. the image of w by the homomorphism Π_Y which is defined by

$$\forall x \in X, \quad \text{if } x \in Y \text{ then } \Pi_Y(x) = x, \text{ else } \Pi_Y(x) = \varepsilon.$$

$u \sqcup\!\sqcup v$ is the *shuffle* of the two words u and v, that is,

$$u \sqcup\!\sqcup v = \{u_1 v_1 u_2 v_2 \ldots u_n v_n \mid u_i \in X^*,\ v_i \in X^* \text{ and } u = u_1 u_2 \ldots u_n,\ v = v_1 v_2 \ldots v_n\}.$$

$u \sqcap\!\sqcap v$ is the *synchronized shuffle* (see [7]) of the two words u and v, that is,

$$u \sqcap\!\sqcap v = \{w \in (\text{alph}(u) \cup \text{alph}(v))^* \mid \Pi_{\text{alph}(u)}(w) = u \text{ and } \Pi_{\text{alph}(v)}(w) = v\}.$$

$D_1^*(x, y)$ is the *Dyck language* on the alphabet $\{x, y\}$, that is,

$$D_1^*(x, y) = \{w \in \{x, y\}^* \mid |w|_x = |w|_y\}.$$

$D_1'^*(x, y)$ is the *semi-Dyck language* on $\{x, y\}$, that is,

$$D_1'^*(x, y) = \{w \in D_1^*(x, y) \mid \forall u \in \text{LF}(w),\ |u|_x \geq |u|_y\}.$$

Finally, Rat will denote the family of *rational languages*, Alg the family of *algebraic languages* and Ocl the one-counter languages family, which is the smallest set of languages which contains $D_1'^*(x, y)$ and which is closed under rational transductions, product, union and star.

2.2. Semicommutations

A *semicommutation relation* defined over an alphabet X is an irreflexive relation: it is a subset of $X \times X \setminus \{(x, x) \mid x \in X\}$.

With each semicommutation relation θ, we associate a rewriting system $S = \langle X, P \rangle$, which is named *semicommutation system* in which P is the set $\{xy \to yx \mid (x, y) \in \theta\}$. We shall write $u \underset{\theta}{\to} v$ if there is a rule $xy \to yx$ in P and two words w and w' such that $u = wxyw'$ and $v = wyxw'$. We shall write $u \overset{*}{\underset{\theta}{\to}} v$ if there are words $w_1, w_2, \ldots, w_n$ $(n \geq 1)$ such that $w_1 = u$, $w_n = v$, and for each $i < n$, $w_i \underset{\theta}{\to} w_{i+1}$. Then we shall write that there is a *derivation* from u to v.

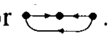
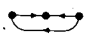
With each semicommutation θ we associate its *commutation graph*, which is the directed graph defined by (X, θ), where X is the vertex set and θ the edge set.

With each semicommutation relation θ, we associate a *semicommutation function* $f_\theta : 2^{A^*} \to 2^{A^*}$, which is defined by

$$\forall L \subset X^*, \quad f_\theta(L) = \bigcup_{w \in L} \left\{ u \in X^* \,\middle|\, w \overset{*}{\underset{\theta}{\to}} u \right\}.$$

3. Algebrico rational functions

Definition 3.1. A semicommutation function f defined on an alphabet X is *algebrico rational* if and only if for any rational language R included in X^*, the language $f(R)$ is algebraic.

Our main result is the following characterization of algebrico rational functions: A semicommutation function f defined on an alphabet X is algebrico rational if and only if the semicommutation graph of f has no subgraph isomorphic to  or . The proof of this result is based on an induction on the cardinality of the alphabet X. Clearly, when X contains only one letter, the semicommutation graph of any semicommutation function f defined on X cannot contain a subgraph isomorphic to or , and it is easy to verify that f is algebrico rational. To give an idea of what happens in the general case, we will also examine the case of a two-letter alphabet.

Let $X=\{a,b\}$. As in the case of a one-letter alphabet, we have to verify that each semicommutation function defined on X is algebrico rational. There are four semicommutation functions on a two-letter alphabet: the identity (no commutation at all), com (the total commutation), $f_{ab\to ba}$ associated with the rule $ab\to ba$, and $f_{ba\to ab}$ associated with the rule $ba\to ab$. At first, we give a necessary and sufficient condition for a word w' to be in the image of a word w of X^* by $f_{ab\to ba}$ (the proof of this result is in [2]).

Lemma 3.2. *Let w and w' be two words of X^*. $w'\in f_{ab\to ba}(w)$ if and only if $w'\in\mathrm{com}(w)$ and $\forall(u,v)\in \mathrm{LF}(w)\times LF(w')$, $|u|=|v|\Rightarrow|u|_b\leqslant|v|_b$.*

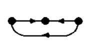
Then we can state the following proposition.

Proposition 3.3. *Any semicommutation function defined on $X=\{a,b\}$ is algebrico rational.*

Proof. If f is the identity, the result is obvious: $f(R)=R\in\mathrm{Rat}\subset\mathrm{Alg}$. Latteux [9] proved that if $f=\mathrm{com}$, $\forall R\in\mathrm{Rat}$, $\mathrm{com}(R)\in\mathrm{Ocl}$. So we have to establish that $f_{ab\to ba}(R)$ is context-free, for each rational language R. The proof is symmetric for $f_{ba\to ab}$. Let h be the morphism defined on $\{a,b,\bar{b}\}$ by $h(a)=a$, $h(b)=b$, $h(\bar{b})=\varepsilon$ and let g be the morphism defined on the same alphabet by $g(a)=a$, $g(b)=\varepsilon$, $g(\bar{b})=b$. We have: $\forall u\in X^*, f(u)=g(h^{-1}(u)\cap(D_1'^*(\bar{b},b)\sqcup\!\sqcup a^*))$. Indeed, set $u'\in f(u)$ and let us denote by $\bar{u}'$ the word u' where each occurrence of the letter b has been marked. ($\bar{u}'=m(u')$ with $m:\{a,b\}\mapsto\{a,\bar{b}\}$, $m(a)=a$, $m(b)=\bar{b}$.) Set $v=u\sqcap\bar{u}'\cap(\bar{b}^*b^*a)^*$. By Lemma 3.2, it is clear that $\Pi_{\{b,\bar{b}\}}(v)\in D_1'^*(\bar{b},b)$; so, $v\in h^{-1}(u)\cap D_1'^*(\bar{b},b)\sqcup\!\sqcup a^*$ and $u'=g(v)$. On the other hand, if $u'\in h^{-1}(u)\cap(D_1'^*(\bar{b},b)\sqcup\!\sqcup a^*)$, each left factor α of $\mathrm{LF}(u')$ satisfies $|\alpha|_{\bar{b}}\geqslant|\alpha|_b$. So, $g(u')\in f(\Pi_{\{a,b\}}(u'))=f(u)$. As $D_1'^*(\bar{b},b)\in\mathrm{Ocl}$, which is a family closed under rational transduction, each rational language has its image by f in Ocl, so f is algebrico rational. □

Remark. If L is an algebraic language, there is a rational language R such that $\mathrm{com}(L)=\mathrm{com}(R)$ (see [9]); so $\mathrm{com}(L)$ is an algebraic language. However, $f_{ab\to ba}(L)$ is not always context-free: Set $L=\{(ba)^n b^n,\ n\geqslant 0\}$. $L\in\mathrm{Alg}$, but $f_{ab\to ba}(L)\cap b^*a^*b^* =\{b^{n+k}a^n b^{n-k},\ n\geqslant k\geqslant 0\}=L_1$. And $\mathrm{LF}(b^*L_1)=\{b^n a^p b^q,\ n\geqslant p\geqslant q\geqslant 0\}\notin\mathrm{Alg}$.

Let us now suppose that the cardinality of the alphabet X is greater than 2.

Definition 3.4. Let f be a semicommutation function defined on the alphabet X. We say that f satisfies the (C) condition if the semicommutation graph of f has no subgraph isomorphic to or .

We may also express this condition in the following way: A semicommutation function f on X associated with the semicommutation relation C_0 satisfies the (C) condition if and only if

$$(y,z)\in C_0 \;\Rightarrow\; ((y,x_1)\in C_0 \text{ and } (x_2,z)\in C_0 \;\Rightarrow\; x_1=z \text{ or } x_2=y).$$

Proposition 3.5. *If a semicommutation function is algebrico rational then it satisfies the (C) condition.*

Proof. Let f be a semicommutation function defined on $X=\{a,b,c\}$ by the commutation graph shown in Fig. 1. Set $R=(abc)^*$. Then

$$f(R)\cap c^*b^*a^* = \{c^n b^n a^n \mid n\in N\}\notin \text{Alg}.$$

Let g be the function defined on $\{a,b,c,d\}$ by the semicommutation graph shown in Fig. 2. Set $R'=(cd)^*(ab)^*$. Then

$$g(R')\cap d^*b^*c^*a^* = \{d^n b^p c^n a^p \mid n,p\in N\}\notin \text{Alg}.$$

A function which does not satisfy the (C) condition would never be algebrico rational. □

We shall now prove the converse of this proposition. We will consider two cases.

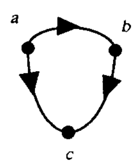


Fig. 1.

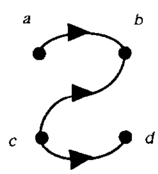


Fig. 2.

Definition 3.6. We say that a semicommutation function f defined on X satisfies the (P) *property* if and only if there exists a letter x in X such that for each letter y in X, $yx \in f(xy)$ or such that for each letter y in X, $xy \in f(yx)$.

So, let f be a semicommutation function defined on X which satisfies both the (C) condition and the (P) property, i.e. there exists a letter x in X such that

$$\forall y \in X, \quad yx \in f(xy).$$

The other case ($xy \in f(yx)$) would be studied in the same way. Let us explain what the function f does:

$$\forall y_1, y_2 \in X \setminus \{x\}, \quad y_1 y_2 \notin f(y_2 y_1).$$

Because the commutation graph of f already contains $y \longleftarrow x \longrightarrow y_2$ it is impossible to add an arrow between y_1 and y_2 since f satisfies the (C) condition. However, we may have commutations of the kind $yx \to xy$, $y \in X \setminus \{x\}$. Thus, the alphabet X may be partitioned into three disjoint subsets: $X = X_1 \cup X_2 \cup \{x\}$, with $X_1 = \{y \in X \setminus \{x\}, xy \in f(yx)\}$, $X_2 = \{y \in X, xy \notin f(yx)\}$. It means that, in a word $w \in X^*$, the occurrences of the letter x are going to move on left or right, in each factor of w which is in X_1^*, but an occurrence of x may move over a letter of X_2 only from left to right. When adding in w the new positions of marked occurrences of the letter x ($\bar{x}$ instead of x) from a word w' in $f(w)$, we get words of $D_1^*(x, \bar{x}) \sqcup\!\sqcup X_1^*$ and words of $D_1'^*(x, \bar{x}) \sqcup\!\sqcup (X_1 \cup X_2)^*$. This is what is formalized in the following lemmas.

Lemma 3.7. *Let f be a semicommutation function such that f satisfies (C) and (P). Let $u, u' \in X^*$. Then $u \xrightarrow{*} u'$ if and only if the following conditions hold*:

(1) $\Pi_{X-\{x\}}(u) = \Pi_{X-\{x\}}(u')$.

(2) $\Pi_x(u) = \Pi_x(u')$.

(3) *for all prefixes $u_1 \in X^* X_2$ of u and $u'_1 \in X^* X_2$ of u', we have*

$$\Pi_{X-\{x\}}(u_1) = \Pi_{X-\{x\}}(u'_1) \;\Rightarrow\; |u_1|_x \geqslant |u'_1|_x.$$

Proof. ($\Rightarrow$): It is clear that conditions (1) and (2) are necessary. We prove condition (3) by induction on the length of the derivation $u \xrightarrow{n} u'$. The value $n = 0$ being obvious, consider $u \xrightarrow{n} u' \to u''$. Two cases may arise:

- $u' = w_1 xy w_2$ and $u'' = w_1 yx w_2$, with $y \in X_1 \cup X_2$.
- $u' = w_1 yx w_2$ and $u'' = w_1 xy w_2$, with $y \in X_1$.

Let u_1, u'_1 and $u''_1 \in X^* X_2$ be prefixes of u, u' and u''. If $\Pi_{X-\{x\}}(u_1) = \Pi_{X-\{x\}}(u'_1)$ then, by induction hypothesis, $|u_1|_x \geqslant |u'_1|_x$. Moreover, in both cases, if $y \in X_1$ then $\Pi_{X-\{x\}}(u'_1) = \Pi_{X-\{x\}}(u''_1)$ implies that $|u'_1|_x = |u''_1|_x$. Now if $y \in X_2$ and $w = \Pi_{X-\{x\}}(u'_1) = \Pi_{X-\{x\}}(u''_1)$ then $|u'_1|_x = |u''_1|_x$ if $w \neq \Pi_{X-\{x\}}(w_1 y)$, and $|u'_1|_x = |u''_1|_x + 1$ otherwise.

($\Leftarrow$): If $u, u' \in X^*$ satisfy the three conditions then $|u| = |u'|$. Let w be the longest common prefix of u and u'. We argue on the integer $n = |u| - |w| = |u'| - |w|$. If $n = 0$

then $u=u'$ and there is nothing to prove; so we assume $n>0$. Because of condition (3) we have only the following different cases:

- $u=wx^r yu_2$ and $u'=wyu'_2$ with $r>0$ and $y\in X_1\cup X_2$.
- $u=wyu_2$ and $u'=wx^r yu'_2$ with $r>0$ and $y\in X_1$.

In the first case we have $u\xrightarrow{*}u''=wyx^r u_2$. Furthermore, u' and u'' satisfy the three conditions of the statement; so, by induction hypothesis, we have $u''\xrightarrow{*}u'$; thus, $u\xrightarrow{*}u'$. In the second case we have $wyx^r u'_2=u''\xrightarrow{*}u'=wx^r yu'_2$. Since u and u'' satisfy the three conditions, we obtain $u\xrightarrow{*}u''$; thus, $u\xrightarrow{*}u'$. □

Notation. If w is a word of X^*, $\bar{w}$ denotes the image of w by the morphism which marks the letter x: $m: X\mapsto X\cup\{\bar{x}\}$, $m(x)=\bar{x}$, and $\forall y\in X\setminus\{x\}$, $m(y)=y$.

Now we are able to prove the following lemma.

Lemma 3.8. *Given $u,u'\in X^*$, let $v\in X^*$ be the word in $u \sqcap \bar{u}' - X^*\bar{x}xX^*$ when it exists. Then $u\xrightarrow{*}u'$ if and only if $v\in((D_1^*(x,\bar{x})\sqcup\!\sqcup X_1^*)(D_1'^*(x,\bar{x})\sqcup\!\sqcup(X_1\cup X_2)^*))^*$.*

Proof. ($\Rightarrow$): By Lemma 3.7, if $u\xrightarrow{*}u'$ holds then the word v exists. Furthermore, we factorize $v=v_1v_2\ldots v_m$ in a unique way as follows. For all prefixes w of v consider the difference $\delta(w)=|w|_x-|w|_{\bar{x}}$. The factors $v_1,v_2,\ldots$ correspond to the positions in the word v where $\delta(w)$ changes sign. Formally, the factor v_1 is the longest prefix of v such that for all prefixes w of v_1, $|\Pi_x(w)|\leqslant|\Pi_{\bar{x}}(w)|$ holds. Set $v=v_1v'$ and define v_2 as the longest prefix of v' such that for all prefixes w of v', $|\Pi_x(w)|\geqslant|\Pi_{\bar{x}}(w)|$ holds. Set $v=v_1v_2v''$ and apply this procedure recursively to v''.

It suffices to verify that for $k=0,1,\ldots$ the word v_{2k+1} belongs to X_1^*. Assume by contradiction that this is not the case, i.e. for some $k=0,1,\ldots$ and some words $t\in(X_1^*\cup\{x\}\cup\{\bar{x}\})^*$, $v'\in X^*$ and $y\in X_2$ we have

$$\Pi_{X-\{x,\bar{x}\}}(v_{2k+1})=tyv'.$$

Then $u_1=\Pi_{X-\{\bar{x}\}}(v_1v_2\ldots v_{2k}ty)$ is a prefix of u and $u'_1=\Pi_{X-\{x\}}(v_1v_2\ldots v_{2k}ty)$ is a prefix of u' that satisfy $\Pi_{X-\{x\}}(u_1)=\Pi_{X-\{x,\bar{x}\}}(v_1v_2\ldots v_{2k}ty)=\Pi_{X-\{\bar{x}\}}(u'_1)$, implying that $|u_1|_x\geqslant|u'_1|_x$. This contradicts the definition of v_{2k+1}.

($\Leftarrow$): If a word v exists then u and u' satisfy conditions (1) and (2) of Lemma 3.7. Furthermore, because of the form of v, if a prefix $u_1\in X^*X_2$ of u and a prefix $u'_1\in X^*X_2$ of u' satisfy

$$\Pi_{X-\{x\}}(u_1)=\Pi_{X-\{x\}}(u'_1)$$

then there exists a prefix w of v such that

$$\Pi_{X-\{x\}}(u_1)=\Pi_{X-\{x\}}(u'_1)=\Pi_{X-\{x,\bar{x}\}}(w).$$

By hypothesis, we have $\Pi_{\{x,\bar{x}\}}(w)\in D_1^*(x,\bar{x})\mathrm{LF}(D_1'^*(x,\bar{x}))$, which is exactly condition (3). □

In this light, we can state that if f satisfies both the (P) property and the (C) condition then f is an algebraic relation.

Proposition 3.9. *Let f be a semicommutation function defined on the alphabet X, which satisfies the (C) condition and for which there exists a letter x such that $\forall y \in X$, $yx \in f(xy)$. Then we can find morphisms h and g and two subsets of X, X_1 and X_2, such that*

$$\forall u \in X^*, \quad f(u) = g(h^{-1}(u) \cap ((D_1^*(x,\bar{x}) \sqcup\!\sqcup X_1^*)(D_1'^*(x,\bar{x}) \sqcup\!\sqcup (X_1 \cup X_2)^*))^*).$$

So f is algebrico rational.

Proof (*sketch*). Set $X_1 = \{y \in X \setminus \{x\} \mid xy \in f(yx)\}$ and $X_2 = X \setminus (X_1 \cup \{x\})$. Let h and g be the morphisms defined on $X \cup \{\bar{x}\}$ by

$$\forall y \in X_1 \cup X_2, \quad h(y) = y, \; g(y) = y,$$

$$h(x) = x, \qquad g(x) = \varepsilon,$$

$$h(\bar{x}) = \varepsilon, \qquad g(\bar{x}) = x.$$

Set $L = ((D_1^*(x,\bar{x}) \sqcup\!\sqcup X_1^*)(D_1'^*(x,\bar{x}) \sqcup\!\sqcup (X_1 \cup X_2)^*))$. It is easy to see that

$$\forall u \in X^*, \quad f(u) = g(h^{-1}(u) \cap L^*).$$

So f is algebrico rational. □

Now we study the case where the (P) property is not satisfied.

Notation. If w is a word of X^*, and $t \in N^+$, we write $w(t)$ as the left factor of length t of w.

Lemma 3.10. *Let $u \in X^+$, $a \notin X$, $w \in a(u \sqcup\!\sqcup a^i)$, $w' \in u \sqcup\!\sqcup a^i a^+$. Then we can find $t_0 \in N^+$ such that*

(1) $\text{com}(w(t_0)) = \text{com}(w'(t_0))$;
(2) $\forall s \in \{1, \ldots, t_0 - 1\}$, $|w'(s)|_a < |w(s)|_a$.

Proof. Let t_0 be the smallest element of $\{t \mid t \in N^+ \text{ and } |w'(t)|_a \geqslant |w(t)|_a\}$. t_0 exists since, if $t = |w|$, we have $|w'(t)|_a \geqslant |w(t)|_a$. Then $|w'(t_0)|_a = |w(t_0)|_a$ and, thus, $|\Pi_X(w'(t_0))| = |\Pi_X(w(t_0))|$, which implies that $\Pi_X(w'(t_0)) = \Pi_X(w(t_0))$ since these two words are left factors of u. Therefore, $\text{com}(w(t_0)) = \text{com}(w'(t_0))$. From the definition of t_0 the second assertion is satisfied. □

For a given semicommutation function f, if no letter may commute with each of the others, shuffles will be local. So, to get the image of a word by f, it is sufficient to make shuffles on factors which are defined on a smaller alphabet. This motivates the following lemma.

Lemma 3.11. *Let f be a semicommutation function defined on X, which satisfies the (C) condition but not the (P) property. Then, for any word u of* X^*, *for any word v of* $f(u)$, *we can find decompositions* $u=u_1u_2$ *and* $v=v_1v_2$ *with* $u_1\neq\varepsilon$, $\mathrm{alph}(u_1)\subsetneq X$ *and* $v_1\in f(u_1)$.

Proof. If $\mathrm{alph}(u)\subsetneq X$, the result is obvious. If not, set $u=au'$, $v=dv'$, $a\in X$, $d\in X$. Then
either $a=d$: we can choose $u_1=v_1=a$;
or $a\neq d$: we set $D=\{z\in X\setminus\{a\}\mid za\in f(az)\}$ and $Y=X\setminus(D\cup\{a\})$.
Then
- $d\in D$; so $D\neq\emptyset$;
- $Y\neq\emptyset$ since f does not verify the (P) property;
- if z_1 and z_2 are in D, the graph of f contains a subgraph $z_1 \leftarrow a \rightarrow z_2$; thus, there is no commutation between z_1 and z_2 because f satisfies the (C) condition.

We have to consider two different cases.

First case: *There do not exist letters* $y\in Y$ *and* $z\in D$ *such that* $zy\in f(yz)$. Then let us set $u=awyu''$ and $v=dw'y'v''$ with $w,w'\in(D\cup\{a\})^*$ and $y,y'\in Y$. No occurrence of letter a in aw can overstep y because $\forall x\in Y$, $xa\notin f(ax)$. So we have $|aw|_a\leqslant|w'|_a$, i.e. $|w'|_a>|w|_a$.

On the other hand, for two letters (d_1,y_1) in $D\times Y$, it is possible to have $y_1d_1\in f(d_1y_1)$, but, since $d_1y_1\notin f(y_1d_1)$, we get $\Pi_D(dw')\in\mathrm{LF}(\Pi_D(dw))$. Then

$$u=u'yu'',\quad u'\in a(\Pi_D(dw')\sqcup a^i)(w''\sqcup a^{i'}),\ \ \Pi_D(w)=\Pi_D(dw')w'',$$

$$v=v'y'v'',\quad v'\in\Pi_D(dw')\sqcup a^j,$$

where $j=|w'|_a$, $i+i'=|w|_a$; so $j>i$.

From Lemma 3.10, it follows that $u'=u_1u'_2$ and $v'=v_1v'_2$, where $u_1\neq\varepsilon$ and $\mathrm{com}(u_1)=\mathrm{com}(v_1)$ (so $v_1\in f(u_1)$). Moreover, $\mathrm{alph}(u_1)\subset D\cup\{a\}\subsetneq X$. Hence, the couple (u_1,v_1) answers the problem.

Second case: *There exist letters* $y\in Y$ *and* $d_1\in D$ *such that* $d_1y\in f(yd_1)$. Then $D=\{d\}$; if two different letters d_1 and d_2 belong to D, we will find in the semicommutation graph of f the subgraph $y \leftarrow d_1 \rightarrow a \leftarrow d_2$, which contradicts the hypothesis.

Then we set $E=\{z\in X\setminus\{d\}\mid dz\in f(zd)\}$ and $Z=X\setminus(E\cup\{d\})$. We have
- $E\neq\emptyset$ because $y\in E$;
- $Z\neq\emptyset$ because f does not verify the (P) property;
- if z_1 and z_2 are two different letters of E, there is no commutation between z_1 and z_2 because the graph of f already contains $z_1 \leftarrow d \rightarrow z_2$ and f satisfies the (C) condition.

Let us set $u=awzu''$ and $v=dw'y'v''$, with $w,w'\in(E\cup\{d\})^*$ and $z,z'\in Z$. No occurrence of the letter d in u'' can overstep z and z', but it is possible to have rules such as $dt\rightarrow td$ for a letter t in Z. So we get $|aw|_d\geqslant|dw'|_d$, i.e. $|w|_d>|w'|_d$.

On the other hand, $\forall x\in E\setminus\{a\}$, $\forall z\in Z$, the graph of f contains $a \leftarrow d \rightarrow x$. No arrow like $x \leftarrow z$ can be added and $za\notin f(az)$ because $D=\{d\}$ and $d\notin Z$. Thus, $\forall x\in E$,

$\forall z \in Z$, $zx \notin f(xz)$. But we can find a rule as $z_1 x_1 \to x_1 z_1$ for $(x_1, z_1) \in E \times Z$. So we have $\Pi_E(aw) \in \mathrm{LF}(\Pi_E(dw')) = \mathrm{LF}(\Pi_E(w'))$. Then we write

$$u = u'zu'', \quad u' \in \Pi_E(aw) \sqcup\!\sqcup\, d^j,$$

$$v = v'z'v'', \quad v' \in d(\Pi_E(aw) \sqcup\!\sqcup\, d^j)(w'' \sqcup\!\sqcup\, d^{i'}), \ \Pi_E(dw') = \Pi_E(aw)w'',$$

where $j = |aw|_d$, $i + i' = |w'|_d$; so $j > i$.

According to Lemma 3.10, we get $u' = u_1 u_2''$ and $v' = v_1 v_2''$, where $u_1 \neq \varepsilon$, $\mathrm{com}(u_1) = \mathrm{com}(v_1)$ (so $v_1 \in f(u_1)$), and $\mathrm{alph}(u_1) \subset E \cup \{d\} \subsetneq X$. The couple (u_1, v_1) agrees with the question, proving the result. □

We are now able to state the main result of the paper.

Proposition 3.12. *Let f be a semicommutation function defined on the alphabet X, and satisfying the (C) condition. Then f is algebrico rational.*

Proof. We proceed by induction on the cardinality of the alphabet X, denoted by $\mathrm{card}(X)$. If $\mathrm{card}(X) = 2$, the result is true; see Proposition 3.3. If $\mathrm{card}(X) > 2$ then if f satisfies the (P) property, the result is true because of Proposition 3.9. If not, we are going to show that for each rational language R, $f(R) \in \mathrm{Alg}$. Let $R \in \mathrm{Rat}$. We can define a deterministic automaton $M = (X, Q, q_0, *, F)$ which accepts R. If $q, q' \in Q$, we set $R_{q,q'} = \{w \in X^*, q * w = q'\}$. Let s be the substitution defined on $Q \times Q$ by

$$\forall (q, q') \in Q \times Q, \quad s((q, q')) = \bigcup_{x \in X} f(R_{q,q'} \cap (X \setminus \{x\})^*).$$

By induction hypothesis, s is an algebraic substitution. Let K be the rational language defined on $Q \times Q$ by

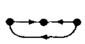
$$K = \{(q_0, q_1)(q_1, q_2) \ldots (q_{p-1}, q_p) \mid p \geqslant 1, \forall i \in \{1, \ldots, p\}, q_i \in Q, q_p \in F\}.$$

We can easily show that $f(R) = s(K)$, and the proof is complete since the image of a rational language by an algebraic substitution is an algebraic language. □

As a matter of fact, Propositions 3.5 and 3.12 permit us to state that the image of a rational language by a semicommutation function is always in Ocl. So we state the following proposition.

Proposition 3.13. *Let f be a semicommutation function defined on X. The following assertions are equivalent:*

(1) *f is algebrico rational.*

(2) *The semicommutation graph of f does not contain any subgraph isomorphic to* *or*  .

(3) *For each rational language R, $f(R) \in$* Ocl.

(4) *For each rational bounded language R, $f(R) \in$* Alg.

Proof. (1)⇒(2) by proposition 2. (2)⇒(3) because constructions which give us the image by f of a rational language use $D_1'^*$ and D_1^* which are in Ocl and operations under which Ocl is closed. (3)⇒(4) is obvious. (4)⇒(1) when looking at the proof of Proposition 3.5. □

In the particular case of partial commutation (associated with irreflexive and symmetrical relations), the results of Propositions 3.5 and 3.12 become Proposition 3.14.

Proposition 3.14. *A partial commutation function is algebrico rational if and only if its commutation graph does not contain a path whose length is* 3.

Acknowledgment

We thank the anonymous referees for suggesting several improvements.

References

[1] P. Cartier and D. Foata, *Problèmes combinatoires de commutations et réarrangements*, Lecture Notes in Mathematics, Vol. 85 (Springer, Berlin, 1969).
[2] M. Clerbout, Commutations partielles et familles de langages, Thèse, Université de Lille 1, 1984.
[3] M. Clerbout and M. Latteux, On a generalization of partial commutations, in: M. Arato, I. Katai and L. Varga, eds. *Proc. Fourthung. Computer Sci. Conf.* (1985) 15–24.
[4] M. Clerbout and M. Latteux, Semi-commutations, *Inform. and Comput.* **73** (1987) 59–74.
[5] R. Cori, Partially abelian monoids, invited lecture, STACS, Orsay, 1986.
[6] R. Cori and D. Perrin, Automates et commutations partielles, *RAIRO Inform. Theor. Appl.* **19** (1985) 21–32.
[7] R. De Simone, Langages infinitaires et produit de mixage, *Theoret. Comput. Sci.* **31** (1984) 83–100.
[8] G. Lampérth, Reducedness of formal languages, in: *Proc. IMYCS Conf.*, Smolenice (1986) 215–219.
[9] M. Latteux, Cônes rationnels commutatifs, *J. Comput. System Sci.* **18** (1979) 307–333.
[10] M. Latteux, Some results on semi-commutations, *Act of the International Colloquium on Words, Languages and Combinatorics*, Kyoto, 1990.
[11] A. Mazurkiewicz, Concurrent program schemes and their interpretations, DAIMI PB 78, University of Aarhus, 1977.
[12] A. Mazurkiewicz, Traces, histories and graphs: instances of process monoids, in: Lecture Notes in Computer Science, Vol. 176 (Springer, Berlin, 1984) 115–133.
[13] Y. Métivier, On recognizable subsets of free partially commutative monoids, in: Lecture Notes in Computer Science, Vol. 226 (Springer, Berlin, 1986) 254–264.
[14] Y. Métivier, Contribution à l'étude des monoides de commutations, Thèse, Université de Bordeaux 1, 1987.
[15] Y. Métivier and E. Ochmanski, On lexicographic semicommutations, *Inform. Process. Lett.* **26** (1987) 55–59.
[16] E. Ochmanski, Regular behaviour of concurrent systems, *Bull. EATCS* **27** (1985) 56–67.
[17] D. Perrin, Words over a partially commutative alphabet, in: NATO ASI Series F12 (Springer, Berlin, 1985) 329–340.
[18] M. Szijarto and D. Van Hung, Synchronized parallel composition of languages, in: *Proc. OALPS Conf.* Salgotarjan (1986) 281–288.
[19] W. Zielonka, Notes on asynchronous automata, *RAIRO Inform. Theor. Appl.* **21** (1987) 99–135.