

YVES MÉTIVIER

## **Une condition suffisante de reconnaissabilité dans un monoïde partiellement commutatif**

*Informatique théorique et applications*, tome 20, n° 2 (1986), p. 121-127.

[http://www.numdam.org/item?id=ITA\\_1986\\_\\_20\\_2\\_121\\_0](http://www.numdam.org/item?id=ITA_1986__20_2_121_0)

© AFCET, 1986, tous droits réservés.

L'accès aux archives de la revue « Informatique théorique et applications » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/legal.php>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme  
Numérisation de documents anciens mathématiques  
<http://www.numdam.org/>

## UNE CONDITION SUFFISANTE DE RECONNAISSABILITÉ DANS UN MONOÏDE PARTIELLEMENT COMMUTATIF (\*)

par Yves MÉTIVIER (1)

Communiqué par J.-E. PIN

---

Résumé. – On généralise un résultat de Cori et Perrin sur le caractère rationnel de certains langages obtenus par clôture suivant des commutations partielles.

Abstract. – We extend a result of Cori and Perrin on the rational behaviour of some languages obtained by closure according partial commutations.

### 0. INTRODUCTION

Un monoïde partiellement commutatif libre est un monoïde engendré par un alphabet fini  $A$  et dans lequel toutes les relations sont de la forme  $ab \simeq ba$  pour certaines lettres  $a$  et  $b$ .

Étant donné un sous-ensemble  $X$  de  $A^*$ , on note  $[X]$  l'ensemble des mots qui sont équivalents à au moins un des mots de  $X$ .

Si  $B$  désigne un sous-ensemble de  $A$ , on définit le graphe des non-commutations associé à  $B$  comme étant le graphe dont les sommets sont les éléments de  $B$  et les arêtes les paires de lettres de  $B$  qui ne commutent pas. De plus, si  $x$  est un mot de  $A^*$ , on note  $\text{alph}(x)$  le sous-ensemble de  $A$  formé des lettres qui apparaissent dans  $x$ .

Dans cette note, nous montrons le résultat suivant :

*Si  $X$  est un ensemble reconnaissable tel que  $[X]=X$  et si pour tout mot  $x$  de  $X$  le graphe des non-commutations associé à  $\text{alph}(x)$  est connexe alors  $[X^*]$  est reconnaissable.*

---

(\*) Reçu Octobre 1984, révisé en février 1985.

(1) Université de Bordeaux-I, U.E.R. de Mathématiques et Informatique, 33405 Talence, France.

Ce théorème généralise d'une part celui de R. Cori et D. Perrin [1] qui le démontrent dans le cas où deux lettres distinctes d'un même mot de  $X$  ne commutent pas et d'autre part celui de [2] pour lequel  $X$  est un ensemble fini de mots tel que chacun contienne au moins une occurrence de toute lettre de  $A$  et vérifie l'hypothèse de connexité du graphe des non-commutations.

La preuve du résultat présentée ici suit très largement celle de [1].

## 1. GÉNÉRALITÉS SUR LES MONOÏDES PARTIELLEMENT COMMUTATIFS LIBRES

Les notations sont celles de [1]; ainsi :  $\varepsilon$  désigne le mot vide.

Soit  $B \subset A$ , on définit le morphisme projection sur  $B$ , noté  $\Pi_B$  par :

$$\begin{aligned} \forall b \in B, \quad \Pi_B(b) &= b, \\ \forall a \in A \setminus B, \quad \Pi_B(a) &= \varepsilon. \end{aligned}$$

Une relation de commutation partielle  $\simeq$  sur  $A^*$  est une congruence engendrée par une relation symétrique  $\Theta$  sur  $A$ . Plus précisément : soit  $\Theta$  une relation symétrique sur  $A$ , on note  $ab \simeq ba$  si  $(a, b) \in \Theta$ ; on définit pour deux mots  $u$  et  $v$  de  $A^*$   $u \simeq v$  s'il existe des mots  $w_1, w_2, \dots, w_k$  tels que :

$$w_1 = u, \quad w_k = v$$

et pour  $i = 1, \dots, k$  :

$$w_i = w'_i abz'_i, \quad w_{i+1} = w'_i baz'_i \quad \text{avec } (a, b) \in \Theta.$$

Soit  $\simeq$  une relation de commutation partielle, comme d'habitude on note :

$$\begin{aligned} \forall x \in A^*, \quad [x] &= \{y \in A^* / x \simeq y\}, \\ \forall L \subset A^*, \quad [L] &= \{y \in A^* / \exists x \in L \text{ et } y \simeq x\}. \end{aligned}$$

Le monoïde quotient  $A^*/\simeq$  sera noté  $M(A, \Theta)$ .

Étant donné une relation symétrique  $\Theta$  sur  $A$  et un sous-ensemble  $B$  de  $A$ , on définit le graphe des non-commutations associé à  $B$  et  $\Theta$  noté  $(B, \Theta)$  par :

- les sommets sont les lettres de  $B$ ;
- les arêtes sont les paires de lettres qui ne commutent pas  $((a, b) \notin \Theta)$ .

On dit que deux mots  $u$  et  $v$  commutent absolument et on note  $u \Gamma v$  si pour tout  $a$  de  $\text{alph}(u)$  et pour tout  $b$  de  $\text{alph}(v)$  on a  $ab \simeq ba$ .

**(b) Rappels de résultats généraux**

Les preuves des résultats que nous allons énoncer se trouvent dans [1], [2]. Ces propositions décrivent les mécanismes de commutation partielle.

**PROPOSITION 1.1 :** Soient  $u$  et  $v$  deux mots de  $A^*$ ,  $u \simeq v$  si et seulement si :

- (i)  $\forall a \in A, |u|_a = |v|_a$ ;
- (ii)  $\forall (a, b) \notin \Theta, \Pi_{a,b}(u) = \Pi_{a,b}(v)$ .

**PROPOSITION 1.2 :** Soient  $x, y, z, t$  des mots de  $A^*$ ,  $xy \simeq zt$  si et seulement si  $x \simeq uv, y \simeq rs, z \simeq ur$  et  $t \simeq vs$  pour des mots  $u, v, r, s$  de  $A^*$  tels que  $v \Gamma r$ .

**COROLLAIRE 1.3 :** Soient  $u, v, x_1, x_2, \dots, x_n$  des mots de  $A^*$ ,  $uv \simeq x_1 x_2 \dots x_n$  si et seulement si  $u \simeq p_1 p_2 \dots p_n, v \simeq q_1 q_2 \dots q_n$  avec  $p_i q_i = x_i$  et  $q_i \Gamma (p_{i+1} \dots p_n)$  ( $1 \leq i < n$ ).

**2. UNE CONDITION SUFFISANTE DE RECONNAISSABILITÉ DANS  $M(A, \Theta)$**

Avant d'établir le théorème principal, nous montrons deux lemmes techniques.

**LEMME 2.1 :** Soient  $\Theta$  une relation symétrique sur  $A$  définissant la relation de commutation partielle  $\simeq$ , soit  $u$  un mot de  $A^*$  dont le graphe  $(\text{alph}(u), \bar{\Theta})$  est connexe. Si  $u \Gamma v$  alors ou bien  $u$  et  $v$  sont puissance d'une même lettre ou bien  $\text{alph}(u) \neq \text{alph}(v)$ . Si de plus  $(\text{alph}(uv), \bar{\Theta})$  est connexe alors  $u$  et  $v$  sont puissance d'une même lettre.

*Preuve :* Supposons que :

- $u$  et  $v$  ne sont pas puissance d'une même lettre;
- $\text{alph}(u) \cap \text{alph}(v) \neq \emptyset$ ;
- $|\text{alph}(u)| > 1$ .

Sachant que  $(\text{alph}(u), \bar{\Theta})$  est connexe, on en déduit qu'il existe  $a$  et  $b$  deux éléments de  $\text{alph}(u)$  tels que  $(a, b) \notin \Theta$  or par hypothèse  $u \Gamma v$  donc au moins une de ces deux lettres ne figure pas dans  $v$  et finalement  $\text{alph}(u) \neq \text{alph}(v)$ .

Si  $(\text{alph}(uv), \bar{\Theta})$  est connexe et  $u \Gamma v$  alors :  $\forall a \in \text{alph}(u) \forall b \in \text{alph}(v) ab \simeq ba$  donc nécessairement  $a = b$ , d'où le résultat.

Le lemme suivant généralise le lemme 2.3 de [1].

**LEMME 2.2 :** Soit  $\simeq$  la relation de commutation partielle engendrée par  $\Theta$ ; soit  $X$  un sous-ensemble de  $A^*$  vérifiant :

- (1)  $[X] = X$ ;
- (2)  $\forall x \in X (\text{alph}(x), \bar{\Theta})$  connexe.

Pour  $u$  et  $v$  deux mots de  $A^*$ ,  $uv \in [X^*]$  si et seulement si :

$$\exists n \geq 0, \quad y_i, z_i \in X^* \quad (0 \leq i \leq n) \quad \text{et} \quad u_i, v_i \in A^+ \quad (1 \leq i \leq n),$$

avec :

- (i)  $u \simeq y_0 u_1 y_1 u_2 \dots u_n y_n$ ;
- (ii)  $v \simeq z_0 v_1 z_1 v_2 \dots v_n z_n$ ;
- (iii)  $u_i v_i \in X$ ;
- (iv)  $z_i \Gamma(u_{i+1} y_{i+1} \dots u_n y_n)$  ( $0 \leq i < n$ );
- (v)  $v_i \Gamma(y_i u_{i+1} \dots u_n y_n)$  ( $1 \leq i < n$ );
- (vi)  $\text{alph}(u_i) \neq \text{alph}(u_j)$  ( $1 \leq i < j \leq n$ ).

*Preuve* : La condition est clairement suffisante; pour montrer qu'elle est nécessaire, on fait une preuve par récurrence sur le plus petit entier  $k$  pour lequel  $uv \in [X^k]$ .

Si  $uv \in [X^k]$  alors par définition  $uv \simeq z$  avec  $z \in X^k$ . Soient  $x$  et  $y$  deux mots tels que :  $x \in X$ ,  $y \in X^{k-1}$  et  $xy = z$ , on a  $uv \simeq xy$ .

D'après la proposition 1.2 :

$$u \simeq pp', \quad v \simeq qq', \quad x \simeq pq \quad \text{et} \quad y \simeq p'q' \quad \text{avec} \quad p' \Gamma q.$$

On peut appliquer l'hypothèse de récurrence à  $p'q'$ , d'où :

- (i)  $p' \simeq xy_1 u_2 y_2 u_2 \dots u_n y_n$ ;
- (ii)  $q' \simeq xz_1 v_2 z_2 v_2 \dots v_n z_n$ ;
- (iii)  $u_i v_i \in X$  ( $2 \leq i \leq n$ );
- (iv)  $z_i \Gamma(u_{i+1} \dots u_n y_n)$  ( $1 \leq i < n$ );
- (v)  $v_i \Gamma(y_i u_{i+1} \dots u_n y_n)$  ( $2 \leq i < n$ );
- (vi)  $\text{alph}(u_i) \neq \text{alph}(u_j)$  ( $2 \leq j < j \leq n$ );

De plus : ( $[X] = X$  et  $x \simeq pq$ )  $\Rightarrow pq \in X$ .

On envisage trois cas :

(a)  $p = \varepsilon$  :

On en déduit que :  $q \in X$ , donc  $qz_1 \in X^*$ .

De plus  $q \Gamma p'$  (i. e. :  $u_2 \dots u_n y_n$ ) et  $z_1 \Gamma(u_2 \dots u_n y_n)$  donc  $qz_1 \Gamma(u_2 \dots u_n y_n)$ . Si l'on pose  $z'_1 = qz_1$ , on obtient une décomposition de  $uv$  qui est de la forme décrite par le lemme ( $z'_1$  étant substitué à  $z_1$ ).

(b)  $q = \varepsilon$  :

Dans ce cas il suffit de poser  $y'_1 = py_1$  pour retrouver une factorisation de la forme (i)-(vi).

(c)  $p \neq \varepsilon$  et  $q \neq \varepsilon$  :

– Si  $\text{alph}(p) \neq \text{alph}(u_i)$  ( $\forall i, 2 \leq i \leq n$ ) il suffit alors de poser  $u_1 = p, v_1 = q, y_0 = z_0 = \varepsilon$  pour se retrouver dans les conditions du lemme.

– Si  $\exists k$  tel  $\text{alph}(p) = \text{alph}(u_k)$ . Remarquons tout d'abord que l'hypothèse de récurrence implique l'unicité de  $k$ .

De plus  $q \Gamma (y_1 u_2 u_2 0 \dots u_n y_n) \Rightarrow q \Gamma u_k$ .

Or  $\text{alph}(p) = \text{alph}(u_k)$  donc  $q \Gamma p$ , de plus  $pq \in X \Rightarrow (\text{alph}(pq), \bar{\Theta})$  est connexe ainsi  $p, q$  et  $u_k$  sont puissance d'une même lettre (lemme 2. 1) :

$$\text{alph}(p) = \text{alph}(q) = \text{alph}(u_k) = \{ a \}.$$

On va montrer que l'on peut réorganiser  $u$  et  $v$  pour retrouver la forme décrite dans le lemme.

On sait que :

$$a \Gamma (y_1 u_2 y_2 \dots u_n y_n) \quad (q \Gamma p'),$$

$$z_1 \Gamma u_k, \dots, z_{k-1} \Gamma u_k \quad (\text{iv}),$$

$$v_1 \Gamma u_k, \dots, v_{k-1} \Gamma u_k \quad (\text{v}).$$

Donc  $a \Gamma (z_1 v_2 z_2 \dots v_{k-1} z_{k-1})$ .

– Si  $|u_k| = |q|$

alors

$u \simeq pu_k y_1 u_2 y_2 \dots u_n y_n, v \simeq z_1 v_2 z_2 \dots z_{k-1} qv_k \dots v_n z_n$  avec  $pu_k \in X$  et  $qv_k \in X$ ; en posant  $y'_1 = pu_k y_1$  et  $z'_{k-1} = z_{k-1} qv_k$  on obtient une décomposition de  $u$  et  $v$  vérifiant (i)-(vi).

– Si  $|u_k| < |q|$  alors :

$$u \simeq pu_k y_1 u_2 y_2 \dots y_{k-1} y_k \dots u_n y_n,$$

$$v \simeq lz_1 v_2 z_2 \dots z_{k-1} l' v_k z_k \dots v_n z_n,$$

avec  $l' = u_k$  et  $ll' = q$ .

En posant  $u_1 = pu_k, v_1 = 1$  et sachant que  $y_{k-1} y_k \in X^\dagger$  ainsi que  $z_{k-1} l' v_k z_k$ , on a obtenu décomposition de  $u$  et  $v$  vérifiant (i)-(vi).

– Si  $|u_k| > |q|$  alors :

$$u \simeq pl y_1 u_2 y_2 \dots y_{k-1} l' y_k \dots u_n y_n,$$

$$v \simeq z_1 v_2 z_2 \dots z_{k-1} qv_k z_k \dots v_n z_n,$$

avec  $l = q$  et  $ll' = u_k$ .

En posant  $y'_1 = pl y_1, u_k = l', v'_k = qv_k$  on obtient le résultat.

Nous pouvons énoncer maintenant le théorème dont la preuve est identique au théorème 2.2 de [1].

THÉORÈME 2.3 : Si  $X$  est une partie reconnaissable de  $A^*$  vérifiant :

- (i)  $[X] = x$ ;
  - (ii)  $\forall x \in X$ ,  $(\text{alph}(x), \bar{\Theta})$  connexe,
- alors  $[X^*]$  est reconnaissable.

Preuve : Pour chaque mot  $u$  de  $A^*$ , on note  $\sigma(u)$  l'ensemble constitué des 3  $n$ -uplets :

$$(\text{alph}(u_1), u_1^{-1}X, \text{alph}(y_1), \text{alph}(u_2), \\ u_2^{-1}X, \text{alph}(y_2), \dots, \text{alph}(u_n), u_n^{-1}X, \text{alph}(y_n)),$$

avec  $(u_i, y_i)$  ( $1 \leq i \leq n$ ) tels que :

- (i)  $u \simeq y_0 u_1 y_1 u_2 \dots u_n y_n$ ,
- (ii)  $y_i \in X^*$  ( $0 \leq i \leq n$ );
- (iii)  $u_i \in X^*$  et  $\text{alph}(u_i) \neq \text{alph}(u_j)$  ( $1 \leq i < j \leq n$ ).

Il n'y a qu'un nombre fini de  $\sigma(u)$  possibles. En effet :

- la condition (iii), entraîne que  $n \leq 2^{\text{Card}(A)}$ ;
- $X$  étant reconnaissable il n'y a qu'un nombre fini de  $u_i^{-1}X$ ;
- l'alphabet  $A$  est fini donc le nombre des  $\text{alph}(y_i)$  est fini.

Pour prouver que  $Y = [X^*]$  est reconnaissable, on montre que la famille  $\{u^{-1}Y/u \in A^*\}$  est fini; pour cela il suffit que :

$$\forall u, t \in A^*, \quad \sigma(u) = \sigma(t) \Rightarrow u^{-1}Y = t^{-1}Y.$$

Soit  $uv \in Y$ . D'après le lemme 2.2 on a :

$$u \simeq y_0 u_1 u_2 \dots u_n y_n \quad \text{et} \quad v \simeq z_0 v_1 y_1 \dots v_n y_n$$

les mots  $u_i, y_i, z_i$  et  $v_i$  satisfaisant les conditions du lemme 2.2. Si  $\sigma(u) = \sigma(t)$  alors  $t \simeq x_0 t_1 x_1 \dots t_n x_n$  avec :

- (1)  $x_i \in X^*$  ( $0 \leq i \leq n$ );
- (2)  $\text{alph}(x_i) = \text{alph}(y_i)$  ( $0 \leq i \leq n$ );
- (3)  $t_i^{-1}X = u_i^{-1}X$  ( $1 \leq i \leq n$ );
- (4)  $\text{alph}(t_i) = \text{alph}(u_i)$  ( $1 \leq i \leq n$ ).

Sachant que  $u_i v_i \in X$  et  $t_i^{-1}X = u_i^{-1}X$  on en déduit que  $t_i v_i \in X$ . De plus  $z_i \Gamma(u_{i+1} y_{i+1} \dots u_n y_n)$ ,  $\text{alph}(u_{i+1} y_{i+1} \dots u_n y_n) = \text{alph}(t_{i+1} x_{i+1} \dots t_n x_n)$  donc  $z_i \Gamma(t_{i+1} x_{i+1} \dots t_n x_n)$ ; de même  $v_i \Gamma(x_i t_{i+1} \dots t_n x_n)$ . Ainsi  $tv \in Y$  ce qui achève la preuve du théorème.

*Remarques* : Posons  $\text{Card}(A) = |A|$ .

Le calcul de normes concernant le nombre d'états des automates obtenus dans [1], cette note et [2] donne :

(a) pour [1] :

Les états de l'automate que l'on déduit de la preuve du théorème sont tous les 3-nuplets qui sont définis par :

$$(\text{alph}(u_i), u_i^{-1} X, \text{alph}(y_1), \text{alph}(u_2), \\ u_2^{-1} X, \text{alph}(y_2) \dots \text{alph}(u_n), u_n^{-1} X, \text{alph}(y_n))$$

avec :

- (i)  $u \simeq y_0 u_1 u_2 \dots u_n y_n$ ;
- (ii)  $y_i \in X^+$ ;
- (iii)  $u_i \in A^+$  et  $\text{alph}(u_i) \cap \text{alph}(u_j) = \emptyset$  ( $1 \leq i < j \leq n$ ).

La condition (iii) implique  $n \leq \text{Card}(A)$ .

Si  $C$  désigne le nombre d'états de l'automate minimal reconnaissant  $X$ , on a  $C$  valeurs possibles pour  $u_i^{-1} X$ . Pour  $\text{alph}(u_i)$  ou  $\text{alph}(y_i)$ , on a au plus  $2^{|A|}$  valeurs possibles. Donc on a au plus :

$$2^{2^{|A|}^2} \cdot C^{|A|} \text{ états différents.}$$

(b) Pour cette Note, la seule différence porte sur la borne de  $n$  qui devient  $2^{|A|}$ , on obtient alors :

$$C^{2^{|A|}} 2^{|A| (2^{|A|+1})}$$

(c) Enfin dans [1], pour montrer que  $L = [w^*]$  est reconnaissable, on prouve :

Pour tout mot  $u$  de  $A^*$  il existe un mot  $s$  avec  $|s| \leq |w| |A|$  tel que :  $u^{-1} L = s^{-1} L$ .

Donc le nombre d'états de l'automate reconnaissant  $L$  est bornée par :

$$|A|^{|A| (|w|+1)}.$$

#### BIBLIOGRAPHIE

1. R. CORI et D. PERRIN, *Sur la reconnaissabilité dans les monoïdes partiellement commutatifs libres*, R.A.I.R.O., Informat. théor. (à paraître).
2. R. CORI et Y. MÉTIVIER, *Rational Subsets of Some Partially Abelian Monoids*, Theoret. Comput. Sc. (à paraître).
3. M. P. FLÉ et G. ROUCAIROL, *Maximal Serializability of Iterated Transactions*, Theoret. Comput. Sc. (à paraître) (voir aussi ACM SIGACT SIGOPS, 1982, p. 194-200).
4. M. LOTHAIRE, *Combinatorics on Words*, Addison Wesley, 1983.
5. J.-E. PIN, *Variétés de langages formels*, Masson, Paris, 1984.