Sur l'étoile des connaissances de Nk

Anne Grazon, Gilles Lesventes

UNITÉ DE RECHERCHE
INRIA-RENNES

Institut National
de Recherche
en Informatique
et en Automatique

Domaine de Voluceau
Rocquencourt
B.P.105
78153 Le Chesnay Cedex
France
Tél.:(1) 39 63 55 11

Rapports de Recherche

N° 786

SUR L'ETOILE DES RECONNAISSABLES DE N^k

Anne GRAZON
Gilles LESVENTES

FEVRIER 1988

IRISA

INSTITUT DE RECHERCHE EN INFORMATIQUE ET SYSTÈMES ALÉATOIRES

Campus Universitaire de Beaulieu
35042 - RENNES CÉDEX
FRANCE
Téléphone : 99 36 20 00
Télex : UNIRISA 950 473 F
Télécopie : 99 38 38 32

Sur l'Etoile des Reconnaissables de N^k

Anne GRAZON (+)
Gilles LESVENTES (++)

Publication Interne n° 386 - 12 pages - Décembre 87

Résumé

Dans un monoïde quelconque M, l'ensemble Rec(M) des parties reconnaissables de M est fermé par union, intersection et complémentation mais pas par produit et opération étoile.

Pour tout entier non négatif k, Rec(N^k) est fermé par produit.

Nous donnons une condition de décision C telle que, Pour tout partie reconnaissable L de N^k,

si $k = 2$ $\quad L^* \in \mathrm{Rec}(N^k) \Leftrightarrow$ C est vraie

si $k \geq 3$ $\quad L^* \in \mathrm{Rec}(N^k) \Rightarrow$ C est vraie.

Star product of Recognizable sets of N^k

Abstract

For an arbitrary monoid M, the set Rec(M) of recognizable subsets of M is closed under union, intersection and complementation but not under product and star operation.

For any non negative integer k, Rec(N^k) is closed by product.

We give a decidable condition C such that, for any recognizable set L of N^k,

if $k = 2$ $\quad L^* \in \mathrm{Rec}(N^k) \Leftrightarrow$ C is true

if $k \geq 3$ $\quad L^* \in \mathrm{Rec}(N^k) \Rightarrow$ C is true.

(+) Faculté des Sciences, 33 rue Pasteur, 84000 AVIGNON
(++) IRISA, Campus de Beaulieu, 35042 RENNES CEDEX

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (L. A. 227)
UNIVERSITÉ DE RENNES 1 — I. N. S. A. DE RENNES

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE (LABORATOIRE DE RENNES)

Introduction

Dans le cadre de l'étude de l'ensemble des états accessibles d'un système d'addition de vecteurs, nous nous intéressons aux parties rationnelles de N^k, et en particulier à la reconnaissabilité de l'étoile des parties reconnaissables de N^k.

Le monoïde N^k a été souvent étudié, notamment par S.Eilenberg et M.P.Schutzenberger [ES] où il est prouvé que tout rationnel ambigu de N^k peut être écrit sous forme inambiguë. Ph.Gohon [G] décide de la reconnaissabilité d'un rationnel de N^k. Il fournit un algorithme utilisant les séries rationnelles et basé sur la transformation de [ES], qui peut être rendue effective.

Nous nous restreignons au calcul de l'étoile d'un reconnaissable de N^k et donnons une condition nécessaire de reconnaissabilité. Celle-ci s'avère suffisante pour $k \leq 2$.

Définitions

Soit **N** le monoïde $(\{1\}^*, ., 0)$ dont la loi binaire représentant la concaténation ou par abus l'addition entière,est commutative.

Soit , pour tout entier k positif , N^k le monoïde $((\{1\}^*)^k, ., 0)$, monoïde produit de k fois N,d'élément neutre le k-uplet **O** tel que :

$\forall\ i \in [k]\ , O(i) = 0$.

Soit le vecteur unité e_i pour tout $i \in [k]$, tel que :

$\forall j \in [k]-\{i\} : e_i(j)=0$ et $e_i(i)=1$.

Ces vecteurs unités sont aussi appelés lettres .

Notations préliminaires

La concaténation sur N^k se fait composante à composante; elle est indifféremment notée . ou $\prod_{i \in [n]}$, suivant qu'elle est binaire ou n-aire .

Par analogie , le produit de langages de N^k est noté . ou $\prod_{i \in [n]}$.

Le produit cartésien de langages du monoïde N sera noté $\times_{i \in [n]}$.

Rappels

Théorème de Mezei ([B])

Soient M_1 et M_2 deux monoïdes et $M = M_1 \times M_2$.

$B \in Rec(M)$ si et seulement si B est une union finie d'ensembles de la forme $A_1 \times A_2$, avec $A_1 \in Rec(M_1)$ et $A_2 \in Rec(M_2)$.

Proposition ([B])

Soit M un monoïde, alors Rec(M) est clos par union,intersection et complémentation finies :

$\forall L_1, L_2 \in Rec(M)$,

$L_1 \cup L_2 \in Rec(M)$; $L_1 \cap L_2 \in Rec(M)$; $M\text{-}L_1 \in Rec(M)$.

Proposition ([B])

Soient X et Y deux alphabets :

$Rec(X^* \times Y^*) \subsetneq Rat(X^* \times Y^*)$.

Proposition

$\forall A \in \mathbf{N}^k$, $A \in Rec(\mathbf{N}^k)$ si et seulement si A est une union finie d'ensembles de la forme $A_1 \times A_2 \times \ldots \times A_k$ avec $A_i \in Rec(\mathbf{N})$ pour $1 \leq i \leq k$.

(par récurrence sur k en appliquant le théorème de Mezei.)

Proposition

$\forall A,B \in Rec(\mathbf{N}^k)$, $A.B \in Rec(\mathbf{N}^k)$

(d'après la proposition ci-dessus et par distributivité du produit par rapport à l'union.)

Conséquence

$Rec(\mathbf{N}^k)$ est fermé par union, produit, intersection et complémentation mais pas par étoile.

Ecriture rationnelle des parties reconnaissables de $\mathbf{N}^k$

Proposition1

$$Rec(\mathbf{N}^k) = \{ \bigcup_{j \in [n]} (\prod_{i \in [k]} e_i^{p_{i,j}} . (e_i^{q_{i,j}})^* \ / \ n,k \in \mathbf{N} , \quad \forall i, \forall j , \ p_{i,j} , q_{i,j} \in \mathbf{N} \} .$$

Preuve

sens ($\supseteq$) :

Soit $L = \bigcup_{j \in [n]} L_j$, où L_j est de la forme $L_j = \prod_{i \in [k]} e_i^{p_{i,j}} . (e_i^{q_{i,j}})^*$.

$Rec(N^k)$ étant clos par union finie, il suffit de montrer que chaque L_j appartient à $Rec(N^k)$.

On a , $\forall j \in [n]$: $L_j = \times_{i \in [k]} L_j(i)$ avec , $\forall i \in [k]$, $L_j(i) = 1^{p_{i,j}} . (1^{q_{i,j}})^* \in Rat(N)$

d'où $L_j(i) \in Rec(N)$.

D'après le théorème de Mezei , comme $N^k = \times_{i \in [k]} N$, on a bien $L_j \in Rec(N^k)$.

sens ($\subseteq$) :

D'après le théorème de Mezei , $\forall L \in Rec(N^k)$,il existe $n \in N$ tel que :

$L = \bigcup_{j \in [n]} \times_{i \in [k]} L_{i,j}$, où $L_{i,j} \in Rec(N)$.

Tout langage rationnel sur une lettre étant un semi-linéaire([B]) $L_{i,j}$ peut s'écrire

$L_{i,j} = \bigcup_{h \in E_{i,j}} (1^{p_{i,j,h}} . (1^{q_{i,j,h}})^*)$ où $E_{i,j}$ est un ensemble fini et $p_{i,j,h}$, $q_{i,j,h} \in N$.

Cette décomposition est effective. ([P])

Donc $L = \bigcup_{j \in [n]} \times_{i \in [k]} \bigcup_{h \in E_{ij}} (1^{p_{i,j,h}} . (1^{q_{i,j,h}})^*)$.

A un renommage près des exposants $p_{i,j,h}$ et $q_{i,j,h}$, on obtient

$L = \bigcup_{j \in E} \times_{i \in [k]} (1^{\alpha_{i,j}} . (1^{\beta_{i,j}})^*)$, où E est un ensemble fini.

Donc $L = \bigcup_{j \in E} \prod_{i \in [k]} (e_i^{\alpha_{i,j}} . (e_i^{\beta_{i,j}})^*)$, ce qui achève la preuve .

Etoile d'un reconnaissable de N^k

Notations

Soit L un reconnaissable de N^k .

Posons $L = \bigcup_{j \in [n]} L_j$ et , pour tout $j \in [n]$, $L_j = \prod_{i \in [k]} e_i^{p_{i,j}} . (e_i^{q_{i,j}})^*$

ou bien $L_j = (\prod_{i \in [k]} e_i^{p_{i,j}}) . (\prod_{i \in [k]} (e_i^{q_{i,j}})^*)$,

Posons également $P_j = \{ i \in [k] / p_{i,j} \neq 0 \}$ et $Q_j = \{ i \in [k] / q_{i,j} \neq 0 \}$.

Alors $L^* = \prod_{j \in [n]} L_j^* = \prod_{j \in [n]} (0 \cup L_j^+)$

$= \prod_{j \in [n]} \Big(0 \cup ((\prod_{i \in P_j} e_i^{p_{i,j}})^+ . \prod_{i \in Q_j} (e_i^{q_{i,j}})^*)\Big)$.

La dernière égalité découle de $(AB^*)^+ = A^+ B^*$, dans un monoïde commutatif.
En développant L^* on obtient :
$L^* = \mathbf{O} \cup L_1^+ \cup L_2^+ \cup ... L_n^+ \cup L_1^+ L_2^+ \cup ... \cup L_{n-1}^+ L_n^+ \cup ... \cup L_1^+ L_2^+ ... L_n^+$,
avec, $\forall j \in [n]$, $L_j^+ = \prod_{i \in P_j} (e_i^{p_{i,j}})^+ . \prod_{i \in Q_j} (e_i^{q_{i,j}})^*$

On utilisera une notation supplémentaire :
pour tout sous-ensemble fini I de [n], soit $\boldsymbol{L}_I = \prod_{j \in I} L_j^+$,
si bien que $L^* = \bigcup_{I \subset [n]} \boldsymbol{L}_I$, avec la convention $\boldsymbol{L}_\varnothing = \mathbf{O}$.

On a alors $\boldsymbol{L}_I = c_I . M_I . S_I$,
où $\quad c_I = \prod_{i \in [k]} e_i^{(\sum_{j \in I} p_{i,j})}$, ce qui représente la partie constante de $\boldsymbol{L}_I$;

$M_I = \prod_{j \in I / \|P_j\| \geq 2} (\prod_{i \in P_j} e_i^{p_{i,j}})^*$, qui représente les étoiles multiples de $\boldsymbol{L}_I$,
i.e. portant sur au moins deux lettres distinctes,

$S_I = \prod_{j \in I} (\prod_{i \in Q_j} (e_i^{q_{i,j}})^*) . \prod_{j \in I / P_j = \{i\}} (e_i^{p_{i,j}})^*$, qui représente les étoiles simples de $\boldsymbol{L}_I$ i.e. portant sur une seule lettre.
c_I et S_I sont reconnaissables.
Pour $L \subset N^k$, soit $E(L) = \{ i \in [k] \ / L \cap N^k e_i N^k \neq \varnothing \}$, qui représente les indices des lettres composant les mots de L.

<u>Proposition 2</u>

Soit $L = \bigcup_{j \in [n]} L_j$ et , pour tout $j \in [n]$, $\quad L_j = \prod_{i \in [k]} e_i^{p_{i,j}} . (e_i^{q_{i,j}})^*$.
Pour tout $I \subset [k]$,
$\boldsymbol{L}_I = \prod_{i \in I} L_i^+ \in Rec(N^k)$ si et seulement si $E(M_I) \subseteq E(S_I)$.

<u>Preuve</u>

<u>sens $\Leftarrow$</u>

Supposons $E(M_I) \subseteq E(S_I)$
Renommons les exposants des lettres de S_I de façon à distinguer d'une part ceux de M_I, d'autre part ceux de S_I :
Pour tout $i \in E(S_I)$, soit $s_i = \| \{ j \in I / P_j = \{i\} \} \| + \| \{ j \in I / i \in Q_j \} \|$
Soient $r_{i,1}, r_{i,2}, ... r_{i,s_i}$ les différents exposants de e_i dans S_I , où $r_{i,1}$ est le plus petit d'entre eux.
Alors $S_I = \prod_{i \in E(S_I)} (\prod_{j \in [s_i]} (e_i^{r_{i,j}})^*)$.
Posons $\alpha = ppcm(\{ r_{i,1} / i \in E(M_I) \})$.

Construisons le sous-ensemble fini F_I de M_I où chaque facteur a été itéré au maximum α fois .

$$F_I = \{ \prod_{j \in I / \|P_j\| \geq 2} (\prod_{i \in P_j} e_i^{p_{i,j}})^{\alpha_j} \ / \ \forall h \in I, \|P_h\| \geq 2 \ , 0 \leq \alpha_h < \alpha \}$$

Nous allons montrer que $\mathcal{L}_I = c_I . M_I . S_I$ est égal à $c_I . F_I . S_I$, produit de 3 reconnaissables . Ceci achèvera la preuve.

1. $F_I . S_I \subseteq M_I . S_I$: évident.

2. Soit $\omega \in M_I . S_I$:

$$\omega = \prod_{j \in I / \|P_j\| \geq 2} (\prod_{i \in P_j} e_i^{p_{i,j}})^{\beta_j} . \prod_{i \in E(S_I)} (\prod_{j \in [s_i]} (e_i^{r_{i,j}})^{\delta_{i,j}})$$

Soit $\beta'_j = (\beta_j \bmod \alpha)$, $\beta'_j < \alpha$

ainsi $\beta_j = \beta'_j . \alpha + (\beta_j \bmod \alpha)$,

soit $\delta'_{i,j} = \delta_{i,j}$ si $j \neq 1$ ou $i \notin E(M_I)$

$\delta'_{i,1} = \delta_{i,1} + (\sum_{h \in I / \|P_h\| \geq 2 \text{ et } i \in P_h} (p_{i,h} . (\beta_h \text{ div } \alpha).\alpha)) / r_{i,1}$

ω peut donc s'écrire :

$$\omega = \prod_{j \in I / \|P_j\| \geq 2} (\prod_{i \in P_j} e_i^{p_{i,j}})^{\beta'_j} . \prod_{i \in E(S_I)} (\prod_{j \in [s_i]} (e_i^{r_{i,j}})^{\delta'_{i,j}})$$

d'où $\omega \in F_I . S_I$.

<u>sens $\Rightarrow$</u>

Supposons que $E(M_I) \not\subseteq E(S_I)$ et , sans perte de généralité , que $e_1 \in E(M_I) \setminus E(S_I)$. Choisissons ,parmi les facteurs de M_I contenant e_1^z pour un entier z non nul,un facteur contenant un minimum de lettres e_i distinctes apparaissant à une puissance non nulle . A un renommage près des lettres,ce facteur est $(e_1^{\alpha_1} e_2^{\alpha_2} ... e_r^{\alpha_r})^*$, pour un entier $r \geq 2$, avec $\forall i \in [r]$, $\alpha_i \neq 0$.

$\mathcal{L}_I$ est reconnaissable ssi $M_I . S_I$ est reconnaissable , montrons que celui-ci ne l'est pas. Soient $A_k = \{a_1,....a_k\}$ et φ l'homomorphisme de A_k^* dans N^k défini par :

$$\varphi : \ A_k^* \rightarrow N^k$$
$$a_i \rightarrow e_i$$

Supposons que $M_I . S_I$ soit reconnaissable. Alors $\varphi^{-1}(M_I . S_I) \in \mathrm{Rec}(A_k^*)$, donc $\varphi^{-1}(M_I . S_I) \in \mathrm{Rat}(A_k^*)$ et $(\varphi^{-1}(M_I . S_I) \cap a_1^* ... a_r^*) \in \mathrm{Rat}(A_k^*)$.

Or $(\varphi^{-1}(M_I . S_I) \cap a_1^* ... a_r^*) = \varphi^{-1}(M_I . S_I \cap e_1^* ... e_r^*) \cap a_1^* ... a_r^*$

et $M_I . S_I \cap e_1^* ... e_r^* = (e_1^{\alpha_1} e_2^{\alpha_2} ... e_r^{\alpha_r})^* . \prod_{j \in [t]} (\prod_{i \in [r]} e_i^{\beta_{i,j}})^*$

$$. \prod_{i=2}^{r} (\prod_{j \in [n_i]} e_i^{\delta_{i,j}})^* ,$$

où

$.\forall\, j \in [t] : \exists\, i,i' \in [r]$, $i \neq i'$ tq $\beta_{i,j} \neq 0$ et $\beta_{i',j} \neq 0$

$.\forall\, j \in [t] : (\beta_{1,j} \neq 0) \Rightarrow (\forall\, i, 2 \leq i \leq r, \beta_{i,j} \neq 0)$

Cette condition est due au choix du facteur de M_I.

$.\forall\, i, 2 \leq i \leq r, \forall\, j \in [n_i]$, $\delta_{i,j} \neq 0$.

Soit $K = \max(\{\alpha_1/\alpha_2\} \cup \{\beta_{1,j}/\beta_{2,j}, j \in [t] \text{ et } \beta_{1,j} \neq 0\})$,
alors pour tout mot w de $M_I.S_I \cap e_1^*...e_r^*$, on a $|w|_{e_1} \leq K.\, |w|_{e_2}$.
Ce même rapport existe entre le nombre de lettres a_1 et le nombre de lettres a_2 des mots de $(\varphi^{-1}(M_I.S_I) \cap a_1^*...a_r^*)$ qui n'est donc pas rationnel.
Par conséquent $M_I.S_I$ n'est pas reconnaissable.

Proposition 3

Soit $L = \bigcup_{j \in [n]} L_j$,

$$\mathcal{L}_{[n]} = \prod_{i \in [n]} L_i^+ \notin \mathrm{Rec}(N^k) \Rightarrow L^* \notin \mathrm{Rec}(N^k)$$

Preuve

$\mathcal{L}_{[n]} = c_{[n]}.M_{[n]}.S_{[n]}$

D'après la proposition 2, si $\mathcal{L}_{[n]}$ n'est pas reconnaissable alors $E(M_{[n]}) \not\subseteq E(S_{[n]})$.
Supposons $e_1 \in E(M_{[n]}) \setminus E(S_{[n]})$.
Choisissons, parmi les facteurs de $M_{[n]}$ contenant e_1^z pour un entier z non nul, un facteur contenant un minimum de lettres e_i distinctes apparaissant à une puissance non nulle.
A un renommage près des lettres, ce facteur est $(e_1^{\alpha_1} e_2^{\alpha_2}...e_r^{\alpha_r})^*$, pour un entier $r \geq 2$, avec $\forall\, i \in [r], \alpha_i \neq 0$.
Reprenons l'homomorphisme φ défini précédemment.
Soit $\Lambda = (\varphi^{-1}(L^*) \cap (\varphi^{-1}(c_{[n]}).\, a_1^*...a_r^*))$, alors si L^* est reconnaissable, Λ est rationnel.
Soit w un mot quelconque non vide de Λ.
Soit $I \subseteq [n]$ tel que $w \in \varphi^{-1}(\mathcal{L}_I)$.
$\mathcal{L}_I$ peut s'écrire sous la forme :

$$\mathcal{L}_I = c_I . \prod_{j \in [t]} \Big(\prod_{i \in [r]} e_i^{\beta_{i,j}} \Big)^* . \prod_{j \in [u]} \Big(\prod_{i=2}^{r} e_i^{\delta_{i,j}} \Big)^* . \prod_{j \in [v]} \Big(\prod_{i \in [k]} e_i^{\mu_{i,j}} \Big)^* . S_I$$

Dans cette écriture, on a factorisé M_I en trois parties :

- les termes faisant apparaître e_1 à une puissance non nulle, et uniquement des e_i tq $i \in [r]$,
- les termes ne faisant pas apparaître e_1 mais uniquement des e_i tq $2 \leq i \leq r$,
- les termes faisant apparaître au moins un e_i, $i \notin [r]$, à une puissance non nulle.

On a par ce choix:

$. \forall j \in [t] \ , \beta_{1,j} \neq 0$ et

$. \forall j \in [v] , \exists i \in [k] \setminus [r] , \mu_{i,j} \neq 0$.

En vertu de l'hypothèse du choix du facteur $(e_1^{\alpha_1} e_2^{\alpha_2} \dots e_r^{\alpha_r})^*$ de $M_{[n]}$, et sachant que tous les facteurs de la forme $(e_{i_1}^{\alpha_{i1}} e_{i_2}^{\alpha_{i2}} \dots e_{i_s}^{\alpha_{is}})^*$, $s \geq 2$, de M_I apparaissent obligatoirement sous la même forme dans $M_{[n]}$, on peut déduire que $\mathcal{L}_I$ ne contient pas de facteur du type $(e_1^{\beta_1} e_2^{\beta_2} \dots e_r^{\beta_r})$ où $\beta_1 \neq 0$ et où l'un des β_i , $2 \leq i \leq r$, serait nul.

D'où $. \forall j \in [t] \ , \forall i \in [r] , \beta_{i,j} \neq 0$.

Soit f un mot de $\mathcal{L}_I$ tel que $\varphi(f)=w$,

$$f= c_I . \prod_{j \in [t]} (\prod_{i \in [r]} e_i^{\beta_{i,j}})^{b_{i,j}} . \prod_{j \in [u]} (\prod_{i=2}^{r} e_i^{\delta_{i,j}})^{d_{i,j}} . \prod_{j \in [v]} (\prod_{i \in [k]} e_i^{\mu_{i,j}})^{m_{i,j}} .h,$$

avec $h \in S_I$.

$w \in \Lambda$ donc $w \in \varphi^{-1}(c_{[n]}) . a_1^* \dots a_r^*$

Pour tout indice $j \geq r+1$, $|w|_{a_j} = |c_{[n]}|_{e_j}$

donc $|c_{[n]}|_{e_j} = |c_I|_{e_j} + \sum_{l \in [v]} \mu_{j,l} . m_{j,l} + |h|_{e_j}$.

On constate donc que tous les nombres $m_{i,j}$ sont bornés par une constante Cte qui ne dépend que du mot $c_{[n]}$.

Par conséquent , on a l'inégalité suivante :

$|w|_{a_1} - |c_I|_{e_1} - Cte. \sum_{l \in [v]} \mu_{1,l} \leq (\max(\{\beta_{1,j} / \beta_{2,j} , j \in [t]\})) . |w|_{a_2}) - |c_I|_{e_2}$.

On peut aisément en déduire que le nombre de lettres a_1 de tout mot de Λ peut être arbitrairement grand mais est linéairement borné par le nombre de lettres a_2 de ce même mot .

Donc Λ n'est pas rationnel .

Proposition 4

Soit $L \in Rec(N^2)$, $L= \bigcup_{j \in [n]} L_j$ et , pour tout $j \in [n]$,

$L_j = e_1^{p_{1,j}} . e_2^{p_{2,j}} . (e_1^{q_{1,j}})^* . (e_2^{q_{2,j}})^*$,

L^* est reconnaissable ssi $\mathcal{L}_{[n]} = \prod_{j \in [n]} L_j^+$ est reconnaissable .

Preuve

sens $\Rightarrow$

Si $\prod_{j \in [n]} L_j^+$ n'est pas reconnaissable, L^* ne l'est pas non plus d'après la proposition 3 sens $\Leftarrow$

$L^* = O \cup L_1^+ \cup L_2^+ \cup \ldots L_n^+ \cup L_1^+ L_2^+ \cup \ldots \cup L_{n-1}^+ L_n^+ \cup \ldots \cup L_1^+ L_2^+ \ldots L_n^+$.

Si tous les termes de cette union sont reconnaissables,alors il est clair que L^* est reconnaissable .

On va supposer maintenant qu'au moins un des termes de cette union n'est pas reconnaissable .

Ceci nous assure que L^* n'est pas écrit sur une seule lettre (e_1 ou e_2) et donc que $\prod_{j \in [n]} L_j^+$ contient au moins un terme de la forme $(e_1^z)^*$, $z > 0$, et au moins un terme de la forme $(e_2^{z'})^*$, $z' > 0$.

On va distinguer par la suite un terme de chacune de ces deux formes :

soit x (resp. y) l'un des $q_{1,j}$ (resp. $q_{2,j}$) non nuls, ou l'un des $p_{1,j}$ non nuls tels que $p_{2,j}$ soit nul . (resp. $p_{2,j}$ et $p_{1,j}$) ,

c'est-à-dire, si $X=\{q_{1,j} / q_{1,j} \neq 0, j \in [n]\} \cup \{p_{1,j} / p_{1,j} \neq 0 \text{ et } p_{2,j}=0, j \in [n]\}$,alors x est un élément quelconque de X

et si $Y=\{q_{2,j} / q_{2,j} \neq 0, j \in [n]\} \cup \{p_{2,j} / p_{2,j} \neq 0 \text{ et } p_{1,j}=0, j \in [n]\}$,alors y est un élément quelconque de Y .

On a finalement :

$$\begin{aligned} \mathcal{L}_{[n]} &= \prod_{j \in [n]} L_j^+ \\ &= \left(e_1^{\sum_{j \in [n]} p_{1,j}} . e_2^{\sum_{j \in [n]} p_{2,j}} \right) . \prod_{j \in [n]} (e_1^{p_{1,j}} . e_2^{p_{2,j}})^* . \prod_{j \in [n]} (e_1^{q_{1,j}})^* \\ &\quad . \prod_{j \in [n]} (e_2^{q_{2,j}})^* . (e_1^x)^* . (e_2^y)^* \end{aligned}$$

La preuve de la condition suffisante découle du lemme suivant :

Lemme

Si $\mathcal{L}_{[n]}$ est reconnaissable alors pour tout sous-ensemble non vide I de [n] , si $\mathcal{L}_I$ n'est pas reconnaissable alors il est l'union de deux langages $\mathcal{L}'_I$ et $\mathcal{L}''_I$ avec $\mathcal{L}'_I$ reconnaissable et $\mathcal{L}''_I \subseteq \mathcal{L}_{[n]}$.

Preuve du lemme

A un renommage près des indices on peut supposer $I = [t]$, $1 \leq t \leq n-1$.

De plus $E(S_I) \neq \{1,2\}$.

Si $\mathcal{L}_I$ n'est pas reconnaissable alors il contient au moins un terme $(e_1^{s} . e_2^{r})^*$, $s.r \neq 0$.

Supposons sans perte de généralité que ce langage ne contienne pas de terme $(e_2^z)^*$, avec $z > 0$.

On a donc $q_{2,j} = 0$, $\forall j \in [t]$.

Posons $p_1=\sum_{j\in[t]} p_{1,j}$; $p_2=\sum_{j\in[t]} p_{2,j}$; $r_1=\sum_{j=t+1}^{n} p_{1,j}$; $r_2=\sum_{j=t+1}^{n} p_{2,j}$.

$$\mathcal{L}_{[t]}= (e_1^{p_1}. e_2^{p_2}). \prod_{j\in[t]} (e_1^{p_{1,j}} . e_2^{p_{2,j}})^*. \prod_{j\in[t]} (e_1^{q_{1,j}})^*$$

$$\mathcal{L}_{[n]}= (e_1^{p_1+r_1}. e_2^{p_2+r_2}). \prod_{j\in[n]} (e_1^{p_{1,j}} . e_2^{p_{2,j}})^*. \prod_{j\in[n]} (e_1^{q_{1,j}})^* . \prod_{j=t+1}^{n} (e_2^{q_{2,j}})^* . (e_1^{x})^*.(e_2^{y})^* .$$

Pour chaque $j\in[t]$ tel que $p_{1,j}\neq 0$ et $p_{2,j}\neq 0$, soit w_j le plus petit entier tel que

$$p_{2,j}.r_1 + x.w_j.p_{2,j} \geq p_{1,j}.r_2 .$$

Soit alors θ un mot quelconque de $\mathcal{L}_{[t]}$:

$$\theta = (e_1^{p_1}. e_2^{p_2}). \prod_{j\in[t]} (e_1^{p_{1,j}} . e_2^{p_{2,j}})^{n_j}. \prod_{j\in[t]} (e_1^{q_{1,j}})^{m_j} .$$

S'il existe $s\in[t]$ tel que $p_{1,s}\neq 0$, $p_{2,s}\neq 0$ et $n_s \geq y.r_1 + x.w_s.y$, alors θ peut se réécrire sous la forme suivante :

$$\begin{aligned}\theta = {} & (e_1^{p_1+r_1}. e_2^{p_2+r_2}). (e_1^{p_{1,1}}. e_2^{p_{2,1}})^{n_1}.(e_1^{p_{1,2}}. e_2^{p_{2,2}})^{n_2}. \dots \\ & . (e_1^{p_{1,s-1}}. e_2^{p_{2,s-1}})^{n_{s-1}} . (e_1^{p_{1,s}}. e_2^{p_{2,s}})^{n_s-y.r_1-x.w_s.y} . (e_1^{p_{1,s+1}}. e_2^{p_{2,s+1}})^{n_{s+1}} . \dots \\ & . (e_1^{p_{1,t}}. e_2^{p_{2,t}})^{n_t} . (e_1^{p_{1,t+1}}. e_2^{p_{2,t+1}})^{(p_{1,s}.y)-1} . \dots . (e_1^{p_{1,k}}. e_2^{p_{2,k}})^{(p_{1,s}.y)-1} \\ & . (e_1^{q_{1,1}})^{m_1} . \dots .(e_1^{q_{1,t}})^{m_t} .(e_1^{q_{1,t+1}})^0 . \dots .(e_1^{q_{1,n}})^0 . (e_2^{q_{2,1}})^0 . \dots .(e_2^{q_{2,n}})^0 \\ & . (e_1^{x})^{p_{1,s}.w_s.y} . (e_2^{y})^{(p_{2,s}.r_1)-(p_{1,s}.r_2)+x.w_s.p_{2,s}} ,\end{aligned}$$

d'où $\theta\in\mathcal{L}_{[n]}$.

Soient alors

$$\mathcal{L}'_{[t]}= (e_1^{p_1}. e_2^{p_2}). \prod_{j\in[t]/p_{1,j}\neq 0 \text{ et } p_{2,j}\neq 0} \{(e_1^{p_{1,j}} . e_2^{p_{2,j}})^{n_j} / n_j < y.r_1 + x.w_j.y \} . \prod_{j\in[t]} (e_1^{q_{1,j}})^* . \prod_{j\in[t]/p_{2,j}=0} (e_1^{p_{1,j}} . e_2^{0})^* . \prod_{j\in[t]/p_{1,j}=0} (e_1^{0} . e_2^{p_{2,j}})^*$$

et $$\mathcal{L}''_{[t]} = \bigcup_{j\in[t]/p_{1,j}\neq 0 \text{ et } p_{2,j}\neq 0} \Big[(e_1^{p_1}. e_2^{p_2}) . \prod_{l\in[t]/l\neq j} (e_1^{p_{1,l}} . e_2^{p_{2,l}})^* .\{(e_1^{p_{1,j}} . e_2^{p_{2,j}})^{n_j} / n_j \geq y.r_1 + x.w_j.y \} . \prod_{l\in[t]} (e_1^{q_{1,l}})^*\Big] .$$

$\mathcal{L}'_{[t]}$ est reconnaissable puisque c'est un produit de langages reconnaissables et de langages finis , $\mathcal{L}''_{[t]}$ est inclus dans $\mathcal{L}_{[n]}$ et $\mathcal{L}_{[t]} = \mathcal{L}'_{[t]} \cup \mathcal{L}''_{[t]}$, d'où le lemme .

Tout terme de l'union L* , non reconnaissable , est donc l'union de deux langages,l'un reconnaissable et l'autre inclus dans $\mathcal{L}_{[n]}$ qui est supposé reconnaissable .
L* est donc une union de langages reconnaissables .

Remarque

Lorsque $k \geq 3$, la condition $\mathcal{L}_{[n]} \in \text{Rec}(N^k)$ n'est pas suffisante pour que L* soit reconnaissable.

On peut le vérifier sur le contre-exemple suivant :

$L_1 = e_1e_3.e_1^*$, $L_2 = e_1e_2.e_2^*e_3^*$ et $L = L_1 \cup L_2$

$L^* = O \cup e_1e_3(e_1e_3)^*e_1^* \cup e_1e_2(e_1e_2)^*e_2^*e_3^* \cup e_1^2e_2e_3(e_1e_3)^*(e_1e_2)^*e_1^*e_2^*e_3^*$,

$\mathcal{L}_{\{1,2\}}$ est reconnaissable .

Soit $K = a_1^*a_3^*$ et φ l'homomorphisme défini plus haut,

alors $\varphi^{-1}(L^*) \cap K = \varepsilon + \{ a_1^{1+n+m}.a_3^{1+n} / n,m \geq 0 \}$, langage non rationnel de A_k^* .

Finalement L* n'est pas reconnaissable.

Références

[B] : J.Berstel , "Transductions and Context-Free Languages" , Teubner 1979.

[P] : R.J.Parikh , "On Context-free Languages" , JACM Vol 13 n°4 p.570-581 1966

[ES] : S.Eilenberg, M.P.Schützenberger , Rational Sets in Commutative Monoïds, Journal of Algebra 13 p.173-191 1969.

[G] : Ph.Gohon , An Algorithm to Decide whether a Rational Subset of N^k is recognizable, Theoritical Computer Science 41 p.51-59 1985.

Imprimé en France

par

l'Institut National de Recherche en Informatique et en Automatique