Lecture Notes in Computer Science

Edited by G. Goos and J. Hartmanis
Series: GI, Gesellschaft für Informatik e.V.

33

Automata Theory and Formal Languages 2nd GI Conference

Kaiserslautern, May 1975

Springer-Verlag
Berlin · Heidelberg · New York

SUR LES RELATIONS RATIONNELLES

M.P. Schutzenberger
IRIA

I. Introduction

Nous faisons référence aux chapitres IX et XI du traité de S. Eilenberg ([1]) pour les résultats de base concernant les relations rationnelles $\rho : A^* \to B^*$ entre monoïdes libres et nous appelons une telle relation *fonctionnelle* ssi l'image $a\rho$ de chaque mot a de A^* est vide (=0) ou un singleton. Nous nous proposons d'établir la propriété suivante qui peut être considérée comme une modification d'un résultat banal concernant les séries rationnelles.

Propriété : Si la relation rationnelle $\rho : A^* \to B^*$ n'est pas une somme finie de relations rationnelles fonctionnelles il existe trois mots $a,a',h \in A^*$ tels que Card $((ah^n a')\rho) \geq n+1$ pour chaque $n \in N$.

Dans ce qui suit nous supposerons d'abord la relation ρ donnée par un transducteur au sens de Nivat ([2]), c'est à dire par un morphisme μ de A^* dans le semi-anneau des QxQ matrices à entrées dans Rat(B), où Q est un ensemble d'indice fini, et la règle que pour chaque mot a, la partie $a\rho$ de B^* est certaine entrée fixe (disons, l'entrée (q_-,q_+)) de la matrice $a\mu$. En outre, on peut supposer que cette représentation est *réduite* en ce sens que pour chaque q de Q il existe des mots a,a' de A^* tels que $a\mu\ (q_-,q)$ et $a'\mu\ (q_-,q_+)$ soient non nuls. En effet s'il n'en était pas ainsi on pourrait, sans changer la valeur de ρ, remplacer par 0 les entrées des lignes et colonnes "q" dans toutes les matrices, et par conséquent, omettre q.

Sous cette hypothèse qui sera toujours faite désormais, il est clair que la propriété serait triviale si l'une des entrées de l'une des matrices génératrices $\alpha\mu$ $(a \in A)$ était une partie infinie du B^*. Nous supposerons donc aussi que toutes ces entrées sont des parties *finies* de B^*. Une autre simplification peut encore être faite : comme l'énoncé ne dépend pas du nombre d'éléments de 1ρ (où 1 est comme d'usage l'élément neutre) nous supposerons toujours que $1\rho = 0$ ou $= 1$, ces deux cas correspondant respectivement aux hypothèses que les états distingués q_- et q_+ sont distincts ou confondus.

II. Preuve de la propriété

Nous notons $\|X\|$ le nombre d'éléments d'un ensemble X quelconque et en particulier nous posons $\|Q\| = d$.

Lemme 1. La relation ρ est fonctionnelle ssi $\|a\rho\| \leq 1$ pour tous les mots a de longueur $|a|$ du plus égal à $L = 1+2\,d(\,d-1\,)$.

Preuve : Soit a un mot de longueur minimum $|a| = n$ parmi ceux pour lesquels $\|a\rho\| \geq 2$. Si l'une des matrices $a'\mu$ ($a' \in A$) a une entrée qui n'est ni vide ni un singleton, l'hypothèse que μ est un transducteur *réduit* implique que $n \leq 1+2(\,d-1\,)$, et le résultat est vérifié dans ce cas. Dans le cas contraire où chaque entrée de chacune des matrices génératrices $a'\mu$ ($a' \in A$) est au plus un singleton, soit $a = a_1 a_2 \ldots a_n$ ($a_i \in A$). Il existe une suite de n-1 paires (q_j, q'_j) d'indices ($1 \leq j \leq n-1, q_j \neq q'_j$) tels que posant $q_0 = q'_0 = q_-$, $q_n = q'_n = q_+$ et $b_j = a_j\mu(\,q_{j-1}, q_j\,)$, $b'_j = a_j\mu(\,q'_{j-1}, q'_j\,)$ ($1 \leq j \leq n$) les mots $b = b_1 \ldots b_n$ et $b' = b'_1 \ldots b'_n$ soient deux éléments *distincts* de $a\rho$. Supposons que $n > L$ et montrons que l'hypothèse de minimalité sur $|a| = n$ conduit à une contradiction.

D'après $L = 1 + 2d(\,d-1\,)$ il existe trois indices $i<j<k$ pour lesquels (q_i, q'_i) = (q_j, q'_j) = (q_k, q'_k), ce qui détermine une factorisation $a = f_1 f_2 f_3 f_4$ où $f_1 = a_1 \ldots a_i$; $f_2 = a_{i+1} \ldots a_j$; $f_3 = a_{j+1} \ldots a_k$; $f_4 = a_k \ldots a_n$ et induit de façon évidente les factorisations correspondantes $b = g_1 g_2 g_3 g_4$ et $b' = g'_1 g'_2 g'_3 g'_4$. Par construction on a les inclusions $g_1 g_4, g'_1 g'_4 \in (\,f_1 f_4\,)\rho$ et $g_1 g_x g_4, g'_1 g'_x g'_4 \in (\,f_1 f_x f_4\,)\rho$ ($x = 2$ ou 3).

En raison du caractère minimal de a, les membres de droite sont des singletons et on a donc les trois équations $g_1 g_4 = g'$ et $g_1 g_x g_4 = g'_1 g'_x g'_4$.

Supposant, par exemple, que $|g_1| \leq |g'_1|$ il en résulte l'existence d'un mot h tel que $g'_1 = g_1 h$ et $g_4 = h g'_4$, d'où en reportant dans les autres équations et en simplifiant, les deux équations $h g_2 = g'_2 h$ et $h g_3 = g'_3 h$.

Il en résulte que

$$b = g_1 g_2 g_3 g_4 = g_1 g_2 g_3 h g'_4 = g_1 g_2 h g'_3 g'_4 = g_1 h g'_2 g'_3 g'_4 = h\, g'_1 g'_2 g'_3 g'_4 = b'$$

en contradiction avec l'hypothèse que b et b' étaient distincts.

Par conséquent $n \leq L$.

Q.E.D.

Nous rappelons maintenant que le *support* $m\beta$ d'une $Q\times Q$ matrice m est la relation binaire sur Q définie par l'ensemble de ses entrées non vides. Autrement dit, β est un morphisme du monoïde $A^*\mu$ dans un monoïde de relations binaires sur Q. On dit que μ est *irréductible* ssi l'union des relations $a\mu\beta$ ($a \in A^*$) est égale à $Q\times Q$ lui-même. Nous avons donc le

<u>Corollaire 2</u>. Quand μ est irréductible et que ρ n'est pas fonctionnelle, il existe trois mots a,h,a' tels que $\|(ah^n a')\rho\| \geq n+1$ ($n \in N$).

<u>Preuve</u> : Supposons que $\|a\rho\| \geq 2$, c'est à dire que $\|a\mu(q_-,q_+)\| \geq 2$.

Puisque μ est irréductible il existe un mot a'' tel que (q_+,q_-) appartienne au support de la matrice $a''\mu$. Posant $h = a''a$ on a donc que l'entrée (q_-,q_+) de $h\mu$ contient au moins deux mots distincts. Il est trivial que la même entrée de $h^n\mu$ contient donc au moins $n+1$ mots distincts pour chaque $n \in N$ et l'on a par conséquent $\|(ah^n)\rho\| = \| (ah^n)\mu(q_-,q_+)\| \geq n+1$ identiquement.

Q.E.D.

Par conséquent la propriété est déjà établie dans le cas particulier où le transducteur μ est irréductible. Comme elle l'est trivialement quand le domaine de ρ est fini nous pouvons désormais procéder par induction sur $\|Q\|$ ou plus exactement, sur le nombre total des entrées non nulles des matrices génératrices $a\mu$ ($a \in A$) du monoïde $A^*\mu$.

Avant de passer au cas général nous rappelons que le produit (de concaténation) de deux relations $\rho,\sigma : A^* \to B^*$ est la relation $\pi = \rho\sigma$ telle que pour chaque mot a on ait :

$$a\pi = \Sigma \{ (a'\rho)(a''\rho) : a',a'' \in A^* ; a = a'a'' \}.$$

Supposons maintenant que μ ne soit pas irréductible. Il existe une partition $Q = Q' \cup Q''$ tel qu'aucun des supports $a\mu\beta$ ne rencontre $Q'' \times Q'$. Nous définissons un morphisme μ' par la condition que pour tout $a \in A$, $q,q' \in Q$ on ait :

$$\begin{aligned} a\mu' (q,q') &= a\mu (q,q') \text{ si } q,q' \in Q' ; \\ &= 0 \qquad\qquad \text{sinon} ; \end{aligned}$$

et un autre morphisme μ'' par la condition que $\mu = \mu' + \mu''$ et que le support des $a\mu''$ ne rencontre pas $Q' \times Q'$. On vérifie facilement que ρ est la somme étendue à tous les $q \in Q'$ des produits $\rho'_q\rho''_q$ où $\rho'_q = \mu'(q_-,q)$ et $\rho''_q = \mu''(q,q_+)$.

Ces relations sont rationnelles et on peut leur appliquer l'hypothèse d'induction. Donc pour conclure la preuve de la propriété il nous suffit de vérifier le

<u>Lemme 3</u> : Si le produit $\pi = \rho\rho'$ de deux relations rationnelles fonctionnelles n'est pas une somme finie de telles relations, il existe trois mots pour lesquels $\|(ah^n a')\pi\| \geq n+1$ ($n \in N$).

<u>Preuve</u> : Nous utilisons les bimachines de Eilenberg ([1], chap XI), c'est à dire que nous associons à ρ un morphisme ϕ de A^* dans un monoïde fini M et une application

partielle $\theta : M \times A \times M \to B^*$ telle que pour chaque mot $a = a_1 \ldots a_n$ ($a_i \in A$) on ait que $a\rho$ est le produit de $i = 1$ à $i = n$ de termes (m'_i, a_i, m''_i)θ où m'_i (resp. m''_i) est l'image par ϕ du facteur gauche (resp. droit) de longueur $i-1$ (resp. $n-i$) de a. Une construction semblable avec $\phi' : A^* \to M'$ et θ' vaut pour ρ'. De fait, on peut remplacer M et M' par leur produit direct et, ceci fait, supposer simplement que $\phi = \phi'$, les deux bimachines ne différant alors que par leurs fonctions θ et θ'. On peut de plus exprimer ρ comme la somme sur tous les $s \in S$ des relations ρ_s qui sont définies comme la restriction de ρ à $s\phi^{-1}$. Il en est de même pour ρ' et il suffit donc d'établir le lemme sous l'hypothèse supplémentaire que le domaine de ρ est $s\phi^{-1} = D$ et que celui de ρ' est $s'\phi^{-1} = D'$.

Le domaine de $\pi = \rho\rho'$ est donc DD' et l'on a que π est fonctionnelle quand chaque mot de A^* admet au plus une factorisation comme produit d'un mot de D par un mot de D'. Dans le cas contraire, l'ensemble des mots $p \neq 1$ satisfaisant les conditions $s'.p\phi = s'$ et $p\phi.s'' = s''$ est un sous semi groupe non vide P^+ de A^* (engendré par la base P) et chaque mot a de DD' admettant plusieurs factorisations admet une factorisation maximale unique $dp_1p_2 \ldots p_nd'$ où $n \geq 1$, $d \in D$, $d' \in D'$ et $p_i, \ldots, p_n \in P$.

Nous étendons θ (et θ') à des fonctions de $M \times A^* \times M$ dans B^* par les identités :

$$(t,xy,t')\theta'' = (t,x,y\phi.t')\theta''. (t.x\phi,y,t')\theta''$$

$$(x,y \in A^*, t,t' \in S , \theta'' = \theta \text{ ou } = \theta').$$

Nous notons σ et σ' les relations rationnelles fonctionnelles de domaine P^+ envoyant respectivement chaque $p_i \in P^+$ sur $g_i = (s,p_i,s')\theta$ et $g'_i = (s,p_i,s')\theta'$.

On vérifie alors que $a\rho$ est l'union pour $i = 0,1,\ldots,n$ des produits $bg_1g_2\ldots g_ig'_{i+1}\ldots g'_nb'$ uù $b = (1,d,s')\theta$ et $b' = (s,d',1)\theta'$. Ces n+1 mots sont les mêmes ssi $g_i = g'_i$ identiquement. Si au contraire $g = p\sigma \neq g' = p\sigma$ (pour un $p \in P$) on voit que les n+1 mots de (ap^na')ρ, c'est à dire les mots $d\,g^ig'^{n-i}\,d'$, ($0 \leq i \leq n$) sont tous distincts. Par conséquent les deux alternatives énoncées dans la propriété correspondent respectivement aux deux cas possibles selon que $\sigma \neq \sigma'$ ou $\sigma = \sigma'$. Nous observons maintenant que comme σ et σ' sont fonctionnelles, on a $\sigma = \sigma'$ ssi la relation rationnelle $\sigma + \sigma'$ l'est aussi, ce que l'on peut vérifier au moyen du lemme 1.

Q.E.D.

Remarque : Un examen plus détaillé des morphismes irréductibles permet de trouver des mots (s'il en existe) tels que Card ((ah^na')ρ) $\geq 2^n$ ($n \in \mathbb{N}$) et, de vérifier que, sinon, Card ($a\rho$) est au plus égal à une fonction polynomiale de la longueur des mots a. ("cas polynomial").

Quand $\|a\rho\|$ est bornée, on peut montrer que ρ est la som me d'une relation ayant un domaine fini et de d relations rationnelles fonctionnelles où d est le plus petit entier tel que $a\rho$ contienne d mots distincts de B^* pour une infinité de mots a de A^*.

Donc pour deux relations de ce type, ρ et ρ', on peut décider si l'on a ou non identiquement $a\rho \subset a\rho'$, et même si cette inclusion est vérifiée dans le complément d'une partie finie de A^*. Je présume qu'il en est de même dans le "cas polymial" mais je ne suis pas parvenu à la démontrer.

Références :

[1] S. Eilenberg. Automata languages and Machins vol. A Academic Press N.Y. 1974.

[2] M. Nivat (1968) Transductions des langages de Chomsky Annales. Inst. Fourier. XVIII. pp. 335 - 455.