Theoretical Computer Science 91 (1991) 71-84
Elsevier

Opérations polynomiales et hiérarchies de concaténation

Mustapha Arfi

Laboratoire d'Informatique de Rouen, Université de Rouen, Faculté des Sciences, B.P. 118, Place Emile Blondel 76134 Mont-Saint-Aignan Cedex, France

Communicated by D. Perrin
Received September 1986
Revised May 1988

Résumé

Arfi, M., Opérations polynomiales et hiérarchies de concaténation, Theoretical Computer Science 91 (1991) 71-84.

Soit $\mathscr{C}$ une classe de langages. Notons Pol($\mathscr{C}$) la fermeture polynomiale de $\mathscr{C}$. Pol($\mathscr{C}$) est la plus petite classe de langages contenant $\mathscr{C}$ et fermée par union finie et produit marqué LaL', où a est une lettre. Nous déterminons les clôtures polynomiales de diverses classes de langages rationnels puis nous étudions les propriétés des fermetures polynomiales. Par exemple, si $\mathscr{C}$ est fermée par quotients (resp. quotients et morphisme inverse), alors il en est de même de Pol($\mathscr{C}$). Notre résultat principal montre que si $\mathscr{C}$ est une algèbre de Boole fermée par résiduels alors Pol($\mathscr{C}$) est fermée par intersection. Comme application, nous affinons la hiérarchie de concaténation introduite par Straubing et nous prouvons la décidabilité des niveaux $\frac{1}{2}$ et $\frac{3}{2}$ de cette hiérarchie.

Abstract

Arfi, M., Opérations polynomiales et hiérarchies de concaténation, Theoretical Computer Science 91 (1991) 71-84.

Given a class $\mathscr{C}$ of languages, let Pol($\mathscr{C}$) be the polynomial closure of $\mathscr{C}$, that is, the smallest class of languages containing $\mathscr{C}$ and closed under the operations union and marked product LaL', where a is a letter. We determine the polynomial closure of various classes of rational languages and we study the properties of polynomial closures. For instance, if $\mathscr{C}$ is closed under quotients (resp. quotients and inverse morphism) then Pol($\mathscr{C}$) has the same property. Our main result shows that if $\mathscr{C}$ is a boolean algebra closed under quotients then Pol($\mathscr{C}$) is closed under intersection. As an application, we refine the concatenation hierarchy introduced by Straubing and we show that the levels $\frac{1}{2}$ and $\frac{3}{2}$ of this hierarchy are decidable.

Introduction

Le célèbre théorème de Kleene (1954) constitue l'un des fondements de la théorie des langages. Il montre qu'un langage sur un alphabet fini est reconnu par un automate fini si et seulement si il peut s'exprimer à partir des lettres en appliquant les opérations union, produit et étoile. Ces trois opérations sont, avec les autres opérations booléennes (intersection et complémentation), les opérations les plus importantes sur les langages rationnels et malgré trente ans d'efforts, leur étude est loin d'être achevée. L'opération étoile mène aux problèmes de hauteur d'étoile; le produit et les opérations booléennes conduisent aux hiérarchies de concaténation de Brzozowski et Straubing. On trouvera d'autres problèmes dans [1] et [3, 4, 5, 6].

Le but de cet article est d'étudier plus particulièrement les opérations "polynomiales", c'est-à-dire l'union et le produit. En fait, comme dans le cas des hiérarchies de concaténation, il est préférable d'utiliser le produit marqué LaL' (où a est une lettre) au lieu du produit de concaténation classique LL'. La fermeture polynomiale, Pol($\mathscr{C}$), d'une classe de langages rationnels $\mathscr{C}$ a été définie par Schützenberger [15]: Pol($\mathscr{C}$) est la plus petite classe contenant $\mathscr{C}$ et fermée par union finie et par produit marqué. Nous déterminons d'abord les fermetures polynomiales de diverses classes de langages, puis nous étudions les propriétés conservées par les opérations polynomiales. Nous montrons par exemple que si $\mathscr{C}$ est fermée par résiduels alors Pol($\mathscr{C}$) est fermée par résiduels et produit classique. Si $\mathscr{C}$ est en plus fermée par morphisme inverse, il en est de même de Pol($\mathscr{C}$). Notre résultat principal montre que si $\mathscr{C}$ est en particulier une algèbre de Boole fermée par résiduels, alors Pol($\mathscr{C}$) est fermée par intersection.

Ce résultat nous permet d'affiner la hiérarchie de concaténation proposée par Straubing de la façon suivante:
- Le niveau 0 est constitué des langages $\emptyset$ et A^*.
- Le niveau $n+\frac{1}{2}$ est la fermeture polynomiale du niveau n.
- Le niveau $n+1$ est la fermeture booléenne du niveau $n+\frac{1}{2}$.

Thomas [22] a mis en évidence le lien entre les niveaux n de cette hiérarchie et les hiérarchies de la logique du premier ordre. Perrin et Pin [8] ont étendu ce résultat aux niveaux $n+\frac{1}{2}$ mais en définissant le niveau $n+\frac{1}{2}$ comme étant la fermeture du niveau n pur les opérations polynomiales *et l'intersection*. Notre résultat principal montre qu'il est inutile de considérer l'intersection et que notre définition coïncide donc avec celle donnée en [8].

Les problèmes de décidabilité concernant les hiérarchies de concaténation sont réputés difficiles. Par exemple, pour la hiérarchie de Straubing, on sait seulement que les niveaux 0 et 1 sont décidables. Straubing [21] a récemment démontré que si L est un langage sur un alphabet à deux lettres, alors on peut décider si L est de niveau 2. D'autres résultats partiels sur ce niveau ont été obtenus par Weil [23]. Tous les autres cas sont ouverts à ce jour. En nous appuyant sur un résultat profond d'Hashiguchi [5], nous montrons que le niveau $\frac{3}{2}$ est également décidable, et ceci quelle que soit la taille de l'alphabet. Au préalable, nous aurons prouvé la décidabilité du niveau $\frac{1}{2}$ grâce à une propriété syntactique remarquable des langages qui le constituent.

Plus précisément, notre étude s'articule de la manière suivante. La première section comporte les définitions et les concepts de base, illustrés par quelques exemples, dont nous dégagerons un théorème important lors des questions de décidabilité. En seconde partie, on mettra en évidence l'action des opérations reconnaissables que sont les résiduels, le morphisme inverse et l'intersection sur les deux opérations polynomiales que l'on vient de définir. Dans la troisième et dernière section sont exposés les problèmes de décidabilité ainsi que leur impact sur la hiérarchie de Straubing, modifiée comme il convient.

1. Opérations polynomiales

Nous renvoyons le lecteur à [2, 9, 12, 13] pour tous les termes non définis. Nous empruntons à [2] la définition suivante: On appelle *classe de langages* une correspondance $\mathscr{C}$ qui, à tout alphabet fini A, associe un ensemble $A^*\mathscr{C}$ de langages reconnaissables sur A^*.

Soit donc $\mathscr{C}$ une classe de langages. Pour tout alphabet fini A, on note $\mathrm{Pol}(A^*\mathscr{C})$ la plus petite famille de langages sur A^* contenant $A^*\mathscr{C}$, fermée par union finie (éventuellement vide) et par l'opération *produit marqué* au sens suivant: $(L, L') \mapsto LaL'$, où a est une lettre. On définit par ce procédé une nouvelle classe de langages, $A \mapsto \mathrm{Pol}(A^*\mathscr{C})$, que l'on désignera par $\mathrm{Pol}(\mathscr{C})$. Ce sera la *fermeture* (ou *clôture*) *polynomiale de* $\mathscr{C}$. On appellera *langage* $\mathscr{C}$*-élémentaire sur* A^* tout langage de la forme $L_0a_1L_1 \ldots a_nL_n$, avec $n \geqslant 0$, $L_i \in A^*\mathscr{C}$ pour $i \in \{0, \ldots, n\}$ et $a_i \in A$ pour $i \in \{1, \ldots, n\}$. Moyennant cette définition, on voit qu'un langage L est dans $\mathrm{Pol}(A^*\mathscr{C})$ si et seulement si L est réunion finie (éventuellement vide) de langages $\mathscr{C}$-élémentaires sur A^*.

Exemple 1.1. Considérons la classe $\mathscr{C}$ définie par $A^*\mathscr{C} = \{\emptyset, A^*\}$ (ou simplement $\{A^*\}$). Alors $\mathrm{Pol}(A^*\mathscr{C})$ est constituée des unions finies de langages de la forme $A^*a_1A^* \ldots a_nA^*$, avec $n \geqslant 0$ et $a_i \in A$ pour $1 \leqslant i \leqslant n$. On a également: $\mathrm{Pol}(A^*\mathscr{C}) = \{L \sqcup\!\sqcup A^* \mid L \subset A^*\}$, où $\sqcup\!\sqcup$ désigne le *produit de mélange* (ou *shuffle*); cf. [7].

Exemple 1.2. Soit la classe de langages $\mathscr{F}$ définie, pour tout alphabet fini A, par $A^*\mathscr{F} = \{B^* \mid B \subset A\}$. On vérifie aisément que $\mathrm{Pol}(A^*\mathscr{F})$ est l'ensemble des unions finies de langages de la forme $A_0^*a_1A_1^* \ldots a_nA_n^*$; $n \geqslant 0$, $A_i \subset A$ pour $i \in \{0, \ldots, n\}$ et $a_i \in A$ pour $i \in \{1, \ldots, n\}$.

Nous verrons ultérieurement l'importance de ces deux exemples, ainsi que celle du théorème suivant, dans l'étude des hiérarchies de concaténation.

Théorème fondamental

Considérons les variétés de monoïdes suivantes:

- $\underline{J}_1$: monoïdes idempotents et commutatifs,
- $\underline{J}$: monoïdes $\mathscr{J}$-triviaux,
- $\underline{R}$: monoïdes $\mathscr{R}$-triviaux,
- $\underline{R}^{\mathrm{r}}$: monoïdes $\mathscr{L}$-triviaux,
- $\underline{DA}$: monoïdes dont les $\mathscr{D}$-classes régulières sont des semigroupes apériodiques.

Nous noterons respectivement $\mathscr{J}_1$, $\mathscr{J}$, $\mathscr{R}$, $\mathscr{R}^{\mathrm{r}}$ et $\mathscr{D}a$ les variétés de langages correspondantes. Reprenons la classe de langages $\mathscr{F}$ définie à l'Exemple 1.2: $A^*\mathscr{F} = \{B^* \mid B \subset A\}$.

Théorème 1.3. *On a les égalités*

$$\mathrm{Pol}(\mathscr{J}_1) = \mathrm{Pol}(\mathscr{J}) = \mathrm{Pol}(\mathscr{R}) = \mathrm{Pol}(\mathscr{R}^r) = \mathrm{Pol}(\mathscr{D}a) = \mathrm{Pol}(\mathscr{F}).$$

Preuve. On a la double série d'inclusions

$$\underline{J}_1 \subset \underline{J} \subset \underline{R} \subset \underline{DA} \quad \text{et} \quad \underline{J}_1 \subset \underline{J} \subset \underline{R}^r \subset \underline{DA}. \tag{1}$$

On sait que $A^*\mathscr{J}_1$ est l'algèbre de Boole engendrée par $A^*\mathscr{F}$. Par conséquent,

$$\mathrm{Pol}(A^*\mathscr{F}) \subset \mathrm{Pol}(A^*\mathscr{J}_1). \tag{2}$$

De plus, Schützenberger [16] a montré que tout langage de $A^*\mathscr{D}a$ est union disjointe de langages de la forme $A_0^* a_1 A_1^* \ldots a_n A_n^*$; où $n \geq 0$; $A_0, \ldots, A_n \subset A$ et $a_1, \ldots, a_n \in A$ et où le produit $A_0^* a_1 A_1^* \ldots a_n A_n^*$ est non ambigu. Ce qui donne l'inclusion $A^*\mathscr{D}a \subset \mathrm{Pol}(A^*\mathscr{F})$. Il en résulte

$$\mathrm{Pol}(A^*\mathscr{D}a) \subset \mathrm{Pol}(A^*\mathscr{F}). \tag{3}$$

Le théorème découle des assertions (1), (2) et (3). □

2. Opérations polynomiales et opérations reconnaissables

Les opérations polynomiales sont l'union et le produit marqué mentionnés au début de cet article. Dans cette section, nous nous proposons d'étudier les rapports entre ces opérations et quelques opérations reconnaissables, à savoir les résiduels, le morphisme inverse et l'intersection.

2.1. Fermeture par résiduels et par produit classique

Soit $\mathscr{C}$ une classe de langages. On dira que $\mathscr{C}$ est fermée par résiduels (resp. résiduels à gauche, resp. résiduels à droite) si pour tout alphabet fini A, tout élément $a \in A$ et tout langage $L \in A^*\mathscr{C}$, $a^{-1}L$ et $La^{-1} \in A^*\mathscr{C}$ (resp. $a^{-1}L \in A^*\mathscr{C}$, resp. $La^{-1} \in A^*\mathscr{C}$). Nous dirons que $\mathscr{C}$ est fermée par produit (classique) si pour tout alphabet fini A et tous langages $L, L' \in A^*\mathscr{C}$, le produit $L \cdot L'$ est dans $A^*\mathscr{C}$.

Proposition 2.1.1. *Soit $\mathscr{C}$ une classe de langages fermée par residuels (resp. à gauche, resp. à droite), alors* $\mathrm{Pol}(\mathscr{C})$ *a la même propriété. De plus,* $\mathrm{Pol}(\mathscr{C})$ *est fermée par produit.*

Preuve. On examinera uniquement le cas des résiduels à gauche, les autres étant similaires.

Fermeture par résiduels à gauche: Supposons donc $\mathscr{C}$ fermée par résiduels à gauche et notons $\mathscr{C}'$ la classe définie par

$$A^*\mathscr{C}' = \{L \in \mathrm{Pol}(A^*\mathscr{C}) \mid \forall a \in A,\ a^{-1}L \in \mathrm{Pol}(A^*\mathscr{C})\}.$$

Il suffit d'établir l'inclusion $\mathrm{Pol}(A^*\mathscr{C}) \subset A^*\mathscr{C}'$. Puisque par hypothèse $A^*\mathscr{C} \subset A^*\mathscr{C}'$, il suffit de vérifier que $A^*\mathscr{C}'$ est stable par union et par produit marqué. Or cela résulte des formules

$$a^{-1}(L \cup L') = a^{-1}L \cup a^{-1}L', \quad \text{et}$$

$$a^{-1}(LbL') = \begin{cases} (a^{-1}L)bL' & \text{si } 1 \notin L, \text{ ou } b \neq a, \\ (a^{-1}L)bL' \cup L' & \text{si } 1 \in L \text{ et } b = a. \end{cases}$$

Fermeture par produit: Soit $\mathscr{C}$ une classe fermée par résiduels à gauche. On sait maintenant que $\mathrm{Pol}(\mathscr{C})$ l'est également. La fermeture par produit de $\mathrm{Pol}(\mathscr{C})$ résulte immédiatement de la formule:

$$LL' = \begin{cases} \bigcup_{a \in A} La(a^{-1}L') & \text{si } 1 \notin L', \\ \left(\bigcup_{a \in A} La(a^{-1}L') \right) \cup L & \text{si } 1 \in L'. \end{cases} \qquad \square$$

Corollaire 2.1.2. *La clôture polynomiale d'une variété de langages est fermée par résiduels et produit de concaténation classique.*

2.2. Fermeture par morphisme inverse

Soit $\mathscr{C}$ une classe de langages. On dira que $\mathscr{C}$ est fermée par morphisme inverse si, pour tout morphisme $\varphi : A^* \to B^*$, $L \in B^*\mathscr{C}$ entraîne $L\varphi^{-1} \in A^*\mathscr{C}$.

Proposition 2.2.1. *Soit $\mathscr{C}$ une classe de langages fermée par résiduels et morphisme inverse. Alors* $\mathrm{Pol}(\mathscr{C})$ *est fermée par morphisme inverse.*

Preuve. Soit $\varphi : A^* \to B^*$ un morphisme de monoïdes libres. Puisque $(L_1 \cup L_2)\varphi^{-1} = L_1\varphi^{-1} \cup L_2\varphi^{-1}$, il suffit de démontrer que si L est un langage $\mathscr{C}$-élémentaire sur B^*, alors $L\varphi^{-1} \in \mathrm{Pol}(A^*\mathscr{C})$. Posons donc $L = L_0 b_1 L_1 \ldots b_n L_n$ avec $n \geq 0$; $L_0, \ldots, L_n \in B^*\mathscr{C}$ et $b_1, \ldots, b_n \in B$. Pour chaque lettre $b \in B$, on note

$$\Delta(b) = \{(u, a, v) \in B^* \times A \times B^* \mid a\varphi = ubv\}.$$

Remarquons que $\Delta(b)$ est toujours un ensemble fini. La Proposition 2.2.1 découle alors immédiatement du lemme qui suit. $\square$

Lemme 2.2.2. *On a la formule*

$$\begin{aligned}(L_0 b_1 L_1 \ldots b_n L_n)\varphi^{-1} &= \bigcup_{(u_i, a_i, v_i) \in \Delta(b_i)} (L_0 u_1^{-1})\varphi^{-1} a_1 (v_1^{-1} L_1 u_2^{-1})\varphi^{-1} \\ &\qquad \ldots (v_{n-1}^{-1} L_{n-1} u_n^{-1})\varphi^{-1} a_n (v_n^{-1} L_n)\varphi^{-1}.\end{aligned}$$

Preuve. Désignons respectivement par K et K' les membres gauche et droit de cette égalité.

(a) $K \subset K'$: Soit $w \in K$. $w\varphi \in L_0 b_1 L_1 \ldots b_n L_n$. On a donc $w\varphi = l_0 b_1 l_1 \ldots b_n l_n$; avec $l_i \in L_i$ pour $i \in \{0, \ldots, n\}$. Puisque les lettres $b_1, \ldots, b_n$ apparaissent dans $w\varphi$, il existe, pour tout $i \in \{1, \ldots, n\}$, $(u_i, a_i, v_i) \in \Delta(b_i)$, il existe $w_0, \ldots, w_n \in A^*$ tels que $w = w_0 a_1 w_1 \ldots a_n w_n$, avec

$$w_0\varphi.u_1 = l_0, \qquad v_n.w_n\varphi = l_n \quad \text{et} \quad v_i.w_i\varphi.u_{i+1} = l_i \text{ pour } i \in \{1, \ldots, n-1\}.$$

Par conséquent, $w \in K'$.

(b) $K' \subset K$: Soit $w \in K'$. Pour tout $i \in \{1, \ldots, n\}$, il existe alors $(u_i, a_i, v_i) \in \Delta(b_i)$ tels que:

$$w \in (L_0 u_1^{-1})\varphi^{-1} a_1 (v_1^{-1} L_1 u_2^{-1})\varphi^{-1} \ldots (v_{n-1}^{-1} L_{n-1} u_n^{-1})\varphi^{-1} a_n (v_n^{-1} L_n)\varphi^{-1}.$$

Par conséquent, w s'écrit: $w = w_0 a_1 w_1 \ldots a_n w_n$; avec

$$w_0 \in (L_0 u_1^{-1})\varphi^{-1}, \qquad w_n \in (v_n^{-1} L_n)\varphi^{-1}$$
$$\text{et} \quad w_i \in (v_i^{-1} L_i u_{i+1}^{-1})\varphi^{-1}, \ \forall i \in \{1, \ldots, n-1\}.$$

Il en résulte

$$\begin{aligned} w\varphi &= (w_0\varphi)(a_1\varphi)(w_1\varphi) \ldots (a_n\varphi)(w_n\varphi) \\ &= (w_0\varphi)(u_1 b_1 v_1)(w_1\varphi) \ldots (u_n b_n v_n)(w_n\varphi) \\ &= (w_0\varphi.u_1) b_1 (v_1.w_1\varphi.u_2) \ldots (v_{n-1}.w_{n-1}\varphi.u_n) b_n (v_n.w_n\varphi). \end{aligned}$$

Donc $w\varphi \in L_0 b_1 L_1 \ldots b_n L_n$; c'est-à-dire $w \in K$. $\square$

Corollaire 2.2.3. *Si $\mathcal{V}$ est une variété de langages, alors* Pol($\mathcal{V}$) *est fermée par morphisme inverse.*

C'est une conséquence immédiate de la Proposition 2.2.1.

2.3. Fermeture par intersection

Pour clore cette section, nous allons donner des conditions suffisantes pour qu'une clôture polynomiale soit fermée par intersection finie. Soit $\mathscr{C}$ une classe de langages. Nous dirons que $\mathscr{C}$ est fermée par union (resp. intersection, resp. différence) si, pour tout alphabet fini A, $L, L' \in A^*\mathscr{C}$ entraîne $L \cup L' \in A^*\mathscr{C}$ (resp. $L \cap L' \in A^*\mathscr{C}$, resp. $L \setminus L' \in A^*\mathscr{C}$). Enonçons tout de suite le théorème fondamental de cette seconde partie.

Théorème 2.3.1. *Supposons $\mathscr{C}$ fermée par union, intersection, différence et résiduels. Alors* Pol($\mathscr{C}$) *est fermée par intersection.*

Nous commencerons d'abord par démontrer un lemme. Soit L un langage rationnel sur un alphabet fini A. Désignons par $\mathscr{A} = (Q, A, ., q^-, F)$ l'automate minimal reconnaissant L. Pour tous états $q, r \in Q$, on définit le langage: $L(q, r) = \{u \in A^* \mid q.u = r\}$.

Lemme 2.3.2. *Soit $\mathscr{C}$ une classe de langages fermée par union, intersection, différence et résiduels. Soient A un alphabet fini, L un langage dans $A^*\mathscr{C}$ et $\mathscr{A} = (Q, A, \cdot, q^-, F)$ l'automate minimal de L. Alors, pour tout couple $(q, r) \in Q \times Q$, $L(q, r) \in A^*\mathscr{C}$.*

Preuve. Par construction de l'automate minimal et par définition de l'action $\cdot$, il existe $x, y \in A^*$ tels que $q = x^{-1}L$ et $r = y^{-1}L$, et $L(q, r)$ s'écrit

$$L(q, r) = \{u \in A^* \mid (x^{-1}L).u = y^{-1}L\} = \{u \in A^* \mid (xu)^{-1}L = y^{-1}L\}.$$

D'où

$$L(q, r) = \{u \in A^* \mid \forall v \in A^*, xuv \in L \Leftrightarrow yv \in L\}.$$

Un petit moment de réflexion convaincra le lecteur de légalité $L(q, r) = L_1 \cap L_2$ avec

$$L_1 = \{u \in A^* \mid \forall v \in y^{-1}L, xuv \in L\} \quad \text{et} \quad L_2 = \{u \in A^* \mid \forall v \notin y^{-1}L, xuv \notin L\}.$$

Or on a

$$L_1 = \bigcap_{v \in y^{-1}L} x^{-1}Lv^{-1}$$

et

$$L_2 = A^* \setminus \{u \in A^* \mid \exists v \notin y^{-1}L, xuv \in L\} = A^* \setminus \bigcup_{v \notin y^{-1}L} x^{-1}Lv^{-1}.$$

D'où

$$L(q, r) = \Big(\bigcap_{v \in y^{-1}L} x^{-1}Lv^{-1}\Big) \cap \Big(A^* \setminus \bigcup_{v \notin y^{-1}L} x^{-1}Lv^{-1}\Big).$$

Et finalement, le lemme est conséquence de la formule:

$$L(q, r) = \bigcap_{v \in y^{-1}L} x^{-1}Lv^{-1} \setminus \bigcup_{v \notin y^{-1}L} x^{-1}Lv^{-1}. \qquad \square$$

Preuve du Théorème 2.3.1. Etant donné la distributivité de l'intersection par rapport à la réunion, il suffit d'établir le résultat pour les langages $\mathscr{C}$-élémentaires sur A^*. Soient donc $K = K_0a_1K_1 \ldots a_nK_n$ et $K' = K'_0a'_1K'_1 \ldots a'_{n'}K'_{n'}$ deux langages de A^*, avec:

$$a_i \in A \quad \text{pour } i \in \{1, \ldots, n\}, \qquad a'_i \in A \quad \text{pour } i \in \{1, \ldots, n'\},$$

$$K_i \in A^*\mathscr{C} \quad \text{pour } i \in \{0, \ldots, n\}, \qquad K'_i \in A^*\mathscr{C} \quad \text{pour } i \in \{0, \ldots, n'\}.$$

Les langages $K_0, \ldots K_n, K'_0, \ldots, K'_{n'}$ sont respectivement reconnus par des automates $\mathscr{A}_0, \ldots, \mathscr{A}_n, \mathscr{A}'_0, \ldots, \mathscr{A}'_{n'}$, supposés tous minimaux. Posons:

$$\mathscr{A}_i = (Q_i, \cdot, q_i^-, F_i) \qquad \text{pour } i \in \{0, \ldots, n\} \quad \text{et}$$

$$\mathscr{A}'_i = (Q'_i, A, \cdot, q_i'^-, F'_i) \quad \text{pour } i \in \{0, \ldots, n'\}.$$

Donnons tout d'abord une idée intuitive de cette preuve. K est reconnu par l'automate non déterministe $\mathcal{A}$ représenté par le diagramme dans la Fig. 1.
On peut de la même manière construire l'automate non déterministe $\mathcal{A}'$ reconnaissant K'. Le produit direct $\mathcal{B}=\mathcal{A}\times\mathcal{A}'$ est un automate reconnaissant $K\cap K'$ et une analyse minutieuse de $\mathcal{B}$ nous permettra d'établir le résultat dont voici la preuve formelle.

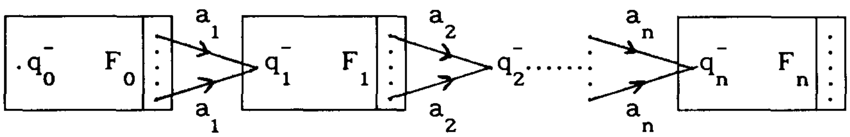

Fig. 1.

Pour tout $j\in\{0,\ldots,n\}$, tout $j'\in\{0,\ldots,n'\}$ et tout quadruplet $(q,q',r,r')\in Q_j\times Q'_{j'}\times Q_j\times Q'_{j'}$, on pose

$$L(q,q',r,r')=\{u\in A^*\mid q.u=r \text{ dans } \mathcal{A}_j \text{ et } q'.u=r' \text{ dans } \mathcal{A}'_{j'}\}.$$

On a, en suivant les notations du début du paragraphe,

$$L(q,q',r,r')=K_j(q,r)\cap K'_{j'}(q',r').$$

Vu les hypothèses et le Lemme 2.3.2, on peut affirmer que tout langage de la forme $L(q,q',r,r')$ appartient à $A^*\mathscr{C}$. Nous allons à présent montrer que $K\cap K'=L$, avec

$$L=\bigcup L(q_0,q'_0,r_0,r'_0)b_1L(q_1,q'_1,r_1,r'_1)\ldots b_sL(q_s,q'_s,r_s,r'_s),$$

où l'union est prise sur tous les langages de la forme ci-dessus satisfaisant les cinq conditons suivantes:

(1) Il existe deux applications strictement croissantes

$$\Theta:\{0,\ldots,n+1\}\to\{0,\ldots,s+1\} \quad\text{et}\quad \Theta':\{0,\ldots,n'+1\}\to\{0,\ldots,s+1\}$$

telles que

$$\{1,\ldots,s\}=\{1\Theta,\ldots,n\Theta\}\cup\{1\Theta',\ldots,n'\Theta'\},$$
$$0\Theta=0\Theta'=0,\qquad (n+1)\Theta=(n'+1)\Theta'=s+1.$$

(2) Pour tout $i\in\{1,\ldots,n\}$, $b_{i\Theta}=a_i$. Pour tout $i\in\{1,\ldots,n'\}$, $b_{i\Theta'}=a'_i$.

(3) Pour tout $i\in\{1,\ldots,s\}$ et tout $j\in\{1,\ldots n\}$ (resp. $\{1,\ldots,n'\}$), si $i=j\Theta$ (resp. $i=j\Theta'$) alors $r_{i-1}\in F_{j-1}$ (resp. $r'_{i-1}\in F'_{j-1}$) et $q_i=q_j^-$ (resp. $q'_i=q'^-_j$).

(4) Pour tout $i\in\{1,\ldots,s\}$ et tout $j\in\{1,\ldots,n\}$ (resp. $\{1,\ldots,n'\}$), si $j\Theta<i<(j+1)\Theta$ (resp. $j\Theta'<i<(j+1)\Theta'$) alors $q_i,r_i\in Q_j$ (resp. $q'_i,r'_i\in Q'_j$) et $r_{i-1}.b_i=q_i$ (resp. $r'_{i-1}.b_i=q'_i$).

(5) $q_0=q_0^-$, $q'_0=q'^-_0$, $r_s\in F_n$ et $r'_s\in F'_{n'}$.

De la condition (1) ressort l'inégalité: $s\leqslant n+n'$. Par conséquent, s est borné. De plus, on a montré que les langages $L(q,q',r,r')$ sont dans $A^*\mathscr{C}$. De ce fait, l'établissement, par double inclusion, de l'égalité $K\cap K'=L$ impliquera l'appartenance de $K\cap K'$ à $\mathrm{Pol}(A^*\mathscr{C})$ et achèvera la preuve.

(a) $L \subset K \cap K'$: Soit $u \in L$. On a $u = u_0 b_1 u_1 \dots b_s u_s$; avec $u_i \in L(q_i, q'_i, r_i, r'_i)$ pour $i \in \{0, \dots, s\}$ et $b_i \in A$ pour $i \in \{1, \dots, s\}$. Montrons que $u \in K$. Pour tout $i \in \{0, \dots, n\}$, on pose

$$v_i = u_{i\Theta} b_{i\Theta+1} u_{i\Theta+1} \dots b_{(i+1)\Theta-1} u_{(i+1)\Theta-1}.$$

En appliquant (1) et (2), on obtient:

$$v_0 = u_0 b_1 u_1 \dots b_{1\Theta-1} u_{1\Theta-1},$$
$$\vdots$$
$$v_n = u_{n\Theta} b_{n\Theta+1} u_{n\Theta+1} \dots b_s u_s.$$

D'où $u = v_0 a_1 v_1 \dots a_n v_n$. De plus, les conditions (3), (4) et (5) impliquent

$$q_0^- . v_0 = q_0.(u_0 b_1 u_1 \dots b_{1\Theta-1} u_{1\Theta-1}) = r_{1\Theta-1} \in F_0,$$
$$r_{1\Theta-1}.b_{1\Theta} = r_{1\Theta-1}.a_1 = q_1^-,$$
$$q_1^- . v_1 = q_{1\Theta}.(u_{1\Theta} b_{1\Theta+1} u_{1\Theta+1} \dots b_{2\Theta-1} u_{2\Theta-1}) = r_{2\Theta-1} \in F_1,$$
$$r_{2\Theta-1}.b_{2\Theta} = r_{2\Theta-1}.a_2 = q_2^-,$$
$$\vdots$$
$$q_n^- . v_n = q_{n\Theta}.(u_{n\Theta} b_{n\Theta+1} u_{n\Theta+1} \dots b_s u_s) = r_s \in F_n.$$

Par conséquent, $v_0 \in K_0, \dots, v_n \in K_n$. Donc $u \in K$. On montre de même que $u \in K'$.

(b) $K \cap K' \subset L$: Réciproquement, soit $u \in K \cap K'$. Alors on a

$$u = u_0 a_1 u_1 \dots a_n u_n, \quad u_i \in K_i \text{ pour } i \in \{0, \dots, n\}, \quad \text{et}$$
$$u = u'_0 a'_1 u'_1 \dots a'_{n'} u'_{n'}, \quad u'_i \in K'_i \text{ pour } i \in \{0, \dots, n'\}.$$

En particulier, $a_1 \dots a_n$ et $a'_1 \dots a'_n$ sont deux sous-mots de u. Il existe donc un sous-mot $b_1 \dots b_s$ de u et deux applications Θ et Θ' vérifiant (1) et (2). On a alors $u = v_0 b_1 v_1 \dots b_s v_s$; avec

$$u_j = v_{j\Theta} b_{j\Theta+1} v_{j\Theta+1} \dots b_{(j+1)\Theta-1} v_{(j+1)\Theta-1} \qquad (0 \leq j \leq n),$$
$$u'_j = v_{j\Theta'} b_{j\Theta'+1} v_{j\Theta'+1} \dots b_{(j+1)\Theta'-1} v_{(j+1)\Theta'-1} \qquad (0 \leq j \leq n').$$

Pour tout $i \in \{0, \dots, s\}$, posons

- $q_i = q_j^-$, s'il existe $j \in \{0, \dots, n\}$ tel que $i = j\Theta$,
- $r_i = q_i.v_i$ dans $\mathscr{A}_j$, s'il existe $j \in \{0, \dots, n\}$ tel que $j\Theta \leq i \leq (j+1)\Theta - 1$,
- $q_{i+1} = r_i.b_{i+1}$ dans $\mathscr{A}_j$, s'il existe $j \in \{0, \dots, n\}$ tel que $j\Theta \leq i < (j+1)\Theta - 1$.

Nous définissons ainsi une famille $(q_i, r_i)_{0 \leq i \leq s}$. La famille $(q'_i, r'_i)_{0 \leq i \leq s}$ se définit de façon analogue. Alors, pour tout $i \in \{0, \dots, s\}$, on a $v_i \in L(q_i, q'_i, r_i, r'_i)$. De plus, s'il existe j tel que $i = j\Theta$, il vient

$$r_{i-1} = q_{i-1}.v_{i-1} = \cdots = q_{j-1}^- . v_{(j-1)\Theta} b_{(j-1)\Theta+1} v_{(j-1)\Theta+1} \dots b_{j\Theta-1} v_{j\Theta-1}$$
$$= q_{j-1}^- . u_{j-1} \in F_{j-1}$$

car $u_{j-1} \in K_{j-1}$.

La condition (3) est donc vraie. On démontre aussi la validité des conditions (4) et (5) et ceci de manière similaire. Par conséquent, $u \in L$. □

Soit $\mathscr{C}$ une classe de langages. On dira que $\mathscr{C}$ est une *algèbre de Boole* si, pour tout alphabet fini A, $A^*\mathscr{C}$ est une algèbre de Boole. Le corollaire ci-dessous constitue une conséquence immédiate du théorème.

Corollaire 2.3.3. *Soit $\mathscr{C}$ une algèbre de Boole stable par résiduels (ce qui est le cas d'une variété de langages). Alors* Pol($\mathscr{C}$) *est fermée par intersection.*

Remarque. La preuve du Théorème 2.3.1 fournit un algorithme permettant la décomposition d'une intersection en réunion de langages $\mathscr{C}$-élémentaires (pour des langages appartenant à la clôture polynomiale d'une classe $\mathscr{C}$), comme l'illustre l'exemple simple suivant.

Exemple. Reprenons la classe $\mathscr{C}$ déjà rencontrée à l'Exemple 1.1 de la première section: $A^*\mathscr{C}=\{\emptyset, A^*\}$. Cette classe n'est autre que la variété triviale. Rappelons que $\mathrm{Pol}(A^*\mathscr{C})$ est l'ensemble des unions finies de langages de la forme $A^*a_1A^*\dots a_nA^*$, $n\geqslant 0$ et $a_i\in A$ pour $i\in\{1,\dots,n\}$. D'après le corollaire précédent, $\mathrm{Pol}(\mathscr{C})$ est stable par intersection. On considère l'alphabet $A=\{a, b\}$ et les deux langages suivants de $\mathrm{Pol}(A^*\mathscr{C})$: $K=A^*aA^*$ et $K'=A^*bA^*$. On obtient la décomposition de $K\cap K'$ en langages $\mathscr{C}$-élémentaires:

$$K\cap K'=A^*aA^*bA^*\cup A^*bA^*aA^*.$$

3. Hiérarchies de concaténation et problèmes de decidabilité

Il existe une méthode générale permettant de construire des hiérarchies de concaténation. On part d'une classe de langages $\mathscr{B}_0$ qui est une algèbre de Boole. On définit le niveau $n+1$ à partir du niveau n comme étant l'algèbre de Boole $\mathscr{B}_{n+1}$ engendrée par les langages $\mathscr{B}_n$-élémentaires. On obtient ainsi une hiérarchie ascendante $\mathscr{B}_0\subset\mathscr{B}_1\subset\cdots\subset\mathscr{B}_n\subset\cdots$, stricte ou non. Lorsque le niveau initial $\mathscr{B}_0$ est une variété de langages, tous les autres niveaux le sont.

La hiérarchie de Straubing $(\mathscr{V}_n, n\in\mathbb{N})$ s'appuie sur ce modèle [20]. Son point de départ est la variété triviale: $A^*\mathscr{V}_0=\{\emptyset, A^*\}$. Cette hiérarchie constitue donc une hiérarchie de variétés de langages. Notons $(\underline{V}_n, n\in\mathbb{N})$ la suite des variétés de monoïdes qui lui correspondent par le théorème d'Eilenberg.

Théorème 3.1. *La hiérarchie de Straubing est stricte: pour tout $n\geqslant 0$, $\mathscr{V}_n$ est strictement contenue dans $\mathscr{V}_{n+1}$. De plus $\bigcup_{n\geqslant 0}\underline{V}_n=\underline{A}$, où $\underline{A}$ désigne la variété des monoïdes apériodiques.*

On considère la hiérarchie de concaténation définie comme suit:

$$A^*\mathscr{V}_0=\{\emptyset, A^*\},\qquad A^*\mathscr{V}_{n+\frac{1}{2}}=\mathrm{Pol}(A^*\mathscr{V}_n),\qquad A^*\mathscr{V}_{n+1}=\mathrm{Bool}(A^*\mathscr{V}_{n+\frac{1}{2}}).$$

Le symbole "Bool" est mis pour "fermeture booléenne". On construit ainsi une hiérarchie ascendante de classes de langages:

$$\mathscr{V}_i : A\mapsto A^*\mathscr{V}_i;\quad i=0,\tfrac{1}{2},1,\tfrac{3}{2},2,\dots$$

Pour les niveaux indexés par des entiers, cette hiérarchie coïncide avec celle de Straubing. Par conséquent, les niveaux $(\mathcal{V}_n, n \in \mathbb{N})$ représentent des variétés de langages.

Proposition 3.2. *La hiérarchie* $\mathcal{V}_i$, $i \in \{0, \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots\}$ *est stricte.*

Preuve. Supposons par l'absurde qu'il existe i tel que $\mathcal{V}_i = \mathcal{V}_{i+1/2}$. Deux situations sont alors à distinguer: $i \in \mathbb{N}$ ou $i \in \mathbb{N} + \frac{1}{2}$. Dans le premier cas, on a $\mathcal{V}_i = \mathrm{Pol}(\mathcal{V}_i)$. D'où

$$\mathrm{Bool}(\mathcal{V}_i) = \mathrm{Bool}(\mathrm{Pol}(\mathcal{V}_i)) = \mathcal{V}_{i+1}.$$

Comme $\mathcal{V}_i$ est une variété de langages, il vient par définition $\mathrm{Bool}(\mathcal{V}_i) = \mathcal{V}_i$. Il en résulte $\mathcal{V}_i = \mathcal{V}_{i+1}$; ce qui contredit le caractère strict de la hiérarchie de Straubing. Si $i \in \mathbb{N} + \frac{1}{2}$, posons $n = i - \frac{1}{2}$. On obtient alors $\mathcal{V}_{n+1/2} = \mathcal{V}_{n+1}$. Puisque $n \in \mathbb{N}$, on a $\mathcal{V}_{n+1/2} = \mathrm{Pol}(\mathcal{V}_n)$. Ce qui donne, en appliquant de nouveau l'opérateur Pol: $\mathrm{Pol}(\mathcal{V}_n) = \mathrm{Pol}(\mathcal{V}_{n+1})$. D'où

$$\mathrm{Bool}(\mathrm{Pol}(\mathcal{V}_n)) = \mathrm{Bool}(\mathrm{Pol}(\mathcal{V}_{n+1}));$$

c'est-à-dire, $\mathcal{V}_{n+1} = \mathcal{V}_{n+2}$. On retrouve là encore la même contradiction. $\square$

Le lien entre les niveaux n de cette hiérarchie et les hiérarchies de la logique du premier ordre a été établi par Thomas [22]. La généralisation aux niveaux $n + \frac{1}{2}$ a été effectuée par Perrin et Pin [8]. Cependant ces auteurs ont défini le niveau $n + \frac{1}{2}$ comme étant la fermeture du niveau n pour les opérations polynomiales *et* l'intersection. Or, le Théorème 2.3.1 indique qu'il est inutile de considérer l'intersection et que, par conséquent, nos définitions sont identiques.

On dit qu'une classe de langages $\mathscr{C}$ est décidable si, pour tout alphabet A, il existe un algorithme pour décider si un langage rationnel donné L de A^* est ou n'est pas élément de $A^*\mathscr{C}$. Il est facile de voir que $\mathcal{V}_0$ est décidable. La décidabilité de $\mathcal{V}_{1/2}$ résulte immédiatement de l'énoncé suivant.

Théorème 3.3. *Soit $L \subset A^*$ un langage rationnel et $\eta : A^* \to M$ le morphisme syntactique de L. On pose $P = L\eta$. Alors $L \in A^*\mathcal{V}_{1/2}$ si et seulement si*

$$\forall r, s, t \in M, \quad rt \in P \Rightarrow rst \in P.$$

Preuve. Si $L \in A^*\mathcal{V}_{1/2}$ on a vu que $L = L \sqcup\!\sqcup A^*$. Soient $r, s, t \in M$ tels que $rt \in P$. Soient x, y, z des mots tels que $x\eta = r$, $y\eta = s$ et $z\eta = t$. Il vient $(xz)\eta = rt \in P$; d'où $xz \in P\eta^{-1} = L$. Il en découle: $xyz \in L \sqcup\!\sqcup A^* = L$ et donc $rst = (xyz)\eta \in L\eta = P$. Réciproquement, supposons que P satisfasse la condition ci-dessus. Soient $x, y, z \in A^*$ tels que $xz \in L$. Alors $(x\eta)(z\eta) \in P$, d'où $(x\eta)(y\eta)(z\eta) \in P$, ou encore $xyz \in P\eta^{-1} = L$. Il en résulte que $L = L \sqcup\!\sqcup A^*$ et donc $L \in A^*\mathcal{V}_{1/2}$. $\square$

La décidabilité de $\mathscr{V}_1$ a été démontrée par Simon [18].

Théorème 3.4. *Un langage rationnel L est dans $A^*\mathscr{V}_1$ si et seulement si son monoïde syntactique est $\mathscr{J}$-trivial. Autrement dit, $\underline{V}_1 = \underline{J}$.*

Nous en arrivons à présent au résultat essentiel de cette dernière section.

Théorème 3.5. *Il existe un algorithme pour décider si un langage rationnel L de A^* est élément de $A^*\mathscr{V}_{3/2}$.*

La démonstration repose sur un résultat d'Hashiguchi [5]. Soient A un alphabet fini et $A^*\mathscr{S}$ une famille *finie* de langages *rationnels* sur A^*. Soit Ω une partie de l'ensemble $\{\cup, ., *\}$. On notera $\mathscr{C}^{\Omega}(A^*\mathscr{S})$ la plus petite famille de langages sur A^* contenant $A^*\mathscr{S}$ et fermée par les opérations de Ω.

Théorème 3.6. *La classe $\mathscr{C}^{\Omega}(\mathscr{S}): A \mapsto \mathscr{C}^{\Omega}(A^*\mathscr{S})$ est décidable.*

La proposition ci-après, ainsi que le corollarie qui va en résulter, nous permettront de donner une preuve simplifiée du Théorème 3.5.

Proposition 3.7. *Soit $\mathscr{C}$ une classe de langages vérifiant les deux conditions suivantes:*
(i) Pol($\mathscr{C}$) *est fermée par produit (classique).*
(ii) *Pour tout alphabet fini A, $A^*\mathscr{C}$ est un ensemble fini et* $\{1\} \in$ Pol($A^*\mathscr{C}$).
Alors Pol($\mathscr{C}$) *est décidable.*

Preuve. Considérons la classe de langages $\mathscr{S}$ définie par

$$A^*\mathscr{S} = A^*\mathscr{C} \cup \{\{a\} \mid a \in A\}.$$

On choisira pour Ω la partie $\{\cup, .\}$. D'après le Théorème 3.6, $\mathscr{C}^{\Omega}(\mathscr{S})$ est décidable. La démonstration consistera à établir l'égalité Pol$(A^*\mathscr{C}) = \mathscr{C}^{\Omega}(A^*\mathscr{S})$. L'inclusion Pol$(A^*\mathscr{C}) \subset \mathscr{C}^{\Omega}(A^*\mathscr{S})$ étant évidente, il suffira de montrer l'inclusion dans l'autre sens. Comme $\{1\} \in$ Pol$(A^*\mathscr{C})$, les lettres appartiennent également à Pol$(A^*\mathscr{C})$, car pour tout $a \in A$, $a = 1.a.1$. Par suite, $A^*\mathscr{S} \subset$ Pol$(A^*\mathscr{C})$. Puisque Pol($\mathscr{C}$) est fermée par réunion (par définition) et par produit (par hypothèse), il vient $\mathscr{C}^{\Omega}(A^*\mathscr{S}) \subset$ Pol$(A^*\mathscr{C})$. □

On reprend maintenant la classe de langages $\mathscr{F}$ définie à l'Exemple 1.2. Rappelons que cette classe est donnée par $A^*\mathscr{F} = \{B^* \mid B \subset A\}$. On a vu que, pour tout alphabet fini A, Pol$(A^*\mathscr{F})$ est l'ensemble des unions finies de langages $\mathscr{F}$-élémentaires: $A_0^* a_1 A_1^* \ldots a_n A_n^*$; $n \geq 0$, $A_i \subset A$ pour $i \in \{0, \ldots, n\}$ et $a_i \in A$ pour $i \in \{1, \ldots, n\}$.

Corollaire 3.8. Pol($\mathscr{F}$) *est décidable.*

Préuve. Pour tout alphabet A, $A^*\mathscr{F}$ est évidemment un ensemble fini. De plus, $\{1\} \in \mathrm{Pol}(A^*\mathscr{F})$ car $\{1\} = \emptyset^* \in A^*\mathscr{F}$. Reste alors à établir la fermeture de Pol($\mathscr{F}$) pour le produit classique. Or, en vertu du Théorème 1.3, on a l'égalité Pol $(\mathscr{F}) = \mathrm{Pol}(\mathscr{J})$, où $\mathscr{J}$ est la variété des langages *testables par morceaux*, et Pol($\mathscr{J}$) est fermée par produit en raison du Corollaire 2.1.2. □

Nous sommes à présent en mesure de fournir la preuve du Théorème 3.5.

Preuve du Théorème 3.5. Désignons par $\underline{J}$ la variété des monoïdes $\mathscr{J}$-triviaux et par $\mathscr{J}$ la variété de langages associée. Par le théorème de Simon (Théorème 3.4), on a $\mathscr{V}_1 = \mathscr{J}$. D'après le Théorème 1.3, on a $\mathrm{Pol}(\mathscr{F}) = \mathrm{Pol}(\mathscr{J})$. D'où $\mathscr{V}_{3/2} = \mathrm{Pol}(\mathscr{V}_1) = \mathrm{Pol}(\mathscr{J}) = \mathrm{Pol}(\mathscr{F})$. Et la décidabilité de $\mathscr{V}_{3/2}$ résulte alors du Corollaire 3.8. □

Concernant le niveau 2, on dispose de l'énconcé suivant, dû à Pin et Straubing. $\underline{PJ}$ désigne la variété de monoïdes engendrée par l'ensemble $\{\mathscr{P}(M) \mid M \in \underline{J}\}$.

Théorème 3.9. *Soit L un langage de A^*. Les conditions suivantes sont équivalentes:*
(i) $L \in A^*\mathscr{V}_2$.
(ii) *L est dans l'algèbre de Boole engendrée par les langages $\mathscr{F}$-élémentaires.*
(iii) *Le monoïde syntactique de L est élément de $\underline{PJ}$.*

En d'autres termes, $\underline{V}_2 = \underline{PJ}$. Mais ni ces diverses caractérisations, ni d'autres travaux sans cesse plus poussés, n'ont pu permettre de dégager un algorithme de décision pour la variété $\mathscr{V}_2$.

Jusqu'à présent, seules les décidabilités des niveaux 0 et 1 étaient connues. On vient d'établir celles des niveaux $\frac{1}{2}$ et $\frac{3}{2}$. Straubing [21] a récemment démontré que si L est un langage sur un alphabet à deux lettres, alors on peut décider si L appartient au niveau 2. Weil [23] a également obtenu des résultats partiels sur ce niveau. Le cas du niveau 2 pour un alphabet fini quelconque, ainsi que ceux des niveaux supérieurs, demeurent à ce jour des problèmes ouverts.

Remerciements

Je tiens tout particulièrement à remercier Jean-Eric Pin pour ses conseils et ses encouragements.

Bibliographie

[1] J.-A. Brzozowski, Open problems about regular languages, dans: R.V. Book, ed., *Formal Language Theory, Perspectives and Open Problems* (Academic Press, New York, 1980) 23-47.

[2] S. Eilenberg, *Automata, Languages and Machines* (Academic Press, New York, Vol. A, 1974, Vol. B, 1976).
[3] K. Hashiguchi, Limitedness theorem on finite automata with distance functions, *J. Comput. System Sci.* **24** (1982) 233–244.
[4] K. Hashiguchi, Regular languages of star height one, *Inform. and Control* **53** (1982) 199–210.
[5] K. Hashiguchi, Representation theorems on regular languages, *J. Comput. System Sci.* **27** (1983) 101–115.
[6] K. Hashiguchi, Algorithms for determining relative star height and star height, *Inform and Comput.* **78** (1988) 124–169.
[7] M. Lothaire, *Combinatorics on Words*, Encyclopedia of Mathematics, Vol. 17 (Addison-Wesley, Reading, 1983).
[8] D. Perrin et J.-E. Pin, First order logic and star free sets, *J. Comput. System Sci.* **32** (1986) 393–406.
[9] J.-E. Pin, Variétés de langages et variétés de semigroupes, Thèse d'Etat, Paris, 1981.
[10] J.-E. Pin, Hiérarchies de concaténation, *Rairo Inform. Théor. Appl.* **18** (1984) 23–46.
[11] J.-E. Pin, Langages rationnels et reconnaissables, Cours Université de Naples, 1984, Rapport LITP 85-60.
[12] J.-E. Pin, *Variétés de Langages Formels* (Masson, Paris, 1984).
[13] J.-E. Pin *Varieties of Formal Languages* (Plenum, New York, 1986).
[14] J.-E. Pin et H. Straubing, Monoids of upper-triangular matrices, *Colloq. Math. Soc. János Bolyai* **39** (1981) 259–272.
[15] M.-P. Schützenberger, Sur certaines opérations de fermeture dans les langages rationnels, dans: *Istituto Nazionale di Alta Matematicà, Symposia Mathematica, Vol.* XV (1975) 245-253.
[16] M.-P. Schützenberger, Sur le produit de concatenation non ambigu, *Semigroup Forum* **13** (1976) 47–75.
[17] I. Simon, Hierarchies of events with dot-depth one, Thèse, Université de Waterloo, 1972.
[18] I. Simon, Piecewise testable events, dans: Lecture Notes in Computer Science, Vol. 33 (Springer, Berlin, 1975) 214–222.
[19] H. Straubing, Aperiodic homomorphisms and the concatenation product of recognizable sets, *J. Pure Appl. Algebra* **15** (1979) 319–327.
[20] H. Straubing, Finite semigroup varieties of the form $\underline{V} * \underline{D}$, *J. Pure Appl. Algebra* **36** (1985) 53–94.
[21] H. Straubing, Semigroups and languages of dot-depth 2, *Theoret. Comput. Sci.* **58** (1988) 351–378.
[22] W. Thomas, Classifying regular events in symbolic logic, *J. Comput. System Sci.* **25** (1982) 360–376.
[23] P. Weil, Inverse monoids of dot-depth 2, *Theoret. Comput. Sci.* **66** (1989) 233–245.