Hiérarchies de concaténation

Jean-Éric Pin

Current address (2008) : LIAFA, Université Paris VII and CNRS, Case 7014, 75205 Paris Cedex 13, France

Résumé. On sait qu'il existe une correspondance bijective entre les variétés de langages reconnaissables et les variétés de semigroupes finis. Dans cet article, je définis des hiérarchies de variétés de langages basées sur le produit de concaténation et je donne une description purement algébrique des hiérarchies de semigroupe correspondantes. On retrouve ainsi comme cas particuliers diverses hiérarchies déjà connues. La construction proposée repose sur le résultat suivant : tout langage reconnu par le produit de Schützenberger des monoïdes $M_0, \ldots, M_n$ est dans l'algèbre de Boole engendrée par les langages de la forme $L_{i_0}a_1L_{i_1}\cdots a_rL_{i_r}$ $(0 \leqslant i_0 < i_1 < \ldots < i_r \leqslant n)$ où les a_k sont des lettres et les L_{i_k} des langages reconnus par M_{i_k} $(0 \leqslant k \leqslant r)$. Enfin je montre un résultat général de décidabilité qui permet de retrouver un résultat de théorie des semigroupes : étant donné un semigroupe fini S et un entier n, on peut décider si S divise un produit en couronne de n demi-treillis.

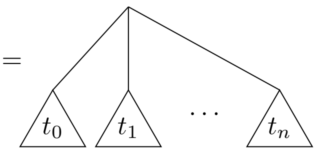
Abstract. As well-known, there exists a one-to-one correspondence between varieties of recognizable languages and varieties of finite semigroups. New hierarchies of varieties of languages (based on the concatenation product) are defined and an algebraic description of the corresponding hierarchies of varieties of semigroups is given. Various well-known hierarchies are obtained as particular cases. The construction is based on the following result : if a language L is recognized by the Schützenberger product of the monoids $M_0, \ldots, M_n$, then L belongs to the Boolean closure of the set of languages of the form $L_{i_0}a_1L_{i_1}\cdots a_rL_{i_r}$ $(0 \leqslant i_0 < i_1 < \ldots < i_r \leqslant n)$ where the a_k are letters and the L_{i_k} are recognized by M_{i_k} $(0 \leqslant k \leqslant r)$. Decidability and inclusion problems are also discussed.

Introduction

Le produit de concatenation est, avec l'étoile, l'opération la plus importante pour l'étude des langages reconnaissables. En particulier, Brzozowski [1] a proposé une hiérarchie des langages apériodiques (« star-free » en anglais) basée sur le produit de concaténation : le niveau 0 est constitué par les langages finis de A^+ et le niveau $n+1$ est l'algèbre de Boole engendrée par les langages de la forme $L_1 \cdots L_k$ avec $k \geqslant 0$, où chaque L_i est un langage de A^+ de niveau n. Brzozowski et Knast [2] on démontré que cette hiérarchie était infinie et Knast [5, 6] a récemment caractérisé les semigroupes syntactiques des langages de hauteur 1. Brzozowski a donné par ailleurs une version plus fine de sa hiérarchie en imposant des bornes supérieures à l'entier k de la construction précédente.

Straubing [14] a défini pour les langages de A^* une hiérarchie de concaténation voisine de celle de Brzozowski. A^* et l'ensemble vide sont de niveau 0 et le niveau $n+1$ est l'algèbre de Boole engendrée par les langages de la forme $L_0a_1L_1a_2\cdots a_kL_k$, $k \geqslant 0$, où les a_i sont des lettres et les L_i des langages de niveau n. Là encore on sait caractériser les langages de niveau 1, appelés aussi langages testables par morceaux : ce sont les langages dont le monoïde est $\mathcal{J}$-trivial [13] (cette classe de langages possède sa propre hiérarchie, proposée par Simon [12]). En revanche, on ne sait toujours pas décider si un langage est de niveau 2 dans la hiérarchie de Straubing, bien que l'on connaisse diverses caractérisations des monoïdes syntactiques correspondant à ces langages [10].

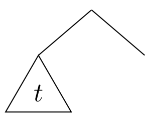
Le but de cet article est de montrer que ces diverses hiérarchies sont des cas particuliers d'une construction plus générale, obtenue en associant des variétés de langages non plus à des entiers, mais à des arbres, selon le procédé suivant. Comme pour les constructions précédentes, on se donne au départ une variété de langages qui est associée par définition à l'arbre réduit à un point. Ensuite, on associe à l'arbre

$t =$ 

l'algèbre de Boole engendrée par les langages de la forme $L_{i_0}a_1L_{i_1}\cdots a_rL_{i_r}$ avec $(0 \leqslant i_0 < i_1 < \ldots < i_r \leqslant n)$ où les a_i sont des lettres et où, pour $0 \leqslant j \leqslant k$, L_{i_j} est élément de la variété de langages associée à l'arbre t_{i_j}.

Il reste à donner l'interprétation algébrique de cette construction, ce qui peut être fait en utilisant le produit de Schützenberger de $(n+1)$ monoïdes $M_0, \ldots, M_n$. Cette opération, introduite par Schützenberger dans le cas $n = 1$ (cf. [4]), puis généralisée par Straubing [15], peut paraître à première vue assez artificielle. C'est en fait un cas particulier d'une construction très naturelle [9]. Je démontre (section 2) que le produit de Schützenberger est, en un certain sens, parfaitement adaptée à l'opération (sur les langages) $(L_0, \ldots, L_n) \to L_0a_1L_1\cdots a_nL_n$ où les a_i sont des lettres. Ce résultat permet de construire, sans référence aux langages, des hiérarchies de variétés de semigroupes (monoïdes) correspondant, via le théorème d'Eilenberg, aux hiérarchies de langages précédemment construites. Autrement dit, partant d'une variété de semigroupes (ou de monoïdes) $\mathbf{V}$, on associe à chaque arbre t — resp. à chaque ensemble d'arbres L — une variété de semigroupes $\Diamond_t(\mathbf{V})$ $(\Diamond_L(\mathbf{V}))$.

Je montre que les variétés de semigroupes ou de monoïdes correspondant aux hiérarchies de Brzozowski, Straubing et Simon s'obtiennent toutes par ce procédé. On retrouve également ainsi d'autres variétés connues telles que la variété $\mathbf{R}$ des monoïdes $\mathcal{R}$-triviaux. je démontre également que l'opération qui à un arbre t associe l'arbre



est liée au produit semi-direct des semigroupes (Theorème 3.7)

Parmi les nombreux problèmes que posent ces hiérarchies, on peut en dégager trois principaux :

1 Comparaison des variétés à l'intérieur d'une hiérarchie

C'est la généralisation du problème « la hiérarchie de Brzozowski est-elle infinie ? », résolu positivement par Brzozowski et Knast. De façon précise, le problème consiste à comparer entre elles les différentes variétés $\diamondsuit_t(\mathbf{V})$ (ou même $\diamondsuit_L(\mathbf{V})$). La fin de la section 3 regroupe quelques résultats partiels et une conjecture sur ce problème.

2 Problème de décidabilité

On dit qu'une variété de semigroupes $\mathbf{V}$ est décidable s'il existe un algorithme qui permet de tester si un semigroupe fini est ou n'est pas dans $\mathbf{V}$. Le problème est de savoir si les variétés étudiées dans cet article sont décidables. La question abordée dans la section 4 où il est montré en particulier que toutes les variétés de la hiérarchie construite à partir de la variété triviale sont décidables. Ce résultat a une conséquence en théorie des semigroupes : étant donné un entier n, on peut décider si un semigroupe fini divise un produit en couronne de n demi-treillis.

3 Équations des variétés

Cette question, liée à la précédente, n'est pas abordée dans cette article et fera l'objet de publications ultérieures.

1 Rappels et notations

Les références de base sont Eilenberg [4] and Lallement [7] dont j'adopterai la plupart les notations. Si S est un semigroupe, on note $E(S)$ l'ensemble des idempotents de S.

Une variété de semigroupes (monoïdes) est une classe de semigroupes (monoïdes) finis $\mathbf{V}$ telle que :

(1) Si $S \in \mathbf{V}$ et si T est un sous-semigroupe de S, alors $T \in \mathbf{V}$.

(2) Si $S \in \mathbf{V}$ et si T est un quotient de S, then $T \in \mathbf{V}$.

(3) Si $(S_i)_{i\in I}$ est une famille d'éléments de $\mathbf{V}$, le produit direct $\prod_{i\in I} S_i$ est dans $\mathbf{V}$.

Eilenberg a montré l'existence d'une bijection entre les variétés de semigroupes (monoïdes) et certaines classes de langages, les $+$-variétés ($*$-variétés) de langages. De façon formelle, une $+$-variété de langages $\mathcal{V}$ associe à chaque alphabet fini A un ensemble $A^+\mathcal{V}$ de langages reconnaissables de A^+ tels que :

(1) Pour tout alphabet A, $A^+\mathcal{V}$ est fermée pour les opérations booléennes finies.

(2) Pour tout alphabet A, si $a \in A$ et $L \in A^+\mathcal{V}$, alors $a^{-1}L, La^{-1} \in A^+\mathcal{V}$.

(3) Pour tout morphisme de semigroupes $\varphi : A^+ \to B^+$, $L \in A$, $L \in B^+\mathcal{V}$ entraîne $L\varphi^{-1} \in A^+\mathcal{V}$.

La définition des $*$-variétés s'obtient en remplaçant $+$ par $*$ et semigroupe par monoïde.

Les hiérarchies de concaténation fournissent des exemples particulièrement intéressants de variétés de langages. La hiérarchie de Brzozowski [1] $\mathcal{B}_n$ est définie comme suit : pour tout alphabet A, $A^+\mathcal{B}_{n+1}$ est l'algèbre de Boole engendrée par les langages de la forme $L_0 \cdots L_k$ avec $k \geqslant 0$ et $L_0, \ldots, L_k \in A^+\mathcal{B}_\backslash$.

On peut définir une sous-hiérarchie à l'intérieur de $\mathcal{B}_{n+1}$ en considérant pour chaque entier $r \geqslant 0$ l'algèbre de Boole $A^+\mathcal{B}_{n+1,r}$ engendrée par les langages de la forme $L_0 \cdots L_k$ avec $L_0, \ldots, L_k \in A^+\mathcal{B}_n$ et $0 \leqslant k \leqslant r$. On définit ainsi une suite de variétés $\mathcal{B}_{n,r}$. On notera que $\mathcal{B}_{n+1,0} = \mathcal{B}_n$ pour tout $n \geqslant 0$. On montre également que $\mathcal{B}_{1,2k} = \mathcal{B}_{1,2k+1}$ pour tout $k > 0$.

La hiérarchie de Straubing [14] $\mathcal{V}_n$ est la suite des $*$-variétés ainsi définie : pour tout alphabet A, $A^*\mathcal{V}_0$ est constitué de $\emptyset$ et de A^* et $A^+\mathcal{V}_{n+1}$ est l'algèbre de Boole engendrée par les langages de la forme $L_0a_1L_1 \cdots a_kL_k$ avec $k > 0$ et $L_0, \ldots, L_k \in A^+\mathcal{V}_n$ et $a_1, \ldots, a_k \in A$.

On démontre que les $\mathcal{V}_n$ (resp. $\mathcal{B}_n, \mathcal{B}_{n,k}$) sont effectivement des $*$-variétés ($+$-variétés). Les variétés de monoïdes (resp. semigroupes) correspondantes sont notées $\mathbf{V}_n$ (resp. $\mathbf{B}_n, \mathbf{B}_{n,k}$). Le problème majeur concernant ces variétés demeure leur caractérisation algébrique. Rappelons les résultats conns à ce jour.

Théorème 1.1 (Simon [13]) *On a $\mathbf{V}_1 = \mathbf{J}$, la variété des monoïdes $\mathcal{J}$-triviaux. En particulier, un monoïde M est dans $\mathbf{J}$ si et seulement si, pour tout $x, y \in M$, $(xy)^n = (yx)^n$ et $x^n = x^{n+1}$ avec $n = \mathrm{Card}(M)$.*

Si $\mathbf{V}$ est une variété de monoïdes, on note $\mathbf{LV}$ la variété de semigroupes *locale* associée à $\mathbf{V}$:

$$\mathbf{LV} = \{S \mid eSe \in \mathbf{V} \text{ pour tout } e \in E(S)\}.$$

On note $\mathbf{J_1}$ la variété des monoïdes idempotents et commutatifs (appelés aussi demi-treillis).

Théorème 1.2 ([3]) *On a $\mathbf{B}_{1,2} = \mathbf{LJ_1}$.*

Les langages de $\mathbf{B}_{1,2}$ sont appelés localement testables. Pour chaque alphabet A, $A^+\mathbf{B}_{1,2}$ est l'algèbre de Boole engendrée par les langages de la forme uA^*, A^*v et A^*wA^* où u, v et w sont des mots de A^+.

Simon avait conjecturé l'égalité $B_1 = \mathbf{LJ}$. En fait, on a bien $B_1 \subseteq \mathbf{LJ}$ mais Knast [5, 6] a montré que l'inclusion était stricte :

Théorème 1.3 *Un semigroupe S est dans $\mathbf{B}_1$ si et seulement si il satisfait la condition suivante :*

(K) *Il existe $m > 0$ tel que pour tout $e_1, e_2 \in E(S)$, pour tout $x, y, u, v \in S$,*

$$(e_1xe_2y)^m e_1xe_2ve_1(ue_2ve_1)^m = (e_1xe_2y)^m e_1(ue_2ve_1)^m$$

Si M est un monoïde, on note $\mathcal{P}(M)$ le monoïde des parties de M, muni du produit usuel des parties :

Théorème 1.4 ([10]) *On a $\mathbf{V}_2 = \mathbf{PJ}$, la variété engendrée par les monoïdes $\mathcal{P}(M)$ où $M \in \mathbf{J}$.*

Malheureusement, ni cette description de $\mathbf{V}_2$ ni les autres caractérisation connues [10] ne permettent de résoudre le problème suivant : peut-on décider si un monoïde fini M est dans $\mathbf{V}_2$.

On ne connaît à ce jour aucun résultat sur les variétés $\mathbf{B}_n$ pour $n \geqslant 2$ et $\mathbf{V}$ pour $n \geqslant 3$, hormis les résultats généraux suivants :

Théorème 1.5 ([2, 15]) *La hiérarchie* $\mathbf{B}_n$ *est infinie.*

Théorème 1.6 ([14]) *On a* $\mathbf{B}_n = \mathbf{V}_n * \mathbf{LI}$ *pour tout* $n > 0$. *En particulier la hiérarchie* $\mathbf{V}_n$ *est infinie.*

Dans ce dernier énoncé, la notation $\mathbf{V}_n * \mathbf{LI}$ désigne la variété engendrée par ls produits semi-directs $M * S$ d'un monoïde $M \in \mathbf{V}_n$ et d'un semigroupe $S \in \mathbf{LI}$. **LI** est la variété des semigroupes localement triviaux, i.e. $S \in \mathbf{LI}$ si et seulement si $eSe = e$ pour tout $e \in E(S)$.

2 Le produit de Schützenberger

Si S est un semigroupe, on note $\mathcal{P}(S)$ le semi-anneau des parties de S, muni de l'union come addition et du produit des parties comme multiplication. On note S^1 le monoïde ainsi défini :

$$S^1 = \begin{cases} S & \text{si } S \text{ est un monoïde} \\ S \cup \{1\} & \text{si } S \text{ n'est pas un monoïde (1 est évidemment alors} \\ & \qquad\qquad\qquad\qquad \text{l'élément neutre de } S^1) \end{cases}$$

Soient $S_1, \ldots, S_n$ des semigroupes. Le *produit de Schützenberger* de $S_1, \ldots, S_n$, noté $\diamondsuit_n(S_1, \ldots, S_n)$ est le semigroupe des matrices $n \times n$ à coefficients dans $\mathcal{P}(S_1^1 \times \cdots \times S_n^1)$ de la forme $p = (p_{i,j})_{1 \leqslant i,j \leqslant n}$ et vérifiant les trois conditions suivantes :

(1) $p_{i,j} = \emptyset$ si $i > j$.
(2) $p_{i,i} = \{(1, \ldots, 1, s_i, 1, \ldots, 1)\}$ pour un certain $s_i \in S_i$.
(3) $p_{i,j} \subseteq \{(s_1, \ldots, s_n) \in S_1^1 \times \cdots \times S_1^n \mid s_1 = \cdots = s_{i-1} = 1 = s_{j+1} = \ldots = s_n\}$.

Il est à noter que le produit de Schützenberger n'est pas associatif, c'est-à-dire qu'en général les semigroupes $\diamondsuit_2(\diamondsuit_2(S_1, S_2), S_3)$, $\diamondsuit_3(S_1, S_2, S_3)$ et $\diamondsuit_2(S_1, \diamondsuit_2(S_2, S_3))$ sont distincts.

Straubing [15] a montré que si les langages L_i ($0 \leqslant i \leqslant n$) de A^* (resp. de A^+) sont reconnus par des monoïdes M_i (resp. des semigroupes S_i), alors le langage $L_0 a_1 L_1 \cdots a_n L_n$, où les a_i sont des lettres, est reconnu par $\diamondsuit_{n+1}(M_0, \ldots, M_n)$. Il est facile de vérifier que si $0 \leqslant i_0 < \ldots < i_r \leqslant n$, alors $\diamondsuit_{r+1}(M_{i_0}, \ldots, M_{i_r})$ est un sous-monoïde de $\diamondsuit_{n+1}(M_0, \ldots, M_n)$. Il en résulte que le monoïde $\diamondsuit_{n+1}(M_0, \ldots, M_n)$ reconnaît tous les langages de la forme $L_{i_0} a_1 L_{i_1} \cdots a_r L_{i_r}$ où L_{i_k} est reconnu par M_{i_k}.

Le but de cette section est d'établir la réciproque de ce résultat. Le cas $n = 1$ a été traité par Reutenauer [11] et la preuve qui suit s'inspire en partie de ses arguments.

Théorème 2.1 *Si un langage* $L \subseteq A^*$ *est reconnu par* $\diamondsuit_{n+1}(M_0, \ldots, M_n)$, *alors* L *est dans l'algèbre de Boole engendrée par les langages de la forme* $L_{i_0} a_1 L_{i_1} \cdots a_r L_{i_r}$ $(0 \leqslant i_0 < \ldots < i_r \leqslant n)$ *où les* a_k *sont des lettres et où* $L_{i_0}, \ldots, L_{i_r}$ *sont des langages reconnus respectivement par* $M_{i_0}, \ldots M_{i_r}$.

Démonstration. Soit $\mu : A^* \to \diamondsuit_{n+1}(M_0, \ldots, M_n)$ le morphisme reconnaissant L et soient $(\mu_{i,j})_{0 \leqslant i,j \leqslant n}$ les composantes de μ considérées comme *applications* de A^* dans $\mathcal{P}(M_0, \ldots, M_n)$.

Lemme 2.2 *Soit $P \subset \diamondsuit_{n+1}(M_0, \ldots, M_n)$. On a la formule :*

$$P\mu^{-1} = \bigcup_{m \in P} \bigcap_{0 \leqslant i,j \leqslant n} m_{i,j}\mu_{i,j}^{-1}$$

C'est immédiat. $\square$

Lemme 2.3 *On a pour tout $i, j \in \{0, 1, \ldots, n\}$*

$$(a_1 \cdots a_r)\mu_{i,j} = \sum_{i=i_0 \leqslant i_1 \leqslant \ldots \leqslant i_{r-1} \leqslant i_r = j} a_1\mu_{i_0,i_1} a_2\mu_{i_1,i_2} \ \cdots \ a_r\mu_{i_{r-1}i_r}$$

La formule est évidente pour $r = 1$. Supposons-là acquise jusqu'au rang $(r-1)$. Il vient :

$$\begin{aligned}
(a_1 \cdots a_r)\mu_{i,j} &= \sum_{k=0}^{n} (a_1 \cdots a_{r-1})\mu_{i,k}\, a_r\mu_{k,j} \\
&= \sum_{k=0}^{n} \sum_{i=i_0 \leqslant i_1 \leqslant \ldots \leqslant i_{r-1}=k} a_1\mu_{i_0,i_1} \ \cdots \ a_{r-1}\mu_{i_{r-2},i_{r-1}} a_r\mu_{k_j} \\
&= \sum_{k=0}^{n} \sum_{i=i_0 \leqslant i_1 \leqslant \ldots \leqslant i_{r-1}=k} a_1\mu_{i_0,i_1} \ \cdots \ a_{r-1}\mu_{i_{r-2},i_{r-1}} a_r\mu_{i_{r-1},j}
\end{aligned}$$

d'où le résultat. $\square$

Désormais, on fixe i et j. D'après le lemme 2.2, il suffit d'établir que le langage $m_{i,j}\mu_{i,j}^{-1}$ s'écrit comme combinaison booléenne de langages de la forme $L_r a_{r+1} \ \cdots \ a_s L_s$ avec $a_k \in A$ et L_k reconnu par M_k pour $r \leqslant k \leqslant s$. Si $i > j$, on a $m_{i,j} = \emptyset$ d'où $m_{i,j}\mu_{i,j}^{-1} = \emptyset$ et le résultat est évident. Si $i = j$, $m_{i,j}$ s'identifie à un élément de M_i et $m_{i,j}\mu_{i,j}^{-1}$ est un langage reconnu pas M_i. On peut donc supposer désormais que $i < j$.

Soit, pour $1 \leqslant k \leqslant j-i+1$, R_k l'ensemble des suites de lettres $(a_1, \ldots, a_k)$ et S_k l'ensemble des suites $(i_0, i_1, \ldots, i_k)$ telles que $i = i_0 < i_1 < \ldots < i_k = j$. Pour $r_k = (a_1, \ldots, a_k) \in R_k$, on définit une application : $\sigma_{r_k} : A^* \to \mathcal{P}(A^* \times \ \cdots \ \times A^*)$ par :

$$u\sigma_{r_k} = \{(u_0, u_1, \ldots, u_k) \in A^* \times \ \cdots \ \times A^* \mid u_0 a_1 u_1 \ \cdots \ a_k u_k = u\}$$

Pour $s_k = (i_0, i_1, \ldots, i_k) \in S_k$ et $r_k = (a_1, \ldots, a_k) \in R_k$, on définit une application :

$$\tau_{r_k,s_k} : (A^* \times \ \cdots \ \times A^*) \to \mathcal{P}(M_0 \times M_1 \ \cdots \ \times M_n)$$

par

$$(u_0, \ldots, u_k)\tau_{r_k,s_k} = u_0\mu_{i_0,i_0} a_1 u_1 \mu_{i_1,i_1} \ \cdots \ u_k\mu_{i_k,i_k}$$

Lemme 2.4 *On a la formule*

$$u\mu_{i,j} = \sum_{1 \leqslant k \leqslant j-i+1} \sum_{r_k \in R_k} \sum_{(u_0,\ldots,u_k) \in u\sigma_{r_k}} \sum_{s_k \in S_k} (u_0, \ldots, u_k)\tau_{r_k,s_k}$$

Démonstration. La formule s'obtient immédiatement à partir du lemme 2.3 en regroupant les produits de la forme

$$a_s\mu_{i_j,i_j}a_{s+1}\mu_{i_j,i_j}\ \cdots\ a_{s+t}\mu_{i_j,i_j}$$

en $(a_sa_{s+1}\cdots a_{s+t})\mu_{i_j,i_j}$. $\square$

Posons, pour $1\leqslant k\leqslant j-i+1$, $r_k\in R_k$, $s_k\in S_k$ et $F\in\mathcal{P}(\mathcal{P}(M_0\times\ \cdots\ M_n))$:

$$L_{F,r_k,s_k}=\{u\in A^*\mid\{(u_0,\ldots,u_k)\tau_{r_k,s_k}\mid(u_0,\ldots,u_k)\in u\sigma_{r_k}\}=F\}$$

Lemme 2.5 *Pour tout* $m\in\mathcal{P}(M_0\times\ \cdots\ M_n)$, $m\mu_{i,j}^{-1}$ *est dans l'algèbre de Boole des langages de la forme* L_{F,r_k,s_k}.

Démonstration. Posons $u\gamma_{r_k,s_k}=\sum_{(u_0,\ldots,u_k)\in u\sigma_{r_k}}(u_0,\ldots,u_k)\tau_{r_k,s_k}$. En intervertissant l'ordre des deux dernières sommations dans le lemme 2.4, on obtient :

$$u\mu_{i,j}=\sum_{1\leqslant k\leqslant j-i+1}\ \sum_{r_k\in R_k}\ \sum_{s_k\in S_k}u\gamma_{r_k,s_k}$$

On en déduit la formule :

$$m\mu_{i,j}^{-1}=\bigcup_{\sum m_{r_k,s_k}=m}\ \bigcap\ m_{r_k,s_k}\gamma_{r_k,s_k}^{-1}\tag{1}$$

La sommation et l'intersection étant étendues à $1\leqslant k\leqslant j-i+1$, $r_k\in R_k$ et $s_k\in S_k$. On a d'autre part, si $m\in\mathcal{P}(M_0\times\ \cdots\ M_n)$:

$$m\gamma_{r_k,s_k}^{-1}=\bigcup_{\sum_{f\in F}f=m}L_{F,r_k,s_k}\tag{2}$$

Démontrons (2) en omettant, pour alléger les notations, les indices r_k et s_k :

$$\sum_{(u_0,\ldots,u_k)\in u\sigma}(u_0,\ldots,u_k)\tau=m$$

et donc en posant :

$$F=\{(u_0,\ldots,u_k)\tau\mid(u_0,\ldots,u_k)\in u\sigma\},$$

on a à la fois

$$\sum_{f\in F}f=m\quad\text{et}\quad u\in L_F.$$

Réciproquement, si $u\in L_F$ avec $\sum_{f\in F}f=m$, on a :

$$u\gamma=\sum_{(u_0,\ldots,u_k)\in u\sigma}(u_0,\ldots,u_k)\tau=\sum_{f\in F}f=m,$$

ce qui établit la formule (2). Le lemme découle alors immédiatement de (1) et (2). $\square$

Les indices k, r_k, s_k et F étant maintenant fixés, il s'agit d'étudier le langage $L_F=\{u\in A^*\mid\{(u_0,\ldots,u_k)\tau\mid(u_0,\ldots,u_k)\in u\sigma\}=F\}$.

Lemme 2.6 *On a la formule* $L_F = \bigcap_{f \in F} L_f \setminus \bigcup_{f \notin F} L_f$, *où :*

$$L_f = \{u \in A^* \mid \exists (u_0, \dots, u_k) \in u\sigma \ (u_0, \dots, u_k)\tau = f\}.$$

Démonstration. Si $u \in L_F$ et si $f \in F$, il existe $(u_0, u_1, \dots, u_k) \in u\sigma$ tel que $(u_0, \dots, u_k)\tau = f$ et donc $u \in L_f$. Si $f \notin F$, on a pour tout $(u_0, \dots, u_k) \in u\sigma$, $u_0, \dots, u_k)\tau \neq f$ et donc $u \notin L_f$.

Réciproquement, si $u \in \bigcap_{f \in F} L_f \setminus \bigcup_{f \notin F} L_f$, on a :

$$F \subset \{(u_0, \dots, u_k)\tau \mid (u_0, \dots, u_k) \in u\sigma\} \subset (F^c)^c = F$$

où F^c désigne le complémentaire de F dans $\mathcal{P}(\mathcal{P}(M_0 \times \ \cdots \ M_n))$, ce qui établit le lemme. $\square$

Désormais, on fixe $f \in \mathcal{P}(M_0 \times \ \cdots \ M_n)$. Rappelons la définition de L_f :

$$L_f = \{u \in A^* \mid \exists u_0, \dots, u_k \in A^* \ u = u_0 a_1 u_1 \ \cdots \ a_k u_k \text{ et } u_0 \mu_{i_0, i_0} a_1 \mu_{i_0, i_1} \ \cdots \ u_k \mu_{i_k, i_k} = f\}.$$

D'après la définition du produit de Schützenberger de n monoïdes, on peut toujours suppoer que f est contenu dans l'ensemble des $(n+1)$-uplets :

$$(m_0, \dots, m_n) \in M_0 \times \ \cdots \ \times M_n \text{ tels que } m_0 = \ldots = m_{i-1} = 1 \text{ et } m_{j+1} = \ldots = m_n = 1.$$

On pose donc :

$$f = \{(1, \dots, 1, m_{i,r}, m_{i+1,r}, \dots, m_{j,r}, 1, \dots, 1) \mid 1 \leqslant r \leqslant |f|\}$$

et, pour $1 \leqslant s \leqslant k$:

$$a_s \mu_{i_{s-1}, i_s} = \{(1, \dots, 1, m_{i_s, 1}(s), \dots, m_{i_s, r}(s), 1, \dots, 1) \mid 1 \leqslant r \leqslant r_s\}$$

où $r_s = |a_s \mu_{i_{s-1}, i_s}|$.

Si $u \in L_f$, il exise par définition des mots $u_0, \dots, u_k \in A^*$ tels que $u = u_0 a_1 u_1 \cdots a_k u_k$ et :

$$f = \{(u_0 \mu_{i_0, i_0} m_{i_0, t_0}(1), m_{i_0+1, t_0}(1), \dots, m_{i_1, t_0}(1) u_1 \mu_{i_1, i_1} m_{i_1, t_1}(2), m_{i_1+1, t_1}(2), \dots, m_{i_k, t_k}(k) u_k \mu_{i_k, i_k}) \mid 1 \leqslant t_j \leqslant r_j, 1 \leqslant j \leqslant k-1\}.$$

Il en résulte qu'il existe deux applications φ et ψ :

$$\varphi : \{1, \dots, |f|\} \to \prod_{0 \leqslant j \leqslant k} \{1, \dots, r_j\} = T,$$

$$\psi : T = \prod_{0 \leqslant j \leqslant k} \{1, \dots, r_j\} \to \{1, \dots, |f|\},$$

telles que pour tout $p \in \{1, \dots, |f|\}$, on ait, en posant

$$(3) \qquad p\varphi = (p\varphi_0, \dots, p\varphi_k)$$

$$(m_{i,p}, \ldots, m_{j,p}) = (u_0\mu_{i_0,i_0}m_{i_0,p\varphi_0}(1), m_{i_0+1,p\varphi_0}(1), \ldots,$$
$$m_{i_1,p\varphi_0}(1)u_1\mu_{i_1,i_1}m_{i_1,p\varphi_1}(2), m_{i_1+1,p\varphi_1}(2), \ldots, m_{i_k,p\varphi_k}(k)u_k\mu_{i_k,i_k})$$

et pour tout $t = (t_1, \ldots, t_k) \in T$:

$$(4) \qquad (u_0\mu_{i_0,i_0}m_{i_0,t_0}(1), \ldots, m_{i_k,t_k}(k)u_k\mu_{i_k,i_k}) = (m_{i,t\psi}, m_{i+1,t\psi}, \ldots, m_{j,t\psi}).$$

Posons, pour $0 \leqslant j \leqslant k$, $t \in \prod_{1\leqslant j\leqslant k-1}\{1, \ldots, r_j\}$, $p \in \{1, \ldots, s\}$,

$$L_{j,t,p} = \big((m_{i_j,t_{j-1}}(j))^{-1}m_{i_j,p}(m_{i_j,t_j}(j+1))^{-1}\big)\mu^{-1}_{i_j,i_j}$$

avec la convention $(m_{i_j,t_{j-1}}(j))^{-1} = 1$ si $j = 0$ et $(m_{i_j,t_j}(j+1))^{-1} = 1$ si $j = k$.

Comme μ_{i_j,i_j} est un morphisme $A^* \to M_{i_j}$, le langage $L(j, t, p)$ est reconnu par M_{i_j}. Il résulte d'autre part des formules (3) et (4) la relation suivante : pour $0 \leqslant j \leqslant k$,

$$(5) \qquad u_j \in \Big(\bigcap_{p\in\{1,\ldots,|f|\}} L(j, p\varphi, p)\Big) \cap \Big(\bigcap_{t\in T} L(j, t, t\psi)\Big)$$

Soit $\mathcal{E}$ l'ensemble des applications de $\{1, \ldots, |f|\}$ dans T et soit $\mathcal{F}$ l'ensemble des applications de T dans $\{1, \ldots, |f|\}$. On va établir la formule

$$(6) \qquad L_f = \bigcup_{\substack{\varphi\in\mathcal{E}\\ \psi\in\mathcal{F}}} K_0(\varphi, \psi)a_1 \ \cdots \ a_kK_k(\varphi, \psi)$$

avec, pour $0 \leqslant j \leqslant k$,

$$K_j(\varphi, \psi) = \Big(\bigcap_{p\in\{1,\ldots,|f|\}} L(j, p\varphi, p)\Big) \cap \Big(\bigcap_{t\in T} L(j, t, t\psi)\Big)$$

L'inclusion de gauche à droite résulte de la relation (5) et de la définition de L_f. Réciproquement, soit u un élément du membre droit de (6). Il existe alors $\varphi \in \mathcal{E}$ et $\psi \in \mathcal{F}$ et $u_0, \ldots, u_k \in A^*$ tels que pour tout $j \in \{0, \ldots, k\}$, $u_j \in K_j(\varphi, \psi)$. On en déduit que, pour tout $p \in \{1, \ldots |f|\}$ (resp. pour tout $t \in T$) la formule (3) [resp. (4)] est satisfaite. Il en découle, en remontant encore les calculs, que $u \in L_f$, et la formule (6) est établie.

Comme chaque $L(, j, t, p)$ est reconnu par M_{i_j}, il en va de même pour $K_j(\varphi, \psi)$ et la formule (6) permet de conclure la preuve du théorème. □

Dans le cas où les langages sont des parties d'un semigroupe libre, on démontre le la même façon le résultat suivant :

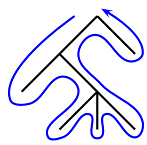
Théorème 2.7 *Si un langage $L \subset A^+$ est reconnu par $\diamondsuit_{n+1}(S_0, \ldots, S_n)$, alors L est dans l'algèbre de Boole engendrée par les langages de la forme $L_{i_0}a_1 \cdots a_rL_{i_r}$ $(0 \leqslant i_0 < i_1 < \ldots < i_r \leqslant n)$, où les a_k sont des lettres et où $L_{i_0}, \ldots, L_{i_r}$ sont des langages reconnus respectivement par $S_{i_0}, \ldots, S_{i_r}$.*

Les conséquences des théorèmes 2.1 et 2.7 seront examinées dans la section suivante.

3 Hiérarchies de concaténation

Dans cette section, je construis des hiérarchies de variétés qui contiennent en particulier les hiérarchies proposées par Brzozowski [1], Simon [12] et Straubing [14].

On note P l'ensemble des arbres (ou mots bien parenthésés) sur l'alphabet $\{a, \bar{a}\}$. De façon formelle, P est l'ensemble des mots de $\{a, \bar{a}\}^*$ congrus à 1 dans la congruence engendrée par la relation $a\bar{a} = 1$. De façon intuitive, les mots de P sont obtenus par le procédé suivant : on dessine un arbre (au sens naïf du terme) et on le parcourt en partant de la racine (suivant le sens trigonométrique) en codant a pour une descente et $\bar{a}$ pour une montée. Par exemple



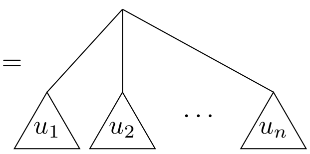
est codé par $aa\bar{a}aa\bar{a}a\bar{a}a\bar{a}a\bar{a}\bar{a}\bar{a}a\bar{a}$.

Le nombre de feuilles d'un mot u de $\{a, \bar{a}\}^*$, noté $d(u)$, est par définition le ombre d'occurrences du facteur $a\bar{a}$ dans u. Dans le cas des arbres, cette définition est bien conforme à l'intuition.

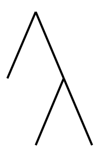
Je rappelle ci-dessous quelques propriétés classiques des arbres que j'utiliserai par la suite :

(1) tout arbre u se factorise de façon unique en $u = au_1\bar{a}au_2\bar{a} \;\cdots\; au_n\bar{a}$ où $n \geqslant 0$ et où les u_i sont des arbres. On a alors $d(u) = \sum_{1 \leqslant i \leqslant n} d(u_i)$. L'interprétation intuitive de cette propriété est illustrée par le schéma suivant :

$u =$



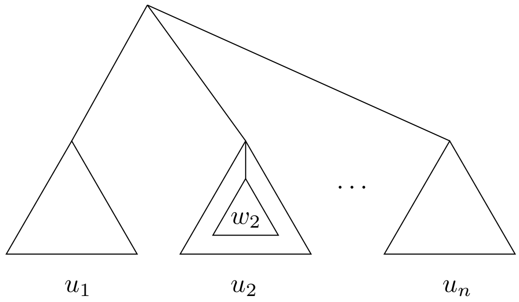
(2) Soit u un arbre et soit $u = u_1au_2\bar{a}u_3$ une factorisation de u. On dit que les occurrences de a et $\bar{a}$ définies par cette factorisation sont *liées* si u_2 st un arbre. On démontre que toute occurrence de a dans u est liée à une *unique* occurrence de $\bar{a}$. L'interprétation intuitive de cette propriété est très simple : les occurrences de a et $\bar{a}$ sont liées si et seulement si elles codent respectivement la descente et la montée sur une même arête de l'arbre. Par exemple,



est codé par $\underbrace{a\bar{a}}\,\underbrace{a\,\underbrace{a\bar{a}}\,\underbrace{a\bar{a}}\,\bar{a}}$, où les occurrences liées sont indiquées par des crochets.

(3) Si $u = au_1\bar{a}au_2\bar{a} \;\cdots\; au_n\bar{a}$ est la factorisation de l'arbre u décrite en (1) et si $u = w_1aw_2\bar{a}w_3$ avec $w_2 \in P$, alors l'arbre $aw_2\bar{a}$ est facteur de l'un

des mots $au_i\bar{a}$ $(1 \leqslant i \leqslant n)$. L'interprétation intuitive est illustrée par le schéma suivant :

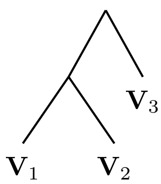


À chaque arbre u et à chaque suite $\mathbf{V}_1, \ldots, \mathbf{V}_{d(u)}$ de variétés de semigroupes (resp. de monoïdes), on associe une variété de semigroupes (monoïdes) $\diamondsuit_u(\mathbf{V}_1, \ldots, \mathbf{V}_{d(u)})$ définie récursivement par :

(a) $\diamondsuit_1(\mathbf{V}) = \mathbf{V}$ pour toute variété $\mathbf{V}$.

(b) Si $u = au_1\bar{a}au_2\bar{a} \, \cdots \, au_n\bar{a}$ avec $n \geqslant 0$ et $u_1, \ldots, u_n \in P$, $\diamondsuit_u(\mathbf{V}_1, \ldots, \mathbf{V}_{d(u)})$ est la variété engendrée par les semigroupes de la forme $\diamondsuit_n(S_1, \ldots, S_n)$ avec $S_1 \in \diamondsuit_{u_1}(\mathbf{V}_1, \ldots, \mathbf{V}_{d(u)})$, $\ldots$, $S_n \in \diamondsuit_u(\mathbf{V}_{d(u_1)+\ldots+d(u_{n-1})+1}, \ldots, \mathbf{V}_{d(u_1)+\ldots+d(u_n)})$.

Exemple. Soit $u = aa\bar{a}a\bar{a}\bar{a}a\bar{a}$ et $\mathbf{V}_1$, $\mathbf{V}_2$, $\mathbf{V}_3$ trois variétés de semigroupes. La variété $\diamondsuit_u(\mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3)$ est la variété engendrée par les semigroupes $\diamondsuit_2(S, S_3)$ où $S_3 \in \mathbf{V}_3$ et où S est dans la variété engendrée par les semigroupes $\diamondsuit_2(S_1, S_2)$ tels que $S_1 \in \mathbf{V}_1$ et $S_2 \in \mathbf{V}_2$.



On observera que $S \in \diamondsuit_u(\mathbf{V}_1, \ldots, \mathbf{V}_{d(u)})$ si et seulement si S divise un semigroupe de la forme $\diamondsuit_n(S_1, \ldots, S_n)$ avec $S_1 \in \diamondsuit_{u_1}(\mathbf{V}_1, \ldots, \mathbf{V}_{d(u)})$, $\ldots$, $S_n \in \diamondsuit_u(\mathbf{V}_{d(u_1)+\ldots+d(u_{n-1})+1}, \ldots, \mathbf{V}_{d(u_1)+\ldots+d(u_n)})$. Ceci résulte de la formule suivante, qui se démontre facilement : pour tout semigroupe $S_{i,j}$, $1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant r$:

$$\diamondsuit_n(S_{1,1}, \ldots, S_{n,1}) \times \, \cdots \, \times \diamondsuit_n(S_{1,r}, \ldots, S_{n,r})$$

divise

$$\diamondsuit_n(S_{1,1} \times \, \cdots \, \times S_{1,r}, \ldots, S_{n,1} \times \, \cdots \, \times S_{n,r})$$

Lorsque $\mathbf{V}_1 = \ldots = \mathbf{V}_{d(u)} = \mathbf{V}$, on note simplement $\diamondsuit_u(\mathbf{V})$ la variété $\diamondsuit_u(\mathbf{V}_1, \ldots, \mathbf{V}_{d(u)})$. Plus généralement, si L est un langage contenu dans P, on note $\diamondsuit_L(\mathbf{V})$ la plus petite variété contenant les variétés $\diamondsuit_u(\mathbf{V})$ avec $u \in L$.

Le théorème ci-dessous permet, par récurrence, de décrire les langages associés aux variétés $\diamondsuit_u(\mathbf{V}_1, \ldots, \mathbf{V}_{d(u)})$ pour tout arbre u.

Théorème 3.1 *Soit n un entier positif et $\mathbf{V}_0, \ldots, \mathbf{V}_n$ des variétés de semigroupes (monoïdes). On note respectivement $\mathcal{V}_j$ et $\mathcal{W}$ les variétés de langages correspondant à $\mathbf{V}_j$ $(0 \leqslant j \leqslant n)$ et à $\diamondsuit_{(a\bar{a})^{n+1}}(\mathbf{V}_0, \ldots, \mathbf{V}_n)$. Alors pour tout alphabet A, $A^+\mathcal{W}$ $(A^*\mathcal{W})$ est l'algèbre de Boole engendrée par les langages de la forme $L_{i_0}a_1L_{i_1} \ \cdots \ a_kL_{i_k}$, avec $0 \leqslant i_0 < i_1 < \ldots < i_k \leqslant n$, $a_1, \ldots, a_k \in A$ et pour $0 \leqslant j \leqslant k$, $L_{i_j} \in A^+\mathcal{V}_{i_j}$ $(A^*\mathcal{V}_{i_j})$.*

Démonstration. Si $L = L_{i_0}a_1L_{i_1} \ \cdots \ a_kL_{i_k}$, avec les conditions précisées par l'énoncé, L est reconnu par le semigroupe $\diamondsuit_{k+1}(S_{i_0}, \ldots, S_{i_k})$ où, pour $0 \leqslant j \leqslant k$, S_{i_j} est le semigroupe syntactique de L_{i_j}. Or

$$\diamondsuit_{k+1}(S_{i_0}, \ldots, S_{i_k}) < \diamondsuit_{n+1}(1, \ldots, 1, S_{i_0}, 1, \ldots, 1, S_{i_1}, \ldots, S_{i_k}, 1, \ldots, 1) = T$$

et T est dans $\diamondsuit_{(a\bar{a})^{n+1}}(\mathbf{V}_0, \ldots, \mathbf{V}_n)$ par définition. Donc $L \in A^+\mathcal{W}$ et $A^+\mathcal{W}$ contient l'algèbre de Boole décrite dans l'énoncé.

Réciproquement soit $L \in A^+\mathcal{W}$. Alors d'après une remarque faite plus haut, le semigroupe syntactique de L divise un semigroupe de la forme $S = \diamondsuit_{n+1}(S_0, \ldots, S_n)$ avec pour $0 \leqslant i \leqslant n$, $S_i \in \mathbf{V}_i$. On en déduit que S reconnaît L et il suffit d'appliquer le théorème 2.7 pour conclure. $\square$

Les propositions qui suivent, qui sont pour l'essentiel des reformulations de résultats déjà connus, me serviront d'exemples. Je commence par la hiérarchie de Simon [12]. Pour tout alphabet A, on note $A^*\mathcal{J}_n$ l'algèbre de Boole engendrée par les langages de la forme $A^*a_1A^*a_2 \ \cdots \ A^*a_kA^*$ avec $0 \leqslant k \leqslant n$ et $a_i \in A$ pour $1 \leqslant i \leqslant k$. On démontre que $\mathcal{J}_n$ est une $*$-variété et on note $\mathbf{J}_n$ la variété de monoïdes correspondante. La variété $\mathbf{J}$ définie dans la section 1 est bien sûr la réunion des variétés $\mathbf{J}_n$.

Proposition 3.2

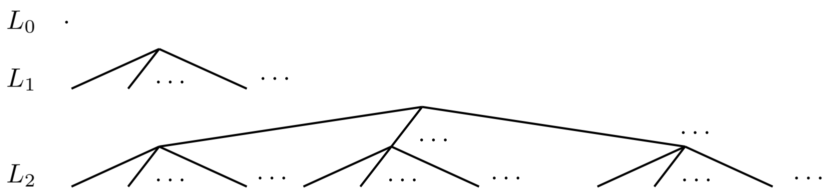
(1) *Si $u = a\bar{a}a\bar{a}$ (arbre $\bigwedge$), $\diamondsuit_u(\mathbf{I}) = \mathbf{J_1}$, la variété des monoïdes idempotents et commutatifs.*

(2) *Si $u = (a\bar{a})^{n+1}$ (arbre*  *à $n+1$ feuilles), $\diamondsuit_u(\mathbf{I}) = \mathbf{J}_n$.*

(3) *Si $L = (a\bar{a})^*$, $\diamondsuit_L(\mathbf{I}) = \mathbf{J}$, la variété des monoïdes $\mathcal{J}$-triviaux.*

Démonstration. C'est une conséquence immédiate du théorème 3.1. $\square$

La hiérarchie de Straubing $\mathbf{V}_n$ peut se décrire de façon analogue. soit L_n la suite de langages définie par $L_0 = \{1\}$ et $L_{n+1} = (aL_n\bar{a})^*$. Pour guider l'intuition, on peut représenter ces langages par des arbres « infinis en largeur » :



Proposition 3.3 *Pour tout $n \geqslant 0$, $\diamondsuit_{L_n}(\mathbf{I}) = \mathbf{V}_n$. en particulier, $\diamondsuit_{L_0}(\mathbf{I}) = \mathbf{I}$, $\diamondsuit_{L_1}(\mathbf{I}) = \mathbf{J}$, $\diamondsuit_{L_2}(\mathbf{I}) = \mathbf{PJ}$.*

Démonstration. Là encore, la proposition résulte immédiatement du théorème 3.1 $\square$

La hiérarchie de Brzozowski s'obtient en considérant des variétés de la forme $\diamondsuit_L(\mathbf{Nil})$ ou $\mathbf{Nil}$ désigne la variété des semigroupes nilpotents. De façon pllus précise :

Proposition 3.4 *Pour tout $n, k \geqslant 0$, on a :*
(a) $\mathbf{B}_n = \diamondsuit_{L_n}(\mathbf{Nil})$.
(b) $\mathbf{B}_{n+1,k} = \diamondsuit_{(aL_n\bar{a})^{k+1}}(\mathbf{Nil})$.

Démonstration. On a $\mathbf{B}_0 = \diamondsuit_{L_1}(\mathbf{Nil})$ par définition. Par récurrence je suppose établie la formule $\mathbf{B}_n = \diamondsuit_{L_n}(\mathbf{Nil})$. Alors pour tout $k \geqslant 0$, $A^+\mathcal{B}_{n+1,k}$ est l'algèbre de Boole engendrée par les langages de la forme $L_0 \ \cdots \ L_r$ avec $r \leqslant k$ et $L_0, \ldots, L_r \in A^+\mathcal{B}_n$. Soit $\mathcal{A}_{n+1,k}$ la variété de langages correspondant à $\diamondsuit_{(aL_n\bar{a})^{k+1}}(\mathbf{Nil})$. Le théorème 3.1 et l'hypothèse de récurrence montrent que pour tout alphabet A, $A^+\mathcal{A}_{n+1,k}$ est l'algèbre de Boole engendrée par les langages de la forme $L_0a_1L_1 \ \cdots \ a_rL_r$ où $0 \leqslant r \leqslant k$, $L_0, \ldots, L_r \in A^+\mathcal{B}_n$ et $a_1, \ldots, a_r \in A$. Il suffit donc d'établir, pour tout $k \geqslant 0$, l'égalité $\mathcal{A}_{n+1,k} = \mathcal{B}_{n+1,k}$. L'inclusion $\mathcal{A}_{n+1,k} \subset \mathcal{B}_{n+1,k}$ résulte du fait [4, p. 259] que si $L \in A^+\mathcal{B}_n$, alors $aL, La \in A^+\mathcal{B}_n$ pour tout $a \in A$. Un produit de la forme $L_0a_1L_1 \ \cdots \ a_rL_r$ avec $L_0, \ldots, L_r \in A^+\mathcal{B}_n$ et $a_1, \ldots, a_r \in A$ s'écrit donc sous la forme $L_0(a_1L_1) \ \cdots \ (a_rL_r)$ avec $L_0, a_1L_1, \ldots, a_rL_r \in A^+\mathcal{B}_n$. Pour démontrer l'inclusion opposée, on montre par récurrence sur k que tout langage de la forme $L_0 \ \cdots \ L_k$ avec $L_0, \ldots, L_k \in A^+\mathcal{B}_n$ est *union* finie de langages de la forme $K_0a_1K_1 \cdots a_rK_r$ avec $0 \leqslant r \leqslant k$, $K_0, \ldots, K_r \in A^+\mathcal{B}_n$ et $a_0, \ldots, a_r \in A$. C'est évident si $k = 0$. Si $L = L_0 \ \cdots \ L_{k+1}$, on observe que :

$$L_{k+1} = (A \cap L_{k+1}) \cup \bigcup_{a \in A} a(a^{-1}L_{k+1}),$$

d'où :

$$L = \bigcup_{a \in A \cap L_{k+1}} (L_0 \ \cdots \ L_k)a \cup \bigcup_{a \in A} (L_0 \ \cdots \ L_k)a(a^{-1}L_{k+1}).$$

Mais d'après l'hypothèse de récurrence, $L_0 \ \cdots \ L_k$ s'écrit comme union de langages de la forme $K_0a_1 \ \cdots a_rK_r$ avec $0 \leqslant r \leqslant k$, $a_1, \ldots, a_r \in A$ et $K_0, \ldots, K_r \in A^+\mathcal{B}_n$. Or puisque $\mathcal{B}_n$ est une variété de langages, on a $a^{-1}L_{k+1} \in A^+\mathcal{B}_n$ pour tout $a \in A$. Ces deux observations permettent d'exprimer L comme union de langages de la forme $K_0a_1 \ \cdots \ a_pK_p$ (avec $0 \leqslant p \leqslant k+1$, $a_1, \ldots, a_p \in A$ et $K_0, \ldots, K_p \in A^+\mathcal{B}_n$), ce qui conclut la récurrence. Donc $\mathcal{A}_{n+1,k} = \mathcal{B}_{n+1,k}$ pour tout k ce qui établit (b).

L'égalité $\mathbf{B}_{n+1} = \diamondsuit_{L_{n+1}}(\mathbf{Nil})$ s'en déduit facilement puisque

$$\mathbf{B}_{n+1} = \bigcup_{k \geqslant 0} \mathbf{B}_{n+1,k} \quad \text{et} \quad \diamondsuit_{L_{n+1}}(\mathbf{Nil}) = \bigcup_{k \geqslant 0} \diamondsuit_{(aL_n\bar{a})^{k+1}}(\mathbf{Nil}). \qquad \square$$

D. Thérien (communication personnelle) a proposé de modifier la hiérarchie de Brzozowski en partant de la variété $\mathbf{LI}$ au lieu de $\mathbf{Nil}$. On définit alors des $+$-variétés $\mathcal{B}'_{n,k}$ de la façon suivante. Pour tout alphabet A,

(1) $A^+\mathcal{B}'_0$ est constitué des langages de la forme $XA^*Y \cup Z$ où X, Y et Z sont des langages finis de A^+.
(2) $A^+\mathcal{B}'_{n+1,r}$ est l'algèbre de Boole engendrée par les langages de la forme $L_0a_1L_1 \ \cdots \ a_kL_k$ avec $0 \leqslant k \leqslant r$, $a_1, \ldots, a_k \in A$ et $L_0, \ldots, L_k \in A^+\mathcal{B}'_n$.
(3) $A^+\mathcal{B}'_{n+1} = \bigcup_{r \geqslant 0} A^+\mathcal{B}'_{n+1,r}$.

On sait que $\mathcal{B}'_0$ est la variété de langages correspondant à $\mathbf{LI}$. Par conséquent, il résulte aisément du théorème 3.1 :

Proposition 3.5 *Pour tout $n, k \geqslant 0$, $\mathcal{B}'_n$ et $\mathcal{B}'_{n,k}$ sont des variétés de langages. Les variétés de semigroupes correspondantes sont respectivement $\mathbf{B}'_n = \Diamond_{L_n}(\mathbf{LI})$ et $\mathbf{B}'_{n+1,k} = \Diamond_{(aL_n\bar{a})^{k+1}}(\mathbf{LI})$.*

Le lien entre les variétés $\mathbf{B}_{n,k}$ et $\mathbf{B}'_{n,k}$ est précisé par l'énoncé suivant, dû à D. Thérien :

Proposition 3.6
(1) *Pour tout $k \geqslant 0$, $\mathbf{B}_{1,2k} = \mathbf{B}_{1,2k+1} = \mathbf{B}'_{1,k}$*
(2) *Pour tout $n > 0$ et $k \geqslant 0$, $\mathbf{B}_{n+1,k} = \mathbf{B}'_{n+1,k}$ et $\mathbf{B}_n = \mathbf{B}'_n$.*

Démonstration. (1) Soit A un alphabet. D'après la définition de $A^+\mathcal{B}'_{1,k}$ et l'expression des éléments de $A^+\mathcal{B}'_0$, on voit que $A^+\mathcal{B}'_{1,k}$ est l'algèbre de Boole engendrée par les langages de la forme uA^*, A^*u et $u_0A^*u_1 \ \cdots \ u_rA^*u_{r+1}$ où $0 \leqslant r \leqslant k$ et $u, u_0, \ldots, u_{r+1} \in A^+$. Or on sait [4, p. 258] que l'algèbre de Boole ainsi définie est $A^+\mathcal{B}_{1,2k} = A^+\mathcal{B}_{1,2k+1}$.

(2) On déduit de (1) la relation :

$$A^+\mathcal{B}'_1 = \bigcup_{k>0} A^+\mathcal{B}'_{1,k} = \bigcup_{k>0} A^+\mathcal{B}_{1,k} = A^+\mathcal{B}_1 \quad \text{d'où } \mathbf{B}'_1 = \mathbf{B}_1.$$

Les autres relations s'en déduisent par récurrence. $\square$

Comme on le voit, les hiérarchies $\mathbf{B}_{n,k}$ et $\mathbf{B}'_{n,k}$ ne diffèrent que pour $n = 1$. Cependant, la hiérarchie $\mathbf{B}'_n$ me paraît plus intéressante. Par exemple, comme me l'a fait remarquer D. Thérien, le théorème 1.5 peut s'écrire

$$\mathbf{B}'_n = \mathbf{V}_n * \mathbf{LI} \quad \text{pour tout } n \geqslant 0,$$

alors que le cas $n = 0$ était curieusement éliminé dans la version donnée plus haut.

Le théorème qui suit donne une caractérisation algébrique intéressante des opérations schématisées par $\mathbf{V} \to \underset{\mathbf{V} \quad \mathbf{I}}{\wedge}$ et $\mathbf{V} \to \underset{\mathbf{I} \quad \mathbf{V}}{\wedge}$.

Théorème 3.7 *Pour toute variété de monoïdes $\mathbf{V}$, on a :*
(a) $\Diamond_{a\bar{a}a\bar{a}}(\mathbf{V}, \mathbf{I}) = \mathbf{J_1} * \mathbf{V}$.
(b) $\Diamond_{a\bar{a}a\bar{a}}(\mathbf{I}, \mathbf{V}) = \mathbf{V} *_r \mathbf{J_1}$.

La démonstration repose sur le « principe du produit semi-direct », qui est une version simplifiée du « wreath product principle » de Straubing (Ph. D. 1978), lequel est lui-même le premier énoncé complet d'un résultat utilisé antérieurement dans des cas particuliers, notamment par A. Meyer (1969) et Brzozowski-Simon (1973). Soit $\eta : A^* \to M * N$ un morphisme de A^* dans un produit semi-direct de monoïdes $M * N$ et soit $\pi : M * N \to N$ la projection canonique. On a alors, en posant $B = N \times A$ et $\varphi = \pi\eta$:

Proposition 3.8 (Principe du produit semi-direct) *Si L est reconnu par η, alors L est réunion de langages de la forme $X \cap Y\sigma^{-1}$ où $X \subset A^*$ est reconnu par N, $Y \subset B^*$ est reconnu par M et où $\sigma : A^* \to B^*$ est la fonction séquentielle définie par $1\sigma = 1$ et $(a_1 \ \cdots \ a_n)\sigma = (1, a_1)(a_1\varphi, a_2) \ \cdots \ ((a_1 \ \cdots \ a_n)\varphi, a_n)$.*

Démonstration. Il suffit d'établir le résultat dans le cas où $L = (s_0, t_0)\eta^{-1}$ pour un certain $(s_0, t_0) \in M * N$. Selon l'usage, je noterai additivement le monoïde M et par juxtaposition l'action de N sur M.

Soit $\alpha : B = N \times A \to M$ défini par $(t, a)\alpha = ts$ où s est la première composante de $a\eta$ (c'est-à-dire $a\eta = (s, u)$ pour un certain $u \in T$). Soit $X = t_0\varphi^{-1}$ et $Y = s_0\alpha^{-1}$: X est donc reconnu par N et Y par M. Il vient

$$X \cap Y\sigma^{-1} = \{a_1 \cdots a_n \in A^* \mid (a_1 \cdots a_n)\varphi = t_0 \text{ et } (a_1 \cdots a_n)\sigma\alpha = s_0\}$$

Or en posant $a_i = (s_i, t_i)$, on a successivement :

$$\begin{aligned}(a_1 \cdots a_n)\sigma\alpha &= \big((1, a_1)(a_1\varphi, a_2) \ \cdots \ ((a_1 \cdots a_{n-1})\varphi, a_n)\big)\alpha \\ &= (1, a_1)\alpha + (a_1\varphi, a_2)\alpha + \ldots + ((a_1 \cdots a_{n-1})\varphi, a_n)\alpha \\ &= s_1 + t_1s_2 + \ldots + t_1 \cdots t_{n-1}s_n\end{aligned}$$

et $(a_1 \ \cdots \ a_n)\varphi = t_1 \cdots t_n$.

On en déduit $((a_1 \cdots a_n)\sigma\alpha, (a_1 \cdots a_n)\varphi) = (a_1 \cdots a_n)\eta$, d'où finalement $X \cap Y\sigma^{-1} = (s_0, t_0)\eta^{-1}$ ce qui démontre la proposition. □

Remarque. La fonction séquentielle σ est réalisée par le transducteur dont les états sont les éléments de N (1 étant état initial) et dont les transitions et la fonction de sortie sont représentés par le diagramme :

$$t \xrightarrow{\ a \mid (t, a)\ } t(a\varphi)$$

On déduit du principe du produit semi-direct l'énoncé suivant, qui est une version améliorée d'un exercice proposé par Eilenberg [4, p. 253].

Proposition 3.9 *Soit* $\mathbf{V}$ *une variété de monoïdes et soit $\mathcal{V}$ la variété de langages correspondante. Soit $A^*\overline{\mathcal{V}}$ l'algèbre de Boole engendrée par les langages de la forme L ou LaA^* avec $a \in A$ et $L \in A^*\mathcal{V}$. Alors $\overline{\mathcal{V}}$ est une variété de langages et la variété de monoïdes correspondante est* $\mathbf{J_1} * \mathbf{V}$.

Démonstration. Tout d'abord, si $L \in A^*\mathcal{V}$, LaA^* est reconnu par $\diamondsuit_2(M(L), 1)$. Or on sait que $\diamondsuit_2(M_1, M_2) = M_2 *_r (V * M_1)$ pour un certain monoïde $V \in \mathbf{J_1}$. On en déduit ici $\diamondsuit_2(M(L), 1) \in \mathbf{J_1} * \mathbf{V}$.

Réciproquement, soit $L \subset A^*$ un langage reconnu par un élément de $\mathbf{J_1} * \mathbf{V}$. Il existe alors deux monoïdes $M \in \mathbf{J_1}$ et $N \in \mathbf{V}$ tels que L soit reconnu par

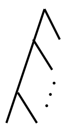
$M * N$. Avec les notations de la proposition précédente, il suffit d'établir que $Y\sigma^{-1}$ est élément de $A^*\overline{\mathcal{V}}$. Or puisque Y est reconnu par M, Y est combinaison booléenne de langages de la forme B^*bB^* avec $b \in B$. Comme σ^{-1} commute aux opérations booléennes, il suffit d'établir que $(B^*bB^*)\sigma^{-1} \in A^*\overline{\mathcal{V}}$. Or si on pose $b = (t, a)$, il vient

$$\begin{aligned}(B^*bB^*)\sigma^{-1} &= \{a_1 \cdots a_n \in A^* \mid \text{il existe } i \text{ tel que } 1 \leqslant i \leqslant n \\ &\qquad\qquad ((a_1 \cdots a_{i-1})\varphi, a_i) = (t, a)\} \\ &= (t\varphi^{-1})aA^*.\end{aligned}$$

Comme $t\varphi^{-1} \in A^*\mathcal{V}$, on a $(B^*bB^*)\sigma^{-1} \in A^*\overline{\mathcal{V}}$ et donc finalement $L \in A^*\overline{\mathcal{V}}$. Il en résulte que $\overline{\mathcal{V}}$ est une variété de langages et que la variété de monoïdes correspondante est $\mathbf{J_1} * \mathbf{V}$. □

Démonstation du théorème 3.7. (a) découle directement de la proposition 3.8 et du théorème 3.1. L'énoncé (b) est dual. □

Corollaire 3.10

(1) *Si* $u = a^n(\bar{a}a\bar{a})^n$ *(arbre*  *à* $n+1$ *feuilles),*

$$\Diamond_u(\mathbf{I}) = \mathbf{J_1^n} = \underbrace{\mathbf{J_1} * \mathbf{J_1} * \cdots * \mathbf{J_1}}_{n \text{ fois}}$$

(2) *Si* $L = \{a^n(\bar{a}a\bar{a})^n \mid n \geqslant 0\}$, $\Diamond_L(\mathbf{I}) = \mathbf{R}$, *la variété des monoïdes* $\mathcal{R}$-*triviaux.*

Démonstration.

(1) Par récurrence à l'aide des théorèmes 3.1 et 3.7.

(2) Un résultat de théorie des semigroupes affirme qu'un monoïde est $\mathcal{R}$-trivial si et seulement si il divise un produit en couronne de la forme $U_1 \circ U_1 \circ \cdots \circ U_1$. En termes de variétés, cet énoncé se traduit par l'égalité $\mathbf{R} = \bigcup_{n \geqslant 0} \mathbf{J}_1^n$ ce qui établit (2). □

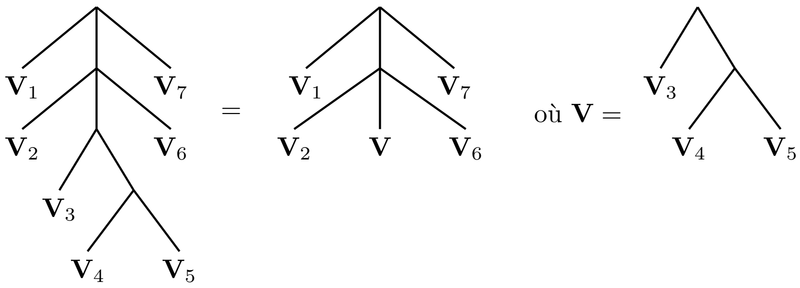
Corollaire 3.11 *Soit* $\mathbf{V}$ *une variété de monoïdes et* $\mathcal{V}$ *la variété de langages correspondante. Soit, pour tout alphabet* A, $A^*\overline{\mathcal{V}}$ *la plus petite algèbre de Boole contenant* $A^*\mathcal{V}$ *et fermée pour les opérations* $L \to LaA^*$ *(resp.* $L \to A^*aL$*) où* $a \in A$. *Alors* $\overline{\mathcal{V}}$ *est une variété de langages et la variété de monoïdes correspondante est* $\mathbf{R} * \mathbf{V}$ *(*$\mathbf{V} *_r \mathbf{R}$*).*

Démonstration. L'énoncé résulte de la proposition 3.9 et de l'égalité $\mathbf{R} = \bigcup_{n \geqslant 0} \mathbf{J}_1^n$. □

Il est intéressant de comparer entre elles les variétés $\Diamond_u(\mathbf{V}_1, \ldots, V_{d(u)})$. Voici une première observation, qui est une conséquence immédiate des définitions.

Proposition 3.12 *Soit* u *un arbre. Alors, pour toute variété* $\mathbf{V}_1, \ldots, \mathbf{V}_{d(u)}$, *on a* $\Diamond_u(\mathbf{V}_1, \ldots, V_{d(u)}) = \Diamond_{au\bar{a}}(\mathbf{V}_1, \ldots, V_{d(u)}) = \Diamond_{a\bar{a}}(\Diamond_u(\mathbf{V}_1, \ldots, V_{d(u)}))$.

la suite repose sur un théorème de substitution qu'on peut illustrer par le schéma suivant :



Théorème 3.13 (théorème de substitution) *Soit $u = u_1 a u_2 \bar{a} u_3$ un arbre avec $u_2 \in P$. alors pour toute variété $\mathbf{V}_1, \ldots, \mathbf{V}_{d(u)}$, on a*

$$\diamondsuit_u(\mathbf{V}_1, \ldots, \mathbf{V}_{d(u)}) = \diamondsuit_{u_1 a \bar{a} u_3}(\mathbf{V}_1, \ldots, \mathbf{V}_{d(u_1)}, \\ \diamondsuit_{u_2}(\mathbf{V}_{d(u_1)+1}, \ldots, \mathbf{V}_{d(u_1)+d(u_2)}), \mathbf{V}_{d(u_1)+d(u_2)+1}, \ldots, \mathbf{V}_{d(u)}).$$

Démonstration. Par récurrence sur $|u|$. Si $u = a\bar{a}$, la relation cherchée s'écrit $\diamondsuit_{a\bar{a}}(\mathbf{V}) = \diamondsuit_{a\bar{a}}(\diamondsuit_1(\mathbf{V}))$ et résulte de l'égalité $\mathbf{V} = \diamondsuit_1(\mathbf{V})$. Dans le cas général, on factorise u en $u = (av_1\bar{a}) \cdots (av_n\bar{a})$ avec $v_1, \ldots, v_n \in P$.

Premier cas : $n = 1$. Si $u_1 u_3 = 1$, il vient $u_2 = v_1$ et d'après la proposition 3.12, on a $\diamondsuit_u(\mathbf{V}_1, \ldots, \mathbf{V}_{d(u)}) = \diamondsuit_{a\bar{a}}(\diamondsuit_{u_2}(\mathbf{V}_2, \ldots, \mathbf{V}_{d(u)}))$. Sinon, on a nécessairement $u_1 = au'_1$, $u_3 = u'_3\bar{a}$ d'où $v_1 = u'_1 a u_2 \bar{a} u'_3$. La relation cherchée s'obtient en utilisant à la fois la proposition 3.12 et l'hypothèse de récurrence appliquée à v_1.

Deuxième cas : $n \geqslant 2$. Alors l'un des arbres $av_i\bar{a}$ $(1 \leqslant i \leqslant n)$ admet $au_2\bar{a}$ comme facteur. On pose $av_i\bar{a} = u' a u_2 \bar{a} u''$ et on applique l'hypothèse de récurrence :

$$\diamondsuit_{av_i\bar{a}}(\mathbf{V}_1, \ldots, \mathbf{V}_{d(v_i)}) = \diamondsuit_{v_i}(\mathbf{V}_1, \ldots, \mathbf{V}_{d(v_i)}) = \\ \diamondsuit_{u' a\bar{a} u''}(\mathbf{V}_1, \ldots, \mathbf{V}_{d(u')}, \diamondsuit_{u_2}(\mathbf{V}_{d(u')+1}, \ldots, \mathbf{V}_{d(u')+d(u_2)}), \ldots, \mathbf{V}_{d(v_i)}).$$

La relation cherchée s'en déduit puisque

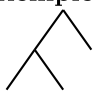
$$\diamondsuit_u(\mathbf{V}_1, \ldots, \mathbf{V}_{d(u)}) = \diamondsuit_{(a\bar{a})^n}(\diamondsuit_{v_1}(\mathbf{V}_1, \ldots, \mathbf{V}_{d(v_1)}), \ldots, \\ \diamondsuit_{v_n}(\mathbf{V}_{d(v_1 \cdots v_{n-1})+1}, \ldots, \mathbf{V}_{d(u)})). \qquad \square$$

On en déduit à l'aide de la proposition 3.12 :

Corollaire 3.14 *Soit $u = u_1 a a u_2 \bar{a}\bar{a} u_3$ un arbre avec $u_2 \in P$. Alors pour toute variété $\mathbf{V}$, $\diamondsuit_u(\mathbf{V}) = \diamondsuit_{u_1 a u_2 \bar{a} u_3}(\mathbf{V})$.*

Soient u et v deux arbres. On dit que u est un arbre *extrait* de v si u s'obtient à partir de u en supprimant dans v un certain nombre d'occurrences liées de a et $\bar{a}$. De façon formelle, la relation « est extrait de » est la fermeture réflexive et transitive de la relation $\leqslant$ définie par $u \leqslant v$ si et seulement si il existe une factorisation $v = v_1 a v_2 \bar{a} v_3$ telle que $v_2 \in P$ et $u = v_1 v_2 v_3$.

Par extension, on notera également $\leqslant$ la relation « est extrait de ». Il est clair que $\leqslant$ est une relation d'ordre.

Exemple. L'arbre $u = a\bar{a}a\bar{a}a\bar{a}$ () est extrait de $v = aa\bar{a}a\bar{a}\bar{a}a\bar{a}$ () puisque $v = (1)a(a\bar{a}a\bar{a})\bar{a}(a\bar{a})$.

Théorème 3.15 *Pour toute variété* $\mathbf{V}$*, si* u *est un arbre extrait de l'arbre* v*, on a* $\diamondsuit_u(\mathbf{V}) \subset \diamondsuit_v(\mathbf{V})$.

Démonstration. D'après la définition de $\leqslant$, on peut se ramener au cas où $v = v_1av_2\bar{a}v_3$ avec $v_2 \in P$ et $u = v_1v_2v_3$. La preuve se fait par récurrence sur $|v|$. Si $v = a\bar{a}$, $u = 1$ et le résultat est évident. Si $d(v) = 1$, on a $v = av'\bar{a}$ avec $v' \in P$ d'où $\diamondsuit_v(\mathbf{V}) = \diamondsuit_{v'}(\mathbf{V})$ d'après (2). Si $v_1v_3 = 1$, il vient $v_2 = v' = u$ et le résultat est établi. Sinon, on a nécessairement $v_1 = av_1'$ et $v_3 = v_3'\bar{a}$ d'où $v' = v_1'av_2\bar{a}v_3'$. On en déduit en appliquant l'hypothèse de récurrence à v' :

$$\diamondsuit_v(\mathbf{V}) = \diamondsuit_{v'}(\mathbf{V}) \supset \diamondsuit_{v_1'v_2v_3'}(\mathbf{V}) = \diamondsuit_{av_1'v_2v_3'a}(\mathbf{V}) = \diamondsuit_u(\mathbf{V})$$

Si $d(v) \geqslant 2$, v se factorise en $v = (aw_1\bar{a})(aw_2\bar{a})\cdots(aw_n\bar{a})$ avec $n \geqslant 2, w_1, \ldots, w_n \in P$ et où l'un des arbres $aw_k\bar{a}$ admet $av_2\bar{a}$ comme facteur. On pose $aw_k\bar{a} = v'av_2\bar{a}v''$ et $w_k' = a^{-1}(v'v_2v'')\bar{a}^{-1}$. On obtient alors en appliquant l'hypothèse de récurrence $\diamondsuit_{w_k}(\mathbf{V}) \supset \diamondsuit_{w_k'}(\mathbf{V})$:

$$\diamondsuit_v(\mathbf{V}) = \diamondsuit_{(a\bar{a})^n}(\diamondsuit_{w_1}(\mathbf{V}), \ldots, \diamondsuit_{w_n}(\mathbf{V})) \supset \diamondsuit_{(a\bar{a})^n}(\diamondsuit_{w_1}(\mathbf{V}), \ldots, \diamondsuit_{w_k'}(\mathbf{V}), \ldots, \diamondsuit_{w_n}(\mathbf{V})) = \diamondsuit_u(\mathbf{V}).$$

Le corollaire 3.14 montre qu'on peut se restreindre à l'ensemble P' des arbres dont chaque noeud est d'arité différente de 1. Bien entendu, si on part d'une variété $\mathbf{V}$ quelconque, la restriction à P' de l'application $u \to \diamondsuit_u(\mathbf{V})$ n'a cependant aucune raison d'être injective. Par exemple si $\mathbf{V} = \mathbf{A}$, la variété des semigroupes apériodiques, on a $\diamondsuit_u(\mathbf{V}) = \mathbf{V}$ pour tout u. En revanche, on peut avancer la conjecture suivante.

Conjecture *Soient* $u, v \in P'$*. Alors* $\diamondsuit_u(\mathbf{I})$ *est contenu dans* $\diamondsuit_v(\mathbf{I})$ *si et seulement si* u *est extrait de* v*.*

Si cette conjecture est vraie, la restriction à P' de l'application $u \to \diamondsuit_u(\mathbf{I})$ est injective. L'étude de l'application $L \to \diamondsuit_L(\mathbf{I})$ paraît encore plus problématique. Il est effet prouvé en [10] que

$$\diamondsuit_{(a\bar{a})^*}(\mathbf{J}) = \diamondsuit_{(a\bar{a})^*}(\mathbf{J_1}) = \diamondsuit_{(a\bar{a})^*}(\mathbf{R})$$

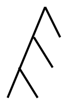
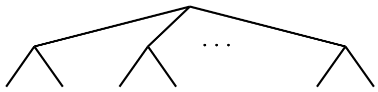
d'où l'on déduit

$$\diamondsuit_{a(a\bar{a})^*\bar{a}}(\mathbf{I}) = \diamondsuit_{(aa\bar{a}a\bar{a}\bar{a})^*}(\mathbf{I}) = \diamondsuit_{(aL\bar{a})^*}(\mathbf{I})$$

où L est le langage $\{a^n(\bar{a}a\bar{a})^n \mid n \geqslant 0\}$.

Je voudrais signaler une petite subtilité relative à cette dernière égalité. On en déduit en particulier :

$$\mathbf{J_1^3} = \diamondsuit_{a^3(\bar{a}a\bar{a})^3}(\mathbf{I}) \subset \bigcup_{n>0} \diamondsuit_{(aa\bar{a}a\bar{a}\bar{a})^n}(\mathbf{I}).$$

Or il est facile de vérifier que l'arbre $a^3(\bar{a}a\bar{a})^3$ () n'est extrait d'aucun des arbres $(aa\bar{a}a\bar{a}\bar{a})^n$ ().

Cependant la relation ci-dessus ne contredit pas la conjecture car on ne peut pas en déduire *a priori* une inclusion du type :

$$\diamondsuit_{a^3(\bar{a}a\bar{a})^3}(\mathbf{I}) \subset \diamondsuit_{(aa\bar{a}a\bar{a}\bar{a})^n}(\mathbf{I}).$$

4 Problèmes de décidabilité

On dit qu'une variété de semigroupes (ou de monoïdes) $\mathbf{V}$ est *décidable* s'il existe un algorithme qui permet de tester si un semigroupe fini donné est ou n'est pas dans $\mathbf{V}$.

Pour les variétés $\mathbf{V}_n$ et $\mathbf{B}_n$ des hiérarchies de Straubing et Brzozowski, le problème de la décidabilité est toujours ouvert puisque seul le cas $n = 1$ a pu être résolu positivement (cf. les théorèmes de Simon et de Knast rappelés dans la section 1). Le résultat qui suit constitue peut-être une première étape vers la solution générale du problème.

Théorème 4.1 *Pour tout arbre u, la variété $\diamondsuit_u(\mathbf{I})$ est décidable.*

Le théorème repose sur une propriété de $\diamondsuit_u(\mathbf{I})$ intéressante pour elle-même.

Proposition 4.2 *Soit u un arbre et soit $\mathcal{V}_u$ la variété de langages associée à $\diamondsuit_u(\mathbf{I})$. Pour tout alphabet A, $A^*\mathcal{V}_u$ est un ensemble fini effectivement descriptible.*

Démonstration. Le résultat est évident si $u = 1$. Si $u = au_1\bar{a}au_2\bar{a}\cdots au_n\bar{a}$ avec $u_1, \dots, u_n \in P$, on a $\diamondsuit_u(\mathbf{V}) = \diamondsuit_{(a\bar{a})^n}(\diamondsuit_{u_1}(\mathbf{V}_1), \dots, \diamondsuit_{u_n}(\mathbf{V}_n))$. Par récurrence les ensembles $A^*\mathcal{V}_{u_1}, \dots, A^*\mathcal{V}_{u_n}$ sont des ensembles finis effectivement descriptibles. Le théorème 3.1 donne alors un algorithme pour construire $A^*\mathcal{V}_u$ qui est un ensemble fini puisque c'est l'algèbre de Boole engendrée par un nombre fini de langages. $\square$

Preuve du théorème 4.1. Soit M un monoïde fini et A un alphabet en bijection avec M. Il existe alors un morphisme surjectif naturel $\pi : A^* \to M$. D'après [4, p. 188], on a pour tout $m \in M$ la double inégalité

$$M(m\pi^{-1}) < M < \prod_{m \in M} M(m\pi^{-1}).$$

On en déduit que M est dans la variété $\diamondsuit_u(\mathbf{I})$ si et seulement si, pour tout $m \in M$, le langage $m\pi^{-1}$ est dans $A^*\mathcal{V}_u$. La proposition 4.2 fournit donc un algorithme pour tester si $M \in \diamondsuit_u(\mathbf{I})$.

On en déduit en particulier à l'aide du corollaire 3.10 un résultat de pure théorie des semigroupes.

Corollaire 4.3 *Pour tout entier n, la variété $\mathbf{J_1}^n = \underbrace{\mathbf{J_1} * \mathbf{J_1} * \cdots * \mathbf{J_1}}_{n \text{ fois}}$ est décidable.*

Références

[1] J. A. Brzozowski, Hierarchies of aperiodic languages, *RAIRO Inform. Théor.* **10**,R–2 (1976), 33–49.

[2] J. A. Brzozowski et R. Knast, The dot-depth hierarchy of star-free languages is infinite, *J. Comput. System Sci.* **16**,1 (1978), 37–55.

[3] J. A. Brzozowski et I. Simon, Characterizations of locally testable events, *Discrete Math.* **4** (1973), 243–271.

[4] S. Eilenberg, *Automata, languages, and machines. Vol. B*, Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1976. With two chapters ("Depth decomposition theorem" and "Complexity of semigroups and morphisms") by Bret Tilson, Pure and Applied Mathematics, Vol. 59.

[5] R. Knast, A semigroup characterization of dot-depth one languages, *RAIRO Inform. Théor.* **17**,4 (1983), 321–330.

[6] R. Knast, Some theorems on graph congruences, *RAIRO Inform. Théor.* **17**,4 (1983), 331–342.

[7] G. Lallement, *Semigroups and combinatorial applications*, John Wiley & Sons, New York-Chichester-Brisbane, 1979. Pure and Applied Mathematics, A Wiley-Interscience Publication.

[8] J.-E. Pin, *Variétés de langages et variétés de semigroupes*, Thèse d'état, Université Paris VI, 1981.

[9] J.-E. Pin et J. Sakarovitch, Une application de la représentation matricielle des transductions. to appear.

[10] J.-E. Pin et H. Straubing, Monoids of upper triangular boolean matrices. to appear.

[11] C. Reutenauer, Sur les variétés de langages et de monoïdes, in *Theoretical computer science (Fourth GI Conf., Aachen, 1979)*, pp. 260–265, *Lecture Notes in Comput. Sci.* vol. 67, Springer, Berlin, 1979.

[12] I. Simon, *Hierarchies of Events with Dot-Depth One*, PhD thesis, University of Waterloo, Waterloo, Ontario, Canada, 1972.

[13] I. Simon, Piecewise testable events, in *Proc. 2nd GI Conf.*, H. Brackage (éd.), pp. 214–222, *Lecture Notes in Comp. Sci.* vol. 33, Springer Verlag, Berlin, Heidelberg, New York, 1975.

[14] H. Straubing, Finite semigroup varieties of the form $\mathbf{V} * \mathbf{D}$. to appear.

[15] H. Straubing, A generalization of the Schützenberger product of finite monoids, *Theoret. Comput. Sci.* **13**,2 (1981), 137–150.