Theoretical Computer Science 9 (1979) 385–398

FAMILLES DE LANGAGES FERMEES PAR CROCHET OUVERT

F. RODRIGUEZ

E.n.s.e.e.i.h.t., 2 rue Camichel, 31071 Toulouse Cedex; Université Paul Sabatier, 31077 Toulouse Cedex, France

Communiqué par M. A. Harrison
Reçu décembre 1976
Revisé octobre 1978

Abstract. We present a new unary operator, the open bracket, and establish it has properties similar to the ones of the bracket of [1] and the syntactic operators of [10]. We can thus study an infinite hierarchy of checking automata languages.

1. Introduction

Un aspect fondamental de la théorie des langages est l'étude des propriétés algébriques des familles de langages. Les théories des cônes rationnels [2] et des familles agréables de langages [5] (Fal en abrégé, 'full AFL' en anglais) ont été développées à cet effet. Cependant la caractérisation de familles de langages classiques a nécessité l'introduction d'outils plus spécifiques tels que les opérateurs 'syntactiques' binaires de Greibach [10]. De même Boasson, Crestin et Nivat ont caractérisé en [1] les langages ultralinéaires et (fortement) quasirationnels [3, 13] de rang p à l'aide de deux opérateurs unaires, le chevron et le crochet, qui a tout langage L associent respectivement les cônes rationnels engendrés par

$$\mathrm{Ch}(L)=\{a^n l b^n \mid l\in L,\ n\in \mathbf{N}\}$$

et

$$[L]=\{wl\tilde{w} \mid l\in L,\ w\in\{a,b\}^*\},$$

où $\{a, b\}$ est un ensemble disjoint de l'alphabet de L et $\tilde{w}$ l'image miroir de w.

Ces deux opérateurs permettent, en outre, d'aborder sous de nouveaux aspects [1] la conjecture de Greibach dans [19] sur la nature du cône rationnel maximal inclus dans les langages algébriques (context-free).

Dans le présent article nous considérons un nouvel opérateur unaire, appelé crochet ouvert, qui à tout langage L associe le cône rationnel engendré par $\mathrm{CO}(L)=\{wlw \mid l\in L,\ w\in\{a,b\}^*\}$. Nous montrons que cet opérateur à des propriétés analogues à celles du chevron et du crochet de [1] ainsi qu'à celles des opérateurs syntactiques de Greibach. Il permet en particulier de construire de nouvelles hiérarchies de cônes rationnels dont l'union est donc un cône non principal.

Nous utilisons alors cet opérateur pour étudier la famille des langages vérifiables [14] ('checking automata languages' dans [9]). Rappelons qu'un langage est vérifiable s'il est reconnu par un accepteur vérificateur [14] ('checking automaton' dans [9]), cas particulier des 'one-way stack automata' introduits par Ginsburg, Greibach et Harrison en [7]. Ces accepteurs unilatères sont munis d'un ruban mémoire sur lequel ils peuvent écrire puis consulter un mot sans plus pouvoir le modifier. La classe des langages vérifiables présente de nombreuses analogies structurelles avec celle des langages algébriques: toutes deux forment des FAL principales, fermées par substitution. On peut y définir des hiérarchies infinies à partir du fonctionnement des automates qui les reconnaissent: familles $\mathcal{F}_{m,n}$ de Greibach dans [8], familles $\mathcal{V}_m$ des cônes rationnels de langages reconnus par des accepteurs vérificateurs faisant au plus m mouvements en mémoire définis par l'auteur en [15] ('finite turn checking automata languages' dans [17]).

Notons enfin que l'étude des langages vérifiables est également justifiée par une conjecture de Greibach dans [19]: le cône rationnel maximal contenu dans la famille des langages algébriques serait l'intersection de cette famille avec celle des langages vérifiables.

Nous montrons dans cet article que le plus petit cône rationnel fermé par union et crochet ouvert est l'union d'une hiérarchie $\mathscr{C}(C_i)$ de cônes rationnels de langages vérifiables que nous comparons à la hiérarchie $\mathcal{V}_m$ de [15]. Enfin nous montrons l'incomparabilité des plus petits cônes rationnels fermés respectivement par crochet (la famille des langages linéaires) et par crochet ouvert.

Un certain nombre de ces résultats a été présenté sans preuve dans [16].

Dans toute la suite les monoïdes libres considérés seront finiment engendrés et ε désignera l'élément neutre. Pour tout langage L, X_L désignera le plus petit alphabet fini tel que $L \subseteq X_L^*$. Les alphabets X_L sont tous partie d'un alphabet dénombrable unique.

2. Notations, définitions

Définition 1. Un homomorphisme φ d'un monoïde libre X^* dans l'ensemble des parties d'un monoïde Y^* (muni de la structure de monoïde induite par celle de Y^*) est une *$\mathscr{L}$-substitution* si et seulement si pour tout $x \in X$, $\varphi(x) \in \mathscr{L}$. Nous noterons $\mathscr{L}\,\mathscr{L}'$ la famille de langages définie par

$$\mathscr{L}\,\mathscr{L}' = \{\varphi(L) \mid L \in \mathscr{L} \text{ et } \varphi \text{ est une } \mathscr{L}'\text{-substitution}\}.$$

Ainsi, si Fin désigne la famille des langages finis et Rat la famille des langages rationnels:

-**Fin** $\mathscr{L}$ est la plus petite de langages contenant $\mathscr{L}$ et fermée par union et produit,

-**Rat** $\mathscr{L}$ est la plus petite famille de langages contenant $\mathscr{L}$ et fermée par union, produit et étoile.

Nous désignerons par $\mathscr{C}(\mathscr{L})$ le *cône rationnel* engendré par $\mathscr{L}$ [2], c'est-à-dire:

$$\mathscr{C}(\mathscr{L}) = \{\tau(L) \mid L \in \mathscr{L}, \tau \text{ est une transduction rationnelle}\}.$$

$\mathscr{C}^{\cup}(\mathscr{L})$ désignera la fermeture par union de $\mathscr{C}(\mathscr{L})$ ('full semi AFL' engendré par $\mathscr{L}$ [6]). La FAL (full AFL, [5]) engendrée par $\mathscr{L}$, notée $\mathscr{F}(\mathscr{L})$, n'est autre que **Rat** $\mathscr{C}(\mathscr{L})$, tandis que **Fin** $\mathscr{C}(\mathscr{L})$ est le plus petit cône rationnel contenant $\mathscr{L}$ et fermé par union et produit.

Rat_2 désignera la famille des parties rationnelles du produit cartésien de deux monoïdes libres [4]. Il est connu [12] que la partie A de $X^* \times Y^*$ est rationnelle si et seulement si il existe un monoïde libre Z^*, un langage rationnel $K \subseteq Z^*$ et deux homomorphismes φ et ψ dans X^* et Y^* respectivement, tels que

$$A = \{(\varphi w, \psi w) \mid w \in K\}.$$

Définition 2. Le crochet ouvert d'un langage L par A appartenant à Rat_2, noté $\mathrm{CO}(L, A)$ est le langage

$$\mathrm{CO}(L, A) = \{a_1 w a_2 \mid (a_1, a_2) \in A,\ w \in L\}.$$

Dans le cas où $A = \{(w, w) \mid w \in \{a, b\}^*\}$, avec $\{a, b\} \cap X_L = \emptyset$, nous désignerons par $\mathrm{CO}(L)$ le langage

$$\mathrm{CO}(L) = \mathrm{CO}(L, \{(w, w) \mid w \in \{a, b\}^*\}).$$

Il est à noter que $\mathrm{CO}(L)$ est défini à un isomorphisme près.

Une famille de langages $\mathscr{L}$ est fermée par crochet ouvert si et seulement si A appartient à Rat_2 et L appartient à $\mathscr{L}$ impliquent $\mathrm{CO}(L, A)$ appartient à $\mathscr{L}$.

Nous désignerons par $\mathrm{CO}(\mathscr{L})$ le plus petite cône rationnel fermé par union et contenant la famille $\mathscr{L}'$ des crochets ouverts de $L \in \mathscr{L}$:

$$\mathscr{L}' = \{\mathrm{CO}(L, A) \mid L \in \mathscr{L}, A \in \mathrm{Rat}_2\}.$$

Dans le cas où $\mathscr{L}$ est un cône rationnel nous pouvons énoncer:

Théorème 1. *Si $\mathscr{L}$ est un cône rationnel, $\mathrm{CO}(\mathscr{L})$ est identique à la fermeture par union de $\mathscr{L}' = \{\mathrm{CO}(L, A) \mid L \in A \in \mathrm{Rat}_2\}$.*

Démonstration. Il est clair que $\mathrm{CO}(\mathscr{L})$ contient la fermeture par union de la famille $\mathscr{L}'$. Il nous suffit donc de vérifier que la fermeture par union de $\mathscr{L}'$ est un cône rationnel.

D'après le théorème de caractérisation des transductions rationnelles de Nivat [12], on sait que pour toute transduction rationnelle τ de X^* dans Y^* il existe un monoïde libre Z^*, un langage rationnel $K \subseteq Z^*$ et deux homomorphismes alphabétiques φ et ψ de Z^* dans X^* et Y^* respectivement, tels que $\forall f \in X^*$, $\tau(f) = \psi(\varphi^{-1} f \cap K)$.

Soit donc $L' \in \mathscr{L}'$, $A \in \mathrm{Rat}_2$ et $L \in \mathscr{L}$ tels que $L' = \mathrm{CO}(L, A)$:

(a) Soit φ un homomorphisme alphabétique de U^* dans T^* avec $L' \subseteq T^*$. Comme φ est alphabétique on a:

$$\varphi^{-1}(L') = \mathrm{CO}(\varphi^{-1} L, \varphi^{-1} A).$$

$\varphi^{-1} L$ appartient au cône rationnel $\mathscr{L}$ et $\varphi^{-1} A$ appartient à Rat_2 [1]. Donc $\varphi^{-1}(L') \in \mathscr{L}'$.

(b) Soit φ un homomorphisme de monoïdes libres. Alors $\varphi(L') = \mathrm{CO}(\varphi L, \varphi A)$ appartient à $\mathscr{L}'$ puisque $\varphi L \in \mathscr{L}$ et $\varphi A \in \mathrm{Rat}_2$ d'après la caractérisation de [12].

(c) Soit $K \in \mathrm{Rat}(X^*)$ reconnu par un automate fini ayant Q pour ensemble d'états, q_0 pour état initial, F pour états finals et δ pour fonction de transition.

Pour tout q et $q' \in Q$, soient

$$K_q = \{w \in X^* \mid \delta(q_0, w) = q\},$$

$$K_{qq'} = \{w \in X^* \mid \delta(q, w) = q'\},$$

$$K'_q = \{w \in X^* \mid \delta(q, w) \in F\}$$

ces ensembles sont rationnels et $K = \bigcup_{qq'} K_q K_{qq'} K'_{q'}$.

Soit $L' = \mathrm{CO}(L, A) \subseteq X^*$. On a clairement

$$L' \cap K = \bigcup_{qq' \in Q} \mathrm{CO}(L \cap K_{qq'}, A \cap (K_q \times K'_{q'}))$$

or $L \cap K_{qq'}$ appartient au cône rationnel $\mathscr{L}$ et si $A = \{(\varphi w, \psi w) \mid w \in R\}$ alors

$$A \cap (K_q \times K'_{q'}) = \{(\varphi w, \psi w) \mid w \in R \cap \varphi^{-1}(K_q) \cap \psi^{-1}(K'_{q'})\}$$

et donc $A \cap (K_q \times K_{q'})$ appartient à Rat_2.

Chacun des crochets ouverts $\mathrm{CO}(L \cap K_{qq'}, A \cap (K_q \times K'_{q'}))$ appartient donc à $\mathscr{L}'$ et leur union finie à la fermeture par union de $\mathscr{L}'$. D'où le résultat.

Dans le cas où $\mathscr{L}$ est un cône rationnel principal nous avons en outre:

Théorème 2. *Si $\mathscr{L}$ est un cône rationnel principal et L un générateur de $\mathscr{L}$, si X est un alphabet d'au moins deux lettres, disjoint de celui de L, alors $\mathrm{CO}(\mathscr{L})$ est un cône rationnel principal dont $\mathrm{CO}(L) = \mathrm{CO}(L, \{(w, w) \mid w \in X^*\})$ est un générateur.*

Démonstration. Soit $L_0 \in \mathrm{CO}(\mathscr{L})$. Nous allons montrer que L_0 est image par une transduction rationnelle de $\mathrm{CO}(L)$. Comme $\mathrm{CO}(L) \in \mathrm{CO}(\mathscr{L})$ le résultat en découlera.

D'après le Théorème 1 il existe $n \in \mathbf{N}$, n ensembles $A_j \in \mathrm{Rat}_2$ et n langages $M_j \in \mathscr{L}$ tels que $L_0 = \bigcup_{j=1}^n \mathrm{CO}(M_j, A_j)$. Comme $\mathscr{L}$ est un cône rationnel principal il existe n transductions rationnelles τ_j, $1 \leqslant j \leqslant n$, telles que $M_j = \tau_j(L_j)$, $1 \leqslant j \leqslant n$, où $L_1, \ldots, L_n$

sont des copies de L sur des alphabets $X_{L_1}, \ldots, X_{L_n}$ disjoints deux à deux. On peut supposer que ε n'appartient pas à L_j, $1 \leqslant j \leqslant n$.

D'autre part pour tout $j \in [1, n]$, il existe n langages rationnels locaux K_j et $2n$ homomorphismes alphabétiques φ_j et ψ_j, $1 \leqslant j \leqslant n$, tels que [12]

$$A_j = \{(\varphi_j w, \psi_j w) \mid w \in K_j\}.$$

On peut supposer les alphabets X_{K_j} et X_{L_j} disjoints deux à deux.

Soit alors X un alphabet d'au moins deux lettres, disjoint des précédents. Soit pour tout j, $1 \leqslant j \leqslant n$, $L_{0_j} = \mathrm{CO}(L_j)$. Il est facile de vérifier, étant donné la disjonction des alphabets qu'il existe des transductions rationnelles τ'_j telles que

$$L'_j = \mathrm{CO}(L_j, \{(w, w) \mid w \in K_j\}) = \tau'_j(L_0).$$

Soit alors pour tout j, $1 \leqslant j \leqslant n$, τ''_j la transduction rationnelle qui rejette les mots n'appartenant pas à $X^*_{K_j} X^*_{L_j} X^*_{K_j}$ et fonctionne comme φ_j sur le préfixe des mots inclus dans $X^*_{K_j}$, comme τ_j sur le facteur inclus dans $X^*_{L_j}$ et comme ψ_j sur le suffixe dans $X^*_{K_j}$. τ''_j se construit facilement étant donné la disjonction des alphabets. On obtient alors:

$$\mathrm{CO}(M_j, A_j) = \tau''_j\, \mathrm{CO}(L_j, \{(w, w) \mid w \in K_j\})$$

et par conséquent

$$L_0 = \bigcup_{j=1}^{n} \tau''_j \circ \tau'_j\, \mathrm{CO}(L_j)$$

et donc L_0 est bien image par transduction rationnelle de $\mathrm{CO}(L)$, puisque le cône rationnel $\mathscr{C}(\mathrm{CO}(L))$ est fermé par union puisque principal.

Corollaire 1. *Un cône rationnel principal $\mathscr{L}$ est fermé pour l'opérateur crochet ouvert si et seulement si pour tout générateur L de $\mathscr{L}$, $\mathrm{CO}(L)$ appartient à $\mathscr{L}$.*

Notation. Nous écrirons $\mathrm{CO}^i(\mathscr{L})$ pour $\mathrm{CO}(\cdots(\mathrm{CO}(\mathscr{L}))\cdots)$, i.e. l'application répétée i fois de l'opérateur crochet ouvert sur la famille $\mathscr{L}$. En particulier nous poserons $\mathrm{CO}^0(\mathscr{L}) = \mathscr{L}$.

On peut alors énoncer:

Corollaire 2. *Soit $\mathscr{L}$ un cône rationnel principal de générateur L. Si X est un alphabet d'au moins deux lettres disjoint de X_L et c un symbole n'appartenant pas à $X \cup X_L$, alors pour tout $i \geqslant 1$, $\mathrm{CO}^i(\mathscr{L})$ est un cône rationnel principal dont*

$$L_i = \{w_1 c w_2 c \cdots w_i c l c w_i \cdots c w_2 c w_1 \mid w_p \in X^*, 1 \leqslant p \leqslant i, l \in L\}$$

est un générateur.

Enfin le Théorème 2 permet de donner une autre caractérisation du cône rationnel $\mathrm{CO}(\mathscr{L})$:

Corollaire 3. *Soit $\mathscr{L}$ une famille de langages, alors*

$$\mathrm{CO}(\mathscr{L}) = \mathscr{C}^{\cup}\{\mathrm{CO}(L) \mid L \in \mathscr{L}\}.$$

3. Propriétés du crochet ouvert

Afin d'étudier les propriétés des cônes rationnels fermés par union et crochet ouvert (resp. union, produit et crochet ouvert) établissons un certain nombre de résultats préliminaires.

Définition 3. Etant donné un langage L nous appellerons *chevron de L*, noté $\mathrm{Ch}(L)$ ($\langle L \rangle$ dans [2]) le langage

$$\mathrm{Ch}(L) = \{a^n l b^n \mid n \in \mathbf{N}, l \in L\}$$

avec $\{a, b\} \cap X_L = \emptyset$.

Ainsi $\mathrm{Ch}(L) = \mathrm{CO}(L, (a, b)^*)$. On a alors le théorème suivant de Latteux [11]:

Théorème 3 [11]. *Si $\mathscr{L}$ est un cône rationnel fermé par union $L \subseteq X^*$ un langage, a et b deux lettres n'appartenant pas à X, la condition* $\mathrm{Ch}(L) = \mathrm{CO}(L, (a, b)^*) \in \mathrm{CO}(\mathscr{L})$ *implique $L \in \mathscr{L}$.*

On en déduit alors immédiatement:

Corollaire 4. *Soit L un langage et X un alphabet d'au moins deux lettres disjoint de X_L. Si $\mathscr{L}$ est un cône rationnel fermé par union, la condition*

$$\mathrm{CO}(L) = \mathrm{CO}(L, \{(w, w) \mid w \in X^*\}) \in \mathrm{CO}(\mathscr{L})$$

implique $L \in \mathscr{L}$.

Démonstration. Si $\varepsilon \notin L$ alors la condition $\mathrm{CO}(L) \in \mathrm{CO}(\mathscr{L})$ implique trivialement $\mathrm{Ch}(L) \in \mathrm{CO}(\mathscr{L})$ et donc $L \in \mathscr{L}$ d'après le Théorème 3.

Si $\varepsilon \in L$ et $\mathrm{CO}(L) \in \mathrm{CO}(\mathscr{L})$, alors

$$\mathrm{CO}(L - \{\varepsilon\}) = \mathrm{CO}(L) \cap X^* X_L X_L^* X^* \in \mathrm{CO}(\mathscr{L})$$

et de même que précédemment $L - \{\varepsilon\} \in \mathscr{L}$. Comme $\mathscr{L}$ est un cône rationnel fermé par union, il en résulte que $L \in \mathscr{L}$.

Corollaire 5. *Soit $\mathscr{L}$ un cône rationnel fermé par union et non clos par crochet ouvert, alors* $\mathrm{CO}^i(\mathscr{L}) \subsetneq \mathrm{CO}^{i+1}(\mathscr{L})$, $\forall i \geq 1$.

Démonstration. Puisque $\mathscr{L}$ n'est pas fermé par crochet ouvert il existe $L \in \mathrm{CO}(\mathscr{L}) \backslash \mathscr{L}$. On a clairement $\mathrm{Ch}(L) \in \mathrm{CO}^2(\mathscr{L})$. Si $\mathrm{Ch}(L)$ appartenait à $\mathrm{CO}(\mathscr{L})$ alors d'après le Théorème 3 on aurait $L \in \mathscr{L}$ ce qui contredit notre hypothèse. La preuve s'achève par une récurrence triviale.

Corollaire 6. *Soit $\mathscr{L}$ un cône rationnel fermé par union et non clos par crochet ouvert (resp. chevron), alors $\mathscr{C}^{\cup}_{\mathrm{CO}}(\mathscr{L})$ (resp. $\mathscr{C}^{\cup}_{\mathrm{Ch}}(\mathscr{L})$) le plus petit cône rationnel fermé par union et crochet ouvert (resp. chevron) contenant $\mathscr{L}$ est non principal.*

Démonstration. Soit $\mathscr{C} = \bigcup_{i \geq 0} \mathrm{CO}^i(\mathscr{L})$. Clairement $\mathscr{C}$ est inclus dans $\mathscr{C}^{\cup}_{\mathrm{CO}}(\mathscr{L})$. Or $\mathscr{C}$ est un cône rationnel fermé par union et crochet ouvert puisque $\mathrm{CO}^i(\mathscr{L})$ est un cône rationnel fermé par union, $\forall i \geq 0$, et si $L \in \mathscr{C}$ alors il existe i tel que $L \in \mathrm{CO}^i(\mathscr{L})$ et donc pour tout $A \in \mathrm{Rat}_2$, $\mathrm{CO}(L, A) \in \mathrm{CO}^{i+1}(\mathscr{L}) \subseteq \mathscr{C}$.

Donc $\mathscr{C}$ est bien égal à $\mathscr{C}^{\cup}_{\mathrm{CO}}(\mathscr{L})$ et le résultat découle du Corollaire 5. La preuve pour le chevron est identique.

Corollaire 7. *Soit $\mathscr{L}$ un cône rationnel fermé par union, L un langage et X un alphabet d'au moins deux lettres disjoint de X_L. Si*

$$[L] = \{wl\tilde{w} \mid l \in L,\ w \in X^*\} \in \mathscr{C}^{\cup}_{\mathrm{CO}}(\mathscr{L})$$

alors $[L] \in \mathscr{L}$.

Démonstration. Il suffit de remarquer qu'il existe une transduction rationnelle τ telle que $\mathrm{Ch}^p(L) = \tau([L])$, où Ch^p est l'application p fois répétée de l'opération chevron sur le langage L, $\forall p \geq 1$. Si donc $[L] \in \mathscr{C}^{\cup}_{\mathrm{CO}}(\mathscr{L})$, il existe $p \in \mathbf{N}$ tel que $[L] \in \mathrm{CO}^p(\mathscr{L})$, donc $\mathrm{Ch}^p(L) \in \mathrm{CO}^p(\mathscr{L})$. Le résultat découle alors de l'application, p fois répétée, du Théorème 3.

Il en résulte alors immédiatement:

Corollaire 8. *Soit $\mathscr{L}$ un cône rationnel fermé par union et $S_2 = \{wc\tilde{w} \mid w \in \{a, b\}^*\}$. Alors $S_2 \in \mathscr{C}^{\cup}_{\mathrm{CO}}(\mathscr{L})$ implique $S_2 \in \mathscr{L}$.*

L'étude des cônes rationnels fermés par union, produit et crochet ouvert repose essentiellement sur un résultat de [2] que nous rappelons tout d'abord:

Théorème 4 [2]. *Soit $\mathscr{L}$ un cône rationnel fermé par union et L un langage. La condition* $\mathrm{Ch}(L) \in \mathbf{Rat}\,\mathscr{L}$ *implique* $L \in \mathscr{L}$.

Corollaire 9. *Soit $\mathscr{L}$ un cône rationnel fermé par union, L un langage et X un alphabet d'au moins deux lettres disjoint de X_L. La condition* $\mathrm{CO}(L) = \mathrm{CO}(L, \{(w, w) \mid w \in X^*\}) \in \mathbf{Rat}\, \mathscr{L}$ *implique $L \in \mathscr{L}$.*

Ceci nous permet alors d'énoncer:

Théorème 5. *Soit $\mathscr{L}$ un cône rationnel fermé par union et crochet ouvert (resp. chevron). Alors $\mathscr{F}(\mathscr{L}) = \mathbf{Rat}\, \mathscr{L}$ est une* FAL *principale si et seulement si $\mathscr{L}$ est une cône rationnel principal.*

Démonstration. Si $\mathscr{L}$ est un cône rationnel principal alors il est connu que $\mathscr{F}(\mathscr{L})$ est une FAL principale [6].

Réciproquement si $\mathscr{F}(\mathscr{L}) = \mathscr{F}(L_0)$ pour quelque langage $L_0 \in \mathscr{F}(\mathscr{L})$ alors, puisque $\mathscr{L}$ est fermé par union, il existe un langage $L_1 \in \mathscr{L}$ tel que $L_0 \in \mathscr{F}(L_1)$.

Soit alors L un langage quelconque appartenant à $\mathscr{L}$ et X un alphabet d'au moins deux lettres disjoint de X_L. Comme $\mathscr{L}$ est fermé pour le crochet ouvert on a:

$$\mathrm{CO}(L) = \mathrm{CO}(L, \{(w, w) \mid w \in X^*\}) \in \mathscr{L} \subseteq \mathscr{F}(\mathscr{L}) = \mathbf{Rat}\, \mathscr{C}(L_1).$$

D'après le Corollaire 9 il en résulte que $L \in \mathscr{C}(L_1)$ puisque ce cône est fermé par union. Donc $\mathscr{L} \subseteq \mathscr{C}(L_1)$ i.e. $\mathscr{L}$ est un cône rationnel principal puisque $L_1 \in \mathscr{L}$.

Exemple 1. Il est connu que la FAL des langages linéaires est principale et engendrée par $S_2 = \{wc\tilde{w} \mid w \in \{a, b\}^*\}$. Si Mét désigne le cône rationnel des langages métalinéaires (fermeture par produit du cône rationnel engendré par S_2), nous avons $\mathscr{F}(S_2) = \mathbf{Rat}$ Mét. Comme Mét est un cône rationnel non principal, il résulte du Théorème 5 que la famille des langages métalinéaires n'est pas fermée par chevron.

De même du Théorème 4 résulte le corollaire:

Corollaire 10. *Soit $\mathscr{L}$ une famille de langages. On a les implications:*

(a) $\mathrm{CO}(\mathscr{L}) \subsetneq \mathbf{Fin}\, \mathrm{CO}(\mathscr{L}) \Rightarrow \mathbf{Fin}\, \mathrm{CO}(\mathscr{L}) \subsetneq \mathrm{CO}(\mathbf{Fin}\, \mathrm{CO}(\mathscr{L}))$,

(b) $\mathrm{CO}(\mathscr{L}) \subsetneq \mathbf{Rat}\, \mathrm{CO}(\mathscr{L}) \Rightarrow \mathbf{Rat}\, \mathrm{CO}(\mathscr{L}) \subsetneq \mathrm{CO}(\mathbf{Rat}\, \mathrm{CO}(\mathscr{L}))$.

Démonstration. Nous établissons (a), la preuve pour (b) est identique .

Soit $L \in \mathbf{Fin}\, \mathrm{CO}(\mathscr{L}) \backslash \mathrm{CO}(\mathscr{L})$, le langage $\mathrm{Ch}(L)$ appartient à $\mathrm{CO}(\mathbf{Fin}\, \mathrm{CO}(\mathscr{L}))$, si $\mathrm{Ch}(L)$ appartenait à $\mathbf{Fin}\, \mathrm{CO}(\mathscr{L}) \subseteq \mathbf{Rat}\, \mathrm{CO}(\mathscr{L})$ alors, d'après le Théorème 4, on aurait $L \in \mathrm{CO}(\mathscr{L})$, puisque $\mathrm{CO}(\mathscr{L})$ est un cône rationnel fermé par union, d'où une contradiction.

Théorème 6. *Soit $\mathscr{L}$ un cône rationnel fermé par union, L_1 et L_2 deux langages sur des alphabets disjoints X_1 et X_2. La condition $L_1L_2 \in \mathrm{CO}(\mathscr{L})$ implique $L_1 \in$* Rat *ou $L_2 \in \mathscr{L}$.*

Démonstration. Si $L_1L_2 \in \mathrm{CO}(\mathscr{L})$, alors d'après le Théorème 1 il existe un entier n, n ensembles $A_j \subseteq \mathrm{Rat}_2$ et n langages $M_j \in \mathscr{L}$, $1 \leqslant j \leqslant n$, tels que

$$L_1L_2 = \bigcup_{j=1}^{n} \mathrm{CO}(M_j, A_j).$$

Nous supposerons que A_j et M_j sont non vides, $1 \leqslant j \leqslant n$. Pour tout $w_1 \in L_1$ et tout $w_2 \in L_2$, il existe donc j, $1 \leqslant j \leqslant n$, $(a, b) \in A_j$ et $l \in M_j$ tels que $w_1w_2 = alb$. Deux cas se présentent:

(a) Pour tout $w_1 \in L_1$, il existe $w_2 \in L_2$, $1 \leqslant j \leqslant n$, $(a, b) \in A_j$, $l \in M_j$ tels que $w_1w_2 = alb$ et w_1 est un facteur gauche propre de a.

Soit alors

$$L = \{w \in X_1^* \mid \exists w' \in X_2X_2^*, b \in X_2^*, 1 \leqq j \leqq n : (ww', b) \in A_j\}.$$

Montrons que $L_1 = L$.

Les conditions de ce cas impliquent $L_1 \subseteq L$. Réciproquement, soit $w \in X_1^*$ tel qu'il existe $1 \leqslant j \leqslant n$, $w' \in X_2^*X_2$, $b \in X_2^*$ vérifiant $(ww', b) \in A_j$. Le langage M_j étant non vide, il existe $l \in M_j$ et le mot

$$ww'lb \in \mathrm{CO}(M_j, A_j) \subseteq L_1L_2.$$

La disjonction des alphabets implique alors $w \in L_1$. Soit alors $\pi_1(A_j) = \{a \mid \exists b : (a, b) \in A_j\}$. Il est clair que

$$L_1 = \left\{ w \in X_1^* \mid \exists w' \in X_2X_2^* : ww' \in \bigcup_{j=1}^{n} \pi_1(A_j) \right\}.$$

Or $\bigcup_{j=1}^{n} \pi_1(A_j)$ est un langage rationnel et L_1 est image de ce dernier dans une transduction rationnelle évidente. Donc L_1 est un langage rationnel.

(b) Il existe $w_1 \in L_1$ tel que pour tout $w_2 \in L_2$, $1 \leqslant j \leqslant n$, $(a, b) \in A_j$ et $l \in M_j$, la condition $w_1w_2 = alb$ implique que a est facteur gauche de w_1.

Vérifions que L_2 est égal au langage

$$L_2' = \{w'b \in X_2^* \mid \exists a, w \in X_1^*, 1 \leqslant j \leqslant n : aw = w_1, ww' \in M_j, (a, b) \in A_j\}.$$

Les conditions de ce cas impliquent $L_2 \subseteq L_2'$. Réciproquement soit $w'b \in X_2^*$ tel qu'il existe a, $w \in X_1^*$, $1 \leqslant j \leqslant n$, vérifiant $aw = w_1$, $ww' \in M_j$ et $(a, b) \in A_j$. On a donc $aww'b \in \mathrm{CO}(M_j, A_j) \subseteq L_1L_2$ et la disjonction des alphabets implique $w'b \in L_2$.

Pour tout w_1' facteur droit de w_1 et tout $1 \leqslant j \leqslant n$, soient

$$P_j(w_1') = \{w_1'' \mid w_1'w_1'' \in M_j\},$$

$$Q_j(w_1') = \{b \mid (a, b) \in A_j, aw_1' = w_1\}.$$

On a alors $L_2 = \bigcup_{w_1'} \{P_j(w_1')Q_j(w_1') \mid 1 \leqslant j \leqslant n\}$.

Or $P_j(w_1')$ est image de M_j par la transduction rationnelle qui rejette les mots ne commençant pas par w_1' et efface w_1' au début des autres. De plus $Q_j(w_1')$ est un langage rationnel image de a tel que $aw'_1 = w_1$ par la transduction associée à A_j. Il en résulte donc que $P_j(w'_1)Q_j(w'_1)$ est image de M_j dans une transduction rationnelle τ_{j,w'_1}. Nous pouvons donc écrire

$$L_2 = \bigcup_{w'_1} \{\tau_{j,w'_1}(M_j) \mid 1 \leq j \leq n\}$$

i.e. L_2 appartient à $\mathscr{L}$ puisque $\mathscr{L}$ est fermé par union.

On déduit immédiatement de ce théorème:

Corollaire 11. *Soit $\mathscr{L}$ un cône rationnel fermé par union. La condition $\mathscr{L} \subsetneq \mathrm{CO}(\mathscr{L})$ implique $\mathrm{CO}(\mathscr{L}) \subsetneq \mathbf{Fin}\,\mathrm{CO}(\mathscr{L})$.*

Démonstration. Soit $L_1 \in \mathrm{CO}(\mathscr{L}) \backslash \mathscr{L}$ et L_2 une copie de L_1 sur un alphabet disjoint de celui de L_1. Le langage L_1L_2 appartient à $\mathbf{Fin}\,\mathrm{CO}(\mathscr{L})$ et n'appartient pas à $\mathrm{CO}(\mathscr{L})$, puisque cela impliquerait $L_1 \in \mathscr{L}$ puisque L_1 n'est pas rationnel.

Des Théorèmes 5 et 6 on tire:

Théorème 7. *Soit $\mathscr{L}$ un cône rationnel fermé par union. La condition $\mathscr{L} \subsetneq \mathrm{CO}(\mathscr{L})$ implique:*

(a) *Le plus petit cône fermé par union, produit et crochet ouvert contenant $\mathscr{L}$ est non principal,*

(b) *La plus petite* FAL *fermée par crochet ouvert contenant $\mathscr{L}$ est non principale.*

Démonstration. Nous établissons (a), la preuve pour (b) est identique.

Soit la famille des cônes rationnels

$$\mathscr{L}_0 = \mathscr{L}, \qquad \mathscr{L}_1 = \mathrm{CO}(\mathscr{L}) \quad \text{et} \quad \mathscr{L}_{p+1} = \mathrm{CO}(\mathbf{Fin}\,\mathrm{CO}(\mathscr{L})).$$

Cette suite est strictement croissante; i.e. $\mathscr{L}_p \subsetneq \mathscr{L}_{p+1}, p \geq 1$. En effet $\mathrm{CO}(\mathscr{L}) \subsetneq \mathbf{Fin}\,\mathrm{CO}(\mathscr{L})$ implique $\mathbf{Fin}\,\mathrm{CO}(\mathscr{L}) \subsetneq \mathrm{CO}(\mathbf{Fin}\,\mathrm{CO}(\mathscr{L}))$ d'après le Corollaire 10 et

$$\mathbf{Fin}\,\mathrm{CO}(\mathscr{L}) \subsetneq \mathrm{CO}(\mathbf{Fin}\,\mathrm{CO}(\mathscr{L}_p))$$

implique

$$\mathrm{CO}(\mathbf{Fin}\,\mathrm{CO}(\mathscr{L})) \subsetneq \mathbf{Fin}\,\mathrm{CO}(\mathbf{Fin}\,\mathrm{CO}(\mathscr{L}))$$

d'après le Corollaire 11 puisque $\mathbf{Fin}\,\mathrm{CO}(\mathscr{L})$ est un cône fermé par union. Ainsi $\mathscr{L}_1 \subsetneq \mathscr{L}_2$ et $\mathscr{L}_2 \subsetneq \mathbf{Fin}\,\mathscr{L}_2$ et la démonstration s'achève par une récurrence simple.

Soit alors $\mathscr{L}_\infty = \bigcup_{p \geq 1} \mathscr{L}_p$. $\mathscr{L}_\infty$ est un cône rationnel fermé par union puisque $\mathscr{L}_p$ en est un, $\forall p \geq 1$. D'autre part si L_1 et L_2 appartiennent à $\mathscr{L}_\infty$ il existe p tel que L_1 et L_2 appartiennent à $\mathscr{L}_p$ et $L_1L_2 \in \mathbf{Fin}\,\mathscr{L}_p \subsetneq \mathscr{L}_\infty$ ainsi que $\mathrm{CO}(L, A) \in \mathscr{L}_{p+1} \subsetneq \mathscr{L}_\infty$ pour tout $A \in \mathrm{Rat}_2$. Donc L_∞ est fermé par produit et crochet ouvert.

Enfin tout cône rationnel fermé par union, produit, crochet ouvert et contenant $\mathscr{L}$ contient trivialement $\mathscr{L}_p$, $\forall p \geqslant 1$, donc $\mathscr{L}$ est le plus petit cône ayant ces propriétés. D'où le résultat.

4. Applications

L'opérateur crochet ouvert permet d'identifier un certain nombre de familles de langages vérifiables [14] (checking automata languages dans [9]). Tout d'abord, il est à noter que cette famille est trivialement fermée pour cet opérateur ainsi que pour les opérateurs crochet et chevron de [1].

Introduisons tout d'abord la famille des langages C_i, $\forall i \geqslant 1$:

Définition 4. Soit pour tout $i \geqslant 1$ les langages C_i définis par:

$$C_i = \{w_i c w_{i-1} \cdots c w_1 d w_1 c \cdots w_{i-1} c w_i \mid w_p \in \{a, b\}^*, 1 \leqslant p \leqslant i\}.$$

Les langages C_i sont vérifiables et non algébriques [14].

Du Corollaire 2 on déduit immédiatement, puisque $\{d\}$ est un générateur de Rat:

Corollaire 12. *Pour tout* $i \geqslant 1$, $\mathrm{CO}^i(\mathrm{Rat}) = \mathscr{C}(C_i)$.

Comme Rat n'est pas fermé pour le crochet ouvert puisque $\mathrm{CO}(\mathrm{Rat}) = \mathscr{C}(\{wcw \mid w \in \{a, b\}^*)$ il résulte du Corollaire 4:

Corollaire 13. *Pour tout* $i \geqslant 1$, $\mathscr{C}(C_i) \subsetneq \mathscr{C}(C_{i+1})$.

Du Corollaire 10 et du fait que $\mathscr{C}(C_i)$ n'est pas fermé par crochet ouvert il résulte également:

Corollaire 14. $\mathscr{C}(C_i) \subsetneq \mathbf{Fin}\ \mathscr{C}(C_i)$, $\forall i \geqslant 1$.

Enfin on peut énoncer:

Théorème 8. (a) $\mathscr{C}^{\cup}_{\mathrm{CO}}(\mathrm{Rat})$, *le plus petit cône rationnel fermé par union et crochet ouvert est non principal.*

(b) $\mathscr{C}^{\cup}_{\mathrm{CO}}(\mathrm{Rat})$, *le plus petit cône rationnel fermé par union, produit et crochet ouvert est non principal.*

(c) $\mathscr{F}_{\mathrm{CO}}(\mathrm{Rat})$, *la plus petite* FAL *fermée par crochet ouvert est non principale.*

Ces familles sont respectivement:

(a) $$\mathscr{C}^{\cup}_{\mathrm{CO}}(\mathrm{Rat}) = \bigcup_{i \geqslant 1} \mathscr{C}(C_i),$$

(b) $\mathscr{C}^{\bigcup\cdot}_{\mathrm{CO}}(\mathrm{Rat}) = \bigcup_{i\geq 1} \mathscr{C}_i$

avec $\mathscr{C}_1 = \mathrm{Rat}$ *et* $\mathscr{C}_{i+1} = \mathrm{CO}(\mathbf{Fin}\ \mathscr{C}_i)$, $i \geq 1$,

(c) $\mathscr{F}_{\mathrm{CO}}(\mathrm{Rat}) = \bigcup_{i\geq 1} \mathscr{F}_i$

avec $\mathscr{F}_1 = \mathrm{Rat}$ *et* $\mathscr{F}_{i+1} = \mathrm{CO}(\mathbf{Rat}\ \mathscr{F}_i)$, $i \geq 1$.

Enfin du Corollaire 8 il résulte directement:

Corollaire 15. *Les langages* $S_2 = \{wc\tilde{w} \mid w \in \{a, b\}^*\}$ *et* C_i, $i \geq 1$, *sont rationnellement incomparables.*

Nous terminerons cet article par la compairson des cônes rationnels $\mathscr{C}(C_i)$ avec les cônes rationnels $\mathscr{V}_j$ des langages vérifiables à mouvements bornés de degré j [15], i.e. tel que l'accepteur vérificateur fait au plus j mouvements en mémoire pour analyser le mot d'entrée.

Rappelons que pour tout $j \geq 1$, $\mathscr{V}_j$ est un cône rationnel principal dont $L_j = \{(wc\tilde{w}d)^j w \mid w \in \{a, b\}^*\}$ est un générateur, avec c et d deux symboles non dans $\{a, b\}$.

Du Corollaire 8 il résulte alors trivialement, puisque $S_2 = \{wc\tilde{w} \mid w \in \{a, b\}^*\} \in \mathscr{V}_j$, $\forall j \geq 1$ et $S_2 \notin \mathrm{Rat}$:

Lemme 1. $\forall i \geq 1$, $L_1 = \{wc\tilde{w}dw \mid w \in \{a, b\}^*\}$ *n'appartient pas à* $\mathscr{C}(C_1)$.

Par contre nous pouvons énoncer:

Lemme 2. $\forall i \geq 1$, $C_i \in \mathscr{V}_j$ *si et seulement si* $i \leq j$.

Démonstration. On construit sans difficulté un accepteur vérificateur à mouvement borné de degré i qui accepte le langage C_i. Réciproquement si C_i appartenait au cône rationnel $\mathscr{V}_{i-1}$ alors d'après un lemme d'itération de [15] pour tout mot $f = w_1cw_2c \cdots cw_idw_ic \cdots cw_2cw_1$ de C_i suffisamment long il existe un entier n et une factorisation de f en $f = x_1y_1x_2y_2 \cdots x_ry_rx_{r+1}$ avec $r \leq 2(i-1)+1$ telle que $f_k = x_1y_1^k \cdots x_ry_r^kx_{r+1} \in C_i$ et quelles que soient les n occurrences de lettres de f que l'on considère comme distinguées, alors il existe un indice i tel que y_i contient au moins une position distinguée. Si l'on considère successivement les lettres de $w_1, w_2, \ldots, w_i$ comme distinguées, il en résulte immédiatement que $r \geq 2i$, ce qui contredit notre hypothèse. D'où le résultat.

Les cônes rationnels $\mathscr{C}(C_i)$ forment donc une nouvelle hiérarchie de cônes de langages vérifiables dont les liens avec la hiérarchie $\mathscr{V}_j$ des langages vérifiables à mouvements bornés de degré j se résument par le théorème:

Théorème 9. (a) $\forall j \geq 1$, *le cône rationnel* $\mathscr{V}_j$ *n'est pas inclus dans*

$$\mathscr{C}^{\sqcup}_{\mathrm{CO}}(\mathrm{Rat}) = \bigcup_{i \geq 1} \mathscr{C}(C_i),$$

(b) $\forall i \geq 1, \forall j \geq i, \mathscr{C}(C_i) \subsetneq \mathscr{V}_j$.

Nous terminerons cet article en énonçant une conjecture sur les liens qui existent entre les trois opérateurs unaires chevron, crochet et crochet ouvert:

Conjecture. *Pour toute famille de langage* $\mathscr{L}$

$$\mathrm{Ch}(\mathscr{L}) = [\mathscr{L}] \cap \mathrm{CO}(\mathscr{L}).$$

Remerciements

L'auteur tient à remercier le rapporteur pour les nombreux conseils et commentaires qu'il a émis sur une précédente version de cet article.

Bibliographie

[1] L. Boasson, J.P. Crestin et M. Nivat, Familles de langages translatables et fermées par crochet, *Acta Informat.* **2** (1973) 383–393.
[2] L. Boasson et M. Nivat, Sur diverses familles de langages fermées par transduction rationnelle, *Acta Informat.* **2** (1973) 180–188.
[3] J.P. Crestin, Sur un langage quasi-rationnel d'ambiguïté non bornée, Thèse de 3ème cycle—Faculté des Sciences de Toulouse (1967) (Miméographiée).
[4] C. Elgot et J. Mezei, On relations defined by generalized finite automata, *IBM J. Res. Develop.* **9** (1965) 47–68.
[5] S. Ginsburg et S. Greibach, Abstract families of languages, *Mem. Amer. Math. Soc.* **87** (1969) 1–32.
[6] S. Ginsburg et S. Greibach, Principal AFL, *J. Comput. System Sci.* **4** (1970) 308–338.
[7] S. Ginsburg, S. Greibach and M. Harrison, One-way stack automata, *J. Assoc. Comput. Mach.* **14** (1967) 389–418.
[8] S. Greibach, An infinite hierarchy of context-free languages, *J. Assoc. Comput. Mach.* **16** (1969) 91–106.
[9] S. Greibach, Checking automata and one way stack languages, *J. Comput. System Sci.* **3** (1969) 196–217.
[10] S. Greibach, Syntactic Operators on full semi AFL's, *J. Computer System Sci.* **6** (1972) 30–76.
[11] M. Latteux, Edtol-Systèmes ultralinéaires et opérateurs associés, Publication No. 100 du laboratoire de Calcul, Université Lille (1977).
[12] M. Nivat, Transductions des langages de Chomsky, *Ann. Inst. Fourier* **18**, (1968) 339–456.
[13] M. Nivat, Transduction des langages de Chomsky, Chapitre VI—Thèse de Doctorat, Paris (1967) (Miméographiée).
[14] F. Rodriguez, Cônes d'accepteurs—application à l'étude d'une hiérarchie infinie de cônes rationnels de langages d'accepteurs vérificateurs, Thèse Docteur-Ingénieur, Toulouse (1973).
[15] F. Rodriguez, Une double hiérarchie infinie de langages vérifiables, *RAIRO* **R1** (1975) 5–20.

[16] F. Rodriguez, Familles de langages fermées par crochet et crochet ouvert, *Proc. 3rd GI conference on Theoretical Computer Science*, Lecture Notes in Computer Science **48** (Springer, Berlin, 1977) 154–168.
[17] R. Siromoney, Finite-turn checking automata, *J. Comput. System Sci.* **5** (1971) 549–559.
[18] S. Ginsburg et E. Spanier, Finite turn pushdown automata, SIAM J. Control (1966) 429–453.
[19] S. Greibach, Chain of full AFL's *Math. Systems Theory* **4** (1970) 231–242.