INFORMATION AND CONTROL 13, 565–593 (1968)

Définition et Etude des Bilangages Réguliers

C. PAIR ET A. QUERE

Institut Universitaire de Calcul Automatique, Université de Nancy, Nancy, France

The paper presents a definition of "arborescent structures" (*or ramifications*) enabling to process them algebraically. "Languages" formed with ramifications are defined and we call them "*bilanguages.*" The notion of regular language is then generalized into the notion of *regular bilanguage*. These regular bilanguages are compared to those generated by context-free grammars. Their properties are demonstrated; especially, a theorem analogous to Kleene's theorem on regular languages is stated.

INTRODUCTION

Dans de nombreux domaines du traitement de l'information, on emploie des "structures arborescentes", avec éventuellement plusieurs racines, souvent orientées "de gauche à droite" et dont chaque noeud est étiqueté par un nom appartenant à un certain alphabet (Fig. 1). En particulier, l'étude des langages conduit à considérer des ensembles de telles structures comme de véritables "langages à deux dimensions", que nous appellerons ici des *bilangages* (cf. par exemple Bar-Hillel *et al.* (1961), Ginsburg (1966), Lucas (1967)).

Dans l'étude qui suit, après un rappel succinct des principales propriétés des langages réguliers, ou langages de Kleene (§ 1 ; à ce sujet le lecteur pourra consulter Gross et Lentin, 1967), on présente une formalisation qui permet de traiter algébriquement ces structures aborescentes nommées *ramifications* (§ 2). On généralise ensuite la notion de langage régulier en celle de *bilangage régulier* (§ 3) ; en particulier on compare les bilangages réguliers à ceux qui sont engendrés par les grammaires "context free" (§ 4).

Parmi les propriétés des bilangages réguliers on peut distinguer celles qui sont des extensions directes des propriétés des langages réguliers : propriétés booléennes, stabilité pour le produit et l'opération étoile (§ 5 et 6), de celles qui sont propres à la "nouvelle dimension" des bilangages et liées à la possibilité de greffer un bilangage sur un autre

Ramification

FIG. 1

(§ 6). Ces dernières interviennent naturellement dans l'extension du théorème de Kleene aux bilangages réguliers (§ 7).

Les définitions et théorèmes 2.4.1, 4.2 et 4.8 sont énoncés par Pair (1967) qui indique aussi d'autres résultats, non repris ici, se rapportant à l'analyse syntaxique. Grâce à l'amabilité de M. Werner, les auteurs ont eu récemment connaissance de l'étude de Thatcher (1967) qui énonce en particulier ces théorèmes, ainsi que du travail de Thatcher et Wright (1966) : ces deux théories entrent dans le cadre qui va être présenté (cf. 4.4, note).

Dans toute l'étude, V désigne un ensemble *fini* appelé *vocabulaire.* V^* est le monoïde libre engendré par V, c'est-à-dire l'ensemble des suites finies d'éléments de V (ou *mots* sur V) ; un *langage* sur V est une partie de V^*.

1. RAPPELS SUR LES LANGAGES REGULIERS

1.1 DÉFINITION ET THÉORÈME. *Etant donné un emsemble V, un langage K sur V est appelé* ***langage régulier*** *ou* ***langage de Kleene*** *s'il satisfait à l'une des deux conditions équivalentes :*

(1) *il existe un monoïde fini M, une partie M' de M et un homomorphisme de monoïdes μ de V^* dans M, tels que $K = \overset{-1}{\mu}(M')$.*

(2) *il existe un automate fini acceptant K.*

On trouvera une définition des automates finis dans Rabin–Scott (1959). La notion d'automate fini peut être formalisée de la manière suivante : appelons *mémoire* sur V un ensemble S muni d'une loi de composition externe à opérateurs dans V notée $\cdot$:

$$(s, A) \in S \times V \mapsto s \cdot A \in S.$$

Par exemple V^* est une mémoire sur V en posant, pour $\alpha \in V^*$ et $A \in V$:

$$\alpha \cdot A = \alpha A.$$

Un homomorphisme de mémoires de V^* dans S, mémoire sur V, est une application ζ de V^* dans S telle que :

$$\zeta(\alpha A) = \zeta(\alpha)\cdot A.$$

Dans ces conditions, (2) est équivalent à l'assertion suivante :

(3) *il existe une mémoire finie S, une partie S' de S et un homomorphisme de mémoires ζ de V^* dans S tels que $K = \zeta^{-1}(S')$.*

1.2 *Langage local.* Etant donné deux sous ensembles D et F de V et un sous-ensemble τ de V^2, l'ensemble des mots non vides $A_1 \dots A_n$ ($A_i \in V$) tels que

$$A_1 \in D, A_n \in F \quad \text{et} \quad (A_i, A_{i+1}) \in \tau \qquad \text{pour } i = 1, \cdots, n-1$$

est un langage régulier. On l'appelle *langage local* ; il sera noté $[D, \tau, F]$. D sera appelé *ensemble des initiales, F ensemble des finales* et τ *ensemble des transitions.*

Les langages réguliers sur un vocabulaire fini sont les transformés des langages locaux par un homomorphisme entre monoïdes libres. Pour les langages réguliers qui ne contiennent pas le mot vide, cet homomorphisme peut toujours être une *transcription,* c'est-à-dire un homomorphisme qui transforme toute lettre en une lettre, autrement dit tout mot en un mot de longueur égale. Il en résulte que, pour prouver une propriété sur les langages réguliers, il suffit de s'assurer qu'elle est vraie sur les langages locaux et qu'elle est conservée par transcription. C'est en s'inspirant de cette méthode, par exemple, qu'on peut prouver le théorème suivant :

1.3 Théorème de Kleene. *L'ensemble des langages réguliers sur un vocabulaire V est le plus petit ensemble $\mathcal{K}$ de langages sur V tel que :*

(a) *$\mathcal{K}$ continent les langages finis ;*

(b) *Si K et K' appartiennent à $\mathcal{K}$, $K \cup K'$, KK' et K^* appartiennent à $\mathcal{K}$.*[1]

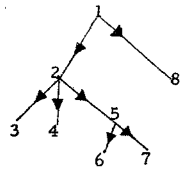
2. DEFINITIONS PRELIMINAIRES : RAMIFICATIONS, BILANGAGES, BILANGAGES GRAMMATICAUX

2.1 Définition des ramifications sur un ensemble V

2.1.1 *Arborescence* (Berge, 1958). On appelle *arborescence* un graphe fini sans circuit tel que :

(a) il existe un point a qui n'est extrémité d'aucun arc ;

[1] Notations : $KK' = \{\alpha\alpha' \mid \alpha \in K, \alpha' \in K'\}$; $\quad K^* = \bigcup_{p=0}^{\infty} K^p$.



Arborescence de racine 1

FIG. 2

(b) tout point $x \neq a$ est l'extrémité d'un arc unique.

a s'appelle la *racine* de l'arborescence.

2.1.2 *Orientation d'une arborescence.* Une orientation d'une arborescence[2] (E, Γ) est un ordre partiel O dans l'ensemble E tel que :

(a) les restrictions de O à chacun des ensembles $\Gamma(x)$ $(x \in E)$ sont des ordres totaux ;

(b) si $y \in \Gamma(x)$ et $z \notin \Gamma(x)$, y et z ne sont pas comparables par la relation O.

Un triplet (E, Γ, O) où (E, Γ) est une arborescence et O une de ses orientations s'appelle *arborescence orientée.*

Soient deux arborescences orientées (E, Γ, O) et (E', Γ', O') ; un *isomorphisme* de la première sur la deuxième est une bijection h de E sur E' telle que

$$(\forall x, y \in E)[(x\Gamma y \Leftrightarrow h(x)\Gamma' h(y)) \quad \text{et} \quad (xOy \Leftrightarrow h(x)O'h(y))].$$

2.1.3. *Pseudo-arborescence sur un ensemble V.* Etiqueter les points d'une arborescence par des éléments d'un ensemble V revient à définir une application de l'ensemble des points de l'arborescence dans V. Mais il est souhaitable que deux arborescences orientées isomorphes, dont les points qui se correspondent dans l'isomorphisme ont même étiquette, déterminent la même ramification. Aussi sommes-nous conduits à la définition suivante :

Soit un ensemble V. Dans l'ensemble des couples (I, f) formés par une arborescence orientée I dont les points sont des entiers, et une applica-

[2] Rappelons qu'un graphe est un couple (E, Γ) formé d'un ensemble E (ensemble des points du graphe) et d'une relation binaire Γ dans E ; on note $\Gamma(x)$ l'ensemble des $y \in E$ tels que $x\Gamma y$.

tion f de l'ensemble des points de I dans V, introduisons la relation d'équivalence :
$(I, f) \sim (I', f') \Leftrightarrow$ il existe un isomorphisme h de I sur I' tel que $f = f' \circ h$.[3]

Une pseudo-arborescence sur V est une classe de cette équivalence.

Elle pourra être représentée par l'un de ses couples (I, f). On appelle *ordre* de cette pseudo-arborescence le nombre des points de l'arborescence I ; pour chaque $A \in V$, on appelle *nombre d'occurrences* de A dans la pseudo-arborescence le nombre de points x de I tels que $f(x) = A$: l'ordre est la somme des nombres d'occurrences des divers éléments de V.

Une pseudo-arborescence d'ordre 1 sera identifiée à l'unique élément de V qui y possède une occurrence. Ainsi l'ensemble V est contenu dans l'ensemble des pseudo-arborescences sur lui-même.

L'ensemble des pseudo-arborescences sur V sera noté $\mathcal{A}(V)$.

2.1.4. *Ramifications sur un ensemble* V . Nous appellerons *ramification* sur V toute suite finie de pseudo-arborescences sur V. L'ensemble des ramifications sur V, c'est-à-dire le monoïde libre engendré par l'ensemble des pseudo-arborescences sur V, sera noté $\hat{V}$. La loi de composition de ce monoïde sera appelée *produit* et notée $\rightarrow$; l'élément neutre de cette loi sera nommée *ramification vide* et noté Λ.

La figure 3 schématise deux ramifications r' et r'' et leur produit $r' \rightarrow r''$.

Par définition :

la longueur d'une ramification est le nombre des pseudo-arborescences qui la constituent ;

l'ordre d'une ramification est la somme des ordres des pseudo-arborescences qui la constituent ;

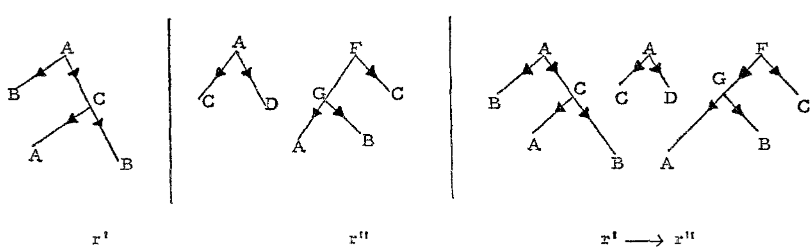


FIG. 3

[3] Le signe $\circ$ désigne la composition des applications.

le nombre d'occurrences de $A \in V$ dans une ramification est la somme des nombres d'occurrences de A dans les pseudo-arborescences qui la constituent.

De même qu'on appelait langage sur V toute partie de V^*, on appellera *bilangage sur V* toute partie de $\hat{V}$. Comme V est inclus dans $\mathcal{A}(V)$, V^* est inclus dans $\hat{V}$ et tout langage est un bilangage.

Le produit et le produit itéré (ou opération étoile, cf. 1.3) sur les langages sont prolongés aux bilangages de la façon suivante : si L et L' sont deux bilangages sur V

$$L \rightarrow L' = \{r \rightarrow r' \mid r \in L, r' \in L'\}$$

est appelé *produit de L et L'* ; en posant $L^0 = \{\Lambda\}$ et $L^p = L^{p-1} \rightarrow L$ pour $p \geqslant 1$, on peut définir

$$L^* = \bigcup_{p=0}^{\infty} L^p$$

appelé *produit itéré de L.*

La définition des ramifications qui vient d'être donnée, si elle est intuitive, n'en reste pas moins relativement complexe. La théorie algébrique qui suit permettra d'engendrer les ramifications de $\hat{V}$ à partir de V grâce au produit et à une loi de composition externe, et ensuite de ne plus guère employer explicitement la définition.

2.2 Etude algébrique de $\hat{V}$

2.2.1. *Enracinement.* Ce sera une loi de composition externe à opérateurs dans l'ensemble V, notée $\uparrow$ avec opérateurs à gauche : intuitivement, la ramification $A \uparrow r$ est la pseudo-arborescence obtenue en adjoignant à r une racine d'étiquette A. La figure 4 présente $B \uparrow r'$ et $A \uparrow r''$, r' et r'' étant prises sur la figure 3.

Soit une ramification r sur V, et un élément A de V.

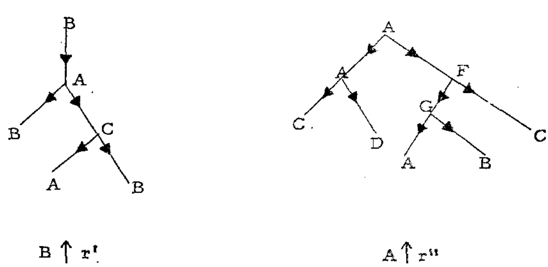


Fig. 4

Si r est la ramification vide, $A \uparrow r = A$.

Supposons que r est une suite de pseudo-arborescences $(E_i, \Gamma_i, O_i, f_i)$, pour $i = 1, 2, \cdots, n$; on peut toujours supposer les ensembles E_i deux à deux disjoints ; désignons par a_i la racine de (E_i, Γ_i). $A \uparrow r$ est la pseudo-arborescence (E, Γ, O, f) définie par :

E est la réunion des E_i et d'un élément a n'appartenant à aucun E_i ;

$$x\Gamma y \Leftrightarrow (\exists i)(x\Gamma_i y) \quad \text{ou} \quad (x = a \quad \text{et} \quad (\exists i)(y = a_i)) \ ;$$

$$xOy \Leftrightarrow (\exists i)(xO_i y) \quad \text{ou} \quad (\exists i, j)(i \leqslant j \quad \text{et} \quad x = a_i \quad \text{et} \quad y = a_j \ ;$$

$$f(x) = f_i(x) \text{ si } x \in E_i, f(a) = A.$$

On voit aisément que $A \uparrow r$ est bien une pseudo-arborescence et que réciproquement :

2.2.2. Proposition. *Pour toute pseudo-arborescence s sur V, il existe un élément A de V et un seul, une ramification r sur V et une seule, tels que $s = A \uparrow r$.*

Si L est un bilangage on notera $A \uparrow L$ le bilangage $\{A \uparrow r \mid r \in L\}$ et, pour $E \subset V$, $E \uparrow L$ désignera le bilangage $\{A \uparrow r \mid A \in E, r \in L\}$.

2.2.3. *Binoïde sur un ensemble V.* On appelle *binoïde* sur un ensemble V un ensemble muni d'une loi de composition interne associative, admettant un élément neutre et d'une loi de composition externe à opérateurs dans V. $\hat{V}$ est donc un binoïde sur V.

Dans la suite les deux lois d'un binoïde sur V seront, en général, notées $\rightarrow$ et $\uparrow$; l'élément neutre sera noté e, sauf pour $\hat{V}$ dont l'élément neutre est Λ.

De même que pour les structures algébriques classiques, on définit un sous-binoïde d'un binoïde $\mathcal{B}$ comme une partie $\mathcal{B}_0$ de $\mathcal{B}$ qui contient e et qui est stable pour $\rightarrow$ et $\uparrow$ (si $r \in \mathcal{B}_0$, $s \in \mathcal{B}_0$, $A \in V$, alors $r \rightarrow s \in \mathcal{B}_0$ et $A \uparrow r \in \mathcal{B}_0$).

Si $\mathcal{B}$ et $\mathcal{B}'$ sont deux binoïdes sur le même ensemble V, un *homomorphisme de binoïdes* (ou, simplement, homomorphisme) de $\mathcal{B}$ dans $\mathcal{B}'$ est une application ψ de $\mathcal{B}$ dans $\mathcal{B}'$, telle que, pour r et s dans $\mathcal{B}$ et A dans V :

$$\psi(r \rightarrow s) = \psi(r) \rightarrow \psi(s) \ ; \qquad \psi(e) = e \ ; \qquad \psi(A \uparrow r) = A \uparrow \psi(r).$$

2.3 Propriétés fondamentales de $\hat{V}$

De la définition des ramifications comme suites de pseudo-arborescences et de la proposition 2.2.2, il résulte immédiatement :

2.3.1 Proposition. *Pour tout* $r \in \hat{V}$, *non vide, il existe* $A \in V$, $r' \in \hat{V}$, $r'' \in \hat{V}$ *uniques tels que*

$$r = r' \to (A \uparrow r'').$$

On en déduit un principe de récurrence dans $\hat{V}$:

2.3.2 *Principe de récurrence.* Soit $\mathcal{P}$ un prédicat tel que

(a) $\mathcal{P}(\Lambda)$ soit vrai,
(b) $(\forall r \in \hat{V}, a \in \mathcal{A}(V))(\mathcal{P}(r)$ et $\mathcal{P}(a) \Rightarrow \mathcal{P}(r \to a))$,
(c) $(\forall r \in \hat{V})(\forall A \in V)(\mathcal{P}(r) \Rightarrow \mathcal{P}(A \uparrow r))$;

alors, $\mathcal{P}(r)$ est vrai pour tout $r \in \hat{V}$.

On montre que $\mathcal{P}(r)$ est vrai en raisonnant par récurrence sur l'ordre de la ramification r, grâce à la proposition 2.3.1.

Soit $\mathcal{B}$ un sous-binoïde de $\hat{V}$; en utilisant le principe de récurrence, on voit que toute ramification sur V appartient à $\mathcal{B}$, autrement dit que $\mathcal{B} = \hat{V}$: $\hat{V}$ ne contient aucun sous-binoïde autre que lui-même. En particulier, l'ensemble des combinaisons finies, par $\to$ et $\uparrow$, d'éléments de V, qui est un sous-binoïde de $\hat{V}$, est égal à $\hat{V}$: *toute ramification sur* V *est une combinaison finie par* $\to$ *et* $\uparrow$, *d'éléments de* V.

Exemples : Pour les ramifications de la figure 3 :

$$r' = A \uparrow (B \to (C \uparrow (A \to B)));$$

$$r'' = (A \uparrow (C \to D)) \to (F \uparrow ((G \uparrow (A \to B)) \to C)).$$

2.3.3. *Les applications de* $\hat{V}$ *dans un ensemble* E seront commodément définies par récurrence. Plus précisément :

Théorème. *Soient un ensemble* E, *un élément* e *de* E, *deux applications* f_1 *de* $\hat{V} \times \mathcal{A}(V) \times E^2$ *dans* E *et* f_2 *de* $V \times \hat{V} \times E$ *dans* E. *Il existe une application* λ *et une seule, de* $\hat{V}$ *dans* E, *telle que* :

(a) $\lambda(\Lambda) = e$;
(b) $(\forall r \in \hat{V})(\forall a \in \mathcal{A}(V))[\lambda(r \to a) = f_1(r, a, \lambda(r), \lambda(a))]$;
(c) $(\forall A \in V)(\forall r \in \hat{V})[\lambda(A \uparrow r) = f_1(A, r, \lambda(r))]$.

On en déduit aisèment :

2.3.4. Théorème. *Pour tout binoïde* $\mathcal{B}$ *sur* V, *il existe un homomorphisme de binoïdes et un seul de* $\hat{V}$ *dans* $\mathcal{B}$.

Ce résultat conduit à nommer $\hat{V}$ *binoïde libre sur* V.

Nous définirons maintenant par récurrence quelques applications simples.

2.3.5. *Mot des racines d'une ramification.* Appelons *mot des racines de* $r \in \hat{V}$ le mot $\rho(r)$ sur le vocabulaire V défini par :

$$\rho(\Lambda) = \Lambda\ ; \qquad \rho(r \to s) = \rho(r)\rho(s)\ ; \qquad \rho(A \uparrow r) = A.$$

Si le mot des racines de r est $A_1 \ldots A_n$, A_i est appelée la *$i^{\text{ème}}$ racine de r.*

Par exemple, le mot des racines de la ramification $r' \to r''$, prise dans la figure 3, est AAF.

2.3.6. PROPOSITION. *Soient α' et α'' deux mots sur V et une ramification r dont le mot des racines est $\alpha'\alpha''$. Il existe deux ramifications r' et r'', chacune unique, telles que $\rho(r') = \alpha'$, $\rho(r'') = \alpha''$ et $r = r' \to r''$.*

(Ce résultat se démontre par récurrence sur la longueur de α'').

2.3.7. *Mot des feuilles d'une ramification.* Appelons *mot des feuilles* de r le mot $\varphi(r)$ sur V tel que :

$$\varphi(\Lambda) = \Lambda\ ; \qquad \varphi(r \to s) = \varphi(r)\varphi(s)\ ;$$
$$\varphi(A \uparrow r) = \text{Si } \varphi(r) \neq \Lambda \text{ alors } \varphi(r) \text{ sinon } A.$$

Si le mot des feuilles de r est $A_1 \ldots A_n$, on dit que A_i est la *$i^{\text{ème}}$ feuille de r.*

Sur la figure 3, par exemple : $\varphi(r' \to r'') = BABCDABC$.

2.3.8. *Familles d'une ramification.* Pour la ramification $r' \to r''$ de la figure 3 nous dirons que le mot CD est *une famille de prédécesseur* A, que le mot vide est une famille de prédécesseur D.

Soit A un élément de V et F_A l'application de $\hat{V}$ dans l'ensemble $\mathfrak{P}(V^*)$ des parties de V^* définie par :

$$F_A(\Lambda) = \varnothing\ ; \qquad F_A(r \to s) = F_A(r) \cup F_A(s)\ ;$$
$$F_A(B \uparrow r) = \text{Si } B = A \text{ alors } F_A(r) \cup \{\rho(r)\} \text{ sinon } F_A(r).$$

Par définition $F_A(r)$ est *l'ensemble des familles de prédécesseur A dans r.*

Sur la figure 3, par exemple, $F_A(r' \to r'') = \{BC, CD, \Lambda\}$.

φ, ρ et F_A garderons la même signification dans tout ce qui suit.

2.3.9. *Ramification réfléchie d'une ramification.* Intuitivement il s'agit d'inverser l'orientation d'une ramification. Introduisons dans $\hat{V}$ la loi interne $\leftarrow$: $r \leftarrow s = s \to r$. L'homomorphisme ψ de $(\hat{V}, \to, \uparrow)$ dans $(\hat{V}, \leftarrow, \uparrow)$ transforme toute ramification r en ce que nous nommerons sa *ramification réfléchie* $\tilde{r}$. La ramification réfléchie de $\tilde{r}$ est r : en effet le carré ψ^2 de l'homomorphisme ψ est l'identité, puisqu'il vérifie

$$\psi^2(r \to s) = \psi^2(r) \to \psi^2(s)\ ; \qquad \psi^2(\Lambda) = \Lambda\ ; \qquad \psi^2(A \uparrow r) = A \uparrow \psi^2(r).$$

Si L est un bilangage sur V le bilangage $\tilde{L} = \psi(L)$ est appelé *bilangage réfléchi de* L. Il est évident que si L est un langage sur V, $\tilde{L}$ est le langage réfléchi (ou image miroir) de L.

2.3.10. *Transcription d'une ramification.* On donne deux ensembles V et V' et une application γ de V dans V'. Remplacer chaque "étiquette" A d'une ramification sur V par $\gamma(A) \in V'$ c'est définir une application $\hat{\gamma}$ de $\hat{V}$ dans $\hat{V}'$ telle que :

$$\hat{\gamma}(\Lambda) = \Lambda \,; \qquad \hat{\gamma}(r \rightarrow s) = \hat{\gamma}(r) \rightarrow \hat{\gamma}(s) \,; \qquad \hat{\gamma}(A \uparrow r) = \gamma(A) \uparrow \hat{\gamma}(r).$$

L'application $\hat{\gamma}$ s'appele *transcription* (associée à γ). Si L est un bilangage sur V nous dirons que $\hat{\gamma}(L)$ est *transcrit* de L.

Il est clair que :

la restriction à V^* d'une transcription $\hat{\gamma}$ est une transcription de monoïdes (1.2) ;

Pour toute ramification r : $\hat{\gamma} \circ \varphi(r) = \varphi \circ \hat{\gamma}(r), \hat{\gamma} \circ \rho(r) = \rho \circ \hat{\gamma}(r)$.

2.4 Bilangages grammaticaux

2.4.1 *Grammaire.* Une grammaire est un triplet $G = (V, ::=, X)$ où

V est un ensemble fini appelé *vocabulaire* de la grammaire;

$::=$ est une relation binaire entre V et le monoïde libre V^* telle que, pour tout $A \in V$, l'ensemble des mots α vérifiant $A ::= \alpha$ soit un langage régulier ; on notera K_A ce langage et ses éléments seront appelés les *productions* de A.

X est un langage régulier sur V, *l'ensemble des axiomes de la grammaire.*

2.4.2 *Bilangage engendré par une grammaire.* Une ramification sur V est *engendrée au sens large par la grammaire* G lorsque chacune de ses familles α de prédécesseur A vérifié $A ::= \alpha$ (notons que le mot α peut être le mot vide Λ ; A est alors une feuille de la ramification).

Pour tout langage régulier K, notons $\mathcal{L}(G, K)$ l'ensemble des ramifications engendrées au sens large par G, dont le mot des racines appartient à K ; c'est un bilangage sur V. En particulier :

$\mathcal{L}(G, V^*)$, aussi noté $\mathcal{L}(G)$, est appelé *bilangage engendré au sens large par* G ;

$\mathcal{L}(G, X)$, aussi noté $\mathcal{S}(G)$, est appelé *bilangage engendré au sens strict* (ou, simplement, *engendré*) par G.

Les bilangages engendrés par une grammaire sont dits *bilangages grammaticaux.*

Pour que la ramification vide appartienne à $\mathcal{S}(G)$ il faut et il suffit que le mot vide appartienne à X. D'autre part il est immédiat que :

2.4.3 PROPOSITION. *Si r et r' appartiennent à $\hat{V}$ et A à V :*
(a) $r \to r' \in \mathcal{L}(G) \Leftrightarrow r \in \mathcal{L}(G)$ et $r' \in \mathcal{L}(G)$;
(b) $A \uparrow r \in \mathcal{L}(G) \Leftrightarrow r \in \mathcal{L}(G)$ et $A ::= \rho(r)$.

On en déduit immédiatement, en utilisant une démonstration par récurrence dans $\hat{V}$, que $\mathcal{L}(G)$ est le plus petit bilangage sur V contenant Λ, stable pour $\to$, contenant $A \uparrow r$ lorsqu'il contient r et que $A ::= \rho(r)$.

2.4.4 *Langage engendré par une grammaire.* L'ensemble des mots des feuilles des ramifications de $\mathcal{S}(G)$ s'appelle *langage engendré par la grammaire G.*

On montre (Pair, 1965) que cette définition du langage engendré par G équivaut à celle de Chomsky (1963) lorsque les langages K_A sont finis et que G admet un seul axiome de longueur 1, et que même lorsque ces hypothèses ne sont pas satisfaites le langage engendré par G est encore un langage *"context-free"* ou *langage de Chomsky.*

Rappelons qu'un homomorphisme de monoïdes libres transforme un langage de Chomsky en un langage de Chomsky (Bar-Hillel, Perles et Shamir, 1961).

3. DEFINITION DES BILANGAGES REGULIERS

3.1 DÉFINITION. Un bilangage L sur V est dit *régulier* s'il existe un binoïde fini $\mathcal{B}$, une partie $\mathcal{B}'$ de $\mathcal{B}$ tels que, si ψ est l'homomorphisme de $\hat{V}$ dans $\mathcal{B}$, $L = \psi^{-1}(\mathcal{B}')$. Nous dirons que L est associé au triplet ($\mathcal{B}$, $\mathcal{B}'$, ψ).

Il s'agit là d'une généralisation immédiate de la définition 1.1 (1). En se basant sur 1.1 (3), on peut envisager la définition qui suit, qui généralise la notion d'automate fini. Nous montrerons qu'elle conduit aux mêmes classes de bilangages que 3.1. Ce résultat ne sera pas utilisé dans la suite.

3.2 *Bilangages réguliers à droite.* Un *semi-binoïde* sur V est un ensemble muni d'une loi de composition interne (notée $\to$) et d'une loi de composition externe à opérateurs dans V (notée $\uparrow$). Appelons *semi-homomorphisme à droite* une application σ de $\hat{V}$ dans un semi-binoïde $\mathcal{R}$ telle qu

pour $r \in \hat{V}$, $a \in \mathcal{A}(V)$ et $A \in V$:

$$\sigma(r \to a) = \sigma(r) \to \sigma(a) \; ; \qquad \sigma(A \uparrow r) = A \uparrow \sigma(r).$$

Un bilangage L sur V est dit *régulier à droite* s'il existe un semi-binoïde fini $\mathcal{R}$, une partie $\mathcal{R}'$ de $\mathcal{R}$ et un semi-homomorphisme à droite σ de $\hat{V}$ dans $\mathcal{R}$, tels que $L = \sigma^{-1}(\mathcal{R}')$. L est dit associé au triplet $(\mathcal{R}, \mathcal{R}', \sigma)$.

3.3 Théorème. *L'ensemble des bilangages réguliers sur V est l'ensemble des bilangages réguliers à droite sur V.*

Il est clair que tout bilangage régulier est un bilangage régulier à droite.

Réciproquement soit L un bilangage régulier à droite associé à $(\mathcal{R}, \mathcal{R}', \sigma)$. L'ensemble $\mathcal{R}^{\mathcal{R}}$ des applications de $\mathcal{R}$ dans $\mathcal{R}$ est fini.

A toute ramification r, on peut associer une application $\psi(r)$ de $\mathcal{R}$ dans $\mathcal{R}$, définie par récurrence sur la longueur de r :

$\psi(\Lambda)$ est l'identité dans $\mathcal{R}$;

pour $r \in \hat{V}$ et $a \in \mathcal{A}(V)$, pour $b \in \mathcal{R}$,

$$\psi(r \rightarrow a)(b) = \psi(r)(b) \rightarrow \sigma(a).$$

Il est alors immédiat, par récurrence sur la longueur de r, que :

(1) si on pose $\sigma(\Lambda) = e$, $\psi(r)(e) = \sigma(r)$;

(2) $\psi(r \rightarrow s) = \psi(s) \circ \psi(r)$.

D'autre part :

$$\psi(A \uparrow r)(b) = b \rightarrow \sigma(A \uparrow r) = b \rightarrow (A \uparrow \sigma(r)).$$

ψ est l'homomorphisme de $\hat{V}$ dans $\mathcal{R}^{\mathcal{R}}$, muni des deux lois suivantes qui en font un binoïde :

$$\lambda \rightarrow \lambda' = \lambda' \circ \lambda$$

$A \uparrow \lambda$ transforme $b \in \mathcal{R}$ en $b \rightarrow (A \uparrow \lambda(e))$.

Enfin :

$$r \in L \Leftrightarrow \sigma(r) \in \mathcal{R}' \Leftrightarrow \psi(r)(e) \in \mathcal{R}'$$

L est l'image réciproque par ψ de l'ensemble des $\lambda \in \mathcal{R}^{\mathcal{R}}$ tels que $\lambda(e) \in \mathcal{R}'$.

On peut définir de manière analogue les bilangages réguliers à gauche et montrer qu'il s'agit encore des bilangages réguliers.

3.4 Proposition. *L'image par transcription d'un bilangage régulier est un bilangage régulier.*

Soient deux ensembles V et V_1, θ une transcription de $\hat{V}$ dans $\hat{V}_1$, et L un bilangage régulier sur V, associé au triplet $(\mathcal{B}, \mathcal{B}', \psi)$. Soit ψ_1 l'application de $\hat{V}_1$ dans $\mathfrak{P}(\mathcal{B})$ définie par :

$$\psi_1(r_1) = \psi(\theta^{-1}(r_1)), \quad \text{pour} \quad r_1 \in \hat{V}_1 .$$

D'après la définition d'une transcription, il est immédiat que :

(1) pour $r_1, s_1 \in \hat{V}_1$, $\quad \theta^{-1}(r_1 \to s_1) = \theta^{-1}(r_1) \to \theta^{-1}(s_1)$

(2) pour $A_1 \in V_1, r_1 \in \hat{V}_1$, $\quad \theta^{-1}(A_1 \uparrow r_1) = \theta^{-1}(A_1) \uparrow \theta^{-1}(r_1)$

Par suite :

$$\psi_1(r_1 \to s_1) = \psi(\theta^{-1}(r_1)) \to \psi(\theta^{-1}(s_1)))$$
$$\psi_1(A_1 \uparrow r_1) = \theta^{-1}(A_1) \uparrow \psi(\theta^{-1}(r_1)).$$

Ainsi, en munissant $\mathfrak{P}(\mathcal{B})$ de la structure suivante de binoïde sur V_1, ψ_1 est l'homomorphisme de $\hat{V}_1$ dans $\mathfrak{P}(\mathcal{B})$:

pour $\mathfrak{M}, \mathfrak{M}' \in \mathfrak{P}(\mathcal{B}), A_1 \in V_1$:

$$\mathfrak{M} \to \mathfrak{M}' = \{m \to m' \mid m \in \mathfrak{M}, \quad m' \in \mathfrak{M}'\}$$
$$A_1 \uparrow \mathfrak{M} = \{A \uparrow m \mid m \in \mathfrak{M}, \quad \theta(A) = A_1\}.$$

De plus :

$$\begin{aligned} r_1 \in \theta(L) &\Leftrightarrow (\exists r \in L)(\theta(r) = r_1) \\ &\Leftrightarrow (\exists r \in \hat{V})(\psi(r) \in \mathcal{B}' \quad \text{et} \quad \theta(r) = r_1) \\ &\Leftrightarrow \psi(\theta^{-1}(r_1)) \cap \mathcal{B}' \neq \varnothing \end{aligned}$$

$\theta(L)$ est l'image réciproque par ψ_1 de l'ensemble des parties $\mathfrak{M}$ de $\mathcal{B}$ qui ne sont pas disjointes de $\mathcal{B}'$.

4. EXEMPLES DE BILANGAGES REGULIERS

4.1 Proposition. *$\hat{V}$, $\{\Lambda\}$, $\{A\}$ pour tout $A \in V$, sont des bilangages réguliers sur V.*

Construisons dans chacun des cas un couple $\mathcal{B}$, $\mathcal{B}'$ satisfaisant à la définition 3.1 :

(1) Pour $\hat{V}: \mathcal{B} = \mathcal{B}' = \{e\}$; $\quad e \to e = e$; $\quad A \uparrow e = e$.

(2) Pour $\{\Lambda\}: \mathcal{B} = \{0, e\}, \mathcal{B}' = \{e\}$; pour b, b' dans $\mathcal{B}$ et A dans V :

$b \to b' =$ si $b = b' = e$ alors e sinon 0 ; $\quad A \uparrow b = 0$.

(3) Pour $\{A\}: \mathcal{B} = \{0, 1, e\}, \mathcal{B}' = \{1\}$, le produit $\to$ est commutatif et, pour $b \in \mathcal{B}$:

$0 \to b = 0, \quad e \to b = b, \quad 1 \to b =$ si $b = e$ alors 1 sinon 0 ;

$B \uparrow b = 0$ pour $B \neq A$ ou $b \neq e, \quad A \uparrow e = 1.$

4.2 THÉORÈME. *Tout bilangage grammatical est régulier.*

Soit $G = (V, ::=, X)$ une grammaire (2.4.1). Remarquons d'abord qu'il existe un monoïde fini $\mathfrak{M}$ tel que les langages X et K_A (ensemble des productions de A), pour tout $A \in V$, sont images réciproques par un homomorphisme de monoïdes, de parties de $\mathfrak{M}$; ceci résulte du lemme suivant :

4.3 LEMME. *Si K_1 et K_2 sont deux langages réguliers sur V, il existe un monoïde fini $\mathfrak{M}$, deux parties $\mathfrak{M}_1$ et $\mathfrak{M}_2$ de $\mathfrak{M}$ et un homomorphisme μ de V^* dans $\mathfrak{M}$ tels que : $K_1 = \mu^{-1}(\mathfrak{M}_1)$, $K_2 = \mu^{-1}(\mathfrak{M}_2)$.*

Par définition, K_1 est associé à un triplet $(\mathcal{C}_1, \mathcal{C}_1', \mu_1)$ et K_2 à $(\mathcal{C}_2, \mathcal{C}_2', \mu_2)$.

Il suffit de prendre pour $\mathfrak{M}$ le monoïde produit $\mathcal{C}_1 \times \mathcal{C}_2$ et pour μ l'homomorphisme produit :

$$\mu(\alpha) = (\mu_1(\alpha), \mu_2(\alpha)).$$

Alors $K_1 = \mu^{-1}(\mathcal{C}_1' \times \mathcal{C}_2)$ et $K_2 = \mu^{-1}(\mathcal{C}_1 \times C_2')$.
Ce lemme s'étend naturellement à un nombre fini de langages réguliers.

Par conséquent, chaque langage K_A de la grammaire G peut être associé à un triplet $(\mathfrak{M}, \mathfrak{M}_A, \mu)$ et X à $(\mathfrak{M}, \mathfrak{M}_X, \mu)$. Ceci nous invite à envisager l'application qui à toute ramification r engendrée au sens large par G associe l'image par μ de son mot des racines ; plus précisément posons pour $r \in \hat{V}$:
$\psi(r) =$ si $r \in \mathcal{L}(G)$ alors $\mu(\rho(r))$ sinon 0 (où 0 est choisi de façon que $0 \notin \mathfrak{M}$).
Il reste à munir $\mathfrak{M} \cup \{0\}$ d'une structure de binoïde telle que ψ soit un homomorphisme. D'après 2.4.3 :

$$\psi(r \to s) = \text{si } r \in \mathcal{L}(G) \quad \text{et} \quad s \in \mathcal{L}(G) \text{ alors } \mu(\rho(r)\rho(s)) \text{ sinon } 0 ;$$

$$\psi(A \uparrow r) = \text{si } r \in \mathcal{L}(G) \quad \text{et} \quad A ::= \rho(r) \text{ alors } \mu(A) \text{ sinon } 0.$$

D'où la définition des lois de binoïde sur $\mathfrak{M} \cup \{0\}$ (la loi de monoïde dans $\mathfrak{M}$ est notée .) :

$$m \to m' = \text{si } m \neq 0 \quad \text{et} \quad m' \neq 0 \text{ alors } m.m' \text{ sinon } 0$$

$$A \uparrow m = \text{si } m \neq 0 \quad \text{et} \quad m \in \mathfrak{M}_A \text{ alors } \mu(A) \text{ sinon } 0.$$

Ainsi ψ est l'homomorphisme de $\hat{V}$ dans $\mathfrak{M} \cup \{0\}$ et :

$$r \in \mathcal{S}(G) \Leftrightarrow r \in \mathcal{L}(G) \quad \text{et} \quad \psi(r) \in \mathfrak{M}_X$$

$$\Leftrightarrow \psi(r) \neq 0 \quad \text{et} \quad \psi(r) \in \mathfrak{M}_X \Leftrightarrow \psi(r) \in \mathfrak{M}_X ,$$

autrement dit $\mathcal{S}(G)$ est le bilangage régulier associé à $(\mathfrak{M} \cup \{0\}, \mathfrak{M}_X, \psi)$.

En particulier tout langage régulier X est le bilangage grammatical engendré par la grammaire $(V, ::=, X)$ où tous les langages K_A sont réduits au mot vide, c'est-à-dire :

4.4 Corollaire. *Les langages réguliers sont des bilangages réguliers.*

On peut aussi retrouver, comme conséquence du théorème 4.2, les résultats de la proposition 4.1.[4]

4.5 *Remarque.* Le lemme 4.3 s'étend immédiatement aux bilangages réguliers : deux bilangages réguliers sur V sont images réciproques par homomorphisme de parties d'un même binoïde fini.

On verra plus loin (6.14) qu'il existe d'autres bilangages réguliers que les bilangages grammaticaux. Le théorème 4.2 admet cependant une réciproque faible (4.8) qui résultera de l'étude suivante.

Parmi les bilangages grammaticaux nous allons envisager les bilangages locaux qui généralisent les langages locaux en ce sens que leurs ramifications ne sont soumises qu'à des contraintes de "voisinage" ; les résultats obtenus sont tout à fait analogues à ceux relatifs aux langages locaux.

4.6 Définition. Etant donné trois sous ensembles de V^2 : τ, d, f, et trois sous ensembles de V : D, F et Φ, appelons bilangage local défini par τ, d, f, D, F, Φ le bilangage grammatical $\mathcal{S}(G)$ engendré par la grammaire $G = (V, ::=, X)$ où :

pour tout $A \in V$ l'ensemble des productions de A est le langage régulier formé du langage local $[\{B \mid (A,B) \in d\}, \tau, \{B \mid (A,B) \in f\}]$, réuni à $\{\Lambda\}$ si $A \in \Phi$;

X est le langage local $[D, \tau, F]$.

Les bilangages locaux sont réguliers. La réciproque est fausse (6.14) ; cependant le théorème suivant donne une nouvelle caractérisation des bilangages réguliers à partir des bilangages locaux :

4.7 Théorème. *Les bilangages réguliers ne contenant pas la ramification vide sont les transformés par transcription des bilangages locaux, et eux seuls.*

Les bilangages transcrits des bilangages locaux sont réguliers (3.4) et ne contiennent pas la ramification vide.

Réciproquement soit L un bilangage régulier, associé à $(\mathcal{B}, \mathcal{B}', \psi)$, ne

[4] D'autre part, si $\langle\Sigma, \sigma\rangle$ est une espèce (*species*) finie au sens de Thatcher et Wright (1966) l'ensemble des *termes* T_Σ peut être identifié à un bilangage régulier (engendré par la grammaire de vocabulaire Σ, d'ensemble d'axiomes Σ où, pour tout $f \in \Sigma$, $K_f = \Sigma^{\sigma(f)}$) et les parties de T_Σ qui sont *reconnaissables* sont les bilangages réguliers contenus dans T_Σ .

contenant pas la ramification vide. Envisageons l'ensemble fini $\mathcal{V} = \mathcal{B}^2 \times V$ et l'application δ de $\hat{V}$ dans $\hat{\mathcal{V}}$ définie par :

$$\delta(\Lambda) = \Lambda,$$
$$\delta(r \to A \uparrow r') = \delta(r) \to [(\psi(r \to A \uparrow r'), \psi(A \uparrow r'), A) \uparrow \delta(r')].$$

Par récurrence sur l'ordre d'une ramification $r \neq \Lambda$ de $\hat{V}$, on déduit immédiatement de cette définition que $\delta(r)$ appartient au bilangage local $\mathcal{L}$ sur $\mathcal{V}$ défini par :

$$D = \{(b, b, A)\}, \qquad \tau = \{(b_1, b_2, A), (b_1 \to b_2', b_2', A')\},$$
$$F = \mathcal{V}, \qquad d = \mathcal{V} \times D, \qquad f = \{(b_1, A \uparrow b_1', A), (b_1', b_2', A')\},$$
$$\Phi = \{(b_1, \psi(A), A)\},$$

où b, b_1, b_2, b_1', b_2' décrivent $\mathcal{B}$ et A, A' décrivent V.
De plus r est déduit de $\delta(r)$ par la transcription θ de $\hat{\mathcal{V}}$ dans $\hat{V}$ définie par $\theta(b_1, b_2, A) = A$:

(1) pour tout $r \in \hat{V}$, $\theta \circ \delta(r) = r$.

Inversement, soit s une ramification de $\mathcal{L}$. Montrons, par récurrence sur l'ordre de s, que :

(2) $\delta \circ \theta(s) = s$,

d'où il résulte que la première composante de la dernière racine de s est $\psi(\theta(s))$.

L' assertion (2) est évidente si s est vide. Sinon,

$$s = s' \to ((b_1, b_2, A) \uparrow s'').$$

D'après l'hypothèse de récurrence :

$$\begin{aligned}\delta \circ \theta(s) &= \delta(\theta(s') \to A \uparrow \theta(s''))\\ &= s' \to [(\psi \circ \theta(s') \to A \uparrow \psi \circ \theta(s''), A \uparrow \psi \circ \theta(s''), A) \uparrow s''].\end{aligned}$$

Toujours d'après l'hypothèse de récurrence, comme s appartient à $\mathcal{L}$. nécessairement :

$$b_2 = A \uparrow \psi \circ \theta(s''),$$
$$b_1 = \psi \circ \theta(s') \to b_2 = \psi \circ \theta(s') \to A \uparrow \psi \circ \theta(s''),$$

et finalement $\delta \circ \theta(s) = s$.
Les assertions (1) et (2) prouvent que δ est une bijection de $\hat{V}$ sur $\mathcal{L}$. dont la bijection réciproque est la restriction de θ à $\mathcal{L}$.

Comme la première composante de la dernière racine de $\delta(r)$ est $\psi(r)$, L est transformé par δ en un bilangage local $\mathcal{L}'$ qui diffère de $\mathcal{L}$ uniquement par son ensemble final

$$F' = \{(b_1, b_2, A) \mid b_1 \in \mathcal{B}'\}.$$

Autrement dit $\delta(L) = \mathcal{L}'$ et $\theta(\mathcal{L}') = L$.

Le théorème est ainsi démontré. La démonstration prouve de plus qu'un bilangage régulier est déduit *bijectivement* d'un bilangage local par une transcription.

D'autre part, $L \cup \{\Lambda\} = \theta(\mathcal{L}' \cup \{\Lambda\})$ et $\mathcal{L}' \cup \{\Lambda\}$ est un bilangage grammatical ; il en résulte la première des deux conséquences suivantes :

4.8 *Conséquences.* 1. Tout bilangage régulier est déduit par transcription d'un bilangage grammatical.

2. Tout bilangage grammatical ne contenant pas la ramification vide est déduit par transcription d'un bilangage local.

4.9 Théorème. *Les langages de Chomsky sont les ensembles des mots des feuilles des bilangages réguliers et eux seuls.*

Les langages de Chomsky sont les ensembles des mots des feuilles des bilangages grammaticaux qui sont réguliers. Réciproquement, tout bilangage régulier L est transcrit d'un bilangage grammatical $\mathcal{S}$ dont les mots des feuilles forment un langage de Chomsky ; donc les mots des feuilles de L, qui sont les transcrits des mots des feuilles de $\mathcal{S}$, forment encore un langage de Chomsky.

Examinons aussi l'ensemble des mots des racines d'un bilangage régulier : c'est le transcrit de l'ensemble des mots des racines d'un bilangage grammatical, engendré par une grammaire $G = (V, ::=, X)$; c'est donc le transcrit du langage $X \cap \{A \in V \mid \mathcal{L}(G, A) \neq \varnothing\}^*$ qui est régulier ; ainsi l'ensemble des mots des racines d'un bilangage régulier est un langage régulier. La réciproque découle de 4.4. D'où l'énoncé :

4.10 Théorème. *Les langages réguliers sont les ensembles des mots des racines des bilangages réguliers et eux seuls.*

En particulier, tout bilangage régulier sur V qui est un langage sur V est un langage régulier.

4.11 *Contrexemple de bilangage régulier.* Soient deux ramifications distinctes r et s, et L l'ensemble des $r^n \to s^n$ (n représentant une puissance entière pour la loi $\to$). Tout homomorphisme de binoïdes est un homomorphisme de monoïdes pour les lois internes $\to$. Si donc L était un bilangage régulier, ce serait aussi un langage régulier du monoïde libre engendré par r et s : L n'est pas un bilangage régulier.

5. PROPRIETES ELEMENTAIRES DES BILANGAGES REGULIERS

5.1 Proposition : Propriétés booléennes. *Le complémentaire dans $\hat{V}$ d'un bilangage régulier sur V est un bilangage régulier. L'intersection et la réunion de deux bilangages réguliers sont des bilangages réguliers.*

Si L est le bilangage régulier associé à $(\mathfrak{B}, \mathfrak{B}', \psi)$ $\complement_{\hat{V}}^{L}$ est le bilangage régulier associé à $(\mathfrak{B}, \complement_{\mathfrak{B}}^{\mathfrak{B}'}, \psi)$.

Si L_1 et L_2 sont deux bilangages réguliers il existe un binoïde fini $\mathfrak{B}$ et deux parties $\mathfrak{B}_1$ et $\mathfrak{B}_2$ de $\mathfrak{B}$ tels que, si ψ est l'homomorphisme de $\hat{V}$ dans $\mathfrak{B}$, $L_1 = \psi^{-1}(\mathfrak{B}_1)$ et $L_2 = \psi^{-1}(\mathfrak{B}_2)$ (cf. 4.5). Dans ces conditions :

$$L_1 \cap L_2 = \psi^{-1}(\mathfrak{B}_1 \cap \mathfrak{B}_2), \qquad L_1 \cup L_2 = \psi^{-1}(\mathfrak{B}_1 \cup \mathfrak{B}_2).$$

5.2 Proposition. *Le bilangage réfléchi d'un bilangage régulier est régulier.*

Soit un bilangage régulier L associé à $(\mathfrak{B}, \mathfrak{B}', \psi)$; définissons l'application $\tilde{\psi}$ de $\hat{V}$ dans $\mathfrak{B}$ par :

$$\tilde{\psi}(r) = \psi(\tilde{r});$$

en particulier

$$\tilde{\psi}(r \to s) = \tilde{\psi}(s) \to \tilde{\psi}(r)$$

$$\tilde{\psi}(A \uparrow r) = A \uparrow \tilde{\psi}(r).$$

Par conséquent $\tilde{\psi}$ est l'homomorphisme de $\hat{V}$ dans $\mathfrak{B}$ muni de la structure de binoïde sur V définie par la loi de composition interne $\leftarrow$:

$$b \leftarrow b' = b' \to b$$

et la loi de composition externe $\uparrow$.

Enfin $\tilde{L} = \tilde{\psi}^{-1}(\mathfrak{B}')$; donc $\tilde{L}$ est régulier.

6. STABILITE DES BILANGAGES REGULIERS RELATIVEMENT A PLUSIEURS OPERATIONS

6.1 Proposition. *Si L et L' sont deux bilangages réguliers sur un vocabulaire V, les bilangages $L \to L'$ et L^* sont réguliers.*

Nous savons que :

tout bilangage régulier est transcrit d'un bilangage grammatical ;

le transcrit d'un bilangage grammatical est un bilangage régulier.

En outre si θ est une transcription, il est clair que

$$\theta(L \to L') = \theta(L) \to \theta(L') \text{ et } \theta(L^*) = \theta(L)^*.$$

Par conséquent il suffit de montrer que si L et L' sont deux bilangages engendrés par les grammaires $G = (V, \ :: =, X)$ et $G' = (V', \ :: \stackrel{\prime}{=}, X')$ sur deux vocabulaires V et V' qu'on peut toujours supposer disjoints, $L \to L'$ (resp. L^*) sont des bilangages grammaticaux sur $V \cup V'$ (resp. V).

Il est évident que $L \to L'$ est inclus dans le bilangage $\mathcal{S}(G_0)$ engendré par $G_0 = (V \cup V', \ :: = \ ou \ :: \stackrel{\prime}{=}, \ X \to X')$. Réciproquement soit $r \in \mathcal{S}(G_0)$; le mot des racines de r appartient à $X \to X'$, donc il existe deux ramifications s et s' telles que $r = s \to s'$, $\rho(s) \in X$, $\rho(s') \in X'$. On voit par récurrence, en utilisant (2.4.3), que, si une ramification t appartient à $\mathcal{L}(G_0)$ et $\rho(t)$ à V^*, alors t appartient à $\mathcal{L}(G)$, car V et V' sont disjoints ; ainsi s est dans $\mathcal{L}(G)$ et, comme $\rho(s)$ est dans X, s appartient à L ; de même façon, s' appartient à L', ce qui achève la démonstration.

Montrons que l'itéré L^* n'est autre que le bilangage engendré par $G_1 = (V, \ :: =, \ X^*)$. Il est évident que L^* est inclus dans $\mathcal{S}(G_1)$. Réciproquement soit $r \in \mathcal{S}(G_1)$; $\rho(r)$ appartient à X^* et r s'écrit $r_1 \to \cdots \to r_p$ où les ramifications $r_1, \cdots, r_p$ ont leur mot des racines dans X. Comme $\mathcal{L}(G_1) = \mathcal{L}(G)$, $r_1, \cdots, r_p$ appartiennent à $\mathcal{L}(G)$ et aussi à $\mathcal{S}(G) = L$; autrement dit, r appartient à L^*.

Introduisons maintenant un nouveau produit entre deux bilangages L et L', qui consiste, intuitivement, à associer à L et L' l'ensemble des ramifications obtenues en "greffant" à toute feuille A d'une ramification de L, une ramification de L'.

6.2 *Greffe, greffe itérée.* La *greffe* en A d'un bilangage L sur une ramification r est le bilangage, noté $\mathcal{G}_A(r, L)$ ou $r_A L$, défini de la façon suivante, par récurrence sur la ramification r : pour $A, B \in V$, $r, s \in \hat{V}$,

$$\mathcal{G}_A(\Lambda, L) = \{\Lambda\}$$

$$\mathcal{G}_A(r \to s, L) = \mathcal{G}_A(r, L) \to \mathcal{G}_A(s, L)$$

$$\mathcal{G}_A(B \uparrow r, L) = B \uparrow \mathcal{G}_A(r, L) \text{ si } B \neq A \text{ ou } r \neq \Lambda$$

$$\mathcal{G}_A(A, L) = A \uparrow L$$

La *greffe* en A d'un bilangage L sur un bilangage L', notée $L'{}_{\dot A}L$, est le bilangage

$$\bigcup_{r \in L'} r_{\dot A} L.$$

La *greffe itérée* en A d'un bilangage L est le bilangage L^{*A} défini par :

$$L^{0,A} = \{\Lambda\}, \qquad L^{p,A} = L_{\dot A} L^{p-1,A} \quad \text{pour} \quad p \geqslant 1 \quad \text{et} \quad L^{*A} = \bigcup_{p=0}^{\infty} L^{p,A}.$$

Le bilangage $\bigcup_{p=1}^{\infty} L^{p,A}$ est noté $\bar{L}^A$.

Dans la suite, lorsque l'élément A sera fixé, on notera parfois $\cdot$ au lieu de $_{\dot A}$.

Avant de montrer que la classe des bilangages réguliers est stable pour ces nouvelles opérations, indiquons quelques propriétés simples qui nous seront utiles.

6.3 Proposition. *Soient A un élément de V, L_1 et L_2 deux bilangages sur V, L_1 ne contenant pas la ramification vide* :

$$r_{\dot A}(L_{1\dot A} L_2) = (r_{\dot A} L_1)_{\dot A} L_2 .$$

La proposition se démontre par récurrence sur r, en quatre étapes :

(a) $r = \Lambda$
(b) $r = r' \to r''$
(c) $r = B \uparrow r'$ avec $B \neq A$ ou $r \neq \Lambda$
(d) $r = A$.

Nous ne traiterons que ce dernier cas, les autres étant immédiats :

$$\begin{aligned} A.(L_1.L_2) &= A \uparrow (L_1.L_2) \\ &= A \uparrow (\bigcup_{r \in L_1} r.L_2) \\ &= \bigcup_{r \in L_1} A \uparrow (r.L_2) \\ &= (\bigcup_{r \in L_1} A \uparrow r).L_2 \quad \text{car } L_1 \text{ ne contient pas } \Lambda \\ &= (A \uparrow L_1).L_2 \\ &= (A.L_1).L_2 \end{aligned}$$

6.4 *Conséquence*. Soient L, L_1, L_2 des bilangages sur V, L_1 ne contenant pas Λ :

$$L_{\dot A}(L_{1\dot A} L_2) = (L_{\dot A} L_1)_{\dot A} L_2 .$$

6.5 Proposition. *Si s appartient à $\mathcal{G}_A(r, L)$, alors $\rho(s) = \rho(r)$.* Ce résultat est immédiat par récurrence sur r.

6.6 *Conséquence.* Si s appartient à $\bar{L}^A$, $\rho(s)$ est le mot des racines d'une ramification de L.

La démonstration est immédiate pour $s \in L^{n,A}$ par récurrence sur $n \geqslant 1$.

6.7 Proposition. *Toute famille de prédécesseur B, différent de A, dans une ramification de $\mathcal{G}_A(r, L)$ est une famille de prédécesseur B dans r ou dans une ramification de L. Toute famille de prédécesseur A dans une ramification de $\mathcal{G}_A(r, L)$ est :*

soit une famille non vide de prédécesseur A dans r,
soit une famille de prédécesseur A dans une ramification de L,
soit le mot des racines d'une ramification de L.

(Immédiat par récurrence sur r).

6.8 *Conséquence.* Toute famille de prédécesseur B dans une ramification de $\bar{L}^A$ est une famille de prédécesseur B dans une ramification de L ou, lorsque $B = A$, le mot des racines d'une ramification de L.

6.9 Lemme. *Si L_1 et L_2 sont deux bilangages réguliers, il existe un binoïde fini $\mathcal{B}$, deux parties $\mathcal{B}_1$ et $\mathcal{B}_2$ de $\mathcal{B}$ tels que, si ψ est l'homomorphisme de $\hat{V}$ dans $\mathcal{B}$ et e l'élément neutre de $\mathcal{B}$:*

$$L_1 = \psi^{-1}(\mathcal{B}_1), \qquad L_2 = \psi^{-1}(\mathcal{B}_2) \quad \text{et} \quad \{\Lambda\} = \psi^{-1}(e).$$

Soit $\mathcal{B}_0$ un binoïde fini tel que L_1 et L_2 soient images réciproques de parties de $\mathcal{B}_0$ par l'homomorphisme ψ_0 de $\hat{V}$ dans $\mathcal{B}_0$ (4.5) ; notons e_0 l'élément neutre de $\mathcal{B}_0$; e étant un élément qui n'appartient pas à $\mathcal{B}_0$, prolongeons à $\mathcal{B} = \mathcal{B}_0 \cup \{e\}$ les lois de binoïde de $\mathcal{B}_0$ en posant :

$$\text{pour} \quad b \in \mathcal{B}, \qquad e \rightarrow b = b \rightarrow e = b$$
$$\text{pour} \quad A \in V, \qquad A \uparrow e = A \uparrow e_0$$

L'application ψ de $\hat{V}$ dans $\mathcal{B}$ définie par :

$$\psi(r) = \psi_0(r) \quad \text{si} \quad r \neq \Lambda$$
$$\psi(\Lambda) = e$$

est l'homomorphisme de $\hat{V}$ dans $\mathcal{B}$. Il est clair que $\psi^{-1}(e) = \{\Lambda\}$ et que L_1 et L_2 sont images réciproques de parties de $\mathcal{B}$.

6.10 Proposition. *Si L_1 et L_2 sont deux bilangages réguliers, $L_1{}_{\dot{A}}L_2$ est régulier.*

L_1 et L_2 peuvent être associés respectivement à $(\mathcal{B}, \mathcal{B}_1, \psi)$ et

$(\mathfrak{B}, \mathfrak{B}_2, \psi)$ où ψ satisfait au lemme précédent. $\mathfrak{C} = \mathfrak{P}(\mathfrak{B}) \times \{0, 1\}$ est un ensemble fini ; soit ξ l'application de $\hat{V}$ dans $\mathfrak{C}$ définie par :

$$\xi(r) = \{(\psi(r), 0)\} \cup \{(\psi(r'), 1) \mid r \in \mathcal{G}_A(r', L_2)\}.$$

Alors $L_1 \acute{A} L_2 = \{r \in \hat{V} \mid \xi(r) \cap (\mathfrak{B}_1 \times \{1\}) \neq \varnothing\}$.

Afin de munir $\mathfrak{C}$ de lois telles que ξ soit un homomorphisme examinons $\xi(\Lambda)$, $\xi(r \to s)$ et $\xi(B \uparrow r)$:

$$\xi(\Lambda) = \{(e, 0), (e, 1)\} \text{ où } e \text{ est l'élément neutre de } \mathfrak{B}.$$

$$\begin{aligned}\xi(r \to s) &= \{(\psi(r \to s), 0)\} \cup \{(\psi(u'), 1) \mid r \to s \in \mathcal{G}_A(u', L_2)\}\\ &= \{(\psi(r) \to \psi(s), 0)\} \cup \{(\psi(r') \to \psi(s'), 1) \mid r \in \mathcal{G}_A(r', L_2)\\ &\qquad \text{et } s \in \mathcal{G}_A(s', L_2)\} \;;\end{aligned}$$

en effet, comme $r \to s$ et u' ont même mot des racines (6.5) il résulte de la proposition 2.3.6 que :

$$r \to s \in \mathcal{G}_A(u', L_2) \Rightarrow (\exists r' \text{ et } s')(r' \to s' = u' \text{ et } r \in \mathcal{G}_A(r', L_2) \text{ et } s \in \mathcal{G}_A(s', L_2)) \;;$$

la réciproque est évidente.

$$\xi(B \uparrow r) = \{(\psi(B \uparrow r), 0)\} \cup \{(\psi(r'), 1) \mid B \uparrow r \in \mathcal{G}_A(r', L_2)\}.$$

Si $B \uparrow r \in \mathcal{G}_A(r', L_2)$, le mot des racines de r' est B. Donc, pour $B \neq A$:

$$B \uparrow r \in \mathcal{G}_A(r', L_2) \Leftrightarrow (\exists r'')(r' = B \uparrow r'' \text{ et } r \in \mathcal{G}_A(r'', L_2)).$$

D'autre part :

$$A \uparrow r \in \mathcal{G}_A(r', L_2) \Leftrightarrow (\exists r'')(r'' \neq \Lambda \text{ et } r' = A \uparrow r'' \text{ et } r \in \mathcal{G}_A(r'', L_2)) \quad \text{ou} \quad (r' = A \text{ et } r \in L_2).$$

Par suite :

$$\xi(B \uparrow r) = \{(B \uparrow \psi(r), 0)\} \cup \{(B \uparrow \psi(r''), 1) \mid r \in \mathcal{G}_A(r'', L_2)\} \quad \text{si} \quad B \neq A,$$

$$\xi(A \uparrow r) = \{(A \uparrow \psi(r), 0)\} \cup \{(A \uparrow \psi(r''), 1) \mid r \in \mathcal{G}_A(r'', L_2) \text{ et } r'' \neq \Lambda\} \quad \text{si} \quad r \notin L_2,$$

$$\xi(A \uparrow r) = \{(A \uparrow \psi(r), 0)\} \cup \{(A \uparrow \psi(r''), 1) \mid r \in \mathcal{G}_A(r'', L_2) \text{ et } r'' \neq \Lambda\} \cup \{(\psi(A), 1)\} \quad \text{si} \quad r \in L_2 .$$

Pour que ξ soit l'homomorphisme de $\hat{V}$ dans $\mathfrak{C}$, il suffit donc de munir $\mathfrak{C}$ de la structure suivante de binoïde sur V :
pour $\mathcal{B}', \mathcal{B}'' \in \mathfrak{C}, A, B \in V$:

$$\mathcal{B}' \to \mathcal{B}'' = \{(b' \to b'', 0) \mid (b', 0) \in \mathcal{B}' \quad \text{et} \quad (b'', 0) \in \mathcal{B}''\} \\ \cup \{(b' \to b'', 1) \mid (b', 1) \in \mathcal{B}' \quad \text{et} \quad (b'', 1) \in \mathcal{B}''\},$$

$$B \uparrow \mathcal{B}' = \{(B \uparrow b', 0) \mid (b', 0) \in \mathcal{B}'\} \cup \{(B \uparrow b', 1) \mid (b', 1) \in \mathcal{B}'\} \\ \text{si} \quad B \neq A,$$

$$A \uparrow \mathcal{B}' = \{(A \uparrow b', 0) \mid (b', 0) \in \mathcal{B}'\} \cup \{(A \uparrow b', 1) \mid (b', 1) \in \mathcal{B}' \\ \text{et} \quad b' \neq e\} \quad \text{si} \quad \mathcal{B}' \cap (\mathcal{B}_2 \times \{0\}) = \varnothing,$$

$$A \uparrow \mathcal{B}' = \{(A \uparrow b', 0) \mid (b', 0) \in \mathcal{B}'\} \cup \{(A \uparrow b', 1) \mid (b', 1) \in \mathcal{B}' \\ \text{et} \quad b' \neq e\} \cup \{(A \uparrow e, 1)\} \quad \text{si} \quad \mathcal{B}' \cap (\mathcal{B}_2 \times \{0\}) \neq \varnothing.$$

$L_1{}_A L_2$ est donc régulier.

En particulier si $L_1 = \{A\}$ le bilangage $A_A L_2 = A \uparrow L_2$ est régulier.

6.11 *Conséquence*. Si L est un bilangage régulier, pour tout $A \in V$, $A \uparrow L$ est un bilangage régulier.

6.12 Proposition. *Si L est un bilangage régulier, L^{*A} est un bilangage régulier.*

L est associé à $(\mathcal{B}, \mathcal{B}_1, \psi)$ où ψ satisfait au lemme 6.9.

Envisageons l'ensemble fini $\mathfrak{C} = \mathfrak{P}(\mathcal{B})$, et soit ξ l'application de $\hat{V}$ dans $\mathfrak{C}$ définie par :

$$\xi(r) = \{\psi(r') \mid r \in \mathcal{G}_A(r', L^{*A})\}.$$

Alors :

$$r \in L^{*A} \Leftrightarrow r = \Lambda \quad \text{ou} \quad (\exists r' \in L)(r \in \mathcal{G}_A(r', L^{*A})) \Leftrightarrow \xi(r) = \{e\} \\ \text{ou} \quad \xi(r) \cap \mathcal{B}_1 \neq \varnothing.$$

De même que précédemment :

$$\xi(\Lambda) = \{e\},$$

$$\xi(r \to s) = \{\psi(r') \to \psi(s') \mid r \in \mathcal{G}_A(r', L^{*A}) \quad \text{et} \quad s \in \mathcal{G}_A(s', L^{*A})\},$$

$$\xi(B \uparrow r) = \{B \uparrow \psi(r'') \mid r \in \mathcal{G}_A(r'', L^{*A})\} \quad \text{si} \quad B \neq A,$$

$$\xi(A \uparrow r) = \{A \uparrow \psi(r'') \mid r \in \mathcal{G}_A(r'', L^{*A}) \quad \text{et} \quad r'' \neq \Lambda\} \quad \text{si} \quad r \notin L^{*A},$$

$$\xi(A \uparrow r) = \{A \uparrow \psi(r'') \mid r \in \mathcal{G}_A(r'', L^{*A}) \quad \text{et} \quad r'' \neq \Lambda\} \\ \cup \{\psi(A)\} \quad \text{si} \quad r \in L^{*A}.$$

$\mathfrak{C}$ est muni d'une structure de binoïde telle que ξ soit l'homomorphisme de $\hat{V}$ dans $\mathfrak{C}$ si on pose :

$$\mathcal{B}' \to \mathcal{B}'' = \{b' \to b'' \mid b' \in \mathcal{B}', b'' \in \mathcal{B}''\},$$

$$B \uparrow \mathcal{B}' = \{B \uparrow b' \mid b' \in \mathcal{B}'\} \quad \text{si} \quad B \neq A,$$

$$A \uparrow \mathcal{B}' = \{A \uparrow b' \mid b' \in \mathcal{B}' \quad \text{et} \quad b' \neq e\} \quad \text{si} \quad \mathcal{B}' \neq \{e\}$$
$$\text{et} \quad \mathcal{B}' \cap \mathcal{B}_1 = \varnothing,$$

$$A \uparrow \mathcal{B}' = \{A \uparrow b' \mid b' \in \mathcal{B}' \quad \text{et} \quad b' \neq e\} \cup \{A \uparrow e\} \quad \text{si} \quad \mathcal{B}' = \{e\}$$
$$\text{ou} \quad \mathcal{B}' \cap \mathcal{B}_1 \neq \varnothing.$$

Les parties finies de $\hat{V}$ sont des réunions finies de ramifications, qui sont des combinaisons finies par $\to$ et $\uparrow$ d'éléments de V ; d'où le résultat suivant (en appliquant 4.1, 5.1, 6.1 et 6.11) :

6.13 Proposition. *Les parties finies de $\hat{V}$ sont des bilangages réguliers.*

6.14 *Remarque.* La proposition 6.13 permet de construire facilement des bilangages réguliers qui ne sont pas grammaticaux. Par exemple, le bilangage $\{A \uparrow A\}$ sur $V = \{A\}$ est régulier ; s'il était engendré par une grammaire $(V, ::=, X)$ nous aurions :

$$A \in X \quad \text{et} \quad A ::= A ;$$

donc $A \uparrow (A \uparrow A)$ appartiendrait au bilangage, ce qui est faux.

7. GENERALISATION DU THEOREME DE KLEENE POUR LES BILANGAGES REGULIERS

7.1 Définition. On appelle bilangages *complètement réguliers* sur un ensemble V les bilangages sur V qui appartiennent au plus petit ensemble de parties de $\hat{V}$ contenant les parties finies et stable par réunion, produit, produit itéré ainsi que par greffe et greffe itérée en tout élément de V.

Il résulte donc des propositions 6.13, 5.1, 6.1, 6.10 et 6.12 que tout bilangage complètement régulier est régulier. Inversement, tout langage régulier est un bilangage complètement régulier d'après le théorème de Kleene. Mais il est faux que tout bilangage régulier soit complètement régulier, comme en témoigne l'exemple suivant.

7.2 Exemple. Soit la grammaire G définie sur le vocabulaire $V_1 = \{A, B\}$ par son unique axiome A et les langages des productions de A et de B : $K_A = \{AB, \Lambda\}$, $K_B = \{\Lambda\}$. L_0 est le bilangage défini sur

$V = \{A\}$ comme transcrit du bilangage grammatical $\mathcal{S}(G)$ engendré par G, dans la transcription qui transforme A en A et B en A. Ainsi L_0 est un bilangage régulier infini. Plus précisément :

$$\mathcal{S}(G) = A \uparrow (\{\Lambda\} \cup (\mathcal{S}(G) \rightarrow B))$$
$$L_0 = A \uparrow (\{\Lambda\} \cup (L_0 \rightarrow A))$$

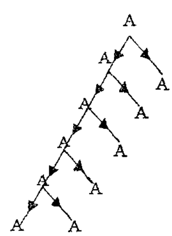
Montrons que L_0 n'est pas complètement régulier. Disons qu'un ensemble E de bilangages non vides se déduit d'un ensemble F par l'opération $\otimes$ ($\otimes$ désignant $\cup$, $\rightarrow$, ou ${}_A$) si on passe de E à F en remplaçant un des bilangages L de E par deux bilangages L' et L'' de F tels que $L = L' \otimes L''$, $L' \neq L$, $L'' \neq L$ (ce qui, si $\otimes$ est ${}_A$, impose $L'' \neq \{\Lambda\}$). Disons de même que E se déduit de F par l'opération $\otimes$, égale à $*$ ou $*^A$, si on passe de E à F en remplaçant L de E par L' de F tel que $L = L'^{\otimes}$ et $L' \neq L$ (ce qui impose $L' \neq \{\Lambda\}$).

Pour qu'un bilangage F_0 non vide soit complétement régulier, il faut et il suffit qu'il existe un entier $n \geqslant 0$, une suite $F_1, \cdots, F_n$ d'ensembles de bilangages non vides et une suite d'opérations $\otimes_i$ telles que, pour $1 \leqslant i \leqslant n$, F_{i-1} se déduise de F_i par $\otimes_i$ et que F_n soit formé de bilangages finis.

Supposons $F_0 = L_0$. Montrons, par récurrence sur i, que si L appartient à F_i,

$$L \subset L_0 \cup (L_0 \rightarrow A) \cup \{\Lambda\}$$

et que, pour $i > 0$, $\otimes_i$ n'est ni $*$ ni $*^A$: comme F_n est formé de bilangages finis, il en résulte que tout F_i est formé de bilangages finis, ce qui contredit le fait que L_0 soit infini.



Ramification de L_0

FIG. 5

Supposons la propriété vraie à l'indice i (évident pour $i = 0$). Nous montrerons que $\otimes_{i+1}$ ne peut être $*$ ni $*^A$ et que les deux bilangages L' et L'' qui, dans F_{i+1}, remplacent $L \in F_i$, sont contenus dans $L_0 \cup (L_0 \rightarrow A) \cup \{\Lambda\}$.

(a) $\otimes_{i+1} = \cup$, $L = L' \cup L''$: l'assertion est évidente car L' et L'' sont contenus dans L.

(b) $\otimes_{i+1} = \rightarrow$, $L = L' \rightarrow L''$: tout $r \in L$ s'écrit $r = r' \rightarrow r''$ avec $r' \in L'$ et $r'' \in L''$; si r appartient à $L_0 \cup \{\Lambda\}$, alors $r = r'$ et $r'' = \Lambda$, ou $r = r''$ et $r' = \Lambda$; si r appartient à $L_0 \rightarrow A$, il existe une autre possibilité : $r' \in L_0$ et $r'' = A$. En tout cas, r' et r'' sont contenus dans $L_0 \cup (L_0 \rightarrow A) \cup \{\Lambda\}$.

(c) $\otimes_{i+1} = {}_A$, $L = L'{}_A L''$, $L'' \neq \{\Lambda\}$: par récurrence sur r, on voit que, pour $r \neq A$, $\mathcal{G}_A(r, L'')$ n'est pas contenu dans $L_0 \cup (L_0 \rightarrow A)$; donc nécessairement $L' = \{A\}$, $L = A \uparrow L''$: L est contenu dans L_0 qui est égal à $A \uparrow (\{\Lambda\} \cup (L_0 \rightarrow A))$, et par suite L'' est dans $\{\Lambda\} \cup (L_0 \rightarrow A)$.

(d) $\otimes_{i+1} = *$, $L = L'^*$, $L' \neq \{\Lambda\}$: L ne peut être contenu dans $L_0 \cup (L_0 \rightarrow A) \cup \{\Lambda\}$ car la longueur de ses ramifications n'est pas bornée.

(e) $\otimes_{i+1} = *^A$, $L = L'^{*A}$, $L' \neq \{\Lambda\}$: L contient $L'{}_A L'$; d'après l'étude faite en (c), nécessairement $L' = \{A\}$ et $L' \subset \{\Lambda\} \cup (L_0 \rightarrow A)$, ce qui est impossible.

7.3 Théorème. *Un bilangage est régulier si, et seulement si, il est déduit d'un bilangage complétement régulier par une transcription.*

Autrement dit, la classe des bilangages réguliers est la plus petite classe contenant les bilangages finis et stable par réunion, produit, produit itéré, greffe, greffe itérée et transcription.

Le transcrit d'un bilangage complètement régulier est un bilangage régulier d'après 3.4. La réciproque résulte immédiatement du théorème 4.8 et du résultat suivant :

7.4 Proposition. *Tout bilangage grammatical est complètement régulier.*

Les notations sont celles de 2.4.2. Soit une grammaire G. Montrons que $\mathcal{S}(G)$ est complètement régulier, par récurrence sur le nombre de $A \in V$ tels que K_A contienne un mot non vide.

(a) Si pour tout $A \in V$, $K_A = \{\Lambda\}$ ou $K_A = \emptyset$, $\mathcal{S}(G)$ est l'ensemble des mots de X formés des lettres A telles que $K_A = \{\Lambda\}$: c'est un langage régulier, donc un bilangage complètement régulier.

(b) Soit A un élément de V tel que K_A contienne un mot non vide. Désignons par G' la grammaire obtenue en remplaçant dans G le langage K_A par $\{\Lambda\}$. Nous montrerons que :

(1) si $\Lambda \notin K_A$, $\mathcal{S}(G) = \mathcal{S}(G')_{\dot{A}}\overline{\mathcal{L}(G', K_A)}^{A}{}_{\dot{A}}\varnothing$ (le produit est alors associatif d'après 6.4) ;

(2) si $\Lambda \in K_A$, $\mathcal{S}(G) = \mathcal{S}(G')_{\dot{A}}\overline{\mathcal{L}(G', K_A)}^{A}$.

L'hypothése de récurrence assure que $\mathcal{S}(G')$ et $\mathcal{L}(G', K_A)$ sont des bilangages complètement réguliers. Il en est donc de même pour $\overline{\mathcal{L}(G', K_A)}^{A} = \mathcal{L}(G', K_A)_{\dot{A}}\mathcal{L}(G', K_A)^{*A}$ et pour $\mathcal{S}(G)$. (1) et (2) résultent des lemmes qui suivent.

Notons $\mathcal{L}_n(G, K)$, resp. $\mathcal{L}_n(G)$, l'ensemble des ramifications de $\mathcal{L}(G, K)$, resp. $\mathcal{L}(G)$, où le nombre d'occurrences de A est strictement inférieur à n ; en particulier $\mathcal{L}_0(G, K) = \varnothing$.

Lemme 1. *Pour tout langage régulier K et tout entier naturel n,*

$$\mathcal{L}_{n+1}(G, K) \subset \mathcal{L}(G', K)_{\dot{A}}\mathcal{L}_n(G, K_A).$$

Soit $s \in \mathcal{L}_{n+1}(G)$. Nous montrerons, par récurrence sur s, qu'il existe $r \in \mathcal{L}(G')$ tel que $s \in \mathcal{G}_A(r, \mathcal{L}_n(G, K_A))$; comme alors r et s ont le même mot des racines (6.5), si $s \in \mathcal{L}_{n+1}(G, K)$, $r \in \mathcal{L}(G', K)$.

(a) Si $s = \Lambda$, $r = \Lambda$ convient.

(b) Soit $s = s' \to s''$: s' et s'' appartiennent à $\mathcal{L}_{n+1}(G)$:

$s' \in \mathcal{G}_A(r', \mathcal{L}_n(G, K_A))$ et $s'' \in \mathcal{G}_A(r'', \mathcal{L}_n(G, K_A))$ entraine

$s' \to s'' \in \mathcal{G}_A(r' \to r'', \mathcal{L}_n(G, K_A))$

$r' \in \mathcal{L}(G')$ et $r'' \in \mathcal{L}(G')$ entraine $r' \to r'' \in \mathcal{L}(G')$.

(c) Soit $s = B \uparrow s'$, avec $B \neq A$: $s' \in \mathcal{L}_{n+1}(G), \rho(s') \in K_B$;

$s' \in \mathcal{G}_A(r', \mathcal{L}_n(G, K_A))$ entraine $B \uparrow s' \in \mathcal{G}_A(B \uparrow r', \mathcal{L}_n(G, K_A))$

et $\rho(r') \in K_B$; $r' \in \mathcal{L}(G')$ entraine alors $B \uparrow r' \in \mathcal{L}(G')$.

(d) Soit $s = A \uparrow s'$: comme $s \in \mathcal{L}_{n+1}(G)$ et $\rho(s) = A$, $s' \in \mathcal{L}_n(G, K_A)$:

$$s \in \mathcal{G}_A(A, \mathcal{L}_n(G, K_A)) \quad \text{et} \quad A \in \mathcal{L}(G').$$

Lemme 2. $\mathcal{S}(G) = \mathcal{S}(G')_{\dot{A}}\mathcal{L}(G, K_A)$.

Du lemme 1, avec $K = X$, résulte que toute ramification de $\mathcal{S}(G)$ appartient à $\mathcal{S}(G')_{\dot{A}}\mathcal{L}(G, K_A)$.

Réciproquement, d'après la proposition 6.7, il est immédiat que si

$r \in \mathcal{L}(G')$ et $s \in \mathcal{G}_A(r, \mathcal{L}(G, K_A))$, alors $s \in \mathcal{L}(G)$. De plus, r et s ont le même mot des racines, ce qui achève de démontrer le lemme puisque G et G' admettent les mêmes axiomes.

Lemme 4. *Si K_A ne contient pas le mot vide*, $\mathcal{L}(G, K_A) = \overline{\mathcal{L}(G', K)}^A{}_{\dot{A}}\, \varnothing$.

(a) Soit $r \in \overline{\mathcal{L}(G', K_A)}^A$ et $s \in \mathcal{G}_A(r, \varnothing)$. Le mot des racines de s appartient à K_A. D'autre part, d'après les propositions 6.7 et 6.8, toute famille de prédécesseur B dans s est une famille de prédécesseur B dans r ; pour $B \neq A$, c'est une famille de prédécesseur B dans une ramification de $\mathcal{L}(G', K_A)$, elle appartient à K_B ; si $B = A$, elle est non vide et c'est le mot des racines d'une ramification de $\mathcal{L}(G', K_A)$; elle appartient à K_A ; s appartient à $\mathcal{L}(G, K_A)$.

(b) Montrons, par récurrence sur n, que

$$\mathcal{L}_n(G, K_A) \subset [\mathcal{L}(G', K_A)]^{n,A}{}_{\dot{A}}\, \varnothing.$$

La propriété est évidente pour $n = 0$. Supposons la vraie pour l'entier n ; d'après le lemme 1 avec $K = K_A$, nous en déduisons

$$\mathcal{L}_{n+1}(G, K_A) \subset \mathcal{L}(G', K_A)_{\dot{A}}[\mathcal{L}(G', K_A)]^{n,A}{}_{\dot{A}}\, \varnothing$$

$$\mathcal{L}_{n+1}(G, K_A) \subset [\mathcal{L}(G', K_A)]^{n+1,A}{}_{\dot{A}}\, \varnothing,$$

car le produit est associatif, puisque la ramification vide n'appartient pas à $\mathcal{L}(G', K_A)$.

Lemme 4. Si *K_A contient le mot vide*, $\mathcal{L}(G, K_A) = \overline{\mathcal{L}(G', K_A)}^A$.

(a) Si $r \in \overline{\mathcal{L}(G', K_A)}^A$ le mot des racines de r appartient à K_A (6.6) et toute famille de prédécesseur B dans r est une famille de prédécesseur B dans une ramification de $\mathcal{L}(G', K_A)$ ou, dans le cas $B = A$, le mot des racines d'une telle ramification : $r \in \mathcal{L}(G, K_A)$.

(b) De la même façon qu'au lemme 3, on montre que

$$\mathcal{L}_n(G, K_A) \subset [\mathcal{L}(G', K_A)]^{n,A}.$$

Reçu le 14 mai 1968; revision reçue le 8 aout 1968.

REFERENCES

Bar-Hillel, Y., Perles, M., et Shamir, E. (1961), On formal properties of simple phrase structure grammars, *Z. Phonetik, Sprachwiss., Kommunikationsforsch.* **14,** 143–172.

Berge, C. (1958), "Théorie des graphes et ses applications." Dunod, Paris.

Chomsky, N. (1963), Formal Properties of Grammars, *in* "Handbook of mathematical psychology" (D. Luce, E. Bush, et E. Galanter, Eds.) Wiley, New York, pp. 323–418.

GINSBURG, S. (1966), "The Mathematical Theory of Context-Free Languages." McGraw-Hill, New York.

GROSS, M. ET LENTIN, A. (1967), "Notions sur les grammaires formelles." Gauthier-Villars, Paris.

LUCAS, P. (1967), Introduction to the Method Used for the Formal Definition of PL/1 [Technical Report TR 25 081]. IBM Laboratory Vienna.

PAIR, C. (1965), Etude de la notion de pile, application à l'analyse syntaxique. Collection des thèses de la faculté des Sciences, Université de Nancy.

PAIR, C. (1967), Préliminaires à l'étude algébrique des ramifications. Communication soumise au congrès 1968 de l'I.F.I.P.

RABIN, M. O. ET SCOTT, D. (1959), Finite automata and their decision problems. *I.B.M. J. Res. Develop.* **3,** 114–125.

THATCHER, J. W. (1967), Characterizing derivation trees of context-free grammars through generalized finite automata theory. I.B.M. Research Note NC 719.

THATCHER, J. W. ET WRIGHT, J. B. (1966), Generalized finite automata theory with an application to a decision problem of second order logic. I.B.M. Research Report RC 1713.