

THESE

présentée à

L'UNIVERSITÉ DES SCIENCES ET TECHNIQUES DE LILLE

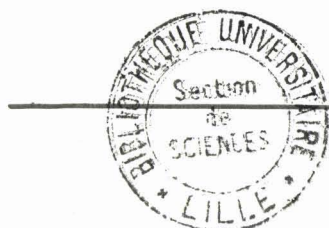
pour obtenir le titre de

DOCTEUR DE SPECIALITÉ

(Mathématiques Appliquées)

par

Michel LATTEUX



CLASSES DE GRAMMAIRES À VARIABLES DISTINGUÉES

Soutenue le 19 Décembre 1973, devant la COMMISSION D'EXAMEN

Membres du Jury : MM. P. BACCHUS

C. CARREZ

B. DRIEUX

Président

Examineur

Rapporteur

UNIVERSITE DES SCIENCES
ET TECHNIQUES DE LILLE.

DOYENS HONORAIRES

MM. H.LEFEBVRE, PARREAU.

PROFESSEURS HONORAIRES

M. ARNOULT, Mme BEAUJEU, MM. BEGHIN, BROCHARD, CAU, CHAPPELON, CHAUDRON, CORDONNIER, DEHEUVELS, DEHORNE, DEHORS, FAUVEL, FLEURY, P.GERMAIN, HEIM DE BALSAC, HOCQUETTE, KAMPE DE FERIET, KOURGANOFF, LAMOTTE, LELONG, Mme LELONG, MM. LIEBAERT, MARTINOT-LAGARDE, MAZET, MICHEL, NORMANT, PARISELLE, PASCAL, PAUTHENIER, PEREZ, ROIG, ROSEAU, ROUBINE, ROUELLE, WIEMAN, ZAMANSKY.

PROFESSEURS TITULAIRES

M. ANGRAND Jean Pierre	Géographie et Aménagement Spatial
M. BACCHUS Pierre	Astronomie et Calcul
M. BEAUFILS Jean Pierre	Chimie Générale
M. BECART Maurice	I.U.T. Lille
M. BLOCH Vincent	Psychophysiologie
M. BIAYS Pierre	Géographie et Aménagement Spatial
M. BONNEMAN Pierre	Chimie Industrielle
M. BONTE Antoine	Géologie Appliquée
M. BOUGHON Pierre	Mathématiques
M. BOURIQUET Robert	Biologie Végétale
M. CAPET Marcel-Francis	Institut de Préparation aux Affaires
M. CELET Paul	Géologie Générale
M. CONSTANT Eugène	Electronique
M. CORSIN Pierre	Paléobotanique
M. DECUYPER Marcel	Mathématiques
M. DEDECKER Paul	Mathématiques
M. DEFRETIN René	Biologie Animale - Directeur de l'Institut de Biologie Maritime de Wimereux
M. DELATTRE Charles	Géologie Générale
M. DURCHON Maurice	Biologie Animale
M. FLATRES Pierre	Géographie et Aménagement Spatial
M. FOURET René	Physique
M. GABILLARD Robert	Electronique
M. GEHU Jean Marie	Institut Agricole
M. GLACET Charles	Chimie Organique
M. GONTIER Gérard	Mécanique des Fluides
M. GUILLAUME Jean	Biologie Végétale
M. HEUBEL Joseph	Chimie Minérale
Mme LENOBLE Jacqueline	Physique
M. MONTREUIL Jean	Chimie Biologique
M. POUZET Pierre	I.U.T. Lille

Mme SCHWARTZ Marie Hélène	Mathématiques
M. TILLIEU Jacques	Physique
M. TRIDOT Gabriel	Chimie Minérale Appliquée
M. VIDAL Pierre	Automatique
M. VIVIER Emile	Biologie Animale
M. WATERLOT Gérard	Géologie et Minéralogie
M. WERTHEIMER Raymond	Physique

PROFESSEURS A TITRE PERSONNEL

M. BOUISSET Simon	Physiologie Animale
M. DELHAYE Michel	Chimie Physique et Minérale 1er Cycle
M. LEBRUN André	Electronique
M. LINDER Robert	Biologie Végétale
M. LUCQUIN Michel	Chimie Physique
M. PARREAU Michel	Mathématiques
M. PRUDHOMME Rémy	Sciences Economiques et Sociales
M. SAVARD Jean	Chimie Générale
M. SCHALLER François	Biologie Animale
M. SCHILTZ René	Physique

PROFESSEURS SANS CHAIRE

M. BELLET Jean	Physique
M. BODARD Marcel	Biologie Végétale
M. BOILLET Pierre	Physique
M. DERCOURT Jean Michel	Géologie et Minéralogie
M. DEVRAINNE Pierre	Chimie Minérale
M. LOMBARD Jacques	Sciences Economiques et Sociales
Mlle MARQUET Simone	Mathématiques
M. MONTARIOL Frédéric	Chimie Minérale Appliquée
M. PROUVOST Jean	Géologie et Minéralogie
M. VAILLANT Jean	Mathématiques

MAITRES DE CONFERENCES (et chargés de fonctions)

M. ADAM Michel	Sciences Economiques et Sociales
M. ANDRE Charles	Sciences Economiques et Sociales
M. AUBIN Thierry	Mathématiques Pures
M. BEGUIN Paul	Mécanique des Fluides
M. BILLARD Jean	Physique
M. BKOUCHE Rudolphe	Mathématiques
M. BOILLY Bénoni	Biologie Animale
M. BONNEMAIN Jean Louis	Biologie Végétale
M. BONNOT Ernest	Biologie Végétale
M. BRIDOUX Michel	I.U.T. Béthune
M. BRUYELLE Pierre	Géographie et Aménagement Spatial

M. CCAPURON Alfred	Biologie Animale
M. CARREZ Christian	Calcul
M. CHOQUET Marcel	I.U.T. Lille
M. CORDONNIER Vincent	Calcul
M. CORTOIS Jean	Physique
M. COULON Jean Paul	Electrotechnique
M. DEBRABANT Pierre	Sciences appliquées
M. ESCAIG Bertrand	Physique
Mme EVRARD Micheline	I.U.T. Lille
M. FAIDHERBE Jacques	Biologie Animale
M. FONTAINE Jacques	I.U.T. Lille
M. FROELICH Daniel	Sciences Appliquées
M. GAMBLIN André	Géographie et Aménagement Spatial
M. GOBLOT Rémi	Mathématiques
M. GOSSELIN Gabriel	Sciences Economiques et Sociales
M. GOUDMAND Pierre	Chimie Physique
M. GRANELLE	Sciences Economiques et Sociales
M. GRUSON Laurent	Mathématiques
M. GUIBAULT Pierre	Physiologie Animale
M. HERMAN Maurice	Physique
M. HUARD de la MARRE Pierre	Calcul
M. JOLY Robert	Biologie (Amiens)
M. JOURNEL Gérard	Sciences Appliquées
Mlle KOSMANN Yvette	Mathématiques
M. LABLACHE-COMBIER Alain	Chimie Générale
M. LACOSTE Louis	Biologie Végétale
M. LANDAIS Jean	Chimie Organique
M. LAURENT François	Automatique
M. LAVAGNE Pierre	Sciences Economiques et Sociales
Mlle LEGRAND Solange	Mathématiques
M. LEHMANN Daniel	Mathématiques
Mme LEHMANN Josiane	Mathématiques
M. LENTACKER Firmin	Géographie et Aménagement Spatial
M. LEROY Jean Marie	ENSCL
M. LEROY Yves	I.U.T. Lille
M. LHENAFF René	Géographie et Aménagement Spatial
M. LOCQUENEUX Robert	Physique
M. LOUAGE Francis	Sciences Appliquées
M. LOUCHEUX Claude	Chimie Physique
M. MAES Serge	Physique
Mme MAILLET Monique	Sciences Economiques et Sociales
M. MAIZIERES Christian	Automatique
M. MALAUSSENA Jean Louis	Sciences Economiques et Sociales
M. MESSELYN Jean	Physique
M. MIGEON Michel	Sciences Appliquées
M. MONTEL Marc	Physique
M. MONTUELLE Bernard	I.U.T. Lille
M. MUSSCHE Guy	Sciences Economiques et Sociales
M. NICOLE Jacques	E.N.S.C.L.
M. OUZIAUX Roger	Sciences Appliquées
M. PANET Marius	Electrotechnique
M. PAQUET Jacques	Sciences Appliquées
M. PARSY Fernand	Mécanique des Fluides
M. PONSOLLE Louis	Chimie (Valenciennes)
M. POVY Jean Claude	Sciences Appliquées
M. RACZY Ladislav	Radioélectricité
Mme RENVERSEZ Françoise	Sciences Economiques et Sociales
M. ROUSSEAU Jean Paul	Physiologie Animale

M. ROYNETTE Bernard
M. SALMER Georges
M. SEGUIER Guy
M. SIMON Michel
M. SMET Pierre
M. SOMME Jean
M. THOMAS Daniel
M. TOULOTTE Jean Marc
M. TREANTON Jean René
M. VANDORPE Bernard
M. VILETTE Michel
M. WATERLOT Michel
Mme ZINN JUSTIN Nicole

Mathématiques
Electronique
I.U.T. Béthune
Sciences Economiques et Sociales
Physique
Géographie et Aménagement Spatial
Chimie Minérale Appliquée
Sciences Appliquées
Sciences Economiques et Sociales
Sciences Appliquées
I.U.T. Béthune
Géologie Générale
Mathématiques.

Je tiens à témoigner toute ma gratitude à
Monsieur le Professeur BACCHUS, directeur de
l'UER Informatique - Electronique - Electrotechnique
et Automatique de l'Université de Lille I, pour
l'honneur qu'il me fait de présider le jury de cette
thèse.

Je remercie également Monsieur le Professeur CARREZ
qui a bien voulu s'intéresser à ce travail et le juger.

Je voudrais exprimer, tout particulièrement, ma
reconnaissance à Monsieur le Professeur DRIEUX qui m'a
fait découvrir la Théorie des Langages et qui n'a cessé
de m'encourager et de me guider de ses précieux conseils
tout au long de ce travail.

Enfin, je tiens à remercier Mademoiselle DRIESSENS,
Monsieur et Madame DEBOCK qui, avec une grande compétence
et beaucoup de gentillesse se sont chargés de la réalisation
matérielle de cette thèse.

A mes Parents

A ma femme

A mon fils

Table des Matières

Première Partie : GRAMMAIRES MONOTONES

- chapitre I Formes standards de grammaires monotones
- chapitre II Grammaires monotones et C-langages
- chapitre III Grammaires monotones et langages contextuels.

Deuxième Partie : CARACTERISATION DE LANGAGES DETERMINISTES

- chapitre IV A_{Pi}-grammaires et C-langages déterministes
- chapitre V LB-grammaires et Automates linéairement
bornés déterministes.

I N T R O D U C T I O N

Depuis l'introduction des grammaires contextuelles (type 1), à contexte-libre (type 2), linéaires à droite (type 3), définies par Chomsky [CH], à partir des grammaires à structure de phrase, en imposant des restrictions aux règles de production, la plupart des travaux relatifs aux grammaires se sont attachés aux C-grammaires et aux langages engendrés par différentes sous-classes de C-grammaires.

Par contre, les résultats concernant les grammaires à structure de phrase qui ne sont pas de l'un des trois types précédents sont beaucoup moins nombreux.

Certains ont été obtenus en considérant des dérivations spéciales [GS2, GS3, MA]. Ainsi, Matthews a montré que le langage engendré par dérivation la plus à gauche dans une grammaire à structure de phrase, est un C-langage.

Deux autres résultats intéressants, obtenus en particulier en analysant la forme des règles de production, mettent en relief le rôle spécial joué par les terminaux. Ginsburg et Greibach [GG1] ont montré, en effet, que les grammaires à structure de phrase dont toutes les règles (resp. toutes les règles de longueur décroissante) contiennent un terminal, engendrent des C-langages (resp. des langages contextuels).

Notre travail consiste à étudier certaines classes de grammaires à structure de phrase, incluses dans l'ensemble L_D des classes de grammaires à variables distinguées, lesquelles sont définies en imposant uniquement des relations entre une partition des variables et la forme des règles de production. Plus précisément, nous dirons que L est une classe de grammaires à variables distinguées, s'il existe un alphabet E et une partie R de $(E^*)^2$ qui permettent de définir L de la manière suivante :

Une grammaire à structure de phrase $G = (N, T, P, A)$ appartient à L si et seulement s'il existe un homomorphisme h de E^* dans le monoïde engendré par l'ensemble des parties de $N \cup T$, tel que $h(E)$ partitionne $N \cup T$ et P soit inclus dans $\{(h(u), h(v)) / (u, v) \in R\}$.

En particulier, toutes les classes de grammaires définies en considérant l'ensemble des variables comme l'union disjointe de T et de N , ensembles, respectivement, des terminaux et des non-terminaux, et en imposant aux règles de production d'appartenir à une partie de $N^+ \times (N \cup T)^*$ qui s'exprime uniquement en fonction de N et T , appartiennent à L_D .

Moins trivialement, Ginsburg et Spanier [GS1] ont introduits les grammaires Ultralinéaires d'ordre k (k entier ≥ 1) en restreignant la forme des règles de production des C-grammaires à l'aide d'une partition de l'ensemble des non-terminaux en k parties.

De même les "standard matching choice set" de rang k (quasi-rationnel de rang k) (k entier ≥ 0) définis par Yntema

[YN] peuvent être caractérisés grammaticalement par un élément de L_D qui est contenu dans la classe des C-grammaires.

Dans la première partie, nous partitionnons l'ensemble des non-terminaux en deux sous-ensembles et nous définissons différentes classes de grammaires appelées monotones. Pour $i \in [1,3]$, les grammaires monotones de type i généralisent strictement les grammaires de même type et engendrent les mêmes classes de langages. Nous obtenons ainsi des conditions suffisantes pour qu'une grammaire à structure de phrase, qui n'est pas de type 1, engendre respectivement un K-langage, un C-langage et un langage contextuel.

En outre, pour $i \in [1,2]$, nous montrons que ces conditions sont compatibles avec les résultats de Ginsburg et Greibach rappelés plus haut ; c'est-à-dire qu'il suffit de considérer l'ensemble des règles de production ne contenant pas de terminaux.

Dans la deuxième partie, les éléments de L_D que nous introduisons nous servent de support pour définir de manière simple d'autres classes de grammaires qui caractérisent divers langages déterministes.

Ainsi, au chapitre IV, nous considérons les API -grammaires ($i \in [1,3]$) qui engendrent les C-langages.

En imposant aux règles de productions des conditions supplémentaires, nous obtenons une nouvelle caractérisation grammaticale pour les langages reconnus par des automates à pile de mémoire déterministes [GG2] ou engendrés par les LR(k) grammaires de Knuth [KN].

Si nous imposons, en outre, aux parties droites des règles de commencer par un terminal, nous définissons les AP3-grammaires déterministes qui engendrent les langages reconnus par un automate à pile de mémoire déterministe, Λ -libre et complet.

Les langages LL(1) de Lewis et Stearns [LS] étant reconnus par une classe particulière de tels automates, il était normal de chercher une caractérisation de ces langages par des AP3-grammaires particulières. C'est ce que nous faisons en ajoutant une autre restriction à l'ensemble des règles de production pour définir les UAP3-grammaires.

Au chapitre V, nous considérons la classe L'_1 des langages reconnus par les automates linéairement bornés [MY]. Les LB-grammaires que nous obtenons à l'aide d'une partition de l'ensemble des non-terminaux en trois sous-ensembles, engendrent les langages contextuels. En ajoutant des conditions simples à vérifier, nous définissons les LB-grammaires déterministes et après avoir montré que la classe L'_1 est close pour les homomorphismes inverses et les homomorphismes Λ -libres, nous établissons l'équivalence entre L'_1 et la classe des langages engendrés par les LB-grammaires déterministes.

Chapitre 0

NOTATIONS ET TERMINOLOGIE

1 NOTATIONS

Pour deux ensembles A et B quelconques, nous noterons :

- $|A|$ = cardinal de A ,
- $P(A) = \{D / D \subset A\}$, l'ensemble des parties de A ,
- $A \setminus B = \{x \in A / x \notin B\}$,
- $AB = \{xy / x \in A, y \in B\}$, le **concatené** de A par B .

Pour $k \geq 1$, définissons A^k par récurrence en posant $A^1 = A$ et $A^{k+1} = A^k A$; si $A = \{a\}$, a^k désignera l'unique élément de A^k . Le **produit cartésien** d'une séquence finie d'ensembles $A_1, \dots, A_n$ est égal à $\{(a_1, \dots, a_n) / a_i \in A_i \text{ pour } i \in [1, n]\}$.

Nous dirons que $\{A_1, \dots, A_p\}$ forme une **partition** de A si $A = \bigcup_{i=1}^p A_i$ et si $A_i \cap A_j \neq \emptyset$ implique $i=j$, $\forall i, j \in [1, p]$.

Afin d'éviter un parenthésage trop important, nous conviendrons que la concaténation est prioritaire par rapport au produit cartésien, lui-même prioritaire vis à vis de l'union et de l'intersection.

Pour un ensemble fini T , appelé **alphabet**, posons $T^+ = \bigcup_{k=1}^{\infty} T^k$; $T^* = \{\Lambda\} \cup T^+$ désignera le monoïde libre engendré par T , le **mot vide**, Λ , étant l'élément neutre pour l'opération dans T^* , notée multiplicativement.

De façon générale, nous appellerons **mots** les éléments de T^* , **lettres** ceux de T et **langage** toute partie de T^* .

Pour tout mot w de T^* , $\ell(w)$, la **longueur** de w est égale au nombre d'occurrences de lettres de T dans w ; en particulier $\ell(\Lambda) = 0$; l'**image miroir**

de W , notée $\tilde{W}$, est définie par $\tilde{\lambda} = \lambda$ et $W'a = a\tilde{W}'$, $\forall a \in T, \forall W' \in T^*$.

Enfin, pour toute partie A de T^* , nous noterons $A^0 = A \cup \{\lambda\}$.

2 GRAMMAIRES ET LANGAGES

- Une grammaire à structure de phrase est un quadruplet $G = (N, T, P, A)$ où :
- N est un alphabet (non terminal),
 - T , disjoint de N , est un alphabet (terminal),
 - P est une partie finie de $N^+ \times (N \cup T)^*$ dont les éléments sont appelés les règles ou les productions,
 - A , l'axiome de la grammaire est un élément particulier de N .

Nous noterons $V = N \cup T$, l'ensemble des variables.

Pour $x, y \in V^*$, nous dirons que x engendre directement y et nous écrirons $x \xrightarrow{G} y$, s'il existe $(u, v) \in P$, $x', x'' \in V^*$, tels que $x = x'ux''$ et $y = x'vx''$. Nous dirons que x se dérive en y , s'il existe une séquence $z_0, z_1, \dots, z_p$ de mots de V^* , appelée dérivation (de longueur p), telle que $z_0 = x$, $z_p = y$ et $z_i \xrightarrow{G} z_{i+1}$ pour $i \in [0, p-1]$.

Si P_1 est inclus dans P , nous écrirons $x \xrightarrow{P_1^*} y$ pour préciser que $x \xrightarrow{G^*} y$ en utilisant uniquement des règles de P_1 .

$L(G) = \{W \in T^* / A \xrightarrow{G^*} W\}$ sera appelé le langage engendré par la grammaire G .

Deux grammaires seront dites équivalente si elles engendrent le même langage.

- Une grammaire "monotonic" est une grammaire à structure de phrase $G = (N, T, P, A)$ vérifiant la propriété :

$$\forall (u, v) \in P, \quad \ell(u) \leq \ell(v).$$

Un langage est contextuel s'il existe une grammaire monotonic qui l'engendre.

- Une grammaire à contexte libre (C-grammaire) [CH] est une grammaire à structure de phrase $G = (N, T, P, A)$ avec $P \subset N \times V^*$. Un langage est un C-langage s'il existe une C-grammaire qui l'engendre.

Une C-grammaire est réduite si pour tout $B \in N$, il existe $u, v \in V^*$, $W \in T^*$ tels que $A \xrightarrow[G]{*} uBv$ et $B \xrightarrow[G]{*} W$.

Une C-grammaire est Λ -libre si $P \subset N \times V^+$.

Pour $u, v \in V^*$, notons $u \xrightarrow[G]{g} v$ s'il existe $W \in T^*$, $u' \in V^*$ et $(B, y) \in P$ tels que $u = WBu'$ et $v = Wyu'$.

$\xrightarrow[G]{g^*}$ désignera la clôture transitive de la relation $\xrightarrow[G]{g}$ et sera appelé dérivation la plus à gauche.

- Une s-grammaire [KH] est une C-grammaire $G = (N, T, P, A)$ vérifiant :

$$(i) \quad P \subset N \times TV^*$$

$$(ii) \quad \forall B \in N, \forall a \in T, \forall y, z \in V^*, \text{ si } (B, ay) \in P \text{ et } (B, az) \in P \text{ alors } y = z.$$

Un langage est un s-langage s'il existe une s-grammaire qui l'engendre.

- Une grammaire linéaire à droite est une C-grammaire $G = (N, T, P, A)$ où P est inclus dans $N \times T^*N^0$.

Un langage est un K-langage (langage régulier) s'il existe une grammaire linéaire à droite qui l'engendre.

3 AUTOMATES ET LANGAGES

- Un automate d'états fini est un 5-uple $M = (Q, T, g, q_0, F)$ où :
 - (i) Q est un ensemble fini non vide (ensemble des états),
 - (ii) T est l'alphabet d'entrée,
 - (iii) g , appelée fonction de transition, est une application de $Q \times T$ dans Q ,
 - (iv) q_0 est un élément particulier de Q (état initial),
 - (v) $F \subset Q$ est l'ensemble des états terminaux.

L'application g est prolongée en une application de $Q \times T^*$ dans Q en posant :

- $\forall q \in Q, g(q, \Lambda) = q$
- $\forall q \in Q, \forall a \in T, \forall W \in T^*, g(q, Wa) = g(g(q, W), a)$

Le langage reconnu par M est $T(M) = \{W \in T^* / g(q_0, W) \in F\}$.

On démontre qu'un langage est un K-langage si et seulement s'il existe un automate d'états fini qui le reconnaisse.

- Un automate à pile de mémoire est un 7-uple

$$M = (Q, T, Z, g, z_0, q_0, F) \text{ où :}$$

- (i) Q, T, q_0 et F ont la même signification que pour les automates d'états finis.
- (ii) Z est l'alphabet de pile.
- (iii) g , la fonction de transition, est une application de $Q \times T^0 \times Z$ dans l'ensemble des parties finies de $Q \times Z^*$.

(iv) z_0 est un élément distingué de Z .

Si pour tout $q \in Q$ et $z \in Z$, $g(q, \Lambda, z) = \emptyset$, nous dirons que M est Λ -libre.

Notons $\mid_{\overline{M}}$ la relation binaire sur $Q \times T^* \times Z^*$ (appelée mouvement simple dans M) définie par :

$\forall q, q' \in Q, \forall W \in T^*, \forall b \in T^0, \forall y, y' \in Z^*, \forall z \in Z :$

$$(q, bW, yz) \mid_{\overline{M}} (q', W, yy') \iff (q', y') \in g(q, b, z)$$

$\mid_{\overline{M}}^*$ (appelé mouvement dans M) désignera la clôture transitive de la relation $\mid_{\overline{M}}$.

Le langage reconnu par l'automate M est :

$$T(M) = \{W \in T^* / \exists q \in F, \exists y \in Z^* \text{ tels que } (q_0, W, z_0) \mid_{\overline{M}}^* (q, \Lambda, y)\}.$$

Si pour tout $W \in T(M)$, il existe $q \in F$ tel que $(q_0, W, z_0) \mid_{\overline{M}}^* (q, \Lambda, z_0)$ nous dirons que M est complet.

On démontre que la classe des langages reconnus par un automate à pile de mémoire coïncide avec la classe des C-langages.

Une particularisation intéressante de ces automates a été introduite par Ginsburg et Greibach [GG2].

Un automate à pile de mémoire M est déterministe si et seulement s'il vérifie les propriétés suivantes :

$$- \forall q \in Q, \forall a \in T, \forall z \in Z, |g(q, a, z)| + |g(q, \Lambda, z)| = 1$$

- $\forall q \in Q, \forall b \in T^0$, si $g(q, b, z_0) \neq \emptyset$ alors il existe $y \in Z^*$ et $q' \in Q$ tels que $g(q, b, z_0) = \{(q', z_0 y)\}$.

Un automate à pile de mémoire M est **quasi-déterministe** si $\forall q \in Q, \forall a \in T, \forall z \in Z, |g(q,a,z)| + |g(q,\Lambda,z)| \leq 1$.

En fait, les automates à pile de mémoire déterministes et quasi-déterministes reconnaissent les mêmes types de langages, appelés **C-langages déterministes**.

- Un **automate linéairement borné déterministe** (sans marqueurs) [My] est un 5-uple $N = (Q, T, g, q_0, F)$ où :

- (i) Q, T, q_0, F ont la même signification que pour les automates d'états finis.
- (ii) g , la fonction de transition, est une application de $Q \times T$ dans $Q \times T \times \{-1, 0, +1\}$.

Notons $\mid_N$ la relation binaire sur $Q \times T^* \times T^*$ (appelée mouvement simple dans N), définie par :

$$\forall q, q' \in Q, \forall W, W_1, W', W'_1 \in T^*, \forall a, a' \in T,$$

$$(q, W, aW_1) \mid_N (q', W', W'_1) \iff Wa'W_1 = W'W'_1 \text{ avec } (q', a', i) = g(q, a) \text{ et } \ell(W') = \ell(W) + i.$$

$\mid_N^*$ (appelé mouvement dans N) désignera la clôture transitive de $\mid_N$.

Le langage reconnu par l'automate N est :

$$T(N) = \{W \in T^* / \exists q \in F, \exists W' \in T^* \text{ tels que } (q_0, \Lambda, W) \mid_N^* (q, W', \Lambda)\}$$

N est **sans boucle** si pour tout $W \in T^*$, il existe $q \in Q, W' \in T^*$ tels que :

$$(q_0, \Lambda, W) \mid_N^* (q, W', \Lambda).$$

Soient $\emptyset$ et $\$$ des symboles spéciaux (marqueurs).

Posons $T' = T \cup \{\emptyset, \$\}$.

Si g' est une application de $Q \times T'$ dans $Q \times T' \times \{-1, 0, +1\}$ vérifiant : $\forall q \in Q, \forall x \in \{\emptyset, \$\}, g'(q, x) = (q', y, i)$ implique $y = x$, alors $N' = (Q, T, g', q_0, F)$ est un automate linéairement borné déterministe avec les margueurs $\emptyset$ et $\$$.

La relation $\mid_{\overline{N}}$, est définie sur $Q \times T'^* \times T'^*$ et le langage reconnu par N' est :

$$T(N') = \{W \in T^* / \exists q \in F, W' \in T'^* \text{ tels que } (q_0, \emptyset, W\$) \mid_{\overline{N}}^* (q, \emptyset W \$, \Lambda)\}.$$

En fait [GH], ces deux types d'automates reconnaissent la même classe de langages notée L'_1 .

4 TRANSDUCTIONS

Soient X et Y deux alphabets.

- Un homomorphisme h de X^* dans Y^* est une application vérifiant : $\forall x, y \in X^*, h(xy) = h(x) h(y)$. Ceci implique que $h(\Lambda) = \Lambda$ et h est complètement déterminé par sa restriction à X .

Si, de plus, $\forall a \in X, h(a) \neq \Lambda$, alors, h est Λ -libre.

Et si $\forall a \in X, h(a) \in Y$ alors, h est strictement alphabétique.

Pour tout homomorphisme h de X^* dans Y^* , l'application, notée h^{-1} , de $P(Y^*)$ dans $P(X^*)$, définie par :

$$h^{-1}(L) = \{x \in X^* / h(x) \in L\} \quad \forall L \in P(Y^*),$$

sera appelée un homomorphisme inverse.

- Une partie R de $X^* \times Y^*$ est une **K-transduction** de X^* dans Y^* [Ni] s'il existe un alphabet Z , un K-langage K inclus dans Z^* et des homomorphismes h_1 et h_2 de Z^* dans, respectivement, X^* et Y^* , tels que

$$R = \{(h_1(W), h_2(W)) / W \in K\}.$$

- Un **transducteur séquentiel** est un 6-uple $T_S = (Q, X, Y, q_0, H, q_f)$ où :
- (i) Q, X, q_0 ont la même signification que pour les automates d'états finis.
 - (ii) Y est l'alphabet de sortie.
 - (iii) $q_f \in Q$ est l'état final.
 - (iv) H est une partie finie de $Q \times X^* \times Y^* \times Q$.

Si H est inclus dans $Q \times X^* \times Y^+ \times Q$, nous dirons que T_S est **Λ -libre**.

Notons $\mid_{T_S}$ la relation binaire sur $Q \times X^* \times Y^*$ définie par :

$$\forall q, q' \in Q, W, u \in X^*, W', v \in Y^*,$$

$$(q, uW, W') \mid_{T_S} (q', W, W'v) \iff (q, u, v, q') \in H$$

$\mid_{T_S}^*$ désignera la clôture transitive de $\mid_{T_S}$.

La transduction S associée à T_S est définie par :

$$S = \{(W, W') \in X^* \times Y^* / (q_0, W, \Lambda) \mid_{T_S}^* (q_f, \Lambda, W')\}.$$

Pour $L \subset X^*$, l'image transductée par T_S ,

$$f_S(L) \text{ est égale à } \{W' \in Y^* / (W, W') \in S, W \in L\}.$$

- Une **machine séquentielle généralisée** [G] est un 6-uple $M_S = (Q, X, Y, q_0, \delta, \lambda)$ où :
- (i) Q, X, Y, q_0 ont la même signification que pour les transducteurs séquentiels.

(ii) δ et λ sont des applications de $Q \times X$, respectivement, dans Q et Y^* .

Les applications δ et λ sont canoniquement étendues à $Q \times X^*$ par :

- $\forall q \in Q, \delta(q, \Lambda) = q ; \lambda(q, \Lambda) = \Lambda$
- $\forall q \in Q, \forall W \in X^*, \forall x \in X, \delta(q, Wx) = \delta(\delta(q, W), x),$
 $\lambda(q, Wx) = \lambda(q, W) \lambda(\delta(q, W), x).$

l'application A_S de X^* dans Y^* définie par

$$\forall X \in X^*, A_S(W) = \lambda(q_0, W),$$

est appelée une application séquentielle.

Première Partie

GRAMMAIRES MONOTONES

Chapitre I

FORMES STANDARDS DE GRAMMAIRES MONOTONES

Nous allons introduire, dans ce chapitre, la notion de grammaire monotone. Pour ces grammaires, il existe un sous-ensemble de variables contenant strictement les terminaux, vis-à-vis desquelles les productions sont de longueur non décroissante. L'ensemble des non-terminaux est ainsi partitionné en deux. En ajoutant des restrictions liant ces deux sous-ensembles aux productions, nous définissons des grammaires monotones de différents types. Le premier travail consiste à trouver pour chacun de ces types, une forme standard équivalente, plus facile à manier, et à laquelle nous nous référons dans les chapitres suivants. Nous en déduisons, en premier lieu, une condition suffisante pour qu'une grammaire à structure de phrase, non régulière, engendre un langage régulier.

Soient X un alphabet et Z une partie de X .

Notation. Pour tout $x \in X^*$, $\ell_Z(x)$ représentera le nombre d'occurrences d'éléments de Z dans x et $\ell(x) = \ell_X(x)$ désignera la longueur de x .

Soient $G = (N, T, P, A)$ une grammaire à structure de phrase, N_2 une partie de N et $V_2 = N_2 \cup T$. (Par la suite, chaque fois que nous introduisons une grammaire G sans spécification, N, T, P désigneront respectivement l'ensemble des non-terminaux, des terminaux et des productions; l'axiome sera noté A et V sera l'ensemble des variables $N \cup T$).

Nous dirons que G est monotone de type 0 (resp. de type 6) en N_2 si et seulement si $(u, v) \in P$ implique $1 \leq \ell_{N_2}(u) \leq \ell_{V_2}(v)$ (resp. $1 = \ell_{N_2}(u) = \ell_{V_2}(v)$).

Remarque Nous supposons par la suite que $A \in N_2$. En effet dans le cas contraire $(u, v) \in P$ entraîne $u \neq A_2$ et donc $L(G) = \emptyset$

Nous dirons que G est monotone de type 0 (resp. de type 6) si et seulement s'il existe une partie N_2 de N telle que G soit monotone de type 0 (resp. de type 6) en N_2 .

Remarquons que la notion de grammaire monotone de type 0 est une généralisation de celle de grammaire "monotonic". (En prenant $N_2 = N$, nous retrouvons la propriété : $(u,v) \in P$ implique $1 \leq l(u) \leq l(v)$). En fait la propriété suivante montre que la classe des langages engendrés par une grammaire monotone de type 0 est beaucoup plus large que la classe des langages contextuels.

Proposition I.1 Pour tout langage récursivement énumérable $L \subset T^*$, il existe une grammaire monotone de type 0 qui engendre $L \setminus \{\Lambda\}$.

Démonstration L étant récursivement énumérable, il existe une grammaire G à structure de phrase telle que $L(G) = L$. A T faisons correspondre biunivoquement l'ensemble $\bar{T} = \{\bar{a} / a \in T\}$ tel que $\bar{T} \cap V = \emptyset$ et considérons l'homomorphisme h de V^* dans $(N \cup \bar{T})^*$ défini par $h(B) = B$ pour tout $B \in N$ et $h(a) = \bar{a}$ pour tout $a \in T$.

Posons $N' = N \cup \bar{T}$, $V' = N' \cup T$, $P'_1 = \{(u, h(v)) / (u, v) \in P\}$

$P'_2 = \{(\bar{a}, a) / a \in T\}$ $P' = P'_1 \cup P'_2$ et $G' = (N', T, P', A)$.

Il est facile de voir que $L(G') = L(G)$.

Prenons maintenant M et $A'' \notin N' \cup T$ et posons

$P''_1 = \{(Mu, Mv) / (u, v) \in P'_1\}$,

$P''_2 = \{(MB, BM) / B \in N'\} \cup \{(BM, MB) / B \in N'\}$,

$P''_3 = \{(M\bar{a}, aM) / a \in T\} \cup \{(M\bar{a}, a) / a \in T\}$,

$P'' = \{(A'', MA)\} \cup P''_1 \cup P''_2 \cup P''_3$,

$N_2 = \{M, A''\}$, $V_2 = N_2 \cup T$, $N'' = N_2 \cup N'$, $V'' = N'' \cup T$ et $G'' = (N'', T, P'', A'')$

Remarquons que $(u'', v'') \in P''$ implique $\ell_{N_2}(u'') = 1$ et $\ell_{V_2}(u'') \geq 1$. Donc G'' est monotone de type 0 en N_2 . Il reste à montrer que $L(G'') = L(G') \setminus \{\Lambda\}$.

- Montrons que $L(G'') \subseteq L(G') \setminus \{\Lambda\}$

$\Lambda \notin L(G'')$, en effet $(u'', v'') \in P''$ entraîne $\ell(v'') \geq \ell_{V_2}(v'') > 0$, donc $v'' \neq \Lambda$. Considérons l'homomorphisme g de V''^* dans V'^* défini par $g(M) = \Lambda$, $g(A'') = A$ et $g(B) = B$ pour tout $B \in V'$. Il est facile de vérifier que $(u'', v'') \in P''$ implique soit $g(u'') = g(v'')$, soit $(g(u''), g(v'')) \in P'$. Nous en déduisons que si $x \xrightarrow[G'']{*} y$ alors $g(x) \xrightarrow[G']{*} g(y)$ et en particulier $A'' \xrightarrow[G'']{*} W$ avec $W \in T^*$ entraîne $g(A'') = A \xrightarrow[G']{*} g(W) = W$.

Donc $W \in L(G'')$ implique $W \in L(G') \setminus \{\Lambda\}$.

- Montrons maintenant que $L(G') \setminus \{\Lambda\} \subseteq L(G'')$.

Soit $x \in L(G') \setminus \{\Lambda\}$. Il existe une dérivation $A \xRightarrow[G']{*} h(x) \xRightarrow[G']{*} x$ (il suffit de regrouper les productions de P'_2 après celles de P'_1 dans la dérivation de x). De plus, si $W_1 \xRightarrow[G']{*} W_2$ avec $W_1, W_2 \in N'^*$, alors $MW_1 \xRightarrow[G'']{*} MW_2$ (en utilisant les productions de P''_1 et P''_2). Donc $A'' \xRightarrow[G'']{*} MA \xRightarrow[G'']{*} Mh(x)$.

Or $Mh(x) \xRightarrow[G'']{*} x$ si $x \in T^+$ (en utilisant les productions de P''_3) ce qui implique $A'' \xRightarrow[G'']{*} x$ et $x \in L(G'')$.

Considérons maintenant une grammaire G monotone de type 6 en $N_2 \subset N$. Le langage engendré par G est inclus dans T . En effet $A \xRightarrow[G]{*} x$ implique $\ell_{V_2}(x) = 1$ avec $V_2 = N_2 \cup T$. Donc $x \in L(G)$ entraîne $\ell(x) = \ell_T(x) = 1$. Pourtant il n'existe pas d'algorithme général pour décider si un mot appartient au langage engendré par une grammaire monotone de type 6. Montrons en effet la propriété suivante :

Proposition I.2 Pour toute classe C de grammaires contenant celle des grammaires monotones de type 6, il est indécidable de savoir si un mot appartient au langage engendré par une grammaire de C .

Démonstration Il suffit de montrer la propriété pour une famille particulière de grammaires monotones de type 6.

Ces grammaires n'auront qu'un seul terminal.

Posons $T = \{a\}$, $N_1 = \{E_1, E_2\}$, $N_2 = \{A, \bar{a}\}$ et

$P = \{(E_1 A E_1, \bar{a}), (E_2 A E_2, \bar{a}), (E_1 \bar{a} E_1, \bar{a}), (E_2 \bar{a} E_2, \bar{a}), (\bar{a}, a)\}$.

A tout $n \geq 1$, à toute liste $x^{(n)} = (x_1, x_2, \dots, x_n)$ et à toute liste $y^{(n)} = (y_1, y_2, \dots, y_n)$ de mots de N_1^* ($x_j, y_j \in N_1^*$ pour $j \in [1, n]$) faisons correspondre la grammaire $G(n, x^{(n)}, y^{(n)}) = (N_1 \cup N_2, T, P(n, x^{(n)}, y^{(n)}), A)$ où

$$P(n, x^{(n)}, y^{(n)}) = P \cup \{(A, x_i \tilde{A} y_i) / i=1 \dots n\}$$

Il est facile de vérifier que $G(n, x^{(n)}, y^{(n)})$ est une grammaire monotone de type 6 en N_2 . Montrons qu'il n'existe pas d'algorithme général pour décider si $a \in L(G(n, x^{(n)}, y^{(n)}))$.

En effet $a \in L(G(n, x^{(n)}, y^{(n)}))$ si et seulement si $A \xRightarrow{*} \bar{a}$ c'est-à-dire si et seulement si il existe une suite de coefficients $i_1, i_2, \dots, i_r$ ($i_j \in [1, n]$ pour $j=1 \dots r$) telle que $x_{i_1} \dots x_{i_r} \tilde{A} y_{i_1} \dots \tilde{y}_{i_r} \xRightarrow{*} \bar{a}$ ou telle que $x_{i_1} \dots x_{i_r} = y_{i_1} \dots y_{i_r}$ ce qui est le problème de correspondance de Post pour deux listes d'éléments de N_1^* . Ce problème est indécidable puisque $|N_1| > 1$ [P0].

Introduisons maintenant un certain nombre de sous-classes de grammaires monotones de type 0, en prenant soin que ces classes ne contiennent pas la classe des grammaires monotones de type 6, ce qui entraînerait, d'après la proposition précédente, l'indécidabilité du problème du mot pour ces classes.

Soit T un alphabet. Pour tous alphabets N_1 et N_2 tels que N_1 , N_2 et T soient disjoints deux à deux, posons : $N = N_1 \cup N_2$, $V_1 = N_1 \cup T$, $V_2 = N_2 \cup T$, $V = N \cup T$ et définissons sur $N^+ \times V^*$ les prédicats suivants :

$$\begin{aligned} M_{N_1, N_2}^0(u, v) &\equiv (1 \leq \ell_{N_2}(u) \leq \ell_{V_2}(v)) \\ M_{N_1, N_2}^4(u, v) &\equiv M_{N_1, N_2}^0(u, v) \text{ et } (u \in N_1^* N_2) \\ M_{N_1, N_2}^5(u, v) &\equiv M_{N_1, N_2}^0(u, v) \text{ et } (v \in V_2^* V_2) \\ M_{N_1, N_2}^1(u, v) &\equiv M_{N_1, N_2}^4(u, v) \text{ et } M_{N_1, N_2}^5(u, v) \\ M_{N_1, N_2}^2(u, v) &\equiv (u \in N_1^* N_2) \text{ et } (v \in V_1^* V_2) \\ M_{N_1, N_2}^3(u, v) &\equiv (\ell_{N_2}(u) = 1) \text{ et } (v \in V_1^* V_2) \\ M_{N_1, N_2}^6(u, v) &\equiv (\ell_{N_2}(u) = \ell_{V_2}(v) = 1) \end{aligned}$$

Remarque Si $N'_1 \supset N_1$, $N'_2 \supset N_2$ et si N'_1 , N'_2 , T , sont encore disjoints deux à deux, alors $M_{N_1, N_2}(u, v)$ implique $M_{N'_1, N'_2}(u, v)$.

Dans tout ce qui suit, nous considérons une grammaire $G = (N, T, P, A)$ à structure de phrase, une partition de N en N_1, N_2 et nous gardons pour V_1 , V_2 et V les définitions données ci-dessus.

Nous dirons qu'une partie P' de P est monotone de type i en N_2 si et seulement si $(u, v) \in P'$ implique $M_{N_1, N_2}^i(u, v)$; G sera dite monotone de type i en N_2 s'il en est ainsi pour P .

Nous dirons que G est **monotone de type i** s'il existe une partie N_2 de N telle que G soit monotone de type i en N_2 .

Remarque Il est décidable de déterminer si une grammaire G est monotone de type i car il existe 2^p sous-ensembles différents de N avec $p = |N|$.

En utilisant quelques lemmes, nous allons trouver, pour chaque type de grammaire monotone, une forme standard équivalente à laquelle nous ferons référence tout au long de la première partie.

Lemme I.1 Pour toute règle $\pi = (u, v)$, monotone de type i en N_2 , il existe N'_1 contenant N_1 et un ensemble P' de règles, monotone de type i en N_2 tel que :

$$- \forall (u', v') \in P', \ell_{N_2}(u') = 1$$

$$- G' = (N_2 \cup N'_1, T, (P \setminus \pi) \cup P', A) \text{ soit équivalente à } G.$$

Démonstration Si $\ell_{N_2}(u) = 1$ (en particulier pour $i=2,3$ ou 6) la propriété est vérifiée en prenant $N'_1 = N_1$ et $P' = \{\pi\}$.

Supposons la propriété vraie pour $\ell_{N_2}(u) \leq k$.

Si $\ell_{N_2}(u) = k+1 > 1$, u et v peuvent se factoriser en : $u = X_1 B Y$ et $v = X'_1 C X'_2 Z$ avec $X_1, X'_1, X'_2 \in N_1^*$, $B \in N_2$, $Y \in N_1^+$, $C \in V_2$ et $Z \in V_2 V_2^*$. Prenons $D \notin V$ et posons $u_1 = X_1 B$, $v_1 = X'_1 C X'_2 D$, $u_2 = D Y$, $v_2 = Z$ et $N''_1 = N_1 \cup \{D\}$.

Il est facile de vérifier que quelque soit la valeur de i , nous obtenons $M_{N''_1, N_2}^4(u_1, v_1)$ et $M_{N''_1, N_2}^5(u_2, v_2)$.

De plus si $i=4$ $Y \in N^* N_2$ ce qui implique $M_{N''_1, N_2}^4(u_2, v_2)$

et si $i=5$ $X'_1 = \Lambda$ donc $M_{N''_1, N_2}^5(u_1, v_1)$.

Nous en déduisons que $M_{N''_1, N_2}^i(u, v)$ entraîne

$$M_{N''_1, N_2}^i(u_1, v_1) \text{ et } M_{N''_1, N_2}^i(u_2, v_2) \quad (\text{pour } i=0,1,4 \text{ ou } 5)$$

Considérons alors la grammaire $G'' = (N''_1 \cup N_2, T, P'', A)$ où $P'' = (P \setminus \pi) \cup \{(u_1, v_1), (u_2, v_2)\}$.

Il est évident que $L(G'') = L(G)$ et comme nous avons $M_{N_1'', N_2}^i(u_2, v_2)$ et $\ell_{N_2}(u_2) = k$, nous pouvons employer l'hypothèse de récurrence. Il existe donc $N_1' \supset N_1''$ et P_1' tels que $G' = (N_1' \cup N_2, T, (P'' \setminus \{(u_2, v_2)\}) \cup P_1', A)$ soit équivalente à G'' et tels que $(u', v') \in P_1'$ entraîne $M_{N_1', N_2}^i(u', v')$ et $\ell_{N_2}(u') = 1$.

En posant $P' = P_1' \cup \{(u_1, v_1)\}$, la propriété est vérifiée car $\ell_{N_2}(u_1) = 1$ et $M_{N_1'', N_2}^i(u_1, v_1)$ donc $M_{N_1', N_2}^i(u_1, v_1)$ ($N_1' \supset N_1''$).

Lemme I.2 Pour toute règle $\pi = (u, v)$, monotone de type i en N_2 telle que $\ell_{N_2}(u) = 1$, il existe $N_2' \supset N_2$ et un ensemble de règles P' monotone de type i en N_2' tels que

- P' est inclus dans $(N_1 N_2' \cup N_2' N_1) \times N_2' \cup N_2' \times (V_1 \cup N_2')^+$
- $G' = (N_1 \cup N_2', T, (P \setminus \pi) \cup P', A)$ soit équivalente à G .

Démonstration Comme $\ell_{N_2}(u) = 1$, u peut se factoriser en EBF avec $B \in N_2$, $E, F \in N_1^*$. Distinguons deux cas :

(1) $F = A$ (en particulier pour $i=1, 2$ ou 4).

Si $E = A$, le lemme est vérifié en prenant $N_2' = N_2$ et $P' = \{\pi\}$. Sinon E se factorise en $E_1 E_2 \dots E_p$ avec $E_j \in N_1$ pour $j \in [1, p]$. Considérons $A_1, \dots, A_p \notin V$. Il est facile de voir que la propriété est vérifiée en posant $N_2' = N_2 \cup \bigcup_{j=1}^p \{A_j\}$ et

$$P' = \{(E_p B, A_p)\} \cup \bigcup_{j=1}^{p-1} \{(E_j A_{j+1}, A_j)\} \cup \{(A_1, v)\}.$$

(2) $F \neq A$ donc se factorise en $F_1 \dots F_k$ avec $F_j \in N_1$ pour $j \in [1, k]$. Considérons $B_1, \dots, B_k \notin V$. Posons $N_2' = N_2 \cup \bigcup_{j=1}^k \{B_j\}$, $P'' = \{(BF_1, B_1)\} \cup \bigcup_{j=1}^{k-1} \{(B_j F_{j+1}, B_{j+1})\}$

et $G'' = (N_1 \cup N_2', T, (P \setminus \pi) \cup P'' \cup \{(EB_k, v)\}, A)$.

Il est clair que $L(G'') = L(G)$ et que l'on a $M_{N_1, N_2'}^i(u'', v'')$ pour toute $(u'', v'') \in P''$.

Il reste à appliquer la transformation du (1) à la règle (EB_k, v) pour achever la démonstration.

Ce résultat nous conduit à poser la définition suivante :

Nous dirons qu'une partie P' de P est monotone standard de type i en N_2 si et seulement si elle est monotone de type i en N_2 et si en outre :

- lorsque i est égal à 3 : $P' \subset (N_1 N_2 \cup N_2 N_1) \times N_2 \cup N_2 \times (TN_2^0 \cup N_1^0 N_2)$
- lorsque i est différent de 3 : $P' \subset (N_1 N_2 \cup N_2 N_1) \times N_2 \cup N_2 \times (T \cup N_1^0 N_2 \cup N_2 N)$

La grammaire G sera dite **monotone standard de type i en N_2** s'il en est ainsi pour P .

Lemme I.3 Pour toute règle $\pi = (u, v)$, monotone de type i en N_2 telle que $u \in N_2$, il existe $N'_2 \supset N_2$ et un ensemble de règles P' monotone standard de type i en N'_2 tels que $G' = (N_1 \cup N'_2, T, (P \setminus \pi) \cup P', A)$ soit équivalente à G .

Démonstration Si $\ell(v) = 1$, il suffit de prendre $N'_2 = N_2$ et $P' = \{\pi\}$. Sinon distinguons deux cas :

(1) $i = 3$.

v se factorise alors en $v_1 \dots v_k B$ avec $B \in V_2$ et $v_j \in V_1$ pour $j \in [1, k]$. Considérons $A_1, \dots, A_k \notin V$. La propriété sera vérifiée en prenant $N'_2 = N_2 \bigcup_{j=1}^k \{A_j\}$ et $P' = \{(u, v_1 A_1)\} \cup \{(A_k, B)\} \bigcup_{j=1}^{k-1} \{(A_j, v_{j+1} A_{j+1})\}$.

(2) $i \neq 3$.

Nous pouvons supposer que $\ell_T(v) = 0$. Dans le cas contraire, il suffit de reprendre l'ensemble $\bar{T}$ et l'homomorphisme h définis dans la démonstration de la proposition I.1. Posons $\bar{P} = \{(\bar{a}, a) / a \in T\}$. $\bar{P}$ est bien monotone standard de type i en $N_2 \cup \bar{T}$. Nous ajoutons $\bar{P}$ à P et nous remplaçons la règle (u, v) par $(u, h(v))$ qui vérifie bien $\ell_T(h(v)) = 0$. Distinguons encore deux cas :

(a) $i=1$ ou 5 .

v se factorise en $B v_1 \dots v_p$ avec $B \in V_2$ et $v_j \in V$ pour $j \in [1, p]$. Considérons $A_1, \dots, A_p \notin V$ et prenons $N'_2 = N_2 \bigcup_{j=1}^p \{A_j\}$, $P' = \{(u, A_p v_p)\} \bigcup_{j=1}^{p-1} \{(A_{j+1}, A_j v_j)\} \cup \{(A_1, B)\}$ qui vérifient les propriétés du lemme.

(b) $i=0, 2, 4$ ou 6

v se factorise en $v_1 \dots v_k B E_1 \dots E_p$ avec $B \in V_2$, $v_j \in V$ pour $j \in [1, p]$ et $E_s \in N_1$ pour $s \in [1, p]$. Considérons $A_0, A_1, \dots, A_k, B_1, \dots, B_p, B_{p+1} \notin V$. Posons $N'_2 = N_2 \bigcup_{j=0}^k \{A_j\} \bigcup_{s=1}^{p+1} \{B_s\}$ et $P' = \{(u, A_0)\} \bigcup_{j=1}^k \{(A_{j-1}, v_j A_j)\} \cup \{(A_k, B_{p+1})\} \bigcup_{s=1}^p \{(B_{s+1}, B_s E_s)\} \cup \{(B_1, B)\}$

Il est clair que pour $i=0, 4, 6$, N'_2 et P' vérifient les propriétés voulues. De plus si $i=2$, $v \in V^* V_2$ donc $p=0$ et $(u', v') \in P'$ entraîne $M_{N_1, N'_2}^2(u', v')$.

Des lemmes I.1, I.2, et I.3, nous déduisons immédiatement le résultat suivant :

Lemme I.4 Pour toute partie P_1 de P , monotone de type i en N_2 , il existe $N'_1 \supset N_1$, $N'_2 \supset N_2$ disjoints et P' monotone standard de type i en N'_2 tels que $G' = (N'_1 \cup N'_2, T, (P \setminus P_1) \cup P', A)$ soit équivalente à G .

En particulier, si nous prenons $P_1 = P$, nous obtenons :

Corollaire I.1 Toute grammaire monotone de type i est équivalente à une grammaire monotone standard de même type.

Etudions maintenant plus particulièrement les grammaires monotones de type 3 et montrons qu'elles engendrent exactement les K -langages.

Lemme I.5 Soit G une grammaire monotone standard de type 3. Pour tout $B, C \in N$, il est décidable de déterminer si $B \xrightarrow[G]{*} C$.

Démonstration Remarquons d'abord que $A \xrightarrow[G]{*} W$ implique $W \in V_1^* V_2$, donc les règles appartenant à $N_2 N_1 \times N_2$ ne pourront jamais être appliquées. Nous pouvons alors supposer que P est inclus dans $N_1 N_2 \times N_2 \cup N_2 \times (TN_2^0 \cup N_1^0 N_2)$.

- Si $B = C$ alors $B \xrightarrow[G]{*} C$
- Si $B \neq C$ et si $B \in N_1$ ou $C \in N_1$, alors C ne peut se dériver de B car $(u, v) \in P$ implique $\ell_{N_2}(u) = 1$ et $\ell_{V_2}(v) \geq 1$.
- Si $B \neq C$ et si B et $C \in N_2$, prenons $Z_0, q_f, M \notin V$ et posons $K = N_2 \cup \{q_f\}$ et $H = N_1 \cup \{Z_0\}$.

Considérons l'automate à pile suivant :

$M = (K, \{a\}, H, g, Z_0, B, \{q_f\})$ ou g est défini par :

Pour tout $A_2 \in N_2$

$$g(A_2, \Lambda, Z_0) = \{(B_2, Z_0 E) / (B, E B_2) \in P \text{ avec } B_2 \in N_2 \text{ et } E \in N_1^0\}$$

Pour tout $A_2 \in N_2$ et $E \in N_1$,

$$g(A_2, \Lambda, E) = \{(B_2, \Lambda) / (E A_2, B_2) \in P \text{ avec } B_2 \in N_2\} \cup \{(B_2, E E') / (A_2, E' B_2) \in P \text{ avec } B_2 \in N_2 \text{ et } E' \in N_1^0\}$$

et $g(C, a, Z_0) = \{(q_f, Z_0)\}$

Il est facile de vérifier que $B \xrightarrow[G]{*} C$ si et seulement si $a \in T(M)$ ce qui est décidable puisque $T(M)$ est un C-langage.

Proposition I.3 $L \subset T^*$ est un K-langage si et seulement s'il existe une grammaire monotone de type 3 qui l'engendre.

Démonstration Si L est un K-langage, il existe une grammaire régulière $G' = (N', T, P', A)$ telle que $L(G') = L$. Comme G' est régulière P' est inclus dans $N' \times TN'^0$ donc G' est monotone de type 3 en N' .

Réciproquement, soit G une grammaire monotone de type 3 en N_2 . D'après le corollaire I.1, nous pouvons supposer que G est monotone standard de type 3 en N_2 . Considérons $G' = (N_2, T, P', A)$ avec

$$P' = \{(u, v) / (u, v) \in P, u \in N_2 \text{ et } \ell_{N_1}(v) = 0\} \\ \cup \{(B, C) / B \xrightarrow[G]{*} C, B, C \in N_2 \text{ et } B \neq C\}.$$

P' est inclus dans $N_2 \times (TN_2^0 \cup N_2)$ donc G' est linéaire à droite et $L(G')$ est un K-langage. De plus $L(G') \subseteq L(G)$ car toutes les règles de G' sont des dérivations dans G . Il reste à montrer que $L(G) \subseteq L(G')$. Pour cela nous utiliserons la propriété (1):

$$(1) \quad A \xrightarrow[G]{*} xB \text{ avec } x \in T^*, B \in N_2 \text{ implique } A \xrightarrow[G']{*} xB.$$

Nous raisonnerons par récurrence sur la longueur de x .

Si $x = \Lambda$, (1) est vérifiée par construction de P' .

Supposons que (1) soit vraie pour $\ell(x) \leq k$ et considérons la dérivation $A \xrightarrow[G]{*} xaB$ avec $x \in T^k$, $a \in T$ et $B \in N_2$.

$$\text{Nous avons } A \xrightarrow[G]{*} yC \xRightarrow[G]{*} yaD \xrightarrow[G]{*} xaB \text{ avec } C, D \in N_2, \\ y \xrightarrow[G]{*} x \text{ et } D \xRightarrow[G]{*} B \text{ (donc } (D, B) \in P').$$

Comme $\ell_{N_2}(y) = 0$, $y=x$ ce qui implique, en utilisant l'hypothèse de récurrence:

$$A \xrightarrow[G']{*} xC \xRightarrow[G']{*} xaD \xRightarrow[G']{*} xaB.$$

Soit, alors, $x \in L(G)$: $x = x'a$ avec $a \in T$. La dérivation de x à partir de A peut s'écrire :

$$A \xrightarrow[G]{*} x'B \xRightarrow[G]{*} x'a \text{ avec } B \in N_2.$$

$(B,a) \in P'$ et la propriété (1) implique :

$$A \xrightarrow[G']{*} x'B \xrightarrow[G']{=} x'a = x.$$

Chapitre II

GRAMMAIRES MONOTONES ET C-LANGAGES

Dans ce chapitre, nous allons montrer que les langages engendrés par une grammaire $G = (N, T, P, A)$ monotone de type 2 sont des C-langages. En fait nous obtenons un résultat plus fort en montrant qu'il suffit de considérer les règles où ne figurent pas de terminaux. Plus précisément, s'il existe une partie N_2 de N telle que $\{(u, v) \notin P / u \in N \setminus N_2 \text{ et } \ell_T(v) = 0\}$ soit monotone de type 2 en N_2 , alors $L(G)$ est un C-langage.

Rappelons d'abord le théorème de Ginsburg et Greibach qui donne une condition suffisante pour qu'une grammaire à structure de phrase engendre un C-langage.

Théorème [GG1] Soit $G = (N, T, P, A)$ une grammaire à structure de phrase. Si $\ell_T(v) \geq 1$ pour tout $(u, v) \in P$, alors $L(G)$ est un C-langage.

En modifiant les démonstrations qui ont permis de montrer ce théorème, nous allons en trouver une première généralisation que nous utiliserons pour obtenir le résultat principal de ce chapitre. Nous conservons les notations du chapitre I. En outre, pour toute grammaire $G = (N, T, P, A)$, nous noterons :

$$P_T = \{(u, v) \in P / \ell_T(v) \geq 1\} \text{ et } P_N = P \setminus P_T.$$

Lemme II.1 Pour toute grammaire $G = (N, T, P, A)$, il existe $N' \supset N$, $T' \supset T$, $P'_0 \subset N \times N^*N''$, $P''_0 \subset N''N \times T'N'' \cup N' \times TV^*$ (en notant $N'' = N' \setminus N$) et un homomorphisme h de T'^* dans T^* tels que la grammaire $G' = (N', T', P_N \cup P'_0 \cup P''_0, A)$ vérifie $h(L(G')) = L(G)$.

Démonstration Soient $r_1, \dots, r_n$ les différentes règles de P_T . Pour $i \in [1, n]$, $r_i = (u_i, v_i)$, u_i se factorisant en $X_{i1} \dots X_{ik_i}$ avec $X_{ij} \in N$ pour $j \in [1, k_i]$. De même $v_i = y_i b_i z_i$ avec $y_i \in N^*$, $b_i \in T$ et $z_i \in V^*$. Prenons $M_{i1}, \dots, M_{ik_i}$, $a_{i1}, \dots, a_{i(k_i-1)} \in V$ et posons $N_i = \{M_{i1}, \dots, M_{ik_i}\}$, $T_i = \{a_{i1}, \dots, a_{i(k_i-1)}\}$, $P'_i = \{(X_{i1}, y_i M_{i1})\}$ et $P''_i = \{(M_{ik_i}, b_i z_i)\} \cup_{j=1}^{k_i-1} \{(M_{ij} X_{i(j+1)}, a_{ij} M_{i(j+1)})\}$.

Considérons la grammaire $G' = (N', T', P_N \cup P'_0 \cup P''_0, A)$ où $T' = T \bigcup_{i=1}^n T_i$,
 $N' = N \bigcup_{i=1}^n N_i$, $P'_0 = \bigcup_{i=1}^n P'_i$ et $P''_0 = \bigcup_{i=1}^n P''_i$.

Nous obtenons bien $P'_0 \subset N \times N^*N''$ et $P''_0 \subset N''N \times T'N'' \cup N' \times TV^*$ avec
 $N'' = \bigcup_{i=1}^n N_i = N' \setminus N$.

De plus, il est clair que $h(L(G')) = L(G)$, h étant l'homomorphisme de T'^* dans T^* défini par $h(a) = a$ pour tout $a \in T$ et $h(a_{ij}) = \Lambda$ pour tout $a_{ij} \in T' \setminus T$.

Lemme II.2 Si P_T est inclus dans $NN^0 \times TV^*$, il existe $T' = \{a_i, a'_i / 1 \leq i \leq n\}$ disjoint de V , avec $n = |P_T| + |T|$, $N' \supset N$, un ensemble $P'_T = \{\pi'_i / 1 \leq i \leq n\}$ avec $\pi'_i \in N'N'^0 \times \{a_i, a'_i\} N'^*$ pour $i \in [1, n]$ et une application séquentielle f_S

tels que la grammaire $G' = (N', T', P_N \cup P'_T, A)$ vérifie $f_S(L(G')) = L(G)$.

Démonstration A T faisons correspondre biunivoquement $\bar{T} = \{\bar{a} / a \in T\}$.
 Posons $N' = N \cup \bar{T}$ et définissons l'homomorphisme g de V^* dans N'^* par $g(B) = B$ pour tout $B \in N$ et $g(a) = \bar{a}$ pour tout $a \in T$. Posons $P_{\bar{T}} = \{(u, ag(z)) / (u, az) \in P_T$ et $a \in T\} \cup \{(\bar{a}, a) / a \in T\}$. $|P_{\bar{T}}| = |P_T| + |T| = n$. Soient $\pi_1, \dots, \pi_n$ les différentes règles de $P_{\bar{T}}$. Si $\pi_i = (u_i, b_i z_i)$ avec $b_i \in T$, posons $\pi'_i = (u_i, a_i a'_i z_i)$. π'_i est inclus dans $NN^0 \times \{a_i, a'_i\} N'^* \cup \bar{T} \times \{a_i, a'_i\}$ donc dans $N'N'^0 \times \{a_i, a'_i\} N'^*$.

Considérons la machine séquentielle généralisée

$S = (K, T', T, q_0, \delta, \lambda)$ avec $K = \{q_0, q_1\}$

δ définie par : pour $i \in [1, n]$

$\delta(q_0, a_i) = \delta(q_1, a_i) = q_1$

$\delta(q_0, a'_i) = \delta(q_1, a'_i) = q_0$

et λ définie par : pour $i \in [1, n]$

$\lambda(q_0, a_i) = \lambda(q_1, a_i) = \Lambda$

$\lambda(q_0, a'_i) = \lambda(q_1, a'_i) = b_i$.

Comme $L(G')$ est inclus dans $\{a_i, a'_i / i \in [1, n]\}^*$, il est facile de vérifier que $f_S(L(G')) = L(G)$.

Lemme II.3 Soit π une règle appartenant à $N^2 \times \{ab\}$ N^* avec $a, b \in T$.
 Si $\forall (u,v) \in P \setminus \{\pi\}$, $\ell_{\{a\}}(v) = \ell_{\{b\}}(v) = 0$, il existe $P' \subset N \times TN^*$ et un
 K-langage R tels que la grammaire $G' = (N, T, (P \setminus \{\pi\}) \cup P', A)$ vérifie $L(G) = L(G') \cap R$.

Démonstration Posons $\pi = (BB', abz)$ avec $B, B' \in N$ et $z \in N^*$,
 $P' = \{\pi_a, \pi_b\}$ avec $\pi_a = (B, a)$ et $\pi_b = (B', bz)$,
 $R = ((T \setminus \{a, b\}) \cup \{ab\})^*$ et $G'' = (N, T, P \cup P', A)$.

Comme R contient $L(G)$ et que $\pi \in P \cup P'$, $L(G) \subseteq L(G'') \cap R$.

Montrons que $L(G'') \cap R \subseteq L(G)$.

Soit $x \in R \cap L(G'')$; il existe une dérivation dans G'' : $A \xRightarrow{*}_{G''} x$.

Notons k le nombre d'utilisations de la règle π_a dans cette dérivation et raisonnons par récurrence sur k .

Si $k=0$, la règle π_b ne peut être utilisée dans la dérivation sinon $x \notin R$. x se dérive en n'utilisant que des règles de P donc $x \in L(G)$.

Supposons que $x \in L(G)$ si $x \in R$ et s'il existe une dérivation de x dans G'' avec $k \leq p$.

Considérons une dérivation $A \xRightarrow{*}_{G''} x \in R$ avec $k = p+1$. Comme la règle π_a a été utilisée au moins une fois et que $x \in R$, x peut se factoriser en $x'abx''$.

Exhibons l'apparition de cette occurrence de a , dans la dérivation :

$$A \xRightarrow{*}_{G''} WBW' \xRightarrow{*}_{G''} WaW' \xRightarrow{*}_{G''} x'abx'' \text{ ce qui entraîne } W \xRightarrow{*}_{G''} x' \text{ et } W' \xRightarrow{*}_{G''} bx''$$

Distinguons deux cas :

$$(1) \quad W' = bW''_1 \text{ donc } A \xRightarrow{*}_{G''} WBbW''_1$$

Considérons l'apparition de cette occurrence de b :

$$A \xRightarrow{*}_{G''} W_1 B' W'_1 \xRightarrow{*}_{G''} W_1 bz W'_1 \xRightarrow{*}_{G''} WBbW''_1 \text{ avec } W_1 \xRightarrow{*}_{G''} WB \text{ et } zW'_1 \xRightarrow{*}_{G''} W''_1 \text{ ce qui nous}$$

permet de transformer la dérivation, en changeant l'ordre de l'emploi des règles :

$$A \xRightarrow{*}_{G''} W_1 B' W'_1 \xRightarrow{*}_{G''} WBB' W'_1 \xRightarrow{*}_{G''} WaB' W'_1 \xRightarrow{*}_{G''} Wabz W'_1 \xRightarrow{*}_{G''} x'abx''$$

mais $WBB' W'_1 \xRightarrow{*}_{G''} Wabz W'_1$ en utilisant la seule règle π donc $A \xRightarrow{*}_{G''} x \in R$

avec $k=p$ et $x \in L(G)$ d'après l'hypothèse de récurrence.

(2) Dans le cas contraire, W' se dérive de la façon suivante :

$$W' \xrightarrow[G'']{*} B'W''_1 \xrightarrow[G'']{=} bzW''_1 \xrightarrow[G'']{*} bx'' \text{ et nous obtenons la dérivation :}$$

$$A \xrightarrow[G'']{*} WBW' \xrightarrow[G'']{*} WBB'W''_1 \xrightarrow[G'']{*} WabzW''_1 \xrightarrow[G'']{*} x'abx''$$

mais $WBB'W''_1 \xrightarrow[G'']{=} WabzW''_1$ en utilisant la règle π donc $A \xrightarrow[G'']{*} x \in R$ avec $k=p$ et $x \in L(G)$ d'après l'hypothèse de récurrence.

Posons maintenant $G' = (N, T, (P \setminus \{\pi\}) \cup P', A)$. Comme π est une dérivation dans G' , $L(G') = L(G'')$ donc $L(G) = L(G') \cap R$. De plus P' est bien inclus dans $N \times TN^*$ et il est clair que R est un K-langage.

Ces trois lemmes nous permettent de généraliser le théorème de Ginsburg et Greibach :

Proposition II.1 Soit $G = (N, T, P, A)$ une grammaire à structure de phrase. Si $(u, v) \in P$ implique $u \in N$ ou $\ell_T(v) \geq 1$, alors $L(G)$ est un C-langage.

Démonstration En appliquant le lemme II.1 à la grammaire G , nous obtenons $N' \supset N$, $T' \supset T$, $P'_0 \subset N' \times N'^+$, $V' = N' \cup T'$, $P''_0 \subset N'N'^0 \times T'V'$, $P' = P_N \cup P'_0 \cup P''_0$ et un homomorphisme h de T'^* dans T^* tels que la grammaire $G' = (N', T', P', A)$ vérifie $h(L(G')) = L(G)$.

$P'_{N'} = \{(u, v) \in P' / \ell_{T'}(v) = 0\} = P_N \cup P'_0$ et comme par hypothèse P_N est inclus dans $N \times N^*$, $P'_{N'}$ est inclus dans $N' \times N'^*$. De plus $P'_{T'} = \{(u, v) \in P' / \ell_{T'}(v) \geq 1\} = P''_0 \subset N'N'^0 \times T'V'^*$, ce qui nous permet d'appliquer le lemme II.2 à la grammaire G' .

Il existe donc $N'' \supset N'$, $T'' = \{a_i, a'_i / 1 \leq i \leq n\}$ avec $n = |P'_{T'}| + |T'|$, $P''_{T''} = \{\pi_i / 1 \leq i \leq n\}$ avec $\pi_i \in N'N'^0 \times \{a_i, a'_i\} N'^*$ pour $i \in [1, n]$ et une application séquentielle f_S de $P(T'^*)$ dans $P(T^*)$ tels que la grammaire $G'' = (N'', T'', P''_{N''} \cup P''_{T''}, A)$ vérifie $f_S(L(G'')) = L(G')$.

En appliquant le lemme II.3 à toutes les règles de $P''_{T''} \cap N''^2 \times T''^2 N''^*$, nous obtenons enfin un K-langage R et une grammaire $G''' = (N'', T'', P'_{N''} \cup P''_{N''}, A)$ tels que $P''' \subset N'' \times T'' N''^*$ et $L(G'') = L(G''') \cap R$. Comme $P'_{N''}$ est inclus dans $N'' \times N''^*$, G''' est une C-grammaire, donc $L(G''')$, $L(G'') = L(G''') \cap R$, $L(G') = f_S(L(G''))$ et $L(G) = h(L(G'))$ sont des C-langages.

Pour toute partie N_2 de N , notons, comme au chapitre I, $N_1 = N \setminus N_2$,
 $V_2 = N_2 \cup T$, $V = N \cup T$.

Posons $N' = N_2 \cup N_2^0 \times N_1 \times N_2^0$, $V' = N' \cup T$ et définissons une fonction H de V^* dans $P(V'^*)$ par

$$\begin{cases} - H(\Lambda) = \{\Lambda\} ; H(tx) = \{t\} H(x), \forall t \in V_2, \forall x \in V^* \\ - H(A_1 x) = \{A_2(A_2, A_1, B_2) y / A_2, B_2 \in N_2^0, B_2 y \in H(x)\}, \forall A_1 \in N_1, \forall x \in V^* \end{cases}$$

Définissons aussi une fonction Q de V'^* dans $P(V^*)$ par

$$\begin{cases} - Q(t) = \{t\}, \forall t \in V_2^0 ; Q(xy) = Q(x) Q(y) \forall x, y \in V' \\ - Q((\Lambda, A_1, B_2)) = \{A_1 B_2\} \text{ pour } A_1 \in N_1 \text{ et } B_2 \in N_2^0 \\ - Q((A_2, A_1, B_2)) = \{x \in V^* / A_1 B_2 \xrightarrow{\star} A_2 x\} \text{ pour } A_1 \in N_1, A_2 \in N_2, B_2 \in N_2^0 \end{cases}$$

La fonction H est introduite pour construire une nouvelle grammaire G' équivalente à G et la fonction Q va nous servir à montrer l'équivalence entre ces deux grammaires.

Montrons d'abord quelques propriétés de la fonction H .

Lemme II.4 H vérifie les quatre propriétés suivantes : pour tout $y \in V^*$

$$\begin{cases} (i) H(xy) = \{t_1 t_2 / t_1 \in H(x A_2), A_2 t_2 \in H(y), A_2 \in N_2^0\} \text{ pour tout } x \in V^* \\ (ii) H(xy) = H(x) H(y) \text{ pour tout } x \in V^* V_2 \\ (iii) H(xay) = H(x) \{a\} H(y) \text{ pour tout } a \in T \text{ et tout } x \in V^* \\ (iv) H(x A_1 y) = \{t_1(A_2, A_1, B_2) t_2 / A_2, B_2 \in N_2^0, t_1 \in H(x A_2), B_2 t_2 \in H(y)\} \\ \text{pour tout } A_1 \in N_1 \text{ et tout } x \in V^* \end{cases}$$

Démonstration

- Montrons (i) par récurrence sur la longueur de x .

Posons $R(x, y) = \{t_1 t_2 / t_1 \in H(x A_2), A_2 t_2 \in H(y), A_2 \in N_2^0\}$.

Si $x = \Lambda$, $t_1 \in H(A_2) = \{A_2\}$ donc $t_1 = A_2$

et $R(\Lambda, y) = \{A_2 t_2 / A_2 t_2 \in H(y), A_2 \in N_2^0\} = H(y)$.

Supposons que (i) soit vérifiée pour $\ell(x) \leq p$. Considérons x de longueur $p+1$ et distinguons deux cas :

(1) $x = Dx'$ avec $D \in V_2$

$H(xy) = \{D\} H(x'y) = \{D\} R(x',y)$ (hypothèse de récurrence).

Or $D \in V_2$ entraîne $R(Dx',y) = \{D\} R(x',y)$ donc $R(x,y) = H(xy)$.

(2) $x = A_1x'$ avec $A_1 \in N_1$.

Soit $z \in H(A_1x'y)$. Alors, $z = A_2(A_2, A_1, B_2) t$ avec $A_2, B_2 \in N_2^0$ et

$B_2 t \in H(x'y) = R(x',y)$ car $\ell(x') \leq p$. Donc $t = t_1' t_2'$ avec $B_2 t_1' \in H(x'C_2)$ et $C_2 t_2' \in H(y)$, C_2 appartenant à N_2^0 . Posons $t_1 = A_2(A_2, A_1, B_2) t_1'$. Par définition de H et R , $t_1 \in H(A_1x'C_2)$ et $z = t_1 t_2' \in R(A_1x',y)$.

Réciproquement, soit $z \in R(A_1x',y)$. Alors, $z = t_1 t_2$ avec $t_1 \in H(A_1x'A_2)$, $A_2 t_2 \in H(y)$, $A_2 \in N_2^0$. $t_1 \in H(A_1x'A_2)$ implique $t_1 = B_2(B_2, A_1, C_2) y$ où B_2 et C_2 appartiennent à N_2^0 et $C_2 y \in H(x'A_2)$. Nous en déduisons que $C_2 y t_2 \in R(x',y)$ et en utilisant l'hypothèse de récurrence, nous obtenons $C_2 y t_2 \in H(x'y)$ et $z = B_2(B_2, A_1, C_2) y t_2 \in H(A_1x',y)$.

- Montrons (ii)

$R(x,y) = H(x) H(y) \cup \{t_1 t_2 / t_1 \in H(xA_2), A_2 t_2 \in H(y), A_2 \in N_2\}$

donc (i) implique $H(xy) \supseteq H(x) H(y)$.

Posons $x = x'C$ avec $C \in V_2$ et soit $z \in H(x'Cy) = R(x',Cy)$. Il existe $A_2 \in N_2^0$, $t_1 \in H(x'A_2)$, $A_2 t_2 \in H(Cy)$ tels que $z = t_1 t_2$. Or $H(Cy) = \{C\} H(y)$ ce qui entraîne

. soit $C = A_2$, $t_2 \in H(y)$ et $t_1 \in H(x'C)$ donc $z \in H(x) H(y)$

. soit $A_2 = \Lambda$, $t_1 \in H(x')$, $t_2 \in H(Cy) = \{C\} H(y)$ donc $t_2 = Ct_2'$ avec $t_2' \in H(y)$.

Alors, $t_1 C \in H(x') \{C\} \subseteq H(x'C)$ ce qui entraîne que $z = t_1 C t_2' \in H(x) H(y)$.

- Montrons (iii)

D'après (ii) $H(xay) = H(xa) H(y)$ et comme $H(xa) \subseteq H(x) \{a\}$ il nous reste à montrer que $H(xa) \subseteq H(x) \{a\}$. Soit $z \in H(xa) = R(x,a)$. Alors, $z = t_1 t_2$ avec

$t_1 \in H(xA_2)$, $A_2 t_2 \in H(a) = \{a\}$ et $A_2 \in N_2^0$.

$A_2 t_2 = a$ implique $A_2 = \Lambda$ et $t_2 = a$ car $A_2 \notin T$. Donc $t_1 \in H(x)$ et $z = t_1 a \in H(x) \{a\}$.

- Montrons (iv)

Pour tout $A_1 \in N_1$, posons $S(x, A_1, y) = \{t_1(A_2, A_1, B_2) t_2 / A_2, B_2 \in N_2^0,$

$t_1 \in H(xA_2), B_2 t_2 \in H(y)\}$. Soit $z \in H(xA_1y) = R(x, A_1y)$. Alors, $z = t_1' t_2'$ avec

$t_1' \in H(xA_2)$, $A_2 t_2' \in H(A_1y)$ et $A_2 \in N_2^0$.

$A_2 t_2' \in H(A_1y)$ implique $A_2 t_2' = A_2'(A_2', A_1, B_2') s$ avec $A_2', B_2' \in N_2^0$ et $B_2' s \in H(y)$.

Distinguons deux cas :

(1) $A_2 = \Lambda$. Alors $t'_2 = A'_2(A'_2, A_1, B'_2)$ s, $t'_1 A'_2 \in H(x) \{A'_2\} \subseteq H(xA'_2)$ ce qui entraîne que $z = t'_1 A'_2(A'_2, A_1, B'_2)$ s $\in S(x, A_1, y)$.

(2) $A_2 \in N_2$. Alors $A'_2 = A_2$ et $z = t'_1(A_2, A_1, B'_2)$ s $\in S(x, A_1, y)$.

Réciproquement, soit $z \in S(x, A_1, y)$. z peut s'écrire $t_1(A_2, A_1, B_2) t_2$ avec $A_2, B_2 \in N_2^0$, $T_1 \in H(xA_2)$ et $B_2 t_2 \in H(y)$ ce qui implique $A_2(A_2, A_1, B_2) t_2 \in H(A_1 y)$ et $Z = t_1(A_2, A_1, B_2) t_2 \in R(x, A_1 y) = H(xA_1 y)$.

Lemme II.5 Soient $x, z \in V^*$ et $y \in V'^*$. Si $y \in H(x)$ et $z \in Q(y)$ alors $x \xrightarrow[G]{*} z$.

Démonstration Si $x = \Lambda$, alors $z = y = \Lambda$. Supposons que le lemme soit vérifié pour tout x de longueur $\leq p$. Considérons x de longueur $p+1$ et distinguons deux cas :

(1) $x = Cx'$ avec $C \in V_2$. Alors, $H(Cx') = \{C\} H(x')$ et $y = Cy'$ avec $y' \in H(x')$, $Q(Cy') = \{C\} Q(y')$ et $z = Cz'$ avec $z' \in Q(y')$. Comme $\ell(x') \leq p$, $x' \xrightarrow[G]{*} z'$ par hypothèse de récurrence ; donc $x = Cx' \xrightarrow[G]{*} Cz' = z$.

(2) $x = A_1 x'$ avec $A_1 \in N_1$.

Alors, $y \in H(A_1 x')$ peut se factoriser en $A_2(A_2, A_1, B_2) t$ avec $A_2, B_2 \in N_2^0$ et $B_2 t \in H(x')$. De même z qui appartient à $Q(A_2(A_2, A_1, B_2)t)$ peut s'écrire $A_2 z_1 z_2$ avec $z_1 \in Q((A_2, A_1, B_2))$ et $z_2 \in Q(t)$. Donc $B_2 z_2 \in Q(B_2 t)$ et comme $B_2 t \in H(x')$, l'hypothèse de récurrence entraîne $x' \xrightarrow[G]{*} B_2 z_2$. Or, si $A_2 = \Lambda$,

$Q((A_2, A_1, B_2)) = \{A_1 B_2\}$ et $z_1 = A_1 B_2$, sinon $A_2 \in N_2$ et $A_1 B_2 \xrightarrow[G]{*} A_2 z_1$.

Dans les deux cas on obtient $x = A_1 x' \xrightarrow[G]{*} A_1 B_2 z_2 \xrightarrow[G]{*} A_2 z_1 z_2 = z$.

Ces résultats sur les fonctions H et Q nous permettent de démontrer le lemme de base de ce chapitre. Posons : $P_{N_1} = \{(u, v) \in P_N \mid u \in N_1\}$ et $\bar{P}_{N_1} = P_N \setminus P_{N_1}$.

Lemme II.6 Si P_T est inclus dans $N \times TV^* \cup N_2 N \times TV^* N_2$ et si $\bar{P}_{N_1}$ est monotone de type 2 en N_2 , alors $L(G)$ est un C-langage.

Démonstration D'après le lemme I.4, nous pouvons supposer, sans nuire à la généralité de la démonstration, que $\bar{P}_{N_1}$ est monotone standard, donc que $\bar{P}_{N_1} \subset N_1 N_2 \times N_2 \cup N_2 \times N_2^0 N_2$.

Nous supposons aussi la propriété suivante :

- (α) $(u, v) \in P_T \cap N_2 \times TV^*$ implique $(uA_2, vA_2) \in P_T$ pour tout $A_2 \in N_2$.

En effet dans le cas contraire, il suffit de considérer la grammaire $G'' = (N, T, P'', A)$ avec $P'' = P \cup P_0''$ où

$$P_0'' = \{(uA_2, vA_2) / (u, v) \in P_T \cap N_2 \times TV^* \text{ et } A_2 \in N_2\}.$$

Comme $(u'', v'') \in P_0''$ entraîne $u'' \xrightarrow{G} v''$ et $(u'', v'') \in N_2 N \times TV^* N_2$, G'' est équivalente à G et vérifie les hypothèses du lemme ainsi que la propriété (α).

Considérons maintenant la grammaire $G' = (N' \cup \{A'\}, T, P', A')$ obtenue à partir de G de la manière suivante :

$$N' = N_2 \cup N_2^0 \times N_1 \times N_2^0, \quad V' = N' \cup T, \quad A' \notin V'$$

$$\bar{P}' = \bigcup_{j=1}^5 P_j' \quad \text{avec :}$$

- $P_1' = \{(s, u) / (s, x) \in P, u \in H(x), s \in N_2 N_2^0\}$
 - $P_2' = \{((A_2, E_1, B_2), u) / A_2 \in N_2^0, E_1 \in N_1, B_2 \in N_2, A_2 u \in H(E_1 x), (B_2, x) \in \bar{P}_{N_1}\}$
 - $P_3' = \{((A_2, E_1, B_2), u) / A_2, B_2 \in N_2^0, E_1 \in N_1, A_2 u \in H(x B_2), (E_1, x) \in P\}$
 - $P_4' = \{((A_2, E_1, B_2), \Lambda) / (E_1 B_2, A_2) \in P; A_2, B_2 \in N_2, E_1 \in N_1\}$
 - $P_5' = \{(A_2(\Lambda, E_1, \Lambda), u) / A_2 \in N_2, E_1 \in N_1, (A_2 E_1, x) \in P_T, u \in H(x)\}$
- et $P' = \bar{P}' \cup \{(A', u) / u \in H(A)\}.$

Remarquons que $(u', v') \in P_2' \cup P_3' \cup P_4'$ entraîne $\ell(u') = 1$.

Si $(u', v') \in P_5'$, il existe x tel que $v' \in H(x)$ et $\ell_T(x) \geq 1$, donc $\ell_T(v') \geq 1$ (propriété (iii) du lemme II.4).

De même, si $(u', v') \in P_1'$ et si $\ell(u') > 1$, alors $u' \in N_2^2$ et il existe x tel que $(u', x) \in P$ et $v' \in H(x)$. Or $(u', x) \notin P_{N_1}$ car $u' \notin N_1$ et $(u', x) \notin \bar{P}_{N_1}$ car $u' \notin N_1 N_2 \cup N_2$. Donc $(u', x) \in P_T$ et $\ell_T(x) \geq 1$, ce qui entraîne $\ell_T(v') \geq 1$.

Donc $(u', v') \in P'$ implique $u' \in N'$ ou $\ell_T(v') \geq 1$ et la proposition II.1 permet d'affirmer que $L(G)$ est un C-langage.

Il nous reste à montrer que G et G' sont des grammaires équivalentes. Pour cela, nous allons nous servir des deux propriétés suivantes :

(β) Si $(s, u) \in \bar{P}'$ et $y \in Q(u)$ alors il existe $z \in Q(s)$ tel que $z \xrightarrow[G]{*} y$.

(γ) Si $u \xrightarrow[G]{*} v$ et $z \in H(v)$ alors il existe $y \in H(u)$ tel que $y \xrightarrow[G]{*} z$.

- Montrons d'abord (β)

Prenons $y \in Q(u)$ et distinguons cinq cas :

(1) $(s, u) \in P'_1$: alors, il existe x tel que $(s, x) \in P$ et $u \in H(x)$. Le lemme II.5 implique $x \xrightarrow[G]{*} y$ et en prenant $z = s \in Q(s) = \{s\}$ car $s \in N_2 N_2^0$, nous obtenons $z \xrightarrow[G]{*} x \xrightarrow[G]{*} y$.

(2) $(s, u) \in P'_2$: alors, $s = (A_2, E_1, B_2)$ et il existe x tel que $(B_2, x) \in \bar{P}_{N_1}$ et $A_2 u \in H(E_1 x)$. Comme $A_2 y \in Q(A_2 u)$, $E_1 x \xrightarrow[G]{*} A_2 y$ (lemme II.5) et $E_1 B_2 \xrightarrow[G]{*} E_1 x \xrightarrow[G]{*} A_2 y$.

Si $A_2 = \Lambda$, $Q(s) = \{E_1 B_2\}$ et $z = E_1 B_2 \xrightarrow[G]{*} y$.

Si $A_2 \in N_2$, il suffit de prendre $z = y$ car $E_1 B_2 \xrightarrow[G]{*} A_2 y$ entraîne $y \in Q(s)$.

(3) $(s, u) \in P'_3$: alors, $s = (A_2, E_1, B_2)$ et il existe x tel que $A_2 u \in H(x B_2)$ et $(E_1, x) \in P$. Comme $A_2 y \in \{A_2\} Q(u) = Q(A_2 u)$, $x B_2 \xrightarrow[G]{*} A_2 y$ (lemme II.5) donc $E_1 B_2 \xrightarrow[G]{*} x B_2 \xrightarrow[G]{*} A_2 y$.

Si $A_2 = \Lambda$, $Q(s) = \{E_1 B_2\}$ et $z = E_1 B_2 \xrightarrow[G]{*} y$.

Si $A_2 \in N_2$, $y \in Q(s)$ et il suffit de prendre $z = y$.

(4) $(s, u) \in P'_4$: alors, $s = (A_2, E_1, B_2)$ avec $(E_1 B_2, A_2) \in P$ et $u = \Lambda$; donc $y = \Lambda$. Comme $E_1 B_2 \xrightarrow[G]{*} A_2$ et que $A_2 \neq \Lambda$, $\Lambda \in Q(s)$.

Il suffit alors de prendre $z = y = \Lambda$.

(5) $(s, u) \in P'_5$: alors $s = A_2(\Lambda, E_1, \Lambda)$ et il existe x tel que $(A_2 E_1, x) \in P_T$ et $u \in H(x)$. Donc, d'après le lemme II.5, $x \xrightarrow[G]{*} y$, donc $A_2 E_1 \xrightarrow[G]{*} x \xrightarrow[G]{*} y$. Comme $Q(s) = \{A_2 E_1\}$, il suffit de prendre $z = A_2 E_1$.

- Montrons maintenant la propriété (γ).

Posons $u = u_1 s u_2$, $v = u_1 x u_2$ avec $(s, x) \in P$. Soit $z \in H(u_1 x u_2)$. Distinguons encore cinq cas :

- (a) $s \in N_2^0$: alors d'après les hypothèses $x \in TV^*N_2$. Dans ce cas $H(u_1 x u_2) = H(u_1 x) H(u_2)$ (propriété (ii) du lemme II.4) donc $z = z_1 z_2$ avec $z_1 \in H(u_1 x)$ et $z_2 \in H(u_2)$.

Posons $x = ax'$ avec $a \in T$. Alors z_1 se factorise en $z'_1 a z''_1$ avec $z'_1 \in H(u_1)$ et $z''_1 \in H(x')$.

Posons $y = z'_1 s z_2$. Comme $H(s) = \{s\}$, $z'_1 s \in H(u_1 s)$ donc $y \in H(u_1 s) H(u_2) = H(u_1 s u_2) = H(u)$ (propriété (ii) du lemme II.4). Enfin, comme $(s, az''_1) \in P'_1$, $y = z'_1 s z_2 \xrightarrow{\overline{G}} z'_1 a z''_1 z_2 = z$.

- (b) $s \in N_2$: alors d'après les hypothèses $x \in TV^* \cup N^0 N_2$.
Distinguons trois cas :

- $x \in N^0 N_2$ et $u_1 \in V^* V_2 \cup \{\Lambda\}$: alors d'après la propriété (ii) du lemme II.4, $H(u_1 x u_2) = H(u_1 x) H(u_2) = H(u_1) H(x) H(u_2)$ et $z = z'_1 z''_1 z_2$ avec $z'_1 \in H(u_1)$, $z''_1 \in H(x)$ et $z_2 \in H(u_2)$. Or $H(s) = \{s\}$ et $(s, z''_1) \in P'_1$ donc $y = z'_1 s z_2 \in H(u)$ et $y \xrightarrow{\overline{G}} z$.

- $x \in N^0 N_2$ et $u_1 \in V^* N_1$: alors $H(u_1 x u_2) = H(u_1 x) H(u_2)$ et $z = z_1 z_2$ avec $z_1 \in H(u_1 x)$ et $z_2 \in H(u_2)$. Posons $u_1 = u'_1 E_1$ avec $E_1 \in N_1$. D'après la propriété (i) du lemme II.4, il existe $A_2 \in N_2^0$ tel que $z_1 = z'_1 z''_1$ avec $z'_1 \in H(u'_1 A_2)$ et $A_2 z''_1 \in H(E_1 x)$. Comme $x \in N^0 N_2$ et $s \in N_2$, $(s, x) \in \bar{P}_{N_1}$, donc $((A_2, E_1, s), z''_1) \in P'_2$. De plus, comme $H(s) = \{s\}$, la propriété (iv) du lemme II.4 implique que $z'_1 (A_2, E_1, s) \in H(u'_1 E_1 s)$. Alors, en posant $y = z'_1 (A_2, E_1, s) z_2$, nous obtenons $y \in H(u'_1 E_1 s) H(u_2) = H(u)$ et $y \xrightarrow{\overline{G}} z'_1 z''_1 z_2 = z$.

- $x \in TV^*$: Posons $x = ax'$ avec $a \in T$. Alors d'après la propriété (iii) du lemme II.4, $z = z_1 a z_2$ avec $z_1 \in H(u_1)$ et $z_2 \in H(x' u_2)$; d'après la propriété (i) du lemme II.4, $\exists A_2 \in N_2^0$, $z' \in H(x' A_2)$ et z'_2 tels que $A_2 z'_2 \in H(u_2)$ et $z_2 = z' z'_2$. Or d'après la propriété (α), $(s A_2, a x' A_2) \in P$. Donc, comme $s A_2 \in N_2 N_2^0$, $(s A_2, a z') \in P'_1$. Posons $y = z_1 s A_2 z'_2$. Alors $y \in H(u_1) \{s\} H(u_2) \leq H(u_1 s u_2)$ et $y \xrightarrow{\overline{G}} z_1 a z' z'_2 = z$.

(c) $s \in N_1$: D'après la propriété (i) du lemme II.4, il existe $B_2 \in N_2^0$, $z_1 \in H(u_1 x B_2)$ et $B_2 z_2 \in H(u_2)$ tels que $z = z_1 z_2$. En appliquant cette propriété une nouvelle fois pour $H(u_1 x B_2)$, nous obtenons $z_1 = z'_1 z''_1$ avec $A_2 \in N_2^0$, $z'_1 \in H(u_1 A_2)$ et $A_2 z''_1 \in H(x B_2)$. Donc $((A_2, s, B_2), z''_1) \in P'_3$ et en posant $y = z'_1(A_2, s, B_2) z_2$, nous obtenons $y \in H(u_1 s u_2)$ (propriété (iv) du lemme II.4) et $y \xrightarrow{G'} z$.

(d) $s \in N_1 N_2$: donc $(s, x) \in \bar{P}_{N_1}$ et $x \in N_2$. Alors d'après la propriété (ii) du lemme II.4, z peut se factoriser en $z_1 z_2$ avec $z_1 \in H(u_1 x)$ et $z_2 \in H(u_2)$. Posons $s = s_1 s_2$ avec $s_1 \in N_1$ et $s_2 \in N_2$. Alors, $((x, s_1, s_2), \Lambda) \in P'_4$ et d'après la propriété (iv) du lemme II.4, $y = z_1(x, s_1, s_2) z_2 \in H(u_1 s_1 s_2 u_2)$ et $y \xrightarrow{G'} z_1 z_2 = z$.

(e) $s \in N_2 N_1$: donc $(s, x) \in P_T$ et $x \in TV^* N_2$. Posons $s = s_2 s_1$ et $x = ax' A_2$ avec $s_2, A_2 \in N_2$, $s_1 \in N_1$ et $a \in T$. z appartenant à $H(u_1 ax' A_2 u_2)$ peut se factoriser en $z_1 az' z_2$ avec $z_1 \in H(u_1)$, $z' \in H(x' A_2)$ et $z_2 \in H(u_2)$. Comme $az' \in H(ax' A_2)$, $(s_2(\Lambda, s_1, \Lambda), az') \in P'_5$. Or $(\Lambda, s_1, \Lambda) \in H(s_1)$ ce qui implique que $y = z_1 s_2(\Lambda, s_1, \Lambda) z_2 \in H(u_1 s_2 s_1 u_2)$ et $y \xrightarrow{G'} z_1 az' z_2 = z$.

Comme nous avons considéré tous les cas possibles, la propriété (γ) est vérifiée.

Montrons maintenant que $L(G) = L(G')$.

Soit $x \in L(G)$. Il existe une dérivation $A \xrightarrow{*}_G x$. La propriété (γ) entraîne immédiatement que pour tout $z \in H(x)$, il existe $y \in H(A)$ tel que $y \xrightarrow{*}_{G'} z$. Or $H(x) = \{x\}$ car $x \in T^*$. De plus $(A', y) \in P'$ car $y \in H(A)$, donc $A' \xrightarrow{*}_{G'} y \xrightarrow{*}_{G'} x$ et $x \in L(G')$.

Réciproquement, si $x \in L(G')$, il existe une dérivation $A' \xrightarrow{*}_{G'} y_0 \xrightarrow{*}_{G'} x$. Or $Q(x) = \{x\}$ car $x \in T^*$. D'après la propriété (β) , et comme les règles utilisées dans la dérivation $y_0 \xrightarrow{*}_{G'} x$, appartiennent à $\bar{P}'$, $\exists z_0 \in Q(y_0)$ tel que $z_0 \xrightarrow{*}_G x$. Comme $y_0 \in H(A)$, y_0 peut se factoriser en $A_2(A_2, A, \Lambda)$ avec $A_2 \in N_2^0$. Si $A_2 = \Lambda$, $z_0 = A$, sinon $z_0 = A_2 z'_0$ avec $A \xrightarrow{*}_G A_2 z'_0$. Dans les deux cas, nous obtenons $A \xrightarrow{*}_G z_0 \xrightarrow{*}_G x$ et $x \in L(G)$.

Nous pouvons maintenant démontrer le résultat le plus général de ce chapitre :

Proposition II.2 Soit $G = (N, T, P, A)$ une grammaire à structure de phrase. S'il existe une partition de N en N_1, N_2 telle que l'ensemble $\bar{P}_{N_1} = \{(u, v) \in P / u \notin N_1 \text{ et } \ell_T(v) = 0\}$ soit monotone de type 2 en N_2 , alors $L(G)$ est un C-langage.

Démonstration D'après le lemme II.1, il existe $N' \supset N, N'' = N' \setminus N, T' \supset T, P'_0 \subset N \times N^*N'', P''_0 \subset N''N' \times T'N'' \cup N' \times TV^*$, un homomorphisme h de T'^* dans T^* et une grammaire $G' = (N', T', P_N \cup P'_0 \cup P''_0, A)$ tels que $h(L(G')) = L(G)$.

Posons $P' = P_N \cup P'_0 \cup P''_0, N'_2 = N_2 \cup N''$ et $V' = N' \cup T'$. Alors $P'_{T'} = P''_0 \subset N' \times T'V'^* \cup N'_2N' \times T'V'^* \subset N'_2$. En outre $\bar{P}'_{N_1} = \{(u, v) \in P'_{N'} / u \notin N_1\} = \bar{P}_{N_1} \cup (P'_0 \cap N'_2 \times N'^*N'_2)$ est monotone de type 2 en N'_2 car $\bar{P}_{N_1}$ est, par hypothèse, monotone de type 2 en $N_2 \subset N'_2$.

Alors, d'après le lemme II.6, $L(G')$ est un C-langage, donc $L(G) = h(L(G'))$ aussi.

Comme P_N contient $\bar{P}_{N_1}$ pour toute partie N_1 de N , P_N monotone de type 2 en N_2 implique $\bar{P}_{N_1}$ monotone de type 2 en N_2 . Nous en déduisons le résultat suivant :

Corollaire II.1 Soit $G = (N, T, P, A)$ une grammaire à structure de phrase. Si l'ensemble des productions de P ne contenant pas de terminaux est monotone de type 2, alors $L(G)$ est un C-langage.

A partir de la proposition II.2, nous obtenons un autre résultat :

Corollaire II.2 Un langage L est un C-langage si et seulement s'il existe une grammaire monotone de type 2 qui engendre $L \setminus \{\Lambda\}$.

Démonstration Soit L un C-langage inclus dans T^* . Il existe une C-grammaire Λ -libre $G = (N, T, P, A)$ qui engendre $L \setminus \{\Lambda\}$. Alors $(u, v) \in P$ implique $v \neq \Lambda$ donc $M_{\emptyset, N}^2(u, v)$. Et G est monotone de type 2 en N .

Réciproquement si $L \setminus \{\Lambda\}$ est engendrée par une grammaire $G = (N, T, P, A)$ monotone de type 2. $P_N \subset P$ est monotone de type 2 et d'après le corollaire II.1, $L \setminus \{\Lambda\}$ est un C-langage donc L aussi.

Chapitre III

GRAMMAIRES MONOTONES et LANGAGES CONTEXTUELS

Le premier résultat de ce chapitre est une relation entre les classes C_4 et C_5 des langages engendrés par les grammaires monotones respectivement de type 4 et 5 : L appartient à C_5 si et seulement si $\tilde{L} = \{\tilde{x} / x \in L\}$ appartient à C_4 . Nous montrons ensuite la récursivité des éléments de C_4 et donc celle des langages appartenant à C_5 . Puis, en reprenant un argument utilisé par Chomsky [CH page 143, note 7^a] pour les langages contextuels, nous établissons l'existence de langages récursifs qui n'appartiennent ni à C_4 ni à C_5 . Nous démontrons, alors, que tout langage engendré par une grammaire monotone de type 1 est un langage contextuel, et nous généralisons ce résultat, en montrant qu'il suffit de considérer les règles où ne figure aucun terminal.

Introduisons une nouvelle classe de langages dont nous montrerons l'équivalence à C_4 . Soit $G = (N, T, P, A)$ une grammaire de type 0 en N_2 .

Nous dirons que G est une grammaire monotone de type 4 bis en N_2 si et seulement si P est inclus dans $N^*N_2 \times V^*V_2$.

Par définition toute grammaire monotone de type 4 bis est monotone de type 4. Réciproquement, montrons que pour tout langage L de C_4 , il existe une grammaire monotone de type 4 bis qui engendre L . Pour obtenir ce résultat nous allons d'abord montrer deux lemmes relatifs à une grammaire $G = (N, T, P, A)$ monotone de type 0 en N_2 .

Lemme III.1 Si $P \setminus (P \cap N_2^2 \times N_2^2)$ est monotone standard de type 4 en N_2 , alors pour toute dérivation $(d): xE_1y \xrightarrow{*} W$ avec $E_1 \in N_1$, $W \in V_2^+$, $x \in V^*$ et $y \in V^+$, il existe $A_2, B_2 \in N_2$ et $Z \in V_2^*$ tels que $y \xrightarrow{*} A_2Z$, $E_1A_2 \Rightarrow B_2$ et $xB_2Z \xrightarrow{*} W$, en utilisant les mêmes règles que pour la dérivation (d) .

Démonstration Comme $W \in V_2^+$, il existe une règle qui fait disparaître l'occurrence de E_1 . D'après les hypothèses, il existe donc $A_2, B_2 \in N_2$ tels que $(E_1 A_2, B_2) \in P$ et (d) peut s'écrire :

$$(d') \quad x E_1 y \xrightarrow{*} x' E_1 A_2 y' \implies x' B_2 y' \xrightarrow{*} W$$

avec $x \xrightarrow{*} x'$ et $y \xrightarrow{*} A_2 y'$.

Soit k la longueur de la dérivation (d').

Si $k = 1$, $x = x'$, $y = A_2 y'$ et $x B_2 y' = W \in V_2^+$ donc $y' \in V_2^*$. Il suffit de prendre $Z = y'$.

Supposons la propriété vraie pour $k \leq t$ et considérons une dérivation (d') de longueur $t+1$.

Si $y' \in V_2^*$, il suffit de prendre $Z = y'$. Dans le cas contraire y' peut se factoriser en $y'_1 E'_1 y'_2$ avec $E'_1 \in N_1$ et $y'_1 \in V_2^*$. La dérivation $x' B_2 y'_1 E'_1 y'_2 \xrightarrow{*} W$ est de longueur inférieure ou égale à t puisque la règle $(E_1 A_2, B_2)$ n'y est pas utilisée. Donc l'hypothèse de récurrence implique l'existence de $A'_2, B'_2 \in N_2$ et de $z' \in V_2^*$ vérifiant $y'_2 \xrightarrow{*} A'_2 z'$, $E'_1 A'_2 \implies B'_2$ et $x B_2 y'_1 B'_2 z' \xrightarrow{*} W$. Posons $Z = y'_1 B'_2 z'$. Nous obtenons les dérivations : $y \xrightarrow{*} A_2 y' = A_2 y'_1 E'_1 y'_2 \xrightarrow{*} A_2 y'_1 E'_1 A'_2 z' \implies A_2 y'_1 B'_2 z' = A_2 Z$ et $x B_2 Z = x B_2 y'_1 B'_2 z' \xrightarrow{*} W$.

Comme $Z \in V_2^*$, la propriété est bien vérifiée.

Lemme III.2 Si P est monotone standard de type 4 en N_2 , il existe une partie P_0 de $N_2^2 \times N_2^2$ telle que $G' = (N, T, P', A)$ soit équivalente à G , avec

$$P' = (P \setminus (P \cap N_2 \times N_2 N_1)) \cup P_0.$$

Démonstration Posons $P_0 = \{(A_2 A'_2, B_2 B'_2) \in N_2^2 \times N_2^2 / (A_2, B_2 E_1), (E_1 A'_2, B'_2) \in P, E_1 \in N_1\}$
 $P' = (P \setminus (P \cap N_2 \times N_2 N_1)) \cup P_0$, $P'' = P \cup P_0$ $G' = (N, T, P', A)$ et $G'' = (N, T, P'', A)$.

Comme toute règle de P_0 est une dérivation dans G et que P'' contient P , $L(G'') = L(G)$. De même, comme P' est inclus dans P'' , $L(G') \subset L(G'')$. Pour montrer l'inclusion inverse, nous utiliserons la propriété suivante :

$$(a) \quad W \xrightarrow[G]{*} W' \text{ avec } W, W' \in V_2^+ \text{ implique } W \xrightarrow[G]{*} W'.$$

Démontrons (a) par récurrence sur la longueur k de la dérivation dans G'' .

Si $k = 0$, $W' = W$ et $W \xrightarrow[G]{*} W$.

Supposons la propriété (a) vraie pour $k \leq t$ et considérons une dérivation de longueur $t+1$.

Si aucune règle de $P \cap N_2 \times N_2 N_1$ n'a été utilisée dans cette dérivation, il est évident que $W \xrightarrow[G']{*} W'$. Dans le cas contraire, exhibons la première utilisation d'une règle appartenant à $P \cap N_2 \times N_2 N_1$. Nous obtenons dans G'' la dérivation suivante : $W \xrightarrow{*} x'A_2'y \Rightarrow x'B_2'E_1y \xrightarrow{*} W'$ avec $A_2', B_2' \in N_2$, $E_1 \in N_1$ et $(A_2', B_2'E_1) \in P$.

Comme $P'' \setminus (P'' \cap N_2^2 \times N_2^2) = P$ est monotone standard de type 4 en N_2 , nous pouvons utiliser le lemme III.1. Il existe donc $A_2, B_2 \in N_2$, $z \in V_2^*$ tels que $E_1A_2 \xrightarrow[G'']{*} B_2$, $y \xrightarrow[G'']{*} A_2z$ et $x'B_2'B_2z \xrightarrow[G'']{*} W'$. Comme $(A_2', B_2'E_1)$ et (E_1A_2, B_2) sont des règles de P , $(A_2'A_2, B_2'B_2)$ appartient à P_0 et nous obtenons dans G'' une dérivation de longueur t :

$$W \xrightarrow{*} x'A_2'y \xrightarrow{*} x'A_2'A_2z \Rightarrow x'B_2'B_2z \xrightarrow{*} W'.$$

L'hypothèse de récurrence permet de conclure que W se dérive en W' dans G' .

En prenant $W = A$ et $W' \in T^*$, la propriété (a) implique que si $W' \in L(G'')$, alors $W \in L(G')$, donc que $L(G'') \subset L(G')$. Et $L(G') = L(G'') = L(G)$.

Il est facile de vérifier que la grammaire G' obtenue ci-dessus est monotone de type 4bis. En utilisant le lemme I.4, nous obtenons :

Lemme III.3 Pour tout langage L appartenant à C_4 , il existe une grammaire monotone de type 4bis qui engendre L .

Nous allons procéder de la même façon pour les grammaires monotones de type 5.

Nous dirons qu'une grammaire $G = (N, T, P, A)$ monotone de type 0 en N_2 est monotone de type 5 bis en N_2 si et seulement si P est inclus dans $N_2 N^* \times V_2 V^*$.

Lemme III.4 Si $P \setminus (P \cap N_2^2 \times N_2^2)$ est monotone standard de type 5 en N_2 , alors pour toute dérivation $(d): W \xrightarrow{*} x E_1 y$ avec $E_1 \in N_1$, $W \in V_2^+$, $x \in V^+$ et $y \in V^*$, il existe $A_2, B_2 \in N_2$, $z \in V_2^*$ tels que $W \xrightarrow{*} z A_2 y$, $A_2 \xRightarrow{*} B_2 E_1$ et $z B_2 \xrightarrow{*} x$, en utilisant les mêmes règles que dans la dérivation (d) .

Démonstration Comme $W \in V_2^+$, il existe une règle qui fait apparaître l'occurrence de E_1 . D'après les hypothèses, il existe donc $A_2, B_2 \in N_2$ tels que $(A_2, B_2 E_1)$ appartienne à P . Et (d) peut s'écrire $(d'): W \xrightarrow{*} x' A_2 y' \xRightarrow{*} x' B_2 E_1 y' \xrightarrow{*} x E_1 y$ avec

$$x' B_2 \xrightarrow{*} x \text{ et } y' \xrightarrow{*} y.$$

Soit k la longueur de la dérivation (d') .

Si $k=1$, $W = x' A_2 y'$, $y'=y$ et $x' B_2 = x$; la propriété est donc vérifiée en prenant $z = x'$.

Supposons la propriété vérifiée pour $k \leq t$ et prenons une dérivation (d') de longueur $t+1$.

Si $x' \in V_2^*$, il suffit de prendre $z=x'$. Dans le cas contraire x' peut se factoriser en $x'_1 E'_1 x'_2$ avec $x'_2 \in V_2^*$, $E'_1 \in N_1$. La dérivation $W \xrightarrow{*} x'_1 E'_1 x'_2 A_2 y'$ est de longueur au plus égale à t ; donc l'hypothèse de récurrence entraîne l'existence de $A'_2, B'_2 \in N_2$, $z' \in V_2^*$ tels que $W \xrightarrow{*} z' A'_2 x'_2 A_2 y'$, $A'_2 \xRightarrow{*} B'_2 E'_1$ et $z' B'_2 \xrightarrow{*} x'_1$. Posons $z = z' A'_2 x'_2$. z appartient à V_2^* . De plus $W \xrightarrow{*} z' A'_2 x'_2 A_2 y' = z A_2 y' \xrightarrow{*} z A_2 y$, $z B_2 = z' A'_2 x'_2 B_2 \xRightarrow{*} z' B'_2 E'_1 x'_2 B_2 \xrightarrow{*} x'_1 E'_1 x'_2 B_2 = x' B_2 \xrightarrow{*} x$ et $A_2 \xRightarrow{*} B_2 E_1$. La propriété est donc vérifiée.

Lemme III.5 Pour toute grammaire G monotone de type 5, il existe une grammaire monotone de type 5 bis équivalente à G .

Démonstration Soit $G = (N, T, P, A)$ une grammaire monotone de type 5 en N_2 . D'après le lemme I.4, nous pouvons supposer qu'elle est standard.

Notons $P_0 = \{(A_2 A'_2, B_2 B'_2) \in N_2^2 \times N_2^2 / (A_2, B_2 E_1), (E_1 A'_2, B'_2) \in P, E_1 \in N_1\}$

$$P' = (P \setminus (P \cap N_1 N_2 \times N_2)) \cup P_0, P'' = P \cup P_0, G' = (N, T, P', A)$$

et $G'' = (N, T, P'', A)$.

Il est évident que G' est monotone de type 5 bis en N_2 .

Comme toute règle de P_0 est une dérivation dans G et que P'' contient P , $L(G'') = L(G)$. Comme P' est inclus dans P'' , $L(G') \subset L(G'')$. Pour montrer l'inclusion dans l'autre sens, établissons la propriété :

$$(a) \quad W \xrightarrow[G'']{*} W' \text{ avec } W, W' \in V_2^+ \text{ implique } W \xrightarrow[G']{*} W'.$$

Démontrons la propriété (a) par récurrence sur la longueur k de la dérivation dans G'' .

$$\text{Si } k=0, W'=W \text{ et } W \xrightarrow[G']{*} W.$$

Supposons la propriété (a) vraie pour toute dérivation de longueur inférieure ou égale à t et considérons une dérivation de longueur $t+1$.

Si aucune règle de production appartenant à $P \cap N_1 N_2 \times N_2$ n'a été utilisée dans cette dérivation, il est évident que $W \xrightarrow[G']{*} W'$. Dans le cas contraire, considérons la première utilisation d'une telle règle. Nous obtenons dans G'' la dérivation suivante : $W \xrightarrow{*} xE_1 A'_2 y' \xrightarrow{*} xB'_2 y' \xrightarrow{*} W'$ avec $A'_2, B'_2 \in N_2, E_1 \in N_1$ et $(E_1 A'_2, B'_2) \in P$. Comme $P'' \setminus (P'' \cap N_2^2 \times N_2^2) = P$ est monotone standard de type 5 en N_2 , nous pouvons utiliser le lemme III.4. Il existe donc $A_2, B_2 \in N_2, z \in V_2^*$ tels que $W \xrightarrow{*} zA_2 A'_2 y', A_2 \xrightarrow{*} B_2 E_1$ et $zB_2 \xrightarrow{*} x$. $(A_2 A'_2, B_2 B'_2)$ appartient à P_0 car $(E_1 A'_2, B'_2)$ et $(A_2, B_2 E_1)$ sont dans P . Nous obtenons, alors, dans G'' , la dérivation suivante : $W \xrightarrow{*} zA_2 A'_2 y' \xrightarrow{*} zB_2 B'_2 y' \xrightarrow{*} xB'_2 y' \xrightarrow{*} W'$.

Cette dérivation est de longueur t puisque nous avons remplacé deux règles par une seule. L'hypothèse de récurrence permet d'en déduire que $W \xrightarrow[G']{*} W'$. La propriété (a) entraîne immédiatement que $L(G'') \subset L(G')$ ce qui implique que G' , monotone de type 5 bis, est équivalente à G .

Soit $G = (N, T, P, A)$ une grammaire monotone de type 5 bis (respectivement 4 bis) en N_2 . Posons $P' = \{(\tilde{u}, \tilde{v}) / (u, v) \in P\}$ et $G' = (N, T, P', A)$. Il est facile de vérifier que $L(G') = \{\tilde{x} / x \in L\}$ et que G' est monotone de type 4 bis (respectivement 5 bis) en N_2 . Les lemmes III.3 et III.5 nous permettent d'en déduire :

Proposition III.1 L appartient à C_5 si et seulement si $\tilde{L} = \{\tilde{x} / x \in L\}$ appartient à C_4 .

Montrons maintenant la récursivité des langages engendrés par les grammaires monotones de type 4 et 5.

Pour obtenir ce résultat, à tout mot x non vide et à toute grammaire G monotone de type 4, nous allons associer un C-langage qui contient le mot vide si et seulement si $x \in L(G)$.

Proposition III.2 Pour tout mot x et tout élément L de la classe C_4 (respectivement C_5), il est décidable de déterminer si x appartient à L .

Démonstration Supposons d'abord que $L \in C_4$. D'après les lemmes I.4 et III.2, il existe une grammaire $G = (N, T, P, A)$ et une partie de N_2 de N telles que $L(G) = L$ et $P = P_1 \cup P_0$ avec P_1 monotone standard de type 2 en N_2 et $P_0 \subset N_2^2 \times N_2^2$.

Si $x = \Lambda$, $x \notin L(G)$ puisque $A \xrightarrow{*} x$ entraîne $\ell_{V_2}(x) \geq 1$.

Soit $x \in T^+$; posons $t = \ell(x)$. A x et à G , faisons correspondre la grammaire $G' = (N', T, P', A')$ définie de la manière suivante :

$$N' = \{(u_1, u_2) \in V_2^+ \times N_1^0 V_2^+ / \ell_{V_2}(u_2) \leq \ell(u_1) \leq t\}, A' = (x, A),$$

$$P' = P'_1 \cup P'_2 \cup P'_3 \text{ avec}$$

$$P'_1 = \{((u_1, y'sy''), (u_1, y'uy'')) \in N' \times N' / (s, u) \in P \setminus (P \cap N_2 \times N_1 N_2), \ell_{V_2}(y'uy'') \leq \ell(u_1)\}$$

$$P'_2 = \{((u_1, y'A_2 y''), (u_1, y'z)(z, E_1 B_2 y'')) \in N' \times N'^2 / (A_2, E_1 B_2) \in P, z \in V_2^+, \ell(u_1) - \ell_{V_2}(y') \geq \ell(z) \geq \ell(B_2 y'')\}$$

$$\text{et } P'_3 = \{(u, u), \Lambda\} / u \in V_2^+, \ell(u) \leq t\}$$

Comme P' est une partie finie de $N' \times N'^*$, G' est une C-grammaire. Montrons d'abord la propriété :

(b) Si $(u_1, u_2) \xrightarrow[G']{*} \Lambda$ avec $(u_1, u_2) \in N'$, alors $u_2 \xrightarrow[G]{*} u_1$.

Soit k la longueur de la dérivation dans G' .

Si $k=1$, $u_1 = u_2$ et donc $u_1 \xrightarrow[G]{*} u_1$.

Supposons la propriété (b) vraie pour une dérivation de longueur au plus égale à s , et considérons une dérivation de longueur $s+1 \geq 2$. Soit π la première règle utilisée. Comme $s+1 > 1$, $\pi \in P'_1 \cup P'_2$.

- Si $\pi \in P'_1$, nous pouvons écrire $\pi = ((u_1, u_2), (u_1, u'_2))$ avec $u_2 \xrightarrow[G]{*} u'_2$.

Comme la dérivation $(u_1, u'_2) \xrightarrow[G']{*} \Lambda$ est de longueur s , l'hypothèse de récurrence implique $u'_2 \xrightarrow[G]{*} u_1$ et nous obtenons $u_2 \xrightarrow[G]{*} u'_2 \xrightarrow[G]{*} u_1$.

- Si $\pi \in P'_2$, posons $\pi = ((u_1, y'A_2y''), (u_1, y'z)(z, E_1B_2y''))$ avec $(A_2, E_1B_2) \in P$ et $u_2 = y'A_2y''$. Comme la dérivation $(u_1, y'z)(z, E_1B_2y'') \xrightarrow[G']{*} \Lambda$ est de longueur s , nous pouvons appliquer l'hypothèse de récurrence aux dérivations

$(u_1, y'z) \xrightarrow[G']{*} \Lambda$ et $(z, E_1B_2y'') \xrightarrow[G']{*} \Lambda$. Et nous obtenons $y'A_2y'' \xrightarrow[G]{*} y'E_1B_2y'' \xrightarrow[G]{*} y'z \xrightarrow[G]{*} u_1$.

Montrons maintenant la propriété réciproque :

(b') Si $y \xrightarrow[G]{*} u_1$ et si $(u_1, y) \in N'$ alors $(u_1, y) \xrightarrow[G']{*} \Lambda$.

Soit k la longueur de la dérivation dans G .

Si $k=0$ $u_1=y$ et $(y, y) \xrightarrow[G']{*} \Lambda$.

Supposons (b') vraie pour $k \leq s$ et considérons une dérivation de longueur $s+1$:

$y \xrightarrow[G]{*} w_0 \xrightarrow[G]{*} u_1$. Comme G est monotone de type 4 en N_2 ,

$\ell_{V_2}(y) \leq \ell_{V_2}(w_0) \leq \ell(u_1) \leq t$ (car $(u_1, y) \in N'$).

Soit π la première règle utilisée.

- Si $\pi \notin N_2 \times N_1N_2$, $w_0 \in N_1^0 V_2^+$ et $(u_1, w_0) \in N'$.

$((u_1, y), (u_1, w_0))$ est une règle de P'_1 et l'hypothèse de récurrence nous donne :

$(u_1, y) \xrightarrow[G']{*} (u_1, w_0) \xrightarrow[G']{*} \Lambda$.

- Si $\pi \in N_2 \times N_1 N_2$, y et W_0 peuvent se factoriser respectivement en $Y'A_2 y''$ et $y'E_1 B_2 y''$ avec $A_2, B_2 \in N_2$, $E_1 \in N_1$, $\pi = (A_2, E_1 B_2) \in P$, $y' \in N_1^0 V_2^*$ et $y'' \in V_2^*$. Etant donné que $P \setminus (P \cap N_2^2 \times N_2^2) = P_1$ est monotone standard de type 2 en N_2 , donc monotone standard de type 4 en N_2 , nous pouvons appliquer le lemme III.1 à la dérivation $y'E_1 B_2 y'' \xrightarrow{*}_G u_1$: il existe alors $z \in V_2^*$, $D_2, C_2 \in N_2$ tels que $B_2 y'' \xrightarrow{*}_G C_2 z$, $E_1 C_2 \xrightarrow{*}_G D_2$ et $y'D_2 z \xrightarrow{*}_G u_1$. Posons $z' = D_2 z$. z' appartient à V_2^+ et comme $y'z' \xrightarrow{*}_G u_1$, $\ell(z') \leq \ell(u_1) - \ell_{V_2}(y')$. De même $E_1 B_2 y'' \xrightarrow{*}_G z'$ implique $\ell(z') \geq \ell(B_2 y'')$. Donc $((u_1, y'A_2 y''), (u_1, y'z'))$ est une règle de P'_2 . Comme les dérivations $y'z' \xrightarrow{*}_G u_1$ et $E_1 B_2 y'' \xrightarrow{*}_G z'$ sont de longueur au plus égale à s , et qu'en outre, $y'z' \in N_1^0 V_2^+$, $E_1 B_2 y'' \in N_1 V_2^+$, $z' \in V_2^+$, $\ell_{V_2}(y'z') \leq \ell(u_1) \leq t$ et $\ell_{V_2}(E_1 B_2 y'') \leq \ell(z') \leq t$, nous pouvons utiliser l'hypothèse de récurrence. Nous obtenons :

$$(u_1, y) = (u_1, y'A_2 y'') \xrightarrow{*}_{G'} (u_1, y'z')(z', E_1 B_2 y'') \xrightarrow{*}_G \Lambda.$$

Des propriétés (b) et (b'), nous déduisons immédiatement: $x \in L(G)$ si et seulement si $\Lambda \in L(G')$ ce qui est décidable car $L(G')$ est un C-langage.

Si L appartient à la classe C_5 , d'après la proposition III.1, $\tilde{L}$ appartient à C_4 . Comme $x \in L$ si et seulement si $\tilde{x} \in \tilde{L}$, pour décider si $x \in L$ il suffit de déterminer si $\tilde{x} \in \tilde{L}$.

Montrons maintenant qu'il existe un ensemble récursif sans mot vide qui n'appartient ni à la classe C_4 ni à la classe C_5 . Nous pouvons, sans nuire à la généralité, supposer que T se réduit à un seul élément. Posons $T = \{a\}$ et notons C la classe des langages de C_4 ou C_5 qui sont inclus dans $\{a\}^+$. Soit $\mathbb{N}$ l'ensemble des entiers naturels. Nous utiliserons une première propriété :

Lemme III.6 Il existe une application surjective H de $\mathbb{N}$ dans C qui peut être construite effectivement.

Démonstration Rappelons d'abord les définitions de quelques fonctions primitives récursives classiques que nous utiliserons par la suite [DA].

- α définie sur $\mathbb{N}$ par $\alpha(x) = 0$ si $x > 0$, 1 si $x = 0$
- Pr définie sur $\mathbb{N}$ par $Pr(0) = 0$, $Pr(1) = 2$, $Pr(x+1) =$ le plus petit nombre premier strictement supérieur à $Pr(x)$.
- kGl où k est un paramètre entier positif définie sur $\mathbb{N}$ par $kGl(0) = 0$
 $kGl(x) =$ l'exposant du $k^{\text{ème}}$ nombre premier dans le développement de x en facteurs premiers, si $x > 0$.
- J définie sur $\mathbb{N} \times \mathbb{N}$ par $J(x,y) = \frac{1}{2} [(x+y)^2 + 3x + y]$ qui est une bijection de $\mathbb{N} \times \mathbb{N}$ sur $\mathbb{N}$.
- Q_1 et Q_2 définies sur $\mathbb{N}$ qui sont les fonctions coordonnées de l'inverse de J et qui vérifient donc $J(Q_1(z), Q_2(z)) = z$ pour tout $z \in \mathbb{N}$,
 $Q_1(J(x,y)) = x$ et $Q_2(J(x,y)) = y$ pour tout $x, y \in \mathbb{N}$.

Pour tout $j \in \mathbb{N}$, posons $i(j) = 4 + \alpha(1Gl(j+1))$, $p_2(j) = 2Gl(j+1) + 2$,
 $p_1(j) = 3Gl(j+1)$, $q(j) = 4Gl(j+1)$, $p(j) = p_1(j) + p_2(j)$.

Soient $N'_2(j) = \{A_1, \dots, A_{p_2(j)}\}$, $N'_1(j) = \{A_{p_2(j)+1}, \dots, A_{p(j)}\}$ deux alphabets et $N'(j) = N'_1(j) \cup N'_2(j)$.

Soit $n_{p(j)}$ une bijection de N'^*_j dans $\mathbb{N}$ définie par $n_{p(j)}(A_{\ell_1} \dots A_{\ell_s}) = \sum_{k=0}^{s-1} \ell_{s-k} \times [p(j)]^k$.

Notons, pour $s \in [1,2]$, $u'_s(j,t) = n_{p(j)}^{-1} [Q_s(tGl(j+1))]$.

Nous pouvons maintenant associer à j , la grammaire $G_j = (N'(j), \{a\}, P'(j), A_2)$ avec $P'(j)$ défini de la manière suivante :

$$P'(j) = \{(u'_1(j,t), u'_2(j,t)) / t \in [5, 4 + q(j)], M_{N'_1(j), N'_2(j)}^{(j)}(u'_1(j,t), u'_2(j,t)) \cup \{(A_1, \bar{a})\}\},$$

où $M_{X,Y}^k$ sont les prédicats définis page 1.4.

Posons $H(j) = L_j = L(G_j)$.

Par construction G_j est une grammaire monotone de type 4 ou 5 en $N'_2(j)$ et $L(G_j)$ est inclus dans $\{a\}^*$. Donc H est bien une application de $\mathbb{N}$ dans C . Il nous reste à montrer que H est surjective. Soit L un langage de la classe C . Il existe donc une grammaire $G = (N, \{a\}, P, A)$ monotone de type i ($i=4$ ou 5) en N_2 .

Posons $i' = 5-i \geq 0$, $p_2 = |N_2| + 1 \geq 2$ car $A \in N_2$, $p_1 = |N_1|$, $p = p_1 + p_2$,
 $q = |P|$, $N'_2 = \{A_1 \dots A_{p_2}\}$, $N'_1 = \{A_{p_2+1}, \dots, A_p\}$ et $N' = N'_1 \cup N'_2$.

Soit h une bijection de V^* dans N'^* vérifiant $h(A) = A_2$, $h(a) = A_1$,
 $h(N_1) = N'_1$ et $h(N_2) = N'_2 \setminus \{A_1\}$. Il est facile de vérifier que la grammaire
 $G' = (N', \{a\}, P' \cup \{(A_1, a)\}, A_2)$ avec $P' = \{(h(u), h(v)) / (u, v) \in P\}$ est équivalente
à G et monotone de type i en N'_2 . Numérotons de façon arbitraire l'ensemble P' ,
en notant : $P' = \bigcup_{s=1}^q \{(u'_s, v'_s)\}$.

$$\text{Posons } j'_0 = 2^{i'} 3^{p_2-2} 5^{p_1} 7^q \prod_{s=1}^q [\text{Pr}(4+s)]^{J(n_p(u'_s), n_p(v'_s))}.$$

Comme tous les coefficients des facteurs premiers sont des entiers positifs ou nuls, j'_0 est strictement positif et $j_0 = j'_0 - 1$ appartient à $\mathbb{N}$.

Montrons que $L(H(j_0)) = L$. Pour obtenir ce résultat, nous allons vérifier que les grammaires G_{j_0} et G' coïncident.

Nous obtenons en effet :

$$i(j_0) = 4 + \alpha(1Gl(j'_0)) = 4 + \alpha(5-i) = i,$$

$$p_2(j_0) = 2Gl(j'_0) + 2 = p_2^{-2+2} = p_2,$$

$$p_1(j_0) = 3Gl(j'_0) = p_1 \text{ et } q(j_0) = 4Gl(j'_0) = q,$$

ce qui entraîne $N'_2(j_0) = N'_2$, $N'_1(j_0) = N'_1$, $N'(j_0) = N'$, $|P'(j_0)| = |P'| + 1$ et G_{j_0} monotone de type i .

Pour tout $s \in [1, q]$, en posant $s' = s+4$, nous obtenons la relation suivante :

$$(1) \quad s'Gl(j'_0) = J(n_p(u'_s), n_p(v'_s)).$$

Soit (u', v') une règle de $P'(j_0) \setminus \{(A_1, a)\}$.

Il existe $s \in [1, q]$ tel que $u' = n_p^{-1}(Q_1(s'Gl(j'_0)))$ et $v' = n_p^{-1}(Q_2(s'Gl(j'_0)))$,
ce qui entraîne d'après la relation (1) et comme n_p est une bijection $u' = u_s$ et $v' = v_s$
donc (u', v') est une règle de P' .

Réciproquement, soit (u'_s, v'_s) une règle de P' . La relation (1) donne $u'_s = n_p^{-1}(Q_1(s'Gl(j'_0)))$ et $v'_s = n_p^{-1}(Q_2(s'Gl(j'_0)))$. En outre $M_{N'_1, N'_2}^i(u'_s, v'_s)$ est vérifiée et $s' \in [5, 4+q]$, donc $(u'_s, v'_s) \in P'(j_0)$.

Nous en déduisons que $P'(j_0) = P'$ et donc $L(H(j_0)) = L(G') = L(G)$.
H est bien une application surjective.

Soit $R = \{x \in \{a\}^+ / x \notin H(L(x))\}$.

$x \in R$ si et seulement si $x \notin H(L(x))$. Comme $x \in \{a\}^+$, nous pouvons effectivement construire la grammaire monotone de type 4 ou 5 qui engendre $H(L(x))$ et déterminer si $x \in H(L(x))$ (proposition III.2). Donc R est récursif. Supposons que R appartienne à C. Comme H est surjective, il existe $j \in N$ tel que $R = H(j)$. En prenant $x = a^j$, nous obtenons une contradiction : $a^j \in H(j) = R$ si et seulement si $a^j \notin H(j)$.

Donc R n'appartient pas à C et nous pouvons énoncer :

Proposition III.3 La classe C est strictement incluse dans la classe des ensembles récursifs de $\{a\}^+$.

Intéressons-nous maintenant aux grammaires monotones de type 1 qui sont des cas particuliers de grammaires monotones de type 4.

Proposition III.4 $L \subset T^*$ est un langage contextuel si et seulement s'il existe une grammaire monotone de type 1 qui l'engendre.

Démonstration Soit $G = (N, T, P, A)$ une grammaire monotone de type 1 en N_2 . D'après le lemme I.4, nous pouvons supposer qu'elle est standard. Nous pouvons donc appliquer le lemme III.2. Soit $G' = (N, T, P', A)$ la grammaire obtenue. Alors $(u, v) \in P' \cap N_2 \times V^*$ implique $v \in T \cup N_2 N_2^0$. Donc, $A \xrightarrow[G']{*} W$ implique $W \in V_2^+$. Les règles appartenant à $N_1 N_2 \times N_2$ ne peuvent donc pas être utilisées dans les dérivations à partir de l'axiome A. Posons $P'' = P' \setminus N_1 N_2 \times N_2$ et $G'' = (N, T, P'', A)$. Alors, $(u'', v'') \in P''$ implique $\ell(v'') \geq \ell(u'')$. Donc $L(G) = L(G') = L(G'')$ est un langage contextuel.

Réciproquement toute grammaire contextuelle $G = (N, T, P, A)$ est monotone de type 1 en N. En effet $(u, v) \in P$ implique $\ell_N(u) = \ell(u) \leq \ell(v) = \ell_V(v)$.

En fait comme au chapitre précédent, nous obtiendrons un résultat plus fort, en montrant qu'il suffit de considérer uniquement les règles où ne figure aucun terminal

Nous utiliserons la proposition III.4 ainsi qu'un lemme préliminaire relatif à une grammaire $G = (N, T, P, A)$ et à $P_N = \{(u, v) \in P / \ell_T(v) = 0\}$.

Lemme III.7 Si P_N est monotone de type 1 en N_2 , il existe une grammaire $G' = (N', T, P', A)$ équivalente à G telle que :

- (1) $P'_{N'} = \{(u, v) \in P' / \ell_T(v) = 0\}$ soit monotone de type 1 en $N'_2 \subset N'$.
- (2) $P'_T = \{(u, v) \in P' / \ell_T(v) > 0\}$ soit inclus dans $N^*N'_2 \times V'^*$ en posant $V' = N' \cup T$.

Démonstration A N_1 faisons correspondre biunivoquement $\bar{N}_1 = \{\bar{E}_1 / E_1 \in N_1\}$ disjoint de V . Posons $N'_2 = N_2 \cup \bar{N}_1$, $V'_2 = N'_2 \cup T$, $N' = N'_2 \cup N_1$, $V' = N' \cup T$ et définissons l'homomorphisme h de V'^* dans V^* par $h(\bar{E}_1) = E_1$ pour tout $\bar{E}_1 \in \bar{N}_1$, $h(X) = X$ pour tout $X \in V' \setminus \bar{N}_1$.

Posons $P'_{N'} = \{(u, v') / (u, v) \in P_N \text{ et } h(v') = v\}$

$P'_0 = \{(u_1 A_2, v') / (u_1 A_2, v) \in P_T, A_2 \in N_2, h(v') = v\}$

$P'_1 = \{(u_1 \bar{E}_1, v') / (u_1 E_1, v) \in P_T, E_1 \in N_1, h(v') = v\}$

$P'_T = P'_0 \cup P'_1$, $P' = P'_{N'} \cup P'_T$ et $G' = (N', T, P', A)$.

La propriété (2) est vérifiée par construction.

De plus, comme pour tout $(u, v') \in P'_{N'}$, $\ell_{V'_2}(v') \geq \ell_{V_2}(h(v')) \geq \ell_{N_2}(u) = \ell_{N'_2}(u)$, $P'_{N'}$ est monotone de type 1 en N'_2 .

Montrons maintenant l'équivalence entre G et G' , en utilisant les deux propriétés :

(a) Si $A \xrightarrow[G]{*} W$ alors $A \xrightarrow[G']{*} W'$ pour tout W' vérifiant $h(W') = W$.

(b) Si $A \xrightarrow[G']{*} W'$ alors $A \xrightarrow[G]{*} h(W')$.

Montrons la propriété (a) par récurrence sur la longueur k de la dérivation dans G .

Si $k=0$, $W = A = h(A)$ et $A \xrightarrow[G']{*} A$.

Supposons la propriété (a) vraie pour $k \leq t$ et considérons une dérivation de longueur $t+1$ en faisant apparaître la dernière règle utilisée :

$A \xrightarrow[G']{*} W_1 u W_2 \xrightarrow[G]{*} W_1 v W_2 = W$ avec $(u, v) \in P$. Soit W' tel que $h(W') = W = W_1 v W_2$.

Comme h est un homomorphisme, W' peut se factoriser en $W'_1 v' W'_2$ vérifiant $h(W'_1) = W_1$, $h(v') = v$ et $h(W'_2) = W_2$.

Si $u \in N^* N_2$, posons $u' = u$; sinon u peut s'écrire $u_1 E_1$ avec $E_1 \in N_1$: posons alors $u' = u_1 \bar{E}_1$. Dans les deux cas (u', v') appartient à P' et $h(W'_1 u' W'_2) = W_1 u W_2$. L'hypothèse de récurrence permet donc d'obtenir : $A \xrightarrow[G']{*} W'_1 u' W'_2 \xrightarrow[G']{*} W'_1 v' W'_2$.

Comme h est un homomorphisme vérifiant $h(A) = A$ et que $(u', v') \in P'$ implique $(h(u'), h(v')) \in P$, la propriété (b) est vraie aussi.

Pour tout $x \in T^*$, $h(x) = x$. Donc les propriétés (a) et (b) entraînent l'équivalence des grammaires G et G' .

Nous pouvons maintenant énoncer :

Proposition III.5 Soit $G = (N, T, P, A)$ une grammaire à structure de phase.

Si $P_N = \{(u, v) \in P / \ell_T(v) = 0\}$ est monotone de type 1, alors $L(G)$ est un langage contextuel.

Démonstration Il existe $N_2 \subset N$ tel que P_N soit monotone de type 1 en N_2 .

D'après le lemme III.7, nous pouvons supposer que P_T est inclus dans $N^* N_2 \times V^*$.

Prenons d et $\bar{d} \notin V$ et posons $T' = T \cup \{d\}$, $N'_2 = N_2 \cup \{\bar{d}\}$, $V'_2 = N'_2 \cup T'$, $N' = N'_2 \cup N'_1$, $V' = N' \cup T'$,

$$P_0 = \{(X\bar{d}, \bar{d}X) / X \in N\}, P'_{N'} = P_N \cup P_0,$$

$$P'_{T'} = \{(u, \bar{d}^t v) / (u, v) \in P_T, t = \ell_{N_2}(u)\} \cup \{(\bar{d}, d)\}$$

$$\text{et } G' = (N', T', P', A).$$

Comme P_T est inclus dans $N^*N_2 \times V^*$ et que $\bar{d} \in N'_2$, $P'_{T'}$ est inclus dans $N^*N'_2 \times V'_2V^*$ et G' est monotone de type 1 en N'_2 . La proposition III.4 permet alors d'affirmer que $L(G')$ est un langage contextuel.

Notons h l'homomorphisme de V'^* dans V^* défini par $h(\bar{d}) = h(d) = \Lambda$ et $h(X) = X$ pour tout $X \in V$.

Comme $(u', v') \in P'$ implique $h(u') \xrightarrow[G]{*} h(v')$, que h est un homomorphisme vérifiant $h(A) = A$ et que $h(T'^*) = T^*$, $W' \in L(G')$ entraîne $h(W') \in L(G)$.

Donc $h(L(G'))$ est inclus dans $L(G)$.

Pour établir l'inclusion inverse nous utiliserons la propriété suivante :

- (a) Si $A \xrightarrow[G]{*} W$, alors il existe $W' \in V'^*$ tel que $\ell_d(W') = 0$, $h(W') = W$ et $A \xrightarrow[G']{*} W'$. (Nous écrivons $\ell_d(W')$ à la place de $\ell_{\{d\}}(W')$).

Démontrons la propriété (a) par récurrence sur la longueur k de la dérivation dans G .

Si $k=0$, $W = A = h(A)$. Comme $\ell_d(A) = 0$ il suffit de prendre $W' = A$.

Supposons la propriété (a) vraie pour $k \leq t$, et considérons une dérivation de longueur $t+1$, où nous faisons apparaître la dernière règle utilisée :

$$A \xrightarrow[G]{*} W_1 u W_2 \xrightarrow[G]{} W_1 v W_2 = W \text{ avec } (u, v) \in P.$$

D'après l'hypothèse de récurrence, il existe W'' vérifiant $A \xrightarrow[G']{*} W''$, $\ell_d(W'') = 0$ et $h(W'') = W_1 u W_2$. Donc W'' peut se factoriser en $W''_1 u'' W''_2$ avec $h(W''_1) = W_1$, $h(u'') = u$ et $h(W''_2) = W_2$. Comme $u \in N^+$ et que $\ell_d(u'') = 0$, u'' appartient à N'^+ et,

en employant les règles de P_0 , nous obtenons $u'' \xrightarrow[G']{*} \bar{d}^s u$ avec $s = \ell_{\bar{d}}(u'')$.

(u, v) est une règle de P , donc il existe $v' \in V'^+$ vérifiant $h(v') = v$, $\ell_d(v') = 0$ et $(u, v') \in P'$, et nous obtenons dans G' , la dérivation suivante :

$$A \xrightarrow{*} W_1'' u'' W_2'' \xrightarrow{*} W_1'' \bar{d}^s u W_2'' \longrightarrow W_1'' \bar{d}^s v' W_2''.$$

Posons $W' = W_1'' \bar{d}^s v' W_2''$. Il est facile de vérifier que $h(W') = W$ et $\ell_d(W') = 0$. Donc la propriété (a) est bien vérifiée.

Si nous prenons $W \in L(G)$, la propriété (a) entraîne l'existence de $W' \in (T \cup \{\bar{d}\})^*$ vérifiant $h(W') = W$ et $A \xrightarrow[G']{*} W'$. Il suffit alors de remplacer dans W' les $\bar{d}$ par d en utilisant la règle $(\bar{d}, d)$, pour obtenir $W'' \in T'^*$ vérifiant $h(W'') = h(W') = W$ et $A \xrightarrow[G']{*} W' \xrightarrow[G']{*} W''$. Donc $L(G)$ est inclus dans $h(L(G'))$ et $L(G) = h[L(G')]$.

Les homomorphismes qui ne sont pas Λ -libres, ne conservent pas, en général, les langages contextuels. Cependant, s'il existe k tel que, pour tout $x \in L$, langage contextuel, $\ell(x) \leq k \times \ell(h(x))$, alors $h(L)$ est encore un langage contextuel [GG1 page 570].

Posons $t_0 = \sup_{(u,v) \in P_T} \ell_{N_2}(u) + 1$ et $\bar{T}' = T \cup \{\bar{d}\}$.

Il est, alors, facile de vérifier que, pour tout $(u', v') \in P' \setminus \{(\bar{d}, d)\}$,

$$\ell_{\bar{T}'}(v') \leq t_0 \times \ell_T(v').$$

Nous en déduisons que, si $A \xrightarrow[G']{*} W'$ avec $\ell_d(W') = 0$, alors

$$\ell_{\bar{T}'}(W') \leq t_0 \times \ell_T(W').$$

En effet, raisonnons par récurrence sur la longueur i de la dérivation.

Si $i=0$, $W' = A$ et $\ell_{\bar{T}'}(A) = \ell_T(A) = 0$.

Supposons la propriété vraie pour $i \leq \ell$ et considérons la dérivation :

$$A \xrightarrow[G']{*} u_1 u' u_2 \xrightarrow[G]{*} u_1 v' u_2 \text{ de longueur } \ell+1 \text{ avec } (u', v') \in P'.$$

Comme $\ell_d(v') = 0$, $u' \neq \bar{d}$, $(u', v') \in P' \setminus \{(\bar{d}, d)\}$,

$$\ell_{\bar{T}'}(v') \leq t_0 \times \ell_T(v') \text{ et } \ell_{\bar{T}'}(u') = \ell_T(u') = 0.$$

En utilisant l'hypothèse de récurrence, nous obtenons :

$$\begin{aligned} \ell_{\bar{T}'}(u_1 u_2) &= \ell_{\bar{T}'}(u_1 u' u_2) \leq t_0 \times \ell_T(u_1 u' u_2) = t_0 \times \ell_T(u_1 u_2) \text{ donc } \ell_{\bar{T}'}(w') = \\ \ell_{\bar{T}'}(u_1 v' u_2) &\leq t_0 \times \ell_T(u_1 v' u_2). \end{aligned}$$

Soit $x' \in L(G')$. Il existe $\bar{x}' \in \bar{T}'^*$ et une dérivation $A \xrightarrow[G']{*} \bar{x}' \xrightarrow[G']{*} x'$.

Comme la dérivation $\bar{x}' \xrightarrow[G']{*} x'$ n'utilise que la règle $(\bar{d}, d)$, nous obtenons

$$\begin{aligned} h(x') &= h(\bar{x}') \text{ et } \ell(x') = \ell(\bar{x}') = \ell_{\bar{T}'}(\bar{x}') \leq t_0 \times \ell_T(\bar{x}') = t_0 \times \ell(h(\bar{x}')) = \\ &t_0 \times \ell(h(x')). \end{aligned}$$

Donc $L(G) = h(L(G'))$ est un langage contextuel.

En prenant le cas particulier où P_N est monotone de type 1 en N (les règles de P_N sont de longueur croissante) nous retrouvons un résultat de Ginsburg et Greibach.

Corollaire III.1 [GG1] Soit $G = (N, T, P, A)$ une grammaire à structure de phrase. Si dans toutes les règles de longueur strictement décroissante, figure au moins un terminal, alors $L(G)$ est un langage contextuel.

Deuxième Partie

CARACTERISATION DE LANGAGES DETERMINISTES

Chapitre IV

APi - GRAMMAIRES ET C-LANGAGES DETERMINISTES.

Dans ce chapitre, nous donnons une double caractérisation grammaticale des C-langages déterministes. Pour cela, nous introduisons les AP1-grammaires qui généralisent les C-grammaires Λ -libres et les AP2, AP3-grammaires qui contiennent les C-grammaires mises sous forme normale de Greibach [GR]. En fait, ces grammaires engendrent les C-langages. Nous définissons sur les dérivations de longueur 1 des conditions de déterminisme et nous obtenons les APi-grammaires déterministes qui généralisent les s-grammaires [KH]. Nous montrons que les notions de dérivation la plus à gauche et de grammaire réduite s'étendent aux APi-grammaires. Ensuite, à l'aide de résultats concernant les automates à pile de mémoire déterministes, nous obtenons la double caractérisation annoncée ainsi que celles des langages LL(1) [LS] et des langages reconnus par les automates complets déterministes Λ -libres.

Soient $G = (N, T, P, A)$ une grammaire à structure de phrase, $N_2 \subset N$, $N_1 = N \setminus N_2$, $V = N \cup T$ et $V_2 = N_2 \cup T$.

Nous dirons que G est une

- AP1-grammaire en N_2 si $P \subset N_2^0 N_1^0 \times V_2 V^*$
- AP2-grammaire en N_2 si $P \subset N_2 N_1 \times N_2 \cup N \times TV^*$
- AP3-grammaire en N_2 si $P \subset N_2^0 N_1^0 \times TV^*$.

G sera une APi-grammaire s'il existe une partie N_2 de N telle que G soit une APi-grammaire en N_2 ($i=1,2,3$).

Toute APi-grammaire est une AP1-grammaire. En fait, ces trois types de grammaires sont équivalentes.

Montrons en effet :

Proposition IV.1 L est un C-langage si et seulement s'il existe une APi-grammaire qui engendre $L \setminus \{\Lambda\}$.

Démonstration Si L est un C-langage, il existe une C-grammaire $G = (N, T, P, A)$ sous forme normale de Greibach, qui engendre $L \setminus \{\Lambda\}$. Donc $P \subset N \times TV^*$ et G est une APi-grammaire en N .

Réciproquement, soit $G = (N, T, P, A)$ une APi-grammaire en N_2 . Considérons la grammaire $\tilde{G} = (N, T, \tilde{P}, A)$ avec $\tilde{P} = \{(\tilde{u}, \tilde{v}) / (u, v) \in P\}$. Posons $N_1 = N \setminus N_2$. L'ensemble $\{(x, y) \in \tilde{P} / x \notin N_1 \text{ et } \ell_T(y) = 0\}$ est monotone de type 2 en N_2 . D'après la proposition II.2, $L(\tilde{G})$ est un C-langage et $L(G) = L(\tilde{G})$ aussi.

Etendons aux APi-grammaires la notion de dérivation la plus à gauche, définie pour les C-grammaires. Soit $G = (N, T, P, A)$ une APi-grammaire. Nous écrirons $u \xrightarrow[G]{g} v$ si u et v se factorisent respectivement en $u_1 x u_2$ et $u_1 y u_2$ avec $u_1 \in T^*$ et $(x, y) \in P$. $\xrightarrow[G]{g^*}$ désignera la clôture transitive de $\xrightarrow[G]{g}$.

Posons $P_g(G) = \{u \in N^+ / \exists v \in V^* \text{ tel que } (u, v) \in P\}$, et pour $x \in N^0 N$ et $a \in T$, notons

$$- D_G(x, a) = \{y \in V^* / x \xrightarrow[G]{g} ay\}$$

$$- D_G(x, \Lambda) = \{y \in N V^* / x \xrightarrow[G]{g} y\}$$

G est une APi-grammaire déterministe (en N_2) si et seulement si c'est une APi-grammaire (en N_2) et si elle vérifie la propriété :

$$(D) \quad \forall x \in P_g(G), \forall a \in T, |D_G(x, a)| + |D_G(x, \Lambda)| \leq 1.$$

Montrons qu'à l'instar des C-grammaires, pour toute APi-grammaire $G = (N, T, P, A)$, si $W \in L(G)$, il existe dans G une dérivation la plus à gauche $A \xrightarrow[G]{g^*} W$. Nous utiliserons un résultat intermédiaire.

Lemme IV.1 Si $u_1 u_2 \xrightarrow[G]{*} W \in T^+$ avec $u_2 \in V_2 V^*$, alors il existe W_1, W_2 tels que $W = W_1 W_2$, $u_1 \xrightarrow[G]{*} W_1$ et $u_2 \xrightarrow[G]{*} W_2$ en utilisant les mêmes règles que dans la dérivation initiale.

Démonstration

Raisonnons par récurrence sur la longueur k de la dérivation.

Si $k=0$, il suffit de prendre $W_1=u_1$ et $W_2=u_2$.

Supposons la propriété vraie pour $k \leq p$ et considérons une dérivation

$u_1 u_2 \xrightarrow[G]{*} W$ de longueur $p+1$. Soit (x,y) la première règle utilisée. Comme $x \in N_2^0 N_1^0$ et que $u_2 \in V_2 V^*$, x est un facteur de u_1 ou de u_2 . Distinguons les deux cas.

- $u_1 = u'_1 x u''_1$. Posons $t_1 = u'_1 y u''_1$. Nous obtenons la dérivation $u_1 u_2 \xrightarrow[G]{*} t_1 u_2 \xrightarrow[G]{*} W$.
D'après l'hypothèse de récurrence appliquée à la dérivation $t_1 u_2 \xrightarrow[G]{*} W$ de longueur p , W se factorise en $W_1 W_2$ tels que $u_1 \xrightarrow[G]{*} t_1 \xrightarrow[G]{*} W_1$ et $u_2 \xrightarrow[G]{*} W_2$.
- $u_2 = u'_2 x u''_2$. Posons $t_2 = u'_2 y u''_2$. $u'_2 \in V_2 V^* \cup \{\Lambda\}$ et $y \in V_2 V^*$ donc $t_2 \in V_2 V^*$ et l'hypothèse de récurrence implique que $W = W_1 W_2$ avec $u_1 \xrightarrow[G]{*} W_1$ et $u_2 \xrightarrow[G]{*} t_2 \xrightarrow[G]{*} W_2$.

Lemme IV.2

Soient $u \in V^*$ et $W \in T^*$. Si $u \xrightarrow[G]{*} W$, alors $u \xrightarrow[G]{g^*} W$ en utilisant les mêmes règles.

Démonstration

Soit k la longueur de la dérivation. Si $k=0$ nous avons immédiatement $u \xrightarrow[G]{g^*} W$.

Supposons la propriété vraie pour $k \leq p$ et considérons une dérivation de longueur $p+1$.

Soit (x,y) la première règle utilisée. u se factorise en $u_1 x u_2$ et nous avons la dérivation : $u = u_1 x u_2 \xrightarrow[G]{*} u_1 y u_2 \xrightarrow[G]{*} W$.

$y \in V_2 V^*$, donc d'après le lemme IV.1, W se factorise en $W_1 W_2$ tels que $u_1 \xrightarrow[G]{*} W_1$ et $y u_2 \xrightarrow[G]{*} W_2$. En utilisant l'hypothèse de récurrence, nous obtenons $u = u_1 x u_2 \xrightarrow[G]{g^*} W_1 x u_2 \xrightarrow[G]{g} W_1 y u_2 \xrightarrow[G]{g^*} W_1 W_2$.

Etendons aussi aux APi-grammaires la notion de grammaire réduite.

Pour $x \in P_g(G)$, nous dirons que A dépend de x s'il existe dans G une dérivation $A \xrightarrow[G]{*} uxv$ avec $u, v \in V^*$. Nous dirons qu'une règle $(x,y) \in P$ est utilisable s'il existe une dérivation $A \xrightarrow[G]{*} uxv \xrightarrow[G]{*} uyv \xrightarrow[G]{*} W$ avec $u, v \in V^*$ et $W \in T^*$.

G sera pré-réduite si et seulement si A dépend de tous les éléments de $P_g(G)$, et sera réduite si et seulement si toutes les règles de P sont utilisables.

Lemme IV.3 Pour toute APi-grammaire (déterministe), il existe une APi-grammaire (déterministe) pré-réduite équivalente.

Démonstration Soient $G = (N, T, P, A)$ une APi-grammaire et $x \in P_g(G)$.

Prenons $d \notin V$ et posons $T' = T \cup \{d\}$,

$$P' = P \cup \{(x,d)\} \cup \{(B,a) / B \in N, a \in T\}.$$

La grammaire $G' = (N, T', P', A)$ est une APi-grammaire donc $L(G')$ est un C-langage.

Si A dépend de x dans G , il existe $u, v \in V^*$ tels que $A \xrightarrow[G]{*} uxv$ donc $A \xrightarrow[G']{*} uxv \xrightarrow[G']{*} udv \xrightarrow[G']{*} W_1 d W_2$ avec $W_1, W_2 \in T^*$ et $L(G') \cap T^* \{d\} T^* \neq \emptyset$.

Réciproquement, si $A \xrightarrow[G']{*} W_1 d W_2$ avec $W_1, W_2 \in T^*$, la règle (x,d) a été utilisée une seule fois et en regroupant les règles de P au début de la dérivation, nous obtenons $A \xrightarrow[G]{*} uxy \xrightarrow[G']{*} W_1 d W_2$. Donc A dépend de x si et seulement si $L = L(G') \cap T^* \{d\} T^* \neq \emptyset$ ce qui est décidable car L est un C-langage.

Notons X_0 l'ensemble des éléments de $P_g(G)$ dont dépend A dans G et considérons la grammaire $G_0 = (N, T, P_0, A)$ avec $P_0 = \{(x,y) \in P / x \in X_0\}$.

$P_0 \subseteq P$ donc G_0 est une APi-grammaire, $L(G_0) \subseteq L(G)$ et si G est déterministe, G_0 l'est aussi.

Toute règle (x,y) utilisée dans une dérivation $A \xrightarrow[G]{*} u$, appartient à P_0 . Donc $L(G) = L(G_0)$ et G_0 est bien pré-réduite.

Nous obtenons de même :

Lemme IV.4 Pour toute API-grammaire (déterministe), il existe une API-grammaire (déterministe) réduite équivalente.

Démonstration Soient $G = (N,T,P,A)$ une API-grammaire en N_2 et (x,y) une règle de P . Montrons qu'il est décidable de déterminer si (x,y) est utilisable dans G . Pour cela, prenons $d \notin V$ et considérons la grammaire $G' = (N,T',P',A)$ avec $T' = T \cup \{d\}$ et $P' = P \cup \{(x,dy)\}$. Si (x,y) est utilisable dans G , nous avons la dérivation $A \xrightarrow[G]{*} uxv \xrightarrow[G]{*} uyv \xrightarrow[G]{*} W \in T^*$. Comme $y \in V_2 V^*$, d'après le lemme IV.1, W se factorise en $W_1 W_2$ tels que $u \xrightarrow[G]{*} W_1$ et $yv \xrightarrow[G]{*} W_2$. Nous obtenons donc dans G' la dérivation $A \xrightarrow[G']{*} uxv \xrightarrow[G']{*} udyv \xrightarrow[G']{*} W_1 dW_2$. Réciproquement si $A \xrightarrow[G']{*} W_1 dW_2$ avec $W_1, W_2 \in T^*$, la règle (x,dy) a été utilisée une fois et une seule et nous avons la dérivation $A \xrightarrow[G']{*} uxv \xrightarrow[G']{*} udyv \xrightarrow[G']{*} W_1 dW_2$ ce qui implique $u \xrightarrow[G]{*} W_1$, $yv \xrightarrow[G]{*} W_2$ et $A \xrightarrow[G]{*} uxv \xrightarrow[G]{*} uyv \xrightarrow[G]{*} W_1 W_2 \in T^*$ donc la règle (x,y) est utilisable dans G .

Donc (x,y) est utilisable dans G si et seulement si $L(G') \cap T^*\{d\} T^* \neq \emptyset$ ce qui est décidable puisque $T^*\{d\} T^*$ est un K-langage, G' une API-grammaire en N_2 donc $L(G')$ un C-langage.

Notons $P_0 \subset P$ l'ensemble des règles de P utilisables dans G et $G_0 = (N,T,P_0,A)$. G_0 est une API-grammaire en N_2 et $L(G_0) \subseteq L(G)$. Comme toute dérivation $A \xrightarrow[G]{*} W \in T^*$ n'utilise que des règles de P_0 , toutes les règles de P_0 sont utilisables dans G_0 et $L(G) \subseteq L(G_0)$. De plus, comme $P_0 \subset P$, si G est déterministe, G_0 l'est aussi.

Montrons un premier résultat liant les APi-grammaires déterministes et les C-langages déterministes.

Lemme IV.5 Tout langage engendré par une APi-grammaire déterministe est un C-langage déterministe.

Démonstration Soit $G = (N, T, P, A)$ une APi-grammaire déterministe en N_2 .

Nous pouvons supposer, sans nuire à la généralité de la démonstration, que

$$P \subset N_2^0 N_1^0 \times V_2 N^*.$$

Considérons un automate à pile de mémoire $M = (Q, T, Z, g, z_0, A, \{z_0\})$ avec $Q = N \cup \{q, z_0\}$ (q, z_0 étant de nouveaux symboles $\notin V$), $Z = N \cup \{z_0\}$ et g définie de la manière suivante :

- $\forall b \in T^0, \forall B, C \in N : g(B, b, C) = \{(q, \tilde{y}) / y \in D_G(BC, b)\}$
- $\forall b \in T^0, \forall B \in N : g(B, b, z_0) = \{(q, z_0 \tilde{y}) / y \in D_G(B, b)\}$
- $\forall B \in Z : g(q, \Lambda, B) = \{(B, \Lambda)\}.$

Comme la grammaire G vérifie la propriété (D), l'automate à pile de mémoire M est quasi-déterministe et $T(M)$ est un C-langage déterministe. Montrons maintenant que $T(M) = L(G)$.

Soient $z \in N$, $b \in T^0$ et $y \in N^*$. S'il existe dans M le mouvement suivant :

$$(q, b, z_0 yz) \vdash_M^* (z, b, z_0 y) \vdash_M^* (q, \Lambda, z_0 y')$$

alors d'après les transitions de $M : z\tilde{y} \xrightarrow{G} b\tilde{y}'$.

Nous en déduisons immédiatement la propriété :

$$(1) \quad \forall W \in T^*, \forall y \in N^* (q, W, z_0 y) \vdash_M^* (q, \Lambda, z_0 y') \text{ implique } \tilde{y} \xrightarrow{G^*} W\tilde{y}'.$$

Réciproquement si $y \xrightarrow{G} by'$ avec $b \in T^0$, $y, y' \in N^*$, nous avons dans M le mouvement $(q, b, z_0 \tilde{y}) \vdash_M^* (q, \Lambda, z_0 \tilde{y}')$ et nous obtenons la propriété :

$$(2) \quad y \xrightarrow{G^*} Wy' \text{ avec } W \in T^*, y, y' \in N^* \text{ implique } (q, W, z_0 \tilde{y}) \vdash_M^* (q, \Lambda, z_0 \tilde{y}')$$

- Prenons $W \in T(M)$. Nous avons dans M le mouvement :

$$(A, bW', z_0) \vdash_M (q, W', z_0 y) \vdash_M^* (q, \Lambda, z_0) \vdash_M (z_0, \Lambda, \Lambda)$$

avec $b \in T^0$, $bW' = W$ et $D_G(A, b) = \{\tilde{y}\}$. Nous obtenons, en utilisant la propriété

$$(1) : A \xrightarrow{G} b\tilde{y} \xrightarrow{G^*} bW' = W \text{ donc } W \in L(G).$$

- Prenons $W \in L(G)$. D'après le lemme IV.2, il existe une dérivation $A \xrightarrow{G} by \xrightarrow{G^*}$

$$bW' = W, \text{ avec } b \in T^0, y \in N^* \text{ et } D_G(A, b) = \{y\}.$$

Donc, d'après la propriété (2), nous obtenons dans M :

$$(A, bW', z_0) \vdash_M (q, W', z_0 \tilde{y}) \vdash_M^* (q, \Lambda, z_0) \vdash_M (z_0, \Lambda, \Lambda) \text{ et } W = bW' \in T(M).$$

Montrons maintenant que tout C-langage déterministe suivi d'un double marqueur de fin, est engendré par une AP2-grammaire déterministe. Pour cela, nous utiliserons un certain nombre de lemmes relatifs aux automates à pile de mémoire déterministes.

Rappelons d'abord quelques notions concernant ces automates. Soit

$M = (Q, T, Z, g, z_0, q_0, F)$ un automate à pile de mémoire déterministe.

Nous écrirons $(q, W, y) \vdash_M^{d^*} (q', W', y')$ si

$$(q, W, y) \vdash_M^* (q', W', y') \text{ et si } g(q', \Lambda, z) = \emptyset \text{ pour } z \in Z \text{ tel que } y' = y''z.$$

$T_d(M)$ désignera $\{W \in T^* / \exists q \in F \text{ tel que } (q_0, W, z_0) \vdash_M^{d^*} (q, \Lambda, y)\}$

et $\bar{M}$ l'automate $(Q, T, Z, g, z_0, q_0, Q \setminus F)$.

Nous dirons que M est **sans boucle** si $T_d(M) \cup T_d(\bar{M}) = T^*$.

Remarquons que M est sans boucle si et seulement si $\forall W \in T^*, \exists q \in Q, y \in Z^*$ tels que $(q_0, W, z_0) \vdash_M^{d^*} (q, \Lambda, y)$. De plus q et y sont uniques.

Considérons maintenant $M = (Q, T, Z, g, z_0, q_0, F)$ un automate à pile de mémoire non nécessairement déterministe.

Le couple $(q, y) \in Q \times Z^*$ sera dit **accessible** dans M s'il existe $W \in T^*$ et $y' \in Z^*$ tels que $(q_0, W, z_0) \stackrel{*}{\vdash}_M (q, \Lambda, y'y)$.

Lemme IV.6 Il est décidable de déterminer si un couple $(q', y) \in Q \times Z^*$ est accessible dans M .

Démonstration Soit $M = (Q, T, Z, g, z_0, q_0, F)$ un automate à pile de mémoire. Si $y = \Lambda$ notons M' l'automate $(Q, T, Z, g, z_0, q_0, \{q'\})$.

Sinon y se factorise en $z_p \dots z_1$ avec $z_i \in Z$ pour $i \in [1, p]$. Prenons $d \notin T$, $q_1, \dots, q_p \in Q$. Posons $T' = T \cup \{d\}$, $Q' = Q \cup \{q_1, \dots, q_p\}$ et considérons l'automate à pile de mémoire $M' = (Q', T', Z, g', z_0, q_0, \{q_p\})$ où g' est définie par

- $\forall q \in Q, \forall b \in T^0, \forall z \in Z \quad g'(q, b, z) = g(q, b, z)$
- $g'(q', d, z_1) = \{(q_1, \Lambda)\}$
- $g'(q_i, \Lambda, z_{i+1}) = \{(q_{i+1}, \Lambda)\}$ pour $i \in [1, p-1]$
- $g'(q, b, z) = \emptyset$ partout ailleurs.

Si (q', y) est accessible dans M , il existe $W \in T^*$ et $y' \in Z^*$ tel que $(q_0, W, z_0) \stackrel{*}{\vdash}_M (q', \Lambda, y'y)$ donc $(q_0, Wd, z_0) \stackrel{*}{\vdash}_{M'} (q', d, y'z_p \dots z_1) \stackrel{*}{\vdash}_{M'} (q_p, \Lambda, y')$ et $Wd \in T(M')$.

Réciproquement si $W' \in T(M')$, alors il existe $y' \in Z^*$ tel que $(q_0, W', z_0) \stackrel{*}{\vdash}_{M'} (q_p, \Lambda, y')$ et nous avons obligatoirement $W' = Wd$ avec $W \in T^*$ et $(q_0, Wd, z_0) \stackrel{*}{\vdash}_M (q', \Lambda, y'z_p \dots z_1) \stackrel{*}{\vdash}_M (q_p, \Lambda, y')$.

Donc (q', y) est accessible dans M si et seulement si $T(M')$ est non vide ce qui est décidable puisque $T(M')$ est un C-langage.

Si M est déterministe, nous obtenons les trois résultats suivants :

Lemme IV.7 Un automate déterministe M est sans boucle si et seulement si
 $\forall (q, z) \in Q \times Z$ accessible dans M , il existe $(q', u') \in Q \times Z^*$ tel que

$$(q, \Lambda, z) \vdash_M^* (q', \Lambda, u')$$

Démonstration Supposons M sans boucle et considérons $(q, z) \in Q \times Z$ accessible dans M . Il existe donc $W \in T^*$ et $y \in Z^*$ tels que

$$(q_0, W, z_0) \vdash_M^* (q, \Lambda, yz) \quad (i)$$

Mais comme M est sans boucle, il existe $q' \in Q$, $y' \in Z^+$ tels que

$$(q_0, W, z_0) \vdash_M^* (q', \Lambda, y') \quad (ii).$$

On ne peut plus faire de mouvement simple à partir de la configuration (q', Λ, y') donc, comme M est déterministe (i) et (ii) entraînent $(q, \Lambda, yz) \vdash_M^* (q', \Lambda, y')$, et il existe $q_1, \dots, q_r \in Q$, $x_1, \dots, x_r \in Z^*$ tels que $q = q_1$, $yz = x_1$, $q_r = q'$, $x_r = y'$ et $(q_i, \Lambda, x_i) \vdash_M (q_{i+1}, \Lambda, x_{i+1})$ pour $i \in [1, r-1]$.

Distinguons deux cas :

- ① $\forall i \in [1, r], x_i = yx'_i$ avec $x'_i \in Z^+$ donc $(q, \Lambda, z) \vdash_M^* (q', \Lambda, x'_r)$.
- ② $\exists i \in [1, r]$ tel que $x_i = y$. Soit i_0 le plus petit. Nous obtenons
 $(q, \Lambda, z) \vdash_M^* (q_{i_0}, \Lambda, \Lambda)$.

Réciproquement, montrons par récurrence sur la longueur i de W , que

$$\forall W \in T^*, \exists q_1 \in Q \text{ et } y_1 \in Z^* \text{ tels que } (q_0, W, z_0) \vdash_M^* (q_1, \Lambda, y_1).$$

Si $i=0$, $W = \Lambda$ c'est vrai par hypothèse.

Supposons la propriété vraie pour $i \leq l$ et prenons $W = W'a$ avec $\ell(W') = l$ et $a \in T$.

D'après l'hypothèse de récurrence, $\exists q_2 \in Q$ et $y_2 \in Z^*$ tels que

$$(q_0, W', z_0) \vdash_M^* (q_2, \Lambda, y_2).$$

De plus, comme M est déterministe, la pile ne peut pas être vide, donc $y_2 \in Z^+$ et se factorise en $y_2' z$ avec $z \in Z$ et $g(q_2, \Lambda, z) = \emptyset$.

Posons $g(q_2, a, z) = (q'', y'')$. Nous obtenons :

$$(q_0, w'a, z_0) \vdash_M^* (q_2, a, y_2' z) \vdash_M (q'', \Lambda, y_2' y'').$$

Si $u'' = y_2' y'' \in Z$, nous avons par hypothèse :

$$(q'', \Lambda, y_2' y'') \vdash_M^{d^*} (q_1, \Lambda, y_1).$$

Supposons la propriété vraie pour $u'' \in Z^k$ et prenons $y_2' y'' = u'' z''$ avec $z'' \in Z$. (q'', z'') est accessible dans M , donc $(q'', \Lambda, z'') \vdash_M^{d^*} (q_1', \Lambda, z_1')$.

Si $z_1'' \in Z^+$ nous avons $(q'', \Lambda, u'' z'') \vdash_M^{d^*} (q_1', \Lambda, u'' z_1')$ sinon $(q'', \Lambda, u'' z'') \vdash_M^* (q_1', \Lambda, u'')$ avec $u'' \in Z^k$ et l'hypothèse de récurrence implique :

$$(q'', \Lambda, u'' z'') \vdash_M^* (q_1', \Lambda, u'') \vdash_M^{d^*} (q_1, \Lambda, y_1)$$

Lemme IV.8 Pour tout automate à pile déterministe M , il existe un automate à pile déterministe sans boucle $M' = (Q', T, Z', g', z_0', q_0', F')$ tel que $T_d(M') = T(M') = T(M)$ et qui vérifie la propriété :

$$(V) \quad \forall q' \in Q', \forall z' \in Z', g'(q', \Lambda, z') \neq \emptyset \implies \exists q_1' \in Q' \text{ tel que } g'(q', \Lambda, z') = (q_1', \Lambda).$$

Démonstration Nous pouvons supposer que M est sans boucle, que $T_d(M) = T(M)$ [GG2page 625] et que $g(q, \Lambda, z) = \emptyset$ pour tout couple $(q, z) \in Q \times Z$ non accessible dans M .

$$\text{Notons } n_0 = |\{(q, z) \in Q \times Z / g(q, \Lambda, z) \neq \emptyset\}|.$$

Si $n_0 = 0$, l'automate est Λ -libre et le lemme est vérifié. Dans le cas contraire, considérons l'ensemble $B = \{y \in Z^* / \exists q, q_1 \in Q, z \in Z \text{ tels que } g(q, \Lambda, z) = (q_1, y)\}$ et notons $\ell_0 = \sup_{y \in B} \ell(y)$.

Si $\ell_0 = 0$ le lemme est vérifié.

Si $\ell_0 > 0$ considérons $(q, z) \in Q \times Z$.

Comme M est sans boucle, d'après le lemme IV.7, il existe $q_1 \in Q$ et $y_1 \in Z^*$ tels que $(q, \Lambda, z) \vdash_M^{d^*} (q', \Lambda, y')$. Montrons que $\ell(y') \leq n_0 \times \ell_0$.

Supposons le contraire. Il existe alors un mouvement :

$$(q, \Lambda, z) = (q_1, \Lambda, y_1 z_1) \vdash_M (q_2, \Lambda, y_2 z_2) \dots \vdash_M (q_n, \Lambda, y_n z_n) = (q', \Lambda, y')$$

avec $y_i \in Z^*$, $z_i \in Z$

et $\ell(y_i) < n_0 \times \ell_0 \leq \ell(y_n)$ pour $i \in [1, n-1]$.

Posons : $I = \{i \in [1, n] / \forall s \in [i, n] \ell(y_s) \geq \ell(y_i)\}$.

Comme $\ell(y_1) = 0 \leq \ell(y_i) \leq \ell(y_n) \forall i \in [1, n]$, $1 \in I$ et $n \in I$.

Notons : $i_1, \dots, i_k$ les éléments de I avec $k = |I|$ et $i_j < i_{j+1}$ pour $j \in [1, k[$.

Montrons que pour $j \in [1, k[$, $\ell(y_{i_{j+1}}) - \ell(y_{i_j}) < \ell_0$.

Posons $I_j = \{i \in [i_j, i_{j+1}] / \ell(y_i) < \ell(y_{i_j}) + \ell_0\}$.

Par définition de i_j , $\ell(y_{i_{j+1}}) \geq \ell(y_{i_j})$, et par définition de ℓ_0 ,

$\ell(y_{i_{j+1}}) < \ell(y_{i_j}) + \ell_0$. Donc, $i_{j+1} \in I_j$ et $I_j \neq \emptyset$.

Soit t_j la borne supérieure de I_j .

Alors $\forall i \in [t_j, i_{j+1}[$, $\ell(y_i) \geq \ell(y_{i_j}) + \ell_0 > \ell(y_{t_j})$.

En outre, si $\ell(y_{i_{j+1}}) - \ell(y_{i_j}) \geq \ell_0$, alors par définition de i_{j+1} ,

$\forall i \in [i_{j+1}, n]$, $\ell(y_i) \geq \ell(y_{i_{j+1}}) \geq \ell_0 + \ell(y_{i_j}) > \ell(y_{t_j})$.

Donc $t_j \in I$ et $i_j < t_j < i_{j+1}$, ce qui est contradictoire.

Comme $\ell(y_{i_1}) = 0$ et $\ell(y_{i_k}) \geq n_0 \times \ell_0$, nous en déduisons que $k > n_0$.

Donc il existe deux entiers $r < r'$ appartenant à la séquence $i_1, \dots, i_k$ et tels que $q_r = q_{r'}$, et $z_r = z_{r'}$. Or d'après (1),

$$(q_r, \Lambda, z_r) \vdash_M^* (q_{r'}, \Lambda, y_{r'} z_{r'}) \text{ avec } y_{r'} = y_r y_r'.$$

Nous obtenons donc, pour tout $p \geq 0$, le mouvement :

$$(q, \Lambda, z) \vdash_M^* (q_r, \Lambda, y_r z_r) \vdash_M^* (q_r, \Lambda, y_r (y_r')^p z_r),$$

ce qui est contradictoire avec $(q, \Lambda, z) \vdash_M^{d*} (q', \Lambda, y')$.

Construisons maintenant l'automate M' de la manière suivante :

$$\text{Prenons } q'_0 \notin Q \text{ et } z'_0 \notin Z \text{ et posons } Y = \bigcup_{i=1}^{n_0 \times \ell_0} Z^i,$$

$$Y' = \{q x z / q \in Q, x z \in Y, z \in Z \text{ et } g(q, \Lambda, z) = \emptyset\}, \quad Q' = Q \cup Y' \cup \{q'_0\} \text{ et } Z' = Z \cup \{z'_0\}.$$

$$\text{Soient } q_1 \in Q \text{ et } x_1 \in Z^* \text{ tels que } (q_0, \Lambda, z_0) \vdash_M^{d*} (q_1, \Lambda, x_1).$$

$$\text{Nous prendrons } F' = F \cup \{q'_0 / \text{ si } q_1 \in F\} \cup (Y' \cap FY).$$

Comme l'automate M ne peut vider sa pile, $x_1 \in Z^+$ et se factorise en $x'_1 z_1$ avec $z_1 \in Z$.

Pour tout $a \in T$, nous poserons alors

$$i) \quad g'(q'_0, a, z'_0) = (q_2, z'_0 x'_1 x_2) \text{ avec } (q_2, x_2) = g(q_1, a, z_1) \text{ et } x'_1 x_2 \in Z^+.$$

Pour tout couple $(q, z) \in Q \times Z$ tel que $g(q, \Lambda, z) = \emptyset$, nous prendrons

$$ii) \quad g'(q, a, z) = g(q, a, z) \text{ pour tout } a \in T$$

$$\text{et } g'(q, \Lambda, z) = \emptyset.$$

Par contre si $g(q, \Lambda, z) \neq \emptyset$ distinguons deux cas :

$$- (q, \Lambda, z) \vdash_M^* (p, \Lambda, \Lambda) \quad \text{alors}$$

$$\text{iii) } g'(q, \Lambda, z) = (p, \Lambda)$$

$$\text{et } g'(q, a, z) = \emptyset \quad \forall a \in T.$$

$$- (q, \Lambda, z) \vdash_M^* (p, \Lambda, xz_2) \text{ avec } z_2 \in Z \text{ et } g(p, \Lambda, z_2) = \emptyset.$$

Dans ce cas $\ell(xz_2) \leq n_Q \times \ell_Q$ et $pxz_2 \in Y'$, nous poserons alors

$$\text{iv) } g'(q, \Lambda, z) = (pxz_2, \Lambda)$$

$$\text{et } g'(q, a, z) = \emptyset \quad \forall a \in T.$$

Soit $qxz \in Y'$ avec $z \in Z$. Pour tout $a \in T$ et tout $z' \in Z'$, nous poserons :

$$\text{v) } g'(qxz, \Lambda, z') = \emptyset$$

$$\text{et } g'(qxz, a, z') = (p, z'xx') \text{ avec } (p, x') = g(q, a, z).$$

Achevons la définition de g' en posant pour tout $q' \in Q'$, $z' \in Z'$, pour lesquels g' n'est pas encore définie, $g'(q', \Lambda, z') = \emptyset$ et $g'(q', a, z') = (q', z')$ $\forall a \in T$.

Par construction M' vérifie la propriété (V) et d'après le lemme IV.7, M' est sans boucle. Il nous reste à montrer que $T_d(M') = T(M') = T(M)$.

Soit t_1 l'homomorphisme de Z'^* dans Z^* défini par $t_1(z'_0) = \Lambda$,
 $T_1(z) = z \quad \forall z \in Z$.

Pour $q \in Q$, $y \in Z^*$, $qy \in Q \cup Y'$, $b \in T^0$ et $x \in Z'^+$,

$H(qy, b, x)$ désignera la configuration $(q, b, t_1(x)y)$. Il est facile de vérifier que pour $q'_1 \in Q \cup Y'$, $b \in T^0$ et $x'_1 \in Z'Z^*$ tels que $x'_1 = z'_0$ entraîne $q'_1 \in Y'$, nous avons la propriété :

$$(2) \quad (q'_1, b, x'_1) \vdash_M^* (q'_2, \Lambda, x'_2) \text{ implique } H(q'_1, b, x'_1) \vdash_M^* H(q'_2, \Lambda, x'_2).$$

Prenons $W \in T(M')$.

Si $W = \Lambda$, $q'_0 \in F'$ donc $(q_0, \Lambda, z_0) \vdash_M^* (q_1, \Lambda, x_1)$ avec $q_1 \in F$ et $\Lambda \in T(M)$.

Si $W \neq \Lambda$, $W = aW'$ avec $a \in T$ et nous avons dans M' le mouvement
 $(q'_0, aW', z'_0) \vdash_{M'} (q_2, W', z'_0 x'_1 x'_2) \vdash_{M'}^* (q', \Lambda, y')$ avec $q' \in F'$, $(q_0, \Lambda, z_0) \vdash_M^* (q_1, \Lambda, x'_1 z'_1)$,
 $g(q_1, a, z_1) = (q_2, x_2)$.

D'après la propriété (2), nous avons dans M :

$$(q_2, W', x'_1 x'_2) = H(q_2, W', z'_0 x'_1 x'_2) \vdash_M^* H(q', \Lambda, y').$$

Comme $q' \in F'$, $H(q', \Lambda, y') = (p, \Lambda, x)$ avec $p \in F$.

Nous obtenons alors

$$(q_0, aW', z_0) \vdash_M^* (q_1, aW', x'_1 z'_1) \vdash_M (q_2, W', x'_1 x'_2) \vdash_M^* (p, \Lambda, x)$$

donc $T(M') \subseteq T(M)$.

Réciproquement, montrons la propriété

$$(3) \quad \left\{ \begin{array}{l} \text{Si } (q, \Lambda, y) \vdash_M^* (p, \Lambda, y_1) \text{ avec } y \in Z^+, \text{ alors il existe } y'_1 \in Z^* \text{ tel que} \\ (q, \Lambda, z'_0 y) \vdash_{M'}^* (py'_1, \Lambda, z'_0 y'_2) \text{ avec } y'_2 y'_1 = y_1. \end{array} \right.$$

Raisonnons par récurrence sur la longueur de la dérivation. Si $\ell=0$ c'est évident puisque $g(q, \Lambda, z) = \emptyset$ entraîne $g'(q, \Lambda, z) = \emptyset$. Supposons la propriété vraie pour $\ell \leq k$ et considérons une dérivation de longueur $k+1$.

Posons $y = y'z$ avec $z \in Z$. Soient $p_1 \in Q$ et $u_1 \in Z^*$ tels que
 $(q, \Lambda, z) \vdash_M^* (p_1, \Lambda, u_1)$.

Distinguons deux cas :

- Si $u_1 \neq \Lambda$, nous avons alors $(q, \Lambda, y'z) \vdash_M^* (p_1, \Lambda, y'u_1)$ et
 $(q, \Lambda, z'_0 y'z) \vdash_{M'}^* (p_1 u_1, \Lambda, z'_0 y')$ par une transition iv).
- Si $u_1 = \Lambda$, nous avons $(q, \Lambda, y'z) \vdash_M^* (p_1, \Lambda, y')$ $\vdash_M^* (p, \Lambda, y_1)$ et
 $g'(q, \Lambda, z) = (p_1, \Lambda)$. En utilisant l'hypothèse de récurrence, nous obtenons bien :

$$(q, \Lambda, z_0' y) \mid_{\overline{M}}, (p_1, \Lambda, z_0' y') \mid_{\overline{M}}^{\frac{d^*}{M}} (py_1', \Lambda, z_0' y_2').$$

Soit $W \in T(M) = T_d(M)$.

Si $W = \Lambda$, $q_1 \in F$ donc $q_0' \in F'$ et comme $g'(q_0', \Lambda, z_0') = \emptyset$, $\Lambda \in T_d(M')$.

Si $W \neq \Lambda$, $W = aW'$ avec $a \in T$ et nous avons dans M :

$$(q_0, aW', z_0) \mid_M^* (q_1, aW', x_1' z_1) \mid_M (q_2, W', x_1' x_2) \mid_M^{\frac{d^*}{M}} (p, \Lambda, y_1) \text{ avec } p \in F.$$

La propriété (3) implique que, si

$$(q, b, y) \mid_M^* (p_2, b, u_2) \mid_M (p_3, \Lambda, u_3) \text{ avec } b \in T, \text{ alors}$$

$$(q, b, z_0' y) \mid_{M'}^* (p_3, \Lambda, z_0' u_3). \text{ Nous en déduisons que}$$

$$(q_2, W', z_0' x_1' x_2) \mid_{M'}^{\frac{d^*}{M}} (py_1', \Lambda, z_0' y_2') \text{ avec } y_2' y_1' = y_1 \text{ et } py_1' \in F'.$$

Comme $(q_0', aW', z_0') \mid_{M'} (q_2, W', z_0' x_1' x_2)$, nous obtenons

$$T_d(M) \subset T_d(M') \quad \text{et} \quad T_d(M') = T(M') = T(M).$$

Lemme IV.9 Pour tout automate à pile déterministe $M = (Q, T, Z, g, z_0, q_0, F)$ et pour $d \notin T$, le langage $T(M) \{d^2\}$ est engendré par une AP2-grammaire déterministe.

Démonstration Nous pouvons supposer que $Z, T \cup \{d\}$ et Q sont disjoints deux à deux et que M vérifie la propriété (v) (lemme IV.8), ce qui implique $g(q_0, \Lambda, z_0) = \emptyset$, puisque l'automate déterministe ne peut pas vider sa pile.

Soient A, E et S de nouveaux symboles.

$$\text{Posons } N = Q \cup \{A, E, S\} \cup Z \cup Q \times Z \cup F \times \{S\}$$

$$N_1 = Z \cup \{S\}, \quad N_2 = N \setminus N_1$$

$$P_0 = \{(A, aq_0'S) / a \in T, g(q_0, a, z_0) = (q, y)\} \cup \{(A, d^2) / \text{si } q_0 \in F\}$$

$$P_1 = \{(qz, q') / q, q' \in Q, z \in Z \text{ et } g(q, \Lambda, z) = (q', \Lambda)\}$$

$$P_2 = \{(qz, (q, z)) / q \in Q, z \in Z \text{ et } g(q, \Lambda, z) = \emptyset\}$$

$$P_3 = \{((q,z), aq'\tilde{y}) / q \in Q, z \in Z, a \in T \text{ et } g(q,a,z) = (q',y)\}$$

$$P_4 = \{((q,z), dE) / q \in F, z \in Z\}$$

$$P_5 = \{(Ez, E) / z \in Z\}$$

$$P_6 = \{(ES, (E,S))\}$$

$$P_7 = \{((E,S), d)\} \quad \text{et} \quad P = \bigcup_{i=0}^7 P_i.$$

Considérons la grammaire $G = (N, T \cup \{d\}, P, A)$.

$P_1 \cup P_2 \cup P_5 \cup P_6 \subset N_2 N_1 \times N_2$, $P_0 \cup P_3 \cup P_4 \cup P_7 \subset N \times TV^*$, donc G est une AP2-grammaire en N_2 . Et comme M est déterministe, il est clair que G vérifie la propriété (D), donc G est une AP2-grammaire déterministe.

Montrons que $L(G) = T(M)$.

Etant donné que M vérifie la propriété (v), si $(q,b,z) \vdash_M (q', \Lambda, y)$ avec $b \in T^0$ et $z \in Z$ alors $qz \xrightarrow{*}_G bq'\tilde{y}$ (en utilisant les règles de $P_1 \cup P_2 \cup P_3$).

Nous en déduisons immédiatement la propriété :

$$(1) \text{ Si } (q,W,y) \vdash_M^* (q', \Lambda, y') \text{ avec } y \in Z^* \text{ et } W \in T^* \text{ alors } q\tilde{y} \xrightarrow{*}_G Wq'\tilde{y}'.$$

De même, si $qz \xrightarrow{*}_G q'$ avec $q, q' \in Q$ et $z \in Z$ alors $(q, \Lambda, z) \vdash_M (q', \Lambda, \Lambda)$ et si $qz \xrightarrow{*}_G (q,z) \xrightarrow{*}_G aq'y$ avec $q, q' \in Q, a \in T$ et $z \in Z$ alors $(q,a,z) \vdash_M (q', \Lambda, \tilde{y})$.

Nous en déduisons la propriété :

$$(2) \text{ Si } qy \xrightarrow{*}_G Wq'y' \text{ avec } q, q' \in Q, W \in T^* \text{ et } y \in Z^*, \text{ alors } (q,W,\tilde{y}) \vdash_M^* (q', \Lambda, \tilde{y}').$$

Prenons $W \in T(M)$.

$$\text{Si } W = \Lambda, A \xrightarrow{P_0} d^2 \text{ et } d^2 \in L(G).$$

Sinon $W = aW'$ avec $a \in T$ et nous avons dans M le mouvement :

$$(q_0, aW', z_0) \vdash_M (p, W', y) \vdash_M^* (q, \Lambda, y') \text{ avec } q \in F \text{ et } y' \in Z^+$$

ce qui entraîne l'existence de la dérivation suivante dans G :

$$A \xrightarrow{P_0} a\tilde{y}S \xrightarrow{(1)^*} aW'q\tilde{y}'S \xrightarrow{P_4 \cup P_5 \cup P_6} aW'd(E,S) \xrightarrow{P_7} Wd^2 \in L(G).$$

Réciproquement, prenons $W_0 \in L(G)$. Il est clair que $L(G) \subset T^*\{d^2\}$, donc W_0 se factorise en Wd^2 avec $W \in T^*$.

Si $W = \Lambda$, $(\Lambda, d^2) \in P$ et $q_0 \in F$ donc $\Lambda \in T(M)$.

Sinon $W = aW'$ avec $a \in T$ et, d'après le lemme IV.2, il existe dans G une dérivation de W' la plus à gauche :

$$A \xrightarrow{P_0} aqyS \xrightarrow{G} aW'q'y'S \xrightarrow{G} aW'(q',z)y''S \xrightarrow{P_4} aW'dEy''S \xrightarrow{P_5 \cup P_6 \cup P_7} aW'd^2$$

avec $q' \in F$ puisque $((q',z), dE) \in P_4$ et nous obtenons dans M :

$$(q_0, aW', z_0) \mid_M (q, W', \tilde{y}) \mid_{(2)^*} (q', \Lambda, \tilde{y}') \quad \text{donc } W \in T(M)$$

$$\text{et } L(G) = T(M) \{d^2\}$$

Soient $L \subset T^*$ et $d \notin L$. Si $L\{d^2\}$ est engendré par une AP2 (AP1)-grammaire, d'après le lemme IV.5, $L\{d^2\}$ est déterministe et $L = \{W \in T^* / Wdu \in L\{d^2\} \text{ avec } u \in \{T \cup d\}^*\}$ est aussi déterministe [GQ2 page 637]. Donc, en utilisant le lemme IV.9 pour la réciproque, nous obtenons une caractérisation grammaticale des C-langages déterministes :

Proposition IV.2 $L \subset T^*$ est un C-langage déterministe si et seulement si, pour tout $d \notin T$, il existe une AP1 (AP2)-grammaire déterministe qui engendre $L\{d^2\}$.

Nous allons, maintenant, montrer qu'en ajoutant des restrictions aux APi-grammaires déterministes, nous obtenons des caractérisations nouvelles pour divers langages déterministes tels que les s-langages, les langages LL(1) et les

langages reconnus par les automates déterministes complets Λ -libres.

En particulier les s-grammaires sont des APi-grammaires déterministes. Réciproquement, si $G = (N, T, P, A)$ est une APi-grammaire déterministe en N , avec $i=2$ ou 3 , alors G est une s-grammaire. Par contre, si G est une AP1-grammaire déterministe en N , G n'est pas nécessairement une s-grammaire, mais $L(G)$ est un s-langage.

En effet, soit $B \in N$.

Si $D_G(B, \Lambda) \neq \emptyset$, $(B, x_1) \in P$ et $(B, x_2) \in P$ entraînent $x_1 = x_2$ (propriété (D)) et si $D_G(B, \Lambda) = \emptyset$, $(B, x_1) \in P$ et $(B, x_2) \in P$, $x_1 \xrightarrow{*}_G ay_1$ et $x_2 \xrightarrow{*}_G ay_2$ avec $a \in T$ impliquent $x_1 = x_2 = ax$ (propriété (D)). Donc G est une LL(1) grammaire. En outre, G est Λ -libre donc $L(G)$ est un s-langage [KS page 229].

Nous pouvons alors énoncer :

Proposition IV.3 L est un s-langage si et seulement s'il existe une APi-grammaire, $G = (N, T, P, A)$, déterministe en N , qui engendre L .

Donnons maintenant une caractérisation nouvelle pour les langages LL(1) introduits par Lewis et Stearns [LS].

Nous dirons que $G = (N, T, P, A)$ est une UAPi-grammaire (en N_2) si G est une APi-grammaire déterministe (en N_2) vérifiant la propriété :

(U) $\forall B, B' \in N, y \in V^*$, si $BB' \in P_g(G)$, alors $(BB', y) \in P \iff (B', y) \in P$.

Remarquons que si G est une UAP2-grammaire en N_2 , alors $L(G)$ est un s-langage. En effet $P_g(G) \cap N_2 N_1 = \emptyset$. Dans le cas contraire, la propriété (U) entraînerait que $P \cap N_1 \times N_2 \neq \emptyset$ ce qui est impossible car $(x, y) \in P$ avec $x \in N$ implique $y \in TV^*$. Donc P est inclus dans $N \times TV^*$ et la propriété (D) caractérise

alors les s-grammaires. Par contre, les UAP1 (UAP3)-grammaires engendrent les langages LL(1).

Rappelons les définitions de deux types particuliers de C-grammaires qui engendrent les langages LL(1). Une grammaire **permutable** [WA] est une C-grammaire $G = (N, T, P, A)$ vérifiant la propriété suivante :

$$\forall W \in T^*, \forall a \in T, \forall u_1, u_2 \in NV^*, \forall v_1, v_2 \in V^* \\ A \xrightarrow[G]{g^*} Wu_i \xrightarrow[G]{g} Wav_i \text{ pour } i \in [1,2] \text{ implique } v_1 = v_2.$$

Pour $B \in N$, notons $S_G(B) = \{a \in T / \exists u, v \in V^* \text{ tels que } A \xrightarrow[G]{*} uBav\}$.

Une **simple LL(1) grammaire** [KS] (grammaire prédictive [WA]) est une C-grammaire vérifiant les trois propriétés suivantes :

- (i) $P \subset N \times (TV^*)^0$,
- (ii) $\forall B \in N, \forall a \in T, \forall y, z \in V^*, \text{ si } (B, ay) \in P \text{ et } (B, az) \in P \text{ alors } y = z$,
- (iii) $\forall B \in N, \forall a \in T, \forall y \in V^*, \text{ si } (B, A) \in P \text{ et } (B, ay) \in P \text{ alors } a \notin S_G(B)$.

Montrons d'abord :

Lemme IV.10 Pour toute UAP1-grammaire, il existe une grammaire permutable équivalente.

Démonstration Soit $G = (N, T, P, A)$ une UAP1-grammaire en N_2 , telle que $P \subset N_2^0 N_1^0 \times V_2 N^*$. Posons $N' = N \cup N \times N_1$ et $V' = N' \cup T$.

Soit f l'application de V^* dans V'^* définie par :

- $\forall B \in V^0, f(B) = B$,
- $\forall a \in T, \forall x \in V^*, f(ax) = af(x)$,
- $\forall B \in N, \forall x \in V_2 V^*, f(Bx) = Bf(x)$,
- $\forall B \in N, \forall A_1 \in N_1, \forall x \in V^*, f(BA_1x) = (B, A_1) f(A_1x)$

Il est facile de montrer que f vérifie les propriétés :

- (1) $\forall x \in V^*, \forall y \in V_2^*, f(xy) = f(x) f(y),$
- (2) $\forall x, y \in V^*, \forall B \in N, f(xBy) = zf(By) \text{ avec } zB = f(xB).$

Posons $N'_2 = \{x / x \in P_g(G) \cap N^2\} \subset N_2 \times N_1,$

$$P'_1 = \{(B, f(v)) / B \in N, (B, v) \in P\},$$

$$P'_2 = \{((B, C), z) / B \in N, C \in N_1, (B, v) \in P \text{ avec } zC = f(vC)\},$$

$$P'_3 = N'_2 \times \{\Lambda\} \text{ et } P' = P'_1 \cup P'_2 \cup P'_3.$$

Considérons la C-grammaire $G' = (N', T, P', A).$

Montrons qu'elle est équivalente à G et qu'elle est permutable.

Soient (x, y) une règle de P et $C \in N_1^0.$

- Si $x \in N,$ il existe, par définition de $P'_1 \cup P'_2, x' \in N'$ et $y' \in V'^+$ tels que $(x', y') \in P'$ avec $x'C = f(xC)$ et $y'C = f(yC).$
- Si $x \in N_2 N_1.$ Posons $x = A_2 A_1$ avec $A_2 \in N_2$ et $A_1 \in N_1.$

Comme $(x, y) \in P, x \in P_g(G)$ et $x \in N'_2$ donc $((A_2, A_1), \Lambda) \in P'_3.$ De plus, comme G vérifie la propriété (U), $(A_1, y) \in P$ donc il existe $x'_1 \in N'$ et $y' \in V'^+$ tels que $(x'_1, y') \in P'$ avec $x'_1 C = f(A_1 C)$ et $y' C = f(yC).$ En posant $x' = (A_2, A_1) x'_1,$ nous obtenons $x' C = f(A_2 A_1 C) = f(xC)$ et $x' \xrightarrow[G']{g} x'_1 \xrightarrow[G']{g} y'.$

Donc si (x, y) est une règle de P et si $C \in N_1^0,$ il existe $x' \in N'^+$ et $y' \in V'^+$ tels que $x' \xrightarrow[G']{g^*} y', x' C = f(xC)$ et $y' C = f(yC).$

Nous en déduisons que si $u \xrightarrow[G]{g} v$ avec $u \in N^+, \text{ alors } f(u) \xrightarrow[G']{g^*} f(v).$

En effet posons $u = xu_1, v = yu_1$ avec $(x, y) \in P.$

Distinguons deux cas :

- (a) $u_1 \in \{\Lambda\} \cup N_2 N^*$. Alors $f(u) = f(x) f(u_1)$, $f(v) = f(y) f(u_1)$ et comme $f(x) \xrightarrow[G']{g^*} f(y)$, nous obtenons $f(u) \xrightarrow[G']{g^*} f(v)$.
- (b) $u_1 \in N_1 N^*$. Alors u_1 se factorise en $C u_2$ avec $C \in N_1$, $f(u) = x' f(C u_2)$, $f(v) = y' f(C u_2)$ avec $x' C = f(xC)$ et $y' C = f(yC)$. Or $x' \xrightarrow[G']{g^*} y'$; donc $f(u) \xrightarrow[G']{g^*} f(v)$.

Prenons $W \in L(G)$. D'après le lemme IV.2, il existe pour W une dérivation la plus à gauche $A \xrightarrow[G]{g^*} W$ et en utilisant la propriété précédente, nous obtenons : $A = f(A) \xrightarrow[G']{g^*} f(W) = W$ donc $L(G') \subseteq L(G)$.

Réciproquement, si $(x', y') \in P'_1 \cup P'_2$, il existe $C \in N_1^0$, $x \in N$, $y \in V^*$ tels que $x' = f(xC)$, $y' = f(yC)$ et $(x, y) \in P$.

Nous en déduisons, en raisonnant de la même façon qu'en (a) et (b), la propriété :

- (3) Si $f(u) \xrightarrow[P'_1 \cup P'_2]{g} v'$ avec $u \in N^+$, alors $\exists v \in V^+$ tel que $u \xrightarrow[G]{g} v$ et $f(v) = v'$.

En effet, posons $u = B u_1$ avec $B \in N$.

- Si $u_1 \in \{\Lambda\} \cup N_2 N^*$, $f(u) = B f(u_1)$, $v' = f(y) f(u_1)$ avec $(B, y) \in P$.
Donc $v' = f(y u_1)$ et $B u_1 \xrightarrow[G]{g} y u_1$.
- Si $u_1 \in N_1 N^*$. Posons $u_1 = C u_2$ avec $C \in N_1$,
 $f(u) = (B, C) f(C u_2)$, $v' = y' f(C u_2)$ avec $((B, C), y') \in P'_2$ donc $y' C = f(yC)$
avec $(B, y) \in P$, $v' = f(yC u_2)$ et $u = BC u_2 \xrightarrow[G]{g} y C u_2$.

Montrons maintenant la propriété suivante :

$$(4) \left\{ \begin{array}{l} f(u) \xrightarrow[G']{g^*} y' \xrightarrow[G']{g} ax' \\ \text{avec} \\ u \in N^+, y' \in N'^+, a \in T \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} \exists x, y \in N^* \text{ tels que } f(x) = x' \\ \text{et} \\ u \xrightarrow[G]{g^*} y \xrightarrow[G]{g} ax \end{array} \right.$$

Raisonnons par récurrence sur la longueur i de la dérivation :

- Si $i=1$, comme $f(u) \in N'^+$ et $a \in T$, la règle utilisée $\in P'_1 \cup P'_2$ et d'après la propriété (3), il existe $v \in V^+$ tel que $u \xrightarrow[G]{g} v$ avec $f(v) = ax'$, donc $v = ax$ avec $f(x) = x'$ et $u \xrightarrow[G]{g^*} y = u \xrightarrow[G]{g} ax$.
- Supposons la propriété vraie pour $i \leq \ell$ et considérons une dérivation de longueur $\ell + 1$:

$$f(u) \xrightarrow[G']{g} y'_1 \xrightarrow[G']{g} y'_2 \xrightarrow[G']{g^*} ax'.$$

Soit π la première règle utilisée. Distinguons deux cas :

- (i) Si $\pi \in P'_1 \cup P'_2$, d'après la propriété (3), il existe $u_1 \in V^+$ tel que $f(u_1) = y'_1$ et $u \xrightarrow[G]{g} u_1$. Comme $y'_1 \in N'^+$, $u_1 \in N^+$ et nous obtenons alors la propriété (4) en utilisant l'hypothèse de récurrence pour la dérivation $y'_1 = f(u_1) \xrightarrow[G']{g^*} ax'$.
- (ii) Si $\pi \in P'_3$. Alors $\pi = ((A_2, A_1), \Lambda)$ avec $A_2 \in N_2$, $A_1 \in N_1$ et $A_2 A_1 \in P_g(G)$; u se factorise en $A_2 A_1 u_1$, $f(u) = (A_2, A_1) f(A_1 u_1)$ et $y'_1 = f(A_1 u_1)$. Comme $A_1 \in N_1$, $y'_1 \notin N'_2 N'^*$ donc la règle utilisée pour passer de y'_1 à y'_2 appartient à $P'_1 \cup P'_2$.

D'après la propriété (3), il existe $u_2 \in V^+$ tel que $f(u_2) = y'_2$ et $A_1 u_1 \xrightarrow[G]{g} u_2$ donc u_2 se factorise en $v_1 u_1$ avec $(A_1, v_1) \in P$.

Comme G vérifie la propriété (U) et que $A_2 A_1 \in P_g(G)$, $(A_2 A_1, v_1) \in P$, donc $A_2 A_1 u_1 \xrightarrow[G]{g} v_1 u_1 = u_2$.

Si $y'_2 = ax'$, alors $u_2 = av_2$ et il suffit de prendre $u = y = A_2 A_1 u_1$ et $x = v_2$.

Si $y'_2 \neq ax'$, alors $u_2 \in N^+$ et il suffit d'appliquer l'hypothèse de récurrence à $y'_2 = f(u_2) \xrightarrow[G']{g^*} y' \xrightarrow[G]{g} ax'$.

De plus, comme G vérifie la propriété (D), le mot x obtenu dans la propriété (4) ne dépend que de u et de a .

Considérons les dérivations :

$$A \xrightarrow{G'}^{g^*} W y_i' \xrightarrow{G'}^g W a x_i' \text{ avec } W \in T^*, a \in T, x_i', y_i' \in N'^* \text{ pour } i \in [1,2].$$

D'après la propriété (4), il existe x_i et $y_i \in N^*$ tels que $A \xrightarrow{G}^{g^*} W y_i \xrightarrow{G}^g W a x_i$ avec $x_i' = f(x_i)$ pour $i \in [1,2]$. Mais x_1 et x_2 ne dépendant que de A , de W et de a , nous obtenons $x_1 = x_2$ et $x_1' = x_2'$, donc G' est permutable.

Soit $W \in L(G')$. Comme $A \xrightarrow{G'}^{g^*} W'y'$ avec $W' \in T^*$ et $y' \in N'^+$ implique que $y' \in N'^*N$, $\Lambda \notin L(G')$. Donc W se factorise en $W'a$ avec $a \in T$ et il existe une dérivation :

$$A \xrightarrow{G'}^{g^*} W'y' \xrightarrow{G'}^g W'a.$$

D'après la propriété (4) $A \xrightarrow{G}^{g^*} Wy \xrightarrow{G}^g W'a$ avec $y \in N^+$.

Donc $L(G') \subseteq L(G)$ et les grammaires G et G' sont équivalentes.

Réciproquement, nous obtenons :

Lemme IV.11 Soit $G = (N, T, P, A)$ une grammaire prédictive. Pour $d \notin T$, il existe une UAP3-grammaire qui engendre $L(G) \setminus \{d\}$.

Démonstration Notons $N_2 = \{B / (B, \Lambda) \in P\}$, $N_1 = N \setminus N_2$, $V = N \cup T$, $V_2 = N_2 \cup T$ et $V_1 = N_1 \cup T$. Nous pouvons supposer que P est inclus dans $N \times (V^* \setminus V^2 V^* N_2 V^*) [WA]$, donc qu'il existe $P_0 = N_2 \times \{\Lambda\}$ et $P_1 \subset N \times T N N_1^*$ tels que $P = P_0 \cup P_1$.

G vérifie alors la propriété :

(1) Si $A \xrightarrow[G]{*} x A_2 y$ avec $A_2 \in N_2$ alors $x \in \{\Lambda\} \cup V^*T$.

Montrons le par récurrence sur la longueur i de la dérivation .

Si $i=0$ c'est évident.

Supposons la propriété (1) vraie pour $i \leq \ell$ et considérons une dérivation de longueur $\ell+1$

$$A \xrightarrow[G]{*} u_1 u u_2 \xrightarrow[G]{} u_1 v u_2 = x A_2 y \text{ avec } (u,v) \in P.$$

•

Distinguons trois cas :

- $u_1 = x A_2 u'_1$. Il suffit alors d'appliquer l'hypothèse de récurrence à la dérivation $A \xrightarrow[G]{*} x A_2 u'_1 u u_2$.

- $u_2 = u'_2 A_2 y$. D'après l'hypothèse de récurrence, $u_1 u u'_2 \in \{\Lambda\} \cup V^*T$. Or $u \in N$ ce qui entraîne $u'_2 \in V^*T$, donc $u_1 v u'_2 = x \in V^*T$.

- $v = v_1 A_2 v_2$. Alors $(u,v) \in P_1$ et $v_1 \in T$.

Prenons $A', F, d \notin V$ et considérons la grammaire $G' = (N', T', P', A')$ avec $N' = N \cup \{A', F\}$, $T' = T \cup \{d\}$, et $P' = P_1 \cup \{(A_2 A_1, x) / A_2 \in N_2, A_1 \in N_1 \text{ et } (A_1, x) \in P_1\} \cup \{(A_2 F, d) / A_2 \in N_2\} \cup \{(A', xF) / (A, x) \in P\} \cup \{(F, d)\}$.

Posons $V' = N' \cup T'$. Soit h l'homomorphisme de V'^* dans V^* défini par $h(d) = h(F) = \Lambda$, $h(A') = A$ et $h(B) = B$, $\forall B \in V$. Comme $(u', v') \in P'$ implique $h(u') \xrightarrow[G]{*} h(v')$ et que $h(A') = A$ et $h(T'^*) = T^*$, nous obtenons $h(L(G')) \subseteq L(G)$. De plus, il est clair que $L(G') \subseteq T^*\{d\}$, donc $L(G') \subseteq L(G) \{d\}$.

Montrons maintenant la propriété :

$$(2) \left\{ \begin{array}{l} A \xrightarrow[G]{g^*} Wxu \xrightarrow[G]{g} Wayu \\ \text{avec } a \in T, W \in T^* \text{ et } (x, ay) \in P \end{array} \right\} \longrightarrow A' \xrightarrow[G']{g^*} WayuF$$

Raisonnons par récurrence sur la longueur i de W .

Si $i=0$, la première règle utilisée appartient à P_1 , sinon $xu = \Lambda$, ce qui est impossible car $x \in P_g(G)$. Donc $x = A$, $u = \Lambda$, $(A, ay) \in P$ et $(A', ayF) \in P'$.

Supposons la propriété vraie pour $i \leq l$ et considérons la dérivation de $Wayu$ avec $l(W) = l+1$, où nous exhibons l'apparition de b , la dernière lettre de W ($W=W'b$) :

$$A \xrightarrow[G]{g^*} W'x_1u'' \xrightarrow[G]{g} W'bu' \xrightarrow[G]{g^*} Wxu \xrightarrow[G]{g} Wayu$$

avec $x_1 \in N$, $(x, ay) \in P$ et $u' \xrightarrow[G]{g^*} xu$.

D'après l'hypothèse de récurrence, $A' \xrightarrow[G']{g^*} W'bu'F$.

Distinguons deux cas :

① $u' = xu$. Comme $(x, ay) \in P'$, nous obtenons $A' \xrightarrow[G']{g^*} W'bu'F = W'bxuF \xrightarrow[G']{g} W'bayuF$.

② $u' \neq xu$. Alors u' se factorise en $u'_1u'_2$ avec $u'_1 \in N$. Comme $u'_1u'_2 \xrightarrow[G]{g^*} xu$ avec $xu \neq u'_1u'_2$ et $xu \in N^+$, $(u'_1, \Lambda) \in P_0$ et $u'_1 \in N_2$. D'après la propriété (1), comme $A \xrightarrow[G]{g^*} W'bu'_1u'_2$, $u'_2 \notin N_2V^*$, donc $u'_2 = xu$ avec $x \in N_1$. Alors $(x, ay) \in P$ implique $(u'_1x, ay) \in P'$. Nous obtenons :

$$A' \xrightarrow[G']{g^*} W'bu'_1xuF \xrightarrow[G']{g} W'bayuF = WayuF.$$

Soit $W \in L(G)$.

Si $W = \Lambda$, $(A, \Lambda) \in P$ donc $A' \xrightarrow[G']{g^*} F \longrightarrow d$ et $d \in L(G')$.

Si $W \neq \Lambda$, $W = W'a$ avec $a \in T$ et il existe une dérivation :

$$A \xrightarrow[G]{g^*} W'u \xrightarrow[G]{g} W'av \xrightarrow[G]{g^*} W'a = W$$

donc $v \xrightarrow[G]{g^*} \Lambda$ et $v \in N_2^*$ mais, d'après la propriété (1), $v \in N_2^0$ et $vF \xrightarrow[G']{g^*} d$.

En utilisant la propriété (2), nous obtenons :

$$A' \xrightarrow[G']{g^*} W'avF \xrightarrow[G']{g^*} W'ad = Wd$$

donc $L(G) \setminus \{d\} \subseteq L(G')$ et $L(G) \setminus \{d\} = L(G')$.

De plus G' est une AP3-grammaire vérifiant la propriété (U) par construction.

Soit $G'' = (N, T, P'', A')$ l'AP3-grammaire en N_2 pré-réduite équivalente à G' (cf. lemme IV.3).

G'' vérifie encore la propriété (U). En effet, si $A_2 A_1 \in P_g(G'')$ avec $A_2 \in N_2$ et $A_1 \in N_1$, A dépend de $A_2 A_1$, donc de A_1 et $(A_2 A_1, x) \in P'' \iff (A_2 A_1, x) \in P' \iff (A_1, x) \in P' \iff (A_1, x) \in P''$.

Il nous reste à montrer que G'' vérifie la propriété (D).

Comme G est une grammaire prédictive, nous avons par construction $|D_G(B, a)| \leq 1$, $\forall B \in N$ et $\forall a \in T$, donc $|D_{G''}(B, a)| \leq 1$ ($P'' \subset P'$).

De même $\forall A_2 \in N_2, A_1 \in N_1, a \in T$, il existe au plus un élément x de N^* tel que $(A_2 A_1, ax) \in P''$. Supposons que G'' ne vérifie pas la propriété (D). D'après ce qui précède, c'est qu'il existe $A_2 \in N_2, A_1 \in N_1, a \in T$ tels que $|D_{G''}(A_2 A_1, a)| > 1$, ce qui implique alors qu'il existe $x_1, x_2 \in N^*$ tels que $(A_2 A_1, ax_1) \in P''$ et $(A_2, ax_2) \in P''$ avec $x_2 A_1 \neq x_1$. Alors, comme $A_2 A_1 \in P_g(G'')$, A' dépend de $A_2 A_1$ dans G' et il existe $u, v \in V'^*$ tels que $A' \xrightarrow[G']{*} u A_2 A_1 v$ ce qui implique $A \xrightarrow[G]{*} h(u) A_2 A_1 h(v)$.

Or $(A_2 A_1, ax_1) \in P''$ implique $(A_2, \Lambda) \in P$ et $(A_1, ax_1) \in P$; de même $(A_2, ax_2) \in P$. Nous obtenons, alors, dans G , la dérivation :

$A \xrightarrow[G]{*} h(u) A_2 A_1 h(v) \xrightarrow[G]{*} h(u) A_2 ax_1 h(v)$, donc $a \in S_G(A_2)$ ce qui est contradictoire puisque G est prédictive et que (A_2, Λ) et (A_2, ax_2) sont des règles de P .

Comme les langages $LL(1)$ sont caractérisés par les grammaires permutables ainsi que par les grammaires prédictives et que pour $L \subset T^*$ et $d \notin T$, $L \setminus \{d\}$ est un langage $LL(1)$ si et seulement si L est un langage $LL(1)$ [WA], les lemmes IV.10 et IV.11, nous donnent le résultat suivant :

Proposition IV.4

$L \subset T^*$ est un langage LL(1) si et seulement si, pour $d \notin T$, il existe une UAP1 (UAP3)-grammaire qui engendre $L\{d\}$.

Etudions maintenant les AP3-grammaires déterministes, pour montrer qu'elles n'engendrent pas tous les C-langages déterministes, mais seulement les langages reconnus par un automate complet déterministe, Λ -libre, suivis d'un marqueur.

Lemme IV.12

Soient $L \subset T^*$ un langage reconnu par un automate complet déterministe Λ -libre et $d \notin T$. Alors $L\{d\}$ est engendré par une AP3-grammaire déterministe.

Démonstration

Soit $M = (Q, T, Z, g, z_0, q_0, F)$ un automate complet, déterministe, Λ -libre tel que Q, T, Z soient disjoints deux à deux et tel que $T(M) = L$ et $d \notin Q \cup Z$.

Prenons $A \notin Q \cup T \cup Z$ et posons $N_2 = Q \cup \{A\}$, $N_1 = Z$, $N = N_1 \cup N_2$ et $P = P_0 \cup P_1 \cup P_2$ avec :

- $P_0 = \{(A, aq\tilde{x}) / a \in T, q \in Q \text{ et } g(q_0, a, z_0) = (q, x)\} \cup \{(A, d) / \text{si } q_0 \in F\}$,
- $P_1 = \{(qz, aq'\tilde{y}) / a \in T, q, q' \in Q, z \in Z \text{ et } g(q, a, z) = (q', y)\}$,
- $P_2 = \{(qz_0, d) / q \in F\}$.

Considérons $G = (N, T \cup \{d\}, P, A)$.

Il est clair que G est une AP3-grammaire déterministe et que pour tout $q, q' \in Q, z \in Z, a \in T, y \in Z^*$,

$$(q, a, z) \vdash_M (q', \Lambda, y) \iff qz \xrightarrow[G]{g} aq'\tilde{y}.$$

Nous en déduisons immédiatement la propriété :

$$(1) \quad \left\{ \begin{array}{l} \forall q, q' \in Q, \forall x, y \in Z^*, \forall W \in T^*, \\ (q, W, x) \vdash_M^* (q', \Lambda, y) \iff q\tilde{x} \xrightarrow[G]{g^*} Wq'\tilde{y} \end{array} \right.$$

Prenons $W \in T(M)$.

- Si $W = \Lambda$, $q_0 \in F$ et $A \xrightarrow[G]{} d$, donc $d \in L(G)$.

- Si $W \neq \Lambda$, W peut se factoriser en aW' avec $a \in T$ et

$$(q_0, aW', z_0) \vdash_M (q, W', x) \vdash_M^* (q', \Lambda, y) \text{ avec } q' \in F \text{ et } y = z_0 \text{ car } M \text{ est complet.}$$

Nous obtenons, alors, en utilisant la propriété (1) :

$$A \xrightarrow[P_0]{} aq\tilde{x} \xrightarrow[G]{g^*} aW'q'z_0 \xrightarrow[P_2]{} aW'd \text{ donc } L\{d\} \subset L(G).$$

Réciproquement si $W \in L(G)$, $W = W'd$ avec $W' \in T^*$.

- Si $W' = \Lambda$, $(A, d) \in P$, donc $q_0 \in F$ et $W' \in T(M)$.

- Si $W' \neq \Lambda$, $W' = aW''$ avec $a \in T$ et $A \xrightarrow[P_0]{} aq\tilde{x} \xrightarrow[P_1]{g^*} aW''q'z_0 \xrightarrow[P_2]{} aW''d$ avec

$(q'z_0, d) \in P_2$, donc $q' \in F$, $(A, aq\tilde{x}) \in P_0$, donc $g(q_0, a, z_0) = (q, x)$. Nous obtenons alors en utilisant la propriété (1) :

$$(q_0, aW'', z_0) \vdash_M (q, W'', x) \vdash_M^* (q', \Lambda, z_0)$$

et $L(G) = L\{d\}$.

Pour montrer la réciproque, nous utiliserons un lemme qui donne une condition suffisante pour qu'un automate à pile déterministe soit équivalent à un automate à pile déterministe Λ -libre.

Lemme IV.13

Soit $M = (Q, T, Z, g, z_0, q_0, F)$ un automate à pile déterministe.

Si M vérifie la propriété :

$$(E) \quad \left\{ \begin{array}{l} \exists n \text{ tel que } \forall (q,zy) \text{ accessible dans } M \text{ avec } q \in Q, z \in Z \text{ et } y \in Z^n, \\ (q, \Lambda, zy) \mid_M^* (q', \Lambda, z) \text{ implique } g(q', \Lambda, z) = \emptyset, \end{array} \right.$$

alors il existe un automate M' déterministe Λ -libre tel que $T(M') = T(M)$. De plus si M est complet, M' l'est aussi.

Démonstration

Raisonnons par récurrence sur n .

Si $n=0$ alors, pour tout $q \in Q$ et $z \in Z, (q,z)$ accessible implique $g(q, \Lambda, z) = \emptyset$. Il suffit de prendre $g(q, \Lambda, z) = \emptyset$ et $g(q, a, z) = (q, z)$ pour tout $a \in T, q \in Q$ et $z \in Z$ tels que (q, z) n'est pas accessible dans M .

Supposons le lemme vrai pour $n \leq l$ et considérons M vérifiant la propriété (E) avec $n=l+1$.

Posons $Z' = Z \cup \{z'_0\}$ avec $z'_0 \notin Z, y_0 = (z'_0 z'_0 z_0), Y = \{(u) / u \in Z^3\} \cup \{y_0\}$ et définissons les applications f de $Z^* \cup \{z'_0 z'_0 z_0\} Z^*$ dans Y^* et r de $Z^* \cup \{z'_0 z'_0 z_0\} Z^*$ dans Z^0 par :

- $\forall u \in Z^0, f(u) = r(u) = \Lambda,$
- $\forall z_1, z_2 \in Z, f(z_1 z_2) = \Lambda \text{ et } r(z_1 z_2) = z_2,$
- $\forall z_1, z_2, z_3 \in Z, \forall u \in Z^*, f(z_1 z_2 z_3 u) = (z_1, z_2, z_3) f(z_3 u) \text{ et } r(z_1 z_2 z_3 u) = r(z_3 u),$
- $\forall u \in Z^*, f(z'_0 z'_0 z_0 u) = y_0 f(z_0 u) \text{ et } r(z'_0 z'_0 z_0 u) = r(z_0 u).$

Posons $\bar{Q} = \{\bar{q} / q \in Q\}$ et $Q' = Q \cup \bar{Q}$.

$$\forall q \in Q, \forall u \in Z^*, \forall z \in Z, \text{ notons } s(q, uz) = \begin{cases} \bar{q} \text{ s'il existe } q_1 \in F, u_1 \in Z^* \text{ tels que} \\ (q, \Lambda, z) \mid_M^* (q_1, \Lambda, u_1) \\ q \text{ sinon.} \end{cases}$$

Notons h l'homomorphisme de Q'^* dans Q^* tel que : $\forall q \in Q, h(q) = h(\bar{q}) = q$
 et t l'homomorphisme de Y^* dans Z^* défini par : $t(y_0) = z_0$ et $\forall z \in Z,$
 $\forall u \in Z^2, t((zu)) = u.$

Considérons l'automate $M' = (K, T, Y, g', y_0, q'_0, F')$ avec $K = Q'Z^0 \cup \{k_p\}$
 (k_p est un nouveau symbole), $q'_0 = s(q_0, z_0)$, $F' = \bar{Q}Z^0$ et g' définie de la manière
 suivante :

Soient $q \in Q, x \in Z^0$ et $y \in Y.$

- Si $(q, t(y)x)$ n'est pas accessible dans M , nous poserons alors :

$$g'(qx, \Lambda, y) = \emptyset \text{ et } g'(qx, a, y) = (k_p, y) \quad \forall a \in T.$$

- Dans le cas contraire, posons $y = (v)$, $vx = uz$ avec $z \in Z$ et distinguons trois cas :

a) Il n'existe pas $q_1 \in Q, u_1 \in Z^*$ tels que $(q, \Lambda, z) \vdash_M^* (q_1, \Lambda, u_1).$

Nous poserons alors :

$$g'(qx, \Lambda, y) = \emptyset \text{ et } g'(qx, a, y) = (k_p, y) \quad \forall a \in T$$

b) $(q, \Lambda, z) \vdash_M^* (q', \Lambda, u'z')$ avec $z' \in Z$ (donc $g(q', \Lambda, z') = \emptyset$).

Nous poserons alors : $g'(qx, \Lambda, y) = \emptyset$

$$\text{et } \forall a \in T \quad g'(qx, \Lambda, y) = (k''r(u'u''), f(u'u''))$$

$$\text{avec } g(q', a, z') = (q'', u'') \text{ et } k'' = s(q'', t(f(u'u''))) r(u'u'').$$

c) $(q, \Lambda, z) \vdash_M^* (q', \Lambda, \Lambda).$ u se factorise en $u'z'$ avec $z' \in Z.$

$$\text{alors } \forall b \in T^0, \quad g'(qx, b, y) = \begin{cases} - \emptyset \text{ si } g(q', b, z') = \emptyset \\ - (k''r(u'u''), f(u'u'')) \text{ avec} \\ \quad k'' = s(q'', t(f(u'u''))) r(u'u'') \\ \quad \text{si } g(q', b, z') = (q'', u'') \end{cases}$$

Achevons la définition de g' en posant :

- $\forall q \in Q, \forall x \in Z^0, \forall y \in Y, \forall b \in T^0, g'(\bar{q}x, b, y) = g'(qx, b, y),$

- $g'(k, \Lambda, y) = \emptyset$ et $\forall a \in T, g'(k, a, y) = (k_p, y)$ pour tout $k \in K$ et $y \in Y$ pour
 lesquels g' n'a pas encore été définie.

Il est clair que M' est déterministe. Montrons, maintenant que

$$T(M) = T(M').$$

Pour $q' \in Q'$, $x \in Z^0$, $W \in T^*$, $y \in Y^+$ notons

$$H(q'x, W, y) = (h(q'), W, t(y)x).$$

Il est facile de vérifier la propriété :

$$(1) \quad \left\{ \begin{array}{l} \text{Si } (k, b, y) \vdash_M (k', \Lambda, y') \text{ avec } b \in T^0, k, k' \in K \setminus \{k_p\}, y, y' \in Y^+ \\ \text{alors } H(k, b, y) \vdash_M^* H(k', \Lambda, y'). \end{array} \right.$$

Prenons $W \in T(M')$. Nous avons alors :

$(q'_0, W, y_0) \vdash_M^* (k, \Lambda, y)$ avec $k \in \bar{Q}Z^0$. Il existe donc $q_1 \in F$ et $u_1 \in Z^*$ tels que

$H(k, \Lambda, y) \vdash_M^* (q_1, \Lambda, u_1)$ et, en utilisant la propriété (1), nous obtenons :

$$(q_0, W, z_0) = H(q'_0, W, y_0) \vdash_M^* H(k, \Lambda, y) \vdash_M^* (q_1, \Lambda, u_1) \text{ et } T(M') \subseteq T(M).$$

Réciproquement pour $q \in Q$, $W \in T^*$, $z \in Z$ et u tel que

$uz \in Z^* \cup \{z'_0 z'_0 z_0\} Z^*$, notons

$$G(q, W, uz) = (s(q, z) r(uz), W, f(uz)).$$

Par définition de g' , l'automate M' vérifie la propriété :

$$(2) \quad \left\{ \begin{array}{l} \text{Si } (q, \Lambda, z) \vdash_M^{d*} (q', \Lambda, u') \text{ et } (q', b, uu') \vdash_M (q'', \Lambda, u'') \text{ avec } q \in Q, b \in T^0, \\ z \in Z, uz, uu' \in Z^3 Z^* \text{ alors, pour tout } W \in T^*, G(q, Wb, uz) \vdash_M G(q'', W, u''). \end{array} \right.$$

Prenons $W \in T(M)$.

Si $W = \Lambda$, il existe $q_1 \in F$, $u_1 \in Z^*$ tels que $(q_0, \Lambda, z_0) \vdash_M^* (q_1, \Lambda, u_1)$

donc $q'_0 = s(q_0, z_0) = \bar{q}_0 \in F'$ et $\Lambda \in T(M')$.

Si $W \neq \Lambda$, nous avons alors, dans M , le mouvement :

(i) $(q_0, W, z_0) \mid_M^* (q, b, uz) \mid_M^* (q_1, b, uu_1) \mid_M (q_2, \Lambda, u_2 z_2)$ avec $b \in T^0$, $z, z_2 \in Z$,
 $(q, \Lambda, z) \mid_M^{d*} (q_1, \Lambda, u_1)$ et $s(q_2, z_2) = \bar{q}_2$.

En ajoutant $z'_0 z'_0$ en base de pile et en posant $u'_2 = z'_0 z'_0 u_2$, nous obtenons, d'après la propriété (2) :

$$(q'_0, W, y_0) = G(q_0, W, z'_0 z'_0 z_0) \mid_M^* G(q_2, \Lambda, u'_2 z_2) = (\bar{q}_2 r(u'_2 z_2), \Lambda, f(u'_2 z_2))$$

donc $W \in T(M')$ et $T(M') = T(M)$.

En outre, si M est complet, il existe $p \in F$ tel que $(q_0, W, z_0) \mid_M^* (p, \Lambda, z_0)$.
 Donc dans le mouvement (i), nous pouvons supposer que $(q_2, \Lambda, z_2) \mid_M^* (p, \Lambda, v_2)$ avec
 $u_2 v_2 = z_0$. Distinguons deux cas :

- Si $u_2 = \Lambda$ alors $z_2 = z_0$ et $G(q_2, \Lambda, u'_2 z_0) = (\bar{q}_2, \Lambda, y_0)$.
- Si $u_2 \neq \Lambda$ alors $u_2 = z_0$ et $G(q_2, \Lambda, u'_2 z_2) = (\bar{q}_2 z_2, \Lambda, y_0)$.

Nous en déduisons que, si $W \in T(M')$, alors $W \in T(M)$ et il existe $k \in F$ tel que $(q'_0, W, y_0) \mid_M^* (k, \Lambda, y_0)$, donc M' est complet.

Il reste à montrer que M' vérifie la propriété (E) avec $n \leq \ell$.

Supposons qu'il existe $k, k' \in K$, $y \in Y$, $u \in Y^\ell$ tels que (k, yu) soit accessible dans M' et $(k, \Lambda, yu) \mid_M^* (k', \Lambda, y)$ avec $g'(k', \Lambda, y) \neq \emptyset$.

$H(k, \Lambda, yu)$ et $H(k', \Lambda, y)$ peuvent s'écrire respectivement $(q, \Lambda, zz't(u)x)$ et $(q', \Lambda, zz'x')$ avec $q, q' \in Q$, $z, z' \in Z$, $x, x' \in Z^0$, $t(y) = zz'$, $k = qx$ ou $\bar{q}x$ et $k' = q'x'$ ou $\bar{q}'x'$.

D'après la propriété (1), $(q, zz't(u)x)$ est accessible dans M . Distinguons deux cas :

- $\ell=0$. Comme $g(k', \Lambda, y) \neq \emptyset$, nous avons, d'après c),

si $x' = \Lambda$, (q', Λ, zz') $\mid_M^*$ (q'', Λ, z) avec $g(q'', \Lambda, z) \neq \emptyset$,

si $x' \in Z$, $(q', \Lambda, zz'x')$ $\mid_M^*$ (q'', Λ, zz') avec $g(q'', \Lambda, z') \neq \emptyset$,

ce qui est impossible, puisque M vérifie la propriété (E) avec $n=1$.

- $\ell \geq 1$. Comme $g(k', \Lambda, y) \neq \emptyset$, nous avons $(q, \Lambda, zz't(u)x) \mid_M^* (q', \Lambda, zz'x') \mid_M^* (q'', \Lambda, zz')$ avec $g(q'', \Lambda, z') \neq \emptyset$, ce qui est impossible puisque $\ell(t(u)) = 2\ell \geq \ell + 1$.

Lemme IV.14

Pour toute AP3-grammaire déterministe G, il existe un

automate M complet déterministe Λ -libre tel que $T(M) = L(G)$.

Démonstration

Soit $G = (N, T, P, A)$ une AP3-grammaire déterministe en N_2 , telle

que $P \subset N_2^0 N_1^0 \times TN^*$.

Prenons $z_0, q, q_1, q_2 \notin N$ et considérons l'automate

$M = (Q, T, Z, g, z_0, A, \{q_1\})$ avec $Q = N \cup \{q, q_1, q_2\}$,

$Z = N \cup \{z_0\}$ et g définie par :

$$\begin{aligned}
 - \forall a \in T, \forall x \in N, g(x, a, z_0) &= \begin{cases} (q, z_0 \tilde{u}) & \text{si } D_G(x, a) = \{u\} \\ (q_2, z_0) & \text{si } D_G(x, a) = \emptyset \end{cases} \\
 - \forall a \in T, \forall x, y \in N, g(x, a, y) &= \begin{cases} (q, \tilde{u}) & \text{si } D_G(xy, a) = \{u\} \\ (q_2, y) & \text{si } D_G(xy, a) = \emptyset \end{cases} \\
 - \forall z \in Z, g(q, \Lambda, z) &= \begin{cases} (z, \Lambda) & \text{si } z \neq z_0 \\ (q_1, z_0) & \text{si } z = z_0 \end{cases}
 \end{aligned}$$

- $\forall a \in T, \forall z \in Z, g(q_1, a, z) = g(q_2, a, z) = (q_2, z)$ et $g(q, a, z) = \emptyset$
- $\forall p \in Q \setminus \{q\}, \forall z \in Z, g(p, \Lambda, z) = \emptyset$.

Comme G vérifie la propriété (D), M est déterministe. Montrons que $T(M) = L(G)$ et que M est complet.

Par construction de g , nous avons les propriétés :

- (1) Si $u \xrightarrow[G]{g} av$ avec $a \in T, u, v \in N^+$, alors $(q, a, z_0 \tilde{u}) \mid_{\overline{M}}^* (q, \Lambda, z_0 \tilde{v})$.
- (2) Si $(q, a, z_0 u) \mid_{\overline{M}} (x, a, z_0 u') \mid_{\overline{M}} (q, \Lambda, z_0 v)$ avec $u \in N^*$ et $a \in T$, alors $v \in N^*$ et $\tilde{u} \xrightarrow[G]{g} a\tilde{v}$.

Soit $W \in L(G)$. $W \neq \Lambda$ puisque G est Λ -libre. Donc W se factorise en aW' avec $a \in T$ et nous avons, dans G , la dérivation $A \xrightarrow[G]{g} au \xrightarrow[G]{g^*} aW'$.

Comme $(A, au) \in P$ $(A, a, z_0) \mid_{\overline{M}} (q, \Lambda, z_0 \tilde{u})$.

Et en utilisant la propriété (1), nous obtenons :

$$(A, aW', z_0) \mid_{\overline{M}} (q, W', z_0 \tilde{u}) \mid_{\overline{M}}^* (q, \Lambda, z_0) \mid_{\overline{M}} (q_1, \Lambda, z_0), \quad \text{donc } W \in T(M).$$

Réciproquement, prenons $W \in T(M)$. Comme $A \neq q_1$ et que $\forall z \in Z, g(A, \Lambda, z) = \emptyset$, W est différent de mot vide et se factorise en aW' avec $a \in T$. Nous avons donc, dans M , le mouvement suivant :

$$(A, aW', z_0) \mid_{\overline{M}} (q, W', z_0 u) \mid_{\overline{M}}^* (q, \Lambda, vz_0) \mid_{\overline{M}} (q_1, \Lambda, vz_0)$$

Mais d'après la propriété (2) $vz_0 \in \{z_0\} N^*$ (donc $v = \Lambda$ et M est complet) et $\tilde{u} \xrightarrow[G]{g^*} W'$.

Nous obtenons alors $A \xrightarrow[G]{g} a\tilde{u} \xrightarrow[G]{g^*} aW' = W$,

et $T(M) = L(G)$.

De plus $(p, \Lambda, zy) \mid_{\overline{M}}^* (p', \Lambda, z)$ avec $y \in Z^+$ implique $p=q$ et $p' = y \in N$

donc $g(p', \Lambda, z) \neq \emptyset$ et d'après le lemme IV.13 il existe un automate M' déterministe complet Λ -libre tel que $T(M') = T(M) = L(G)$.

Lemme IV.15 Soient $L \subset T^*$ et $d \notin T$. Si $L\{d\}$ est reconnu par un automate complet déterministe Λ -libre, alors L l'est aussi.

Démonstration Soit $M = (Q, T \cup \{d\}, Z, g, z_0, q_0, F)$ un automate complet déterministe Λ -libre qui reconnaît $L\{d\}$.

Posons $Q' = QZ^0$, $F' = \{qz / q \in Q, z \in Z, g(q, d, z) = (q_1, u_1) \text{ avec } q_1 \in F\}$ et notons $h(z) = z$, $\forall z \in Z \setminus \{z_0\}$ et $h(z_0) = \Lambda$.

Considérons l'automate $M' = (Q', T, Z, g', q_0, z_0, F')$ avec g' définie de la manière suivante :

- $\forall q \in Q, \forall z \in Z,$

$$g'(q, \Lambda, z) = \begin{cases} (qz, \Lambda) & \text{si } z \neq z_0 \\ (qz_0, z_0) & \text{sinon} \end{cases}$$

et $\forall a \in T, g'(q, a, z) = \emptyset$

- $\forall q \in Q, \forall z, z' \in Z,$

$$g'(qz, \Lambda, z') = \emptyset$$

et $\forall a \in T, g'(qz, a, z') = (q', h(z')u)$ avec $(q', u) = g(q, a, z)$.

L'automate M' est bien déterministe et vérifie les deux propriétés suivantes :

(1) Si $(q, a, y) \mid_{\overline{M}} (q', \Lambda, y')$ avec $a \in T, y, y' \in Z^*$ alors $(q, a, y) \mid_{\overline{M}}^* (q', \Lambda, y')$

(2) Si $(qz, a, y) \mid_{\overline{M}} (q', \Lambda, y'z')$ $\mid_{\overline{M}} (q'z', \Lambda, y'')$ avec $q, q' \in Q, z, z' \in Z, a \in T$ et $y, y', y'' \in Z^*$ alors $(q, a, yh(z)) \mid_{\overline{M}} (q', \Lambda, y''h(z'))$.

Prenons $W \in L$. Comme $Wd \in T(M)$, nous avons dans M le mouvement suivant :

$$(q_0, Wd, z_0) \mid_{\overline{M}}^* (q, d, yz) \mid_{\overline{M}} (q_1, \Lambda, z_0) \text{ avec } z \in Z \text{ et } q_1 \in F.$$

$$\text{D'après la propriété (1)} \quad (q_0, W, z_0) \mid_{\overline{M}}^* (q, \Lambda, yz).$$

Distinguons deux cas :

- Si $y = \Lambda$, $z = z_0$ et $(q, \Lambda, z_0) \mid_{\overline{M}} (qz_0, \Lambda, z_0)$ avec $qz_0 \in F'$
- Si $y \neq \Lambda$, $y = z_0$ et $(q, \Lambda, z_0 z) \mid_{\overline{M}} (qz, \Lambda, z_0)$ avec $qz \in F'$.

Donc $L \subseteq T(M')$ et M' est complet.

Réciproquement, si $W \in T(M')$, nous avons

$$(q_0, W, z_0) \mid_{\overline{M}} (q_0 z_0, W, z_0) \mid_{\overline{M}}^* (qz, \Lambda, y) \text{ avec } qz \in F', \text{ donc}$$

$$g(q, d, z) = (q_1, u_1) \text{ avec } q_1 \in F.$$

D'après la propriété (2)

$$(q_0, Wd, z_0) \mid_{\overline{M}}^* (q, d, yh(z)).$$

$$\text{Si } z \neq z_0, (q, d, yh(z)) = (q, d, yz) \mid_{\overline{M}} (q_1, \Lambda, y_1) \text{ avec } q_1 \in F.$$

$$\text{Si } z = z_0, \text{ alors } y = y'z_0 \text{ et}$$

$$(q, d, yh(z)) = (q, d, y'z_0) \mid_{\overline{M}} (q_1, \Lambda, y_1) \text{ avec } q_1 \in F,$$

$$\text{donc } T(M') = L.$$

En outre, si $g'(q', \Lambda, z) \neq \emptyset$, alors $q' \in Q$ et

$$g'(q', \Lambda, z) = (q'z, u) \text{ avec } g'(q'z, \Lambda, z') = \emptyset \quad \forall z' \in Z.$$

Donc, d'après le lemme IV.15, il existe un automate M'' déterministe complet Λ -libre tel que $T(M'') = T(M') = L$.

Des lemmes IV.12, IV.14 et IV.15, nous déduisons immédiatement le résultat suivant :

Proposition IV.5 $L \subset T^*$ est reconnu par un automate complet déterministe Λ -libre si et seulement si, pour $d \in T$, $L\{d\}$ est engendré par une AP3-grammaire déterministe.

Chapitre V

LB-GRAMMAIRES ET AUTOMATES LINEAIREMENT BORNES DETERMINISTES

Nous allons nous intéresser, dans ce chapitre, à la classe L'_1 des langages reconnus par un automate linéairement borné déterministe. L'_1 forme une algèbre de Boole contenant les C-langages [KU] et incluse dans la classe L_1 des langages contextuels [LA]. Mais l'inclusion stricte n'a pu être démontrée jusqu'à présent.

Nous donnerons une caractérisation grammaticale de L'_1 en introduisant les LB-grammaires qui engendrent exactement les langages de L_1 , et en ajoutant à ces grammaires des conditions de déterminisme pour obtenir les LB-grammaires déterministes.

Pour démontrer l'équivalence entre L'_1 et la classe des langages engendrés par ces grammaires, nous utiliserons quelques résultats relatifs aux transformations qui conservent L'_1 . Nous montrerons d'abord, qu'à l'instar de L_1 , L'_1 est clos pour les homomorphismes inverses et les homomorphismes Λ -libres. Nous généraliserons ensuite ces résultats en établissant que l'image transduite par un transducteur séquentiel Λ -libre d'un langage de L'_1 , appartient encore à L'_1 .

Considérons un alphabet T contenant une lettre c et l'homomorphisme h_c de T^* dans $(T \setminus \{c\})^*$ défini par $h_c(c) = \Lambda$ et $h_c(a) = a$ pour tout $a \in T \setminus \{c\}$.

Lemme V.1 Soit $L \subset (T \setminus \{c\})^*$. Si L appartient à L'_1 , alors $h_c^{-1}(L)$ appartient aussi à cette classe.

Démonstration Soit $M = (Q, V, q_0, g, F)$ un automate linéairement borné déterministe qui reconnaît L avec $c \notin V$. Nous pouvons supposer que M est sans boucle [KU] et sans marqueur [GH]. A Q , faisons correspondre biunivoquement $\bar{Q} = \{\bar{q} / q \in Q\}$. Posons $Q' = Q \cup \bar{Q}$, $V' = V \cup \{c\}$, étendons h_c à V'^* en posant $h_c(v) = v$ pour tout $v \in V$ et considérons l'automate $M' = (Q', V', q_0, g', F)$ où g' est définie de la manière suivante :

$\forall q \in Q \text{ et } v \in V :$

$$- g'(\bar{q}, v) = (q, v, 0)$$

$$- g'(q, v) = \begin{cases} (q_1, v_1, i) & \text{si } g(q, v) = (q_1, v_1, i) \text{ avec } i = 0 \text{ ou } 1 \\ (\bar{q}_1, v_1, -1) & \text{si } g(q, v) = (q_1, v_1, -1) \end{cases}$$

$$- g'(q, c) = (q, c, +1) \text{ et } g'(\bar{q}, c) = (\bar{q}, c, -1)$$

Montrons que $T(M') = h_c^{-1}(L)$. Il est facile de vérifier que, pour tout mouvement simple dans M : $(q, x, y) \vdash_M^* (q_1, x_1, y_1)$ et pour tout x', y' appartenant respectivement à $h_c^{-1}(x)$ et $h_c^{-1}(y)$, il existe $x'_1 \in h_c^{-1}(x_1)$ et $y'_1 \in h_c^{-1}(y_1)$ tels que $(q, x', y') \vdash_{M'}^* (q_1, x'_1, y'_1)$. Nous en déduisons immédiatement que si $(q_0, \Lambda, u) \vdash_M^* (q, v, \Lambda)$, alors, pour tout $u' \in h_c^{-1}(u)$, il existe v'_1 et $p \geq 0$ tels que $(q_0, \Lambda, u') \vdash_{M'}^* (q, v'_1, c^p)$. Comme $u' \in h_c^{-1}(h_c(u'))$ et que $g'(q, c) = (q, c, +1)$, nous obtenons, pour tout $u' \in V'^*$, la propriété suivante :

$$(1) \quad (q_0, \Lambda, h_c(u')) \vdash_M^* (q, y, \Lambda) \text{ implique } \exists z \text{ tel que } (q_0, \Lambda, u') \vdash_{M'}^* (q, z, \Lambda)$$

De plus $h_c(u') \in V^*$ et M est sans boucle. Donc, pour tout $u' \in V'^*$, il existe $q \in Q$, $y \in V^*$ tels que $(q_0, \Lambda, h_c(u')) \vdash_M^* (q, y, \Lambda)$ avec $q \in F$ si et seulement si $h_c(u') \in L$ ce qui implique d'après (1) $u' \in T(M')$ si et seulement si $h_c(u') \in L$ et $T(M') = h_c^{-1}(L)$.

Lemme V.2 Soient T et T' deux alphabets, h un homomorphisme Λ -libre de T'^* dans T^* et L une partie de T^* . Si $L \in L_1'$ alors $h^{-1}(L)$ appartient aussi à L_1' .

Démonstration Soit $M = (Q, V, q_0, g, F)$ un automate linéairement borné déterministe sans marqueur et sans boucle qui reconnaît L .

Posons $\bar{Q} = \{\bar{q} / q \in Q\}$, $Q'' = Q \cup \bar{Q}$, $\ell = \sup_{a' \in T'} \ell(h(a'))$,

$V_0 = \{(v) / v \in V^k \text{ avec } 1 \leq k \leq \ell\}$ et $V' = V_0 \cup T'$.

Comme M est sans boucle, pour tout $q \in Q$ et tout $u, v \in V^*$, nous avons :

- soit $(q,u,v) \models_M^* (q',u',\Lambda)$, dans ce cas, nous poserons $U(q,u,v) = (q',u',+1)$.
- soit $(q,u,v) \models_M^* (q'',\Lambda,ZV')$ avec $z \in V$ et $g(q'',z) = (q',z',-1)$; nous poserons alors $U(q,u,v) = (\bar{q}',(z'v'),-1)$.

Soient q'_0, q'_1 et $q'_2 \notin Q''$. Posons $Q' = Q'' \cup \{q'_0, q'_1, q'_2\}$ et considérons l'automate $M' = (Q', V', q'_0, g', F)$ avec les marqueurs $\emptyset$ et $\$$, où g' est définie de la façon suivante :

$$\forall q \in Q, (v) \in V_0 \text{ et } a' \in T'$$

- $g'(q, (v)) = U(q, \Lambda, v)$
- $g'(\bar{q}, (v)) = U(q, v_1 \dots v_{k-1}, v_k)$ si $v = v_1 \dots v_k$
avec $v_j \in V$ pour $j \in [1, k]$
- $g'(q'_0, a') = (q'_0, (h(a')), +1)$
- $g'(q'_0, \$) = (q'_1, \$, -1)$
- $g'(q'_1, (v)) = (q'_1, (v), -1)$
- $g'(q'_1, \emptyset) = (q'_0, \emptyset, +1)$
- $g'(q', \$) = (q', \$, +1)$ pour $q' \in Q' \setminus \{q'_0\}$
- $g'(q', v') = (q'_2, v', +1)$ pour tout $q' \in Q'$ et tout $v' \in V'$

pour lesquels g' n'est pas encore définie.

Par construction M' est un automate linéairement borné déterministe sans boucle et $T(M') \subseteq T'^*$.

Montrons que $T(M') = h^{-1}(L)$.

Soit t l'homomorphisme de V_0^* dans V^* défini par $t((v)) = v$ pour tout $(v) \in V_0$.

Pour $q \in Q, u'$ et $v' \in V_0^*$, notons :

$$H(q, u', v') = (q, t(u'), t(v'))$$

$$H(\bar{q}, u', v') = \begin{cases} (q, t(u')v_1 \dots v_{k-1}, v_k t v'') \text{ si } v' = (v_1 \dots v_k) v'' \text{ avec } v_j \in V \\ \text{pour } j \in [1, k] \\ (q, \Lambda, \Lambda) \text{ si } v' = \Lambda. \end{cases}$$

Il est facile de vérifier que $(q', u', v') \vdash_{M'}^* (q'_1, u'_1, v'_1)$ avec $q' \in Q''$, $u', v' \in V_0^*$ implique $q'_1 \in Q''$, $u'_1, v'_1 \in V_0^*$ et $H(q', u', v') \vdash_M^* H(q'_1, u'_1, v'_1)$.

Nous en déduisons la propriété :

$$(2) \quad \begin{cases} \text{si } (q_0, \Lambda, v') \vdash_{M'}^* (q, y, \Lambda) \text{ avec } v' \in V_0^*, \text{ alors } q \in Q, y \in V_0^* \text{ et} \\ (q_0, \Lambda, t(v')) = H(q_0, \Lambda, v') \vdash_M^* H(q, y, \Lambda) = (q, t(y), \Lambda). \end{cases}$$

Prenons $y' \in T'^*$. y' se factorise en $y'_1 \dots y'_n$ avec $y'_j \in T'$ pour $j \in [1, n]$.

Comme M' est sans boucle, nous avons dans M' les mouvements suivants :

$$\begin{aligned} (q'_0, \emptyset, y' \$) &\vdash_{M'}^* (q_0, \emptyset, (h(y'_1)) \dots (h(y'_n)) \$), \\ (q_0, \Lambda, (h(y'_1)) \dots (h(y'_n))) &\vdash_{M'}^* (q, z, \Lambda) \text{ avec } z \in V_0^n \text{ et} \\ (q, \emptyset z, \$) &\vdash_{M'} (q, \emptyset z \$, \Lambda) \end{aligned}$$

D'après la propriété (2), nous obtenons, dans M :

$$(q_0, \Lambda, h(y')) \vdash_M^* (q, t(z), \Lambda).$$

Si $y' \in T(M')$, $q \in F$ et $h(y') \in L$ donc $y' \in h^{-1}(L)$.

Si $y' \notin T(M')$, $q \in Q \setminus F$ et $h(y') \notin L$ donc $y' \notin h^{-1}(L)$.

$T(M')$ et $h^{-1}(L)$ sont inclus dans T'^* ; nous en déduisons que $T(M') = h^{-1}(L)$.

Comme tout homomorphisme peut être obtenu par composition d'un homomorphisme Λ -libre et d'homomorphismes du type h_c défini pour le lemme V.1, nous obtenons :

Proposition V.1 L'_1 est clos pour les homomorphismes inverses.

L'_1 n'est pas clos en général pour les homomorphismes. En effet tout ensemble récursivement énumérable est l'image homomorphe de l'intersection de deux C-langages déterministes [GGH page 405], donc, aussi l'image homomorphe d'un langage de L'_1 . Montrons cependant, à l'aide des deux lemmes suivants, que L'_1 est clos pour les homomorphismes Λ -libres.

Lemme V.3 L'_1 est clos pour les homomorphismes strictement alphabétiques.

Démonstration Soient deux alphabets T et T' et un homomorphisme h de T^* dans T'^* , strictement alphabétique. Nous pouvons supposer h surjectif. En effet dans le cas contraire, il suffirait de remplacer T' par $h(T)$ qui est un alphabet inclus dans T' . Soit $L \subset T^*$ reconnu par un automate linéairement borné déterministe sans boucle et sans marqueur, $M = (Q, V, q_0, g, F)$.

Posons $T' = \{a'_1, \dots, a'_p\}$, $n_i = |h^{-1}(a'_i)|$ pour $i \in [1, p]$ et $\ell = \sup_{i \in [1, p]} n_i$. Comme h est surjectif, tous les n_i sont strictement positifs. Numérotons les éléments des ensembles $h^{-1}(a'_i)$. Pour $j \in [1, n_i]$, $h_j^{-1}(a'_i)$ désignera le $j^{\text{ème}}$ élément de $h^{-1}(a'_i)$.

Soient $q'_0, q'_1, q'_2, q'_3 \notin Q$. Notons $Q' = Q \bigcup_{i=0}^3 \{q'_i\}$, $V_0 = T' \times V \times [1, \ell]$ et $V' = T' \cup V_0$.

Considérons l'automate linéairement borné déterministe avec les marqueurs $\emptyset$ et $\$$, $M' = (Q', V', q'_0, g', F)$ où g' est définie par :

- $g'(q'_0, a'_i) = (q'_0, (a'_i, h_1^{-1}(a'_i), 1), +1)$ pour $i \in [1, p]$
- $g'(q'_0, \$) = (q'_1, \$, -1)$
- $g'(q'_1, (a'_i, v, j)) = (q'_1, (a'_i, h_j^{-1}(a'_i), j), -1)$ pour $(a'_i, v, j) \in V_0$
- $g'(q'_1, \emptyset) = (q_0, \emptyset, +1)$
- $g'(q, (a'_i, v, j)) = (q_1, (a'_i, v_1, j), s)$ pour $q \in Q$ et $(a'_i, v, j) \in V_0$ tels que $g(q, v) = (q_1, v_1, s)$
- $g'(q, \$) = \begin{cases} (q, \$, +1) & \text{si } q \in F \\ (q'_2, \$, -1) & \text{si } q \notin F \end{cases}$
- $g'(q'_2, (a'_i, v, j)) = \begin{cases} (q'_1, (a'_i, h_{j+1}^{-1}(a'_i), j+1), -1) & \text{si } j < n_i \\ (q'_2, (a'_i, h_1^{-1}(a'_i), 1), -1) & \text{si } j = n_i \end{cases}$ pour $(a'_i, v, j) \in V_0$
- $g'(q', v') = (q'_3, v', +1)$ pour tout $q' \in Q'$ et $v' \in V'$ pour lesquels g' n'est pas encore définie.

Il est facile de voir que M' est sans boucle et que $T(M') \subset T'^*$. Montrons maintenant que $T(M') = h(L)$.

Posons $X_0 = \{(a_i, j) / i \in [1, p], j \in [1, n_i]\}$ et $X_1 = \{(a_i, n_i) / i \in [1, p]\}$.

Soient t_2 , l'homomorphisme de V_0^* dans V^* défini par $t((a_i', v, j)) = v$ et f , l'homomorphisme de X_0^* dans V_0^* défini par $f((a_i', j)) = (a_i', h_j^{-1}(a_i'), j)$. Pour $y = a_{i_1}' \dots a_{i_k}'$, notons $n(y) = |h^{-1}(y)| = \sum_{\ell=1}^k n_{i_\ell}$ et définissons, par récurrence $s(y, j)$ avec $j \in \mathbb{N}$:

- $s(y, 1) = u_1' \dots u_k'$ avec $u_t' = (a_{i_t}', 1)$ pour $t \in [1, k]$
- si $s(y, j) \in X_1^*$ alors $s(y, j+1) = s(y, j)$
- sinon $s(y, j)$ se factorise en $v'(a_{i_t}', j) v''$ avec $v' \in X_0^{t-1}$, $j < n_{i_t}$ et $v'' \in X_1^*$, nous poserons alors $s(y, j+1) = v'(a_{i_t}', j+1) u_{t+1}' \dots u_k'$.

Notons $G(y, j) = f(s(y, j))$.

A partir de t_2 et G , nous obtenons, pour tout $y \in T'^*$, une numérotation bijective de $h^{-1}(y)$. C'est-à-dire que pour $j \in [1, n(y)]$, $t_2(G(y, j)) \in h^{-1}(y)$ et pour tout $x \in h^{-1}(y)$, il existe $j \in [1, n(y)]$ tel que $x = t_2(G(y, j))$.

Prenons $y \in T'^*$. Par construction de g' , et comme M est sans boucle, nous avons dans M' les mouvements suivants :

- (1) $(q_0', \varnothing, y\$) \xrightarrow{*}_{M'} (q_0, \varnothing, G(y, 1)\$)$
- (2) $(q_0, \varnothing, G(y, j)\$) \xrightarrow{*}_{M'} (q, \varnothing z\$, \Lambda)$ avec $q \in F$ si $j \leq n(y)$ et si $t_2(G(y, j)) \in L$
- (3) $(q_0, \varnothing, G(y, j)\$) \xrightarrow{*}_{M'} (q_0, \varnothing, G(y, j+1)\$)$ si $j < n(y)$ et $t_2(G(y, j)) \notin L$
- (4) $(q_0, \varnothing, G(y, n(y))\$) \xrightarrow{*}_{M'} (q_3', \varnothing z'\$, \Lambda)$ si $t_2(G(y, n(y))) \notin L$.

- Si $y \in h(L)$, il existe $j \in [1, n(y)]$ tel que $t_2(G(y, j)) \in L$. Notons j_0 le plus petit j vérifiant cette propriété. Nous avons dans M' le mouvement suivant :

$$(q_0', \varnothing, y\$) \xrightarrow{*}_{(1)} (q_0, \varnothing, G(y, 1)\$) \xrightarrow{*}_{(3)} (q_0, \varnothing, G(y, j_0)\$)$$

$\vdash_{(2)}^* (q, \emptyset z \$, \Lambda)$ avec $q \in F$, donc $y \in T(M')$.

- Si $y \in T'^* \setminus h(L)$, pour tout $j \in [1, n(y)]$, $t_2(G(y, j)) \notin L$.

Nous avons donc dans M' :

$$\begin{aligned} (q'_0, \emptyset, y \$) &\vdash_{(1)}^* (q_0, \emptyset, G(y, 1) \$) \vdash_{(3)}^* (q_0, \emptyset, G(y, n(y)) \$) \\ &\vdash_{(4)}^* (q'_3, \emptyset z \$, \Lambda), \text{ donc } y \in T'^* \setminus h(L) \end{aligned}$$

et nous obtenons bien $h(L) = T(M')$.

Considérons un homomorphisme h Λ -libre de T^* dans T'^* et h_1 l'homomorphisme strictement alphabétique de T^* dans T'^* , construit à partir de h , en prenant pour tout $a \in T$, $h_1(a) = a'$ si $h(a) = a'x$ avec $a' \in T'$. Nous dirons que h est **1-injectif** si et seulement si h_1 est injectif.

Lemme V.4 L'_1 est clos pour les homomorphismes 1-injectifs.

Démonstration Soient $C \notin T$ et $M = (Q, V, g, q_0, F)$ un automate linéairement borné déterministe sans boucle, sans marqueur, qui reconnaît $h_C^{-1}(L)$ (cf. Lemme V.1) avec $Q \cap T'^* = \emptyset$.

Prenons $q'_0, q'_1, q'_2 \notin Q$ et posons $\ell = \sup_{a \in T} \ell(h(a)) - 1$, $Q_0 = \bigcup_{i=1}^{\ell} T^i$,
 $Q' = Q \cup \{q'_0, q'_1, q'_2\} \cup Q_0$, $V' = V \cup T' \cup \{c\}$.

Considérons l'automate avec marqueurs,

$M' = (Q', V', q'_0, g', F)$ avec g' définie par :

$$- \forall a' \in T', g'(q'_0, a') = \begin{cases} (q'_0, a, +1) & \text{si } h(a) = a' \\ (x', a, +1) & \text{si } h(a) = a'x' \text{ avec } x' \in T'^+ \end{cases}$$

$$- \forall x' \in Q_0, a' \in T', g'(x', a') = \begin{cases} (x'', c, +1) & \text{si } x' = a'x'' \text{ avec } x'' \in T'^+ \\ (q'_0, c, +1) & \text{si } x' = a' \end{cases}$$

$$- g'(q'_0, \$) = (q'_1, \$, -1)$$

$$- \forall z \in T \cup \{c\}, g'(q'_1, z) = (q'_1, z, -1)$$

$$- g'(q'_1, \emptyset) = (q_0, \emptyset, +1)$$

$$- \forall q \in Q, v \in V, g'(q, v) = g(q, v)$$

$$- \forall q \in Q \quad g'(q, \$) = (q, \$, +1)$$

- $g'(q', v') = (q'_2, v', +1)$ pour tout $q' \in Q'$, $v' \in V'$ pour lesquels g' n'est pas encore définie.

Comme h est 1-injectif et que M est sans boucle, M' est bien déterministe sans boucle.

Montrons que $T(M') = h(L)$.

Prenons $y' \in h(T)$. Alors il existe un seul $x \in T$ tel que $y' = h(x)$ (h est 1-injectif).

Il est facile de vérifier, par construction de g' , que nous avons dans M' le mouvement suivant :

$$(q'_0, \varnothing, y' \$) \stackrel{*}{\vdash}_{M'} (q'_0, \varnothing', z \$) \text{ avec } h_c(z) = x.$$

Et comme M est sans boucle, $(q_0, \Lambda, z) \stackrel{*}{\vdash}_M (q, z', \Lambda)$. Si $y' \in h(L)$, $x \in L$ donc $z \in h_c^{-1}(L)$ et $q \in F$ ce qui implique $y' \in T(M')$.

Si $y' \in h(T) \setminus h(L)$ alors $x \notin L$ donc $z \notin h_c^{-1}(L)$ et $q \notin F$, ce qui implique $y' \notin T(M')$.

Pour terminer la démonstration, il suffit de remarquer que si $y' \notin h(T)$ alors $(q'_0, \varnothing, y' \$) \stackrel{*}{\vdash}_{M'} (q'_2, \varnothing z \$, \Lambda)$ et donc $y' \notin T(M')$.

Ces deux lemmes permettent d'énoncer :

Proposition V.2 L'_1 est clos pour les homomorphismes Λ -libres.

Démonstration Soit h un homomorphisme Λ -libre de T^* dans T'^* . Posons $T'' = T' \cup T \times T'$ et définissons les homomorphismes h'' de T^* dans T''^* et h' de T''^* dans T'^* par :

$$- \forall a \in T, h''(a) = (a, a') \times' \text{ si } h(a) = a' \times' \text{ avec } a' \in T'$$

$$- h'((a, a')) = a' \text{ pour tout } (a, a') \in T \times T' \text{ et } h'(a') = a' \text{ pour tout } a' \in T'$$

Il est clair que h'' est 1-injectif, que h' est strictement alphabétique et que $h = h'oh''$. Donc les lemmes V.3 et V.4 impliquent immédiatement la propriété énoncée.

Les propositions V.1 et V.2 vont nous servir, pour trouver un résultat plus général : l'image transduite d'un langage de L'_1 , par un transducteur séquentiel Λ -libre, appartient à L'_1 .

Pour obtenir ce résultat, particularisons la notion de K-transducteur introduite par Nivat [Ni]. Nous dirons qu'une K-transduction $T_K = \{(h_1(v), h_2(v)) / v \in R\}$ est Λ -libre si et seulement si h_2 est un homomorphisme Λ -libre.

Cette définition permet d'énoncer le résultat suivant :

Lemme V.5 T_K est K-transduction Λ -libre si et seulement si T_K est la transduction associée à un transducteur séquentiel Λ -libre.

Démonstration Prenons T_K une K-transduction Λ -libre de X^* dans Y^* . Il existe donc un alphabet Z , un K-langage $R \subset Z^*$ et des homomorphismes h_1 et h_2 respectivement de Z^* dans X^* et de Z^* dans Y^* avec h_2 Λ -libre, tels que

$$T_K = \{(h_1(v), h_2(v)) / v \in R\}.$$

Soit $M = (Q, Z, q_0, g, F)$ l'automate d'état fini qui reconnaît R . Prenons $q_f \notin Q$, posons $Q' = Q \cup \{q_f\}$ et considérons le transducteur séquentiel :

$$T_S = (Q', X, Y, q_0, H, q_f) \text{ avec } H = \{(q, h_1(v), h_2(v), g(q, v)) / q \in Q, v \in Z\}$$

$$\cup \{(q, h_1(v), h_2(v), q_f) / q \in Q, v \in Z \text{ et } g(q, v) \in F\}$$

Comme h_2 est Λ -libre, H est inclus dans $(Q' \times X^* \times Y^+ \times Q')$ et T_S est Λ -libre. Appelons S la transduction associée à T_S .

Soit $(x, y) \in T_K$. Il existe alors $v \in R$ tel que $x = h_1(v)$ et $y = h_2(v)$. v se factorise en $v_1 \dots v_n$ avec $v_i \in Z$ pour $i \in [1, n]$ et il existe $q_1, \dots, q_n \in Q$ tels que $g(q_i, v_{i+1}) = q_{i+1}$ pour $i \in [0, n-1]$ et $q_n \in F$, ce qui implique $(q_i, h_1(v_{i+1}), h_2(v_{i+1}), q_{i+1}) \in M$ pour $i \in [0, n-2]$ et $(q_{n-1}, h_1(v_n), h_2(v_n), q_f) \in H$.

Donc $(q_0, h_1(v), \Lambda) \xrightarrow{T_S^*} (q_f, \Lambda, h_2(v))$ et $(x, y) = (h_1(v), h_2(v)) \in S$.

Prenons maintenant $(x, y) \in S$. Il existe, alors, $x_1, \dots, x_n \in X^*$ et $y_1, \dots, y_n \in Y^+$ tels que $(q_i, x_{i+1}, y_{i+1}, q_{i+1}) \in H$, $q_i \in Q$ pour $i \in [0, n-1]$ et $q_n = q_f$. Il existe donc pour $i \in [0, n-1]$, $v_i \in Z$ vérifiant $h_1(v_i) = x_i$, $h_2(v_i) = y_i$, $g(q_i, v_{i+1}) = q_{i+1}$ si $i \neq n-1$ et $g(q_{n-1}, v_n) \in F$. Donc $v \in R$, $x = h_1(v)$, $y = h_2(v)$ et $(x, y) \in T_K$.

Réciproquement, considérons un transducteur séquentiel Λ -libre

$T_S = (Q, X, Y, q_0, H, q_f)$. Soient h_1 et h_2 les homomorphismes de H^* dans, respectivement, X^* et Y^* , définis par :

pour tout $W = (q, u, v, q') \in H$, $h_1(W) = u$ et $h_2(W) = v$.

Comme H est inclus dans $(Q \times X^* \times Y^+ \times Q)$ h_2 est Λ -libre.

Considérons l'automate d'état fini $M = (Q, H, q_0, g, \{q_f\})$ où g est défini par $g(q, (q, u, v, q')) = q'$ pour tout $q \in Q$ et tout $(q, u, v, q') \in H$. Soit R le K -langage reconnu par M . Montrons que la transduction S associée à T_S est égale à $T_K = \{(h_1(e), h_2(e)) / e \in R\}$. Prenons $(u, v) \in S$. Il existe alors $q_1, \dots, q_n \in Q$, $u_1, \dots, u_n \in X^*$ et $v_1 \dots v_n \in Y^+$ tels que $e_{i+1} = (q_i, u_{i+1}, v_{i+1}, q_{i+1}) \in H$ pour $i \in [0, n-1]$, $u = u_1 \dots u_n$, $v = v_1 \dots v_n$ et $q_n = q_f$.

Posons $e = e_1 \dots e_n$. $g(q_0, e) = q_f$, donc $e \in R$ et comme $h_1(e) = u$, $h_2(e) = v$, $(u, v) \in T_K$.

Prenons maintenant $(u, v) \in T_K$. Il existe alors $e \in R$ tel que $u = h_1(e)$ et $v = h_2(e)$, donc $e_1, \dots, e_n \in H$ et $q_1, \dots, q_n \in Q$ tels que $e = e_1 \dots e_n$, $q_n = q_f$ et $g(q_i, e_{i+1}) = q_{i+1}$ pour $i \in [0, n-1]$ ce qui implique, par définition de g , $e_{i+1} = (q_i, h_1(e_{i+1}), h_2(e_{i+1}), q_{i+1})$, donc $(u, v) = (h_1(e), h_2(e)) \in S$.

Considérons un transducteur séquentiel Λ -libre T_S et S la transduction de X^* dans Y^* associée à T_S . Pour $L \subset X^*$, l'image transduite par T_S , $f_S(L)$ est égale à $\{y \in Y^* / (x,y) \in S \text{ et } x \in L\}$. D'après le lemme précédent, il existe un alphabet fini Z , un K -langage $R \subset Z^*$, des homomorphismes h_1 de Z^* dans X^* , h_2 Λ -libre de Z^* dans Y^* tels que $S = \{(h_1(v), h_2(v)) / v \in R\}$. Donc $f_S(L) = \{h_2(v) / v \in R \text{ et } h_1(v) \in L\} = h_2(R \cap h_1^{-1}(L))$. Des propositions V.1 et V.2 nous déduisons immédiatement :

Proposition V.3 L'image transduite par un transducteur séquentiel Λ -libre, d'un élément de L'_1 appartient à L'_1 .

Introduisons maintenant un nouveau type de grammaire qui va nous permettre de donner une caractérisation grammaticale des langages de L'_1 . Nous dirons qu'une grammaire $G = (N, T, P, A)$ est une **LB-grammaire** en N_1, N_2, N_3 si et seulement si $\{N_1, N_2, N_3\}$ forme une partition de N et si $P = P_1 \cup P_2 \cup P_3$ avec $P_1 \subset N_1 \times V^+$, $P_2 \subset (N_2 N_3 \cup N_3 N_2) \times (N_2 N_3 \cup N_3 N_2)$ et $P_3 \subset N_3^* N_2 \times N_2^0 T^+$.

Dans ce cas, pour $y \in V^+$, nous noterons $R_2(y) = \{z \in V^* / y \xrightarrow{P_2} z\}$ c'est-à-dire l'ensemble des mots obtenus à partir de y en utilisant une règle de P_2 .

Nous dirons que G est une **LB-grammaire déterministe** en N_1, N_2, N_3 si, en plus, les trois propriétés suivantes sont vérifiées :

- (1) $A \xrightarrow{P_1}^* W$ implique $\ell_{N_2}(W) \leq 1$
- (2) $\forall A_3, B_3 \in N_3, E_2 \in N_2, |R_2(A_3 E_2 B_3)| \leq 1$
- (3) Si $E_2 \in N_2$ figure dans une règle de P_3 , $\forall A_3 \in N_3, R_2(A_3 E_2) = \emptyset$

Remarque Les conditions (2) et (3) sont décidables car elles ne concernent qu'un nombre fini de mots.

Notons $L_1 = \{W / A \xrightarrow{P_1}^* W\}$ et $R = (V \setminus N_2)^* N_2^0 (V \setminus N_2)^*$

L_1 est un C -langage et R un K -langage. La condition (1) correspond à $L_1 \subset R$ c'est-à-dire à $L_1 \cap (V^* \setminus R) = \emptyset$ ce qui est aussi décidable.

Nous dirons que G est une LB-grammaire (déterministe) si et seulement si il existe une partition de N en $\{N_1, N_2, N_3\}$ telle que G soit une LB-grammaire (déterministe) en N_1, N_2, N_3 .

Nous obtenons une première propriété :

Proposition V.4 Tout langage de L_1 (resp. L'_1) est engendré par une LB-grammaire (resp. LB-grammaire déterministe).

Démonstration Soit $M = (Q, X, q_0, g, F)$ un automate linéairement borné. Prenons $A \notin Q \cup X$, posons $\bar{Q} = \{\bar{q} / q \in Q\}$, $N_1 = \{A\}$, $N_2 = Q \cup \bar{Q}$, $N_3 = X \times X$, $P_1 = \{(A, C(x, x)) / x \in X \text{ et } C = A \text{ ou } q_0\}$ et $P_3 = \{((x, y)q, q'x) / x, y \in X, q \in F \text{ et } q' = q \text{ ou } \Lambda\}$.

Construisons P_2 de la manière suivante :

- Si $g(q, x) = (q', x', +1)$, alors pour tout $y \in X$,

$$(q(y, x), (y, x')q') \in P_2 \text{ et } ((y, x)\bar{q}, (y, x')q')' \in P_2$$

- Si $g(q, x) = (q', x', i)$ avec $i = 0$ ou -1 , posons $q'' = \begin{cases} q' & \text{si } i = 0 \\ \bar{q}' & \text{si } i = -1 \end{cases}$

Alors, pour tout $y \in X$, $(q(y, x), q''(y, x'))$ et $((y, x)\bar{q}, q''(y, x'))$ appartiennent à P_2 .

Il est facile de vérifier que la grammaire $G = (N, X, P, A)$ avec $N = N_1 \cup N_2 \cup N_3$ et $P = P_1 \cup P_2 \cup P_3$ est bien une LB-grammaire. Montrons que $L(G) = T(M)$.

Définissons les homomorphismes t_1 et t_2 de $(N_2 \cup N_3)^*$ dans X^* , par $t_1(q) = t_1(\bar{q}) = t_2(q) = t_2(\bar{q}) = \Lambda$ pour tout $q \in Q$ et $t_1((x, y)) = x$, $t_2((x, y)) = y$ pour tout $x, y \in X$.

Pour $q \in Q$ et $u, v \in X^*$. Posons :

$$H(q, u, v) = \begin{cases} - \{sq / s \in N_3^*, t_2(s) = u\} & \text{si } v = \Lambda \\ - \{s_1qs_2 / s_1, s_2 \in N_3^*, t_2(s_1) = u, t_2(s_2) = v\} \\ \cup \{s_1\bar{q}s_2 / s_1, s_2 \in N_3^*, t_2(s_1) = uv, t_2(s_2) = v'\} \\ & \text{si } v = v_1v' \text{ avec } v_1 \in X \end{cases}$$

Cette notation nous permet d'énoncer la propriété suivante :

Si $(q, u, v) \vdash_M^* (q', u', v')$ alors pour tout $z \in H(q, u, v)$ il existe $z' \in H(q, u', v')$ tel que $z \xrightarrow{P_2}^* z'$ et $t_1(z') = t_1(z)$.

Il suffit, en effet, de vérifier, en examinant P_2 , que cette propriété est vraie pour tout mouvement simple. Prenons $W \in T(M)$. Il existe alors un mouvement dans M : $(q_0, \Lambda, W) \vdash_M^* (q, W', \Lambda)$ avec $q \in F$.

Nous avons dans G la dérivation suivante :

$$A \xrightarrow{P_1}^* q_0 W_1 \text{ avec } t_1(W_1) = t_2(W_1) = W.$$

$q_0 W_1 \in H(q_0, \Lambda, W)$, donc il existe $W_2 q \in H(q, W', \Lambda)$ vérifiant $t_1(W_2 q) = t_1(q_0 W_1) = W$, $q_0 W_1 \xrightarrow{P_2}^* W_2 q$ et, comme $q \in F$, $W_2 q \xrightarrow{P_3}^* t_1(W_2) = W$.

Nous obtenons donc $A \xrightarrow{P_1}^* q_0 W_1 \xrightarrow{P_2}^* W_2 q \xrightarrow{P_3}^* W$ et $W \in L(G)$.

Pour tout $W, W' \in N_3^*$, $W_1 \in N_3$ et $q \in Q$ notons :

$$H'(WqW') = (q, t_2(W), t_2(W'))$$

$$H'(WW_1 \bar{q} W') = (q, t_2(W), t_2(W_1 W')).$$

Il est facile de vérifier que si $z_1 \xrightarrow{P_2}^* z_2$ avec z_1 et $z_2 \in N_3^* N_2 N_3^* \setminus \bar{Q} N_3^*$ alors $H'(z_1) \vdash_M^* H'(z_2)$.

Prenons $W \in L(G)$. Il existe dans G , une dérivation de la forme

$$A \xrightarrow{P_1}^* q_0 W_0 \xrightarrow{P_2}^* W_1 q \xrightarrow{P_3}^* W, \text{ avec } t_1(W_0) = t_2(W_0) = t_1(W_1) = W, q \in F$$

$$\text{et } (q_0, \Lambda, W) = H'(q_0 W_0) \vdash_M^* H'(W_1 q) = (q, t_2(W_1), \Lambda).$$

Donc $W \in T(M)$ et $L(G) = T(M)$.

La grammaire G vérifie les propriétés (1) et (3) de déterminisme. Si M est déterministe, la propriété (2) est aussi vérifiée ; G est alors une LB-grammaire déterministe.

Pour toute LB-grammaire G , $P_N = \{(u,v) \in P / \ell_T(v) = 0\}$ est inclus dans $P_1 \cup P_2$. Donc toutes les règles de P_N sont de longueur croissante et d'après le corollaire III.1, $L(G)$ est un langage contextuel.

La proposition V.4 nous permet alors d'énoncer :

Proposition V.5 La classe des langages engendrés sur une LB-grammaire est équivalente à la classe L_1 .

Montrons maintenant que tout langage engendré par une LB-grammaire déterministe, appartient à L_1 . Nous utiliserons trois résultats intermédiaires.

Lemme V.6 Soit $M = (Q, V, q_0, g, F)$ un automate linéairement borné déterministe avec les marqueurs $\varnothing$ et $\$$, tel que $V \cap Q = \emptyset$. Il existe un automate $M' = (Q', V', q'_0, g', \{q'_f\})$ linéairement borné déterministe avec les marqueurs $\varnothing$ et $\$$, vérifiant pour tout $u, v \in V^*$ et tout $q \in Q$ la propriété :

$$(A) \left\{ \begin{array}{l} (q, \varnothing u, v\$) \xrightarrow{*}_M (q_1, \varnothing y\$, \Lambda) \\ \text{avec } q_1 \in F \end{array} \right\} \iff (q'_0, \varnothing, uqv\$) \xrightarrow{*}_{M'} (q'_f, \varnothing yq_1\$, \Lambda)$$

Démonstration Posons $V' = V \cup Q$, $Q' = Q \cup V \cup \{q'_0, q'_1, q'_f, q'_p\}$ avec $V' \cap \{q'_0, q'_1, q'_f, q'_p\} = \emptyset$ et considérons l'automate linéairement borné déterministe $M' = (Q', V', q'_0, g', \{q'_f\})$ où g' est définie par :

$$(1) \left\{ \begin{array}{l} - g'(q'_0, v') = (q'_0, v', +1) \quad \forall v' \in V' \\ - g'(q'_0, \$) = (q'_1, \$, 0) \end{array} \right.$$

$$(2) \left\{ \begin{array}{l} - g'(q'_1, \$) = (q'_1, \$, -1) \\ - \text{pour tout } v' \in V' \quad g'(q'_1, v') = (v', q_0, i) \text{ avec } i = \begin{cases} 0 & \text{si } v' \in Q \\ -1 & \text{si } v' \in V \end{cases} \\ - \text{pour tout } v \in V \text{ et tout } v' \in V' \quad g'(v, v') = (v', v, i) \text{ avec } i = \begin{cases} 0 & \text{si } v' \in Q \\ -1 & \text{si } v' \in V \end{cases} \end{array} \right.$$

$$(3) \quad \begin{array}{l} - \text{ pour tout } q \in Q \quad g'(q, v) = g(q, v) \\ \text{et tout } v \in V \cup \{\emptyset\} \end{array}$$

$$(4) \quad \begin{array}{l} - \text{ pour tout } q \in Q \quad g'(q, q_0) = \begin{cases} (q'_f, q_1, +1) & \text{si } g(q, \$) = (q_1, \$, +1) \\ & \text{avec } q_1 \in F \\ (q_2, q_0, i) & \text{si } g(q, \$) = (q_2, \$, i) \\ & \text{avec } i=0 \text{ ou } -1 \end{cases} \end{array}$$

$$(5) \quad \left\{ \begin{array}{l} - g'(q'_f, \$) = (q'_f, \$, +1) \\ - g'(q', x) = (q'_p, x, +1) \text{ pour tout } q' \in Q' \text{ et } x \in V' \cup \{\$, \emptyset\} \text{ pour} \\ \hspace{10em} \text{lesquels } g' \text{ n'est pas encore définie.} \end{array} \right.$$

Pour tout $u, v \in V^*$ et tout $q \in Q$, nous avons, dans M' , le mouvement suivant :

$$(q'_0, \emptyset, uqv\$) \mid_{M'}^* \textcircled{1} (q'_1, \emptyset uqv, \$) \mid_{M'}^* \textcircled{2} (q, \emptyset u, vq_0\$).$$

Si $(q, \emptyset u, v\$) \mid_M^* (q_1, \emptyset y \$, \Lambda)$ avec $q_1 \in F$, $\exists q' \in Q$ tel que

$$(q, \emptyset u, v\$) \mid_M^* (q', \emptyset y, \$) \text{ et } g(q', \$) = (q_1, \$, +1) \text{ donc } g'(q', q_0) = (q'_f, q_1, +1) \text{ (cf. } \textcircled{4} \text{)}.$$

$$\text{Alors } (q, \emptyset u, vq_0\$) \mid_{M'}^* \textcircled{3} (q', \emptyset y, q_0\$) \mid_{M'}^* \textcircled{4} (q'_f, \emptyset yq_1, \$) \mid_{M'}^* \textcircled{5} (q'_f, \emptyset yq_1 \$, \Lambda).$$

Réciproquement, si $(q'_0, \emptyset, uqv\$) \mid_{M'}^* (q'_f, \emptyset yq_1 \$, \Lambda)$, il existe $q_2 \in Q$ tel

que :

$$(q, \emptyset u, vq_0\$) \mid_{M'}^* (q_2, \emptyset y, q_0\$) \mid_{M'}^* (q'_f, \emptyset yq_1, \$) \mid_{M'}^* (q'_f, \emptyset yq_1 \$, \Lambda).$$

Donc $(q, \emptyset u, v\$) \mid_M^* (q_2, \emptyset y, \$)$ et comme

$$g'(q_2, q_0) = (q'_f, q_1, +1), \quad g(q_2, \$) = (q_1, \$, +1)$$

$$\text{et } (q, \emptyset u, v\$) \mid_M^* (q_1, \emptyset y \$, \Lambda).$$

Lemme V.7 Soient $M = (Q, V, q_0, g, F)$ un automate linéairement borné déterministe, avec les marqueurs $\emptyset, \$$ et $L_0 \subset V^*$. Si $L_0 \in L'_1$, alors
 $L_1 = \{v \in V^* / (q_0, \emptyset, u\$) \mid_M^* (q, \emptyset v \$, \Lambda) \text{ avec } u \in L_0 \text{ et } q \in F\} \in L'_1.$

Démonstration Posons $V' = V \times V$, $Q' = Q \cup \{q'_v, q'_f, q'_p\}$ avec $Q \cap \{q'_v, q'_f, q'_p\} = \emptyset$ et considérons $M' = (Q', V', q_0, g', \{q'_f\})$ où g' est définie par :

- 1- pour $x, y \in V$ et $q \in Q$ $g'(q, (x, y)) = (q_1, (x_1, y), i)$ si $g(q, x) = (q_1, x_1, i)$
- 2- pour $q \in Q$ $g'(q, \emptyset) = g(q, \emptyset)$
- 3- pour $q \in Q$ $g'(q, \$) = \begin{cases} (q_1, \$, i) & \text{si } g(q, \$) = (q_1, \$, i) \\ & \text{avec } i=0 \text{ ou } 1 \\ (q'_V, \$, -1) & \text{si } g(q, \$) = (q_1, \$, +1) \\ & \text{avec } q_1 \in F \end{cases}$
- 4- pour $x \in V$, $g'(q'_V, (x, x)) = (q'_V, (x, x), -1)$
- 5- $g'(q'_V, \emptyset) = (q'_F, \emptyset, +1)$
- 6- pour $v' \in V' \cup \{\$, \}$, $g'(q'_F, v') = (q'_F, v', +1)$
- 7- $g'(q', v') = (q'_P, v', +1)$ pour tout $q' \in Q'$ et $v' \in V' \cup \{\$, \}$ pour lesquels g' n'est pas encore définie.

Il est clair que M' est un automate linéairement borné déterministe.

Posons $L' = \{W' \in V'^* / (q_0, \emptyset, h_1(W'))\$ \mid_M^* (q, \emptyset h_2(W'))\$, \Lambda) \text{ avec } q \in F\}$,
 h_1 et h_2 étant les projections canoniques homomorphes de V'^* dans V^* définies par
 $h_1((x, y)) = x$, $h_2((x, y)) = y$ pour tout $(x, y) \in V'$ et montrons que $L' = T(M')$.

En effet, si $(q, \emptyset u', v'\$) \mid_{M'}^* (q_1, \emptyset u'_1, v'_1\$)$ avec q et $q_1 \in Q$, il est facile de vérifier que

$$h_2(u', v') = h_2(u'_1 v'_1) \text{ et que } (q, \emptyset h_1(u'), h_1(v')\$) \mid_M^* (q_1, \emptyset h_1(u'_1), h_1(v'_1)\$).$$

Prenons $W' \in T(M')$; nous avons dans M' le mouvement suivant :

$$(q_0, \emptyset, W'\$) \mid_{M'}^* (q_2, \emptyset W'_1, \$) \mid_{M'}^* (q'_F, \emptyset, W'_1\$) \mid_{M'}^* (q'_F, \emptyset W'_1, \$, \Lambda)$$

avec $g'(q_2, \$) = (q'_V, \$, -1)$. Ce qui implique :

$$(q_0, \emptyset, h_1(W')\$) \mid_M^* (q_2, \emptyset h_1(W'_1), \$), \quad h_2(W') = h_2(W'_1),$$

$$g(q_2, \$) = (q, \$, +1) \text{ avec } q \in F \text{ et } h_1(W'_1) = h_2(W'_1) \quad (\text{cf. 4})$$

Donc $(q_0, \emptyset, h_1(W')\$) \mid_M^* (q, \emptyset, h_2(W')\$, \Lambda)$ avec $q \in F$ et $W' \in L'$.

Réciproquement, pour tout $W', W'_1, z', z'_1 \in V'^*$ vérifiant $h_2(W'z') = h_2(W'_1 z'_1)$,
si $(q, \emptyset h_1(W'), h_1(z')\$) \mid_M^* (q_1, \emptyset h_1(W'_1), h_1(z'_1)\$)$ alors

$$(q, \varnothing W', z' \$) \mid_{M'}^* (q_1, \varnothing W'_1, z'_1 \$).$$

Prenons $W' \in L'$. Alors il existe $q_2 \in Q$ tel que $(q_0, \varnothing, h_1(W') \$) \mid_M^* (q_2, \varnothing h_2(W') \$)$, $g(q_2, \$) = (q, \$, +1)$ avec $q \in F$. En prenant W'_1 défini par $h_1(W'_1) = h_2(W'_1) = h_2(W')$, nous obtenons $(q_0, \varnothing, W' \$) \mid_{M'}^* (q_2, \varnothing W'_1 \$) \mid_{M'}^* (q_f, \varnothing W'_1 \$, \Lambda)$ donc $W' \in T(M')$.

$$\begin{aligned} L_1 &= \{v \in V^* / (q_0, \varnothing, v \$) \mid_M^* (q, \varnothing v \$, \Lambda) \text{ avec } q \in F \text{ et } v \in L_0\} \\ &= \{h_2(W') / W' \in V'^*, (q_0, \varnothing, h_1(W') \$) \mid_M^* (q, \varnothing h_2(W') \$, \Lambda) \text{ avec } q \in F \text{ et } h_1(W') \in L_0\} \\ &= \{h_2(W') / W' \in L' \text{ et } h_1(W') \in L_0\}. \end{aligned}$$

$$\text{Donc } L_1 = h_2[h_1^{-1}(L_0) \cap L'].$$

Comme L_0 et $L' \in L'_1$ et que h_2 est un homomorphisme Λ -libre, nous déduisons des propositions V.1 et V.2 que $L_1 \in L'_1$.

Lemme V.8 Pour toute LB-grammaire déterministe, il existe une LB-grammaire équivalente $G' = (N', T, P', A)$ déterministe en N'_1, N'_2, N'_3 telle que, en posant $T'_1 = V' \setminus N'_1$ et $P'_1 = P' \cap (N'_1 \times V'^+)$, le langage engendré par la grammaire $G'_1 = (N'_1, T'_1, P'_1, A)$ soit inclus dans $N_3^* N'_2 N_3^*$.

Démonstration Soient $G = (N, T, P, A)$ une LB-grammaire déterministe en N_1, N_2, N_3 et $\{P_1, P_2, P_3\}$ la partition de P correspondante. Considérons $G = (N_1, T_1, P_1, A)$ avec $T_1 = V \setminus N_1$. Posons $L_1 = L(G_1) \cap T^*$ et $L_2 = L(G_1) \cap T^* N_3^* N_2 N_3^* T^*$. Comme $x \in L(G_1)$ implique $\ell_{N_2}(x) \leq 1$ (cf. propriété 1 de déterminisme), si $W \in L(G_1) \setminus (L_1 \cup L_2)$, alors W ne peut pas se dériver dans G , en un mot de T^* . Donc, en posant $L'_2 = \{z \in T^* / y \xrightarrow[G]{*} z \text{ avec } y \in L_2\}$, nous obtenons $L(G) = L_1 \cup L'_2$.

Prenons $\bar{T} = \{\bar{a} / a \in T\}$ disjoint de V , $E \notin \bar{T} \cup V$ et posons

$$N'_2 = N_2 \cup \{E\}, N'_3 = N_3 \cup \bar{T}, T'_1 = N'_2 \cup N'_3 \text{ et } L_3 = h(L_1) \{E\} \cup h(L_2),$$

h étant l'homomorphisme de T_1^* dans T'^*_1 défini par :

$$\forall x \in T_1 \quad h(x) = \begin{cases} \bar{x} & \text{si } x \in T \\ x & \text{sinon.} \end{cases}$$

Comme L_3 est un C-langage, il existe une C-grammaire $G'_1 = (N'_1, T'_1, P'_1, A)$ telle que $L(G'_1) = L_3 \subset N'_1 N'_2 N'_1$.

Posons $R = \{E_2 \in N_2 / R_2(A_3 E_2) = \emptyset \text{ pour tout } A_3 \in N_3\}$,

$$P'_2 = P_2 \cup \{(E_2 \bar{a}, \bar{a} E_2) / a \in T \text{ et } E_2 \in R\},$$

$$P'_3 = P_3 \cup \{(\bar{a} E_2, E_2 a), (\bar{a} E_2, a) / a \in T \text{ et } E_2 \in R\}$$

$$\cup \{(\bar{a} E, E a), (\bar{a} E, a) / a \in T\}$$

$$\cup \{(xE_2, EW) / x \in N_3^*, E_2 \in N_2, W \in T^+ \text{ et } (xE_2, W) \in P_3\},$$

$$N' = N'_1 \cup N'_2 \cup N'_3, P' = P'_1 \cup P'_2 \cup P'_3 \text{ et } G' = (N', T, P', A).$$

Il est clair que G' est une LB-grammaire déterministe et que

$$L(G') = \{z' \in T^* / y' \xrightarrow[G']{*} z' \text{ avec } y' \in L_3\}.$$

En outre, il est facile de vérifier que :

$$L_1 = \{z' \in T^* / y' \xrightarrow[G']{*} z', y' \in h(L_1) \setminus \{E\}\}$$

$$\text{et } L'_2 = \{z' \in T^* / y' \xrightarrow[G']{*} z', y' \in h(L_2)\}.$$

$$\text{Donc } L(G') = L_1 \cup L'_2 = L(G).$$

Soient $G = (N, T, P, A)$ une LB-grammaire déterministe en N_1, N_2, N_3 , $\{P_1, P_2, P_3\}$ la partition de P obtenue à partir de N_1, N_2, N_3 , $T_1 = V \setminus N_1$ et $G_1 = (N_1, T_1, P_1, A)$. D'après le lemme V.8, nous pouvons supposer que $L(G_1)$ est inclus dans $N_3^* N_2 N_3^*$. Notons $N'_2 = \{E_2 \in N_2 / R_2(A_3 E_2) = \emptyset \ \forall A_3 \in N_3\}$.

D'après la condition 3 de déterminisme, pour tout $W \in L(G)$, il existe

$$W_1 \in L(G_1) \text{ et } W_2 \in N_3^* N'_2 \text{ tels que } A \xrightarrow[P_1]{*} W_1 \xrightarrow[P_2]{*} W_2 \xrightarrow[P_3]{*} W.$$

$$\text{Posons } L_2 = \{W_2 \in N_3^* N'_2 / A \xrightarrow[G]{*} W_2\} = \{W_2 \in N_3^* N'_2 / W_1 \xrightarrow[P_2]{*} W_2, W_1 \in L(G_1)\}$$

$$L(G) \text{ est alors égal à } \{W \in T^* / W_2 \xrightarrow[P_3]{*} W, W_2 \in L_2\}.$$

Montrons que $\widetilde{L(G)}$ est l'image transductée de $\widetilde{L_2}$ par un transducteur séquentiel Λ -libre.

Soient k_0 et $k_f \notin N'_2$. Posons $K = N'_2 \cup \{k_0, k_f\}$, $X = N_3 \cup N'_2$ et considérons le transducteur séquentiel $T_S = (K, X, T, k_0, H, k_f)$ avec H défini par :

$$\begin{aligned} H = & \{(k_0, q\tilde{w}, \tilde{z}, q') / w \in N_3^*, q \in N'_2, z \in T^+, q' \in N'_2, (Wq, q'z) \in P_3\} \\ & \cup \{(k_0, q\tilde{w}, \tilde{z}, k_f) / w \in N_3^*, q \in N'_2, z \in T^+, (Wq, z) \in P_3\} \\ & \cup \{(q, \tilde{w}, \tilde{z}, q') / w \in N_3^*, q, q' \in N'_2, z \in T^+, (Wq, q'z) \in P_3\} \\ & \cup \{(q, \tilde{w}, \tilde{z}, k_f) / w \in N_3^*, q \in N'_2, z \in T^+, (Wq, z) \in P_3\}. \end{aligned}$$

T_S est Λ -libre et l'image de $\tilde{L}_2$ par T_S est $\tilde{L}(G)$.

Pour montrer que $L(G) \in L'_1$, il suffit de montrer que $L_2 \in L'_1$. En effet, $L_2 \in L'_1$ implique $\tilde{L}_2 \in L'_1$ et d'après la proposition V.3, $\tilde{L}(G) \in L'_1$ donc $L(G) \in L'_1$.

Soient $q_p \notin N_2$ et $\bar{N}_2 = \{\bar{E}_2 / E_2 \in N_2\}$.

Posons $Q = N_2 \cup \bar{N}_2 \cup \{q_p\}$ et considérons l'automate linéairement borné déterministe $M = (Q, N_3, q_p, g, N'_2)$ avec les marqueurs $\emptyset$ et $\$$ où g est définie par :

$$\text{- pour } E_2 \in N_2 \text{ et } A_3 \in N_3 \quad g(E_2, A_3) = \begin{cases} (F_2, B_3, 0) \text{ si } (E_2 A_3, F_2 B_3) \in P_2 \text{ avec } F_2 \in N_2 \\ (F_2, B_3, +1) \text{ si } (E_2 A_3, B_3 F_2) \in P_2 \text{ avec } F_2 \in N_2 \\ (\bar{E}_2, A_3, -1) \text{ si } R_2(E_2 A_3) = \emptyset \end{cases}$$

$$\text{- pour } E_2 \in N_2 \quad g(E_2, \$) = \begin{cases} (E_2, \$, +1) \text{ si } E_2 \in N'_2 \\ (\bar{E}_2, \$, -1) \text{ sinon} \end{cases}$$

$$\text{- pour } \bar{E}_2 \in \bar{N}_2 \text{ et } A_3 \in N_3 \quad g(\bar{E}_2, A_3) = \begin{cases} (F_2, B_3, 0) \text{ si } (A_3 E_2, F_2 B_3) \in P_2 \text{ avec } F_2 \in N_2 \\ (F_2, B_3, +1) \text{ si } (A_3 E_2, B_3 F_2) \in P_2 \text{ avec } F_2 \in N_2 \end{cases}$$

- $g(q, x) = (q_p, x, +1)$ pour tout $q \in Q$ et tout $x \in N_3 \cup \{\$, \emptyset\}$ pour lesquels g n'a pas encore été définie.

Pour $u, v \in N_3^*$ et $E_2 \in N_2$, notons $C(uE_2v) = (E_2, \varnothing u, v\$)$ et $D((E_2, \varnothing u, v\$)) = uE_2v$. Pour $\bar{E}_2 \in \bar{N}_2$,

$$D((\bar{E}_2, \varnothing u, v)) = \begin{cases} u.A_3E_2v' & \text{si } v = A_3v' \text{ avec } A_3 \in N_3 \\ u\bar{E}_2 & \text{si } v = \Lambda \end{cases}$$

Par construction de g , nous obtenons les propriétés suivantes :

- Si $W \xrightarrow{P_2} W'$ alors $C(W) \mid_M^* C(W')$
- Si $Z \mid_M Z'$ alors $D(Z) \xrightarrow{P_2}^* D(Z')$.

Nous en déduisons que pour tout $u, u', v' \in N_3^*$, $E_2 \in N_2$, $E'_2 \in N'_2$, $uE_2v \xrightarrow{P_2}^* u'E'_2$ si et seulement si $(E_2, \varnothing u, v\$) \mid_M^* (E'_2, \varnothing u', v'\$, \Lambda)$.

$$\begin{aligned} \text{Donc } L_2 &= \{u'E'_2 / uE_2v \xrightarrow{P_2}^* u'E'_2 \text{ avec } E_2 \in N_2, E'_2 \in N'_2 \text{ et } uE_2v \in L_1\} \\ &= \{u'E'_2 / (E_2, \varnothing u, v\$) \mid_M^* (E'_2, \varnothing u', v'\$, \Lambda) \text{ avec } E_2 \in N_2, E'_2 \in N'_2, uE_2v \in L_1\}. \end{aligned}$$

Comme $Q' \cap N_3 = \emptyset$, nous pouvons utiliser le lemme V.6 et nous obtenons un automate $M' = (Q', V', q'_0, g', \{q'_f\})$ tel que $L_2 = \{u'E'_2 / (q'_0, \varnothing, uE_2v\$) \mid_{M'}^*, (q'_f, \varnothing u'E'_2\$, \Lambda) \text{ avec } E_2 \in N_2, E'_2 \in N'_2, u, v, u' \in N_3^* \text{ et } uE_2v \in L_1\}$.

Mais L_1 est inclus dans $N_3^*N_2N_3^*$, donc

$$L_2 = \{y' / (q'_0, \varnothing, y\$) \mid_{M'}^* (q'_f, \varnothing y'\$, \Lambda) \text{ avec } y \in L_1\} \cap N_3^*N'_2$$

et d'après le lemme V.7, $L_2 \in L'_1$.

Nous pouvons alors énoncer, en tenant compte de la proposition V.4 :

Proposition V.6 La classe des langages engendrés par une LB-grammaire déterministe est équivalente à la classe L'_1 .

B I B L I O G R A P H I E

- [CH] N. CHOMSKY : On certain formal properties of grammars,
Information and Control, vol.2, 1959, pp.137-167.
- [DA] M. DAVIS : Computability and Unsolvability, Mc Graw-Hill
Book Compagny, New York, 1958.
- [GG1] S. GINSBURG et S.A. GREIBACH : Mappings which preserve
Context sensitive languages, Information and Control,
vol.9, 1966, pp.563-582.
- [GG2] S. GINSBURG et S.A. GREIBACH : Deterministic Context Free
Languages, Information and Control, vol.9, 1966, pp.620-648.
- [GGH] S. GINSBURG, S.A. GREIBACH et M.A. HARRISON : One-Way
Stack Automata, Journal of the Association for Computing
Machinery, vol.14, 1967, pp.389-418.
- [GH] S. GINSBURG et M.A. HARRISON : On the Elimination of End-
markers, Information and Control, vol.12, 1968, pp.103-115.
- [GI] S. GINSBURG : The mathematical theory of context-free languages,
1966, Mc Graw-Hill, New York.
- [GR] S.A. GREIBACH : A new normal-form theorem for context-free
phrase structure grammars, Journal of the Association for
Computing Machinery, vol.12, 1965, pp.42-52.

- [GS1] S. GINSBURG et E.M. SPANIER : Finite-Turn Pushdown Automata,
Journal SIAM Control, vol.4, 1966, pp.429-453.
- [GS2] S. GINSBURG et E.M. SPANIER : Control Sets on grammars,
Mathematical Systems Theory, vol.2, 1968, pp.159-177.
- [GS3] S. GINSBURG et E.M. SPANIER : Derivation-Bounded Languages,
Journal of Computer and System Sciences, vol.2, 1968,
pp.228-250.
- [KH] A.J. KORENJAK et J.E. HOPCROFT : Simple deterministic languages,
Proc. 7th. Symp. Switching and Automata Theory, I.E.E.E., 1966,
pp. 36-46.
- [KN] D.E. KNUTH : On the Translation of Languages from Left to Right,
Information and Control, vol.8, 1965, pp.607-639.
- [KS] R. KURKI-SUONIO : Notes on Top-Down Languages, BIT, vol.9, 1969,
pp.225-238.
- [KU] S.Y. KURODA : Classes of Languages and Linear-bounded automata,
Information and Control, vol.7, 1964, pp.207-223.
- [LA] P.S. LANDWEBER : Three theorems on phrase structure grammars
of type 1, Information and Control, vol.6, 1963, pp.131-136.
- [LS] P.M. LEWIS et R.E. STEARNS : Syntax-Directed Transduction,
Journal of the Association for Computing Machinery, vol.15,
1968, pp.465-488.

- [MA] G.M. MATTHEWS : A Note on Asymmetry in Phrase Structure Grammars, Information and Control, vol.7, 1964, pp.360-365.
- [MY] J. MYHILL : Linear bounded Automata, Wright Air Development Division, 1960, Tech.Note No. 60-165, Cincinnati, Ohio.
- [NI] M. NIVAT : Transductions des langages de Chomsky, Ann. Inst. Fourier, Grenoble, 1968, pp.339-456.
- [PO] E.L. POST : A variant of a recursively unsolvable problem, Bulletin of the American Mathematical Society, vol.52, 1946, pp. 264-268.
- [YN] M.K. YNTEMA : Inclusion Relations among Families of Context-Free Languages, Information and Control, vol.10, 1967, pp.572-597.
- [WA] P.M. WALLEZ : C-langages permutables et Analyse syntaxique, Thèse, Faculté des Sciences de Lille, 1971.
