PUBLICATIONS DE L'UNIVERSITÉ DE NANCY

COLLECTION DES THÈSES DE LA FACULTÉ DES SCIENCES

MATHÉMATIQUES

Première Thèse

Etude de la notion de pile application à l'analyse syntaxique

Deuxième Thèse

Algèbre différentielle

Présentées à la Faculté des Sciences de l'Université de Nancy
pour obtenir le grade de Docteur ès-Sciences Mathématiques
par **Claude PAIR**
soutenues le 21 décembre 1965
devant la Commission d'Examen

M. LEGRAS, Président
M^lle HUET
MM. ARSAC
EYMARD
HERZ
Examinateurs

1966

PUBLICATIONS DE L'UNIVERSITÉ DE NANCY

COLLECTION DES THÈSES DE LA FACULTÉ DES SCIENCES

MATHÉMATIQUES

Première Thèse

Etude de la notion de pile application à l'analyse syntaxique

Deuxième Thèse

Algèbre différentielle

Présentées à la Faculté des Sciences de l'Université de Nancy
pour obtenir le grade de Docteur ès-Sciences Mathématiques

par **Claude PAIR**

soutenues le 21 décembre 1965
devant la Commission d'Examen

M. LEGRAS, Président
Mlle HUET
MM. ARSAC
EYMARD
HERZ
Examinateurs

1966

Doyen : M. AUBRY

Assesseur : M. GAY

Doyens honoraires : MM. CORNUBERT - DELSARTE - URION - ROUBAULT.

Professeurs honoraires : MM. CROZE - RAYBAUD - LAFFITTE - LERAY - JULY - LAPORTE - EICHHORN - GODEMENT - DUBREIL - L. SCHWARTZ - DIEUDONNE - DE MALLEMANN - LONGCHAMBON - LETORT - DODE - GAUTHIER - GOUDET - OLMER - CORNUBERT - CHAPELLE - GUERIN - WAHL - ECHEVIN.

Maîtres de conférences honoraires : MM. LIENHART - PIERRET.

Professeurs

MM. URION	Chimie biologique	*GAYET	Physiologie
DELSARTE	Analyse supérieure	*MALAPRADE	Chimie
ROUBAULT	Géologie	HADNI	Physique
CAPELLE	Mécanique rationnelle	BONVALET	Mécanique appliquée
VEILLET	Biologie animale	*KERN	Minéralogie
BARRIOL	Chimie théorique	*BASTICK	Chimie
BIZETTE	Physique	DUCHAUFOUR	Pédologie
GUILLIEN	Electronique	NEEL	Chimie organique industrielle
GIBERT	Chimie physique		
LEGRAS	Mécanique rationnelle	GARNIER	Agronomie
BOLFA	Minéralogie	*WEPPE	Minéralogie appliquée
NICLAUSE	Chimie	BERNARD	Géologie appliquée
FAIVRE	Physique appliquée	*CHAMPIER	Physique
AUBRY	Chimie minérale	*REGNIER	Physico-chimie
DUVAL	Chimie	*GAY	Chimie biologique
COPPENS	Radiogéologie	*WERNER	Botanique
FRUHLING	Physique	*CONDE	Zoologie
SUHNER	Physique expérimentale	STEPHAN	Zoologie
HILLY	Géologie	EYMARD	Calcul différentiel et intégral
LE GOFF	Génie chimique		
CHAPON	Chimie biologique	LEVISALLES	Chimie organique
HEROLD	Chimie miné. industrielle	MANGENOT	Botanique
SCHWARTZ	Exploitation minière	Mme HERVE	Méthodes mathématiques de la physique

* Professeur à titre personnel.

Maîtres de conférences

MM.	GOSSE	Mécanique physique
	ROCCI	Géologie
	VUILLAUME	Psychophysiologie
Mme	BASTICK	Chimie M.P.C. (Epinal)
	GUDEFIN	Physique
	HORN	Physique propédeutique
	FRENTZ	Biologie animale
	AUROUZE	Géologie
	MARI	Chimie (I.S.I.N.)
	LAFON	Physique (I.S.I.N.)
	FELDEN	Phys. théorique et nucléaire
	FLECHON	Physique M.P.C.
	VIGNES	Métallurgie
Mle	HUET	Mathématiques S.P.C.N.
	DEVIOT	Physique du solide
	BLAZY	Minéralogie appliquée (ENSG)
	BALESDENT	Thermodynamique chimique appliquée
	JANOT	Physique M.P.C. (Epinal)
	JACQUIN	Pédologie et ch. agricole

N... Probabilités et statistiques
N... Mécanique (I.S.I.N.)
N... Physique M.G.P.
N... Mathématiques
N... Mécanique expérimentale
N... Génie chimique
N... Mathématiques propédeutiques
N... Phytopathologie
N... Mathématiques S.P.C.N.

Cette thèse doit son existence à Monsieur le Professeur J. LEGRAS. Je le remercie de m'avoir orienté vers la théorie des langages, d'avoir facilité au maximum mon travail,et surtout de la confiance qu'il m'a toujours témoignée.

Que Monsieur J. ARSAC, Professeur à la Faculté des Sciences de Paris, et Monsieur J. C. HERZ, Maître de Conférences à la Faculté des Sciences de Lille, trouvent ici l'expression de ma gratitude pour le nouvel encouragement que constitue pour moi leur présence au Jury.

Je veux remercier non moins vivement Mademoiselle D. HUET, et Monsieur P. EYMARD, Professeurs à la Faculté des Sciences de Nancy : la participation à ce jury de deux mathématiciens d'une telle compétence est un honneur pour moi.

J'ai demandé aux Secrétaires du Centre Universitaire de Calcul un travail ingrat ; leur concours m'a été précieux et je les en remercie.

Je remercie enfin le personnel de l'Atelier de Reproduction Universitaire qui a apporté tous ses soins à l'impression de ce travail.

A ma Femme.

TABLE DES MATIERES

INTRODUCTION

De nombreux auteurs ont défini et utilisé, sous des noms divers (symbolkeller, last in first out, push down store, pile), des dispositifs qui rappellent la notion courante de pile d'objets. Employée en 1959 par [52] pour traduire un langage de programmation, cette notion a depuis été très féconde en théorie des langages. On peut penser qu'elle sera utile aussi dans d'autres domaines.

Nous étudierons d'abord cette notion de pile en elle-même. En particulier, nous mettrons en évidence la relation qui existe entre piles et ramifications orientées : une ramification est un graphe dont toute composante connexe est une arborescence ; elle est orientée lorsqu'on a défini un ordre total dans chacune de ses "familles". A toute ramification orientée, nous associerons trois piles (piles attachée, adjointe, conjointe) dont chacune caractérise la ramification. Au passage, nous indiquerons rapidement quelques applications des piles à des problèmes qui se posent en traitement de l'information.

Une grande partie de ce travail est consacrée à l'analyse syntaxique pour les grammaires de Chomsky. La formalisation des grammaires due à Chomsky [12], qui rend compte des principaux aspects syntaxiques des langues naturelles aussi bien que des langages de programmation, conduit à associer à chaque phrase une ou plusieurs "structures arborescentes" : on dit alors qu'on a analysé la phrase. Si on se réfère à l'analyse grammaticale scolaire, il s'agit de décomposer la phrase en propositions, de distinguer dans chacune d'elles un groupe sujet, un groupe verbal, un groupe complément, à leur tour décomposés en nom, article, adjectif,... et même racine, désinence... . L'analyse d'une expression algébrique du genre $(a+b)\times c-d\times \sin(a\times y+b)$ est d'ailleurs de même nature.

Nous introduirons ces "structures arborescentes" sous la forme précise de pseudo-arborescences dans la définition même des structures de Chomsky. Chaque pseudo-arborescence, comme chaque ramification orientée, sera caractérisée par l'une quelconque des trois piles, attachée, adjointe, conjointe à la pseudo-arborescence ; ainsi le problème de l'analyse se ramène à un problème de construction de piles, et nous distinguerons trois familles d'algorithmes d'analyse selon celle de ces trois piles qu'on cherche pour déterminer une pseudo-arborescence. La plupart des algorithmes décrits dans la littérature entrent d'ailleurs dans l'une de ces trois familles. Malheureusement, ils sont souvent publiés sans être rattachés à aucune idée générale, sans hypothèse de validité et sans démonstration. La raison en est que les outils de démonstration ne sont guère forgés. Nous essaierons d'y remédier en précisant au chapitre 5 cette notion d'algorithme d'analyse, dans la ligne de la théorie des automates, sous la forme de ce que nous nommerons les *générateurs de piles*. Les algorithmes introduits seront donc définis sous cette forme précise, qui nous permettra de les justifier et de les étudier ; le plus souvent, ils seront aussi décrits rapidement et approximativement par un organigramme qui en permettra une vue d'ensemble.

CHAPITRE I

DEFINITIONS ET NOTATIONS

1. 1. Vocabulaire utilisé à propos des relations binaires, des relations d'ordre.

On pourra consulter [5], [18]. Pour une relation (binaire) R dans un ensemble E, si x est un élément de E, on note R (x) l'ensemble des $y \in E$ tels que x R y ; l'ensemble des $y \in E$ tels que y R x sera donc désigné par $R^{-1}(x)$.

Soit R une relation d'ordre dans un ensemble E. R (x) est l'ensemble des *majorants* de x, $R^{-1}(x)$ celui des *minorants* de x. Si la *restriction* R_A de R à un sous-ensemble A de E est un *ordre total*, A est une *R-chaîne* (ou simplement chaîne). Si x R y, on appelle *intervalle fermé* $[_R\, x,y]$ l'ensemble des éléments z de E tels que x R z R y (autrement dit l'intersection $R(x) \cap R^{-1}(y)$), et intervalle ouvert $]_R\, x,y[$, le complémentaire dans $[_R\, x,y]$ de l'ensemble formé de x et y (la lettre R mise en indice du premier crochet pourra être enlevée quand il n'y aura pas ambiguïté). Si l'intervalle]x,y[est vide, y *couvre* x. Une partie F de E est *convexe* pour R si, chaque fois qu'elle contient deux éléments x et y, elle contient aussi $[_R\, x,y]$; si F est une partie convexe de E pour R et G une partie convexe de F pour la restriction de R à F, alors G est une partie convexe de E pour R. Lorsque l'ensemble des minorants (resp. majorants) *stricts* de x, c'est-à-dire différents de x, n'est pas vide et possède un *plus grand* (resp. *plus petit*) élément y, y est le *prédécesseur* (resp. *successeur*) de x ; alors x couvre y (resp. y couvre x).x et y sont *comparables* si x R y ou y R x.

Pour toute relation R, on notera $\bar{R}$ sa négation " non R ".

1. 2. Graphes.

[3] définit un graphe comme un couple (E, Γ) d'un ensemble E (ensemble des points du graphe ; nous le supposerons toujours non vide) et d'une application multivoque Γ de E dans E, c'est-à-dire d'une relation Γ dans l'ensemble E. Une suite $(x_0, \dots, x_n)$ de points d'un graphe (E, Γ) telle que $n \geqslant 0$ et pour $i = 1, \dots n$, $x_{i-1}\, \Gamma\, x_i$ est un *chemin* du graphe qui *joint* x_0 à x_n ; x_0 est son *origine*, x_n son *extrémité*. Si $n \neq 0$ et $x_0 = x_n$, le chemin est un *circuit*.

Nous nommons *pseudo-chemin*(1) du graphe (E, Γ) toute suite ($x_0, x_1, \dots x_n$) de points distincts telle que pour i = 1, 2, ... n, $x_{i-1} \Gamma x_i$ ou $x_i \Gamma x_{i-1}$. x_0 et x_n sont les extrémités du pseudo-chemin. Une *composante connexe* du graphe (E, Γ) est un sous-graphe maximal parmi ceux où tout couple de points distincts est le couple des extrémités d'un pseudo-chemin.

A tout graphe (E, Γ) on associe un *préordre* R qui est la *fermeture transitive* de Γ : x R y si, et seulement si, il existe un chemin d'origine x et d'extrémité y. R est un ordre si, et seulement si, le graphe n'admet pas de circuit contenant deux points distincts.

Nous nous intéresserons surtout dans la suite aux graphes finis (nombre fini de points) et sans circuit. Soit E un ensemble fini, et R une relation d'ordre dans E. Toute partie non vide de E possède des éléments maximaux et minimaux. Par conséquent, pour tout élément x de E qui est strictement majoré, il existe y qui couvre x. Tout graphe (E, Γ) auquel l'ordre R est associé vérifie :

$$y \text{ couvre } x \Longrightarrow x \Gamma y.$$

Inversement, étant donné une relation d'ordre R, le graphe (E, Γ) défini par

$$x \Gamma y \Longleftrightarrow y \text{ couvre } x \text{ pour la relation } R$$

possède R pour relation d'ordre associée. Parmi les graphes auxquels R est associée c'est celui pour lequel la relation Γ est minimale (autrement dit, qui possède le minimum d'*arcs*). Ainsi, toute relation d'ordre R dans E définit un graphe sans circuit, que nous noterons [E, R]. En particulier, la relation Γ de ce graphe est antiréflexive : pour tout x de E, x $\bar{\Gamma}$ x.

Un graphe fini (E, Γ) est une *arborescence* si :

a) il n'admet pas de circuit ;

b) il existe un élément r de E, tel que $\Gamma^{-1}(r)$ soit vide ;

c) si x ≠ r, $\Gamma^{-1}(x)$ est un ensemble à un élément.

La notion d'arborescence sera généralisée au paragraphe 1. 4 en celle de ramification.

On appelle *isomorphisme* d'un graphe (E, Γ) sur un graphe (E', Γ') une application bijective φ de E dans E', telle que

$$(\forall x \in E) \ (\forall y \in E) \ (x \Gamma y \iff \varphi(x) \Gamma' \varphi(y)).$$

Nous dirons que φ transforme le graphe (E, Γ) en (E', Γ') et nous noterons parfois φ (E, Γ) = (E', Γ'). Si R et R' sont les préordres associés aux deux graphes (E, Γ) et (E', Γ'), φ est aussi un isomorphisme du graphe (E, R) sur (E', R'). Inversement, soit φ un isomorphisme d'un graphe (E, R) sur (E', R'), où R est une relation d'ordre : R' est aussi une relation d'ordre et φ est un isomorphisme du graphe [E, R] sur [E', R']. Un isomorphisme transforme un graphe sans circuit en un graphe sans circuit, une arborescence en une arborescence.

(1) Nous éviterons le mot *chaîne*, employé par [3], pour ne pas créer de confusion avec une partie totalement ordonnée d'un ensemble (1. 1).

1. 3. Monoïde libre. (cf. [6], [11]).

Une suite finie $\alpha = (a_0, a_1, \dots a_n)$ d'éléments d'un ensemble E est appelée *mot sur* E ; nous la noterons '$a_0 a_1 \dots a_n$' ou parfois simplement $a_0 a_1 \dots a_n$. n+1 est la *longueur* du mot α : n+1 = |α|. Par abus de langage, on confond souvent le mot 'a_0' de longueur 1 avec l'élément a_0 de E. Si $a_i = b$, nous dirons que i est une *occurrence* ou un *rang* de b dans α, et que b *possède une occurrence dans* α.

L'ensemble E* des mots sur l'ensemble E est muni d'une loi de composition interne, nommée *concaténation*, définie par :

$$'a_0 a_1 \dots a_n' \ 'b_0 b_1 \dots b_p' = 'c_0 c_1 \dots c_{n+p+1}'$$

où $c_0 = a_0, \ c_1 = a_1, \dots, c_n = a_n, \ c_{n+1} = b_0, \dots, c_{n+p+1} = b_p$.

E* est un monoïde : on dit que c'est le *monoïde libre déduit de* E. Dans la suite de ce travail, pour tout ensemble E, E* désignera le monoïde libre déduit de E. Le *mot vide*, noté Λ (de longueur 0) est élément neutre de E*. Si α,β,γ,δ sont des mots tels que α = β δ γ, nous dirons que β est *facteur gauche*, γ *facteur droit* et δ *sous-mot* de α.

Si α = '$a_0 a_1 \dots a_{n-1} a_n$', le mot $\tilde{\alpha}$ = '$b_0 b_1 \dots b_{n-1} b_n$' où $b_i = a_{n-i}$ est le *réfléchi* de α.

Soient E' et E'' deux sous-ensembles complémentaires de E. Tout mot α de E* s'écrit, de manière unique $\alpha = \beta_1 \gamma_1 \cdots \beta_p \gamma_p$ où p ≥ 1, $\beta_i \in E'^*$, $\gamma_j \in E''^*$ et aucun mot β_i ou γ_j n'est vide sauf éventuellement β_1 et γ_p. Nous nommerons *trace* de α sur le monoïde libre E'* le mot $\beta_1 \dots \beta_p$.

Toute application g de E dans le monoïde libre E'* déduit d'un ensemble E' induit un homomorphisme g* de E* dans E'* par g*('$a_0 a_1 \dots a_n$') = g(a_0) g(a_1) ... g(a_n). Nous nous servirons surtout du cas particulier où g est défini par une application f de E dans E' : g(a) = 'f(a)' pour tout a de E. Nous dirons que l'homomorphisme induit, encore noté f*, est une *transcription* de E* dans E'*. Si f est une bijection, f* est un isomorphisme. D'autre part si f est une application de E dans E', et f' une application de E' dans E'', (f' o f)* = f'* o f*. Le produit de deux transcriptions est une transcription.

1. 4. Ramifications.

1. 4. 1. Définitions. Un graphe fini (E, Γ) est une *ramification* si c'est un graphe sans circuit et si pour tout x de E, il existe au plus un y tel que $y \Gamma x$.

Soient a et b deux points d'une ramification tels que $\Gamma^{-1}(a)$ et $\Gamma^{-1}(b)$ soient vides. Si a et b appartenaient à la même composante connexe du graphe, il existerait un pseudo-chemin formé de points distincts ($a = x_0, x_1, \dots, x_{n-1}$, $x_n = b$) où $a \Gamma x_1$ et $x_{n-1} \bar{\Gamma} b$, et donc un premier entier positif i tel que $x_i \bar{\Gamma} x_{i+1}$; alors $x_{i-1} \Gamma x_i$ et $x_{i+1} \Gamma x_i$, ce qui est impossible. Donc toute composante connexe d'une ramification est une arborescence. Réciproquement, un graphe fini dont toute composante connexe est une arborescence est une ramification.

Le préordre R associé à une ramification est un ordre. Pour la relation R, tout élément strictement minoré possède un prédécesseur : en effet, si y est strictement minoré, il existe un x et un seul tel que $x \Gamma y$; lorsque $z R y$ et $z \neq y$, n points $v_1, v_2, \dots, v_n$ vérifient $z \Gamma v_1 \Gamma v_2 \dots \Gamma v_n \Gamma y$; nécessairement, $v_n = x$ et $z R x$: x est le prédécesseur de y. Nous noterons $x = pd(y)$. D'autre part, y couvre x. D'après le paragraphe 1. 2, (E, Γ) est donc le graphe $[E, R]$.

Comme tout élément strictement minoré possède un prédécesseur, l'ensemble $R^{-1}(x)$ des minorants de x est, pour tout x de E, une R-chaîne. Réciproquement, soit un ensemble E muni d'une relation d'ordre R telle que, pour tout x de E, $R^{-1}(x)$ soit une chaîne, c'est-à-dire où tout élément strictement minoré a un prédécesseur. $[E, R]$ est un graphe (E, Γ) sans circuit. Tout x de E couvre un élément au plus (son prédécesseur), il existe donc au plus un y tel que $y \Gamma x$. $[E, R]$ est une ramification.

Théorème 1. 1. *Les propositions suivantes sont équivalentes :*

a) *le graphe* G *est une ramification ;*
b) *toute composante connexe de* G *est une arborescence ;*
c) G *est un graphe* $[E, R]$ *où, pour tout élément* x *de* E, $R^{-1}(x)$ *est une chaîne ;*
d) G *est un graphe sans circuit où tout élément strictement minoré possède un prédécesseur.*

Les points maximaux d'une ramification pour son ordre associé seront nommés *feuilles*, les points non maximaux *nœuds*, et les points minimaux *racines*. Une arborescence est donc une ramification à une seule racine ; pour son ordre associé, c'est un $\cap$ - demi - treillis [19]. On appellera *familles* d'une ramification (E, Γ) l'ensemble de ses racines et les ensembles $\Gamma(x)$ non vides (c'est-à-dire où x est un nœud) : les familles forment une partition de E.

Soit $[E, R]$ une ramification, E_1 un sous-ensemble de E, R_1 la restriction de R à E_1. D'après le théorème 1. 1, $[E_1, R_1]$ est une ramification, qui sera dite *sous-ramification* de $[E, R]$ (1). On peut toujours considérer une ramification $[E, R]$ comme sous-ramification d'une arborescence $[E_0, R_0]$, obtenue en complétant E par un élément Δ (qui sera la racine de l'arborescence):

$$\Delta \notin E \; ; \; E_0 = E \cup \{\Delta\} \; ; \; (\forall x \in E_0)(\Delta R_0 x).$$

Une sous-ramification de $[E, R]$ qui est une arborescence sera appelée *sous-arborescence* de $[E, R]$. Une sous-arborescence $[E_1, R_1]$ de racine x telle que E_1 soit une partie convexe de E pour R et avec tout élément $y \neq x$ contienne la famille de y, sera dite *complète*.

Une *orientation* d'une ramification (E, Γ) est une relation d'ordre dans E pour laquelle chaque famille est une chaîne et deux éléments appartenant à deux familles différentes ne sont pas comparables ; une ramification munie d'une orientation sera dite *orientée* (2). Dans une ramification orientée, à tout point x est associé un ensemble $\Gamma(x)$ vide ou totalement ordonné, c'est-à-dire un mot $\gamma(x)$ du monoïde libre E^* déduit de l'ensemble E des points. Si on convient que la ramification est complétée en une arborescence comme il vient d'être dit, on notera $\gamma(\Delta)$ la suite de ses racines. Une ramification orientée peut donc être donnée par un couple (E, γ) d'un ensemble E et d'une application γ de $E \cup \{\Delta\}$ dans E^* : nous dirons qu'il s'agit de la ramification orientée (E, γ) ; il est clair cependant que γ n'est pas une application arbitraire de $E \cup \{\Delta\}$ dans E^* : *pour qu'une application de* $E \cup \{\Delta\}$ *dans* E^* *soit une application* γ, *il faut que tout élément de* E *ait une, et une seule, occurrence dans les mots* $\gamma(x)$, *pour* $x \in E \cup \{\Delta\}$ *; si on définit alors* $\Gamma(x)$, *pour* $x \in E$, *comme l'ensemble des points ayant une occurrence dans* $\gamma(x)$ *et si* (E, Γ) *est un graphe sans circuit, c'est une ramification.*

Soit φ un isomorphisme d'une ramification (E, Γ) dans une ramification (E', Γ'), O une orientation de (E, Γ), O' une orientation de (E', Γ'). Si φ est également un isomorphisme du graphe (E, O) dans le graphe (E', O'), nous dirons que c'est un isomorphisme de la ramification (E, Γ) orientée par O dans la ramification (E', Γ') orientée par O'.

1. 4. 2. Réunion horizontale de deux ramifications orientées disjointes. Soient deux ramifications orientées (E_1, γ_1) et (E_2, γ_2) dont les ensembles de points E_1 et E_2 sont disjoints. Désignons par E la réunion de E_1 et E_2 et définissons une application γ de $E \cup \{\Delta\}$ dans le monoïde libre E^*, par :

$$\begin{aligned} \gamma(x) &= \gamma_1(x) && \text{si } x \in E_1, \\ \gamma(x) &= \gamma_2(x) && \text{si } x \in E_2, \\ \gamma(\Delta) &= \gamma_1(\Delta) \; \gamma_2(\Delta). && \end{aligned}$$

(1) $[E_1, R_1]$ n'est pas en général un sous-graphe de $[E, R]$ au sens de [3].

(2) Certains auteurs disent *ordonnée*.

Définissons la relation Γ à partir de γ comme en 1. 4. 1. Soient $(E_1, \Gamma_1) = [E_1, R_1]$ et $(E_2, \Gamma_2) = [E_2, R_2]$ les ramifications obtenues à partir de (E_1, γ_1) et (E_2, γ_2) en faisant abstraction de l'orientation. Lorsque a Γ b, a et b appartiennent tous deux à E_1 ou tous deux à E_2 et selon le cas a Γ_1 b ou a Γ_2 b ; par suite, (E, Γ) est un graphe $[E, R]$ sans circuit : (E, γ) est une ramification orientée, que nous nommerons *réunion horizontale* de (E_1, γ_1) et (E_2, γ_2). L'opération de réunion horizontale est associative, mais non commutative.

$[E_1, R_1]$ et $[E_2, R_2]$ sont des sous-ramifications de $[E, R]$; E_1 et E_2 sont parties convexes de E pour l'ordre R ; toute partie convexe de E_i ($i = 1$ ou 2) pour R_i est partie convexe de E pour R ; *toute sous-arborescence complète de* $[E_i, R_i]$ *est une sous-arborescence complète de* $[E, R]$.

1. 4. 3. Détermination d'une ramification orientée. Un graphe quelconque peut être défini par la donnée, pour tout point x, de l'ensemble $\Gamma(x)$, ou de l'ensemble $\Gamma^{-1}(x)$, ou encore, ce qui revient au même, par un ordre total sur ses points et sa *matrice associée* [3]. Il en est ainsi en particulier pour les ramifications ; alors $\Gamma^{-1}(x)$ contient un point au plus. Si chaque $\Gamma(x)$ est donné sous forme d'une suite $\gamma(x)$, il suffit de connaître de plus un ordre sur les racines pour orienter la ramification. L'ordre total des points qui correspond à l'ordre des lignes et colonnes de la matrice associée définit aussi une orientation par ses restrictions aux différentes familles de la ramification.

Une ramification orientée est aussi déterminée lorsqu'on donne pour tout point x le premier élément x' de $\gamma(x)$, si $\gamma(x)$ n'est pas vide, et le successeur x" de x dans sa famille, si x n'est pas le dernier élément de sa famille, ainsi que la première racine. On obtient une matrice à trois colonnes contenant sur chaque ligne x, x', x", et dont certaines cases sont vides. Nous nommons une telle matrice, *matrice d'enchaînement* ; x' est le *lien vertical* et x" le *lien horizontal* de x. Tout x est caractérisé par le numéro de la ligne associée ; on peut remplacer les liens x' ou x" par le numéro de leur ligne.

1. 4. 4. Extension à un graphe quelconque. On peut tenter d'orienter un graphe quelconque (E, Γ) en se donnant un ordre total O_x dans chaque ensemble $\Gamma(x)$, et un ordre total O_0 dans l'ensemble des éléments y pour lesquels $\Gamma^{-1}(y)$ est vide ; soit O la relation dans E obtenue par réunion des O_x et de O_0. Si O est un ordre dans E pour lequel tout élément majoré a un successeur (autrement dit tel que le graphe $[E, O^{-1}]$ soit une ramification), on peut, comme au paragraphe précédent, définir une matrice d'enchaînement qui détermine le graphe. Pour que O possède ces propriétés, il faut et il suffit que, quels que soient x et x' dans E, les mots obtenus en ordonnant $\Gamma(x)$ par O_x et $\Gamma(x')$ par $O_{x'}$,

s'écrivent $\gamma_1\ \gamma_2$ et $\gamma'_1\ \gamma_2$ où γ_1 et γ'_1 ne possèdent aucun élément commun. Il en est ainsi en particulier si les ensembles $\Gamma(x)$ distincts sont deux à deux disjoints.

1. 4. 5. Ramification des index d'une ramification orientée. Tout point x d'une ramification $[E, R]$ est déterminé lorsqu'on connaît la chaîne de ses R-minorants $(y_0, y_1, \ldots, y_{q-1}, y_q = x)$. L'orientation définit un rang pour chaque élément dans sa famille. Ces rangs peuvent être comptés à partir d'une origine arbitraire ; nous emploierons l'origine 0, un rang sera un entier positif ou nul. x est donc déterminé par la suite des rangs $r_0, r_1, \ldots, r_{q-1}, r_q$ de ses minorants $y_0, y_1, \ldots, y_{q-1}$, x. Cette suite $r_0\, r_1 \ldots r_q$ est un mot du monoïde libre $\mathbf{N}^*$ déduit de l'ensemble $\mathbf{N}$ des entiers naturels, qu'on nomme *index* du point x. Dans deux ramifications orientées isomorphes, les points associés ont même index (démonstration par récurrence sur le nombre des minorants). Ainsi, pour toute ramification orientée, on définit un ensemble I d'index, et deux ramifications orientées isomorphes ont même ensemble d'index. I est un sous-ensemble fini de $\mathbf{N}^*$ qui, avec tout mot de dernier élément r, $\lambda = \lambda_1 r$, contient λ_1 si $\lambda_1 \neq \Lambda$ et les mots $\lambda_1 r'$ pour $0 \leqslant r' \leqslant r$.

Inversement, soit I un sous-ensemble fini de $\mathbf{N}^*$ qui possède cette propriété. Définissons une relation $\sqsubset$ dans I :

$$\lambda \sqsubset \lambda' \text{ si, et seulement si, } \lambda \text{ est facteur gauche de } \lambda'.$$

$[I, \sqsubset]$ est une ramification (théorème 1. 1, c). Il est facile de l'orienter : les mots ayant λ pour prédécesseur sont de la forme λr où $r \in \mathbf{N}$; il suffit de les ordonner suivant les r croissants ; désignons par o cette orientation. On voit aisément par récurrence sur le nombre de minorants de λ que tout λ de I est son propre index : I est un ensemble d'index.

Soit une ramification $[E, R]$, orientée par O. Chacun de ses points x a un index i (x) et

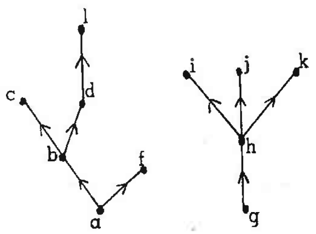
$$x\,R\,y \iff i(x) \sqsubset i(y) \quad ; \quad x\,O\,y \iff i(x)\ o\ i(y).$$

La ramification orientée des index est isomorphe à la ramification orientée donnée.

Deux ramifications orientées ayant même ensemble d'index sont donc isomorphes. Comme réciproquement deux ramifications orientées isomorphes ont même ensemble d'index, la ramification des index caractérise une ramification orientée à un isomorphisme près ; une ramification orientée, dont l'ensemble de points est E, est définie par son ensemble I d'index et une bijection de I sur E ; tout couple formé d'un sous-ensemble I de $\mathbf{N}^*$ qui possède la propriété indiquée plus haut, et d'une bijection de I sur un ensemble quelconque, définit ainsi une ramification.

Certains auteurs, tels que [37], déterminent l'ensemble E et la bijection i par une matrice dite *matrice des index* ; sa première colonne contient les éléments de E, et si p est le nombre maximum de minorants d'un point de la ramification, on peut placer les index dans p autres colonnes, certaines cases restant vides. [37] joint à cette matrice, pour tout point x, le nombre d'éléments de $\Gamma(x)$ et nomme la matrice obtenue "full list matrix". De même, en joignant ce renseignement à la matrice d'enchaînement, il obtient sa "filial-heir matrix". D'autre part, il ordonne les points de la ramification, c'est-à-dire les lignes des matrices, suivant certains ordres privilégiés : nous y reviendrons (2. 1. 2). La même notion d'index a aussi été introduite par [26].

La structure de ramification est fort utilisée en traitement de l'information non numérique : codage, questionnaires, analyse, stockage et recherche d'information, tri... ([33], [48], [51], [55], [56], [9], [58], [31], [37], etc...).



Ramification
(orientée de gauche à droite)

j	-	k
f	-	-
a	b	g
c	-	d
k	-	-
i	-	j
l	-	-
b	c	f
g	h	-
h	i	-
d	l	-

Matrice d'enchaînement

j	1 0 1 -
f	0 1 - -
a	0 - - -
c	0 0 0 -
k	1 0 2 -
i	1 0 0 -
l	0 0 1 0
b	0 0 - -
g	1 - - -
h	1 0 - -
d	0 0 1 -

Matrice des index

Figure 1. 1

1. 5. Quasi isomorphismes de graphes sans circuit.

1. 5. 1. Soient deux ensembles E et E' munis respectivement de relations R et R', et φ une application de E *sur* E' telle que

$$(\forall x \in E)\ (\forall y \in E) \left\{ \begin{array}{l} [x\, R\, y \Longrightarrow \varphi(x)\, R'\, \varphi(y)] \quad \text{et} \\ [\varphi(x)\, R'\, \varphi(y) \Longrightarrow x\, R\, y \text{ ou } (\varphi(x) = \varphi(y) \text{ et } y\, R\, x)] \end{array} \right\}$$

Si R est réflexive, R' l'est aussi et alors

$$(1) \qquad \varphi(x)\, R'\, \varphi(y) \Longleftrightarrow x\, R\, y \text{ ou } \varphi(x) = \varphi(y)$$

$$(2) \qquad \varphi(x) = \varphi(y) \Longrightarrow x\, R\, y \text{ ou } y\, R\, x.$$

Si R est une relation d'ordre, il en est de même pour R'. Nous dirons alors que φ est un *quasi-isomorphisme* du graphe [E, R] sur le graphe [E', R'] et que [E', R'] est quasi-isomorphe à [E, R]. Remarquons que φ est aussi un quasi-isomorphisme de $[E, R^{-1}]$ sur $[E', R'^{-1}]$.

Tout isomorphisme est un quasi-isomorphisme. Tout quasi-isomorphisme injectif est un isomorphisme. Si φ est un quasi-isomorphisme de [E, R] sur [E', R'], à tout x de E', associons un a tel que $x = \varphi(a)$. L'ensemble des a est un sous-ensemble E_1 de E ; si φ_1 et R_1 sont les restrictions de φ et R à E_1, φ_1 est injective et $\varphi_1(x)\, R'\, \varphi_1(y)$ équivaut à $x\, R_1\, y$: les graphes $[E_1, R_1]$ et [E', R'] sont isomorphes.

Si y est minimal pour la relation R, $\varphi(y)$ est minimal pour R' ; en effet, dans ce cas $x\, R\, y$ entraîne $x = y$; donc, d'après (1), $\varphi(x)\, R'\, \varphi(y)$ entraîne $\varphi(x) = \varphi(y)$. Au contraire, si y n'a la même image par φ qu'aucun élément minimal de E, il existe $x \in E$ tel que $x\, R\, y$ et $\varphi(x) \neq \varphi(y)$; d'où $\varphi(x)\, R'\, \varphi(y)$: $\varphi(y)$ n'est pas minimal. Enfin, deux éléments minimaux pour la relation R ne peuvent, d'après (2), avoir même image. φ établit une bijection entre les points minimaux des deux graphes. En remplaçant R et R' par leurs inverses, on en déduit que φ établit une bijection entre les points maximaux. On peut aussi montrer que les deux graphes ont le même nombre de composantes connexes et que φ applique chaque composante connexe de [E, R] sur une composante connexe de [E', R'].

1. 5. 2. φ **définit une relation d'équivalence dans E :**

$$x \sim y \Longleftrightarrow \varphi(x) = \varphi(y).$$

D'après (2), les classes d'équivalence sont des chaînes. D'autre part, si $\varphi(x) = \varphi(y)$ et $x\, R\, z\, R\, y$, alors $\varphi(x)\, R'\, \varphi(z)\, R'\, \varphi(y)$, c'est-à-dire $\varphi(x)\, R'\, \varphi(z)\, R'\, \varphi(x)$; comme la relation R' est antisymétrique, $\varphi(x) = \varphi(z)$: les classes d'équivalence sont des chaînes convexes. Enfin, supposons $x\, R\, y$ et $x\, R\, z$ et $\varphi(x) = \varphi(y)$; alors $\varphi(x)\, R'\, \varphi(z)$, c'est-à-dire $\varphi(y)\, R'\, \varphi(z)$, d'où ($y\, R\, z$ ou $z\, R\, y$). Il en résulte que si deux points couvrent x, x n'est équivalent à aucun d'eux. En remplaçant les relations R et R' par leurs inverses, si x couvre deux points, il n'est équivalent à aucun d'entre eux. En conclusion, les classes d'équivalence sont des chaînes convexes dont tout élément, sauf peut-être le plus grand, a un successeur dans E et, sauf peut-être le plus petit, un prédécesseur dans E : ce sont des suites $(x_1, x_2, \ldots, x_n)$ telles que x_{i+1} soit successeur de x_i et x_i prédécesseur de x_{i+1}, pour $i = 1, 2, \ldots, n$ (figure 1. 2).

Munissons l'ensemble quotient E'' de la relation R'' définie par :

$$C(x)\, R''\, C(y) \Longleftrightarrow \varphi(x)\, R'\, \varphi(y)$$

où C (x) désigne la classe de x : le graphe [E', R'] est isomorphe à [E'', R''].

Réciproquement, soit un graphe $[E, R] = (E, \Gamma)$ et une relation d'équivalence dont toute classe est une suite $(x_1, x_2, \ldots, x_n)$ où x_{i+1} est successeur de x_i et x_i prédécesseur de x_{i+1} : x_{i+1} est le seul élément u tel que $x_i \Gamma u$, x_i est le seul v tel que $v \Gamma x_{i+1}$. D'après la définition de R, fermeture transitive de Γ, si x n'est pas équivalent à y et si $x \sim x'$, x R y entraîne x' R y, y R x entraîne y R x'. La relation (x R y ou $x \sim y$) est donc compatible avec la relation d'équivalence étudiée, ce qui permet de définir une relation dans l'ensemble quotient E' :

$$C(x)\ R'\ C(y) \iff x\ R\ y \text{ ou } x \sim y.$$

Si C (x) R' C (y) et $x \bar{R} y$, alors C (x) = C (y) et donc y R x puisque toute classe est une chaîne. C est un quasi-isomorphisme de [E, R] sur [E', R'].

En particulier, tout graphe [E, R] admet un graphe [E', R'] (et un seul à un isomorphisme près) qui lui soit quasi-isomorphe et qui possède un nombre minimum de points : pour le définir, il suffit de choisir les classes d'équivalence maximales, c'est-à-dire telles que : x_1 ne possède pas de prédécesseur ou son prédécesseur n'a pas de successeur ; x_n ne possède pas de successeur ou son successeur n'a pas de prédécesseur (figure 1. 2). Nous dirons alors que [E', R'] est *graphe réduit* de [E, R]. Il est aussi graphe réduit de tout graphe quasi-isomorphe à [E, R], en particulier de lui-même.

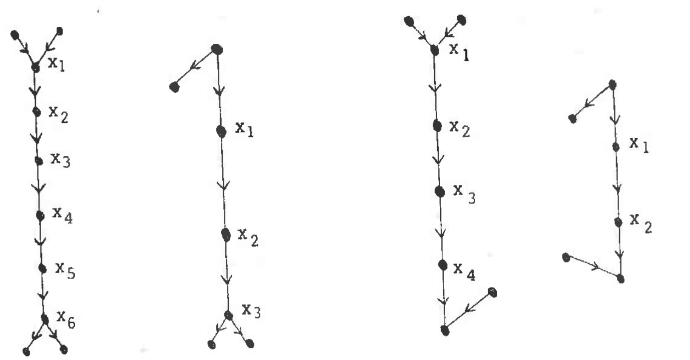


Figure 1.2- Classes d'équivalence maximales

1. 5. 3. Application aux ramifications. Soit φ un quasi-isomorphisme d'une ramification [E, R] sur un graphe [E', R']. Pour tout e de E, $R^{-1}(e)$ est une chaîne. Il résulte alors de (1) que $R'^{-1}[\varphi(e)]$ est aussi une chaîne. Un quasi-isomorphisme transforme une ramification [E, R] en une ramification [E', R'].

D'après le paragraphe 1. 5. 1., [E', R'] est isomorphe à une sous-ramification $[E_1, R_1]$ de [E, R]. φ transforme de manière bijective l'ensemble des feuilles de [E, R] en l'ensemble des feuilles de [E', R'], et toute famille de [E, R] en une famille de [E', R'] : la famille des racines en la famille des racines, la famille de prédécesseur x, lorsqu'elle contient plusieurs éléments, en la famille de prédécesseur $\varphi(x)$.

Soit E_2 l'ensemble des points d'une ramification orientée [E, R] qui ne sont pas des racines. Tout élément x de E_2 possède un prédécesseur pd(x) dans E. L'application pd est un quasi-isomorphisme de $[E_2, R_2]$ sur la sous-ramification des nœuds de [E, R], si on définit R_2 par :

$$x\ R_2\ y \iff (pd(x) \neq pd(y) \text{ et } pd(x)\ R\ pd(y)) \text{ ou } (pd(x) = pd(y) \text{ et } y\ O\ x)$$

où O désigne l'orientation de la ramification.

Pour étudier le graphe réduit (ou *ramification réduite*) d'une ramification [E, R], nous reprendrons les notations du paragraphe 1. 5. 2. Tout point qui n'est pas une racine possède un prédécesseur. Les classes d'équivalence sont donc telles que x_1 soit une racine, ou que plusieurs points couvrent son prédécesseur, et que x_n soit une feuille ou soit couvert par plusieurs points. En associant à toute classe son dernier point x_n, on voit qu'alors [E', R'] est isomorphe à la sous-ramification de [E, R] formée par les feuilles et les nœuds prédécesseurs de plusieurs points. Si tout nœud est le prédécesseur de plusieurs points, [E, R] et [E', R'] sont isomorphes.

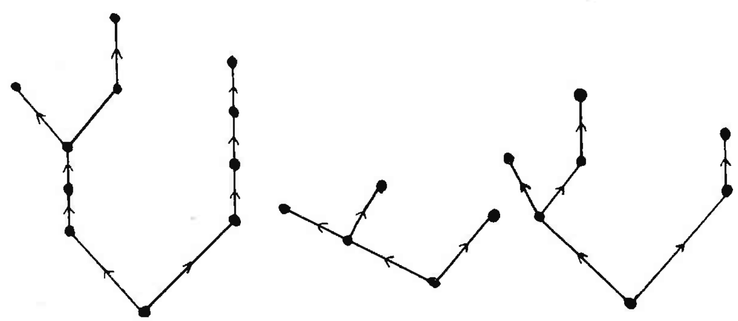
Soit une ramification [E, R]. A tout x de E, associons l'ensemble r (x) des feuilles qui majorent x ; si x est un nœud, r (x) est la réunion des images par r des éléments qui couvrent x. Soit E' l'ensemble des r (x). x R y entraîne $r(x) \supseteq r(y)$. Réciproquement, supposons $r(x) \supseteq r(y)$: x et y minorent tout élément de r (y), donc x R y ou y R x ; et y R x entraîne $r(x) \subseteq r(y)$, d'où r(x) = r(y). r est un quasi-isomorphisme de [E, R] sur $[E', \supseteq]$. D'autre part, si x a un successeur y, r(x) = r(y). Il en résulte que les éléments de toute suite $(x_1 \ldots x_n)$ où tout x_i a x_{i+1} pour successeur, ont même image par r : les classes d'équivalence associées à r sont maximales, au sens du paragraphe 1. 5. 2. ; $[E', \supseteq]$ est une ramification réduite de [E, R].

Construisons maintenant une nouvelle ramification quasi-isomorphe à [E, R], en conservant les classes d'équivalence $(x_1, \ldots, x_n)$ maximales où x_n n'est pas une feuille de [E, R], mais en scindant celles pour lesquelles x_n est une feuille et $n > 1$ en $(x_1, \ldots, x_{n-1})$ et (x_n). La ramification ainsi obtenue est isomorphe à $[\tilde{E}, \tilde{R}]$ où $\tilde{E}$ est la réunion de l'ensemble E" des r (x) associés aux nœuds x et de l'ensemble F des feuilles de [E, R], et où $\tilde{R}$ est défini par :

$$u\ \tilde{R}\ v \iff u \in E'' \text{ et } [(v \in E'' \text{ et } u \supseteq v) \text{ ou } (v \in F \text{ et } v \in u)].$$

Nous appellerons $[\tilde{E}, \tilde{R}]$ *ramification sous-réduite* de $[E, R]$.

Deux ramifications isomorphes ont mêmes ramifications réduites et sous-réduites. Un exemple est schématisé figure 1. 3.



Ramification $[E, R]$ — Ramification réduite $[E', R']$ — Ramification sous-réduite $[\hat{E}, \hat{R}]$

Figure 1. 3.

1. 6. Piles. (1)

1. 6. 1. Définitions. Soit un ensemble E. On appelle *pile sur* E toute suite finie $U = (u_0, u_1, \dots, u_n)$ d'éléments du monoïde libre E^*, telle que,

1) $u_0 = u_n = \Lambda$;

2) pour $i = 1, 2, \dots, n$: - u_{i-1} est facteur gauche de u_i et $|u_i| = |u_{i-1}| + 1$

ou - u_i est facteur gauche de u_{i-1} et $|u_i| = |u_{i-1}| - 1$.

$u_0, u_1, \dots, u_n$ sont les *états de la pile* ; le dernier élément de E dans le mot u_i est le *sommet de l'état* u_i *de la pile*.

Une pile sur l'ensemble E est donc un élément du monoïde libre E^{**} déduit de E^*. Le passage de u_{i-1} à u_i peut se faire de deux manières différentes :

a) il existe $e \in E$ tel que $u_i = u_{i-1}\,'e'$: on dit que i est une *entrée* de e dans la pile (e est le sommet de u_i).

(1) En anglais : *stack, push-down store* ; le mot de *pile* semble avoir été introduit pour la première fois par [25].

b) il existe $e \in E$ tel que $u_{i-1} = u_i\,'e'$: on dit que i est une *sortie* de e de la pile (e est le sommet de u_{i-1}).

Théorème 1. 2. *Soit* u_i *un état non vide d'une pile* U, *et* j *le plus petit entier supérieur à* i *tel que* $|u_j| < |u_i|$.

a) *pour* $i \leqslant k < j$, u_i *est facteur gauche de* u_k ;

b) j *est une sortie du sommet de* u_i.

a se démontre par récurrence sur k ; b en est une conséquence.

Soit une pile U sur un ensemble E. Définissons une relation R dans l'ensemble des éléments de E qui possèdent une entrée dans U :
e R e' si, et seulement si, e possède une occurrence dans tout état de U où e' en possède une.
R est une relation d'ordre ; nous dirons que c'est *l'ordre associé* à la pile U.

Le maximum des longueurs des états d'une pile sera appelé *hauteur* de la pile.

1. 6. 2. Mots liés à une pile. Introduisons un ensemble $\bar{E}$ disjoint de E et ayant même cardinal : à tout élément x de E, on associe $\bar{x}$ de $\bar{E}$, de manière injective ; cette application définit comme il a été dit au paragraphe 1. 3 une transcription de E^* dans $\bar{E}^*$, que nous appellerons *conjugaison*. Soit E_c la réunion de E et $\bar{E}$. A toute pile $U = (u_0, u_1, \dots, u_n)$ sur E, associons une suite de mots sur E_c de la manière suivante :

$$m_0 = \Lambda,$$

$$\text{pour } i = 1, 2, \dots n \quad \begin{cases} m_i = m_{i-1}\,'e_i' & \text{si i est une entrée de } e_i, \\ m_i = m_{i-1}\,'\bar{e}_i' & \text{si i est une sortie de } e_i. \end{cases}$$

Le mot $m = m_n$ caractérise la pile U : nous dirons que m est le *mot attaché* à U.

Les mots m_i et u_i représentent le même élément du groupe libre [18] dont les générateurs sont les éléments de E et où l'inverse de $x \in E$ est $\bar{x} \in \bar{E}$. On montre facilement que pour qu'il existe une pile sur E à laquelle soit attaché un mot donné m sur E_c, il faut et il suffit que tout facteur gauche de m représente le même élément du groupe libre G qu'un mot sur E, et que m représente l'élément neutre du groupe. Les mots attachés aux piles sur E appartiennent au langage de Dyck sur E_c [15]. Nous n'utiliserons pas ces résultats.

La trace (1. 3) de m sur E^* sera appelée *mot d'entrée* de la pile U. La trace de m sur $\bar{E}^*$ est transformée par conjugaison en un mot sur E, qui sera nommé *mot de sortie* de la pile U.

Remplaçons dans m tous les éléments de $\bar{E}$ par un unique caractère σ n'appartenant pas à E. Nous obtenons un mot μ qui caractérise aussi la pile U. Nous dirons que μ est le *mot de description* de la pile U. μ s'écrit $\mu'\mu''$ où μ'' est formé uniquement de caractères σ et où le dernier caractère de μ' n'est pas σ : μ' caractérise aussi la pile ; nous dirons que c'est son *mot de description incomplet*. La donnée du mot de description incomplet est la manière la plus "économique" de définir une pile : si on choisit un code de longueur minimum pour représenter σ, c'est celle qui occupe le moins de place. Pour qu'il existe une pile sur un ensemble E dont un mot donné μ sur $E \cup \{\sigma\}$ soit le mot de description, il faut et il suffit que, dans tout facteur gauche de μ, le nombre d'occurrences de σ soit au plus égal au nombre total d'occurrence des éléments de E, et que dans μ ces deux nombres soient égaux. Pour qu'un mot μ' ne se terminant pas par σ soit le mot de description incomplet d'une pile sur E, il faut et il suffit que la première de ces conditions soit réalisée.

Pour préciser la manière dont le mot de description caractérise une pile, étant donné un mot u sur E et un mot v sur $R \cup \{\sigma\}$ tels que tout facteur gauche de $u\,v$ contienne au moins autant d'occurrences d'éléments de E que d'occurrences de σ, nous définirons $u \boxtimes v$ par les conditions suivantes : si $v = {}'a'\,(a \in E)$, $u \boxtimes v = u\,a$; si $v = {}'\sigma'$ et $u = u'\,a\,(a \in E)$, $u \boxtimes v = u'$; si $v = v'x$, $x \in E \cup \{\sigma\}$, $u \boxtimes v = (u \boxtimes v) \boxtimes x$. Alors $u \boxtimes v'\,v'' = (u \boxtimes v') \boxtimes v''$. Si μ_{ij} est le sous-mot d'un mot de description μ d'une pile U, formé des éléments dont le rang r vérifie $i \leqslant r < j$, $u_j = u_i \boxtimes \mu_{ij}$.

Signalons deux autres mots qu'on peut lier à une pile, bien qu'ils ne soient pas utilisés dans les applications qui se trouvent dans ce travail. Soit i une entrée d'un élément x dans une pile U. Associons-lui le plus petit entier $j(i) > i$ tel que $|u_j| < |u_i|$ et le nombre $r(i) > 0$ des états u_k, pour $i < k < j(i)$, qui ont x pour sommet (c'est-à-dire tels que $u_k = u_i$). $j(i)$ est une sortie de x (théorème 1.2), donc $|u_{i-1}| = |u_{j(i)}| = |u_i| - 1$; l'application j est injective car $i' < i$ et $j(i') = j(i)$ serait contraire à la définition de j. Les $r(i)$, ordonnés suivant les i croissants, forment un mot sur l'ensemble des entiers naturels, que nous appellerons *mot direct des poids* de la pile U ; le mot obtenu en ordonnant les $r(i)$ suivant les $j(i)$ croissants sera appelé *mot inverse des poids*. Toute pile est déterminée par le couple de son mot d'entrée et de son mot direct des poids : l'élément x ayant pour entrée i sort au $r(i)$-ème état qui suit u_i et qui possède x pour sommet.

On trouvera au paragraphe 2. 1. 3 un exemple de pile U avec les divers mots introduits ici.

1. 6. 3. Pile réciproque d'une pile donnée. Soit une pile $U = (u_0, u_1, \ldots, u_n)$. La pile $U' = (u_n, u_{n-1}, \ldots, u_1, u_0)$ (qui est la réfléchie de U dans le monoïde libre E^{**}) sera nommée *pile réciproque* de U. U est la pile réciproque de U'. Si i est une entrée (resp. sortie) de e dans U, $n-i+1$ est une sortie (resp. entrée) de e dans U'. Le mot m' attaché à U' est le réfléchi du conjugué du mot m attaché à U : m et m' représentent deux éléments inverses du groupe libre G. Le mot d'entrée de U' est le réfléchi du mot de sortie de U, le mot direct des poids de U' est le réfléchi du mot inverse des poids de U. Il résulte du paragraphe précédent qu'une pile est déterminée par son mot de sortie et son mot inverse des poids. U et U' ont le même ordre associé.

1. 6. 4. Trace d'une pile. Soit une pile $U = (u_0, u_1, \ldots, u_n)$ sur un ensemble E, et E' un sous-ensemble de E. Désignons par u'_i la trace de u_i sur E' (1.3). Envisageons la suite des entiers qui sont des entrées dans U ou des sorties de U d'un élement de E' : supposons qu'il existe p tels entiers et soit j_k le k-ème, enfin posons $j_0 = 0$. Si $j_k \leqslant i < j_{k+1}$, $u'_i = u'_{j_k}$; si $j_p \leqslant i \leqslant n$, $u'_i = u'_{j_p}$. On passe de u'_{j_k} à $u'_{j_{k+1}}$ par l'entrée ou la sortie d'un élément de E' ; $u'_{j_0} = u'_0 = \Lambda$ et $u'_{j_p} = u'_n = \Lambda$. La suite $U' = (u'_{j_0}, u'_{j_1}, \ldots, u'_{j_p})$ est donc une pile sur E' : nous dirons que c'est la *trace* de U sur E'.

A la pile U est associé un ordre R, et à U' un ordre R'. D'après la définition de R et R', si x et y appartiennent à E', x R y entraîne x R' y. Réciproquement, supposons x R' y ; soit i un entier tel que u_i contienne une occurrence de y, et k le plus grand entier pour lequel $j_k \leqslant i$: il existe une occurrence de y dans u_{j_k}, donc aussi une occurrence de x, puisque x R' y ; par suite, x a une occurrence dans u_i : x R y. R' est donc la restriction de R à l'ensemble des éléments de E' qui possèdent une entrée dans U'.

1. 6. 5. Transcription d'une pile. Nous avons signalé au paragraphe 1. 3 que toute application f d'un ensemble E dans un ensemble E' définit un homomorphisme f^* (transcription) du monoïde libre E^* dans E'^*, d'où un homomorphisme f^{**} de E^{**} dans E'^{**}. Une pile U sur E est ainsi transformée en une pile sur E' :

$$U = (u_0, u_1, \ldots, u_n) \longrightarrow U' = (f^*(u_0), f^*(u_1), \ldots, f^*(u_n)).$$

Nous dirons que la pile U' est *transcrite* (par f) de la pile U. Le mot d'entrée et le mot de sortie de U' sont les images par f^* des mots d'entrée et de sortie de U.

Si, ainsi qu'on a associé à E un ensemble $\bar{E}$, on associe à E' un ensemble de conjugués $\bar{E}'$, et si on prolonge f à $\bar{E}$ par $f(\bar{x}) = \overline{f(x)}$, f^* est prolongé à $\bar{E}^*$, et le mot attaché à U' est l'image par f^* du mot attaché à U ; de même, si on prolonge f par $f(\sigma) = \sigma$, le mot de description et le mot de description incomplet de U' sont les images par f^* du mot de description et du mot de description incomplet de U.

Lorsque f est une bijection, f^* et f^{**} sont des isomorphismes. Nous dirons alors que les piles U et U' sont *isomorphes*. Notons encore que, si g est une application de E' dans E'', $(g \circ f)^{**} = g^{**} \circ f^{**}$.

1. 7. Piles simples.

Définition. *Une pile* U *sur un ensemble fini* E *est* simple *lorsque tout élément de* E *possède une entrée et une seule dans* U.

Tout élément de E possède une occurrence et une seule dans le mot d'entrée d'une pile simple sur E : l'ordre de ces occurrences détermine un ordre total P dans E, que nous appellerons *ordre d'entrée* de la pile simple. Le mot m attaché à une pile simple représente l'élément neutre du groupe libre G défini au paragraphe 1. 6. 2 ; il contient une fois et une seule chaque élément de E, donc une fois et une seule chaque élément de $\bar{E}$. Il en résulte que le mot de sortie d'une pile simple sur E contient une occurrence et une seule de tout élément de E et définit de même un ordre total S dans E, qui sera nommé *ordre de sortie* de la pile. Si E est formé de k éléments, une pile simple sur E possède k entrées et k sorties, donc 2k+1 états. Tout élément de E a au plus une occurrence dans chacun de ces états. Ceci permettra de dire, par abus de langage, que e appartient à u_i, ou que u_i contient e, au lieu de dire que e possède une occurrence dans u_i.

La réciproque d'une pile simple U est une pile simple U' ; l'ordre d'entrée (resp. de sortie) de U' est l'inverse de l'ordre de sortie (resp. d'entrée) de U. La trace, sur un sous-ensemble E' de E, d'une pile simple U sur E est une pile simple U' ; les ordres d'entrée et de sortie de U' sont les restrictions à E' des ordres d'entrée et de sortie de U.

Toute pile est transcrite (1. 6. 5) d'une pile simple. Soit en effet U une pile sur un ensemble E, et n la longueur de son mot d'entrée : l'application f qui associe à l'entier i $(0 \leqslant i \leqslant n-1)$ l'élément de E dont i est occurrence dans le mot d'entrée de U, définit une transcription f^{**} qui transforme en U une pile simple sur l'ensemble des entiers de 0 à n-1. De cette propriété vient l'importance des piles simples. Nous y reviendrons au chapitre 3. Si une pile simple U est transcrite d'une pile simple U', U et U' sont isomorphes.

Soit U une pile simple sur un ensemble E. Désignons respectivement par $\varepsilon(e)$ et $\sigma(e)$ l'entrée et la sortie d'un élément e de E. e appartient aux états u_i tels que $\varepsilon(e) \leqslant i < \sigma(e)$ et à eux seuls. L'ordre associé à la pile U (qui est ici défini dans l'ensemble E lui-même) vérifie donc :

$$e \, R \, e' \iff \varepsilon(e) \leqslant \varepsilon(e') \text{ et } \sigma(e') \leqslant \sigma(e).$$

Théorème 1. 3. $e \, R \, e' \iff e \, P \, e' \text{ et } e' \, S \, e.$

Nous déduirons de ce théorème une expression de S en fonction de R et P, et une expression de P en fonction de R et S.

Théorème 1. 4. $e \, S \, e' \iff e' \, R \, e \text{ ou } (e \, P \, e' \text{ et } e \, \bar{R} \, e').$

En effet, grâce au théorème 1. 3, les propositions suivantes sont équivalentes :

$e' \, R \, e$ ou $(e \, P \, e'$ et $e \, \bar{R} \, e')$
$(e' \, P \, e$ et $e \, S \, e')$ ou $[e \, P \, e'$ et $(e \, \bar{P} \, e'$ ou $e' \, \bar{S} \, e)]$
$(e' \, P \, e$ et $e \, S \, e')$ ou $(e \, P \, e'$ et $e' \, \bar{S} \, e)$
$(e' \, P \, e$ et $e \, S \, e')$ ou $(e \, P \, e'$ et $e \, S \, e'$ et $e \neq e')$
$e \, S \, e'$ et $(e' \, P \, e$ ou $e' \, \bar{P} \, e)$
$e \, S \, e'$.

On a utilisé le fait que P et S étaient des ordres totaux.

La démonstration prouve d'autre part :

Théorème 1. 5. *Etant donnés une relation* R *et un ordre total* P, *il existe au plus un ordre total* S *tel que* $e \, R \, e' \iff e \, P \, e'$ et $e' \, S \, e$.

Théorème 1. 6. $e \, P \, e' \iff e \, R \, e' \text{ ou } (e \, S \, e' \text{ et } e' \, \bar{R} \, e).$

Il suffit d'appliquer le théorème 1. 4. à la pile réciproque de la pile simple étudiée : son ordre d'entrée est S^{-1}, son ordre de sortie P^{-1}, son ordre associé R.

De plus, étant donnés une relation R et un ordre total S, il existe au plus un ordre total P tel que $e \, R \, e' \iff e \, P \, e'$ et $e' \, S \, e$.

Soit maintenant un élément e de E, $i = \varepsilon(e)$, $j = \sigma(e)$. D'après le théorème 1. 2 , pour $i \leqslant k < j$, u_i est le facteur gauche de u_k qui se termine par e. Par définition de R, l'ensemble $R^{-1}(e)$ des minorants de e pour la relation R est l'ensemble des éléments de u_i. Si e' et e'' sont deux minorants de e et si e' précède e'' dans u_i, c'est donc un minorant de e''.

Théorème 1. 7. *Soit* R *la relation d'ordre associée à une pile simple* U *sur un ensemble* E, *et* e *un élément de* E. *L'ensemble des* R*-minorants de* e *est une* R*-chaîne, et la suite de ces minorants, ordonnés par* R, *est le facteur gauche se terminant par* e *de tout état de la pile* U *qui contient* e.

Le théorème 1. 7 va nous permettre de rapprocher les notions de pile simple et de ramification. Les théorèmes 1. 3, 1. 4 et 1. 6 prouvent que si on donne deux des ordres R, P, S, le troisième est déterminé. Nous montrerons que ces ordres déterminent la pile, autrement dit qu'il existe au plus une pile simple possédant un ordre R pour ordre associé et un ordre total P pour ordre d'entrée : cette démonstration sera faite au paragraphe 2. 2 où on donnera un algorithme de construction d'une telle pile. Nous étudierons ensuite (2. 3 et 2. 8) à quelles conditions un ordre R et un ordre total P, ou un ordre R et un ordre total S, ou deux ordres totaux P et S sont effectivement l'ordre associé et l'ordre d'entrée, ou l'ordre associé et l'ordre de sortie, ou l'ordre d'entrée et l'ordre de sortie d'une pile simple.

CHAPITRE II

PILES SIMPLES ET RAMIFICATIONS

2. 1. Pile attachée à une ramification. Pile strictement attachée à une ramification orientée.

2. 1. 1. D'après les théorèmes 1. 1 et 1. 7, *si* R *est la relation d'ordre associée à une pile simple* U *sur un ensemble* E, [E, R] *est une ramification.* De plus :
(PR1) si x a pour entrée j, ou bien u_{j-1} est vide et x est une racine de [E, R], ou bien le sommet de u_{j-1} est le R-prédécesseur de x ; dans les deux cas, x n'a pas d'entrée inférieure à j.
(PR2) si x a pour sortie j, tout élément dont x est prédécesseur a une entrée inférieure à j (car si k est cette entrée, x doit appartenir à u_k).
(PR3) toute racine de [E, R] possède une entrée dans la pile U.

Inversement, soit une ramification [E, R]. Montrons que les piles U sur E qui satisfont à ces trois consitions sont simples et que R est la relation d'ordre qui leur est associée. Toute racine possède une entrée dans U. Par récurrence sur le nombre des minorants de x, il en est de même pour tout point x de E ; car si le prédécesseur de x possède une entrée, il a aussi une sortie j, puisque le dernier état de la pile est vide ; d'après PR2, x a une entrée inférieure à j. D'autre part, d'après PR1, tout élément possède au plus une entrée : U est une pile simple. Une relation d'ordre R' lui est associée. D'après le théorème 1. 7 et la condition PR1, si un élément de E est une racine de [E, R], c'est aussi une racine de [E, R'], et sinon il a un prédécesseur pour R, qui est aussi son prédécesseur pour R'. Les relations R et R' sont confondues.

Ainsi lorsque [E, R] est une ramification, les piles simples U auxquelles la relation d'ordre R est associée sont les piles qui satisfont aux conditions PR1, PR2, PR3. On dira que ces piles sont *attachées* à [E, R]. L'ordre d'entrée et l'ordre de sortie d'une pile attachée à [E, R] sont, par définition, *un ordre d'entrée et un ordre de sortie de la ramification* [E, R] ; ces deux ordres seront dits *associés*. Si une pile simple est attachée à une ramification, sa pile réciproque l'est aussi : les ordres de sortie d'une ramification sont les inverses de ses ordres d'entrée. Les piles attachées à une ramification [E, R] ont toutes même hauteur, le nombre maximum de R-minorants d'un point de E : on appelle ce nombre *hauteur* de la ramification.

Les conditions PR1, PR2, PR3 conduisent à un algorithme de construction des piles U attachées à une ramification (E, Γ) = [E, R] :

a) $u_0 = \Lambda$;
b) lorsque u_{j-1} n'est pas vide et a pour sommet e :
- si tout point de Γ(e) possède une entrée inférieure à j, j est la sortie de e ;
- sinon, on choisit un point y de Γ(e) qui n'a pas d'entrée inférieure à j : j est l'entrée de y.

c) lorsque u_{j-1} est vide :
- si toute racine possèdeune entrée inférieure à j, u_{j-1} est le dernier état de la pile ;
- sinon, on choisit une racine y qui n'a pas d'entrée inférieure à j : j est l'entrée de y.

2. 1. 2. Cet algorithme prouve l'existence, pour toute ramification, de piles simples qui lui sont attachées. L'une d'elles est déterminée lorsqu'on impose l'ordre dans lequel entrent les éléments d'une même famille, c'est-à-dire lorsqu'on oriente la ramification ; ainsi, à toute ramification orientée, on associe l'une des piles qui lui sont attachées, et qui sera dite *strictement attachée* à cette ramification orientée : celle où les points de chaque famille entrent dans l'ordre d'orientation. On appelle *ordre d'entrée* et *ordre de sortie de la ramification orientée* l'ordre d'entrée et l'ordre de sortie de la pile simple qui lui est strictement attachée. Réciproquement, on a vu que toute pile simple U est attachée à une ramification [E, R] ; son ordre d'entrée (ou de sortie) définit une orientation de [E, R] grâce à ses restrictions aux différentes familles : U est strictement attachée à la ramification [E, R] ainsi orientée. La pile réciproque de U est d'ailleurs strictement attachée à la même ramification, mais munie de l'orientation inverse. L'algorithme prouve aussi que les piles simples strictement attachées à deux ramifications orientées isomorphes par une application φ, sont transcrites par φ (1. 6. 5) et donc aussi isomorphes.

Théorème 2. 1. *Toute pile simple est strictement attachée à une ramification orientée unique : la pile et la ramification ont le même ordre associé. A deux ramifications orientées isomorphes sont strictement attachées des piles simples isomorphes.*

Une ramification orientée est donc définie par le mot attaché m, ou le mot de description μ, ou le mot de description incomplet μ' de la pile simple qui lui est strictement attachée ; μ' donne une représentation qui occupe un minimum de place. Elle est aussi définie par le mot d'entrée et le mot direct des poids (ou le mot de sortie et le mot inverse des poids) de sa pile attachée : le "poids" r(i) associé à l'élément x d'entrée i (1. 6. 2) est le nombre de points de la ramification dont x est le prédécesseur. Le mot direct (ou inverse) des poids, seul, détermine une ramification orientée à un isomorphisme près (cf. [26]). Dans les matrices qui représentent une ramification orientée (1. 4. 3), on peut adopter l'ordre d'entrée pour ordre des noeuds. [37] qualifie de "gauche" cet ordre et les matrices ainsi construites. Le lien vertical x' de x, s'il existe, suit alors x : dans la matrice d'enchaînement, il suffit d'indiquer son existence.

La connaissance de la pile strictement attachée à une ramification orientée permet de résoudre facilement les problèmes suivants :

- recherche des points de prédécesseur donné, détermination des familles ordonnées suivant l'orientation, construction de la matrice associée.
- construction de la matrice d'enchaînement où les points sont ordonnés dans l'ordre d'entrée ; cette matrice peut être aisément obtenue à partir du mot m attaché à la pile : si dans m l'occurrence de x est suivie d'une occurrence d'un élément x' de E, x' est le premier élément de $\Gamma(x)$ (lien vertical de x), sinon $\Gamma(x)$ est vide ; si l'occurrence de $\bar{x}$ est suivie d'une occurrence d'un élément x'' de E, x'' est le successeur de x dans sa famille (lien horizontal de x), sinon x est le dernier élément de sa famille.
- construction de la matrice des index où les points sont ordonnés par l'ordre d'entrée : on place dans la pile, avec chaque élément, son rang dans sa famille.
- recherche des minorants d'un élément donné : tout état u_i de la pile est le chemin de la ramification qui a une racine pour origine et pour extrémité le sommet de u_i. Pour que les piles attachées à une ramification ne possèdent aucun état, sauf le premier et le dernier, qui soit vide, il faut et il suffit que cette ramification soit une arborescence : la première entrée et la dernière sortie sont celles de la racine de l'arborescence.

Soit une ramification orientée [E, R], x un de ses points, U la pile strictement attachée. Les majorants de x constituent une sous-arborescence complète de [E, R], de racine x : on peut orienter cette arborescence par la restriction de l'orientation de [E, R]. La pile U_x qui lui est alors strictement attachée est construite par l'algorithme indiqué plus haut : on voit de proche en proche que ses états non vides sont les facteurs droits qui commencent par x des états de U dont les indices sont compris entre l'entrée et la sortie de x. Le mot m_x attaché à U_x est l'intervalle $[x, \bar{x}]$ du mot m attaché à U. Il en résulte qu'il est facile de traduire sur les mots m des opérations telles que le transfert d'un sous-graphe dans une arborescence ou le produit d'arborescences définis par [48]. Notons d'autre part que R(x) *est formé des éléments de* E *dont une occurrence se trouve dans* $m_x = [x, \bar{x}]$.

2. 1. 3. Construction de la pile U strictement attachée à une ramification orientée (E, γ). Lorsque la ramification donnée est une arborescence, l'algorithme indiqué plus haut se simplifie, car aucun état de la pile n'est vide, sauf le premier et le dernier. Dans le cas général, nous compléterons la ramification en une arborescence par l'introduction d'une racine Δ (1. 4. 1) : les racines de la ramification donnée forment alors $\gamma(\Delta)$. Supposons que pour tout x de $E \cup \{\Delta\}$ on donne le mot $\gamma(x)$ du monoïde libre E^*. L'algorithme qui construit

la pile simple U est décrit par l'organigramme de la figure 2. 1 : e désigne le sommet de la pile.

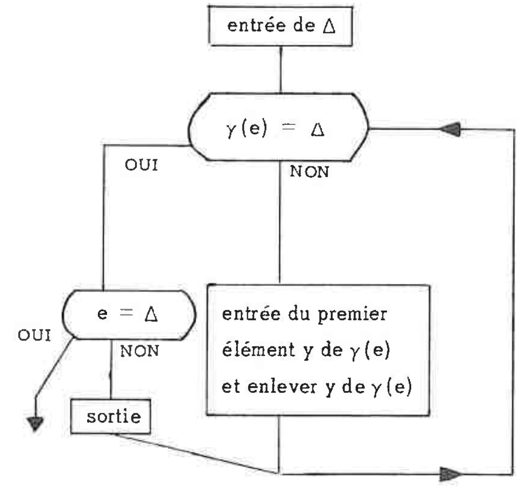


Figure 2. 1

Exemple (figure 2. 2.) :

$\gamma(\Delta)$ = 'a b'
$\gamma(a)$ = Λ
$\gamma(b)$ = 'c d'
$\gamma(c)$ = Λ
$\gamma(d)$ = 'f g h'
$\gamma(f)$ = Λ
$\gamma(g)$ = Λ
$\gamma(h)$ = Λ

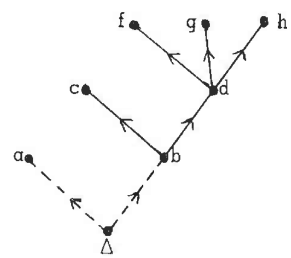


Figure 2. 2

Etats successifs de la pile :

							f		g		h			
				c		d	d	d	d	d	d	d		
	a		b	b	b	b	b	b	b	b	b	b	b	
Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ
u_1	u_2	u_3	u_4	u_5	u_6	u_7	u_8	u_9	u_{10}	u_{11}	u_{12}	u_{13}	u_{14}	u_{15}

Ordre d'entrée P : a, b, c, d, f, g, h.
Ordre de sortie S : a, c, f, g, h, d, b.
Mot attaché à la pile strictement attachée à la ramification orientée :
$m =$ '$a\,\bar{a}\,b\,c\,\bar{c}\,d\,f\,\bar{f}\,g\,\bar{g}\,h\,\bar{h}\,\bar{d}\,\bar{b}$'.
Mot de description : $\mu =$ '$a\,\sigma\,b\,c\,\sigma\,d\,f\,\sigma\,g\,\sigma\,h\,\sigma\,\sigma\,\sigma$'.
Mot de description incomplet : $\mu' =$ '$a\,\sigma\,b\,c\,\sigma\,d\,f\,\sigma\,g\,\sigma\,h$'.
Mot direct des poids : '0 2 0 3 0 0 0'.
Mot inverse des poids : '0 0 0 0 0 3 2'.

2. 1. 4. Modification de l'algorithme. Pour construire la pile U cherchée, on peut utiliser une pile auxiliaire U' sur l'ensemble des $\gamma(x)$: u'_i contient les $\gamma(x)$ associés aux x de u_i (cf. paragraphe 1. 6. 5 : U' est une pile transcrite de U) ; quand x entre dans U, $\gamma(x)$ entre dans U'. D'autre part, pour éviter que $\gamma(x)$ ne devienne vide, on peut le terminer par un signe auxiliaire ǂ : quand x sort de U, $\gamma(x)$ qui se réduit à ǂ, sort de U'. Il est en réalité plus commode de remplacer U' par une pile U'' sur E∪{ǂ}, en faisant entrer dans U'', lorsque x entre dans U, ǂ et, dans l'ordre, les éléments du mot réfléchi $\tilde{\gamma}(x)$ de $\gamma(x)$. La figure 2. 3 décrit l'algorithme ainsi obtenu ; e et e'' désignent les sommets respectifs de U et U''. Une justification directe de cet algorithme sera donnée au paragraphe 2. 9. 1. Pour l'exemple déjà étudié plus haut, on a figuré au dessus de chaque état de U un état correspondant de U'' :

						f	ǂ							
						g	g	g	ǂ					
			c	ǂ		h	h	h	h	h	ǂ			
a	ǂ		d	d	d	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ		
b	b	b	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	
ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ
u''_3	u''_5	u''_6	u''_{10}	u''_{12}	u''_{13}	u''_{18}	u''_{20}	u''_{21}	u''_{23}	u''_{24}	u''_{26}	u''_{27}	u''_{28}	u''_{29}

							f		g		h			
				c		d	d	d	d	d	d	d		
	a		b	b	b	b	b	b	b	b	b	b	b	
Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ

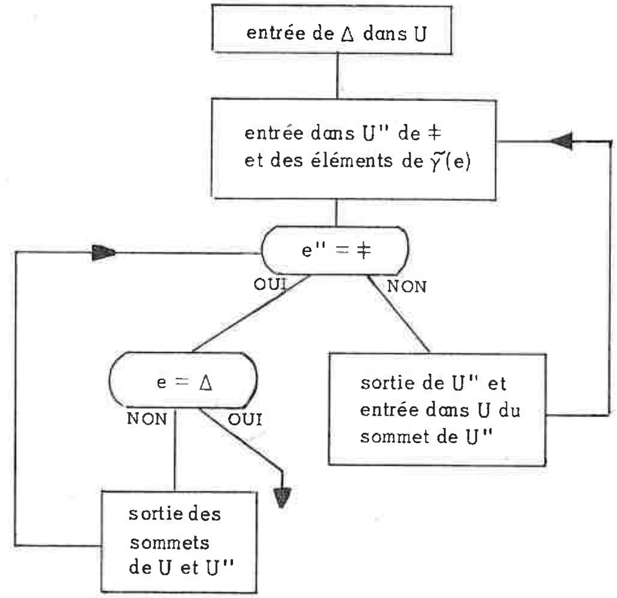


Figure 2. 3.

On peut encore modifier l'algorithme : on teste si l'élément e'' qui va sortir de U'' est une feuille ; dans ce cas, il sortira de U dès son entrée ; sinon, le premier élément de γ(e'') entre dans U aussitôt après e'' ; il est inutile de le faire entrer d'abord dans U''. En particulier, dans le cas d'une *ramification binaire*, c'est-à-dire d'une ramification où, pour tout noeud x, $\gamma(x)$ a pour longueur 2, la moitié seulement des points entre dans U''.

Il est plus facile de construire la pile U si la ramification est donnée par une matrice d'enchaînement : après l'entrée d'un élément x entre son lien vertical s'il existe, sinon x sort ; après la sortie de x entre son lien horizontal s'il existe, sinon se produit une nouvelle sortie.

2. 1. 5. Piles attachées à un graphe quelconque. On peut étendre à un graphe quelconque la notion de pile attachée à une ramification, en partant de l'algorithme de construction (2. 1. 1). Nous disons qu'une pile U sur un ensemble E est *attachée à un graphe* (E, Γ), si elle possède les propriétés suivantes :

a) lorsqu'un état u_{j-1} de U n'est pas vide et a pour sommet e :
 - si tout point de Γ(e) possède une entrée inférieure à j, j est la sortie de e ;
 - sinon, j est l'entrée d'un point de Γ(e) ne possédant pas d'entrée inférieure à j.

b) lorsque u_{j-1} est vide :
 - si tout élément de E a une entrée inférieure à j, u_{j-1} est le dernier état de la pile ;
 - sinon, j est une entrée d'un élément qui ne possède pas d'entrée inférieure à j.

Toute pile attachée à (E, Γ) est une pile simple sur E. Elle est attachée à une ramification qui possède les mêmes points que (E, Γ) et en est un graphe partiel (1).

On pourra trouver dans [22] des applications de cette notion à la détermination de la fermeture transitive d'une relation ou des chemins élémentaires d'un graphe.

2. 2. Construction d'une pile simple connaissant son ordre associé et son ordre d'entrée.

2. 2. 1. Soit une pile simple U sur un ensemble E, R et P son ordre associé et son ordre d'entrée. Au passage de u_{j-1} à u_j, deux cas sont possibles :

a) j est l'entrée d'un élément z, qui est le premier élément (dans l'ordre P) non encore entré.

b) le sommet de u_{j-1} sort de la pile.

a se présente lorsque u_{j-1} est vide, et b lorsque tous les éléments de E sont entrés. Ecartons ces deux possibilités, et soit e le sommet de u_{j-1}, z le premier élément non encore entré. Dans le cas a, eRz (théorème 1. 7). Dans le cas b, la sortie de e précède celle de z : eSz ; comme e ≠ z, eSz n'est pas compatible avec eRz (théorème 1. 4) ; donc e$\bar{R}$z. Ainsi, pour que z entre dans la pile, il faut et il suffit que eRz.

(1) Un graphe partiel de (E, Γ) est un graphe (E, Γ') tel que, pour tout x de E, Γ(x) contienne Γ'(x) [3].

Pour éviter le cas u_{j-1} vide, on complète la ramification [E, R] en une arborescence en lui adjoignant une racine Δ (paragraphe 1. 4. 1) : Δ est le premier élément à entrer. D'autre part, pour que la dernière entrée marque la fin du processus, on introduit un nouvel élément Φ tel que :

$$(\forall x \in E)\ (x \bar{R} \Phi \quad \text{et} \quad x P \Phi) \quad \text{et} \quad \Delta R \Phi.$$

A l'entrée de Φ, tous les éléments de E sont sortis ; le déroulement de l'algorithme est terminé. Cet algorithme est schématisé sur l'organigramme, figure 2. 4.

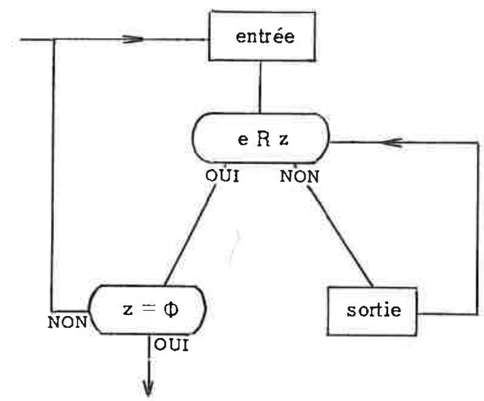


Figure 2. 4.

L'étude faite prouve *qu'il existe au plus une pile simple ayant un ordre associé* R *et un ordre d'entrée* P *donnés.*

On peut utiliser cet algorithme pour construire la pile strictement attachée à une ramification orientée [E, R] dont on donne la matrice des index où les points sont ordonnés par P. Alors e R z équivaut à : l'index de e est un facteur gauche de l'index de z ; pour que z entre, il suffit (et il faut) d'ailleurs que son index soit plus long que celui de e, car l'index du premier élément qui entre après la sortie de e a une longueur au plus égale à celle de l'index de e.

2. 2. 2. Etude des couples (e, z) sur lesquels porte le test. *Il s'agit des couples tels qu'il existe un entier* j *pour lequel* e *est le sommet de* u_{j-1} *et* z *l'élément dont l'entrée est la plus petite, supérieure ou égale à* j. Alors :

a) e P z et e ≠ z ;

b) soit $v \in\,]_P\, e, z[$; ε(v) < j d'après la définition de z ; σ(v) < j car sinon v appartiendrait à u_{j-1} et il en résulterait vRe (th. 1. 7), qui contredit ePv et e ≠ v (théorème 1. 3) ; d'autre part, σ(e) ⩾ j ; d'où vSe, et donc eRv (théorème 1. 3).

$c)\ \sigma(v) < \sigma(z)$: $v\ S\ z$; par suite $v\ \bar{R}\ z$.

Réciproquement, soit un couple (e,z) d'éléments de E, tel que :

$$e\ P\ z \quad \text{et} \quad e \neq z \quad \text{et} \quad (\forall v \in \,]_P\, e,z[) \quad (e\ R\ v \quad \text{et} \quad v\ \bar{R}\ z).$$

Si l'intervalle $]_P\ e,z[$ n'est pas vide, soit j-1 la dernière sortie d'un de ses éléments, v. L'entrée de z est supérieure ou égale à j, car sinon on aurait $\varepsilon(v) < \varepsilon(z) < \sigma(v)$, d'où $v\ R\ z$; l'entrée de z est la première qui est supérieure ou égale à j. e se trouve dans u_{j-1}, car $e\ R\ v$ (théorème 1. 7) ; le sommet e' de u_{j-1} vérifie donc $e\ R\ e'$, d'où $e\ P\ e'$: e' ne peut être que e.

Théorème 2. 2. *Pour qu'un couple (x,y) d'éléments de E soit un couple (e,z) sur lequel porte un test dans l'algorithme du paragraphe 2. 2. 1, il faut et il suffit que*

$$x\ P\ y \quad \text{et} \quad x \neq y \quad \text{et} \quad (\forall v \in \,]_P\, x,y[) \quad (x\ R\ v \quad \text{et} \quad v\ \bar{R}\ y).$$

2. 3. Conditions sur l'ordre associé et l'ordre d'entrée ou de sortie d'une pile simple.

2. 3. 1. Théorème 2. 3. *Soit P un ordre total et R une relation transitive dans un ensemble fini E. Pour qu'il existe une pile simple dont P soit l'ordre d'entrée et R l'ordre associé, il faut et il suffit que, pour tout x appartenant à E, R(x) soit un P-intervalle ayant x pour premier élément. Cette pile simple est alors unique.*

Nous avons vu au paragraphe 2. 1. 2 que si m est le mot attaché à une pile simple sur l'ensemble E, pour tout x de E, R(x) est l'intersection avec E de l'intervalle $[x, \bar{x}]$ du mot m. Comme le mot d'entrée est la trace de m sur E, R(x) est un intervalle du mot d'entrée possédant x pour premier élément.

Réciproquement, soit un ensemble E muni d'une relation transitive R et P un ordre total sur E, tel que pour tout x de E, R(x) soit un P-intervalle ayant x pour premier élément. Construisons une pile simple U sur E dont P soit l'ordre d'entrée, en utilisant l'algorithme du paragraphe 2. 2. 1 : au passage de u_{j-1} à u_j, désignons par z le premier élément de E (dans l'ordre P) non encore entré dans la pile, s'il existe de tels éléments, et soit e le sommet de u_{j-1} si u_{j-1} n'est pas vide.

a) lorsque u_{j-1} n'est pas vide et que z existe : si $e\ R\ z$, z entre dans u_j, si $e\ \bar{R}\ z$, e sort de la pile ;
b) lorsque u_{j-1} n'est pas vide et que z n'existe pas, e sort de la pile ;
c) lorsque u_{j-1} est vide : si z existe, il entre dans u_j ; si z n'existe pas, u_{j-1} est le dernier état de la pile.

Soit R' l'ordre associé à la pile U. Si y couvre x pour la relation R', il existe un entier j tel que x soit le sommet de u_{j-1} et que j soit l'entrée de y, et donc $x\ R\ y$. Par suite, puisque l'ensemble E est fini et que la relation R est transitive, $x\ R'\ y$ entraîne $x\ R\ y$.

D'après la partie directe du théorème, R'(x) est un P-intervalle dont x est le premier élément. Il en est de même pour R(x) et de plus R'(x) est contenu dans R(x). Si le dernier élément de E pour l'ordre P appartient à R'(x), R'(x) = R(x). Sinon, soit y le successeur, pour l'ordre total P, du dernier élément de R'(x) : si v appartient au P-intervalle]x, y[, $x\ R'\ v$ mais $v\ \bar{R}'y$ puisque $x\ \bar{R}'\ y$. D'après le théorème 2. 2, (x, y) est donc un couple (e, z) pour la pile construite ; comme $x\ \bar{R}'y$, x sort de la pile lors de sa comparaison avec y : donc $x\ \bar{R}\ y$. Les intervalles R(x) et R'(x) sont encore confondus. R est l'ordre associé à la pile construite.

L'unicité de la pile simple ayant un ordre associé et un ordre d'entrée donnés a été démontrée au paragraphe 2. 2. 1.

2. 3. 2. L'ordre de sortie d'une pile simple est l'inverse de l'ordre d'entrée de la pile simple réciproque. Et une pile simple a même ordre associé que sa réciproque. D'où :

Théorème 2. 4. *Soit S un ordre total et R une relation transitive dans un ensemble fini E. Pour qu'il existe une pile simple dont S soit l'ordre de sortie et R l'ordre associé, il faut et il suffit que, pour tout x appartenant à E, R(x) soit un S-intervalle ayant x pour dernier élément. Cette pile simple est alors unique.*

2. 3. 3. La condition énoncée au théorème 2. 3 (resp. 2. 4) est aussi nécessaire et suffisante pour que [E, R] soit une ramification et que P (resp. S) en soit un ordre d'entrée (resp. de sortie).

Une première conséquence est que si y est un noeud d'une ramification [E, R] orientée, d'ordre d'entrée P, et si $x_1, x_2, \ldots, x_n$ sont les points qui ont y pour prédécesseur, dans l'ordre d'orientation, le P-intervalle R(y) est formé de y suivi de $R(x_1)$, puis de $R(x_2), \ldots$ de $R(x_n)$; le premier élément de $R(x_i)$ est x_i, le dernier n'a pas d'autre R-majorant que lui-même, car chacun de ses R-majorants devrait appartenir à $R(x_i)$: c'est donc une feuille ; la première feuille f de R(y) est aussi la première feuille de $R(x_1)$; en répétant ce raisonnement, on voit qu'il existe une suite $z_0 = y,\ z_1 = x_1, \ldots z_p = f$, telle que, pour $1 \leqslant i \leqslant p$, z_i soit le premier, dans l'ordre d'orientation, des éléments qui couvrent z_{i-1}. On a aussi $x_1\ S\ x_2 \ldots S\ x_n$ (théorème 1.4) ; le S-intervalle R(y) est formé de $R(x_1)$, suivi de $R(x_2), \ldots R(x_n)$ et de y ; dans l'ordre S, le dernier élément de $R(x_i)$ est x_i et le premier est une feuille. Ces résultats étaient d'ailleurs presque évidents d'après la construction de la pile simple strictement attachée à une ramification orientée (2. 1).

Dans les paragraphes suivants, nous tirerons d'autres conséquences des théorèmes 2. 3. et 2. 4.

2. 4. Ordres d'entrée et de sortie d'une ramification et de ramifications qui lui sont simplement liées.

2. 4. 1. Sous-ramifications. Soit P_1 la restriction d'un ordre P d'entrée d'une ramification [E, R] à une sous-ramification $[E_1, R_1]$; P et R vérifient l'hypothèse du théorème 2. 3, donc P_1 et R_1 la vérifient aussi : P_1 est un ordre d'entrée de $[E_1, R_1]$. *Une orientation O d'une ramification [E, R] détermine donc une orientation O_1 de chacune de ses sous-ramifications $[E_1, R_1]$, de manière que l'ordre d'entrée de $[E_1, R_1]$ orientée par O_1 soit restriction de l'ordre d'entrée de [E, R] orientée par O.*

On fait le même raisonnement pour la restriction à E_1 d'un ordre de sortie de [E, R]. De plus d'après le théorème 1. 4, les restrictions de deux ordres d'entrée et de sortie de [E, R] associés sont pour $[E_1, R_1]$ des ordres d'entrée et de sortie associés.

Les traces (1. 6. 4) des piles simples attachées à [E, R] sont des piles simples attachées à $[E_1, R_1]$.$[E_1, R_1]$ peut avoir d'autres ordres d'entrée ou de sortie que les restrictions des ordres d'entrée ou de sortie de [E, R]. Autrement dit, une pile simple attachée à $[E_1, R_1]$ n'est pas toujours la trace d'une pile simple attachée à [E, R].

Exemple : $E = \{a,b,c,d\}$; $a R b$, $a R c$, $a \bar{R} d$; b, c, d, sont des feuilles de [E, R], $E_1 = \{b,c,d\}$. b d c est un ordre d'entrée de $[E_1, R_1]$, mais ce n'est la restriction d'aucun ordre d'entrée de [E, R].

Supposons cependant que toute famille de $[E_1, R_1]$ soit contenue dans une famille de [E, R]. L'orientation O_1 de $[E_1, R_1]$ déterminée par une orientation O de [E, R] est alors la restriction de O à E_1, et $[E_1, R_1]$ ne possède pas d autre orientation, donc pas d'autre ordre d'entrée (ou de sortie) que les restrictions à E_1 des ordres d'entrée (ou de sortie) de [E, R].

Soit par exemple une ramification orientée (E, γ), réunion horizontale (1. 4. 2) de (E_1, γ_1) et (E_2, γ_2), P, P_1, P_2 les ordres d'entrée respectifs de ces trois ramifications orientées. L'orientation O_1 de (E_1, γ_1) est la restriction à E_1 de l'orientation O de (E, γ), donc P_1 est la restriction de P à E_1 ; de même P_2 est la restriction de P à E_2. Comme dans l'ordre P, toute racine de (E_1, γ_1) précède toute racine de (E_2, γ_2), d'après le théorème 2. 3, tout élément de E_1 précède tout élément de E_2. On fait le même raisonnement pour les ordres de sortie.

2. 4. 2. Ramifications isomorphes. Si φ est un isomorphisme d'une ramification [E, R] sur une ramification [E', R'], les ordres d'entrée de [E', R'] sont les ordres P' tels que

$$\varphi(x)\, P'\, \varphi(y) \iff x\, P\, y$$

où P est un ordre d'entrée quelconque de [E, R]. Les ordres de sortie S' de [E', R'] sont définis par

$$\varphi(x)\, S'\, \varphi(y) \iff x\, S\, y$$

où S est un ordre de sortie de [E, R]. Si P et S sont associés, il en est de même pour P' et S'. D'ailleurs φ transcrit (1. 6. 5) les piles attachées à [E, R] en les piles attachées à [E', R'].

2. 4. 3. Transformation par quasi-isomorphisme. Soit φ un quasi-isomorphisme d'une ramification [E, R] sur une ramification [E', R']. [E', R'] est isomorphe à une sous-ramification $[E_1, R_1]$ de [E, R] : à tout $\varphi(x)$ de E', cet isomorphisme associe x_1 de E_1 tel que $\varphi(x_1) = \varphi(x)$. Alors, si P est un ordre d'entrée de [E, R], P' défini par

$$\varphi(x)\, P'\, \varphi(y) \iff x_1\, P\, y_1$$

est un ordre d'entrée de [E', R']. D'autre part, lorsqu'un élément u de E a un successeur v pour la relation R, v est aussi successeur de u pour l'ordre d'entrée P d'après le paragraphe 2. 3. 3. Par suite, les classes de la relation d'équivalence définie en 1. 5. 2 sont des P-intervalles ; x_1 et x, y_1 et y sont équivalents. Donc

$$x_1\, P\, y_1 \iff x\, P\, y \quad \text{ou} \quad x_1 = y_1 \iff x\, P\, y \quad \text{ou} \quad \varphi(x) = \varphi(y).$$

Il en résulte

$$(3) \qquad \varphi(x)\; P'\; \varphi(y) \iff x\, P\, y \quad \text{ou} \quad \varphi(x) = \varphi(y).$$

Réciproquement, soit P' un ordre d'entrée de [E', R']. Montrons qu'*il existe un, et un seul, ordre P d'entrée de [E, R], qui vérifie* (3). Nécessairement, pour un tel ordre P :

- lorsque $\varphi(x) \neq \varphi(y)$, $x P y$ équivaut à $\varphi(x)\, P'\, \varphi(y)$
- lorsque $\varphi(x) = \varphi(y)$, $x P y$ équivaut à $x R y$, d'après la forme des classes de la relation d'équivalence définie en 1. 5. 2.

Il existe donc au plus un ordre P répondant à la question, défini par :

$$x\, P\, y \iff [\varphi(x) \neq \varphi(y) \text{ et } \varphi(x)\, P'\, \varphi(y)] \text{ ou } [\varphi(x) = \varphi(y) \text{ et } x\, R\, y].$$

Cette relation P est réflexive et antisymétrique ; elle est transitive : en effet, supposons $x P y$ et $y P z$:

a) $\varphi(x) = \varphi(y) = \varphi(z)$: $x R y$ et $y R z \Longrightarrow x R z$;

b) $\varphi(x) \neq \varphi(y) \neq \varphi(z)$: $\varphi(x)\, P'\varphi(y)$ et $\varphi(y)\, P'\varphi(z) \Longrightarrow \varphi(x)\, P'\ \varphi(z)$;
c) $\varphi(x) = \varphi(y) \neq \varphi(z)$: $\varphi(x) \neq \varphi(z)$ et $\varphi(x)\, P'\ \varphi(z)$;
d) $\varphi(x) \neq \varphi(y) = \varphi(z)$: $\varphi(x) \neq \varphi(z)$ et $\varphi(x)\, P'\ \varphi(z)$.

Deux éléments sont toujours comparables car on a vu (1. 5. 2) que les classes de la relation d'équivalence associée au quasi-isomorphisme étaient des R-chaînes : P est un ordre total. On voit aisément que P vérifie (3).

D'autre part,

$$x\, R\, y \Longrightarrow \varphi(x)\, R'\ \varphi(y) \Longrightarrow \varphi(x)\, P'\ \varphi(y).$$

Par suite, d'après la définition de P, x R y entraîne x P y. Enfin

$$\begin{aligned} xRy \text{ et } xPzPy &\Longrightarrow \varphi(x)\, R'\varphi(y) \text{ et } \varphi(x)\, P'\ \varphi(z)\, P'\varphi(y) \\ &\Longrightarrow \varphi(x)\, R'\ \varphi(z) \\ &\Longrightarrow x\, R\, z \text{ ou } (z\, R\, x \text{ et } \varphi(x) = \varphi(z)). \end{aligned}$$

Mais zRx et xPz sont incompatibles si $x \neq z$, donc (xRy et xPzPy) entraîne xRz. P est un ordre d'entrée de [E, R] (théorème 2. 3).

La correspondance entre les ordres d'entrée P et P' montre que *toute orientation* O *de* [E, R] *détermine une orientation* O' *de* [E', R'], *et inversement*; si xOy, $\varphi(x)$ et $\varphi(y)$ appartiennent à la même famille, et $\varphi(x)\, O'\ \varphi(y)$.

Puisque les ordres de sortie d'une ramification sont les inverses de ses ordres d'entrée, si S est un ordre de sortie de [E, R], l'ordre S' défini par

(4) $\qquad \varphi(x)\, S'\ \varphi(y) \Longleftrightarrow x\, S\, y$ ou $\varphi(x) = \varphi(y)$

est un ordre de sortie de [E', R']. Si les ordres P et S sont associés, P' et S' sont associés. Réciproquement, si S' est un ordre de sortie de [E', R'] la relation S définie dans E par

$$x\, S\, y \Longleftrightarrow [\varphi(x) \neq \varphi(y) \text{ et } \varphi(x)\, S'\varphi(y)] \text{ ou } [\varphi(x) = \varphi(y) \text{ et } y\, R\, x]$$

est un ordre de sortie de [E, R] qui vérifie (4), et il n'en existe pas d'autre.

Tout ceci s'applique en particulier lorsque [E', R'] est ramification réduite (ou sous-réduite) de [E, R]. Si [E, R] est orientée, on en déduit une orientation de [E', R'] et on définit ainsi les *ramifications orientées réduites* (ou *sous-réduites*) d'une ramification orientée donnée : ces ramifications orientées sont toutes isomorphes.

2. 5. Ramification d'intervalles.

2. 5. 1. Soit un ensemble G, q un ordre total strict de G et E un ensemble fini d'intervalles de G tel que :

$$(\forall\, \alpha \in E)\ (\forall\, \beta \in E)\ (\alpha \subseteq \beta \text{ ou } \beta \subset \alpha \text{ ou } \alpha \cap \beta = \phi).$$

[E, $\supseteq$] est une ramification car ($\alpha \supseteq \gamma$ et $\beta \supseteq \gamma$) entraîne $\alpha \cap \beta \neq \phi$, d'où $\alpha \supseteq \beta$ ou $\beta \supseteq \alpha$. On peut la compléter en une arborescence en adjoignant G à l'ensemble E. Lorsque α et β appartiennent à E, on dira que α *précède* β si les intervalles α et β sont disjoints et si un élément de α précède, dans l'ordre q, un élément de β (donc tout élément de α précède tout élément de β). Définissons la relation P :

$$\alpha\ P\ \beta \Longleftrightarrow \alpha \supseteq \beta \text{ ou } \alpha \text{ précède } \beta.$$

P est une relation réflexive et antisymétrique ; deux éléments sont toujours comparables. Elle est aussi transitive, car :

- $\alpha \supseteq \beta$ et $\beta \supseteq \gamma \Longrightarrow \alpha \supseteq \gamma$;
- α précède β et β précède $\gamma \Longrightarrow \alpha$ précède γ ;
- α précède β et $\beta \supseteq \gamma \Longrightarrow \alpha$ précède γ ;
- si $\alpha \supseteq \beta$ et β précède γ, $\gamma \supseteq \alpha$, qui entraînerait $\gamma \supseteq \beta$, est impossible ; et pour $\alpha \cap \beta = \phi$, α précède γ.

P est un ordre d'entrée, d'après le théorème 2. 3 ; en effet, $\alpha \supseteq \beta$ entraîne $\alpha\ P\ \beta$; d'autre part supposons $\alpha \supset \beta$ et $\alpha\ P\ \gamma\ P\ \beta$: si $\gamma \supseteq \beta$, alors $\alpha \cap \gamma \neq \phi$ et donc α ne précède pas γ ; si γ précède β, α ne précède pas γ. D'après la définition de P, $\alpha \supseteq \gamma$.

Si on oriente la ramification [E, $\supseteq$] en ordonnant les racines et les intervalles couvrant un même élément, qui sont disjoints, par la relation "précède", P est l'ordre d'entrée strictement attaché à la ramification orientée.

L'ordre de sortie S associé à P est obtenu grâce au théorème 1. 4 :

$$\begin{aligned} \alpha\ S\ \beta &\Longleftrightarrow \alpha \subseteq \beta \text{ ou } (\alpha\ P\ \beta \text{ et } \alpha \not\supseteq \beta) \\ \alpha\ S\ \beta &\Longleftrightarrow \alpha \subseteq \beta \text{ ou } (\alpha \text{ précède } \beta). \end{aligned}$$

Si les intervalles α de l'ensemble E sont repérés par leurs extrémités :
$\alpha = [\delta(\alpha), \omega(\alpha)]$ ou $\alpha =]\,\delta(\alpha), \omega(\alpha)\,[$, on voit aisément que

$$\begin{aligned} \alpha\ P\ \beta &\Longleftrightarrow \delta(\alpha)\ q\ \delta(\beta) \text{ ou } (\delta(\alpha) = \delta(\beta) \text{ et } \alpha \supseteq \beta) \\ \alpha\ S\ \beta &\Longleftrightarrow \omega(\alpha)\ q\ \omega(\beta) \text{ ou } (\omega(\alpha) = \omega(\beta) \text{ et } \alpha \subseteq \beta). \end{aligned}$$

Si on parcourt G dans l'ordre q, P est en gros l'ordre où on entre dans les intervalles formant E, et S l'ordre où on en sort.

Exemple : G étant l'ensemble des signes qui composent un programme Algol, ordonné par l'ordre q d'écriture, l'ensemble E des blocs de ce programme possède la structure qui vient d'être décrite. Si donc on lit le programme dans l'ordre q, si au début de chaque bloc α on fait entrer dans une pile une information $\mathcal{J}(\alpha)$ (cf. 1. 6. 5) relative à α (déclaration, indications d'occupation de la mémoire) et qu'à la fin on fait sortir $\mathcal{J}(\alpha)$ de la pile, lorsque $\mathcal{J}(\alpha)$ est au sommet, la pile est constituée des $\mathcal{J}(\beta)$ relatifs aux blocs β contenant α, ordonnés par l'inclusion. La relation d'ordre P est la relation "est antérieur à" de [45].

2. 5. 2. Intervalles consécutifs. Deux intervalles α et β sont *consécutifs* si, pour l'ordre q, α précède β et la réunion de α et β est encore un intervalle. Supposons $[E, \supseteq]$ complétée en une arborescence et soit γ le plus petit intervalle de E qui contient α et β. Il existe dans E deux chaînes convexes pour l'inclusion :

$$\gamma \supseteq \alpha_1 \supseteq \ldots \supseteq \alpha_n = \alpha \quad \text{et} \quad \gamma \supseteq \beta_1 \supseteq \ldots \supseteq \beta_p = \beta \ ;$$

dans la ramification $[E, \supseteq]$ orientée comme il a été dit, tout α_i, pour $1 < i \leqslant n$, est le dernier élément de sa famille, tout β_j, pour $1 < j \leqslant p$, le premier de sa famille ; α_1 est le prédécesseur de β_1 dans leur famille commune. Réciproquement, ces propriétés entraînent que les intervalles α et β sont consécutifs, dans le cas où tout intervalle appartenant à E est réunion des intervalles de E qui le couvrent (pour $\supseteq$) : car alors α_1 et β_1 sont consécutifs, α_{i-1} et α_i ont même dernier élément, β_{i-1} et β_i ont même premier élément.

2. 6. Restriction d'un ordre d'entrée ou de sortie à l'ensemble des feuilles d'une ramification.

2. 6. 1. Soit une ramification $[E, R]$, P un de ses ordres d'entrée, q la restriction de P à l'ensemble F de ses feuilles. D'après le théorème 1. 4, l'ordre de sortie S associé à P possède la même restriction à F, car deux feuilles ne sont pas comparables par la relation R.

Pour tout élément x de E, $R(x)$ est un P-intervalle, donc $r(x) = R(x) \cap F$ est un q-intervalle. Réciproquement, soit q un ordre total de F pour lequel les $r(x)$ soient des intervalles ; montrons que q est la restriction à F d'un ordre d'entrée. Les $r(x)$ sont les points de la ramification réduite $[E', \supseteq]$ de $[E, R]$ (1. 5. 3) ; l'ensemble F' des feuilles de cette ramification est composé d'intervalles formés seulement d'une feuille de $[E, R]$. Si $r(x)$ et $r(y)$ ne sont pas disjoints, x et y ont un minorant commun, et donc x R y ou y R x (théorème 1. 1) ; pour x R y par exemple, $r(x)$ contient $r(y)$. L'ensemble d'intervalles E' vérifie donc l'hypothèse faite au paragraphe 2. 5. 1, ce qui nous permet de considérer l'ordre d'entrée P' de la ramification réduite :

$$r(x) \ P' \ r(y) \iff r(x) \supseteq r(y) \quad \text{ou} \quad (r(x) \text{ précède } r(y)).$$

Si x et y appartiennent à F

$$r(x) \ P' \ r(y) \iff x \, q \, y.$$

D'après le paragraphe 2. 4. 3, il existe un ordre d'entrée P de $[E, R]$ pour lequel :

$$r(x) \ P' \ r(y) \iff x \, P \, y \quad \text{ou} \quad r(x) = r(y).$$

Lorsque x et y appartiennent à F, $r(x) \neq r(y)$ si $x \neq y$, et donc

$$(\forall x \in F) \quad (\forall y \in F) \quad (x \, P \, y \iff x \, q \, y).$$

D'autre part, soient P et P_1 deux ordres d'entrée de $[E, R]$ distincts. Il existe deux éléments distincts x et y de E tels que $x P y$ et $y P_1 x$: x et y ne sont pas comparables par la relation R. Soit f et g deux feuilles telles que x R f et y R g. Comme $R(x)$ et $R(y)$ sont des intervalles disjoints pour P et pour P_1, il en résulterait $f P g$ et $g P_1 f$, ce qui est impossible.

Théorème 2. 5. *Pour qu'un ordre total q, dans l'ensemble F des feuilles d'une ramification $[E, R]$, soit la restriction à F d'un ordre d'entrée P de $[E, R]$, il faut et il suffit que, pour tout x de E, $R(x) \cap F$ soit un q-intervalle. L'ordre P est alors unique et q est aussi la restriction à F de l'ordre de sortie associé à P.*

Un ordre q qui satisfait aux hypothèses de ce théorème détermine un ordre d'entrée de la ramification $[E, R]$, donc aussi une orientation de $[E, R]$; et inversement toute orientation de $[E, R]$ détermine un ordre d'entrée (2.1.1), donc sa restriction q à l'ensemble des feuilles. Nous appellerons *suite de feuilles* de la ramification orientée l'ensemble de ses feuilles ordonné par cet ordre q.

2. 6. 2. Réunion verticale de deux ramifications orientées. Au paragraphe 1.4.2 on a défini la réunion horizontale (E, γ) de deux ramifications orientées (E_1, γ_1) et (E_2, γ_2) disjointes. D'après l'étude faite au paragraphe 2. 4. 1, la suite de feuilles de (E, γ) est obtenue par concaténation des suites de feuilles de (E_1, γ_1) et (E_2, γ_2).

Soient ici (E_1, γ_1) et (E_2, γ_2) deux ramifications orientées telles que $E_1 \cap E_2$ soit à la fois l'ensemble des feuilles de (E_1, γ_1) et la famille des racines de (E_2, γ_2), et que l'ordre q des feuilles de (E_1, γ_1) soit aussi l'ordre d'orientation des racines de (E_2, γ_2). Désignons par E la réunion de E_1 et E_2 et définissons une application γ de $E \cup \{\Delta\}$ dans le monoïde libre E^*, par :

$$\begin{array}{lll} \gamma(x) = \gamma_1(x) & \text{si} & x \in E_1 \setminus E_2, \\ \gamma(x) = \gamma_2(x) & \text{si} & x \in E_2, \\ \gamma(\Delta) = \gamma_1(\Delta). & & \end{array}$$

Désignons par Γ la relation dans E pour laquelle $\Gamma(x)$ est l'ensemble des éléments qui possèdent une occurrence dans $\gamma(x)$, et par $(E_1, \Gamma_1) = [E_1, R_1]$, $(E_2, \Gamma_2) = [E_2, R_2]$ les ramifications obtenues à partir de (E_1, γ_1) et (E_2, γ_2) en faisant abstraction de l'orientation. Lorsque $a \, \Gamma \, b$, si $b \in E_1$, alors $a \in E_1$ et $a \, \Gamma_1 \, b$, si $a \in E_2$, alors $b \in E_2$ et $a \, \Gamma_2 \, b$. On en déduit que (E, Γ) est un graphe $[E, R]$ sans circuit. Comme tout élément de E a au plus une occurrence dans les mots $\gamma(x)$, (E, γ) est une ramification orientée (1. 4. 1), que nous nommerons *réunion verticale* de (E_1, γ_1) et (E_2, γ_2).

De l'étude précédente résulte aussi que R_1 et R_2, sont les restrictions de R à E_1 et E_2, que $[E_1, R_1]$ et $[E_2, R_2]$ sont des sous-ramifications de [E, R], que E_1 et E_2 sont des parties convexes de E pour R. Toute partie convexe de E_i pour R_i($i = 1$ ou 2) est donc partie convexe de E pour R. Par suite, *toute sous-arborescence complète* (1. 4. 1) *de* $[E_i, R_i]$ *est une sous-arborescence complète de* [E, R].

Comme pour la réunion horizontale au paragraphe 2. 4. 1, on démontre que les ordres d'entrée (resp. de sortie) de (E_1, γ_1) et de (E_2, γ_2) sont les restrictions à E_1 et E_2 de l'ordre d'entrée (resp. de sortie) de (E, γ). En particulier, l'ordre des racines de (E, γ) est celui des racines de (E_1, γ_1), la suite des feuilles de (E, γ) est celle de (E_2, γ_2). D'après le théorème 2. 3 (resp. 2. 4), la suite des éléments de E dans l'ordre d'entrée (resp. de sortie) de (E, γ) est obtenue à partir de la suite des éléments de E_1 dans l'ordre d'entrée (resp. de sortie) de (E_1, γ_1) en insérant après (resp. avant) chaque feuille la suite de ses majorants stricts dans la ramification (E_2, γ_2), dans leur ordre d'entrée (resp. de sortie).

L'opération de réunion verticale est associative, mais non commutative. Le produit d'arborescences de [48] est une réunion verticale d'une arborescence et d'une réunion horizontale d'un certain nombre d'arborescences toutes isomorphes.

2. 6. 3. Points consécutifs dans une ramification orientée. Deux points x et y d'une ramification $(E, \Gamma) = [E, R]$ orientée sont dits *consécutifs* si les intervalles r(x) et r(y) sont consécutifs dans la suite de feuilles de la ramification.

Pour qu'il en soit ainsi, il faut et suffit que, dans la suite des éléments de E ordonnés par l'ordre d'entrée (resp. de sortie), l'intervalle R(x) précède l'intervalle R(y) et qu'aucune feuille ne soit comprise entre eux.

Soit alors z le plus grand minorant commun à x et y, dans la ramification complétée en une arborescence. Il existe deux suites de points telles que $z \Gamma x_1 \Gamma \dots \Gamma x_m = x$ et $z \Gamma y_1 \dots \Gamma y_p = y$. D'après les propriétés des intervalles consécutifs (2. 5. 2) et celles de la transformation d'une ramification orientée par quasi-isomorphisme (1. 5. 3 et 2. 4. 3), y_1 suit immédiatement x_1 dans la famille des points qui ont z pour prédécesseur, pour $1 < i \leqslant m$, x_i est le dernier élément de sa famille, pour $1 < j \leqslant p$, y_j est le premier élément de sa famille ; et réciproquement ces propriétés entraînent que x et y sont consécutifs.

Il en résulte que si $[E_1, R_1]$ est une sous-arborescence complète de [E, R] (1. 4. 1), deux points consécutifs de $[E_1, R_1]$ sont aussi consécutifs dans [E, R].

2. 7. Application du théorème 2. 4 à des problèmes de minimisation.

2. 7. 1. Soit une ramification [E, R]. *Nous dirons qu'un ordre total* N *dans* E *vérifie la condition* (OR) *si* vRw *entraîne* wNv. A tout u de E associons alors l'ensemble $A_N(u)$ des $v \in E$ tels que vNu et pd(v) $\bar{N}$u. (pd(v) désigne le prédécesseur de v pour la relation R) ; soit $n_N(u)$ le cardinal de $A_N(u)$.

Les ordres de sortie de la ramification vérifient la condition (OR). Montrons que *si* N *satisfait à* (OR), *mais n'est pas un ordre de sortie, il existe un autre ordre total* N' *satisfaisant à* (OR) *pour lequel :*

$$(5) \qquad (\forall u \in E)\ (n_{N'}(u) \leqslant n_N(u)) \quad \text{et} \quad (\exists x \in E)\ (n_{N'}(x) < n_N(x)).$$

a) D'après le théorème 2. 4, si N n'est pas un ordre de sortie, il existe un élément x de E pour lequel R(x) n'est pas un N-intervalle ; nous choisirons pour x le premier, dans l'ordre N, de ces éléments. Si y est le premier, dans l'ordre N, des R-majorants de x, il existe un entier positif p et p éléments $z_1, z_2, \dots, z_p$ de E tels que

$$y N z_1 N z_2 \dots N z_p N x \quad \text{et} \quad x \bar{R} z_j \ (j = 1, 2, \dots, p).$$

Alors i) z_j R u et x R u sont incompatibles car $x \bar{R} z_j$, $z_j \bar{R} x$ d'après l'hypothèse (OR), et [E, R] est une ramification. En particulier $z_j \bar{R} y$.

ii) z_j R u, u N y et $u \neq y$ sont aussi incompatibles, car $z_j \bar{R} y$: z_j serait un élément précédant x dans l'ordre N, dont l'ensemble des R-majorants n'est pas un intervalle .

Finalement, si z_j R u, u est un z_i.

b) Construisons une nouvelle permutation de E à partir de la permutation dans l'ordre N, en plaçant des z_j immédiatement après x ; d'où un ordre N' ; plus précisément :

si v et w ne sont pas des z_j, $\quad w N' v \iff w N v$;

si $v \bar{N} x$, $z_p N' v$; $x N' z_1 N' z_2 \dots N' z_p$.

Montrons que N' vérifie la condition (OR) :

i) si v et w ne sont pas des z_i, ou sont tous deux des z_i, wNv entraîne wN'v.

ii) si $v = z_j$ et si w n'est pas un z_i, vRw est impossible d'après a.

iii) supposons que $w = z_j$, que v n'est pas un z_i et que vRw : vRz_j entraîne z_jNv, d'où yNv ; de vNx résulterait alors xRv, d'où xRw, qui est impossible ; donc $v\bar{N}x$ et par suite z_jN'v.

c) Montrons que N' vérifie la condition (5) :

i) si $u \notin [_N y, x]$, vNu équivaut à vN'u ; donc $A_N(u) = A_{N'}(u)$.

ii) si u appartient à $[_N y, x]$, mais n'est pas un z_i, uN'x ; vN'u entraîne donc que v n'est pas un z_i ; d'après a, il en est alors de même pour pd(v).

D'où $\qquad$ vN'u et pd(v) $\bar{N}$'u $\implies$ vNu et pd(v) $\bar{N}$u :

$$A_{N'}(u) \subseteq A_N(u).$$

De plus, z_p appartient à $A_N(x)$, mais pas à $A_{N'}(x)$:

$$A_{N'}(x) \subset A_N(x).$$

iii) si $u = z_i$, $pd(v)\,\bar{N}'u$ entraîne $pd(v)\,\bar{N}u$. D'autre part,

$$pd(v)\,\bar{N}'u \Longrightarrow pd(v)\,\bar{N}'x \Longrightarrow x\bar{R}\,pd(v) \Longrightarrow x\bar{R}v \quad \text{ou} \quad x = v.$$

Lorsque $x\bar{R}v$, $vN'u$ entraîne soit vNy d'où vNu, soit $v = z_j (j \leqslant i)$, d'où encore vNu. Mais x appartient à $A_{N'}(u)$, alors qu'il n'appartient pas à $A_N(u)$. Cependant, soit y_1 l'élément couvrant x, tel que $y_1 R y$; $y_1 \bar{R} z_i$. A cause de la définition de x, $R(y_1)$ est un intervalle et donc $u = z_i$ n'appartient pas à l'intervalle $[_N y, y_1]$. Par suite $y_1 NuNx$: y_1 appartient à $A_N(u)$ et il n'appartient pas à $A_{N'}(u)$.

$$A_{N'}(u) \subset A_N(u) \cup \{x\}.$$

Il résulte de l'étude faite que N' répond à la question.

Si N' n'est pas un ordre de sortie, on peut recommencer. Comme l'un des $n_N(u)$ diminue strictement à chaque fois, on parviendra à un ordre de sortie au bout d'un nombre fini d'étapes. Donc si un ordre N satisfaisant à la condition (OR) n'est pas un ordre de sortie, il existe un ordre de sortie S tel que :

$$(\forall u \in E)\ (n_S(u) \leqslant n_N(u)) \quad \text{et} \quad (\exists x \in E)\ (n_S(x) < n_N(x)).$$

Par conséquent, le minimum, pour tous les ordres N satisfaisant à (OR), de fonctions telles que :

$$mx(N) = \max_{u \in E} n_N(u)$$

ou $\qquad k(N) = \sum_{u \in E} \lambda(u)\, n_N(u)$ où les $\lambda(u)$ sont positifs ou nuls, est réalisé par un ordre de sortie (il peut être aussi réalisé par d'autres ordres N).

2. 7. 2. Etudions ce cas où N est un ordre de sortie ; les points de toute composante connexe de la ramification forment un N-intervalle ; pour tout u, $A_N(u)$ est constitué de points de la composante connexe de u, car si v appartient à $A_N(u)$, $pd(v)Ru$ (théorème 2. 4) ; on peut donc se borner à étudier chaque composante, c'est-à-dire qu'on est ramené au cas où la ramification est une arborescence. Alors sa racine a est le plus grand élément de E pour l'ordre N ; convenons que $A_N(a) = \{a\}$; les autres éléments de E forment une ramification à p composantes connexes $E_0, E_1, \ldots, E_{p-1}$: si $i < j$, E_i précède E_j dans l'ordre N (et aussi dans l'ordre d'entrée associé) ; i est le rang de la racine de E_i dans sa famille, les rangs étant comptés à partir de l'origine 0 (paragraphe 1. 4. 5). Soit N_i la restriction de N à E_i. Si u appartient à E_i, $A_N(u)$ est la réunion de $A_{N_i}(u)$ et des derniers éléments (racines) de $E_0, E_1, \ldots, E_{i-1}$; il en est ainsi en particulier si u est la racine v_i de E_i, avec la convention

$A_{N_i}(v_i) = \{v_i\}$. Donc

$$n_N(u) = n_{N_i}(u) + i. \tag{6}$$

Soit $(r_0, r_1, \ldots, r_q)$ l'index de u (1. 4. 5). D'après (6) et avec des notations évidentes :

$$n_N(u) = n_{N_{r_1}}(u) + r_1$$

$$n_{N_{r_1}}(u) = n_{N_{r_1 r_2}}(u) + r_2$$

$$\ldots$$

$$n_{N_{r_1 \ldots r_{q-1}}} = n_{N_{r_1 \ldots r_q}} + r_q$$

$$n_{N_{r_1 \ldots r_q}} = 1$$

Donc $\qquad n_N(u) = r_1 + r_2 + \ldots + r_q + 1.$

Le maximum $mx(N)$ de $n_N(u)$ est réalisé par une feuille de la ramification.

Exemple : si N est l'ordre de sortie S donné dans l'exemple du paragraphe 2. 1. 3, $n_N(d) = 1 + 1 + 1 = 3$, $n_N(f) = 1 + 1 + 0 + 1 = 3$. Effectivement $A_S(d) = \{a, c, d\}$, $A_S(f) = \{a, c, f\}$.

Si, sans modifier l'ordre relatif des E_i, on réordonne chacun d'eux de manière à réaliser le minimum de

$$mx(N_i) = \max_{u \in E_i} n_{N_i}(u) \quad \text{ou de} \quad k(N_i) = \sum_{u \in E_i} \lambda(u)\, n_{N_i}(u)$$

$mx(N)$ ou $k(N)$ décroît (au sens large). Le minimum de $mx(N)$ ou $k(N)$ sera donc réalisé pour un ordre de sortie N tel que chaque $mx(N_i)$ ou $k(N_i)$ soit minimum. De plus, d'après (6)

$$mx(N) = \max_i\, [mx(N_i) + i]$$

$$k(N) = \sum_{i=0}^{p-1} [k(N_i) + \lambda_i\, i] \quad \text{avec} \quad \lambda_i = \sum_{u \in E_i} \lambda(u).$$

Le minimum de $mx(N)$ sera donc réalisé si la suite de ces $mx(N_i)$ est décroissante ; et celui de $k(N)$ si la suite des λ_i est décroissante. Nous donnerons au paragraphe 2. 10. 6 un algorithme qui permet de passer d'un ordre de sortie quelconque à un ordre de sortie qui réalise le minimum de $mx(N)$ ou de $k(N)$. [37] donne un tel algorithme pour $mx(N)$ à propos de la transformation d'une formule écrite en notation polonaise.

2. 7. 3. Si la ramification $[E, R] = (E, \Gamma)$ représente l'organigramme d'une tâche à accomplir, qui nécessite des résultats intermédiaires ($x \Gamma y$ lorsque y entre dans la construction de x), $n_N(u)$ est le nombre d'éléments qui se trouvent stockés lorsque l'on vient d'obtenir u. On cherche l'ordre N de construction des éléments qui minimise l'espace nécessaire au stockage. Cet espace est mesuré par $mx(N)$ si tous les éléments v de E demandent le même espace. Pour généraliser, il faudrait affecter chaque v d'un coefficient positif ou nul $\chi(v)$ et remplacer $n_N(u)$ et $mx(N)$ par

$$n'_N(u) = \sum_{v \in A_N(u)} \chi(v) \quad \text{et} \quad mx'(N) = \max_{u \in E} n'_N(u).$$

La formule (6) est alors remplacée par

$$n'_N(u) = n'_{N_i}(u) + \sum_{j=0}^{i-1} \chi(v_j)$$

(v_j est la racine de E_j). Le minimum de $mx'(N)$ parmi les ordres de sortie est encore réalisé lorsque la suite des $mx'(N_j)$ est décroissante. Mais le minimum de $mx'(N)$ parmi les ordres qui satisfont à la condition (OR) n'est plus nécessairement réalisé par un ordre de sortie.

Exemple (figure 2. 5) : l'ordre N_0 ci-dessous réalise les n'_N indiqués :

N_0	h	e	k	f	d	b	g	c	a
χ	5	1	5	1	3	3	3	3	1
n'_{N_0}	5	1	6	2	5	4	7	6	1

$mx'(N_0) = 7$

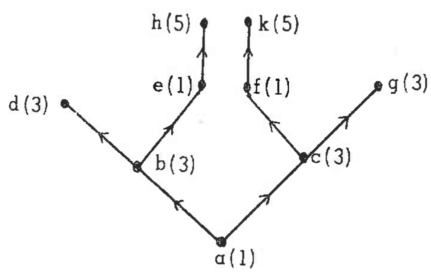


Figure 2. 5.

Le minimum de $mx'(N)$ parmi les ordres de sortie est réalisé par l'ordre S : h, e, d, b, l, f, g, c, a et $n'_S(l) = 8$.
On peut faire le même raisonnement avec $k'(N) = \sum_{u \in E} \lambda(u)\ n'_N(u)$.

L'ordre S réalise le minimum de $k'(N)$ parmi les ordres de sortie et $k'(S) = 39$. En revanche $k'(N_0) = 37$.

2. 8. Utilisation d'une pile simple pour permuter un ensemble.

2. 8. 1. Soit un ensemble fini E, P et S deux ordres totaux de E. Existe-t-il une pile simple sur E dont P et S soient l'ordre d'entrée et l'ordre de sortie ?

Si une telle pile existe, la relation d'ordre associée R vérifie (théorème 1. 3)

(7) $$e R e' \iff e P e' \text{ et } e' S e \,;$$

pour tout élément e de E, $R(e)$ est un P-intervalle (théorème 2. 3) et un S-intervalle (théorème 2. 4).

Réciproquement, supposons que pour la relation R définie par (7), tout $R(e)$ soit un P-intervalle, e est nécessairement son premier élément ; d'après le théorème 2. 3, il existe une pile simple dont P est l'ordre d'entrée, et R l'ordre associé. Son ordre de sortie S_1 vérifie aussi

$$e R e' \iff e P e' \text{ et } e' S_1 e.$$

D'après le théorème 1. 5, S et S_1 sont confondus. P et S sont l'ordre d'entrée et l'ordre de sortie d'une pile simple. De même, on montre qu'il suffit que tout $R(e)$ soit un S-intervalle pour qu'il existe une pile simple dont P soit l'ordre d'entrée et S l'ordre de sortie.

2. 8. 2. Toute pile simple sur E définit une permutation de E, par le passage de son ordre d'entrée à son ordre de sortie. Le nombre des permutations ainsi engendrées est égal au nombre des piles simples sur E ayant un mot d'entrée donné. Nous calculerons ce nombre par récurrence sur le nombre k d'éléments de E, c'est-à-dire la longueur du mot d'entrée $\lambda_k = {}' x_1 \ldots x_k {}'$. Une pile simple U_k ayant λ_k pour mot d'entrée est déterminée par ceux de ses états dont l'indice précède l'entrée $\varepsilon(x_k)$ de x_k, donc par une pile simple ayant λ_{k-1} pour mot d'entrée, et par le nombre d'éléments contenus dans l'état $u_{\varepsilon(x_k)}$; posons $u_{\varepsilon(x_q)} = v_q$: si v_{k-1} contient i éléments ($i \leqslant k-1$), v_k peut en contenir 1, 2, 3, ... ou $i+1$. Soit N_j^k le nombre de piles simples U_k où v_k contient j éléments :

$$N_j^k = \sum_{i=j-1}^{k-1} N_i^{k-1} \qquad \text{pour } 1 \leqslant j \leqslant k \text{ et } k \geqslant 2$$

$$N_0^k = 0 \text{ pour } k \geqslant 1 \;;\; N_1^1 = 1.$$

D'où une autre définition des N_j^k :

$$N_j^k = N_{j+1}^k + N_{j-1}^{k-1} \quad \text{pour } 1 \leqslant j < k$$
$$N_k^k = N_{k-1}^{k-1} \quad \text{pour } k \geqslant 1 \; ; N_0^0 = 1$$
$$N_0^k = 0 \quad \text{pour } k \geqslant 1.$$

Le nombre de piles simples ayant λ_k pour mot d'entrée est alors

$$\alpha_k = \sum_{j=1}^{k} N_j^k = N_2^{k+1} = N_1^{k+1}.$$

On peut calculer α_k en cherchant des fonctions génératrices (1) des N_j^k ; supposons que les séries entières $\sum_{k=j}^{+\infty} N_j^k t^k$ soient convergentes, et désignons leurs sommes par $\varphi_j(t)$. Il résulte des relations définissant les N_j^k que :

$$(8) \qquad \varphi_j(t) = \varphi_{j+1}(t) + t\,\varphi_{j-1}(t) \quad \text{pour } j > 0$$
$$\varphi_0(t) = 1.$$

De plus, lorsque t tend vers 0, $\varphi_j(t)$ est équivalent à t^j. Les suites solutions de l'équation (8) sont de la forme $\lambda r_1^j + \mu r_2^j$ où λ et μ sont indépendants de j, et r_1, r_2 sont les racines de l'équation $r^2 - r + t = 0$. La racine $\frac{1-\sqrt{1-4t}}{2}$ est équivalente à t quand t tend vers 0 ; l'autre racine tend alors vers 1. On est ainsi conduit à envisager la fonction

$$\varphi_j(t) = \left(\frac{1-\sqrt{1-4t}}{2}\right)^j.$$

Réciproquement, cette fonction est analytique dans un voisinage de 0. Elle vérifie l'équation (8), de sorte que

$$N_j^k = N_{j+1}^k + N_{j-1}^{k-1} \quad \text{pour } 1 \leqslant j < k$$
$$N_j^j = N_{j-1}^{j-1} \quad \text{pour } j \geqslant 1$$
$$\varphi_0(t) = 1 \text{ donc } N_0^0 = 1 \text{ et } N_0^k = 0 \quad \text{pour } k \geqslant 1.$$

Il est alors facile de calculer

$$\alpha_k = N_1^{k+1} = \frac{1.3.5\ldots(2k-1)}{(k+1)!}\,2^k$$

$$\alpha_k = \frac{(2k)!}{k!\,(k+1)!}$$

(1) cf. [39]. Je dois cette idée à M.M. DEPAIX.

Le nombre des piles simples sur un ensemble à k éléments est

$$k!\,\alpha_k = \frac{(2k)!}{(k+1)!}.$$

Par la formule de Stirling,

$$\alpha_k \sim \frac{4^k}{\sqrt{\pi}\,k\,\sqrt{k}} \qquad \text{lorsque } k \longrightarrow +\infty.$$

La probabilité $\frac{\alpha_k}{k!}$ pour qu'une permutation donnée de E soit engendrée tend vers 0 quand k tend vers l'infini.

On montre cependant qu'on peut engendrer toute permutation en itérant la transformation ; de sorte qu'on peut ainsi trier une suite, c'est-à-dire la réordonner dans un ordre imposé ([46], [50]). Il ne semble pas que cette méthode de tri se classe parmi les meilleures.

2. 9. Pile adjointe à une ramification orientée.

2. 9. 1. Justification de l'algorithme du paragraphe 2. 1. 4. Cet algorithme construit la pile simple U strictement attachée à une ramification orientée [E, R], complétée en une arborescence par l'introduction d'une racine Δ. La trace sur l'ensemble E de la pile U'' introduite au paragraphe 2. 1. 4 est une pile simple. Son ordre de sortie est l'ordre d'entrée P de la ramification orientée donnée. Son ordre d'entrée Q peut être défini par :

$$x\,Q\,y \iff [pd(x) \neq pd(y) \text{ et } pd(x)\;P\;pd(y)] \text{ ou } [pd(x) = pd(y) \text{ et } y\,P\,x].$$

Nous montrerons directement, comme application du paragraphe 2. 8. 1, *qu'il existe une pile simple dont l'ordre d'entrée est* Q *et l'ordre de sortie* P. Montrons d'abord que Q est un ordre total. D'après 1. 5. 3, l'application pd est un quasi-isomorphisme d'une ramification $[E, R_2]$ sur l'arborescence des noeuds de $[E \cup \{\Delta\}, R]$: R_2 a été définie au paragraphe 1. 5. 3 en utilisant l'orientation O de la ramification donnée ; comme P et O ont même restriction à toute famille de la ramification :

$$x\,R_2\,y \iff [pd(x) \neq pd(y) \text{ et } pd(x)\;R\;pd(y)] \text{ ou } [pd(x) = pd(y) \text{ et } y\,P\,x]$$

Alors $\quad pd(x) = pd(y) \text{ et } y\,P\,x \iff pd(x) = pd(y) \text{ et } x\,R_2\,y.$

Q est l'ordre d'entrée de la ramification $[E, R_2]$ déduit de l'ordre d'entrée P d'une ramification quasi-isomorphe comme il a été dit au paragraphe 2. 4. 3 ; Q est un ordre total.

Soit R' la relation définie par

$$x \, R' \, y \iff x \, Q \, y \text{ et } y \, P \, x$$

c'est-à-dire

$$x \, R' \, y \iff pd(x) \, P \, pd(y) \text{ et } y \, P \, x. \tag{9}$$

$R'(x)$ est un P-intervalle car ($x R' y$ et $y P z P x$) entraîne $pd(x) \, P \, pd(y) \, P \, y \, P \, z \, P \, x$, d'où $pd(x) R z$ (théorème 2. 3) et $pd(x) \neq z$, d'où $pd(x) R \, pd(z)$ et enfin $pd(x) P \, pd(z)$. Il existe une pile simple π dont l'ordre d'entrée est Q et l'ordre de sortie P.

Etudions π. Soit y un noeud de $[E, R]$, $x_1, x_2, \dots, x_n$ les éléments dont y est le R-prédécesseur, avec : $x_1 \, P x_2 \dots P x_n$.
$y \, P \, x_1$, $y \, Q \, x_n \, Q \dots Q \, x_1$, $y \, \bar{R}' \, x_n$, la sortie de y précède l'entrée de x_n, et $x_n, \dots, x_2, x_1$ forment un Q-intervalle, car il s'agit d'une classe de la relation d'équivalence définie par le quasi-isomorphisme pd (2. 4. 3) ; dans la suite, *nous dirons que deux éléments de* E, x *et* y, *sont équivalents, et nous noterons* $x \sim y$, *lorsqu'ils appartiennent à la même famille, c'est-à-dire que* $pd(x) = pd(y)$. Dans l'ordre de sortie P de la pile π, x_1 est le successeur de y (2. 4. 3). Entre la sortie de y et celle de x_1 ne se place donc aucune sortie, mais, nécessairement, dans l'ordre, les entrées de $x_n, x_{n-1}, \dots, x_1$, et entre l'entrée de x_n et celle de x_1 n'existe aucune autre entrée. D'autre part, entre l'entrée et la sortie de x_1, n'entre aucun élément v, sinon il sortirait avant x_1 (théorèmes 1. 7 et 1. 3). Enfin, si l'entrée d'un élément v se produisait entre la sortie de y et l'entrée de x_n, la sortie de y se placerait entre celle de $pd(v)$ et celle de v_1, premier élément (dans l'ordre P) équivalent à v. Le même raisonnement montre que la sortie d'une feuille z ne peut être immédiatement suivie de l'entrée d'un élément v. En résumé, *pour un noeud* y, *on trouve successivement la sortie de* y, *les entrées de* $x_n, x_{n-1}, \dots, x_1$ *et la sortie de* x_1 ; *pour une feuille* z, *sa sortie est suivie immédiatement d'une autre sortie ; les premiers éléments à entrer sont les racines, qui ont pour prédécesseur* Δ. π est ainsi parfaitement définie.

Pour construire la pile U strictement attachée à la ramification donnée, on fait entrer un élément dans U lors de sa sortie de π, puisque P est l'ordre de sortie de π. Reste à placer les sorties de la pile U ; pour cela on a seulement besoin de savoir si un élément sortant de π est le dernier point x_n d'une famille. Cette reconnaissance peut être facilement effectuée en faisant entrer un symbole $\ddagger$ avant x_n : π est alors remplacé par une pile U'' et l'algorithme du paragraphe 2. 1. 4 est ainsi de nouveau complètement justifié.

2. 9. 2. Définition, étude et construction de la pile adjointe. La pile π sur E, d'ordre d'entrée Q et d'ordre de sortie P est dite *adjointe* à la ramification orientée $[E, R]$.

Le mot m attaché à π a pour facteur gauche $\tilde{\gamma}(\Delta)$; pour tout x de E, l'intervalle $]\bar{x}, \bar{x}_1[$ de m qui a pour bornes l'occurrence du conjugué $\bar{x}$ de x et la première occurrence suivante d'un élément $\bar{x}_1$ n'appartenant pas à E, est $\tilde{\gamma}(x)$; il peut d'ailleurs être vide. Une pile est adjointe à une ramification orientée au plus.

Montrons que réciproquement, pour toute pile simple π sur un ensemble E, il existe une ramification orientée dont elle est la pile adjointe. Définissons un graphe (E, Γ) : pour tout x de E, $\Gamma(x)$ est l'ensemble des éléments de E dont l'entrée est comprise entre la sortie de x et la sortie suivante. $x \, \Gamma \, z$ entraîne que x entre avant z dans la pile : le graphe est sans circuit. Comme d'autre part, pour tout z, il existe au plus un x tel que $x \, \Gamma \, z$, (E, Γ) est une ramification (théorème 1. 1) ; ses racines sont les éléments de E qui entrent dans π avant la première sortie. Orientons cette ramification en utilisant pour ordre, dans chaque famille, l'inverse de la restriction de l'ordre d'entrée de π ; soit π_1 sa pile adjointe. Les états de π_1 dont l'indice précède la première sortie sont les mêmes que les états correspondants de π, puis se produit une sortie dans π_1 et dans π ; supposons que les états de π_1 et π jusqu'à la sortie de x (incluse) sont les mêmes ; alors se produisent les entrées dans π_1 et dans π des éléments de $\Gamma(x)$, dans le même ordre, puis une nouvelle sortie, du même élément y. Les piles π et π_1 sont confondues.

Théorème 2. 6. *Toute pile simple est adjointe à une ramification orientée unique.*

Notons que deux ramifications orientées isomorphes par φ ont des piles adjointes transcrites par φ.

La construction de π lorsque les mots $\gamma(x)$ sont donnés a été exposée au paragraphe 2. 9. 1. Elle est schématisée figure 2. 6 : u désigne l'état courant de la pile, e son sommet, v est une variable auxiliaire prenant ses valeurs dans l'ensemble E.

Lorsque la ramification est donnée par une matrice d'enchaînement correspondant à l'orientation inverse, l'algorithme de construction de π est schématisé sur la figure 2. 7 : les cases vides de la matrice d'enchaînement sont supposées contenir σ ; e, u, v ont le même sens que pour la figure 2. 6 ; e'' désigne le lien horizontal de e, v' le lien vertical de la valeur de v.

2. 9. 3. Hauteur de la pile adjointe. Transformons la définition (9) de R' : $pd(x) \, P \, pd(y)$ entraîne $pd(x) \, P \, y$ et $pd(x) \neq y$, soit $y \, \bar{P} \, pd(x)$. Réciproquement

$$y \, P \, x \text{ et } y \, \bar{P} \, pd(x) \implies pd(x) \, P y P x \text{ et } pd(x) \neq y$$

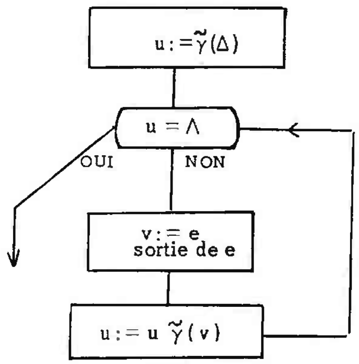


Figure 2. 6

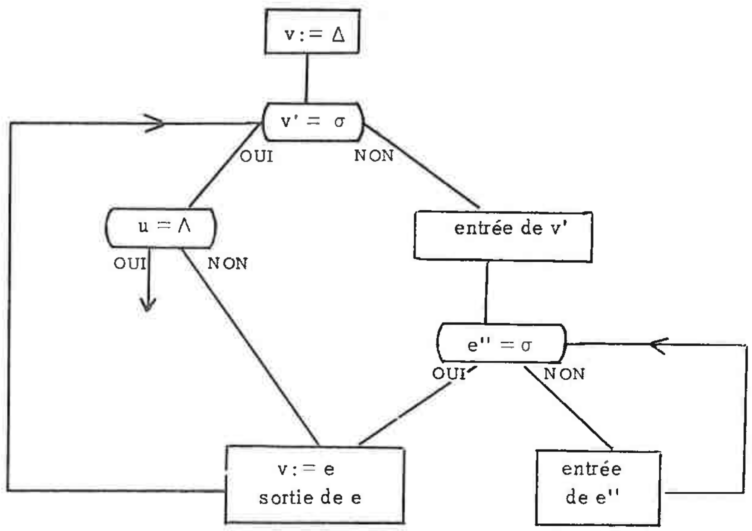


Figure 2. 7.

$$y P x \text{ et } y \bar{P} pd(x) \Longrightarrow y P x \text{ et } pd(x) R y \text{ et } pd(x) \neq y \text{ (théorème 2. 3)}$$
$$\Longrightarrow y P x \text{ et } pd(x) R pd(y)$$
$$\Longrightarrow y P x \text{ et } pd(x) P pd(y).$$

Finalement $\quad x R' y \Longleftrightarrow y P x \text{ et } y \bar{P} pd(x)$

ou encore $\quad x R' y \Longleftrightarrow x P^{-1} y \text{ et } pd(x) \bar{P}^{-1} y.$

P^{-1} est un ordre de sortie de la ramification $[E, R]$. Un état de la pile π qui a pour sommet e est formé des éléments de l'ensemble $A_{P^{-1}}(e)$ défini au paragraphe 2. 7. 1. Le nombre maximum d'éléments contenus dans un état de la pile adjointe est $mx(P^{-1})$.

Les propositions qui ont servi d'intermédiaire dans la démonstration de la réciproque ci-dessus sont aussi équivalentes à $x R' y$:

(10) $\quad x R' y \Longleftrightarrow y P x \text{ et } pd(x) R y \text{ et } y \neq pd(x)$

(11) $\quad x R' y \Longleftrightarrow y P x \text{ et } pd(x) R pd(y).$

2. 9. 4. Construction de ramifications de racines données dans un ensemble donné. Soit un ensemble D et une relation ρ entre D et le monoïde libre D^* déduit de D. On se propose de trouver une ramification orientée (E, γ) telle que E soit un sous-ensemble de D, que pour tout élément x de E, $x \rho \gamma(x)$ et que la suite $\nu(\Delta)$ des racines soit donnée. Il suffit d'utiliser l'organigramme décrit figure 2. 6, en choisissant les $\gamma(v)$ tels que $v \rho \gamma(v)$ de manière à construire une pile simple, c'est-à-dire de manière qu'un élément de D n'ait deux occurrences ni dans un même mot $\gamma(v)$, ni dans deux mots $\gamma(v)$ distincts : on obtient alors la pile adjointe à une ramification orientée qui répond à la question (paragraphe 2. 9. 2, réciproque).

Soit par exemple E' un ensemble de nombres entiers, et D l'ensemble des parties de E'. On appelle *arborescence binaire de recherche dans* E' une arborescence (E, Γ) telle que :

a) E est un sous-ensemble de D ;

b) la racine de l'arborescence est E' ;

c) pour tout $x \in E$, il existe $u \in x$ tel que si $y(x, u)$ est l'ensemble des éléments de x inférieurs à u et $z(x, u)$ l'ensemble des éléments de x supérieurs à u, $\Gamma(x)$ est formé de ceux des ensembles $y(x, u)$ et $z(x, u)$ qui ne sont pas vides (en particulier $x \in E$ est une feuille de l'arborescence si, et seulement si, x est un ensemble à un élément).

La construction d'une arborescence binaire de recherche entre dans le cadre qui a été tracé ; pour tout sous-ensemble x de E' contenant plus d'un élément, les mots en relation avec x sont ceux qui sont formés de $y(x, u)$ et

$z(x,u)$ (dans un ordre quelconque) lorsque u est un élément arbitraire de x ni maximum ni minimum, le mot '$y(x,u_1)$' si u_1 est le nombre maximum dans x, le mot '$z(x,u_2)$' si u_2 est le nombre minimum ; lorsque x a un seul élément, il n'est en relation qu'avec Λ. La condition qui assure qu'on obtient une pile simple est réalisée quel que soit le choix des mots $\gamma(v)$. En réalité, comme on l'a déjà signalé au paragraphe 2. 1. 4, pour chaque x prédécesseur de deux sous-ensembles y et z, il suffit de faire entrer l'un d'eux dans la pile, l'autre devant sortir dès son entrée ; si x n'est prédécesseur que d'un seul sous-ensemble, y par exemple, y n'entre pas dans la pile ; si x est réduit à un seul élément, après sa sortie intervient une nouvelle sortie. On montre, par récurrence sur le nombre n d'éléments de E', que si on oriente l'arborescence de manière que, lorsque $\gamma(x)$ contient deux sous-ensembles y et z, le cardinal du premier (qui entre seul dans la pile) soit au moins égal à celui du second, la hauteur de la pile ne dépasse pas $\log_2 n$.

L'arborescence binaire ainsi construite permet la recherche binaire ("binary search", [37]). Si à tout x on associe l'élément $u \in x$ de la définition (c), on transforme l'arborescence trouvée en une arborescence (E', Γ') dont l'ensemble des points est E' : c'est plutôt cette dernière arborescence qui sera utilisée. Lorsqu'on écrit un programme de recherche d'un nombre a, où à chaque test portant sur un couple (u,a), l'égalité fait aller à une sortie déterminée, $a > u$ fait effectuer un saut à un nouveau test et $a < u$ poursuivre la séquence, l'ordre des u pour lesquels sont écrites les instructions de test est l'ordre d'entrée de l'arborescence avec l'orientation pour laquelle, lorsque $\gamma(x)$ est de longueur 2, $\gamma(x) =$ '$y(x,u)\, z(x,u)$'.

Si l'ensemble E' est donné par ordre croissant, $y(x,u)$ et $z(x,u)$ sont des intervalles dont il suffit de placer les bornes dans la pile. On peut choisir pour u la médiane de x ou, si chaque élément de E' possède une probabilité, égaliser autant que possible la probabilité pour un élément d'être dans y ou dans z.

Lorsque E' est donné dans un ordre quelconque, on choisit u dans x de manière quelconque (par exemple le premier ou le dernier élément), et on permute x de manière à ce que $y(x,u)$ et $z(x,u)$ en deviennent des intervalles : ils sont ainsi faciles à repérer. L'arborescence (E', Γ') est déterminée par une matrice à trois colonnes, qui donne pour chaque nombre 0,1 ou 2 nombres dont il est le prédécesseur. [31] montre qu'il est alors facile, aussi bien de chercher un nombre de E' que d'enlever ou d'introduire de nouveaux points (nombres) dans l'arborescence. De sorte que l'arborescence peut aussi être construite par adjonctions successives ; on peut voir que cette construction demande les mêmes tests que celle qui vient d'être donnée, et elle évite d'utiliser une pile. En réalité, la construction de la pile par la méthode qui vient d'être indiquée permet surtout de trier l'ensemble E', car chaque élément u prend par permutation sa place dans l'ordre croissant de E' : c'est la méthode de tri proposée par T. Hibbard.

La méthode de tri dite "radix-exchange" [32] revient aussi à la construction d'une arborescence binaire : soit un ensemble E' de nombres entiers écrits en numération binaire, qu'on peut supposer de même longueur λ. Soit E l'ensemble de leurs facteurs gauches (1.3) et R la relation dans E : $x R y$ si, et seulement si, x est un facteur gauche de y ; [E, R] est une ramification. A tout x de E, on associe le sous-ensemble $\varphi(x)$ de E' formé des nombres ayant x pour facteur gauche : [$\varphi(E)$, $\supseteq$] est une ramification isomorphe à [E, R]. On construit la pile adjointe à cette ramification : pour chaque x de E n'ayant pas λ pour longueur, on est amené à partager l'ensemble $\varphi(x)$ en les deux sous-ensembles $y(x)$ et $z(x)$ des nombres qui ont x'0' et x'1' respectivement pour facteur gauche ; et on permute $\varphi(x)$ de manière que tout élément de $y(x)$ précède tout élément de $z(x)$. A la fin de la construction de la pile, l'ensemble E' donné est trié. Comme dans l'étude précédente, on peut se borner à faire entrer dans la pile celui des deux intervalles $y(x)$ ou $z(x)$ qui a le plus grand nombre d'éléments. Cette méthode diffère légèrement de celle proposée par [32], qui utilise cependant une partie de mémoire qui se rapproche d'une pile.

La détermination d'une arborescence graphe partiel d'un graphe donné entre aussi dans le cadre du problème posé au début du paragraphe. D'ailleurs la construction d'une arborescence binaire dans l'ensemble D des parties d'un ensemble E', que nous avons faite, est la construction d'une arborescence, graphe partiel d'un treillis.

2. 10. Pile conjointe à une ramification orientée.

2. 10. 1. Constructions de ramifications orientées dont on donne la suite de feuilles. Soit un ensemble D. Un premier problème consiste à chercher toutes les ramifications orientées dont l'ensemble des points est contenu dans D et qui ont une suite de feuilles donnée. On peut aussi, comme au paragraphe 2. 9. 4, se donner une relation ρ entre D et D* et chercher les ramifications orientées (E, γ) avant une suite de feuilles données et telles que, pour tout x de E, $x \,\rho\, \gamma(x)$.

Soit [E, R] une ramification orientée à déterminer. Nous utiliserons comme outil une pile π' où, si $x R y$, x entre après y : il suffit que l'ordre d'entrée de π' soit l'ordre de sortie de la ramification. Or la pile adjointe à une ramification orientée a pour ordre de sortie l'ordre d'entrée P de la ramification ; sa pile réciproque a pour ordre d'entrée P^{-1}, qui est l'ordre de sortie de la même ramification, mais munie de l'orientation inverse. Appelons *pile conjointe* à une ramification orientée la réciproque de la pile adjointe à la même ramification, munie de l'orientation inverse. D'après les résultats analogues sur la pile adjointe (2. 9. 2) :

Théorème 2. 7. *Toute pile simple est conjointe à une ramification orientée unique. Lorsqu'une ramification orientée est transformée par un isomorphisme* φ, *sa pile conjointe est transcrite par* φ.

2. 10. 2. Etude et construction de la pile conjointe à une ramification orientée donnée. Soit une ramification [E, R] orientée, P son ordre d'entrée, S son ordre de sortie. En utilisant les résultats du paragraphe 2. 9. 1, moyennant le passage à la pile réciproque et le changement d'orientation, on obtient la description suivante de sa pile conjointe π' : dans le mot de sortie de π', les familles de la ramification forment des intervalles, où les éléments de chaque famille sont rangés dans l'ordre inverse de l'ordre d'orientation ; après la sortie de π' d'un des éléments d'une famille, viennent immédiatement les sorties des autres éléments de cette famille, puis l'entrée du prédécesseur des points de la famille : les noeuds entrent tous de cette façon. Après la dernière entrée, sortent les racines. L'entrée d'une feuille n'est pas immédiatement précédée d'une sortie. On peut dissocier le mot d'entrée en ses deux traces sur les ensembles des noeuds et des feuilles : la trace sur l'ensemble des feuilles est la suite de feuilles de la ramification orientée donnée.

Complètons [E, R] en une arborescence en lui adjoignant une racine δ: aussi tout point aura un prédécesseur. π' possède S pour ordre d'entrée ; son ordre associé R' est celui qui a été défini au paragraphe 2. 9, moyennant le changement de notation qui consiste à remplacer P par S^{-1} (à cause du changement d'orientation) : d'après (9) (2.9.1), (10) et (11) (2.9.3) :

(12) $x R' y \Longleftrightarrow x S y$ et $pd(y) S pd(x) \Longleftrightarrow x S y$ et $pd(x) R y$ et $y \neq pd(x)$
$\Longleftrightarrow x S y$ et $pd(x) R pd(y)$.

Nous construirons la pile π' grâce à un algorithme déduit de celui du paragraphe 2. 2. 1. Comme après la sortie d'un élément sortent tous ceux de sa famille, il sera inutile de revenir au test qui permet la sortie, pourvu qu'on sache reconnaître le premier des éléments de la famille qui est entré dans la pile (donc le dernier à sortir). Cette reconnaissance sera aisée si à l'entrée du premier élément d'une famille, on fait d'abord entrer un séparateur $\ddagger$. Alors, quand un élément sortira de la pile, sortiront à la suite tous les sommets successifs jusqu'au premier $\ddagger$ rencontré, qui sortira aussi et pourra servir, dans le mot de sortie, de séparateur entre deux familles différentes. A l'entrée d'un élément z sur un sommet e, pour savoir si on doit d'abord faire entrer $\ddagger$, on testera si e est équivalent à z : si $e \not\sim z$ (1), $\ddagger$ entre avant z. Comme $e S z$, $e \sim z$ entraîne d'ailleurs $e R' z$, donc que z entre dans la pile.

Le test d'entrée ($e R' z$) n'a besoin d'être effectué que lorsque z est une feuille. Alors z n'est pas le prédécesseur de e ; d'autre part, $e S z$. D'après (12), le test se réduit donc à celui de ($pd(e) R z$).

(1) $\not\sim$ signifie "non équivalent à". Rappelons que e est dit équivalent à z si e et z appartiennent à la même famille.

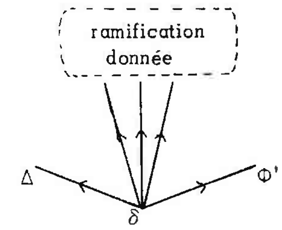


Figure 2. 8.

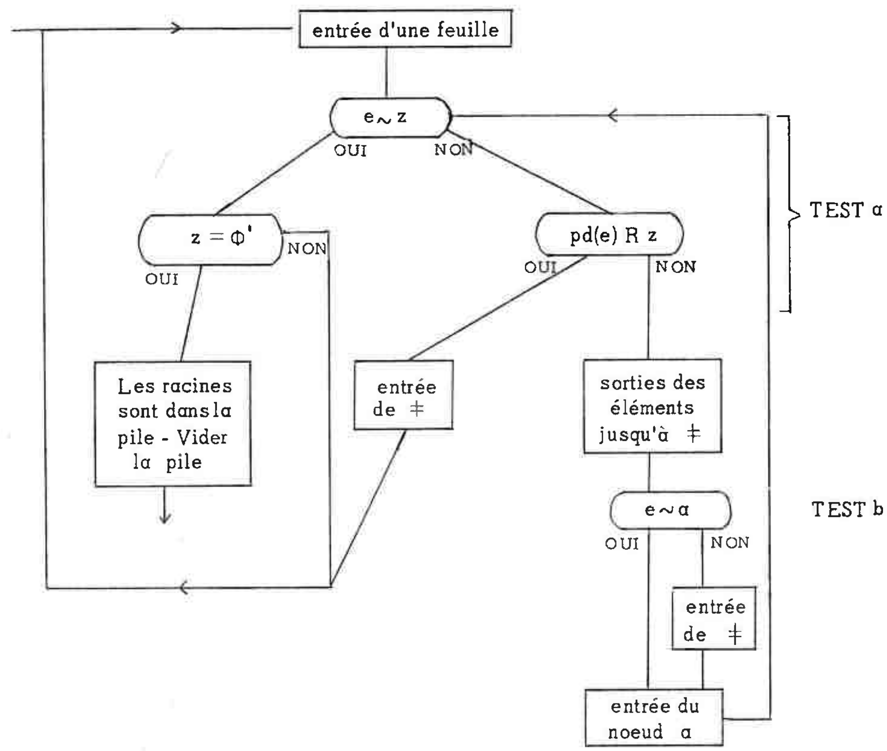


Figure 2. 9.

La ramification [E, R] a été complétée en une arborescence par l'introduction d'une racine δ. D'après le paragraphe 2. 2. 1, on doit aussi adjoindre à E un élément Δ tel que Δ R'x pour tout x de E : il suffit de convenir que Δ a δ pour prédécesseur et est, dans l'ordre d'orientation, le premier des points qui possèdent cette propriété (Δ est équivalent aux racines de la ramification donnée). Enfin, au paragraphe 2. 2. 1, on introduisait un élément Φ dont l'entrée dans la pile marquait la fin du déroulement de l'algorithme. Ici, nous le remplacerons par un élément Φ' qui soit une feuille équivalente à Δ et le dernier élément, dans l'ordre d'orientation, de la famille de Δ : Φ' ne possède pas les propriétés indiquées pour Φ, mais à son entrée dans la pile, celle-ci contient les racines de la ramification donnée et Δ : en effet Φ' devrait ressortir aussitôt avec toute sa famille et déterminer ainsi δ ; l'entrée de Φ' marque donc encore la fin du processus, δ n'entre pas dans la pile, et pendant le déroulement de l'algorithme, il existe toujours une première feuille non encore entrée. Cette feuille est notée z sur l'organigramme de la figure 2. 9, qui décrit l'algorithme ; e est le sommet de la pile à chaque étape ; α est le prédécesseur de la famille qui vient de sortir.

On ne construit pas en réalite une pile simple, mais une pile sur l'ensemble E $\cup$ {Δ, ǂ} dont la trace sur E est la pile simple π'. Le mot de sortie est formé des divers $\gamma(x)$, où x est un noeud de la ramification, séparés par ǂ, et obtenus dans l'ordre de sortie des prédécesseurs x.

Exemple (figure 2. 10) : suite de feuilles : Δ v w x y z Φ'.

					x								z							
				ǂ	ǂ		α_1					y	y			α_3				
			w	w	w	w	w				ǂ	ǂ	ǂ		ǂ	ǂ		α_4		
		v	v	v	v	v	v			α_2	α_2	α_2	α_2	α_2	α_2	α_2	α_2	α_2		
	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ		ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ		α_5
Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ

Mot de sortie : x ǂ α_1 w v ǂ z y ǂ α_3 ǂ $\alpha_4 \alpha_2$ ǂ α_5.

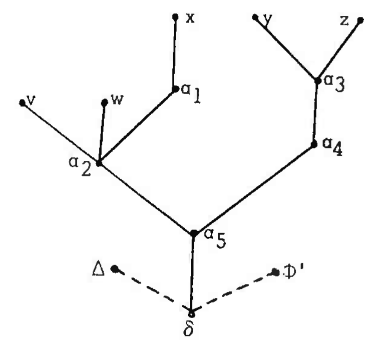


Figure 2. 10.

Remarque : Dans l'algorithme précédent, remplaçons le signe ǂ par σ : le mot de sortie obtenu contient chaque élément de E qui y est placé lors de sa sortie de la pile, et un σ qui est placé à chaque entrée d'un noeud. Plaçons aussi dans ce mot un σ lors de l'entrée d'une feuille : on obtient ainsi un mot λ. Le réfléchi $\tilde{\lambda}$ de λ n'est autre que le mot de description de la pile réciproque de la pile conjointe à la ramification orientée, c'est-à-dire de la pile adjointe à la même ramification, munie de l'orientation inverse.

2. 10. 3. Couples sur lesquels portent les tests. Deux sortes de tests sont effectués :

a) (e $\sim$ z) et (pd(e) R z) ; ce test est effectué immédiatement après l'entrée de e, que e soit une feuille ou un noeud. z est donc la première feuille qui suit e dans l'ordre S.

b) (e $\sim$ α) ; pour chaque noeud α, il existe un seul test portant sur e sommet de la pile et α premier élément à entrer, car α entre immédiatement. Identifions ce sommet e grâce au théorème 2. 2. Soit x le plus grand minorant strict de l'intervalle R(α) pour l'ordre S ; pd(x) n'appartient pas à R(α), et donc x S α S pd(x), d'où pd(x) Rα (théorème 2. 4). Soit v un élément de $]_S$ x, α[; α R v et par suite pd(x)R v, avec pd(x)$\neq$v : xR'v d'après (12). D'autre part, α R pd(v), d'où α = pd(v) ou pd(v) $\bar{R}$ α : v $\bar{R}'\alpha$ d'après (12). (x, α) est un couple comparé, donc x est le sommet cherché.

En résumé, on a à effectuer les tests suivants :
(test a) pour tout noeud ou toute feuille b, $(b \sim z)$ et $(\mathrm{pd}(b)\ R\ z)$ où z est la première feuille qui majore strictement b dans l'ordre S ($b \sim z$ entraîne $\mathrm{pd}(b)\ R\ z$).
(test b) pour tout noeud b, $(x \sim b)$ où x est le plus grand minorant strict de l'intervalle R(b) dans l'ordre S.
Les couples de points ainsi comparés sont des couples de points consécutifs (2. 6. 3). Inversement, tout couple de points consécutifs b, z où z est une feuille, est l'objet d'un test a.

Revenons à la pile simple $\pi' = (u_0, \dots, u_u)$ conjointe à [E, R] :
1) si y est le sommet de u_{i-1} alors que i est l'entrée d'un noeud ou d'une feuille z, le couple (y, z) fait l'objet d'un test, donc y et z sont consécutifs et $\mathrm{pd}(y)\ R\ z$.
2) si u_{i-1} est vide alors que i est l'entrée d'un point z, le couple (Δ, z) fait l'objet d'un test, Δ et z sont des points consécutifs de la ramification complétée : il existe un chemin de [E, R] ayant une racine pour origine, z pour extrémité, et dont tout point soit le premier de sa famille.
3) si i est une sortie d'un point y, i-1 son entrée et z la première feuille dont l'entrée est supérieure à i, le couple (y, z) fait l'objet d'un test a, y et z sont consécutifs et $\mathrm{pd}(y)\ \bar{R}\ z$.
4) si i est une sortie d'un point y, i-1 son entrée et si l'entrée d'aucune feuille n'est supérieure à i, le couple (y, Φ) fait l'objet d'un test a, y et Φ sont des points consécutifs de la ramification complétée : il existe un chemin de [E, R] ayant une racine pour origine, y pour extrémité, et dont tout point soit le dernier de sa famille.

2. 10. 4. Piles conjointes à certaines sous-ramifications d'une ramification orientée.

Théorème 2. 8. *Soit [E, R] une arborescence orientée et $[E_1, R_1]$ une sous-arborescence complète de même racine, dont l'orientation est induite par celle de [E, R] (2. 4. 1). Tout état de la pile conjointe à $[E_1, R_1]$ est un état de la pile conjointe à [E, R].*

Soient $y \in E_1$ et $x \in E$ tels que $x\ R'\ y$. D'après (12), $\mathrm{pd}(x)\ R\ y$ et $y \neq \mathrm{pd}(x)$; un point x' de la famille de x vérifie $\mathrm{pd}(x)\ R\ x'\ R\ y$. Comme la racine de [E, R] appartient à E_1, d'après la définition d'une pseudo-arborescence complète (1. 4. 1), pd(x), x' et x appartiennent à E_1. R_1 et l'ordre de sortie S_1 de $[E_1, R_1]$ sont les restrictions à E_1 de R et de S, ordre de sortie de [E, R] ; (12) montre alors que les piles conjointes aux deux arborescences orientées ont même état de sommet y.

Théorème 2. 9. *Soient une ramification orientée [E, R] et la sous-ramification $[E_1, R_1]$ formée de ses racines, de ses feuilles et de ceux de ses points qui ne sont pas seuls dans leur famille ; l'orientation de $[E_1, R_1]$ est induite par celle de [E, R]. Tout état de la pile U_1 conjointe à $[E_1, R_1]$ est un état de la pile U conjointe à [E, R]. Pour tout état u de U, il existe un état u_1 de U_1 qui en diffère au plus par le sommet, le sommet de u étant transformé en celui de u_1 par le quasi-isomorphisme qui transforme [E, R] en $[E_1, R_1]$. Si y est un point de E_1 autre qu'une racine, la suite des éléments dont la sortie est comprise entre la sortie de y et la première entrée suivante est la même pour U et U_1.*

$[E_1, R_1]$ est une ramification sous-réduite de [E, R], donc déduite de [E, R] par un quasi-isomorphisme φ (1. 5). Soient S et S_1 les ordres de sortie de [E, R] et $[E_1, R_1]$, R' et R'_1 les ordres associés aux piles U et U_1. S_1 est restriction de S. Envisageons $y \in E_1$ et $x \in E$ tels que $x\ R'\ y$; x appartient à E_1, sinon sa sortie de U suivrait immédiatement son entrée, ce qui contredirait $x R' y$. Soient z et z_1 les prédécesseurs de x pour R et R_1 : $z_1\ R\ z$, $z_1 = \varphi(z)$ (1. 5. 3) ; $z\ R\ y$ et $y \neq z$ entraîne $z_1\ R\ y$ et $y \neq z_1$; réciproquement, $z_1\ R\ y$ et $y \neq z_1$ s'écrit $\varphi(z)\ R\ \varphi(y)$ et $\varphi(y) \neq \varphi(z)$, qui entraîne $z\ R\ y$ (1. 5. 1, et $y \neq z$. Donc, d'après (12), $R'(y) = R'_1(y)$.

Si y' n'appartient pas à E_1, il existe y dans E_1 tel que $\varphi(y') = y$, et un chemin $y, y_1, \dots, y_p, y'$ de [E, R] dont chaque point, sauf y, est seul dans sa famille. L'étude de la pile conjointe à [E, R] montre que cette pile contient des états $u'y', u'y_p, \dots, u'y_1, u'y$: u'y est un état de U_1 d'après la première partie du théorème.

La dernière partie résulte du fait qu'avec un point, E_1 contient toute sa famille dans [E, R].

2. 10. 5. Retour au problème posé en 2. 10. 1. Soit un ensemble D et une suite $f_1, f_2, \dots, f_p$ d'éléments de D. Utilisons $\Delta, f_1, f_2, \dots, f_p, \Phi'$ comme suite de feuilles à l'entrée de l'algorithme de 2. 10. 2, et faisons entrer d'autre part dans la pile comme noeuds a des éléments de D distincts n'appartenant pas à cette suite. Imposons aux tests les seules conditions qui permettent d'assurer $\Delta \sim \Phi'$ et $\mathrm{pd}(\Delta)\ R\ z$, c'est-à-dire : si la pile ne contient aucun séparateur ‡ au moment d'un test (a), $e \sim z$ si $z = \Phi'$, et $\mathrm{pd}(e)\ R\ z$ (avec éventuellement $e \sim z$) si $z \neq \Phi'$; si la pile contient un séparateur à ce moment, $e \not\sim \Phi'$ et $\mathrm{pd}(e)\ \bar{R}\ \Phi'$. Dans les autres cas, les réponses aux tests seront choisies arbitrairement. Δ entre le premier dans la pile et ne sort jamais ; Φ' n'entre pas dans la pile ; $f_1, f_2, \dots, f_p$ y entrent tous si on envisage une exécution de l'algorithme qui soit finie.

L'algorithme engendre alors une pile sur un ensemble $E \cup \{\Delta, ‡\}$, où E est un sous-ensemble de D ; sa trace sur E est une pile simple π'. D'après le théorème 2. 7, π' est la pile conjointe à une ramification orientée, dont les feuilles sont caractérisées par le fait que leur entrée dans la pile n'est pas immédiatement précédée d'une sortie : la ramification à laquelle π' est conjointe admet $f_1, \dots f_p$ comme suite de feuilles.

En résumé, l'algorithme du paragraphe 2. 10. 2 engendre les piles conjointes à toutes les ramifications orientées ayant une suite donnée comme suite de feuilles, et elles seules, lorsque les éléments de D qui entrent dans la pile sont tous distincts et que les réponses aux tests satisfont à la seule condition qui a été imposée ci-dessus. Comme toute pile simple est conjointe à une ramification orientée et une seule, on peut dire qu'on détermine ainsi toutes les ramifications orientées (E, γ) ayant une suite de feuilles donnée.

Supposons que l'on donne de plus une relation ρ entre D et le monoïde libre D^*. Si à la sortie d'une suite d'éléments formant un mot λ, on fait entrer un a tel que $a \rho \tilde{\lambda}$, pourvu qu'il existe un tel élément a non encore entré, les ramifications engendrées sont les solutions du deuxième problème posé au paragraphe 2. 10. 1.

2. 10. 6. Application de la pile conjointe aux problèmes de minimum du paragraphe 2. 7. Au paragraphe 2. 7, on a vu que, pour une ramification $[E, R]$, le minimum de la fonction $mx(N)$, pour les ordres totaux N vérifiant la condition (OR), était réalisé par un ordre de sortie. Un tel ordre de sortie S est obtenu en orientant la ramification de proche en proche : pour tout point y de la ramification, notons S_y la restriction de S à l'ensemble $R(y)$ des majorants de y, et $mx(S_y)$ la valeur pour l'ordre S_y de la fonction mx associée à la sous-arborescence de $[E, R]$ ayant $R(y)$ pour ensemble de points ; si pour chaque famille $\{x_1, x_2, \dots, x_n\}$ de prédécesseur x, on connaît les S_{x_i} et les $mx(S_{x_i})$, il suffit d'ordonner la famille pour que les $mx(S_{x_i})$ soient décroissants ; on connaît alors S_x (2. 3. 3) et $mx(S_x)$ (2. 7). Si y est une feuille, $R(y)$ est réduit à y et $mx(S_y) = 1$.

Partons d'un ordre de sortie quelconque S' de la ramification (qui est ainsi orientée), et transformons le en S, en réordonnant les familles l'une après l'autre comme il vient d'être dit : à chaque étape, on obtient une nouvelle orientation, d'où un nouvel ordre de sortie, pour lequel tous les ensembles $R(y)$ sont des intervalles (théorème 2. 4). Nous utiliserons la pile conjointe à la ramification, orientée par l'ordre S' donné : cette pile sera construite par l'algorithme du paragraphe 2. 10. 2, et nous y noterons, avec chaque point y, $mx(S_y)$ et les extrémités de l'intervalle $R(y)$ dans la permutation actuelle de E ; à la sortie de la famille $\{x_1, x_2, \dots, x_n\}$ de prédécesseur x, on possède les renseignements nécessaires pour la réordonner : on peut ainsi calculer $mx(S_x)$, connaître les extrémités de l'intervalle $R(x)$ dans la permutation de E, et réordonner les intervalles $R(x_i)$ dans $R(x)$ pour obtenir la nouvelle permutation. Dans la mémoire d'une calculatrice, les permutations successives pourront être stockées séquentiellement, et il faudra effectuer des transferts, ou bien on pourra éviter ces transferts en les stockant sous forme de "liste" (cf. par exemple [44]).

Tout ce qui vient d'être dit reste valable en remplaçant la fonction mx par la fonction k définie au paragraphe 2. 7, ou par la fonction mx' si on cherche seulement à minimiser $mx'(N)$ pour les ordres de sortie N.

2. 11. Construction de la pile conjointe grâce à des tests portant uniquement sur des feuilles.

2. 11. 1. Etude générale. La construction de la pile conjointe à une ramification orientée (ou la détermination d'une ramification orientée parmi celles qui possèdent une suite de feuilles donnée) est donc ramenée à la connaissance de deux relations, liées à la ramification : la relation d'équivalence et la relation $pd(x) R y$, pour les couples étudiés au paragraphe 2. 10. 3. Ces couples ne sont pas formés uniquement de feuilles. Peut-on se contenter des restrictions de ces deux relations à l'ensemble F des feuilles ?

Soit π'' la trace sur l'ensemble F (1. 6. 4) de la pile π' conjointe à une ramification orientée. Chaque test (a) (2. 10. 3) porte sur un couple (b, z) et provoque l'entrée de la feuille z ou la sortie de la famille de b ; on peut lui associer un test sur le couple (y, z), où y est le sommet de la trace sur F de l'état de π' dont b est le sommet : si b est un noeud, b est distinct de y ; peut-on alors ramener le test (a) ($b \sim z$, puis $pd(b) R z$) aux tests analogues ($y \sim z$) et($pd(y) R z$) ? On doit aussi effectuer le test (b) ($x \sim b$) : peut-on encore le ramener aux tests précédents ?

Soit une ramification $[E, R]$. Ses feuilles et ceux de ses noeuds qui sont prédécesseurs d'une feuille forment une sous-ramification $[\mathring{E}, \mathring{R}]$, qu'on appellera le *feuillage* de $[E, R]$. $[\mathring{E}, \mathring{R}]$ est son propre feuillage. La ramification sous-réduite de $[E, R]$ (1. 5. 3) a même feuillage que $[E, R]$. Un feuillage est sa propre ramification sous-réduite. Une orientation de $[E, R]$ détermine une orientation de son feuillage (2. 4. 1).

Si, pour construire la pile conjointe à une ramification orientée, on se limite aux tests sur les feuilles qui viennent d'être proposés, deux ramifications distinctes ayant même feuillage fourniront les mêmes réponses aux tests, alors qu'elles n'ont pas la même pile conjointe. On devra faire une hypothèse supplémentaire qui ne puisse être satisfaite à la fois par deux ramifications distinctes ayant le même feuillage. Le plus simple est de se limiter à étudier (ou à chercher) une ramification qui soit un feuillage [1].

(1) On montre aisément qu'on obtient l'une quelconque des ramifications orientées (E, γ) ayant un feuillage orienté $(\mathring{E}, \mathring{\gamma})$ donné, de la manière suivante : pour chaque noeud y du feuillage, on construit une arborescence orientée $A(y)$ de racine y, qui a pour suite de feuilles $\mathring{\gamma}(y)$, où tout élément de $\mathring{\gamma}(y)$ qui est une feuille de $(\mathring{E}, \mathring{\gamma})$ a y pour prédécesseur, les points de $A(y)$ autres que la racine ou les feuilles n'appartenant ni à $\mathring{E}$, ni à l'arborescence $A(y')$ associée à un autre noeud y' ; on construit une ramification A_0 qui a pour suite de feuilles la suite des racines du feuillage donné, dans l'ordre d'orientation, les noeuds de A_0 n'appartenant ni à $\mathring{E}$, ni à aucune arborescence $A(y)$. Alors E est formé des points de A_0 et des $A(y)$; si x est un noeud de A_0 ou de $A(y)$, il n'est noeud d'aucune autre de ces ramifications : $\gamma(x)$ est le mot qui a pour prédécesseur x dans la ramification orientée dont x est noeud ; pour les autres points x de E, $\gamma(x) = \Lambda$; le mot des racines de (E, γ) est celui de A_0.

Soit donc un feuillage [E, R] orienté : tout noeud est le prédécesseur d'une feuille. Etudions l'algorithme de construction de la pile π' conjointe à ce feuillage. Nous conserverons les notations qui ont déjà été utilisées : R' est l'ordre associé à la pile π' ; il a été défini au paragraphe 2. 10. 2, en (12) ; Q est l'ordre de sortie de cette pile ; pour tout noeud b, x est le sommet de la pile π' lors de l'entrée de b, autrement dit l'élément qui, dans l'ordre S, précède immédiatement R(b) ; y est le sommet de la pile π'' à ce moment ; il en résulte que $y\ R'\ x\ R'\ b$ (théorème 1. 7) ; z est la première feuille qui majore b dans l'ordre S. Nous poserons de plus $a = pd(b)$ et $r = pd(y)$. Démontrons d'abord quelques lemmes :

Lemme 1. $r\ R\ pd(x)\ R\ a$.

Ceci résulte directement de $y\ R'\ x\ R'\ b$ d'après (12) (paragraphe 2. 10. 2).

Lemme 2. *Pour que* x *soit équivalent à* b, *il faut et il suffit que* b *ne soit pas le premier élément, dans l'ordre* S, *de sa famille.*

La condition est évidemment nécessaire. Réciproquement, si b n'est pas le premier élément de sa famille, x qui précède immédiatement R(b) dans l'ordre S est équivalent à b (2. 3. 3).

Lemme 3. *S'il existe une feuille* w *équivalente à* b *et telle que* $w\ S\ b$, w, y, x *et* b *sont équivalents.*

x est équivalent à b d'après le lemme 2. D'après (12), $w\ R'\ b$; b est le sommet d'un état u_i de π' : w appartient à u_i (théorème 1. 7). Le sommet de la trace de u_i sur F est y ; donc $w\ R'\ y\ R'\ b$ (théorème 1. 7) et par suite $b\ Q\ y\ Q\ w$. Comme les familles sont des Q-intervalles, y est équivalent à b.

Lemme 4. *Pour que* $a\ R\ z$, *il faut et il suffit que* b *ne soit pas le dernier élément, dans l'ordre* S, *de sa famille.*

Si b est le dernier élément de sa famille, son S-successeur est a (2. 3. 3) : donc $z\ \bar{S}\ a$, d'où $a\ \bar{R}\ z$. Si b n'est pas le dernier élément de sa famille, son S-successeur appartient à R(a) et c'est une feuille (2. 3. 3) ; c'est donc z.

En particulier, s'il existe une feuille w équivalente à b et telle que $b\ S\ w$, alors $a\ R\ z$.

Lemme 5. *S'il n'existe aucune feuille* w *équivalente à* b *et telle que* $w\ S\ b$, y *n'est pas équivalent à* z *et* $r\ R\ z$.

Comme la ramification étudiée est un feuillage, il existe une feuille w équivalente à b ; ici elle vérifie $b\ S\ w$; d'après les lemmes 1 et 4, $r\ R\ a\ R\ z$. $a \neq z$ car a est un noeud et z une feuille ; $r \neq a$ car y n'est pas équivalent à b ; donc $r \neq pd(z)$.

Etudions maintenant si les réponses aux tests $(y \sim z)$ et $(r\ R\ z)$ déterminent les réponses à $(x \sim b)$, $(b \sim z)$ et $(a\ R\ z)$:

1) Hypothèse : $y \sim z$. D'après le lemme 5, il existe une feuille w équivalente

à b telle que $w\ S\ b$. Alors, d'après le lemme 3 : $x \sim b \sim z$. b et z entrent dans la pile sans séparateur.

2) Hypothèse : $r\ \bar{R}\ z$ (d'où $y \not\sim z$). D'après le lemme 5, il existe une feuille w équivalente à b telle que $w\ S\ b$. Donc (lemme 3) $y \sim x \sim b$; d'où $r = a$, $b \not\sim z$ et $a\ \bar{R}\ z$. b entre dans la pile sans séparateur, z n'entre pas ; la famille de b sort donc.

3) Hypothèse : $y \not\sim z$ et $r\ R\ z$. De $r\ R\ a$ (lemme 1) et $r\ R\ z$, on déduit $a\ R\ z$, car z est une feuille de la ramification. Mais nous allons voir qu'ici on ne peut déterminer le résultat du test $(x \sim b)$ sans hypothèses supplémentaires. En effet, l'hypothèse est réalisée dès que b n'est précédé (dans l'ordre S) par aucune feuille équivalente (lemme 5). Mais selon que b est ou non le premier élément de sa famille, $b \not\sim x$ ou $b \sim x$ (lemme 2). Il est donc nécessaire qu'il n'existe aucune famille dont les deux premiers éléments soient des noeuds. D'autre part, supposons qu'une famille possède deux noeuds consécutifs (pour l'ordre S) b et b', précédés d'une feuille. Etudions b : d'après le lemme 4, $a\ R\ z$; d'après le lemme 3, $y \sim x \sim b$; par suite $r\ R\ z$. z vérifie $b'\ R\ z$ et $b' \neq z$. Donc z n'est pas équivalent à b ni à y. L'hypothèse est satisfaite, et $b \sim x$. Il est donc nécessaire que s'il existe une telle famille, aucune autre famille ne commence par un noeud.

Les réponses aux tests $(y \sim z)$ *et* $(r\ R\ z)$ *ne peuvent donc déterminer les réponses à* $(x \sim b)$, $(b \sim z)$ *et* $(a\ R\ z)$ *que dans les cas :*

(A) *si deux noeuds appartiennent à la même famille, dans cette famille une feuille se trouve entre eux (pour l'ordre d'orientation),*

ou

(B) *pour tout noeud, le premier, dans l'ordre d'orientation, des éléments dont il est le prédécesseur est une feuille.*

Dans (A) on a introduit une condition sur les racines parce qu'elles forment, avec Δ et Φ', la classe des éléments de prédécesseur δ (2. 10. 2). Aucune condition sur les racines ne se trouve dans (B), car Δ est le premier élément de la classe des racines et il a été considéré comme une feuille.

Pour qu'une ramification soit un feuillage et satisfasse à la condition (A), il faut et il suffit qu'elle vérifie (A) et

(A1) *si un noeud est le prédécesseur d'un seul élément* c, c *est une feuille.*
Les ramifications sous-réduites satisfaisant à (A) sont toutes des feuillages.

L'hypothèse (B) suffit pour qu'une ramification soit un feuillage. Inversement, il est clair que tout feuillage possède une orientation telle que (B) soit satisfaite.

2. 11. 2. Etude du cas (A).

Lemme 6. *Dans le cas (A),* x *est une feuille et donc* $x = y$.

D'après le lemme 1, $pd(x)\ R\ a$, d'où $pd(x)\ R\ b$; il existe donc b_1 tel que

$x \sim b_1$ et b_1 R b : $x \neq b_1$ car x $\bar{R}$ b, x n'appartient pas à $R(b_1)$ et $R(b_1)$ contient R(b). x est donc aussi l'élément qui, dans l'ordre S, précède $R(b_1)$; dans leur famille, x et b_1 sont consécutifs. x est donc une feuille.

Lemme 7. *Dans le cas (A), pour que* b *ne soit pas le dernier élément de sa famille, il faut et il suffit que* b *soit équivalent à* z.
Si b n'est pas le dernier élément de sa famille, l'élément suivant est une feuille z'. R(z') est réduite à z' : z' est donc le S-successeur de b (2. 3. 3). D'où z' = z, b est équivalent à z. Réciproquement, $b \sim z$ entraîne a R z et il suffit d'appliquer le lemme 4.

Revenons à l'hypothèse 3 du paragraphe 2. 10. 2 : $y \not\sim z$ et r R z. Supposons que b ne soit pas le premier élément de sa famille. Alors b est équivalent à x = y (lemme 2), b n'est donc pas équivalent à z, b est le dernier élément de sa famille (lemme 7) et a $\bar{R}$ z (lemme 4) : comme a = r, ceci contredit r R z, b est donc le premier élément de sa famille. On en déduit $x \not\sim b$ (lemme 2) et $b \sim z$ (lemme 7) : un séparateur entre, puis b, puis z. *Les tests sur* y *et* z *sont donc suffisants dans le cas (A).*

Il est facile d'adapter à ce cas l'algorithme de 2. 10. 2. Après avoir déterminé la classe des éléments dont un noeud a est le prédécesseur, on placera a dans une "mémoire auxiliaire" μ, en attendant de savoir s'il doit être précédé dans la pile par un séparateur ǂ. On a vu (lemme 6) que y est le sommet de la pile construite. "entrée de (μ)" ne modifie pas la pile lorsque $\mu = \Lambda$.

2. 11. 3. Etude du cas (B). L'hypothèse du lemme 3 est toujours réalisée : x est équivalent à b et à y. Pour $y \not\sim z$ et r R z, $b \not\sim z$ et a R z : b entre, puis un séparateur, puis z.

Les tests sur y *et* z *suffisent donc aussi dans le cas (B).* Et tout noeud entre sans séparateur. L'algorithme de détermination de la ramification est décrit figure 2. 12.

2. 12. Ramification gauche d'une ramification orientée.

2. 12. 1. Définition. Au chapitre 6, nous serons amenés à déterminer une ramification orientée (E, γ) par une autre, qui a le même ensemble de points E et où toute famille, sauf celle des racines, commence par une feuille. Pour cela, nous allons ici associer à (E, γ), de manière injective, une telle ramification (E, γ').

Désignons par E_1 l'ensemble formé des racines de (E, γ) et de tout point qui n'est pas le premier de sa famille, par R l'ordre associé à (E, γ), par S son ordre de sortie. La relation R' définie dans E par

$$x \, R' \, y \iff (x \, R \, y \text{ et } x \in E_1) \text{ ou } x = y$$

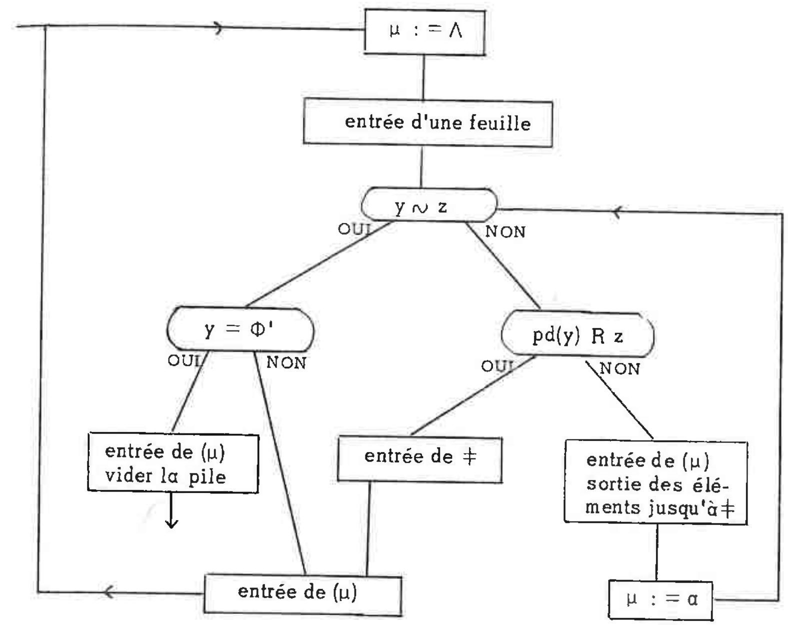


Figure 2. 11

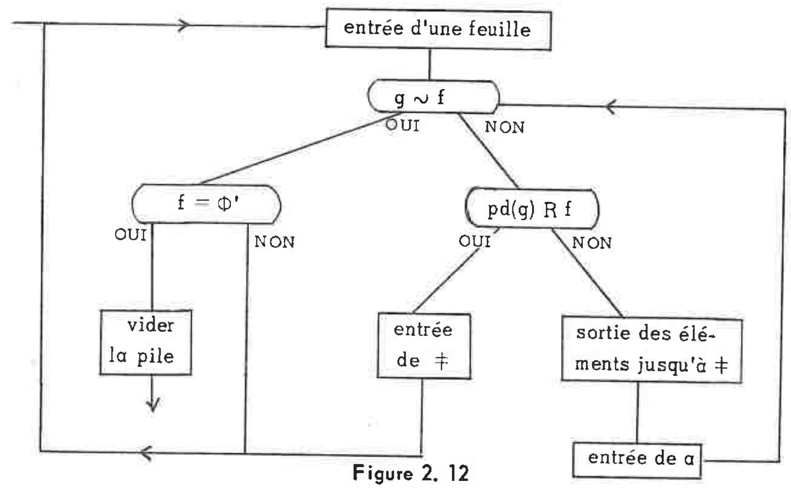


Figure 2. 12

est transitive ; $R'(x) = R(x)$ si $x \in E$, $R'(x) = \{x\}$ sinon. L'ordre total S et la relation R satisfont aux hypothèses du théorème 2. 4, il en est de même pour S et R'. Il existe donc une ramification orientée (E, γ') unique dont R' est l'ordre associé et S l'ordre de sortie. Nous l'appellerons *ramification gauche* de (E, γ).

Si un isomorphisme φ transforme (E, γ) en (E'', γ''), φ est aussi un isomorphisme de (E, γ') sur la ramification gauche de (E'', γ''). Si toute famille de (E, γ) autre que celle des racines commence par une feuille, les relations R et R' sont confondues, $(E, \gamma) = (E, \gamma')$.

2. 12. 2. Propriétés de (E, γ'). (E, γ') a mêmes racines que (E, γ) ; ses feuilles sont les feuilles de (E, γ) et les noeuds de (E, γ) n'appartenant pas à E_1 ; ses noeuds sont ceux de (E, γ) qui appartiennent à E_1. Identifions le R'-prédécesseur z d'un point x qui n'est pas une racine : zRx et $z \neq x$; de plus, dans le R-intervalle $]z, x[$ n'existe aucun élément z' de E_1, sinon z' vérifierait $zR'z'R'x$. Le R'-prédécesseur de x est, pour R, le plus grand minorant strict de x qui appartient à E_1. De ce résultat, nous déduirons diverses conséquences :

a) Si $z \in E_1$ et si y est la première feuille de R(z), d'après l'étude faite en 2. 3. 3, z est le R'-prédécesseur de y, la famille $\gamma'(z)$ a y pour premier élément : (E, γ') est sa propre ramification gauche.

b) Deux points x et y qui appartiennent à la même famille λ de (E, γ) appartiennent aussi à la même famille λ' de (E, γ'). Si y suit immédiatement x dans λ, y appartient à E_1 et $R'(y) = R(y)$; dans l'ordre de sortie S, x précède immédiatement R(y) : y suit aussi immédiatement x dans λ'.

c) Si le R-prédécesseur y d'un point x appartient à E_1, y est aussi le R'-prédécesseur de x; sinon , dans (E, γ'), x et y appartiennent à la même famille. Supposons que x soit le dernier point de sa famille dans (E, γ) : pour l'ordre S, x est le prédécesseur de y ; si y appartient à E_1, x est le dernier point de $\gamma'(y)$; sinon, x et y sont consécutifs dans leur famille commune.

On déduit de cette étude la forme des familles de (E, γ') : si z est un noeud de (E, γ) appartenant à E_1, $\gamma'(z)$ commence par la première feuille α_0 de R(z), qui est suivie des autres éléments de sa famille $\gamma(\alpha_1)$; si $\alpha_1 \in E_1$, $\alpha_1 = z$, $\gamma'(z) = \gamma(\alpha_1)$; sinon $\gamma(\alpha_1)$ est suivi de la famille $\gamma(\alpha_2)$ dont α_1 est le premier élément, etc... Finalement, $\gamma'(z) = \gamma(\alpha_1) \ldots \gamma(\alpha_p)$ où : $p \geqslant 1$; $\alpha_p = z$; pour $1 \leqslant i \leqslant p-1$, α_i est un noeud de (E, γ) et une feuille de (E, γ') ; tout autre élément de $\gamma(z)$ est soit α_0, feuille de (E, γ), soit un point de E_1 : c'est une feuille de (E, γ) ou un noeud de (E, γ') ; pour $1 \leqslant i \leqslant p$, α_{i-1} est le premier élément de $\gamma(\alpha_i)$.

La figure 2. 13 schématise une arborescence orientée et sa ramification (arborescence) gauche (orientation de "gauche à droite").

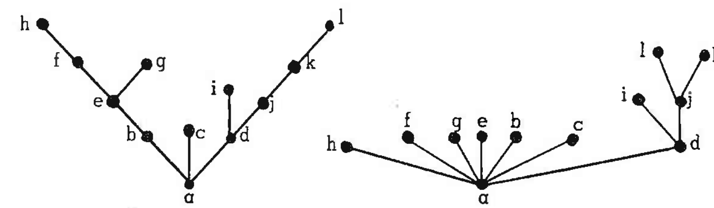


Figure 2. 13

2. 12. 3. Existence et unicité d'une ramification orientée ayant une ramification gauche et un ensemble de feuilles donnés. La donnée de la ramification (E, γ') et de l'ensemble F des feuilles de (E, γ) permet de déterminer pour chaque mot $\gamma'(z)$ la suite des α_i ; on connaît alors les $\gamma(\alpha_i)$ et $\gamma(z)$, c'est-à-dire pour tout noeud u de (E, γ) le mot $\gamma(u)$. Par suite, il existe au plus une ramification orientée (E, γ) qui a un ensemble de feuilles F donné et dont une ramification orientée donnée (E, γ') est ramification gauche.

Pour que (E, γ) existe, il est nécessaire que toute famille de (E, γ'), sauf celle des racines, commence par une feuille et que F soit un ensemble de feuilles de (E, γ') qui contienne le premier élément de chacune de ces familles. Supposons qu'il en soit ainsi. Envisageons un noeud z de (E, γ') et, dans $\gamma'(z)$, la suite (éventuellement vide) des feuilles de (E, γ') n'appartenant pas à F : $\alpha_1, \ldots, \alpha_{p-1}$; posons $\alpha_p = z$ $(p \geqslant 1)$: on peut écrire $\gamma'(z) = \gamma(\alpha_1) \ldots \gamma(\alpha_p)$ où, pour $2 \leqslant i \leqslant p$, α_{i-1} est le premier élément de $\gamma(\alpha_i)$; notons d'autre part que, d'après l'hypothèse, le premier élément α_0 de $\gamma(\alpha_1)$ appartient à F. On définit ainsi, pour toute feuille u de (E, γ') n'appartenant pas à F, et pour tout noeud u de (E, γ'), un mot $\gamma(u)$: chacun de ces mots commence par une feuille de (E, γ') : si $u \in F$, posons $\gamma(u) = \Lambda$. Tout point de E qui n'est pas une racine de (E, γ') appartient à un, et un seul, des mots $\gamma(u)$: d'après 1. 4. 1, ces mots, avec $\gamma(\Delta)$, mot des racines de (E, γ'), définissent une ramification orientée (E, γ) dont F est l'ensemble des feuilles ; soit E_1 l'ensemble de ses racines et de ceux de ses points qui ne sont pas premiers de leur famille : les noeuds de (E, γ) qui appartiennent à E_1 sont les noeuds de (E, γ'). Si R est l'ordre associé à (E, γ), d'après 2. 3. 3, la première feuille de $R(z) = R(\alpha_p)$ est α_0. Il résulte alors de la construction des familles de la ramification gauche de (E, γ) que cette ramification est (E, γ').

Théorème 2. 10. *Pour qu'il existe une ramification orientée* (E, γ) *d'ensemble de feuilles* F *donné, dont une ramification orientée* (E, γ') *donnée soit la ramification gauche, il faut et il suffit que tout élément de* F *soit une feuille de* (E, γ') *et que chaque famille de* (E, γ') *autre que celle des racines commence par un élément de* F. (E, γ) *est alors unique.*

2. 12. 4. Entrées et sorties de la pile strictement attachée à (E, γ'). L'entrée d'un noeud de (E, γ') est suivie de celle d'une feuille de (E, γ). Si un noeud x de (E, γ) qui n'appartient pas à E_1 a pour successeur, dans l'ordre S, un autre noeud y de (E, γ), x est le premier point de sa famille, mais aussi le dernier et y est le R-prédécesseur de x (2. 3. 3), donc $\gamma(y) = {}'x'$. Soit u_i un état de la pile, i' le plus petit entier supérieur à i qui est l'entrée d'une feuille de (E, γ) ou l'indice du dernier état de la pile. i'-1 est la seule entrée possible d'un noeud de (E, γ) comprise entre i et i'. Entre i et i' se trouvent m $(\geqslant 0)$ sorties d'éléments de u_i, $j_1 < \ldots < j_m$; posons $i = j_0$ et $i' = j_{m+1}$; j_k et j_{k+1} $(k = 0, \ldots, m)$ sont séparés par l'entrée et la sortie de q $(\geqslant 0)$ feuilles de (E, γ'), qui sont aussi des noeuds de (E, γ), $x_1, \ldots, x_q$: dans l'ordre S, $x_1, \ldots, x_q$ forment un intervalle et, pour $2 \leqslant k \leqslant q$, $\gamma(x_k) = {}'x_{k-1}'$.

CHAPITRE III

PILES ET PSEUDO-RAMIFICATIONS

3. 1. Piles simples dont une pile donnée est transcrite.

Nous avons vu au paragraphe 1. 7 que toute pile U sur un ensemble E est transcrite d'une pile simple U_0 sur un ensemble E_0, par une application f de E_0 dans E ; le mot d'entrée de U est transcrit du mot d'entrée de U_0 par l'application f ; cette condition définit d'ailleurs f, puisque tout élément de E_0 a une occurrence dans le mot d'entrée de U_0, de sorte qu'il existe au plus une transcription transformant une pile simple donnée U_0 en une pile donnée U sur E.

Soient U_0 et U_1 deux piles simples, sur deux ensembles E_0 et E_1, dont une pile U sur un ensemble E est transcrite. Les mots d'entrée α_0 et α_1 de U_0 et U_1 ont même longueur, celle du mot d'entrée de U ; tout élément de E_0 a une occurrence et une seule dans α_0 ; associons-lui l'élément de E_1 qui a même rang dans α_1 : nous définissons ainsi une bijection g de E_1 sur E_0. Par récurrence sur j, le $j^{\text{ème}}$ état de U_0 est transcrit par g du $j^{\text{ème}}$ état de U_1 ; la pile U_1 est donc transcrite en U_0 par g, et par suite U_1 est isomorphe à U_0.

Réciproquement, soit une pile U transcrite d'une pile simple U_0 par une application f, et U_1 une pile simple isomorphe à U_0, c'est-à-dire que U_0 est transcrite de U_1 par une bijection g. Comme le produit de deux transcriptions est une transcription ($f^{**} \circ g^{**} = (f \circ g)^{**}$), U est transcrite de U_1 (par $f \circ g$).

Théorème 3. 1. *Toute pile* U *sur un ensemble* E *est transcrite d'une pile simple* U_0. *Si* E, U, *et* U_0 *sont donnés, la transcription est unique. Les piles simples dont* U *est transcrite sont les piles isomorphes à* U_0 *et elles seules.*

3. 2. Pseudo-Ramifications.

3. 2. 1. Définitions. Au paragraphe 1. 4. 5, nous avons vu qu'une ramification orientée est déterminée par la donnée d'un ensemble d'index I et d'une *bijection* définie sur I : un ensemble d'index est un sous-ensemble du monoïde libre déduit de l'ensemble des entiers naturels, qui, avec tout mot $\lambda = \lambda_1 r$ de dernier élément r, contient le mot λ_1 s'il n'est pas vide et les mots $\lambda_1\, r'$ pour $0 \leqslant r' \leqslant r$. Tout ensemble d'index I définit une ramification orientée, que nous noterons simplement I par abus de langage. Deux ramifications orientées isomorphes ont même ensemble d'index I, et la ramification orientée I leur est isomorphe.

Soit un ensemble E non vide. Nous appelons *pseudo-ramification dans* E tout couple d'un ensemble d'index I et d'une application quelconque f de I dans E. Nous dirons que la pseudo-ramification (I, f) est *l'image* par f de la ramification orientée I.

Une ramification orientée quelconque, d'ensemble de points E', détermine son ensemble d'index I et une bijection g de I sur E' ; si on donne de plus une application f' de E' dans E, on peut définir la pseudo-ramification $(I, f' \circ g)$: nous dirons encore que *cette pseudo-ramification est l'image par* f' *de la ramification orientée donnée* g(I) *et nous la noterons* $((g(I), f'))$. Pour que $((\mathcal{R}_1, f'_1)) = ((\mathcal{R}_2, f'_2))$, il faut que les ramifications orientées $\mathcal{R}_1$ et $\mathcal{R}_2$ aient même ramification d'index I, donc soient isomorphes : $\mathcal{R}_2 = h(\mathcal{R}_1)$; alors, si $\mathcal{R}_1 = g(I)$, $((\mathcal{R}_2, f'_2)) = ((h \circ g(I), f'_2)) = (I, f'_2 \circ h \circ g)$; pour que $((\mathcal{R}_1, f'_1)) = ((h(\mathcal{R}_1), f'_2))$, il faut et il suffit que $f'_1 = f'_2 \circ h$:

$$((\mathcal{R}, f \circ h)) = ((h(\mathcal{R}), f)) \text{ si } h \text{ est un isomorphisme.}$$

Plus généralement, nous dirons qu'un graphe (E', Γ) et une application f de E' dans E définissent un *pseudo-graphe* dans E $(E', \Gamma ; f)$, avec la convention que $(E_1, \Gamma_1 ; f_1) = (E_2, \Gamma_2 ; f_2)$ si, et seulement si, le graphe (E_2, Γ_2) est image de (E_1, Γ_1) par un isomorphisme h et $f_1 = f_2 \circ h$. Le pseudo-graphe sera dit *fini* si l'ensemble E' l'est.

Si (I, f) est une pseudo-ramification dans E et g une application de E dans un ensemble E_1, nous dirons que g *transforme* (I, f) en la pseudo-ramification $(I, g \circ f)$ dans E_1.

Exemple de pseudo-ramification : soit un ensemble A, et D un sous-ensemble fini du monoïde libre déduit de A, tel que tout facteur gauche d'un mot de D appartienne à D. Soit $\sqsubset$ la relation dans D :

$\lambda \sqsubset \lambda'$ *si, et seulement si, le mot* λ *est facteur gauche du mot* λ'.

$[D, \sqsubset]$ est une ramification ; orientons-la d'une manière quelconque. A tout mot de D, associons l'élément de A qui le termine : l'image, par l'application f ainsi définie, de la ramification $[D, \sqsubset]$ orientée est une pseudo-ramification ; f transcrit les chemins de $[D, \sqsubset]$ dont l'origine est une racine en les mots de D. On peut utiliser une telle pseudo-ramification pour stocker un dictionnaire (figure 3. 1, voir par exemple [56]) : D est alors l'ensemble des mots du dictionnaire, terminés par un symbole Φ qui permet de reconnaître leur fin, et de leurs facteurs gauches.

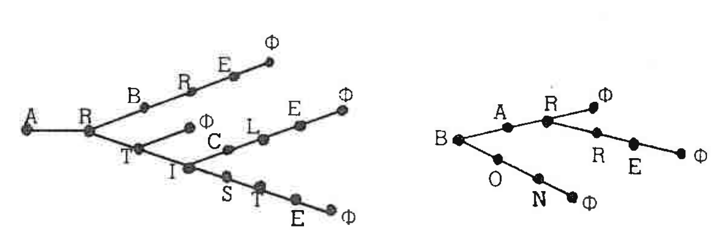


Figure 3. 1

Soit une pseudo-ramification (I, f) dans un ensemble E. L'application f transcrit les familles de la ramification orientée I, qui sont des mots du monoïde libre I^*, en des mots du monoïde libre E^*, qu'on appelle *familles de la pseudo-ramification*. La famille des racines de I est transcrite en la *famille des racines* de (I, f) ; si I est une arborescence, (I, f) s'appelle une *pseudo-arborescence* : elle a une seule racine. Toute autre famille de I est la suite $\gamma(x)$ des éléments de I ayant le même prédécesseur x : on dit que f(x) est *prédécesseur* de la famille transcrite de $\gamma(x)$. La suite de feuilles (2.6.1) de la ramification orientée I est transcrite en un mot de E^*, appelé *suite de feuilles de la pseudo-ramification*. On nomme *hauteur* de (I, f) la hauteur de I. Si J est une sous-ramification de I, elle est orientée par l'orientation de I (2. 4. 1). Son image par la restriction de f à J est une pseudo-ramification, dite *sous-pseudo-ramification* de (I, f). Si J est une sous-arborescence complète (1.4.1) de I, son image est appelée *sous-pseudo-arborescence complète* de (I, f).

La pseudo-ramification (I, f) peut être donnée par sa *matrice des index*, définie comme pour une ramification (1.4.5). Elle est aussi déterminée par toute matrice qui définit à la fois I (ou une ramification orientée isomorphe) et f. Il en est ainsi pour une matrice associée à I bordée d'un vecteur qui pour chaque ligne, correspondant à x de I, donne f(x). Il en est de même pour une matrice à trois colonnes qui, pour chaque x de I, donnent f(x), le numéro de la ligne de la matrice qui correspond au lien vertical de x, et le numéro de celle qui correspond au lien horizontal, lorsque ces liens existent : une telle matrice est appelée *matrice d'enchaînement de la pseudo-ramification*. En particulier, on peut ranger les éléments de I dans l'ordre d'entrée de la ramification : on a vu qu'alors il suffisait d'indiquer dans la deuxième colonne l'existence du lien vertical.

3. 2. 2. Réunion horizontale ou verticale de deux pseudo-ramifications. Soient deux pseudo-ramifications (I_1, f_1) et (I_2, f_2) dans un ensemble E ; désignons par k le nombre des racines de (I_1, f_1). A tout index $\lambda \in I_2$, associons la suite d'entiers $t(\lambda)$ obtenue en ajoutant k au premier des entiers qui composent λ ; $t(I_2)$ est un ensemble, disjoint de I_1, de suites d'entiers qui sont les points d'une ramification orientée I'_2 isomorphe à I_2. La réunion de I_1 et de $t(I_2)$ est un ensemble d'index I, qui définit une ramification d'index : cette ramification n'est autre que la réunion horizontale (1. 4. 2) des ramifications orientées I_1 et I'_2. Définissons une application f de I dans E, par :

$$f(x) = f_1(x) \text{ si } x \in I_1 \quad ; \quad f(x) = f_2 \circ t^{-1}(x) \text{ si } x \in t(I_2).$$

La pseudo-ramification (I,f) s'appelle *réunion horizontale* de (I_1, f_1) et (I_2, f_2). On définit ainsi une loi de composition interne de l'ensemble des pseudo-ramifications dans un ensemble E ; il est facile de voir que cette loi de composition interne est associative (mais non commutative). On note

$$(I, f) = (I_1, f_1) \underset{H}{\cup} (I_2, f_2).$$

Soient φ_1 et φ_2 les mots des racines de (I_1, f_1) et (I_2, f_2) : $\varphi_1 = f_1^*(\Psi_1)$ et $\varphi_2 = f_2^*(\Psi_2)$ où Ψ_1 et Ψ_2 sont les mots des racines de I_1 et I_2. Le mot des racines de I est $\Psi_1 \; t^*(\Psi_2)$; celui de la pseudo-ramification (I, f) est donc :

$$f^*[\Psi_1 \; t^*(\Psi_2)] = f^*(\Psi_1)\,(f \circ t)^*(\Psi_2) = f_1^*(\Psi_1)\, f_2^*(\Psi_2) = \varphi_1 \varphi_2.$$

Si α_1 et α_2 sont les suites de feuilles de (I_1, f_1) et (I_2, f_2), on montre de la même façon que la suite de feuilles de (I, f) est $\alpha_1 \alpha_2$; on utilise le fait que la suite de feuilles de la réunion horizontale de deux ramifications orientées est obtenue par concaténation des suites de feuilles de ces ramifications (2. 6. 2).

Envisageons maintenant deux pseudo-ramifications (I_1, f_1) et (I_2, f_2) telles que la suite de feuilles de la première soit la famille des racines de la seconde. Tout λ de I_2 s'écrit $r\,\lambda'$, où r est un entier et λ' une suite d'entiers ; soit λ'' la feuille de I_1 qui a pour rang r ; posons ici $t(\lambda) = \lambda'' \lambda'$. $t(I_2)$ est l'ensemble des points d'une ramification orientée I'_2 isomorphe à I_2, dont le mot des racines est la suite de feuilles de I_1. La réunion verticale (2.6.2) de I_1 et I'_2 est une ramification d'index I. Définissons une application f de I dans E, par $f(x) = f_1(x)$ si $x \in I_1$, $f(x) = f_2 \circ t^{-1}(x)$ si $x \in t(I_2)$: si $x \in I_1 \cap t(I_2)$, x est une feuille de I_1 et une racine de I'_2 ; comme la suite de feuilles de (I_1, f_1) est la famille des racines de (I_2, f_2), on a bien $f_1(x) = f_2 \circ t^{-1}(x)$. La pseudo-ramification (I, f) s'appelle *réunion verticale* de (I_1, f_1) et (I_2, f_2). On note

$$(I, f) = (I_1, f_1) \underset{V}{\cup} (I_2, f_2).$$

Soit φ_1 le mot des racines de (I_1, f_1) : $\varphi_1 = f^*_1(\Psi_1)$ où Ψ_1 est le mot des racines de I_1. Le mot des racines de I est Ψ_1, et celui de (I, f), $f^*(\Psi_1) = f_1^*(\Psi_1) = \varphi_1$. De même, la suite de feuilles de (I,f) est celle de (I_2, f_2).

Soit (I, f) la réunion horizontale (resp. verticale) de deux pseudo-ramifications (I_1, f_1) et (I_2, f_2) : une sous-pseudo-arborescence complète de (I_2, f_2) est l'image d'une sous-arborescence complète J de I_2 par la restriction $\bar{f}_2$ de f_2 à J. Désignons par $\bar{t}$ et $\bar{f}$ les restrictions respectives de t et f à J et à t(J) ; $\bar{t}(J)$ est une sous-arborescence complète de I'_2 et $((J, \bar{f}_2)) = ((\bar{t}(J), f_2 \circ \bar{t}^{-1})) = ((t(J), \bar{f}))$. $\bar{t}(J)$ est une sous-arborescence complète de I (1. 4. 2 et 2. 6. 2); donc $((J, \bar{f}_2))$ est une sous-pseudo-arborescence complète de (I, f). Il est immédiat de montrer que toute sous-pseudo-arborescence complète de (I_1, f_1) est une sous-pseudo-arborescence complète de (I, f).

Théorème 3. 2. *Soient deux pseudo-ramifications dans un même ensemble. Toute sous-pseudo-arborescence complète de l'une d'elles est une sous-pseudo-arborescence complète de leur réunion horizontale, et de leur réunion verticale lorsque celle-ci existe.*

3. 2. 3. Pseudo-ramification gauche d'une pseudo-ramification. On appelle *pseudo-ramification gauche* d'une pseudo-ramification (I, f) l'image par f de la ramification gauche I_G de I (2. 12).

I_G est isomorphe à sa ramification d'index I' : $I' = h(I_G)$; la pseudo-ramification gauche de (I, f) est $(I', f \circ h^{-1})$. I' est la ramification gauche de h(I), et (I, f) est l'image de h(I) par $f \circ h^{-1}$: $(I', f \circ h^{-1})$ est la pseudo-ramification gauche de l'image par $f \circ h^{-1}$ d'une ramification orientée dont I' est ramification gauche.

Réciproquement, soit une pseudo-ramification (I', f') telle que I' soit la ramification gauche d'une ramification orientée I_1. I_1 est isomorphe à sa ramification d'index I : $I_1 = h(I)$; la ramification gauche de I est $h^{-1}(I')$. La pseudo-ramification gauche de $(I, f' \circ h)$ est $((h^{-1}(I'), f' \circ h)) = (I', f')$.

Les pseudo-ramifications dont une pseudo-ramification (I', f') est pseudo-ramification gauche sont les images par f' des ramifications orientées dont I' est la ramification gauche. On déduit alors du théorème 2. 10 :

Théorème 3. 3. *Soit un ensemble E, E_0 un sous-ensemble de E, $\mathcal{C}_0$ l'ensemble des pseudo-ramifications (I, f) dans E telles que f transforme les feuilles de I en des éléments de E_0, les noeuds de I en des éléments du complémentaire de E_0. Pour qu'il existe une pseudo-ramification (I, f) $\in \mathcal{C}_0$ dont une pseudo-ramification dans E, (I', f'), donnée soit la pseudo-ramification gauche, il faut et il suffit que tout point x de I' tel que $f'(x) \in E_0$ soit une feuille et que chaque famille de I', autre que celle des racines, commence par un tel point.* (I, f) *est alors unique.*

En effet, les pseudo-ramifications dont (I', f') est pseudo-ramification gauche sont les images par f' des I_1 dont I' est ramification gauche ; leur imposer d'appartenir à $\mathcal{E}_0$ revient à déterminer les feuilles de I_1 : les points x tels que $f'(x) \in E_0$.

La pseudo-ramification gauche de la transformée d'une pseudo-ramification (I, f) par une application g est la transformée par g de la pseudo-ramification gauche de (I, f).

Dans le cas où (I, f) est une pseudo-arborescence, il en est de même pour sa pseudo-ramification gauche : nous parlerons alors de la *pseudo-arborescence gauche* de (I, f).

3. 3. Piles attachée, adjointe, conjointe à une pseudo-ramification.

On appelle pile *attachée* (resp. *adjointe, conjointe*) à une pseudo-ramification (I, f) la pile transcrite par f de la pile strictement attachée (resp. adjointe, conjointe) à la ramification orientée I.

Soit une pile U. D'après le théorème 3. 1, elle est transcrite d'une pile simple U_0, définie à un isomorphisme près. Chaque pile U_0 est strictement attachée (resp. adjointe, conjointe) à une ramification orientée, et aux diverses piles U_0 toutes isomorphes correspondent des ramifications orientées isomorphes (théorèmes 2. 1, 2. 6, 2. 7), qui ont toutes la même ramification d'index I, qui leur est isomorphe. U est transcrite de la pile strictement attachée (resp. adjointe, conjointe) à la ramification I, et la transcription f est unique (théorème 3. 1).

Théorème 3. 4. *Toute pile sur un ensemble* E *est attachée (resp. adjointe, conjointe) à une pseudo-ramification sur* E *et une seule.*

Exemple : pseudo-ramification représentée figure 3. 2.

Matrice d'enchaînement (dans l'ordre d'entrée) : 1 indique la présence d'un lien vertical, 0 son absence.

n° de ligne	élément	lien vertical	lien horizontal
1	D	1	-
2	A	0	3
3	B	1	-
4	C	0	5
5	B	1	-
6	A	0	7
7	C	0	8
8	A	0	-

Pile attachée :

```
              A   C   A
         C   B B B B B B B
    A   B B B B B B B B B B B
  D D D D D D D D D D D D D D D
```

Pile adjointe :

```
                          A
        A       C       C C C
  D    B B B   B B B   A A A A A
   Λ         Λ       Λ
```

Pile conjointe :

```
         A
       C C C
     A A A A A   B
   C C C C C C C C C   B
  A A A A A A A A A A A A A   D
                             Λ
```

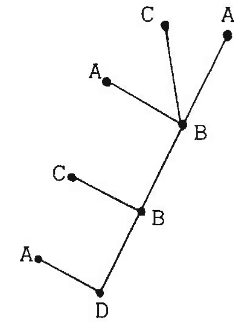


Figure 3. 2.

Une pseudo-ramification est donc déterminée par sa pile attachée, sa pile adjointe ou sa pile conjointe, c'est-à-dire par leur mot attaché ou leur mot de description ou leur mot de description incomplet ou le couple de leur mot d'entrée et de leur mot direct des poids ou le couple de leur mot de sortie et de leur mot inverse des poids.

Soit un ensemble E et une application p de E dans l'ensemble des entiers naturels. Envisageons l'ensemble des pseudo-ramifications dans E où toute famille de prédécesseur x ait pour longueur p(x), et dont chaque feuille y vérifie $p(y) = 0$. Dans cet ensemble, une pseudo-ramification est définie par le mot d'entrée (resp. de sortie) de sa pile attachée, qui détermine le mot direct (resp. inverse) des poids : on peut appeler ce mot d'entrée (resp. de sortie), *mot de Lukasiewicz direct* (resp. *inverse*) de la pseudo-ramification. Les mots de Lukasiewicz directs des pseudo-ramifications de l'ensemble forment un langage préfixé au sens de [27].

CHAPITRE IV

STRUCTURES, GRAMMAIRES ET LANGAGES DE CHOMSKY

4. 1. Définitions.

4. 1. 1. Structures de Chomsky. Soit un ensemble fini V et une relation binaire entre V et le monoïde libre V^* déduit de V : nous noterons ::= cette relation. Nous appelons *structure de Chomsky* $\mathcal{S}$ (V, ::=) l'ensemble des pseudo-ramifications dans V telles que, pour toute famille φ de prédécesseur A, A::= φ. Nous disons que $\lambda \in V^*$ *dérive* de $\lambda' \in V^*$ lorsque $\lambda = \lambda' = \Lambda$ ou lorsqu'il existe dans $\mathcal{S}$ (V, ::=) une pseudo-ramification dont λ est la suite de feuilles et λ' la famille des racines ; si la hauteur de cette pseudo-ramification est supérieure à 1, λ *dérive strictement* de λ'. λ dérive de λ' sera noté $\lambda' \xrightarrow{*} \lambda$; une pseudo-ramification dont la famille des racines est λ' et la suite de feuilles λ sera dite *de type* $(\lambda' \xrightarrow{*} \lambda)$. Si $\mathcal{R}'$ est une pseudo-ramification de type $(\lambda' \xrightarrow{*} \lambda)$ et $\mathcal{R}''$ une pseudo-ramification de type $(\lambda \xrightarrow{*} \lambda'')$, la réunion verticale de $\mathcal{R}'$ et $\mathcal{R}''$ est de type $(\lambda' \xrightarrow{*} \lambda'')$, d'après 3. 2. 2. Si $\mathcal{R}_1$ est de type $(\lambda'_1 \xrightarrow{*} \lambda_1)$ et $\mathcal{R}_2$ de type $(\lambda'_2 \xrightarrow{*} \lambda_2)$ la réunion horizontale de $\mathcal{R}_1$ et $\mathcal{R}_2$ est de type $(\lambda'_1 \lambda'_2 \xrightarrow{*} \lambda_1 \lambda_2)$, d'après 3. 2. 2. Dans le monoïde libre V^*, la relation $\xrightarrow{*}$ est un préordre, compatible avec la concaténation.

Nous supposerons que le nombre de couples qui sont dans la relation ::= est fini et non nul (voir généralisation en 4. 6). Alors ::= peut être donnée par énumération des couples en relation, A::= φ ; nous dirons que A::= φ *est la règle de premier membre* A, *de second membre* φ. On abrège l'écriture A::= φ et A::= φ' en A::= φ | φ' [1].

Dans la suite nous supposerons que V est donné comme réunion de deux ensembles disjoints T et N, N contenant tout premier membre de règle, c'est-à-dire tout prédécesseur d'une famille dans une pseudo-ramification de $\mathcal{S}$(V, ::=). Nous appellerons alors *proposition* de la structure tout mot du monoïde libre T^* qui dérive d'un élément de N.

4. 1. 2. Grammaires et langages de Chomsky. Soit $X \in N$. L'ensemble des propositions qui dérivent de X est le *langage de Chomsky engendré par la grammaire* (N, T, ::=, X). Ses éléments sont les *phrases* du langage : ce sont les

suites de feuilles des pseudo-arborescences de $\mathcal{S}$ (V, ::=) qui ont X pour racine et dont les feuilles appartiennent à T ; on notera $\mathcal{C}$ (N, T, ::=, X) l'ensemble de ces pseudo-arborescences. V est le *vocabulaire*, T le *vocabulaire terminal*, les éléments de T les *symboles terminaux*, N le *vocabulaire non terminal* ou *auxiliaire*, ses éléments les *symboles non terminaux* ou *auxiliaires*. ::= la *relation de production*, X *l'axiome* de la grammaire. Deux grammaires qui engendrent le même langage sont dites *équivalentes*.

Si on peut trouver deux mots λ et λ' de V^* tels qu'il existe plusieurs pseudo-arborescences de la structure de Chomsky $\mathcal{S}$ (V, ::=) qui ont λ pour suite de feuilles et λ' pour famille des racines, la *structure* est dite *ambiguë*. Si on peut trouver un mot λ de T^* qui soit suite de feuilles de plusieurs pseudo-arborescences ayant pour racine X, on dit que la *grammaire* (N, T, ::=, X) est *ambiguë*. La structure d'une grammaire ambiguë est ambiguë.

Soit une pseudo-arborescence d'une structure de Chomsky, qui définit une proposition α = '$\alpha_0 \ldots \alpha_p$'. Elle est l'image par une application f d'une arborescence I. Pour tout noeud x de I, on a montré (2. 6. 1) que les feuilles qui majorent x forment un intervalle r(x) de la suite de feuilles de I ; f transcrit cet intervalle en un sous-mot β (x) de α, qui est une proposition de la structure : on dit que β (x) est une *sous-proposition* de α. Les intervalles r(x) et les feuilles de I forment une arborescence sous-réduite de I (1. 5. 3), orientée grâce à l'orientation de I (2. 4. 3) ; l'image de cette arborescence par l'application qui transforme r(x) en β (x) et transforme les feuilles de I comme f sera nommée *pseudo-arborescence des sous-propositions* de α. Une proposition de la structure peut posséder plusieurs pseudo-arborescences des sous-propositions ; la structure est alors ambiguë : on dit qu'elle est *fortement ambiguë*. De même une grammaire qui engendre une phrase possédant plusieurs pseudo-arborescences des sous-propositions est dite fortement ambiguë.

Exemple (cf. instructions d'affectation Algol, [43], [4]) :

N = {A, E, P, F, B, C, G, L}   T = {:=, +, ×, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, a, b}

A ::= G:=E	E ::= E+P | P	P ::= P×F | F
F ::= G | B	B ::= BC | C	G ::= GL | L
C ::= 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9		L ::= a | b

Axiome A.

b:=a×bb+42 est une phrase, définie par la pseudo-arborescence qui est représentée figure 4. 1. La figure 4 2 schématise la pseudo-arborescence des sous-propositions de cette phrase. a×bb+42, a×bb, bb, a sont des propositions de la structure, + n'en est pas une.

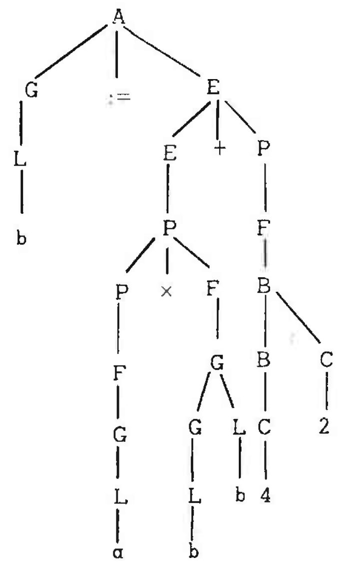


Figure 4. 1

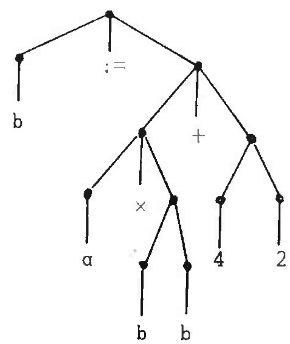


Figure 4. 2

4. 1. 3. Comparaison à d'autres définitions. La définition des langages de Chomsky qui vient d'être donnée est équivalente à celle des langages du type 2 de [12] ou des langages "context-free" de [13]. En effet, d'après les propriétés de la relation "dérive de", s'il existe une Ψ -dérivation d'un mot ω au sens de [12], ω dérive de Ψ. Réciproquement, supposons que ω dérive de Ψ, c'est-à-dire qu'il existe une pseudo-ramification (I, f) de la structure qui soit de type ($\Psi \xrightarrow{*} \omega$) ; envisageons la suite des états u'_j de la pile conjointe à I tels que j soit entrée d'un noeud ou j=0 : $u'_{j_0} = u'_0 = \Lambda$, $u'_{j_1}, \dots, u'_{j_p}$: on passe de u'_{j_i} à $u'_{j_{i+1}}$ en faisant entrer un certain nombre (éventuellement nul) de feuilles, puis en remplaçant une famille par son prédécesseur (2.10.2) ; associons à chaque u'_{j_i} la suite θ'_{j_i} des feuilles dont l'entrée est supérieure à j_i ; f transcrit u'_{j_i} et θ'_{j_i} en u_{j_i} et θ_{j_i} : pour i=0,.., p-1, il existe un symbole non terminal A_i et trois mots φ_i, λ_i, λ'_i de V* tels que $u_{j_i}\,\theta_{j_i} = \lambda_i\,\varphi_i\,\lambda'_i$, $u_{j_{i+1}}\,\theta_{j_{i+1}} = \lambda_i\,A_i\,\lambda'_i$ et $A_i ::= \varphi_i$; d'autre part, $u_{j_0} = \Lambda$, $\theta_{j_0} = \omega$, $u_{j_p} = \Psi$, $\theta_{j_p} = \Lambda$: la suite des $u_{j_i}\,\theta_{j_i}$ est, dans l'ordre inverse, une Ψ -dérivation de ω. Cette Ψ -dérivation présente la particularité que les λ'_i sont formés de feuilles : on "réécrit" à chaque étape le dernier noeud. On aurait pu aussi utiliser la pile adjointe à (I, f), et obtenir une dérivation où on réécrive à chaque étape le premier noeud.

[13] montre que les langages de Chomsky sont les langages "acceptés par un automate à mémoire pile" : la pile utilisée pour engendrer une phrase α est ici, à peu de chose près, la pile attachée à une pseudo-arborescence de type ($X \xrightarrow{*} \alpha$).

Les langages SPL de [2] sont les langages de Chomsky et leurs réunions avec le mot vide. On remarquera qu'avec la définition que nous avons donnée, l'adjonction ou la suppression de règles $A ::= \Lambda$ ne modifie pas une structure de Chomsky, car Λ ne peut être une famille d'une pseudo-ramification. Dans la suite, on supposera que la relation de production ne possède aucune règle de cette forme.

4. 1. 4. Définition de langages de Chomsky par un système d'équations. Soit Q(A) le langage des propositions qui dérivent d'un symbole auxiliaire A donné. Q(A) est formé des mots de T* qui peuvent s'écrire $\alpha_0\,\beta_1\,\alpha_1 \dots \beta_p\,\alpha_p$ où : $p \geqslant 0$, $\alpha_0 \in T^*$; pour $1 \leqslant i \leqslant p$, $\alpha_i \in T^*$, $\beta_i \in T^*$, β_i dérive de $B_i \in N$; $A ::= \alpha_0\,B_1\,\alpha_1 \dots B_p\,\alpha_p$. Q(A) est la réunion des sous-ensembles du monoïde libre T^*, $\alpha_0\,Q(B_1)\,\alpha_1 \dots Q(B_p)\,\alpha_p$ tels que les α_i et les B_i satisfassent aux hypothèses précèdentes. Précisons cet énoncé : toute application Θ de N dans l'ensemble $\mathfrak{P}(T^*)$ des parties de T^*, prolongée en une application Θ' de V dans $\mathfrak{P}(T^*)$ par $\Theta'(a) = \{a\}$ pour tout a de T, induit un homomorphisme de monoïde de V* dans $\mathfrak{P}(T^*)$ (1. 3), que nous notons Θ'^* ; avec cette notation, pour tout $A \in N$, Q(A) est la réunion des ensembles $Q'^*(\varphi)$ tels que $A ::= \varphi$. Ce résultat peut être complété en :

Théorème 4. 1 [54]. *Soient* T *et* N *les vocabulaires terminal et auxiliaire d'une structure de Chomsky* $\mathcal{S}(T \cup N, ::=)$ *telle qu'aucun couple de symboles auxiliaires* A, A' *ne vérifie* $A ::= A'$. *Il existe une, et une seule, application* Θ *de* N *dans* $\mathfrak{P}(T^*)$ *pour laquelle*

$$(\forall A \in N) \quad (\Lambda \notin \Theta(A) \text{ et } \Theta(A) = \bigcup_{A ::= \varphi} \Theta'^*(\varphi)),$$

l'application Q *qui transforme tout* A *de* N *en le langage des propositions dérivant de* A.

Exemple : $T = \{a, b, c\}$; l'unique solution dans $\mathfrak{P}(T^*)^2$ du système de deux équations

$$\left|\begin{array}{l} X = Y \cup Xc \\ Y = \{ab\} \cup aYb \end{array}\right.$$

est formée des deux ensembles des propositions de la structure de Chomsky $\mathcal{S}(T \cup N, ::=)$ qui dérivent respectivement de X et Y, avec :

$N = \{X, Y\}$; $T = \{a, b, c\}$; $X ::= Y \mid Xc$; $Y ::= ab \mid aYb$.

Il s'agit des ensembles $\{a^n b^n c^p ; n > 0, p \geqslant 0\}$, $\{a^n b^n ; n > 0\}$.

4. 1. 5. Image homomorphe et transcription d'un langage de Chomsky. Soit L un langage de Chomsky, engendré par une grammaire (N, T, ::=, X), et f une application de T dans le monoïde libre T'^* déduit d'un ensemble fini T' : f induit un homomorphisme f^* de T^* dans T'^* Prolongeons f en une application f_1 de $T \cup N$ dans $T'^* \cup N$, par l'identité dans N. f^* transforme L en le langage de Chomsky $f^*(L)$ engendré par la grammaire (N, T', ::=', X), où la relation de production ::=' est définie par les règles A ::=' $f_1^*(\varphi)$ si A ::= φ. En particulier, si f est une application de T dans T' (identifié à un sous-ensemble de T'^*), f^* est une transcription : nous dirons que f *transcrit* L en $f^*(L)$.

4. 2. Grammaires réduites. (d'après [2]).

Une grammaire (N, T, ::=, X) *qui engendre un langage de Chomsky non vide est dite* réduite *lorsque*

a) *pour tout* $A \in V$, *il existe* φ_1 *et* $\varphi_2 \in V^*$ *tels que* $\varphi_1 A \varphi_2$ *dérive de* X ;

et b) *pour tout symbole non terminal* A, *il existe une proposition qui dérive de* A.

Soit alors $\lambda \in V^*$ qui dérive de $A \in V$; il existe, dans V^*, φ_1 et φ_2 tels que $\varphi_1 A \varphi_2$ dérive de X ; des symboles non terminaux qui possèdent une occurrence dans φ_1, λ et φ_2 dérivent des propositions ; de φ_1, λ, φ_2 dérivent des mots sur T, α_1, β, α_2. Il existe donc des pseudo-ramifications $\mathcal{R}_1$ de type $(X \xrightarrow{*} \varphi_1 A \varphi_2)$, $\mathcal{R}_2$ de type $(A \xrightarrow{*} \lambda)$, $\mathcal{R}_3$ et $\mathcal{R}_4$ de types $(\varphi_1 \xrightarrow{*} \alpha_1)$, $(\varphi_2 \xrightarrow{*} \alpha_2)$, $\mathcal{R}_5$ de type $(\lambda \xrightarrow{*} \beta)$. La pseudo-ramification

$$\mathcal{R}_1 \underset{V}{\cup} [\mathcal{R}_3 \underset{H}{\cup} (\mathcal{R}_2 \underset{V}{\cup} \mathcal{R}_5) \underset{H}{\cup} \mathcal{R}_4]$$

est de type $(X \xrightarrow{*} \alpha_1 \beta \alpha_2)$ et $\mathcal{R}_2$ en est une sous-arborescence complète (théorème 3. 2).

Théorème 4. 2. *Si la grammaire* (N, T, ::=, X) *est réduite, toute pseudo-arborescence de la structure de Chomsky* $\mathcal{S}(T \cup N, ::=)$ *est une sous-pseudo-arborescence complète d'une pseudo-arborescence de* $\mathcal{C}$ (N, T, ::=, X). *De plus deux pseudo-arborescences de même racine qui ont deux suites de feuilles terminales* β *et* β' *sont sous-pseudo-arborescences complètes de deux pseudo-arborescences de* $\mathcal{C}$ (N, T, ::=, X), *dont les suites de feuilles* $\alpha_1 \beta \alpha_2$ *et* $\alpha'_1 \beta' \alpha'_2$ *vérifient* $\alpha_1 = \alpha'_1$ *et* $\alpha_2 = \alpha'_2$.

[2] donne un algorithme qui permet, pour toute grammaire G, de décider si elle engendre un langage de Chomsky vide et, sinon, de construire une grammaire réduite G' telle que $\mathcal{C}(G) = \mathcal{C}(G')$:

1) On cherche les symboles non terminaux A dont dérivent des propositions ; en particulier, on reconnaît si X est un tel symbole, c'est-à-dire si le langage n'est pas vide ; dans ce cas, on écarte les symboles auxiliaires dont ne dérive aucune proposition et les règles contenant de tels symboles ; d'où une grammaire (N', T, ::=', X) ; soit $V' = N' \cup T$.

2) Pour réaliser la condition a, définissons dans V' la relation ρ :

$$Z \rho Y \iff (\exists \varphi_1 \in V'^*) (\exists \varphi_2 \in V'^*) (Z ::=' \varphi_1 Y \varphi_2).$$

On construit la fermeture transitive R de ρ. Les règles qui ne contiennent que des symboles de l'ensemble R(X) définissent une relation ::=". La grammaire $(R(X) \cap N', R(X) \cap T, ::=", X)$ répond à la question.

4. 3. Initiales, finales, couples vicinaux.

4. 3. 1. Soit une structure de Chomsky $\mathcal{S}$ (V, ::=) et un élément Y de V. On dit que $Z \in V$ est une *initiale* de Y s'il existe un mot φ de V^* tel que $Z\varphi$ dérive de Y. On dit que $Z' \in V$ est une *finale* de Y s'il existe un mot φ' de V^* tel que $\varphi' Z'$ dérive de Y.

Il existe alors dans $\mathcal{S}$ (V, ::=) une pseudo-arborescence qui a Y pour racine et $Z\varphi$ (resp. $\varphi' Z'$) pour suite de feuilles ; si l'une de ces pseudo-arborescences a une hauteur supérieure à 1, ce qui est toujours le cas pour $Y \neq Z$ (resp. $Y \neq Z'$), on dit que Z (resp. Z') est une *initiale (resp. finale) stricte* de Y.

La relation "Z est une initiale de Y" est la fermeture transitive de la relation in définie dans V par :

$$Z \text{ in } Y \iff (\exists \psi \in V^*) (Y ::= Z \psi).$$

"Z est une finale de Y" est la fermeture transitive de fn :

$$Z \text{ fn } Y \iff (\exists \psi \in V^*) (Y ::= \psi Z).$$

Le graphe (V, in) sera nommé *graphe des initiales* de la grammaire.

Soit (I, f) une pseudo-ramification de $\mathcal{S}$ (V, ::=). Etudions le couple formé par un point x de I et la première feuille y qui majore x dans l'ordre d'entrée de I. D'après le paragraphe 2. 3. 3, f(y) est une initiale de f(x). Inversement, soit (N, T, ::=, X) une grammaire réduite, A un symbole non terminal, a une initiale terminale de A. Il existe dans $\mathcal{S}$ ($N \cup T$, ::=) une pseudo-arborescence (I, f) de racine A et de première feuille a ; nommons x la racine de I, y sa première feuille : f(x) = A, f(y) = a. (I, f) est une sous-pseudo-arborescence complète d'une pseudo-arborescence (I', f') de $\mathcal{C}$ (N, T, ::=, X) ; dans l'arborescence orientée I', tout point, autre que x, du chemin joignant x à y est le premier de sa famille : y est la première feuille qui majore x pour l'ordre d'entrée de I'.

4. 3. 2. Envisageons maintenant une pseudo-ramification (I, f) de $\mathcal{S}$ (V, ::=), y, z un couple de points consécutifs de la ramification orientée I, Y = f(y) et Z = f(z) ; une telle situation se présente en particulier lorsqu'il existe deux mots φ_1 et φ_2 de V^* tels que $\varphi_1 Y Z \varphi_2$ dérive d'un symbole auxiliaire. D'après le paragraphe 2. 6. 3, on peut trouver trois symboles A, Y', Z' de V et deux mots λ et λ' de V^* tels que Y soit une finale de Y', Z une initiale de Z' et A ::= λ Y' Z' λ'. Réciproquement, soit une règle A ::= λ Y' Z' λ', Y une finale de Y' et Z une initiale de Z'. Il existe un mot $\lambda \varphi$ Y Z $\varphi' \lambda'$ qui dérive de A, donc une pseudo-arborescence (I, f) de $\mathcal{S}$ (V, ::=) qui a A pour racine et deux points consécutifs y, z de I tels que Y = f(y), Z = f(z). Si, de plus, (N, T, ::=, X) est une grammaire réduite, (I, f) est une sous-pseudo-arborescence complète de (I', f') appartenant à $\mathcal{C}$ (N, T, ::=, X), et y et z sont aussi consécutifs dans I' (2. 6. 3) ; d'autre part, il existe φ_1 et φ_2 dans V* tels que φ_1 Y Z φ_2 dérive de X.

Théorème 4. 3. *Soit $\mathcal{S}$ (V, ::=) une structure de Chomsky, Y et Z deux éléments de V. Les propositions suivantes sont équivalentes :*

a) *il existe dans $\mathcal{S}$ (V, ::=) une pseudo-ramification (I, f) et un couple de points consécutifs y, z de I tels que Y = f(y), Z = f(z) ;*

b) *il existe un symbole non terminal A et deux mots φ_1, φ_2 de V* tels que φ_1 Y Z φ_2 dérive de A ;*

c) *il existe trois symboles A, Y', Z' de V et deux mots λ, λ' de V* tels que Y soit une finale de Y', Z une initiale de Z', et A ::= λ Y' Z' λ'.*

Si (N, T, ::=, X) est une grammaire réduite et V = N ∪ T, les propositions suivantes sont aussi équivalentes à a, b, c :

d) *il existe dans $\mathcal{C}$ (N, T, ::=, X) une pseudo-arborescence (I, f) et un couple de points consécutifs y, z de I tels que Y = f(y), Z = f(z) ;*

e) *il existe deux mots φ_1, φ_2 de V* tels que φ_1 Y Z φ_2 dérive de X.*

Définition : Le couple Y, Z est appelé *couple vicinal* lorsque les propositions a, b, c sont vraies.

Exemples : pour la grammaire donnée en exemple au paragraphe 4. 1. 2 : L est une initiale de G et A, b est une initiale de L et E, B est une finale de A, b une finale de L, L une finale de G, G une finale de P ; L, L est un couple vicinal, ainsi que L, b et G, x.

4. 4. Représentation d'une relation de production par une pseudo-ramification.

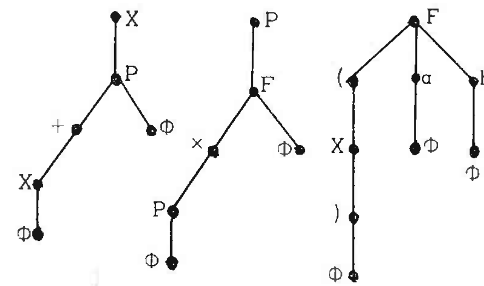
Chaque règle A ::= φ est connue lorsqu'on connaît le mot A φ . Ces mots, en nombre fini, peuvent être donnés par une pseudo-ramification (I, f) (3. 2. 1, exemple) ; pour les distinguer de leurs facteurs gauches, on pourra remplacer A φ par un mot A φ Φ, où Φ n'appartient pas à V : les mots A φ Φ tels que A ::= φ sont les mots transcrits par f des chemins de I dont l'origine est une racine et dont l'extrémité a Φ pour image.

La pseudo-ramification peut être représentée par une matrice d'enchaînement. Seuls les points d'image Φ ne possèdent pas de lien vertical. Aussi, si l'on emploie la matrice d'enchaînement où les points sont dans l'ordre d'entrée, la colonne qui sert à indiquer l'existence du lien vertical (2. 1. 2) peut

être omise. Chaque symbole non terminal peut être représenté par le numéro de la ligne où il se trouve en tant que racine ou, pour la plupart des utilisations, par le numéro de la ligne suivante (celle du lien vertical) ; dans ce cas, il n'y a en général pas d'inconvénient à omettre les lignes qui correspondent aux racines.

Exemple : N = {X, P, F} , T = {+, ×, a, b, (,)}

X ::= P + X | P   P ::= F × P | F   F ::= (X) | a | b



Dans la matrice d'enchaînement, comprimée comme il vient d'être dit, Φ est représentée par 0, et les symboles terminaux par des entiers négatifs :

+	×	a	b	(	)
-1	-2	-3	-4	-5	-6

Figure 4. 3.

n° de ligne	n° d'élément	lien horizontal
1	6	-
2	-1	5
3	1	-
4	0	-
5	0	-
6	11	-
7	-2	10
8	6	-
9	0	-
10	0	-
11	-5	15
12	1	-
13	-6	-
14	0	-
15	-3	17
16	0	-
17	-4	-
18	0	-

Pour certaines utilisations, on préfère représenter une règle $A ::= \varphi$ par le mot φA ou $\tilde{\varphi} A$ (1). Il est commode de distinguer des symboles composant φ la dernière occurrence de A, en associant biunivoquement à tout symbole auxiliaire A un $\bar{A}$ n'appartenant pas à V, et en représentant $A ::= \varphi$ par le mot $\varphi \bar{A}$ ou $\tilde{\varphi} \bar{A}$; soit $\bar{N}$ l'ensemble des $\bar{A}$. Les mots $\varphi \bar{A}$ ou $\tilde{\varphi}\bar{A}$ peuvent être donnés par une pseudo-ramification (I, f) : ce sont les mots transcrits par f des chemins de I dont l'origine est une racine et dont l'image de l'extrémité appartient à $\bar{N}$. Comme plus haut, dans la matrice d'enchaînement où les points sont classés dans l'ordre d'entrée, la colonne du lien vertical peut être omise. Des exemples seront donnés aux paragraphes 6. 4. 5 et 7. 5. 4.

Nous reviendrons sur ces représentations d'une relation de production au paragraphe 4. 6. 4.

4. 5. Structures et langages de Kleene.

4. 5. 1. Définitions. Une structure de Chomsky dont toutes les règles sont du type $A ::= aB$ (resp. $A ::= Ba$) ou $A ::= a$, où A et B sont des symboles non terminaux et a un symbole terminal, s'appelle *structure droite (resp. gauche) de Kleene*. Une grammaire de Chomsky dont la structure est une structure droite (resp. gauche) de Kleene est une *grammaire droite (resp. gauche) de Kleene*; un langage engendré par une grammaire de Kleene est un *langage de Kleene*.

4. 5. 2. Langages de Kleene et chemins de pseudo-graphes. Soit un graphe fini (E, Γ). Ses chemins sont des mots du monoïde libre E^*. Etudions l'ensemble $c(x)$ des chemins dont l'origine x est fixée et dont l'extrémité appartient à un sous-ensemble E' de E : si x est dans E', x appartient à $c(x)$; les autres chemins de $c(x)$ sont les $x\lambda$ où $\lambda \in c(y)$ et $x \Gamma y$. D'après le théorème 4. 1, $c(x)$ est un ensemble de propositions de la structure de Kleene définie de la manière suivante : E est le vocabulaire terminal, le vocabulaire non terminal N a autant d'éléments que E ; il est obtenu en associant biunivoquement un symbole $h(x)$ à tout x de E ; la relation de production est donnée par les règles : $h(x) ::= x$ pour $x \in E'$, $h(x) ::= x h(y)$ pour tout couple (x, y) tel que $x \Gamma y$. $c(x)$ est le langage de Kleene engendré par la grammaire $(N, E, ::=, h(x))$. Comme la réunion d'un nombre fini de langages de Kleene est un langage de Kleene [14], les chemins de (E, Γ) dont l'origine est dans un ensemble E'' et l'extrémité dans un ensemble E' forment un langage de Kleene ; il en est ainsi en particulier pour ceux qui ont une extrémité donnée et une origine dans E'' : un raisonnement direct analogue au précédent aurait conduit ici à une structure gauche de Kleene.

(1) Rappelons que $\tilde{\varphi}$ est le mot réfléchi de φ (1. 3).

Soit f une application de E dans un ensemble E_1 ; $(E_1, \Gamma ; f)$ est un pseudo-graphe (3. 2. 1). Appelons *chemin du pseudo-graphe, d'origine x et d'extrémité y*, tout mot transcrit par f d'un chemin du graphe (E, Γ) ayant x pour origine et y pour extrémité. Si E' et E'' sont deux sous-ensembles de E, les chemins du pseudo-graphe dont l'origine est dans E'' et l'extrémité dans E' forment le langage de Kleene transcrit par f(4.1.5) de celui des chemins de (E, Γ) qui ont leur origine dans E'' et leur extrémité dans E'.

Nous montrerons maintenant que réciproquement tout langage de Kleene peut être considéré comme un tel ensemble de chemins d'un pseudo-graphe. Soit une structure droite de Kleene, T et N ses vocabulaires, terminal et non terminal, $::=$ sa relation de production. Associons lui une autre structure droite, dont le vocabulaire terminal T' et le vocabulaire non terminal N' aient le même nombre d'éléments :

- T' est l'ensemble des couples (a, A) de l'ensemble produit $T \times N$ tels que $A ::= a$ ou qu'il existe $B \in N$ pour lequel $A ::= aB$;
- N' est l'ensemble des couples (A, a), pour $(a, A) \in T'$;
- la relation de production $::='$ est définie par les règles : $(A, a) ::=' (a, A)$ pour toute règle $A ::= a$; $(A, a) ::=' (a, A)\,(B, b)$ pour toute règle $A ::= aB$ et tout $b \in T$ tel que $(B, b) \in N'$.

Désignons par T'' le sous-ensemble de T' formé des couples (a, A) tels que $A ::= a$. D'après l'étude directe, les propositions de cette structure sont les chemins d'un graphe (T', Γ) dont l'extrémité appartient à T'' : la relation Γ est définie dans T' par :

$$(a, A)\,\Gamma\,(b, B) \iff (A, a) ::=' (a, A)\,(B, b).$$

Envisageons l'application f de $T' \cup N'$ dans $T \cup N$:

$$f(a, A) = a \quad \text{et} \quad f(A, a) = A \quad \text{si} \quad (a, A) \in T'.$$

f transforme toute pseudo-arborescence de la structure $\mathcal{S}(T' \cup N', ::=')$, ayant (A, a) pour racine et α pour suite de feuilles, en une pseudo-arborescence de $\mathcal{S}(T \cup N, ::=)$, ayant A pour racine et $f^*(\alpha)$ pour suite de feuilles, et cette transformation est surjective. D'autre part, f transforme le graphe (T', Γ) en un pseudo-graphe. Les propositions de la structure de Kleene donnée sont les chemins de ce pseudo-graphe dont l'extrémité appartient à T''. Les phrases d'un langage de Kleene K engendré par une grammaire $(N, T, ::=, X)$ sont les chemins de ce pseudo-graphe dont l'extrémité appartient à T'' et dont l'origine est l'un des couples de T' se terminant par X : on notera T''' l'ensemble de ces couples.

On ferait une démonstration analogue pour une structure gauche de Kleene.

Théorème 4. 4. *Soit un ensemble fini T. Les ensembles suivants sont égaux :*
- celui des langages de Kleene de vocabulaire terminal T, engendrés par une grammaire gauche,

- *celui des langages de Kleene de vocabulaire terminal T, engendrés par une grammaire droite,*
- *celui des ensembles de chemins des pseudo-graphes finis dans T, dont l'origine et l'extrémité appartiennent à deux ensembles donnés.*

4. 5. 3. Etude du graphe (T', Γ). Chaque phrase du langage K qui a x pour premier élément est transcrite par f d'un chemin généralisé d'origine (x, X) dans T'''. Si, pour tout point (a, A) de T', les points (b, B) de $\Gamma(a, A)$ ont des images b distinctes, ce chemin est unique, car ses points sont déterminés de proche en proche de manière unique. Il suffit pour cela qu'il n'existe pas deux règles A ::= aB et A ::= aB' telles que $B \neq B'$. Tout langage de Kleene peut être engendré par une grammaire droite qui possède cette propriété (et une grammaire gauche qui possède la propriété analogue) ([2], [49]) ; cette grammaire peut être réduite sans perdre la propriété.

Un chemin β quelconque de (T', Γ) se présente sous la forme $(a_0, A_0) \dots (a_p, A_p)$, avec $p \geqslant 0$; si $p > 0$, $A_i ::= a_i A_{i+1}$ pour $0 \leqslant i < p\text{-}1$; $A_p ::= a_p$ ou il existe $B \in N$ tel que $A_p ::= a_p B$: selon le cas, $a_0 \dots a_p$ ou $a_0 \dots a_p B$ dérive de A_0. Supposons la grammaire réduite : A_0 possède une occurrence dans un mot qui dérive de X et il existe une proposition qui dérive de B. On en déduit qu'il existe un chemin de (T', Γ), $(x, X) \dots (a_0, A_0) \dots (a_p, A_p) \dots (c, C)$ tel que C::=c , autrement dit un chemin $\beta_1\ \beta\ \beta_2$ dont l'origine est dans T''' et l'extrémité dans T''. Pour tout chemin φ du pseudo-graphe (T', Γ ; f), il existe deux mots φ_1 et φ_2 sur T tels que $\varphi_1\ \varphi\ \varphi_2$ soit une phrase de K. Alors, les points de T''' sont les seuls points w de T' pour lesquels l'ensemble $\Gamma^{-1}(w)$ peut être vide. Il l'est effectivement si X n'a d'occurrence à droite d'aucune règle. Toute grammaire de Kleene réduite possède une grammaire de Kleene réduite équivalente vérifiant cette propriété : il suffit de remplacer X, dans le second membre des règles, par un nouveau symbole non terminal X', puis de poser X' ::= Ψ lorsque X ::= Ψ. De même, les seuls points w pour lesquels Γ (w) peut être vide sont ceux de T''.

4. 5. 4. Remarque. La caractérisation des langages de Kleene par le théorème 4.4 se rapproche de caractérisations classiques ([14], [40]), que l'étude précédente permet d'ailleurs de retrouver :

1) le graphe (T', Γ) possède la propriété suivante :

(a, A) Γ (b, B) entraîne (a, A) Γ (b', B) pour tout b' tel que $(b', B) \in T'$.

On peut définir un *multigraphe* d'ensemble de points N, d'ensemble de noms d'arêtes T, comme un graphe sur une partie T' de T × N, qui possède cette propriété ; si (a, A) Γ (b, B), on dira que (A, a, B) est une arête de nom a qui lie le point A au point B. Un tel multigraphe (on dit aussi *graphe d'état fini*) définit les transitions d'un automate fini (non déterministe). Au chapitre 5, nous reviendrons sur cette notion d'automate fini et nous énoncerons le théorème qui relie structures de Kleene et automates finis.

2) les chemins élémentaires, c'est-à-dire sans répétition, d'un graphe fini sont en nombre fini. Les circuits élémentaires, c'est-à-dire sans répétition autre que celle de leur origine, sont aussi en nombre fini. On obtient tous les chemins d'origine x et d'extrémité x' en substituant dans les chemins élémentaires d'origine x, d'extrémité x', à

certains éléments y un circuit d'origine y. Les circuits d'origine y sont de la forme $c_1^{p_1}\ c_2^{p_2} \dots c_n^{p_n}$ y où les p_i sont des entiers positifs et les c_i des circuits élémentaires d'origine y privés de leur extrémité. Un langage de Kleene peut ainsi être défini par une "somme d'expressions représentantes" ([14], [13]). Réciproquement, pour toute expression représentante, on peut définir un pseudo-graphe dont elle détermine l'ensemble des chemins qui ont leur origine et leur extrémité dans deux ensembles donnés. Un langage fini est obtenu à partir d'un pseudo-graphe sans circuit : aux paragraphes 3. 2. 1 et 4. 4, nous avons utilisé une pseudo-ramification pour stocker les mots d'un dictionnaire ou les règles d'une relation de production : ces mots sont les chemins de la pseudo-ramification, dont l'origine est une racine et l'extrémité une feuille ; on aurait aussi pu par exemple utiliser un pseudo-graphe ne contenant qu'un seul point d'image Φ.

4. 5. 5. Détermination d'un pseudo-graphe par une matrice d'enchaînement. Un pseudo-graphe fini (T', Γ ; f), image d'un graphe (T', Γ), peut être déterminé par une matrice qui définit à la fois (T', Γ) et f. Si en particulier (T', Γ) possède une matrice d'enchaînement (1. 4. 4), le pseudo-graphe pourra lui aussi être déterminé par une matrice d'enchaînement : il s'agit, comme pour une pseudo-ramification, d'une matrice à trois colonnes qui, pour chaque x de T', donnent respectivement f(x) et les numéros des lignes qui correspondent au lien vertical et au lien horizontal de x, lorsque ces liens existent. On a vu au paragraphe 1. 4. 4 qu'il suffit, pour que (T', Γ) possède une matrice d'enchaînement, que les ensembles Γ (x) soient deux à deux confondus ou disjoints. Nous allons montrer qu'à tout pseudo-graphe fini (T', Γ ; f), qui ne possède pas cette propriété, on peut associer un pseudo-graphe fini, dans le même ensemble E, qui la possède et qui a les mêmes chemins.

Soit M l'ensemble des Γ(x), pour $x \in E'$, Ω et Ω' deux éléments de M, distincts mais non disjoints ; introduisons un ensemble Σ disjoint de T', qui a même cardinal que leur intersection : à tout y de cette intersection, on associe bijectivement un élément y' de Σ ; soit $\Omega_1 = (\Omega \setminus \Omega') \cup \Sigma$. Définissons alors :

$T'_1 = T' \cup \Sigma$; pour tout $w \in E'$ tel que $\Gamma(w) = \Omega$, $\Gamma_1(w) = \Omega_1$; pour les autres points w de T', $\Gamma_1(w) = \Gamma(w)$; pour $y' \in \Sigma$, $\Gamma_1(y') = \Gamma_1(y)$; $f_1(w) = f(w)$ si $w \in T'$, $f_1(y') = f(y)$ si $y' \in \Sigma$.

Parmi les $\Gamma_1(w)$, aucun n'est Ω, tous sauf Ω_1 appartiennent à M ; Ω_1 et Ω' sont disjoints et si $\Omega'' \in M$, $\Omega_1 \cap \Omega'' \neq \phi$ entraîne $\Omega \cap \Omega'' \neq \phi$: le graphe (T'_1, Γ_1) possède moins de couples d'ensembles $\Gamma_1(x)$ distincts et non disjoints que le graphe (T', Γ). D'autre part, les pseudo-graphes $(T'_1, \Gamma_1 ; f_1)$ et (T', Γ ; f) ont les mêmes chemins. Si de plus T'' et T''' sont deux sous-ensembles de T' et si on définit T''_1 (resp. T'''_1) comme la réunion de T'' (resp. T''') et de l'ensemble des points y' de Σ tels que y appartienne à T'' (resp. T'''), les chemins de $(T'_1, \Gamma_1 ; f_1)$ d'extrémité dans T''_1, d'origine dans T'''_1 sont les chemins de (T', Γ ; f) d'extrémité dans T'', d'origine dans T'''. Il suffit d'itérer cette construction.

Cependant, il n'est pas nécessaire que les $\Gamma(x)$ soient disjoints, pour que (E, Γ) ait une matrice d'enchaînement, de sorte que ce résultat peut parfois être obtenu plus simplement.

Exemple : $N = \{X, A, B, C, D, F, H\}$, $T = \{b, c, d, e, h, k\}$, axiome X ;

$X ::= d\ B \mid d\ C$ $\quad$ $B ::= b \mid k \mid b\ D$ $\quad$ $D ::= e\ H$

$H ::= e\ B$ $\quad$ $C ::= c \mid c\ A$ $\quad$ $A ::= b\ F$ $\quad$ $F ::= h\ C$

Le graphe (T', Γ) et le pseudo-graphe $(T', \Gamma\ ;\ f)$ associés sont représentés sur la figure 4. 4, où on a aussi noté l'origine et les extrémités des chemins qui forment le langage. Le graphe (T', Γ) n'a pas de matrice d'enchainement ; en effet :

$\Gamma(d,X) = \{(b,B), (k,B), (c,C)\}$; $\Gamma(e,H) = \{(k,B), (b,B)\}$; $\Gamma(h,F) = \{(c,C)\}$;

pour satisfaire au critère d'existence de la matrice d'enchaînement (1. 4. 4), il faudrait placer (c, C) à la fois au premier et au dernier rang de $\Gamma(d, X)$. On peut y remédier en introduisant un seul nouveau couple (c, C') ; on obtient alors le pseudo-graphe représenté figure 4. 5 avec sa matrice d'enchaînement, qui a les mêmes chemins. Il peut être commode d'introduire un nouveau signe terminal Φ à la fin de toutes les phrases du langage ; pour le pseudo-graphe, cela revient à lier tout point de T'' à un point d'image Φ ; dans la matrice d'enchaînement, Φ peut être repéré par un 0, à placer colonne 2, ligne 4, et colonne 3, lignes 5 et 8.

4. 5. 6. Borne d'intersection de deux langages de Kleene disjoints. Soient K et K_1 deux langages de Kleene de même vocabulaire terminal T. S'il existe un entier naturel q tel qu'aucun mot sur T, de longueur q, ne soit facteur gauche d'une phrase de K et d'une phrase de K_1, nous appellerons *borne d'intersection* de K et K_1 le plus petit de ces entiers q.

Supposons K et K_1 définis par deux grammaires droites (par exemple) de Kleene, dont les vocabulaires terminaux N et N_1 contiennent respectivement n et n_1 éléments. Nous allons montrer que si K et K_1 possèdent une borne d'intersection, elle est au plus égale à $n n_1 + 1$. Par conséquent, l'existence d'une borne d'intersection est décidable.

K est l'ensemble des chemins d'un pseudo-graphe $(T', \Gamma\ ;\ f)$ où $T' \subset T \times N$, d'extrémité dans T'', d'origine dans T''' ; K_1 l'ensemble des chemins d'un pseudo-graphe $(T'_1, \Gamma_1\ ;\ f_1)$ où $T'_1 \subset T \times N_1$, d'extrémité dans T''_1, d'origine dans T'''_1. Supposons que '$\alpha_1 \ldots \alpha_r$' soit facteur gauche d'une phrase de K et d'une phrase de K_1. Il existe un chemin $(\alpha_1, A_1) \ldots (\alpha_r, A_r) \ldots (\alpha_{r+q}, A_{r+q})$ de (T', Γ), d'extrémité dans T'', d'origine dans T''', et un chemin $(\alpha_1, B_1) \ldots (\alpha_r, B_r) \ldots (\alpha'_{r+q'}, B_{r+q'})$ de (T'_1, Γ_1) d'extrémité dans T''_1, d'origine dans T'''_1. Si $r > n n_1$, nombre des éléments de $N \times N_1$, il existe

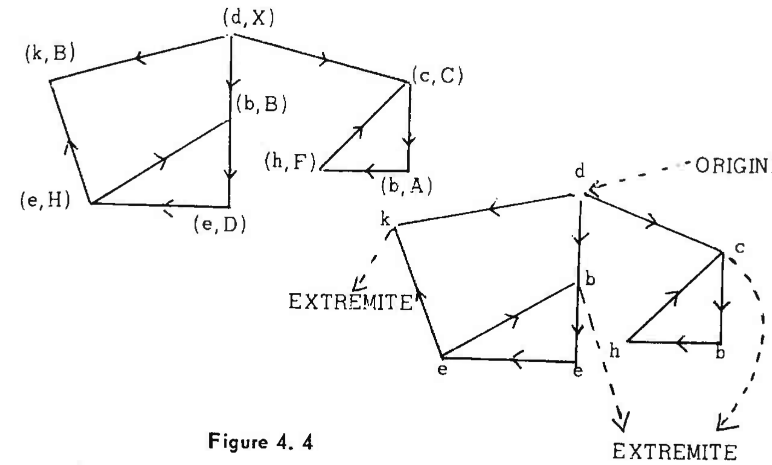
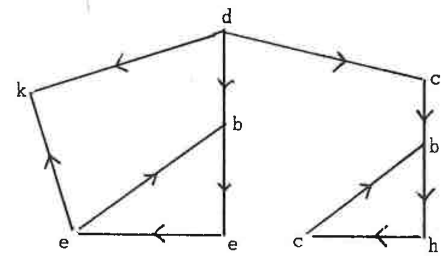


Figure 4. 4



n° ligne	élément	lien vertical	lien horizontal
1	d	2	-
2	c	5	3
3	b	8	4
4	k	-	-
5	b	6	-
6	h	7	-
7	c	5	-
8	e	9	-
9	e	3	-

Figure 4. 5

deux entiers i et j tels que $1 < i < j < r$, $A_i = A_j$, $B_i = B_j$. Quel que soit l'entier naturel k, $(\alpha_1, A_1) \dots [(\alpha_i, A_i) \dots (\alpha_j, A_j)]^k \dots (\alpha_{r+q}, A_{r+q})$ est un chemin de (T', Γ) et $(\alpha_1, B_1) \dots [(\alpha_i, B_i), \dots, (\alpha_j, B_j)]^k \dots (\alpha'_{r+q'}, B_{r+q'})$ un chemin de (T'_1, Γ_1) ; par suite $\alpha_1 \dots (\alpha_i \dots \alpha_j)^k \dots \alpha_r$ est facteur gauche d'une phrase de K et d'une phrase de K_1 : K et K_1 n'ont pas de borne d'intersection.

L'existence d'une borne d'intersection pour deux langages de Kleene K et K_1 contenus dans T^* équivaut à celle d'un sous-ensemble fini Q de T^*, tel que Q contienne un facteur gauche de toute phrase de K et qu'aucun mot de Q ne soit facteur gauche d'une phrase de K_1. Nous dirons que Q est un ensemble *séparant* K *de* K_1.

On peut généraliser la notion de borne d'intersection à p (> 2) langages de Kleene, $K_1, K_2, \dots K_p$, de vocabulaire terminal T : il s'agit du plus petit entier q, s'il existe, tel qu'aucun mot sur T de longueur q ne soit facteur gauche d'une phrase de deux distincts de ces p langages. Pour que $K_1, K_2, \dots, K_p$ aient une borne d'intersection, il faut et il suffit que pour tous les entiers i, j tels que $1 \leqslant i < j \leqslant p$, K_i et K_j aient une borne d'intersection, ou encore que pour tout i ($1 \leqslant i < p$), K_i et la réunion de $K_{i+1}, \dots, K_p$ aient une borne d'intersection.

4. 6. Extension de la définition des structures et grammaires de Chomsky.

4. 6. 1. Nous avons supposé fini le nombre de couples qui sont dans la relation de production ::=. En réalité, cette hypothèse n'est intervenue qu'à deux reprises : pour l'algorithme de construction d'une grammaire réduite (4. 2) et pour la représentation d'une relation de production (4. 4). En particulier, le théorème 4. 4 reste vrai lorsque le nombre de couples en relation est infini. Bornons-nous à supposer que, pour tout symbole non terminal A, l'ensemble L_A des φ tels que $A ::= \varphi$ est décidable : les langages ainsi définis sont décidables car la démonstration donnée par [12] reste valable.

Mais il importe de définir les ensembles L_A. Le plus simple est de prendre pour chacun d'eux un langage de Chomsky. On n'étend pas de cette manière la classe des langages de Chomsky, mais seulement celles des structures et des grammaires. En effet, soit $G = (N, T, ::=, X)$ une grammaire ainsi généralisée. On peut supposer que les langages L_A sont engendrés par des grammaires $G_A = (N_A, T_A, ::=_A, X_A)$ dont les vocabulaires non terminaux N_A sont deux à deux disjoints, et disjoints de $V = N \cup T$; T_A est un sous-ensemble de V. Désignons par N_0 la réunion de N et des N_A et définissons une relation de production $::=_0$ en réunissant les règles des G_A et les règles $A ::=_0 X_A$: on obtient ainsi une grammaire $G_0 = (N_0, T, ::=_0, X)$. Pour toute

dérivation dans G, au sens de [12], entre deux étapes $\delta_i = \Psi_1 A \Psi_2$ et $\delta_{i+1} = \Psi_1 \varphi \Psi_2$ où $\varphi \in L_A$, on peut placer une $\Psi_1 X_A \Psi_2$ -dérivation de $\Psi_1 \varphi \Psi_2$ dans G_0 ; on obtient ainsi une dérivation dans G_0 : le langage engendré par G est contenu dans le langage engendré par G_0. Réciproquement, soit (I, f) une pseudo-arborescence de $\mathfrak{T}(G_0)$, R son ordre associé, et I' la sous-arborescence de I formée des points x tels que $f(x) \in V$: I' a même racine et même suite de feuilles que I. Envisageons un noeud x de I' et $A = f(x)$; dans I, x est prédécesseur d'un seul point y et $f(y) = X_A$; dans I', soit ζ la famille dont x est prédécesseur. La réunion des R-intervalles [y, z] pour z décrivant ζ est une sous-arborescence I_x de I dont y est la racine et ζ la suite de feuilles ; supposons qu'il existe un point u de I_x, tel que f(u) n'appartienne pas à $V_A = N_A \cup T_A$, et choisissons-le minimal pour R parmi les points qui possèdent cette propriété ; u et y sont distincts ; si v est le prédécesseur de u, $f(v) \in V_A$ et $f(v) \notin V$; nécessairement, $f(u) \in V_A$. f transcrit donc ζ en une phrase de L_A. Il en résulte que l'image de I' par une restriction de f appartient à $\mathfrak{T}(G)$, et donc que G et G_0 engendrent le même langage.

4. 6. 2. Cas particulier. Un cas particulier intéressant est celui où chacun des langages L_A est réunion d'un langage de Kleene L'_A contenu dans T^* et d'un nombre fini de langages BL_{AB} obtenus en faisant suivre $B \in N$ des phrases d'un langage de Kleene L_{AB} contenu dans T^*. [54] montre qu'alors la grammaire G engendre un langage de Kleene. Indiquons rapidement comment on peut retrouver ici ce résultat : on engendre chaque L_{AB} par une grammaire gauche de Kleene ; on engendre BL_{AB} en remplaçant par $B\alpha$ le second membre de toute règle de cette grammaire qui se réduit à un symbole terminal α ; les règles de la relation $::=_0$ sont de l'un des types $C ::=_0 D\alpha$, $C ::=_0 \alpha$, $C ::=_0 D$ où C et D sont des symboles auxiliaires et α un symbole terminal. On verra en 7. 10. 2 qu'on peut modifier cette relation de production pour obtenir une grammaire de Kleene équivalente. Ce résultat sera utilisé sous la forme suivante : pour tout symbole auxiliaire X, les mêmes propositions dérivent de X dans $\mathcal{S}(V, ::=)$ et dans une structure gauche de Kleene.

4. 6. 3. La première partie de l'algorithme de construction d'une grammaire réduite équivalente à une grammaire donnée (4. 2, cf [2]) demande seulement qu'on puisse décider, étant donné un ensemble fini E et un symbole non terminal A, si l'intersection de L_A avec le monoïde libre E^* déduit de E est vide. Or E^* est un langage de Kleene ; si les L_A sont des langages de Chomsky, cette propriété est décidable et il en est de même pour la relation ρ introduite dans la seconde partie de l'algorithme.

Sont aussi décidables, les relations in, fn, la relation dans V :

$$(\exists \lambda \in V^*) \quad (\exists \lambda' \in V^*) \quad (\lambda \; Y' \; Z' \; \lambda' \in \bigcup_{A \in N} L_A)$$

et donc la relation : le couple (Y, Z) est vicinal (4. 3, théorème 4. 3, c). Pour la justification de ces propriétés de décidabilité, voir [2].

4. 6. 4. Pseudo-graphes de production lorsque les L_A sont des langages de Kleene. Les règles d'une structure de Chomsky, en nombre fini, sont représentées par des chemins d'un pseudo-graphe, par exemple une pseudo-ramification (3. 2. 1, 4. 4). Plus généralement, si les L_A sont des langages de Kleene, la relation de production peut encore être représentée par un pseudo-graphe ; nous utiliserons trois tels pseudo-graphes, selon qu'il sera commode de représenter $A ::= \varphi$ par l'un des mots $A\varphi$, $\varphi\bar{A}$ ou $\tilde{\varphi}\bar{A}$. Plus précisément, associons biunivoquement à tout symbole auxiliaire A un élément $\bar{A}$ n'appartenant pas à V : soit $\bar{N}$ l'ensemble des $\bar{A}$. Les mots $A\varphi$ (resp. $\varphi\bar{A}$, $\tilde{\varphi}\bar{A}$) où $A \in N$ et $\varphi \in L_A$ forment un langage de Kleene K (réunion de langages de Kleene) : ce sont les chemins d'un pseudo-graphe $(J, \Gamma ; g)$ dans V (resp. $V \cup \bar{N}$), dont l'origine et l'extrémité appartiennent respectivement à deux sous-ensembles J' et J'' de J. Nous dirons que $(J, \Gamma ; g)$ est un *pseudo-graphe de production du premier* (resp. *deuxième, troisième*) *type*.

On a vu (4. 5. 3, 4. 5. 5) que ce pseudo-graphe pouvait être construit de manière à posséder les propriétés suivantes, d'ailleurs vérifiées pour les pseudo-ramifications introduites en 3. 2. 1 (exemple) :
(PG1) Pour tout $Y \in V$, J' contient au plus un point w tel que $g(w) = Y$; ce point, lorsqu'il existe, sera noté $o(Y)$.
(PG2) Si w, y_1 et y_2 sont des points de J tels que $w\Gamma y_1$, $w\Gamma y_2$ et $y_1 \neq y_2$, alors $g(y_1) \neq g(y_2)$.
De PG1 et PG2 résulte : (PG2') il existe au plus un chemin de (J, Γ) d'origine dans J', dont un mot donné sur V (resp. $V \cup \bar{N}$) est transcrit.
(PG3) Pour tout chemin β du graphe (J, Γ), il existe un chemin $\beta_1\,\beta\,\beta_2$ dont l'origine est dans J' et l'extrémité dans J'', c'est-à-dire qui est transcrit en une phrase de K. En particulier, pour les pseudo-graphes du deuxième ou troisième type, si l'image $g(z)$ d'un point z de J appartient à $\bar{N}$, il existe un tel chemin qui contient z, nécessairement comme extrémité, et donc z appartient à J'' ; réciproquement tout point z de J'' est l'extrémité d'un chemin d'origine dans J', donc $g(z)$ appartient à $\bar{N}$. D'où :
(PG3') Pour les pseudo-graphes du deuxième et troisième type, J'' est l'ensemble des points de J dont l'image appartient à $\bar{N}$; si $w \in J''$; $\Gamma(w) = \phi$.
(PG4) J' est l'ensemble des points w de J tels que $\Gamma^{-1}(w)$ soit vide.
(PG5) Le pseudo-graphe peut être déterminé par une matrice d'enchaînement. La transformation éventuellement effectuée pour réaliser PG5 (4. 5. 5) n'altère pas PG2 et PG3, ni PG4 et donc pas PG1 qui ne porte que sur les points de J'.

4. 6. 5. Lorsque, dans la suite, nous parlerons d'une *grammaire généralisée*, il s'agira d'une grammaire de ce type, où les L_A sont des langages de Kleene ; sauf avis contraire, les études qui suivent sont valables pour ces grammaires. Dans toutes les réalisations pratiques, il est souhaitable que le pseudo-graphe de production possède la propriété PG5 ; pour les pseudo-graphes du premier type, il peut être plus commode, à la place de $o(A)$ de connaître la

ligne, dans la matrice d'enchaînement, du premier élément de $\Gamma|o(A)|$ et de retirer de la matrice la ligne de $o(A)$.

Ce type de grammaire généralisée est utilisé par [8], [36], [23].

Une autre généralisation consiste à remplacer l'axiome unique d'une grammaire par un ensemble décidable d'axiomes, comme dans un système logique [17]. Une telle extension est proposée par [20]. Les phrases d'un langage sont alors définies par des pseudo-ramifications qui ne sont plus en général des pseudo-arborescences. Ici encore, si l'ensemble des axiomes est un langage de Chomsky, le langage engendré reste un langage de Chomsky. Nous nous contenterons d'envisager au chapitre 7 le cas où la grammaire possède un ensemble fini d'axiomes contenu dans son vocabulaire auxiliaire.

4. 7. Problème de l'analyse.

Soit $G = (N, T, ::=, X)$ une grammaire de Chomsky, ou une grammaire généralisée. A son propos se posent :
- le problème de la *reconnaissance* : étant donné un mot α sur T, α appartient-il au langage engendré par G ?
- le problème plus général de *l'analyse* : trouver dans $\mathcal{C}(G)$ les pseudo-arborescences dont α est la suite de feuilles.
C'est essentiellement ce dernier problème qui sera étudié dans la suite de ce travail. Nous conserverons les notations G, N, T, ::=, X, α. Nous déterminerons les pseudo-arborescences par leur pile attachée, adjointe ou conjointe (3. 3). La plupart des méthodes d'analyse exposées dans la littérature, notamment celles qui sont séquentielles, c'est-à-dire se contentent d'une seule lecture du mot α, se ramènent à la construction de l'une de ces piles. Nous aurons donc à engendrer les piles attachées, adjointes ou conjointes aux pseudo-arborescences de $\mathcal{C}(G)$ qui ont α pour suite de feuilles, ou à des pseudo-arborescences qui les déterminent. Pour préciser la notion d'algorithme d'analyse, nous introduirons au chapitre 5 la notion de générateur de piles. Puis nous déduirons de l'étude des piles attachée , adjointe, conjointe à une ramification orientée (chapitre 2) divers types de générateurs de piles résolvant le problème de l'analyse.

CHAPITRE 5

GENERATEURS DE PILES

5. 1. Automates finis et transducteurs finis. (cf. [49], [13])

5. 1. 1. Définitions. Une application h d'un ensemble F dans l'ensemble $\mathfrak{P}(F')$ des parties d'un ensemble F' induit une application $\hat{h}$ de $\mathfrak{P}(F)$ dans $\mathfrak{P}(F')$ par : $\hat{h}(G) = \bigcup_{z \in G} h(z)$. Une application k d'un ensemble produit $E \times E'$ dans $\mathfrak{P}(E)$ induit une application $\bar{k}$ de $\mathfrak{P}(E) \times E'$ dans $\mathfrak{P}(E)$ par :

$$\bar{k}(G, y) = \bar{k}(G \times \{y\}) = \bigcup_{x \in E} k(x, y).$$

Définition 1. *Etant donné un ensemble* S *et un ensemble fini* A, une fonction d'automate d'alphabet A et d'ensemble d'états S *est une application* k *du produit cartésien* $S \times A^*$ *dans* $\mathfrak{P}(S)$, *telle que pour tout* s *de* S, *tout* a *de* A *et tout mot* v *non vide de* A^*, $k(s, v\,a) = \bar{k}[k(s, v), 'a']$. *Si l'ensemble* S *est fini,* k *est* une fonction d'automate fini.

Une fonction d'automate d'alphabet A et d'ensemble d'états S est définie par sa restriction au produit de S par l'ensemble des mots de A^* de longueur 1 ou 0. Elle possède les propriétés suivantes :

1) si $s \in S$, si v et v' sont deux mots non vides sur A, par récurrence sur $|v'|$:

$$k(s\ v\,v') = \bar{k}[k(s, v), v'].$$

2) pour $G \subset S \times A^*$ et $v \in A^*$, notons G. v la partie de $S \times A^*$ formée des cou- $(s, v'\,v)$ où $(s, v') \in G$; en revenant à la définition de $\hat{k}$, on voit que :

$$\hat{k}(G.\,v) = \hat{k}[\hat{k}(G) \times \{v\}] = \bar{k}[\hat{k}(G), v] ;$$

on en déduit que si $F \subset S$ et si v et v' sont des mots non vides sur A,

$$\bar{k}(F, v\,v') = \bar{k}[\bar{k}(F, v), v'].$$

Un exemple simple de fonction d'automate est $k(s, v) = \{s\}$ pour tout couple (s, v). Si k et k' sont deux fonctions d'automate de même alphabet A, d'ensembles d'états S et S', l'application k'' de $S \times S' \times A^*$ dans $\mathfrak{P}(S \times S')$ définie par $k''(s, s', v) = k(s, v) \times k'(s, v)$ est une fonction d'automate d'alphabet A, d'ensemble d'états $S \times S'$, appelée *produit* de k et k' et notée $k \times k'$.

Nous utiliserons le résultat classique suivant, sans le redémontrer :

Théorème 5. 1. *Pour toute structure droite (resp. gauche) de Kleene, il existe une fonction d'automate fini* k *dont l'alphabet est le vocabulaire terminal* T, *dont l'ensemble d'états est formé par le vocabulaire non terminal* N *et un autre élément* s, *et tels que, quels que soient* $A \in N$, $v \in T^*$:

$$v \text{ dérive de } A \iff s \in k(A, v) \quad (\text{resp. } A \in k(s, v)).$$

Réciproquement, pour toute fonction d'automate fini k *d'alphabet* A *et tout couple d'états* s, s', *les mots* v *de* A^* *tels que* s' *appartienne à* k(s,v) *forment un langage de Kleene.*

Définition 2. *Soit une fonction d'automate fini* k, *d'alphabet* A, *d'ensemble d'états* S *et deux sous-ensembles* D *et* F *de* S ; *l'application* k' *de* $D \times A^*$ *dans* $\mathfrak{P}(F)$ *définie par* $k'(s,v) = k(s,v) \cap F$ *est un* automate fini, d'alphabet A, d'ensemble d'états S, d'ensemble initial D, d'ensemble final F. *Nous noterons* $k' = (A, S, D, F, k)$.

Toute fonction d'automate fini est un automate fini dont l'ensemble d'états, l'ensemble initial et l'ensemble final sont confondus, Soient deux automates finis $k' = (A, S, D, F, k)$ et $k'_1 = (A, S_1, D_1, F_1, k_1)$; l'automate fini $(A, S \times S_1, D \times D_1, F \times F_1, k \times k_1)$ s'appelle *produit* de k' et k'_1 : il est noté $k' \times k'_1$.

5. 1. 2. Composition à droite et à gauche. Théorème 5. 2. *Soit un automate fini* $k' = (A, S, D, F, k)$, D' *un ensemble fini disjoint de* S, h *une application de* D' *dans l'ensemble des parties finies de* $D \times A^*$; *l'application* k'' *de* $D' \times A^*$ *dans* $\mathfrak{P}(F)$, *définie par* $k''(s, v) = \hat{k}'[h(s).\,v]$[1] *est un automate fini d'alphabet* A, *d'ensemble d'états* $S \cup D'$, *d'ensemble initial* D', *d'ensemble final* F. *Nous dirons que* k'' *est obtenu par* composition à gauche *de* k' *par* h.

Définissons une application k_1 de $(S \cup D') \times A^*$ dans $\mathfrak{P}(S \cup D')$ par :

$k_1(s,v) = k(s,v)$ si $s \in S$; $k_1(s', v) = \hat{k}[h(s').\,v]$ si $s' \in D'$;

pour $s' \in D'$, $k_1(s',v) \cap F = k''(s,v)$. Il reste à vérifier que k_1 est une fonction d'automate :

$k_1(s', v\,a) = \hat{k}[h(s').\,v\,a] = \bar{k}[\hat{k}(h(s').\,v), 'a'] = \bar{k}_1[k_1(s',v), 'a']$.

Théorème 5. 3. *Soit un automate fini* $k' = (A, S, D, F, k)$, F' *un ensemble fini disjoint de* S, h *une application de* F *dans l'ensemble des parties finies de* F' ; $\hat{h} \circ k'$ *est un automate fini d'alphabet* A, *d'ensemble d'états* $S \cup F'$, *d'ensemble initial* D, *d'ensemble final* F' : *nous dirons qu'il est obtenu par* composition à droite *de* k' *par* h.

(1) G. v, où $G \subset S \times A^*$ et $v \in A^*$ a été défini en 5. 1. 1.

Prolongeons h en une application h' définie dans S, par $h'(s) = \emptyset$ si $s \notin F$; et définissons une application k_1 de $(S \cup F') \times A^*$ dans $\mathfrak{P}(S \cup F')$:

$$k_1(s,v) = [\hat{h}' o k(s,v)] \cup k(s,v) \text{ si } s \in S \; ; \; k_1(s', v) = \emptyset \text{ si } s' \in F'.$$

Alors $k_1(s,v) \cap F' = \hat{h}' o k(s,v) = \hat{h} o k'(s,v)$ pour $s \in D$; k_1 est une fonction d'automate :

$$\bar{k}_1[k_1(s,v), \text{'}\alpha\text{'}] = \bar{k}_1[k(s,v), \text{'}\alpha\text{'}] = [\hat{h}' o \bar{k}(k(s,v), \text{'}\alpha\text{'})] \cup \bar{k}(k(s,v), \text{'}\alpha\text{'}) = k_1(s, v\alpha).$$

5. 1. 3. Transducteurs finis bornés. Un *transducteur borné* sur un ensemble fini A est un algorithme qui transforme tout mot v du monoïde libre A^* en une partie finie $t(v)$ de A^*. Nous emploierons en particulier des transducteurs finis bornés : intuitivement, un transducteur fini [13](1) est un automate fini muni d'un dispositif qui imprime de gauche à droite.

Définition 3. Un transducteur fini borné *est un quintuplet* (A, S, s_d, s_f, k) *où* S *et* A *sont des ensembles finis (ensemble des états et alphabet),* s_d *et* s_f *deux éléments de* S *(états initial et final) et* k *une application de* $S \times A^*$ *dans l'ensemble de ses parties finies, telle que l'application* k' *de* $(S \times A^*) \times A^*$ *dans* $\mathfrak{P}(S \times A^*)$ *définie par :*

$$k'[(s, \rho), \lambda] = \{(s', \rho\mu) \; ; \; (s', \mu) \in k(s, \lambda)\}$$

soit une fonction d'automate d'alphabet A *et d'ensemble d'états* $S \times A^*$. *Il lui est associé une application* t *de* A^* *dans* $\mathfrak{P}(A^*)$ *par :*

$$\mu \in t(\lambda) \iff (s_f, \mu) \in k(s_d, \lambda).$$

k' est déterminée par sa restriction à $(S \times A^*) \times A_1$ où A_1 est l'ensemble des mots sur A de longueur 0 ou 1 ; comme $k(s, \lambda) = k'[(s, \Lambda), \lambda]$, k est déterminée par sa restriction à l'ensemble fini $S \times A_1$.

5. 2. Générateurs de piles.

On se donne un ensemble fini W, un élément σ n'appartenant pas à W, $E = W \cup \{\sigma\}$, un entier positif cd et une application t du monoïde libre E^* dans l'ensemble de ses parties finies. Ces éléments vont nous permettre d'associer à tout mot $\alpha = \text{'}\alpha_0 \ldots \alpha_{p-1}\text{'}$ du monoïde libre W^* un ensemble fini $\pi(\alpha)$ de piles sur W.

De t, nous déduisons d'abord une application t' de E^{*2} dans l'ensemble des parties de E^* par $t'(v', v'') = t(v'' \tilde{v}')$(2). Nous emploierons les notations suivantes : $\alpha_{p+i} = \sigma$ pour tout entier $i \geqslant 0$; $\alpha_i = \text{'}\alpha_i \ldots \alpha_{i+cd-1}\text{'}$ pour

(1) Si on use du vocabulaire de [53], il n'est question ici que de transducteurs droits.

(2) Rappelons que $\tilde{v}$ désigne le mot réfléchi du mot v (1. 3).

tout entier $i \geqslant 0$; $U = (u_0, u_1, \ldots, u_n)$ désignant une pile, un *p-partage* de U est une suite d'entiers j_i telle que $j_0 = 0 \leqslant j_1 \leqslant \ldots \leqslant j_p \leqslant j_{p+1} = n$; posons alors $v_i = u_{j_i}$ et soit μ_i le mot formé des éléments du mot de description de U (1. 6. 2) dont le rang r vérifie $j_i \leqslant r < j_{i+1}$.

Une pile U appartient à $\pi(\alpha)$ si, et seulement si, il existe un p-partage de U tel que $\mu_i \in t'(v_i, \alpha_i)$ pour $i = 0, 1, \ldots, p$. v_i et μ_i déterminent un v_{i+1} au plus : $v_{i+1} = v_i \boxtimes \mu_i$ (1. 6. 2).

Si l'application t est définie par un transducteur borné sur E, on construit, à partir de α, les piles de l'ensemble $\pi(\alpha)$ par l'algorithme suivant : pour $i = 0, 1, \ldots, p+1$ on détermine un ensemble F_i de couples (v_i, ρ_i) appartenant à $W^* \times E^*$:

a) $F_0 = \{\Lambda, \Lambda\}$;

b) F_{i+1} est l'ensemble des couples (v', ρ') tels qu'il existe $(v, \rho) \in F_i$ et $\mu \in t'(v, \alpha_i)$ pour lesquels $v' = v \boxtimes \mu$ et $\rho' = \rho\mu$.

Les mots ρ tels que $(\Lambda, \rho) \in F_{p+1}$ sont les mots de description des piles de l'ensemble $\pi(\alpha)$.

Nous appellerons cet algorithme un *générateur de piles* ; α est sa *donnée* ; lorsque le générateur est désigné par gn, nous notons $gn(\alpha)$ l'ensemble $\pi(\alpha)$ des piles associées à α. Si les ensembles $t(v)$ ont tous un élément au plus, les ensembles F_i ont tous au plus un élément : nous dirons alors qu'il s'agit d'un *générateur de pile unique* ; lorsqu'il n'en est pas ainsi, il s'agit d'un *générateur de piles multiples*. Tous les éléments de F_{p+1} sont de la forme (Λ, ρ) lorsqu'est vérifiée l'hypothèse :

(GP1) si $v' \in W^*$, $\mu \in t(\sigma^{cd} \tilde{v}')$ entraîne que $v' \boxtimes \mu$ est le mot vide ou n'existe pas.

Par récurrence sur i, la réunion, pour i fixé, des ensembles F_i associés à tous les mots α est finie. Il en résulte que, *pour i fixé,* j_i *est borné.*

Les algorithmes d'analyse qui vont être présentés sont des générateurs de piles ayant pour donnée le mot α à analyser. Les entiers j_i du p-partage, pour $1 \leqslant i \leqslant p$, ou les $j_i + 1$, seront les entrées ou les sorties des feuilles de la pseudo-arborescence cherchée ; l'entier cd représentera la longueur du "contexte droit" utile à l'analyse ; le plus souvent cd = 1. Un générateur de pile unique ne peut résoudre le problème de l'analyse que pour des grammaires non ambiguës.

Un générateur de piles est donc défini par un ensemble W, un entier cd et un transducteur borné. En pratique, cette notion de transducteur borné est trop générale, et nous restreindrons la classe des transducteurs bornés utilisés.

5. 3. Générateurs de piles utilisant un transducteur fini borné.

5. 3. 1. On se donne :

a) un ensemble fini M, et un élément s_f n'appartenant pas à M ; on pose $M' = M \cup \{s_f\}$ (M' est contenu dans l'ensemble des états du transducteur, et s_f est l'état final) ;

b) un entier $q \geqslant cd$ et une application l d'une partie D de E^q dans M ;

c) une application d de $M \times E$ dans l'ensemble des parties de M' ;

d) une application m de M dans l'ensemble des parties finies de $M' \times E^*$.

Nous appellerons *calcul de premier état* $s_0 \in M$ *et d'entrée* $v \in E^*$ toute suite finie de triplets (s_i, y_i, λ_i), $i = 1, \dots, r$ (r est la *longueur du calcul*) telle que :

- $s_i \in M'$; y_i est un mot de E^* de longueur 0 ($y_i = \Lambda$) ou 1 ; $\lambda_i \in E^*$;
- pour tout i : ou bien $y_i \neq \Lambda$, $s_i \in d(s_{i-1}, y_i)$ et $\lambda_i = {}'\sigma'$; ou bien $y_i = \Lambda$ et $(s_i, \lambda_i) \in m(s_{i-1})$;
- $y_1 \dots y_r$ est facteur gauche de v.

$\lambda_1 \dots \lambda_r$ est le *résultat* et s_r est le *dernier état* du calcul. Si $\rho \in D$ et $v \in E^*$, nous appellerons *calcul complet d'entrée* ρv tout calcul de premier état $l(\rho)$, d'entrée v, de dernier état s_f.

Pour que la longueur des calculs d'entrée donnée soit bornée, il faut et il suffit que :

(GP2) il n'existe aucun calcul d'entrée Λ dont le premier et le dernier état sont égaux.

Autrement dit, si on désigne par $m_1(s)$ la projection sur M' de m(s), sous-ensemble de $M' \times E^*$, il n'existe aucune suite $s_1, \dots, s_n$ d'éléments de M tels que $s_1 = s_n$ et, pour $i = 2, \dots, n$, $s_i \in m_1(s_{i-1})$. Nous imposerons cette condition GP2.

A partir de ces éléments, nous définirons l'application t qui, avec W et cd, détermine le générateur de piles, par : t(v) est l'ensemble des résultats des calculs complets d'entrée $v\,\sigma^{q-cd}$: t(v) est fini puisque la longueur des calculs d'entrée donnée est bornée.

Lorsque, pour tout couple (s,x) de $M \times E$, la réunion des ensembles m(s) et d(s,x) possède un élément au plus, nous dirons que le transducteur est *déterministe*. Il définit alors un générateur de pile unique.

5. 3. 2. Montrons qu'il existe effectivement un transducteur fini borné (E, S, s_d, s_f, k_0) tel que, pour son application associée t_0 et $|v| \geqslant q$, $t_0(v)$ soit l'ensemble des résultats des calculs complets d'entrée v. $E = W \cup \{\sigma\}$ et s_f ont déjà été définis ; $S = M \cup \{s_f\} \cup E_{q-1}$ où E_{q-1} est l'ensemble des mots sur E de longueur inférieure à q ; s_d est le mot vide. Pour définir k_0, nous ne nous intéresserons qu'aux calculs d'entrée v dont le dernier état est s_f ou tels que $y_1 \dots y_r = v$ (avec les notations de 5. 3. 1) ; ces calculs seront dits

exhaustifs. k_0 est défini de la manière suivante :

- Si $s \in M$, $k_0(s,v)$ est formé des couples du dernier état et du résultat des calculs exhaustifs de premier état s et d'entrée v : cet ensemble est fini à cause de GP2.
- $k_0(s_f, v) = \{(s_f, \Lambda)\}$.
- Si $s \in E_{q-1}$, $k_0(s,v) = \{(s\,v, \Lambda)\}$ pour $|s\,v| \leqslant q-1$,

  $k_0(s,v) = \{(l(s\,v), \Lambda)$ pour $|s\,v| = q$ et $s\,v \in D$,

  $k_0(s,v) = \phi$ pour $|s\,v| = q$ et $s\,v \notin D$;

pour $|sv| > q$, on peut écrire $s\,v = v_1\,v_2$ où $|v_1| = q$: si $v_1 \in D$, $k_0(s,v)$ est formé des couples du dernier état et du résultat des calculs exhaustifs de premier état $l(v_1)$ et d'entrée v_2 ; $k_0(s,v) = \phi$ si $v_1 \notin D$.

Il reste à montrer que l'application k'_0 associée à k_0 dans la définition d'un transducteur fini borné (5. 1. 3, définition 3) vérifie $k'_0[(s,\rho), v\,\alpha] = \overline{k'}_0[k'_0((s,\rho), v), {}'\alpha']$. C'est évident pour $s = s_f$. Pour $s \in M$ ou $s \in E_{q-1}$ et $|s\,v| > q$, cela résulte immédiatement du fait que les calculs exhaustifs de premier état s et d'entrée $v\,\alpha$ sont : les calculs de premier état s et d'entrée v dont s_f est le dernier état ; tout autre calcul exhaustif (C) de premier état s et d'entrée v, suivi des calculs exhaustifs d'entrée 'α' dont le premier état est le dernier état de (C) : le résultat est obtenu par concaténation de ceux des deux calculs. Pour $s \in E_{q-1}$ et $|sv| = q$, les calculs exhaustifs de premier état s et d'entrée $v\,\alpha$ sont ceux de premier état $l(sv)$ et d'entrée 'α' : les deux membres de l'égalité à démontrer sont vides ou valent $k'_0[(l(s,v), \rho), {}'\alpha']$. Pour $s \in E_{q-1}$ et $|sv| = q-1$, d'où $|s\,v\,\alpha| = q$, les deux membres valent $\{l(s\,v\,\alpha)\}$; pour $s \in E_{q-1}$ et $|sv| < q-1$, ils valent $\{s\,v\,\alpha\}$.

Remarque : Soit m^* l'application de M dans l'ensemble des parties finies de $M' \times E^*$, définie par $m^*(s) = k_0(s, \Lambda)$. Il est facile de voir que le transducteur reste le même si on remplace l'application m par m^* ; on peut alors imposer que dans les calculs qui définissent le transducteur, il n'existe jamais deux triplets consécutifs de la forme $(s_i, \Lambda, \lambda_i)$.

5. 3. 3. Intuitivement, l'application l permet de déduire des cd caractères lus dans le mot α (contexte droit) et, si $q > cd$, des q-cd derniers caractères de l'état courant v d'une pile engendrée, un état initial du transducteur. Puis, à chaque transition, ou bien un caractère sort de la pile et l'état du transducteur est modifié par l'application d, grâce à un nouveau caractère lu dans v, ou bien les états du transducteur et de la pile sont modifiés par m, sans tenir compte de v. En particulier, pour les transducteurs qui seront définis dans la suite, chaque caractère y_i lu dans v est le (q-cd)^{ème} avant le sommet de l'état v ■ $(\lambda_1 \dots \lambda_{i-1})$ de la pile ; si $q > cd$, les éléments compris dans la pile entre y_i et le sommet ont pu être "mémorisés" grâce aux états du transducteur. Il suffit pour cela que, quels que soient $s \in M$ et $(s', \lambda) \in m(s)$, λ contienne autant d'occurrences d'éléments de W que de σ, ou que dans tout calcul complet où un triplet est (s', Λ, λ) les triplets suivants soient tous de la forme (s'', Λ, λ').

5. 4. Généralisation.

5. 4. 1. Il est utile de généraliser les transducteurs finis bornés qui viennent d'être étudiés, en adjoignant aux éléments qui les définissent un automate fini k d'ensemble initial M, d'ensemble final $M \times \{0,1\} \times H$ où H est une partie finie de E^*, et en ajoutant dans la définition d'un calcul d'entrée ν une troisième possibilité pour les triplets (s_i, y_i, λ_i) :

$$(s_i, |y_i|, \lambda_i) \in k(s_{i-1}, \nu_{i-1}) \quad \text{où} \quad y_1 \dots y_{i-1} \; \nu_{i-1} = \nu$$

Appelons $k_1(s, \nu)$ la projection de $k(s, \nu)$ sur $M \times \{0,1\}$; ici, la condition GP 2 est équivalente à : quel que soit $\nu \in E^*$, il n'existe aucune suite $s_1, \dots, s_n$ d'éléments de M telle que $s_1 = s_n$ et, pour $i = 2, \dots, n$, $s_i \in m_1(s_{i-1})$ ou $(s_i, 0) \in k_1(s_{i-1}, \nu)$. L'application t est définie comme en 5. 3. 1.

On n'obtient plus ainsi des transducteurs finis. Intuitivement, avant certaines transformations des états de la pile et du transducteur, on cherche, grâce à un automate fini, une information sur la partie de l'état donné de la pile qui n'a pas encore été lue. (1) Il est facile de voir qu'on peut définir les mêmes transducteurs généralisés sans utiliser d'applications d et m, avec seulement un automate fini k. Nous conserverons pourtant la définition ci-dessus, qui permet d'évaluer la complexité du transducteur par l'ensemble des éléments s de M tels qu'il existe un mot ν pour lequel $k(s, \nu)$ n'est pas vide ; le cas particulier des transducteurs finis étudiés plus haut est celui où $k(s, \nu)$ est vide pour tout couple s, ν. Pour simplifier l'étude de l'ensemble des résultats des calculs, nous allons attacher au transducteur une nouvelle application χ, qui récapitule en quelque sorte les applications d, m, k.

5. 4. 2. χ est une application de $M \times E^*$ dans l'ensemble des parties finies de $M' \times E^* \times E^*$: pour $s \in M$ et $\nu \in E^*$, $\chi(s, \nu)$ est formé des éléments :

- $(s', \nu', '\sigma')$ si $\nu = x\nu'$ $(x \in E)$ et $s' \in d(s, x)$,
- (s', ν, λ) si $(s', \lambda) \in m(s)$ ou $(s', 0, \lambda) \in k(s, \nu)$,
- (s', ν', λ) si $\nu = x\nu'$ $(x \in E)$ et $(s', 1, \lambda) \in k(s, \nu)$.

Dans un calcul de premier état s_0, d'entrée ν_0, de longueur r, pour $i = 1, \dots, r$, $(s_i, \nu_i, \lambda_i) \in \chi(s_{i-1}, \nu_{i-1})$ (ν_i a été défini ci-dessus : $y_1 \dots y_i \; \nu_i = \nu$). Réciproquement, soit $s_0 \in M$, $\nu_0 \in E^*$ et une suite de triplets (s_i, ν_i, λ_i), $i = 1, \dots, r$, tels que, pour tout i, $(s_i, \nu_i, \lambda_i) \in \chi(s_{i-1}, \nu_{i-1})$: d'après la définition de χ, les cas suivants sont possibles :

- $\nu_{i-1} = y_i \nu_i$, $y_i \in E$, $\lambda_i = '\sigma'$ et $s_i \in d(s_{i-1}, y_i)$;
- $\nu_{i-1} = \nu_i$, $(s_i, \lambda_i) \in m(s_{i-1})$ ou $(s_i, 0, \lambda_i) \in k(s_{i-1}, \nu_{i-1})$; on pose alors $y_i = \Lambda$;
- $\nu_{i-1} = y_i \nu_i$, $y_i \in E$ et $(s_i, 1, \lambda_i) \in k(s_{i-1}, \nu_{i-1})$.

(1) cf. la définition d'un automate à pile, donnée par [54].

Pour tout i, $\nu_{i-1} = y_i \nu_i$; on en déduit $\nu_0 = y_1 \dots y_i \nu_i$. Les triplets (s_i, y_i, λ_i) forment un calcul de premier état s_0 et d'entrée ν_0.

5. 4. 3. Déterminisme du transducteur. Un premier usage de l'application χ est la généralisation de la notion de déterminisme aux transducteurs définis en 5. 4. 1 ; on dira qu'un tel transducteur est déterministe lorsque, quels que soient $s \in M$ et $\nu \in E^*$, l'ensemble $\chi(s, \nu)$ contient au plus un élément ; le transducteur définit alors un générateur de pile unique. Dans le cas d'un transducteur fini, on retrouve la définition du déterminisme donnée en 5. 3. 1 ; dans le cas général, pour que le transducteur soit déterministe, il suffit que pour tous $s \in M$, $x \in E$, $\nu' \in E^*$, la réunion de $m(s)$, $d(s, x)$, $k(s, x\nu')$ contienne au plus un elément, mais ce n'est pas nécessaire.

5. 5. τ -projection et projection d'un transducteur dans un autre.

Envisageons deux générateurs de piles gn^0 et gn^1, définis par deux ensembles W^0 et W^1 ($E^0 = W^0 \cup \{\sigma\}$, $E^1 = W^1 \cup \{\sigma\}$), deux entiers cd^0 et cd^1 et deux transducteurs tr^0 et tr^1, donnés respectivement par M^0, q^0, l^0, d^0, m^0, k^0, définissant χ^0 et M^1, q^1, l^1, d^1, m^1, k^1 définissant χ^1. Supposons qu'on donne une application τ de E^1 dans E^0 pour laquelle $\tau(\sigma) = \sigma$; τ définit une transcription τ^* de E^{1*} dans E^{0*}. Soit ϖ une application de $M^1 \times E^{1*}$ dans $M^0 \times E^{0*}$ telle que, pour $s \in M^1$, $\nu \in E^{1*}$, $s' \in M^1$, $\nu' \in E^{1*}$, $\lambda \in E^{1*}$:

a) $(s', \nu', \lambda) \in \chi^1(s, \nu)$ entraîne $[\varpi(s', \nu'), \tau^*(\lambda)] \in \chi^0[\varpi(s, \nu)]$;
b) $(s_f, \nu', \lambda) \in \chi^1(s, \nu)$ entraîne qu'il existe $\nu'' \in E^{0*}$ pour lequel $[s_f, \nu'', \tau^*(\lambda)] \in \chi^0[\varpi(s, \nu)]$.

D'après 5. 4. 2, τ transcrit le résultat de tout calcul de tr^1, dont le premier état est s^1, l'entrée ν^1 et le dernier état s_f, en le résultat d'un calcul de tr^0 dont le premier état et l'entrée vérifient $(s^0, \nu^0) = \varpi(s^1, \nu^1)$, et dont le dernier état est s_f. D'autre part, si un calcul de tr^1, d'entrée Λ, a son premier et son dernier état égaux, il en est de même pour un calcul de tr^0 : lorsque tr^0 vérifie la condition GP 2, tr^1 la vérifie également.

Supposons $cd^1 \geqslant cd^0$. A tout mot $\gamma^1 \nu \sigma^{q^1 - cd^1}$ où $\gamma^1 \in (E^1)^{cd^1}$ et $\nu \in E^{1*}$, associons $\gamma^0 \tau^*(\nu) \sigma^{q^0 - cd^0}$ où γ^0 est le facteur gauche de $\tau^*(\gamma^1)$ qui a pour longueur cd^0. Ecrivons $\gamma^0 \tau^*(\nu) \sigma^{q^0 - cd^0} = \rho^0 \nu^0$ où $|\rho^0| = q^0$, et $\gamma^1 \nu \sigma^{q^1 - cd^1} = \rho^1 \nu^1$ où $|\rho^1| = q^1$. Si ϖ vérifie aussi :

c) pour tout $\gamma^1 \in (E^1)^{cd^1}$ et tout $\nu \in E^{1*}$, $l^0(\rho^0)$ existe lorsque $l^1(\rho^1)$ existe et $\varpi(l^1(\rho^1), \nu^1) = (l^0(\rho^0), \nu^0)$,

τ transcrit le résultat de tout calcul complet de tr^1, d'entrée $\gamma^1 \nu \sigma^{q^1 - cd^1}$ en le résultat d'un calcul complet de tr^0 d'entrée $\gamma^0 \tau^*(\nu) \sigma^{q^0 - cd^0}$. τ transcrit toute pile engendrée par gn^1, pour une donnée α, en une pile engendrée par gn^0, pour la donnée $\tau^*(\alpha)$. D'autre part, si gn^0 vérifie GP 1, il en est de même pour gn^1. Nous dirons alors que ϖ *τ-projette* tr^1 dans tr^0.

Dans le cas où $W^0 = W^1$ et où τ est l'identité, nous dirons simplement que ϖ *projette* tr^1 dans tr^0 ; alors, pour tout mot α de T^*, l'ensemble $gn^1(\alpha)$ est contenu dans $gn^0(\alpha)$: nous dirons que gn^1 *est contenu dans* gn^0.

5. 6. Note pour une généralisation.

La notion de générateur de piles qui vient d'être introduite ne rend pas compte des algorithmes tels que ceux de [10], [35], [34], [57], (cf. [7]) : ces algorithmes permettent des retours en arrière dans la donnée α et résolvent le problème de la reconnaissance pour une classe de grammaires de Chomsky et le problème de l'analyse pour celles de ces grammaires qui ne sont pas ambiguës, en utilisant un nombre borné de piles. L'algorithme de [10] construit la pile attachée à une pseudo-arborescence de $\mathcal{C}(G)$ ayant α pour suite de feuilles ; ceux de [35], [34], [57] construisent la pile attachée à la pseudo-arborescence gauche d'une pseudo-arborescence de $\mathcal{C}(G)$ ayant α pour suite de feuilles. Pour rendre compte de ces algorithmes et en construire d'autres du même genre, on serait amené à généraliser la notion de générateur de pile unique de manière à pouvoir associer à certains générateurs de piles gn, définis à l'aide d'un transducteur non déterministe, un générateur ainsi généralisé gn' tel que gn'(α) soit vide si, et seulement si, gn(α) l'est, et que gn'(α) soit contenu dans gn(α). On pourrait alors justifier du domaine d'application de ces algorithmes.

CHAPITRE 6

ANALYSE PAR CONSTRUCTION D'UNE PILE ADJOINTE OU ATTACHEE AUX PSEUDO-ARBORESCENCES CHERCHEES

6. 1. Analyse par construction d'une pile adjointe.

Nous nous limitons ici au cas d'une grammaire de Chomsky non généralisée.

D'après l'étude de la pile adjointe à une ramification orientée (2.9.2), la pile U adjointe à une pseudo-arborescence (I, f) de $\mathcal{C}(N, T, ::=, X)$ possède les propriétés suivantes :

1) la première entrée est celle de la racine X ; elle est suivie d'une sortie ;
2) après une sortie de symbole terminal, se produit une nouvelle sortie si la pile n'est pas vide ; si la pile est vide, il s'agit de la dernière sortie ;
3) la suite de feuilles de (I, f) est la trace sur T^* du mot de sortie de la pile ;
4) après toute sortie d'un symbole non terminal A, entrent les éléments d'un mot Ψ tel que $A ::= \tilde{\Psi}$, puis se produit une sortie.

Soit $\alpha = `\alpha_0 \ldots \alpha_{p-1}'$ la suite de feuilles de (I, f). Les entiers j_i tels que j_i+1 soit une sortie de symbole terminal, $j_0 = 0$ et l'indice $j_{p+1} = n$ du dernier état de la pile U, forment un p-partage de U(j_i+1 est sortie de α_{i-1}). Pour cé p-partage, avec les notations du paragraphe 5. 2 :

- μ_0 est de la forme $X \sigma \varphi_1 B_1 \sigma \ldots \sigma \varphi_k B_k \ldots \sigma \varphi_r \alpha_0$ où $r \geqslant 1$ et $B_{k-1} ::= B_k \tilde{\varphi}_k$ pour $1 \leqslant k \leqslant r$, en posant $B_0 = X$ et $B_r = \alpha_0$ (d'après les propriétés 1,4 ci-dessus) ; $v_0 = \Lambda$.
- Pour $0 < i < p$, si j_i+2 est sortie d'un symbole non terminal C, $v_i = \eta C \alpha_{i-1}$, μ_i est de la forme $\sigma \sigma \varphi_1 B_1 \sigma \ldots \sigma \varphi_k B_k \sigma \ldots \sigma \varphi_r \alpha_i$ où $r \geqslant 1$ et $B_{k-1} ::= B_k \tilde{\varphi}_k$ pour $1 \leqslant k \leqslant r$, en posant $B_0 = C$ et $B_r = \alpha_i$ (d'après 4) ; sinon, j_i+2 est sortie d'un symbole terminal (d'après 2), $j_{i+1} = j_i+1$, $v_i = \eta \alpha_i \alpha_{i-1}$ et $\mu_i = `\sigma'$.
- $v_p = `\alpha_{p-1}'$ et $\mu_p = `\sigma'$.

Cette étude va nous permettre de construire un générateur de piles gn dont la donnée est un mot α de T^* et qui engendre les piles adjointes aux pseudo-arborescences de $\mathcal{C}(N, T, ::=, X)$ dont α est la suite de feuilles : l'ensemble W est V, l'entier cd est 1 ; l'application t est définie par :

- pour $a \in T$, $t('a')$ est l'ensemble des mots $X \sigma \varphi_1 B_1 \sigma \dots \sigma \varphi_r a$ pour lesquels $r \geqslant 1$ et $B_{k-1} ::= B_k \tilde{\varphi}_k$ pour $1 \leqslant k \leqslant r$, en posant $B_0 = X$ et $B_r = a$ (a est une initiale de chaque B_k) ;
- pour $a \in T$, $b \in E$, $C \in N$, $\eta \in V^*$, $t(a\, b\, C\, \eta)$ est l'ensemble des mots $\sigma \sigma \varphi_1 B_1 \sigma \dots \sigma \varphi_r a$ pour lesquels $r \geqslant 1$ et $B_{k-1} ::= B_k \tilde{\varphi}_k$ pour $1 \leqslant k \leqslant r$, en posant $B_0 = C$ et $B_r = a$ (a est une initiale de chaque B_k) ;
- pour $a \in T$, $b \in E$, $\eta \in V^*$, $t(a\, b\, a\, \eta) = \{'\sigma'\}$;
- pour $b \in E$, $t('\sigma\, b') = \{'\sigma'\}$;
- dans les autres cas $t(\nu) = \phi$.

Notons que si $v \in V^*$, $a \in E$ et $\mu \in t'(v, a) = t(a\, \tilde{v})$, $v \boxtimes \mu$ est toujours défini, et que t vérifie la condition GP 1. Il résulte de l'étude précédente que l'ensemble $gn(\alpha)$ contient les piles cherchées.

Réciproquement, soit U une pile de $gn(\alpha)$, pour $\alpha = 'a_0 \dots a_{p-1}'$. Elle est adjointe à une pseudo-ramification (I, f) dans V. Chaque mot μ_i appartient à l'un des ensembles $t(\nu)$ qui viennent d'être décrits. μ_0, qui appartient à $t('a_0')$, commence par $X\sigma$: (I, f) a X pour seule racine. Pour $0 < i \leqslant p$, j_i+1 est une sortie du sommet de $v_i = v_{i-1} \boxtimes \mu_{i-1}$; si $\mu_{i-1} \neq '\sigma'$, $\mu_{i-1} = t(a_{i-1}\, \eta')$ se termine par a_{i-1} ; si $\mu_{i-1} = '\sigma'$, $\mu_{i-1} = t(a_{i-1}\, b\, a_{i-1}\, \eta)$ et $v_{i-1} = \eta\, a_{i-1}\, b$; dans les deux cas, le sommet de v_i est a_{i-1} ; dans les deux cas aussi, j_i+2 n'est pas une entrée : a_{i-1} est l'image par f d'une feuille de J. Si $j_i+2 \neq j_{i+1}$, $|\mu_i| > 1$ et j_i+2 est sortie de l'élément C qui précède le sommet de v_i : $a_i\, \tilde{v}_i = a_i\, a_{i-1}\, C\, \eta$; après cette sortie, entre un mot Ψ tel que $C ::= \tilde{\Psi}$: $C \in N$, $\tilde{\Psi}$ est une famille de prédécesseur C. Si k est une sortie de la pile qui n'est ni un j_i+1, ni un j_i+2, c'est la sortie de l'élément B d'entrée $k-1$; k est suivi des entrées des éléments d'un mot Ψ tel que $B ::= \tilde{\Psi}$: $B \in N$, $\tilde{\Psi}$ est une famille de prédécesseur B. La suite de feuilles de (I, f) est donc α et (I, f) appartient à $\mathfrak{C}(N, T, ::=, X)$. $gn(\alpha)$ est l'ensemble des piles adjointes aux pseudo-arborescences de $\mathfrak{C}(N, T, ::=, X)$ dont α est la suite de feuilles.

L'ensemble $t(\nu)$ ne dépend que du premier et du troisième élément du mot $\nu\, \sigma\, \sigma\, \sigma$; t est défini par une application d'un ensemble fini dans l'ensemble des parties de E^*. t est l'application associée à un transducteur borné lorsque les ensembles $t(\nu)$ sont tous finis (1). Les suites ($B_0 = X$ ou C, $B_1, \dots, B_k, \dots, B_r = a$) rencontrées dans la définition de t sont des chemins du graphe des initiales de la grammaire. Pour que t soit défini par un transducteur borné, il suffit donc que *ce graphe ne contienne pas de circuit*, autrement dit *qu'aucun symbole ne soit initiale stricte de lui-même*. Nous désignerons dans la suite cette condition par AD.

(1) On peut d'ailleurs définir un transducteur fini borné dont t soit l'application associée.

Supposons inversement que la grammaire soit réduite et que A soit initiale stricte de lui-même. De toute pseudo-arborescence de la structure $\mathcal{S}(V, ::=)$ qui a A pour racine, on peut en déduire une autre de hauteur supérieure. A possède une initiale a appartenant à T. Les pseudo-arborescences de $\mathcal{S}(V, ::=)$ dont A est la racine et a la première feuille, ont leur hauteur h non bornée. Elles sont sous-pseudo-arborescences complètes de pseudo-arborescences de $\mathfrak{C}(N, T, ::=, X)$ dont les suites de feuilles ont même facteur gauche $b_0 \dots b_{i-1}$ avec $b_{i-1} = a$ (théorème 4. 2). La sortie σ_i de la $i^{\text{ème}}$ feuille a est précédée d'au moins h-1 sorties : σ_i n'est donc pas bornée. Les piles adjointes aux pseudo-arborescences de l'ensemble ne peuvent être engendrées par un générateur tel que $\frac{\sigma_i}{j_i}$ soit borné (1).

La condition (AD) est en particulier réalisée si toute règle est du type $A ::= a\, \varphi$ ou $a \in T$ et $\varphi \in N^*$; [30] montre que pour toute grammaire, il existe une grammaire équivalente qui possède cette propriété. Alors, les chemins du graphe des initiales sont tous du type (A, a). Les entiers r utilisés lors de la définition de l'application t sont tous égaux à 1. Ce cas est étudié par [29]. L'algorithme de [41] revient aussi à engendrer la pile adjointe à une pseudo-arborescence.

Pour qu'on obtienne un générateur de pile unique, il suffit que :
(AD') *pour tout* $A \in N$ *et tout* $a \in T$, *il existe au plus une règle* $A ::= B\, \varphi$ *telle que* a *soit une initiale de* B.
Lorsque la grammaire est réduite, cette condition entraîne AD ; en effet, si A est initiale stricte de lui-même, il existe dans le graphe des initiales plusieurs chemins d'origine A ayant la même extrémité $a \in T$; deux de ces chemins ont un premier point non commun, B sur l'un, B' sur l'autre, précédés d'un point commun A' ; il existe donc deux règles $A' ::= B\, \varphi$ et $A' ::= B'\, \varphi'$, B et B' ayant tous deux a pour initiale.

Supposons maintenant que la grammaire est réduite, que AD' n'est pas satisfaite, et envisageons un générateur des piles adjointes tel que cd = 1 et que j_i+1 soit la sortie de la $i^{\text{ème}}$ feuille, pour $0 < i \leqslant p$. Il existe dans $\mathcal{S}(V, ::=)$ deux pseudo-arborescences de même racine A, de même première feuille a, où les deux familles qui ont la racine pour prédécesseur sont différentes. Elles sont sous pseudo-arborescences complètes de deux pseudo-arborescences de $\mathfrak{C}(N, T, ::=, X)$ dont les suites de feuilles ont même facteur gauche $b_0 \dots b_{i-1}$, avec $b_{i-1} = a$. Par conséquent, pour une donnée $b_0 \dots b_{i-1}\, \eta$, l'ensemble F_i contient au moins deux éléments : le générateur est un générateur de piles multiples. En résumé :

(1) Mais il est cependant possible qu'elles le soient par un autre générateur : par exemple, si $T = \{a\}$, $N = \{X\}$ et $X ::= X\, a | a$, elles le sont par le générateur pour lequel cd = 1 et $t(a\, \eta) = \{'X \sigma a'\}$, $t(\sigma\, \eta) = \{\sigma^{|\eta|}\}$; l'application t est associée à un transducteur fini déterministe.

Théorème 6. 1. *Soit une grammaire* $(N, T, ::=, X)$ *réduite. Pour que l'application* t *définisse, avec l'ensemble* V *et* cd = 1, *un générateur de piles* gn, *il faut et il suffit que le graphe des initiales de la grammaire ne possède pas de circuit. Pour tout mot* α *sur* T, gn(α) *est alors l'ensemble des piles conjointes aux pseudo-arborescences de* $\mathscr{C}(N, T, ::=, X)$ *dont* α *est la suite de feuilles.* gn *est un générateur de pile unique si, et seulement si, la grammaire vérifie la condition* AD'.

6. 2. Utilisation dans la suite du chapitre d'un pseudo-graphe de production.

Nous emploierons dans la suite du chapitre un pseudo-graphe de production $(J, \Gamma ; g)$ du premier ou du second type (4.6.4). Nous pourrons ainsi traiter le cas d'une grammaire généralisée. Les mots $A\varphi$ (resp. $\varphi\bar{A}$) tels que $A ::= \varphi$ sont les chemins de $(J, \Gamma ; g)$ dont l'origine et l'extrémité appartiennent respectivement à deux sous-ensembles J' et J''. Nous supposerons que le pseudo-graphe possède les propriétés PG1, PG2, PG3 ; l'origine dans J' d'un chemin commençant par Z est o(Z).

Soit (I, f) une pseudo-arborescence de $\mathscr{C}(N, T, ::=, X)$. Un noeud z de I a pour image $f(z) = A$, la famille β de prédécesseur z est transcrite par f en φ et $A ::= \varphi$; $A\varphi$ (resp. $\varphi\bar{A}$) est un chemin de $(J, \Gamma ; g)$, transcrit par g d'un chemin (unique) $w\beta'$ (resp. $\beta' w$) de (J, Γ) dont l'origine appartient à J' : $w \in J$, $|\beta| = |\varphi| = |\beta'|$; en associant à tout élément y de β l'élément $f_1(y)$ de même rang dans β', on définit une application de I, privé de sa racine r, dans J. Soit δ un élément n'appartenant pas à J ; prolongeons f_1 par $f_1(r) = \delta$ et g par $g(\delta) = X$; alors $f = g \circ f_1$. (I, f_1) est une pseudo-arborescence dans l'ensemble $J_1 = J \cup \{\delta\}$. Soit $\mathscr{C}_1(N, T, ::=, X)$ l'ensemble des pseudo-arborescences (I, f_1) ainsi associées aux pseudo-arborescences de $\mathscr{C}(N, T, ::=, X)$. Comme (I, f) est transformée de (I, f_1) par g, l'application qui à (I, f) associe (I, f_1) est une bijection de $\mathscr{C}(N, T, ::=, X)$ sur $\mathscr{C}_1(N, T, ::=, X)$; (I, f_1) détermine (I, f).

Dans la pile attachée à une pseudo-ramification, une sortie d'un élément Y est suivie d'une entrée d'un élément Z qui suit Y dans une famille, ou d'une sortie si Y est dernier élément d'une famille. Les familles β' de (I, f_1) sont des chemins d'un graphe ; les éléments w prolongeant un chemin β_0 d'un graphe en un chemin $\beta_0 w$ ne dépendent que du dernier élément de β_0. Cette propriété est fausse pour les chemins d'un pseudo-graphe ; or les familles φ de (I, f) sont des chemins d'un pseudo-graphe. Plutôt que de construire les piles attachées aux pseudo-arborescences (I, f) de $\mathscr{C}(N, T, ::=, X)$, nous construirons les piles attachées aux (I, f_1) associées.

6. 3. Analyse par construction des piles attachées aux pseudo-arborescences de $\mathscr{C}(G)$.

Dans la pile strictement attachée à une ramification orientée, un noeud entre avant les points dont il est le prédécesseur. Il sera commode d'utiliser ici un pseudo-graphe de production $(J, \Gamma ; g)$ du premier type.

6. 3. 1. Caractérisation des piles cherchées. D'après l'étude de la pile strictement attachée à une ramification orientée (2. 1), la pile attachée à une pseudo-ramification (I, f_1) de $\mathscr{C}_1(G)$ possède les propriétés suivantes :

(AT1) La première entrée est une entrée de δ.

(AT2) Si i est une entrée de x et i+1 une sortie, $g(x) \in T$.

(AT3) Si i est une entrée de x et i+1 une entrée de y, $o[g(x)] \;\Gamma\; y$ (d'où $g(x) \in N$ et $y \in J$).

(AT4) Si i est une sortie de x et i+1 une entrée de y, x et y sont des points de J tels que $x \,\Gamma\, y$.

(AT5) Si i est une sortie de x et i+1 une sortie, x appartient à J''.

Réciproquement, soit U_1 une pile sur l'ensemble J_1, qui possède ces cinq propriétés. Elle est attachée à une pseudo-ramification (I, f_1), et transcrite par g en la pile U attachée à $(I, g \circ f_1) = (I, f)$. (I, f) est une pseudo-ramification sur V ; montrons qu'elle appartient à $\mathscr{C}(N, T, ::=, X)$. D'après AT3 et AT4, toute entrée autre que la première est celle d'un point de J ; d'après AT4 et AT5, il en est de même pour toute sortie autre que la dernière ; d'après AT1, la première entrée, et donc la dernière sortie, portent sur δ : (I, f) est une pseudo-arborescence de racine $g(\delta) = X$. Si z est une feuille de I, $g[f_1(z)] = f(z) \in T$ (AT2). Pour un noeud z de I, posons $x = f_1(z)$ et $A = g(x) = f(z)$: la famille β de prédécesseur z dans I est transcrite par f_1 en une famille β' de prédécesseur x dans (I, f_1), elle-même transcrite par g en une famille φ de prédécesseur A dans (I, f) ; d'après AT3, AT4, AT5, le mot $A\varphi$ est un chemin du pseudo-graphe $(J, \Gamma ; g)$ dont l'origine est o(A) et l'extrémité dans J'' : $A ::= \varphi$. (I, f) appartient à $\mathscr{C}(N, T, ::=, X)$.

Complétons les propriétés précédentes par trois autres propriétés de la pile U_1 attachée à $(I, f_1) \in \mathscr{C}_1(G)$; g transcrit U_1 en U, attachée à $(I, f) \in \mathscr{C}(G)$:

(AT6) Si i est une entrée de y dans U_1 et si dans U la première entrée de symbole terminal supérieure ou égale à i est une entrée de b, b est initiale de g(y) : en effet, il existe deux points y_1 et b_1 de I tels que $y = f_1(y_1)$, $b = f(b_1)$ et b_1 soit la première feuille qui majore y_1 dans l'ordre d'entrée de I ; b est initiale de $f(y_1) = g(y)$ (4. 3. 1).

(AT7) Si i est une sortie de y de la pile U_1 et si la première sortie de U d'un symbole terminal qui est supérieure à i est une sortie de b, g(y) et b, images par f de deux points consécutifs de I, forment un couple vicinal (4. 3. 2).

(AT8) Si i est une sortie de x et i+1 une sortie de y, g(x) fn g(y) (4. 3. 1).

6. 3. 2. Générateur. De l'étude qui précède, déduisons *un générateur de piles* gn *tel que, pour tout mot* α *de* T^*, g *transcrive les piles de l'ensemble* gn(α) *en les piles attachées aux pseudo-arborescences de* $\mathcal{C}$(G) *dont* α *est la suite de feuilles.* Nous emploierons les notations du chapitre 5 ; les j_i seront les entrées de feuilles. Ce générateur est défini par $W = T \cup J_1$ (rappelons que $E = W \cup \{\sigma\}$), cd = 1 et par un transducteur fini borné (5. 3) tr ; à tout élément x de J_1, associons bijectivement un élément $\bar{x}$ n'appartenant pas à J_1 ; soit $\bar{J}_1$ l'ensemble de ces $\bar{x}$; appelons d'autre part T' l'ensemble $T \cup \{\sigma\}$.

a) $M = T' \times (J_1 \cup \bar{J}_1)$.

b) q = 2 ; pour $\alpha \in T$, $l(\alpha, \sigma) = (\alpha, \delta)$; pour $\alpha \in T'$ et $x \in J_1$, $l(\alpha, x) = (\alpha, \bar{x})$.

c) Pour $\alpha \in T'$, $x \in J_1$ et $x' \in J_1$, d[(α,$\bar{x}$), x'] contient les seuls éléments :

- (α, y) pour tout $y \in J$ tel que x Γ y et α soit initiale de g(y) (d'où $x \neq \delta$ et $\alpha \neq \sigma$),
- (α, $\bar{x}'$) si $x \in J''$, g(x) fn g(x') (1) et si, ou bien le couple g(x'), α est vicinal, ou bien $\alpha = \sigma$;
  d [(σ, $\bar{\delta}$), σ] = {s_f}.

d) Pour $\alpha \in T$ et $x \in J_1$, m (α, x) contient les seuls couples suivants :

- (s_f,' x') si g(x) = α, '
- [(α, y),' x'] si $g(x) \in N$, pour tout $y \in J$ tel que o[g(x)] Γ y et α soit initiale de g(y).

Dans tous les cas où $l(\alpha, b)$ n'a pas été défini ($\alpha \in E$, $b \in E$), $l(\alpha, b)$ n'existe pas ; dans les cas où d (s,x') ou m (s) n'a pas été défini ($s \in M$, $x' \in E$), cet ensemble est vide. *Ces conventions seront valables pour les applications* l, d, m, k *des transducteurs définis dans toute la suite.*

Le résultat d'un calcul du transducteur est un mot sur $J_1 \cup \{\sigma\}$. Etudions divers calculs complets ; on remarque que si s' est la projection sur M d'un couple de m(s), d(s', x') est vide pour tout x'.

1) Si l'entrée est 'α σ ' ($\alpha \in T$), $s_0 = (\alpha, \delta)$ et les calculs complets sont les suites : ((α, y_1), Λ, 'δ') ...((α, y_k), Λ, 'y_{k-1}' ' ...((α, y_{r-1}), Λ, 'y_{r-2}') (s_f, 'y_{r-1}') où $r \geqslant 2$, $\alpha = g(y_{r-1})$ et o[g(y_{k-1})] Γ y_k pour $1 \leqslant k \leqslant r-1$, en posant $y_0 = \delta$; ces conditions entraînent que α est initiale des g(y_k). Le résultat du calcul est 'δ y_1 ... y_{r-1}'.

2) Si l'entrée est 'α x_1 ... $x_{r'}\sigma$ ' ($r' \geqslant 1$, $\alpha \in T$, $x_k \in J_1$), $s_0 = (\alpha, \bar{x}_1)$; les calculs complets sont les suites : ((α, $\bar{x}_2$), x_2, 'σ') ... ((α, $\bar{x}_{r''}$), $x_{r''}$, 'σ') ((α, $y_{r''}$), $x_{r''+1}$, 'σ') ((α, $y_{r''+1}$), Λ, '$y_{r''}$') ... ((α, y_{r-1}), Λ, 'y_{r-2}') (s_f, 'y_{r-1}') où : $1 \leqslant r'' < r$, $r'' \leqslant r'$; pour $1 \leqslant k \leqslant r''-1$, $x_k \in J''$ et g(x_k) fn g(x_{k+1}) ; $x_{r''}$ Γ $y_{r''}$; pour $r''+1 \leqslant k \leqslant r-1$, o[g($y_{k-1}$)] Γ y_k ; g(y_{r-1}) = α ; il en résulte que α est une initiale de g($y_{r''}$), ... , g(y_{r-1}) et que g(x_2), ... , g($x_{r''}$) forment avec α des couples vicinaux (théorème 4. 3). Le résultat du calcul est $\mu = \sigma^{r''}$ $y_{r''}$ ... y_{r-1} . '$x_{r'}$ ... x_1' ⊠ μ existe puisque $r'' \leqslant r'$, et ce n'est pas le mot vide.

(1) Voir la remarque qui suit l'étude des calculs du transducteur.

3) Si l'entrée est 'σ x_1 ... $x_{r'}$ σ' ($r' \geqslant 1$, $x_k \in J_1$), $s_0 = (\sigma, \bar{x}_1)$; puisqu'il existe un calcul complet, il faut que $x_{r'} = \delta$; ce calcul est ((σ, $\bar{x}_2$), x_2, 'σ') ... ((σ, $\bar{x}_k$), x_k, 'σ') ... ((σ, $\bar{\delta}$), δ, 'σ'), (s_f, σ, 'σ') si, pour $1 \leqslant k \leqslant r'-1$, $x_k \in J''$ et g(x_k) fn g(x_{k+1}). Le résultat du calcul est $\sigma^{r'}$. '$x_{r'}$ ... x_1' ⊠ $\sigma^{r'}$ existe et c'est le mot vide : le générateur gn vérifie GP 1.

Remarque : envisageons les calculs du transducteur étudiés en 2 et 3. Si on suppose que, pour $1 \leqslant k \leqslant r-1$, il existe un chemin de (J, Γ) d'origine o[g(x_{k+1})] et d'extrémité x_k, $x_k \in J''$ entraîne g(x_k) fn g(x_{k+1}). L'hypothèse est vérifiée lorsque '$x_{r'}$ ... x_1' est l'un des états v_i d'une pile engendrée : on le voit facilement par récurrence sur i. Il sera donc inutile de tenir compte de la condition g(x) fn g(z), imposée lors de la définition de l'application d pour que (α, $\bar{x}'$) appartienne à d[(α, $\bar{x}$), x'].

Soit α = 'α_0 ... α_{p-1}' un mot sur T. Envisageons une pile U_1 de l'ensemble gn(α). Le premier calcul du transducteur conduisant à cette pile est du type 1 ci-dessus ; v_1 appartient à J_1^* et n'est pas vide ; le second calcul est donc du type 2, et, par récurrence, il en est de même de tous jusqu'au $p^{\text{ème}}$; le dernier calcul est du type 3. La première entrée est donc une entrée de δ : la pile vérifie donc la condition AT 1. Pour $1 \leqslant i \leqslant p$, j_i est une entrée d'un élément y tel que g(y) = α_{i-1} ; toute autre entrée est suivie d'une entrée ; U_1 vérifie aussi AT 2.Toute entrée de x suivie d'une entrée de y est distincte d'un j_i ; d'après les résultats des calculs de type 1 ou 2, o[g(x)] Γ y. Aucune sortie n'est un j_i, ni inférieure à j_1 : si la sortie de x est suivie de l'entrée de y, x Γ y d'après les résultats des calculs du type 2 ; si elle est suivie d'une sortie, $x \in J''$. La pile U_1 vérifie donc aussi AT 3, AT 4, AT 5 ; elle est transcrite par g en la pile attachée U à une pseudo-arborescence (I, f) de $\mathcal{C}$(N, T, ::=, X). La suite de feuilles de (I, f) est la suite des symboles terminaux qui entrent dans U, c'est-à-dire α.

Réciproquement, la pile U attachée à une pseudo-arborescence de $\mathcal{C}$(N, T, ::=, X), ayant α pour suite de feuilles, est transcrite par g d'une pile U_1 qui vérifie AT 1, AT 2, AT 3, AT 4, AT 5, AT 6, AT 7, AT 8. Soit j_i l'entrée dans U de la $i^{\text{ème}}$ feuille, α_{i-1} ; $j_0 = 0$, j_1, ... , j_p et j_{p+1}, indice du dernier état de U, forment un p-partage de U_1. D'après AT 2, toute entrée qui n'est pas un j_i est suivie d'une autre entrée. Pour le p-partage considéré :

- μ_0 est de la forme 'δ y_1 ... y_{j_1} ', avec $j_1 \geqslant 2$, g(y_{j_1-1}) = α_0, o[g(y_{k-1})] Γ y_k pour $1 \leqslant k \leqslant j_1-1$, en posant $y_0 = \delta$ (d'après AT 1, AT 3) ; c'est le résultat d'un calcul complet d'entrée 'α_0 σ'.
- lorsque $1 \leqslant i \leqslant p-1$, v_i est de la forme '$x_{r'}$ ... x_1' avec $r' \geqslant 1$; j_i+1 est une sortie, j_{i+1} une entrée. μ_i est donc de la forme $\sigma^{r''}$ $y_{r''}$...y_{r-1} : g(y_{r-1}) = α_i; $1 \leqslant r'' < r$, $r'' \leqslant r'$; d'après AT 5, AT 8, pour $1 \leqslant k \leqslant r''-1$, $x_k \in J''$ et g(x_k) fn g(x_{k+1}) ; d'après AT 4, $x_{r''}$ Γ $y_{r''}$; pour $r''+1 \leqslant k \leqslant r-1$, o[g($y_{k-1}$)] Γ y_k, d'après AT 3. μ_i est le résultat d'un calcul complet d'entrée α_i $\tilde{v}_i$ σ.

- $v_p = `\delta\, x_{r'-1} \ldots x_1$' où $r' \geqslant 2$; $\mu_p = \sigma^{r'}$ et pour $1 \leqslant k \leqslant r'-1$, $x_k \in J''$ et $g(x_k)$ fng(x_{k+1}) (AT5, AT8). μ_p est le résultat d'un calcul complet d'entrée $\sigma\, \tilde{v}_p\, \sigma$. U_1 appartient à gn(α).

Théorème 6. 2. *Pour que le transducteur* tr *vérifie la condition* GP2, *il faut et il suffit que le graphe des initiales de la grammaire, supposée réduite, ne contienne pas de circuit (condition* AD*).*

La condition est suffisante, car $o[g(x)]\ \Gamma\ y$ entraîne qu'il existe un mot φ tel que $g(x) ::= g(y)\, \varphi$. Lorsque le graphe des initiales contient un circuit et que la grammaire est réduite, on voit par une démonstration analogue à celle donnée pour les piles adjointes (6. 1), qu'il n'existe aucun générateur des piles attachées aux pseudo-arborescences de $\mathfrak{C}(G)$ ou $\mathfrak{C}_1(G)$ pour lequel, si ε_i est l'entrée de la i ème feuille, $\frac{\varepsilon_i}{j_i}$ est borné ; ce qui achève de démontrer et étend le théorème.

Revenons au cas où le graphe des initiales ne contient pas de circuit. L'ensemble $m_2(\alpha, x)$ des mots '$y_0 \ldots y_n$' sur J_1 tels que $n > 0$, $y_0 = x$, $g(y_n) = \alpha$ et $o[g(y_{i-1})]\ \Gamma\ y_i$ pour $1 \leqslant i \leqslant n$, est fini. Les résultats des calculs complets du transducteur ne changent pas si on remplace la définition de l'ensemble $m(\alpha, x)$, pour $\alpha \in T$ et $x \in J_1$, donnée plus haut, par $m'(\alpha, x) = \{s_f\} \times m_2(\alpha, x)$ (cf 5. 3. 2, remarque). Les ensembles $m'(\alpha, x)$ peuvent être déterminés une fois pour toutes ; c'est ainsi que procède [36] qu'on pourra consulter pour obtenir plus de détails sur la pratique de la méthode.

6. 3. 3. Etude du déterminisme du transducteur. Pour tout couple (s, x') de $M \times E$, $d(s, x')$ ou $m(s)$ est vide. Si le transducteur n'est pas déterministe, l'un de ces ensembles a plusieurs éléments, ce qui ne peut se produire que dans deux cas :

1) Il existe dans J trois points w, y_1, y_2 tels que $w\ \Gamma\ y_1$, $w\ \Gamma\ y_2$, $y_1 \neq y_2$, $g(y_1)$ et $g(y_2)$ ont une initiale commune. D'après la propriété PG1 du pseudo-graphe de production, $B_1 = g(y_1)$ est distinct de $B_2 = g(y_2)$; d'après PG3, le chemin $w\, y_1$ du graphe (J, Γ) est contenu dans un chemin $\lambda\, w\, y_1\, \lambda_1$ d'origine dans J', d'extrémité dans J'' ; le chemin $\lambda\, w\, y_2$ est contenu dans un chemin $\lambda\, w\, y_2\, \lambda_2$ d'extrémité dans J''. g transcrit $\lambda\, w\, y_1\, \lambda_1$ en un mot $A\, \Psi\, B_1\, \varphi_1$ et $\lambda\, w\, y_2\, \lambda_2$ en $A\, \Psi\, B_2\, \varphi_2$: $A ::= \Psi\, B_1\, \varphi_1$, $A ::= \Psi\, B_2\, \varphi_2$.
2) Il existe trois points de J, x, y, x', et un symbole terminal α tels que $x\ \Gamma\ y$, α est initiale de $g(y)$, x' est l'origine dans J' et x l'extrémité dans J'' d'un chemin $x'\, \lambda\, x$ de (J, Γ), et $g(x'), \alpha$ est un couple vicinal. Alors $x'\, \lambda\, x\, y$ est un chemin de (J, Γ), d'origine dans J' ; il est contenu dans un chemin $x'\, \lambda\, x\, y\, \lambda_1$ d'extrémité dans J'' ; g transcrit ce dernier chemin en $g(x')\, \varphi\, g(x)\, g(y)\, \varphi_1$: $g(x') ::= \varphi\, g(x)$ et $g(x') ::= \varphi\, g(x)\, g(y)\, \varphi_1$

Soient un symbole non terminal A et un mot Ψ (éventuellement vide) sur V, tels qu'il existe des mots ω pour lesquels $A ::= \Psi\, \omega$; désignons par $B_1, \ldots, B_p$ les premiers symboles des mots ω non vides et par $Q_k(A, \Psi)$ l'ensemble des initiales terminales de B_k ; enfin, si l'un des mots ω est vide, soit $Q_0(A, \Psi)$ l'ensemble des symboles terminaux α tels que le couple A, α soit vicinal. Pour que le transducteur soit déterministe, il suffit que soit vérifiée :

(AT) *pour tout couple* A, Ψ *les ensembles* $Q_i(A, \Psi)$ *sont deux à deux disjoints.*

AD' entraîne AT ; si la grammaire est réduite, AT entraîne AD (même démonstration que pour AD' entraîne AD, 6. 1).

Supposons AT non satisfaite et la grammaire réduite. Si pour un couple A, Ψ deux ensembles $Q_i(A, \Psi)$ et $Q_k(A, \Psi)$ ($i \neq 0$, $k \neq 0$, $i \neq k$) ne sont pas disjoints, il existe dans $\mathcal{S}(V, ::=)$ deux pseudo-arborescences (I, f) et (I', f') de racine A dont les suites de feuilles ont un facteur gauche commun γ non vide ; si $Q_0(A, \Psi)$ et $Q_k(A, \Psi)$ ont en commun un élément α, d'après le théorème 4. 3, il existe un symbole auxiliaire A' et deux mots φ_1, φ_2 sur V tels que $\varphi_1\, A\, \alpha\, \varphi_2$ dérive de A', d'où deux pseudo-arborescences (I, f) et (I', f') de racine A', dont les suites de feuilles ont un facteur gauche commun γ non vide ; dans les deux cas, les sous-pseudo-arborescences de (I, f) et (I', f'), images des sous-arborescences de I et I' qui sont formées des noeuds majorés par l'une des $|\gamma|$ premières feuilles étant différentes. On en déduit de la même manière qu'au paragraphe 6. 1 :

Théorème 6. 3. *Soit une grammaire* G *réduite. Pour que le transducteur* tr *soit déterministe, il faut et il suffit que* G *vérifie la condition* AT. *Lorsque* AT *n'est pas satisfaite, tout générateur des piles attachées aux pseudo-arborescences de* $\mathfrak{C}(G)$ *(ou de* $\mathfrak{C}_1(G)$*), pour lequel la donnée est la suite de feuilles de la pseudo-arborescence (ou sa transcrite par l'application* g*),* cd = 1 *et* j_i *est l'entrée de la* i ème *feuille, est un générateur de piles multiples.*

On remarquera que pour prouver plus haut que toute pile engendrée était l'une des piles cherchées, on n'a pas utilisé toutes les conditions imposées au transducteur. Il existe donc des transducteurs plus simples conduisant aux mêmes piles ; ils sont seulement "plus indéterministes", et de ce fait mènent à des réalisations pratiques moins efficaces ; en particulier, ils peuvent n'être pas déterministes même lorsque AT est vérifiée. Une remarque analogue peut être faite pour les autres algorithmes d'analyse que nous étudions ; elle sera illustrée par les divers générateurs qui, au chapitre 7, engendreront tous les mêmes piles.

6. 3. 4. Exemples. Un langage de Dyck sur 2n lettres a_i ($1 \leqslant i \leqslant n$ et $-n \leqslant i \leqslant -1$) [15] peut être défini par la grammaire suivante :

$T = \{ a_i ; 1 \leqslant i \leqslant n, -n \leqslant i \leqslant -1 \}$;

$N = \{ X, A_i, B_i ; 1 \leqslant i \leqslant n, -n \leqslant i \leqslant -1 \}$; axiome X ;

$$\left.\begin{array}{l} A_i ::= a_i B_i a_{-i} \mid a_i a_{-i} \\ B_i ::= A_j B_i \mid A_j \\ X ::= A_i X \mid A_i \end{array}\right\} \begin{array}{l} \text{pour } 1 \leqslant i \leqslant n \text{ et } -n \leqslant i \leqslant -1, \\ 1 \leqslant j \leqslant n \text{ et } -n \leqslant j \leqslant -1, \\ \text{et } j \neq -i. \end{array}$$

Il est facile de voir que la condition $j \neq -i$ assure AT. Rappelons que tout langage de Chomsky est image homomorphe de l'intersection d'un langage de Dyck avec un langage de Kleene [54]. La relation de production donnée en exemple au paragraphe 4. 1. 2 vérifie aussi (AT). D'autres exemples pourront être trouvés dans [25], qui étudie des cas où l'hypothèse (AT) est réalisée.

Certains auteurs n'utilisent pas de pseudo-graphe de production, et le remplacent par un jeu de procédures, récursives en général, une procédure étant associée à certains symboles auxiliaires (ces procédures jouent parfois un rôle en dehors de l'analyse syntaxique) : [28], [42] étudient ainsi Algol. Cependant, pour éviter d'avoir à introduire un trop grand nombre de procédures ou un trop grand nombre d'éléments ("états syntaxiques") dans l'alphabet utilisé, ces auteurs transforment souvent implicitement la grammaire donnée en une grammaire généralisée équivalente.

6. 4. Analyse par construction des piles attachées aux pseudo-arborescences gauches des pseudo-arborescences de $\mathcal{C}(G)$.

6. 4. 1. Introduction. L'hypothèse AD est fort restrictive : en particulier, elle interdit les règles de la forme $A ::= A\varphi$. Son rôle est d'assurer qu'un état u_i de la pile attachée à une pseudo-arborescence de $\mathcal{C}(G)$ étant donné, la première entrée après i d'un symbole terminal, si elle existe, ou l'indice du dernier état de la pile sinon, est borné. Cette condition sera plus facile à réaliser pour les piles attachées aux pseudo-arborescences gauches (2. 12) des pseudo-arborescences de $\mathcal{C}(G)$: toute famille commence par une feuille ; plus précisément, d'après l'étude faite en 2. 12. 4, il suffit qu'il n'existe aucune suite de symboles auxiliaires $A_1, \ldots, A_n$ tels que $A_1 = A_n$ et, pour $2 \leqslant k \leqslant n$, $A_k ::= A_{k-1}$, autrement dit, qu'aucun symbole auxiliaire ne dérive strictement de lui-même.

L'étude des familles d'une ramification gauche conduit à penser qu'il est plus commode de disposer d'une règle $A ::= \varphi$ sous la forme d'un mot $\varphi\tilde{A}$ plutôt que $A\varphi$; nous utiliserons donc un pseudo-graphe de production $(J, \Gamma ; g)$ du deuxième type. Nous construirons les piles attachées aux pseudo-arborescences gauches des pseudo-arborescences (I, f_1) de $\mathcal{C}_1(G)$. D'après le théorème 3. 3, (I, f_1) est déterminée de manière unique par sa pseudo-arborescence gauche.

Dans le cas déjà envisagé d'une grammaire où toute règle est du type $A ::= a\varphi$, $a \in T$, $\varphi \in N^*$ (ou V^*), toute pseudo-arborescence de $\mathcal{C}_1(G)$ est sa propre pseudo-arborescence gauche, puisque chacune de ses familles commence par une feuille.

6. 4. 2. Caractérisation des piles cherchées. D'après l'étude des ramifications gauches (2. 12) et celle de la pile strictement attachée à une ramification orientée (2. 1), la pile U_1 attachée à $((I_G, f_1))$, pseudo-arborescence gauche de $(I, f_1) \in \mathcal{C}_1(N, T, ::=, X)$ possède les propriétés suivantes :

(G1) La première entrée est une entrée de δ.

(G2) Si i est une entrée de x et i+1 une entrée de y, y appartient à J', g(x) appartient à N et g(y) à T, car dans $((I_G, f_1))$ y est le premier élément d'une famille de prédécesseur x ; y est le premier élément d'une famille de (I, f_1) et l'image d'une feuille de I.

(G3) Si i est une sortie de x et i+1 une entrée de y, deux cas sont possibles (et ils s'excluent) :
  a) $x \Gamma y$ (d'où $x \neq \delta$, $y \neq \delta$) et $g(y) \in V$; si $g(y) \in N$, $i + 2$ est une entrée ;
  b) $y \in J'$ et il existe $z \in J$ tel que $x \Gamma z$ et $g(z) = \overline{g(y)} \in \overline{N}$; $i + 2$ est une sortie de y.

En effet, x et y sont consécutifs dans une famille de $((I_G, f_1))$, et ce qui précède résulte de l'étude des familles d'une ramification gauche : dans le cas a, x et y sont consécutifs dans une famille de (I, f_1) ; dans le cas b, x est l'image par f_1 du dernier point d'une famille de I de prédécesseur y_1, $f_1(y_1) = y$ et y_1 est le premier point de sa famille.

(G4) Si i est une sortie de x et i+1 une sortie de y, il existe z tel que $x \Gamma z$ et $g(z) = \overline{g(y)} \in \overline{N}$ (d'où $x \neq \delta$ et $z \in J''$), car x est le dernier élément d'une famille de $((I_G, f_1))$ et d'une famille de (I, f_1), dont le prédécesseur est y.

Réciproquement, soit U_1 une pile sur J_1 qui vérifie G1, G2, G3, G4. Elle est attachée à une pseudo-ramification (I', f') dans J_1 ; les points w de I' tel que $g \circ f'(w) \in T$ sont des feuilles, car, d'après G2, $f'(w)$ sort de la pile dès son entrée ; toute famille de I' commence par un tel point (G2) ; d'après le théorème 3. 3, (I', f') est la pseudo-ramification gauche d'une unique pseudo-ramification (I, f_1) dans J_1 telle que les feuilles de I soient ceux de ses points w_1 qui vérifient $g \circ f_1(w_1) \in T$. D'après G1, G2 et G3, pour toute entrée dans U_1 d'un élément y, g(y) appartient à V : g transforme (I, f_1) en une pseudo-ramification (I, f) dans V ; g transforme aussi (I', f') en la pseudo-ramification gauche de (I, f) et transcrit la pile U_1 en la pile attachée à cette pseudo-ramification gauche. D'après G1, G2, G3, δ n'a qu'une entrée dans U_1, la première ; d'après G3 et G4, la sortie de δ est nécessairement la dernière, (I', f') a une seule racine δ, donc (I, f_1) a une seule racine δ, (I, f) a une seule racine X. Etudions les familles de (I, f_1) : d'après 2. 12. 2 et 3. 2. 3, pour les construire, on considère dans toute famille ω de I', de prédécesseur z, les feuilles $a_1, \ldots, a_{p-1}$ telles que $g \circ f'(a_i) \in N$ (si elles existent) : les a_i sont caractérisés par : $g \circ f'(a_i) \in N$ et, dans la pile U' strictement attachée à I', l'entrée de α_i est suivie d'une sortie ; ω s'écrit $\alpha_0 \, \omega_0 \, \alpha_1 \, \omega_1 \dots \alpha_{p-1} \, \omega_{p-1}$ où $g \circ f'(\alpha_0) \in T$; si un élément x d'un mot ω_i vérifie $g \circ f'(x) \in N$, son entrée est suivie d'une entrée ; $f'^*(\omega) = \alpha'_0 \, \omega'_0 \, \alpha'_1 \, \omega'_1 \dots \alpha'_{p-1} \, \omega'_{p-1}$ est une famille de (I', f') : $\alpha'_0 \, \omega'_0$ est une famille de (I, f_1) qui a pour prédécesseur α'_1, $\alpha'_1 \, \omega'_1$ une famille de prédécesseur $\alpha'_2, \dots, \alpha'_{p-1} \, \omega'_{p-1}$ une famille de prédécesseur $f'(z)$ et on obtient ainsi toutes les familles de (I, f_1) sauf 'δ'. Pour $1 \leqslant i \leqslant p-1$ l'entrée dans U' de α_i est précédée et suivie d'une sortie : d'après G3, $\alpha'_i \in J'$; d'autre part, d'après G2, $\alpha'_0 \in J'$. Pour $0 \leqslant i \leqslant p-1$, après la sortie d'un élément x de $\alpha_i \, \omega_i$ qui n'est pas le dernier vient l'entrée de l'élément suivant y ; d'après G3, $f'(x) \,\Gamma\, f'(y)$. Pour $0 \leqslant i \leqslant p-2$, la sortie du dernier élément x de $\alpha_i \, \omega_i$ est suivie de l'entrée, puis de la sortie, de α_{i+1} : il existe y tel que $f'(x) \,\Gamma\, f'(y)$ et $g \circ f'(y) = \overline{g \circ f'(\alpha_{i+1})}$ (G3). La sortie du dernier élément x de $\alpha_{p-1} \, \omega_{p-1}$ est suivie de la sortie de z : il existe y tel que $f'(x) \,\Gamma\, f'(y)$ et $g \circ f'(y) = \overline{g \circ f'(z)}$ (G4). $g^*(\alpha'_i \, \omega'_i) \, \overline{g(\alpha'_{i+1})}$, pour $0 \leqslant i \leqslant p-2$, et $g^*(\alpha'_{p-1} \, \omega'_{p-1}) \, \overline{g[f'(z)]}$ sont des chemins du pseudo-graphe de production, d'origine dans J', d'extrémité dans J''. (I, f) est une pseudo-arborescence de $\mathcal{G}(N, T, ::=, X)$.

Nous allons préciser G1, G2, G3, G4 par d'autres propriétés de la pile U_1 attachée à $((I_G, f_1))$ pseudo-arborescence gauche de $(I, f_1) \in \mathcal{T}_1(G)$. Notons d'abord que l'entrée de δ est suivie de l'entrée d'un élément b tel que $g(b) \in T$. Désignons par U' la pile strictement attachée à I_G, arborescence gauche de I ; soient x_1, y_1 deux points de I (et I_G), b_1 une feuille de I, $x = f_1(x_1)$, $y = f_1(y_1)$, $b = f_1(b_1)$.

Supposons d'abord que l'entrée de x_1 dans U' soit suivie immédiatement de l'entrée de y_1 : y_1 est la première feuille qui majore x_1 dans l'ordre d'entrée de I ; d'après 4. 3. 1, $f(y_1) = g(y)$ est une initiale de $f(x_1) = g(x)$:

(G5) Si i est une entrée de x dans U_1 et i+1 une entrée de y, g(y) est initiale de g(x).

Supposons maintenant que b_1 soit la première feuille de I qui sort de U' après y_1, donc qui entre après la sortie de y_1 : comme I et I_G ont le même ordre de sortie, y_1 et b_1 forment un couple de points consécutifs de I ; d'après 4. 3. 2, le couple g(y), g(b) est vicinal. Si après y_1 ne sort plus aucune feuille de I, les points qui sortent après y_1 sont tous les derniers de leur famille (2. 3. 3), et g(y) est une finale de l'axiome X :

(G6) S'il existe une feuille de (I, f_1) qui entre dans U_1 après une sortie de y, et si b est la première de ces feuilles, le couple g(y), g(b) est vicinal ; sinon g(y) est une finale de X.

Supposons ensuite que la sortie i de x_1 soit suivie par l'entrée i+1 de y_1 et que b_1 soit la première feuille de I dont l'entrée est supérieure à i, si elle existe ; étudions les deux cas qui, d'après (G3) peuvent se présenter :

cas a) si y_1 est une feuille, $y_1 = b_1$, $y = b$; sinon, i+2 est une entrée de b dans U_1 (G2) et g(b) est initiale de g(y) (G5).
cas b) y_1 sort dès son entrée, y_1 est une feuille de I_G et un noeud de I. Désignons par R l'ordre associé à l'arborescence I, par e_1 le sommet de l'état u'_i de la pile U', et posons $e = f_1(e_1)$. D'après 2. 12. 2, il existe un chemin de I, $e_1 = w_0, w_1 \dots w_q, y_1 = w_{q+1}$, où $w_1, \dots, w_q, y_1$ sont les premiers de leur famille (mais pas e_1). Si aucune feuille de I n'entre dans U' après y_1, aucune ne sort après $w_q, \dots, w_1, e_1$, ni après les autres éléments de u'_i : chacun d'eux est le dernier de sa famille et $w_1, \dots, w_q, y_1$ sont seuls dans leur famille ; $g(e) = f(e_1)$ est une finale de X, $g(y) = f(y_1)$ dérive strictement de g(e). Si une feuille de I entre après y_1, y_1 et b_1 sont des points consécutifs de I. Deux possibilités s'offrent alors (figure 6. 1) :
i) $e_1 \bar{R} b_1$: dans l'ordre de sortie de I, l'intervalle $R(e_1)$ contient $R(y_1)$ et précède b_1 ; d'après l'étude faite en 2. 6. 3, e_1 et b_1 sont aussi consécutifs et de plus, $w_1, \dots, w_q, y_1$ sont derniers de leur famille, donc seuls dans leur famille : g(y) dérive strictement de g(e) ; d'autre part, g(e), g(b) est un couple vicinal.
ii) $e_1 R b_1$: le plus grand R-minorant commun à y_1 et b_1 est l'un des points $w_0, \dots, w_q$, appelons le w_j ; si $j \neq q$, d'après 2. 6. 3, $w_{j+2}, \dots, w_q, y_1$ sont derniers de leur famille : $f(y_1) = g(y)$ dérive de $f(w_{j+1})$; il existe une règle $f(w_j) ::= f(w_{j+1}) \, Z \, \varphi$ telle que g(b) soit initiale de Z. D'autre part, $f(w_j)$ est une initiale de $f(e_1) = g(e)$.

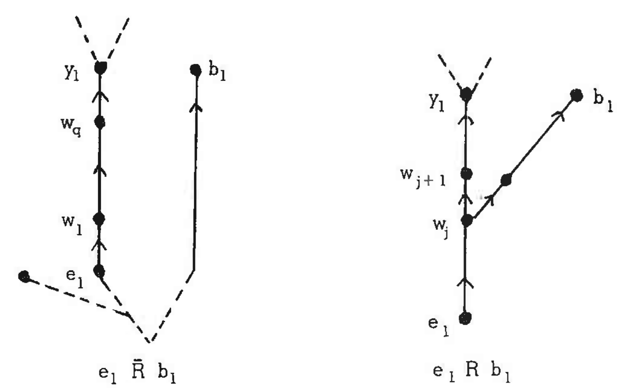


Figure 6. 1.

Définition. *Les propositions en* $A \in N$, $B \in N$, $\alpha \in T$:
- A *dérive strictement de* B *et le couple* B, α *est vicinal,*
- *il existe* A', B', Z *appartenant à* V *et* φ *appartenant à* V* *tels que* B' ::= A' Z φ, A *dérive de* A', α *soit initiale de* Z *et* B' *initiale de* B[1],

sont décidables. Lorsque l'une d'elles est vraie, nous disons que A, α *est un* couple initial *de* B.

De l'étude qui précède résulte la propriété de la pile U_1 :

(G7) Si i est une sortie de x et i+1 une entrée de y :
- dans le cas (a) de (G3), y est l'image d'une feuille de I, ou bien i+2 est entrée de b, image d'une feuille de I et g(b) est initiale de g(y) ;
- dans le cas (b) de (G3), e étant le sommet de la pile après la sortie de x, si aucun b pour lequel $g(b) \in T$ n'a d'entrée supérieure à i, $g(y) \in N$ dérive strictement de g(e), finale de X ; si au contraire b est le premier élément pour lequel $g(b) \in T$, dont une entrée est supérieure à i, g(y), g(b) est un couple initial de g(e).

Pour simplifier la description du générateur de piles qui sera étudié plus bas, *lorsque* $A \in N$ *dérive strictement d'une finale* B *de* X, *nous dirons encore que* A, σ *est un couple initial de* B.

Supposons enfin que i soit une sortie de x_1 et i+1 une sortie de y_1 : y_1 est un noeud de la ramification gauche I_G, donc n'est pas le premier de sa famille dans I. Si après ces deux sorties aucune feuille de I n'entre plus dans la pile U', y_1 et les points qui sortent après lui sont les derniers de leur famille ; donc y_1 est la racine de I, $y = \delta$, ou bien il existe une règle B ::= $\varphi g(y)$ où $\varphi \neq \Lambda$ et B est une finale de X. Si une feuille de I entre après la sortie de y et si b_1 est la première telle feuille, y_1 et b_1 sont des points consécutifs de I ; il existe une règle B ::= φ A' Z φ_1 où $f(b_1) = g(b)$ est initiale de Z et ou bien A' = g(y) et $\varphi \neq \Lambda$, ou bien il existe une règle A'' ::= $\varphi_2 g(y)$ où A'' est une finale de A' et $\varphi_2 \neq \Lambda$.

Définition. *Si, pour deux éléments* $A \in N$, $\alpha \in T$, *il existe* B, A', Z *appartenant à* V, φ *et* φ_1 *appartenant à* V* *tels que* B ::= φ A' Z φ_1, α *est initiale de* Z *et soit* A' = A *et* $\varphi \neq \Lambda$, *soit une finale* A'' *de* A' *et un mot non vide* φ_2 *de* V* *vérifient* A'' ::= φ_2 A, *nous dirons que le couple* A, α, *qui est vicinal, est un* couple vicinal post-initial.

D'où pour la pile U_1 :

(G8) Si i est une sortie de x et i+1 une sortie de y, si b est le premier élément pour lequel $g(b) \in T$ dont une entrée est supérieure à i, g(y), g(b) est un couple vicinal post-initial ; si aucun b tel que $g(b) \in T$ n'a d'entrée supérieure à i, $y = \delta$ ou une finale B de X et un mot φ non vide vérifient B ::= $\varphi g(y)$ avec $g(y) \in N$.

(1) Cette proposition équivaut à : il existe $\Psi \in V^*$ tel que A α Ψ dérive de B.

Lorsque A est une finale stricte de X, nous conviendrons de dire que le couple A, σ est vicinal. *Lorsque pour un symbole auxiliaire* A, *une finale* B *de* X *et un mot* φ *non vide vérifient* B ::= φ A, *nous dirons encore que* A, σ *est un couple vicinal post-initial.*

6. 4. 3. Générateur. Cette étude va nous permettre de définir un générateur de piles gn tel que, pour tout mot α de T^*_1 les piles de l'ensemble gn(α) soient transcrites par g en les piles attachées aux pseudo-arborescences gauches des pseudo-arborescences de $\mathcal{C}$(N, T, ::=, X) qui ont α pour suite de feuilles. Pour soutenir la compréhension, la figure 6. 2 schématise la construction de ces piles ; les références entre parenthèses sont celles qui justifient ses diverses parties. Ici encore, les j_i sont les entrées des éléments dont l'image par g est un symbole terminal ; le générateur est donné par $W = T \cup J_1$, cd = 1 et par le transducteur tr qui va être défini ; rappelons que $T' = T \cup \{\sigma\}$, $J_1 = J \cup \{\delta\}$ et posons $J'_1 = J_1 \cup \{\sigma\}$:

a) $M = T' \times J'_1 \times J'_1$.

b) q = 2 ; $l(\alpha, e) = (\alpha, e, \sigma)$ pour $\alpha \in T'$ et $e \in J'_1$.

c) Pour $\alpha \in T'$, $e \in J$, $e' \in J_1$, d[(α, e, σ), e'] contient le seul triplet (α, e', e) si g(e), α est un couple vicinal ;

pour $\alpha \in T'$, $e \in J$, $x \in J$, $e' \in J_1$, d[(α, e, x), e'] contient le seul triplet (α, e', e) s'il existe $w \in \Gamma(x)$ tel que g(w) = g(e) et si g(e), α est un couple vicinal post-initial ;

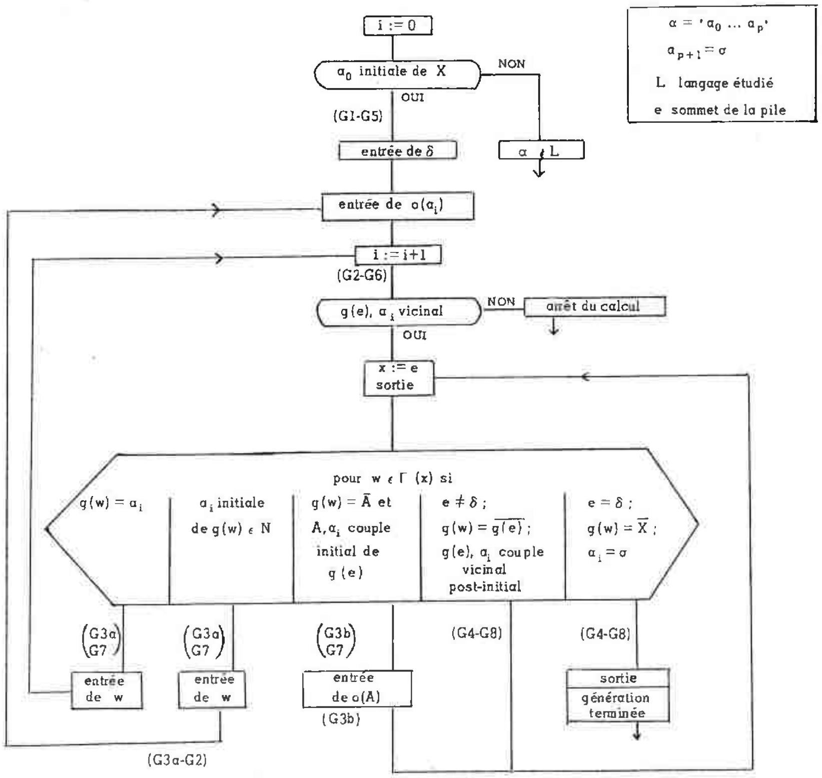
pour $x \in J$, d[(σ, δ, x), σ] contient le seul élément s_f s'il existe $w \in \Gamma(x)$ tel que g(w) = $\bar{X}$;

d) Pour $\alpha \in T$, m(α, σ, σ) contient le seul couple (s_f,' δ o(α)') si α est une initiale de X ;

pour $\alpha \in T'$, $e \in J'_1$, $x \in J$, m(α, e, x) contient les couples suivants et eux seuls:
- (s_f,' w') si $w \in \Gamma(x)$ et g(w) = α,
- (s_f,' w o(α)') si $w \in \Gamma(x)$, $g(w) \in N$ et α est initiale de g(w),
- [(α, e, o(A)), ' o(A) σ'] s'il existe $w \in \Gamma(x)$ tel que g(w) = $\bar{A}$ et A, α soit un couple initial de g(e).

La condition GP 2 est-elle réalisée ? Envisageons une suite $s_1, \ldots, s_n$ d'éléments de M tels que $s_1 = s_n$ et, pour i = 2, ..., n, $s_i = m_1(s_{i-1})$. Cette suite est nécessairement du type (α, e, x_1), ..., (α, e, x_n) où $x_1 = x_n$; $x_2 = o(A_2)$, ..., $x_n = o(A_n)$ appartiennent à l'ensemble J' des origines des chemins qui représentent les règles de la grammaire ; pour $1 \leqslant i \leqslant n-1$, $g(x_i) = A_i$ est le dernier élément du second membre d'une règle de premier membre A_{i+1} : cette règle est donc A_{i+1} ::= A_i. Il en résulte que A_1 dérive strictement de lui-même. On retrouve la condition prévue en 6. 4. 1 : pour que GP 2 soit réalisée, il suffit que :

(DA) *aucun symbole auxiliaire ne dérive strictement de lui-même.*



Certains choix peuvent conduire à un nombre d'étapes suivantes différent de 1 (lorsque ce nombre est 0, le calcul s'arrête). Dans toutes les figures schématisant un générateur de piles, nous représentons ces choix par ⬡, pour les distinguer des tests, qui conduisent dans tous les cas à une et une seule étape suivante.

Figure 6. 2

Si DA n'est pas vérifiée, il existe dans $\mathcal{Y}$(V, ::=) une infinité de pseudo-arborescences de racine A, de seule feuille A ; si on suppose de plus la grammaire réduite, $\mathcal{Y}$(V, ::=) contient une infinité de pseudo-arborescences de racine A ayant la même suite de feuilles terminales β, et d'après le théorème 4. 2, $\mathcal{C}$(N, T, ::=, X) contient une infinité de pseudo-arborescences qui ont la même suite de feuilles.

Théorème 6. 4. *Pour que* tr *vérifie la condition* GP 2, *il suffit que la grammaire vérifie* DA. *Si une grammaire réduite* G *ne satisfait pas à* DA, $\mathcal{C}$(G) *contient une infinité de pseudo-arborescences ayant la même suite de feuilles ; aucun générateur de piles ne peut engendrer pour toute donnée* α *un ensemble de piles en correspondance biunivoque avec l'ensemble des pseudo-arborescences de* $\mathcal{C}$(G) *qui ont* α *pour suite de feuilles.*

En suivant la même démarche qu'au paragraphe 6.3.2, on démontre que pour tout mot α sur T, les piles de gn(α) sont transcrites par g en les piles attachées aux pseudo-arborescences gauches des pseudo-arborescences de $\mathcal{C}$(G) dont α est la suite de feuilles.

6. 4. 4. Etude du déterminisme du transducteur. Pour tout couple (s, z) de M $\times$ E, d(s, z) contient un élément au plus. d(s, z) et m(s) sont tous deux non vides dans trois cas :

1) s = (a, e, x) où $a \in T'$, $e \in J$, $x \in J$; il existe $w \in \Gamma(x)$ tel que $g(w) = \overline{g(e)}$ et g(e), a est un couple vicinal post-initial ; il existe $w' \in \Gamma(x)$ tel que a soit initiale de g(w') (avec comme cas particulier g(w') = a) ; z est quelconque dans J_1. *Il existe donc trois éléments de* V, A, A', Z, *un élément* a *de* T' *et deux mots* φ *et* φ_1 *de* V^* *tels que* A ::= φ, A' ::= φ Z φ_1, A, a *est un couple vicinal post-initial,* a *est une initiale de* Z.

2) s = (a, e, x) où $a \in T'$, $e \in J$, $x \in J$; il existe $w \in \Gamma(x)$ tel que $g(w) = \overline{g(e)}$ et g(e), a est un couple vicinal post-initial ; il existe $w' \in \Gamma(x)$ et $A' \in N$ tels que $g(w') = \bar{A}'$ et A', a est un couple initial de g(e) ; $z \in J_1$. *Il existe donc deux symboles non terminaux* A, A', *un* a *de* T' *et un mot* φ *de* V^* *tels que* A ::= φ, A' ::= φ, A, a *est un couple vicinal post-initial,* A', a *est un couple initial de* A.

3) s = (σ, δ, x) où $x \in J$; il existe $w \in \Gamma(x)$ tel que $g(w) = \bar{X}$; il existe $w' \in \Gamma(x)$ et $A \in N$ tels que $g(w) = \bar{A}$ et A dérive strictement de X (d'après l'extension de la notion de couple initial) ; z = σ. *Il existe donc un symbole non terminal* A *et un mot* φ *de* V^* *tels que* A ::= φ, X ::= φ, A *dérive strictement de* X (d'où A $\neq$ X d'après DA).

m(s) contient plusieurs éléments dans trois cas où s = (a, e, x) avec $a \in T'$, $e \in J'_1$, $x \in J$:

4) Il existe deux points w et w' de Γ (x) tels que a soit initiale de g(w) (avec comme cas particulier g(w) = a) et de g(w'). *Il existe donc quatre éléments* A, A', Z, Z' *de* V *et trois mots* φ, φ_1, φ'_1 *de* V^* *tels que* $A ::= \varphi Z \varphi_1$, $A' ::= \varphi Z' \varphi'_1$, $\varphi \neq \Lambda$, $Z \neq Z'$, Z *et* Z' *ont une initiale terminale commune.*

5) Il existe $w \in \Gamma$ (x) et $A \in N$ tels que $g(w) = \bar{A}$ et A, a est un couple initial de g(e) ; il existe $w' \in \Gamma$ (x) tel que a soit initiale de g(w'). *Il existe donc trois éléments* A, A', Z *de* V, *un élément* a *de* T' *et deux mots* φ *et* φ_1 *de* V^* *tels que* $A ::= \varphi$, $A' ::= \varphi Z \varphi_1$, A, a *est couple initial d'un symbole non terminal,* a *est initiale de* Z.

6) Il existe deux points w et w' de Γ (x) et deux symboles non terminaux A et A' tels que $g(w) = \bar{A}$, $g(w') = \bar{A}'$, A, a et A', a sont des couples initiaux de g(e). *Il existe donc trois symboles auxiliaires* A, A', B, *un terminal* a *et un mot* φ *de* V^* *tels que* $A ::= \varphi$, $A' ::= \varphi$, $A \neq A'$, A, a et A', a *sont deux couples initiaux de* B.

De la définition d'un couple initial d'un symbole auxiliaire résulte que tout tel couple est vicinal. Réciproquement, soit $A \in N$, $a \in T'$ un couple vicinal qui ne soit pas post-initial. Si $a \neq \sigma$, d'après 4.3.2, il existe une règle $B ::= \varphi A' Z \varphi_1$. A est une finale de A', a une initiale de Z ; si $A = A'$, $\varphi = \Lambda$ sinon A, a serait post-initial : A, a est un couple initial de B ; si $A \neq A'$, A est une finale stricte de A', il existe une finale A'' de A' et un mot φ_2 tels que $A'' ::= \varphi_2 A$; comme le couple A, a n'est pas post-initial, $\varphi_2 = \Lambda$; d'autre part A'', a est un couple vicinal : A, a est un couple initial de A''. Si $a = \sigma$, A est une finale stricte de X, il existe une finale B de X et un mot φ tels que $B ::= \varphi A$; $\varphi = \Lambda$ et A, σ est un couple initial de B. La réunion de l'ensemble des couples vicinaux post-initiaux et de celui des couples initiaux des divers symboles de N est donc l'ensemble des couples vicinaux appartenant à $N \times T'$. On peut réunir les propriétés trouvées en 1 et 5 dans : *il existe trois éléments* A, A', Z *de* V, *un élément* a *de* T' *et deux mots* φ *et* φ_1 *de* V^* *tels que* $A ::= \varphi$, $A' ::= \varphi Z \varphi_1$, A, a *est un couple vicinal,* a *est une initiale de* Z.

Si la grammaire ne possède pas deux règles distinctes ayant le même second membre, les cas 3 et 6 ne peuvent se présenter ; dans le cas 2, il existe un symbole non terminal A et un a de T' tels que A, a soit à la fois un couple vicinal post-initial et un couple initial de A.

On peut voir, par une méthode analogue à celle qui a été employée en 6.1 et 6.3.3 que, réciproquement, dans le cas où l'une des propriétés signalées en 2, 3, 6 est vérifiée, tout générateur des piles cherchées où cd = 1 et où j_i est l'entrée du $i^{ème}$ élément dont l'image par g est terminale, est un générateur de piles multiples. Mais il n'en est plus de même lorsqu'est vérifiée une des propriétés 1, 5 ou 4. Par exemple, envisageons la grammaire :

$$N = \{X, A, A'\} ; T = \{a, c\} ; X ::= a A a \mid c A', A ::= c, A' ::= ca.$$

Elle engendre un langage à deux phrases a c a et c c a ; le couple A, a est vicinal et a est une initiale de a ; le transducteur n'est pas déterministe. Or le second membre de toute règle commence par un symbole terminal, les pseudo-arborescences de $\mathcal{C}_1$(N, T, ::= X) sont donc leurs propres pseudo-arborescences gauches ; leurs piles attachées sont engendrées par le générateur décrit en 6.3.2, générateur de pile unique puisque la grammaire vérifie la condition A T. On peut montrer que si la grammaire est réduite et si elle vérifie l'une des propriétés énoncées en 1, 5 ou 4, il existe une phrase α telle que, dans la génération de l'ensemble gn(α), l'un des calculs du transducteur possède un triplet (s, y, λ) pour lequel la réunion de m(s) et d(s, z) (pour tout z de J_1) contienne effectivement deux éléments.

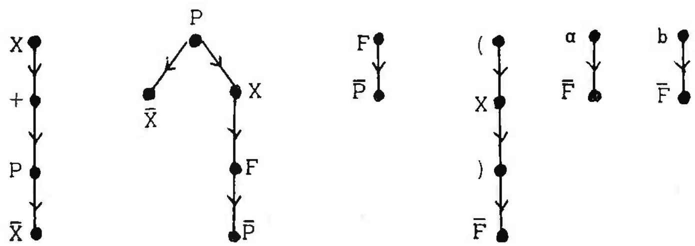
6. 4. 5. Etude d'un exemple. $N = \{X, P, F\}$; $T = \{+, \times, (,), a, b\}$;

$X ::= X + P \mid P$ $\qquad$ $P ::= P \times F \mid F$ $\qquad$ $F ::= (X) \mid a \mid b.$

Le pseudo-graphe de production du deuxième type (J, Γ ; g), qui est une pseudo-ramification, et sa matrice d'enchaînement sont représentés sur la figure 6.3. L'ensemble J est l'ensemble des numéros de ligne de la matrice, c'est-à-dire des entiers de 1 à 19. Dans la matrice, les points de la ramification (J, Γ) orientée par l'ordre de ces entiers, sont classés par ordre d'entrée, de sorte qu'on peut omettre la colonne qui indique le lien vertical : le lien de x est x+1 si $g(x) \in V$, il n'existe pas si $g(x) \in \bar{N}$. Le lien horizontal des racines est inutile ici et aurait pu être omis.

(, a, b sont les initiales terminales de X, P, F ; o (( ) = 12, o (a) = 16, o (b) = 18. Les couples vicinaux appartenant à $T \times T'$ sont donnés par la matrice :

	+	×	(	)	a	b	σ
+			.		.	.	
×			.		.	.	
(			.		.	.	
)	.	.		.			.
a	.	.		.			.
b	.	.		.			.



ligne x	g(x)	lien horizontal
1	X	5
2	+	-
3	P	-
4	$\bar{X}$	-
5	P	10
6	$\bar{X}$	7
7	×	-
8	F	-
9	$\bar{P}$	-
10	F	12
11	$\bar{P}$	-
12	(	16
13	X	-
14	)	-
15	$\bar{F}$	-
16	α	18
17	$\bar{F}$	-
18	b	-
19	$\bar{F}$	

Figure 6. 3

Les couples vicinaux appartenant à $N \times T'$ sont :

X + , P + , F + , P × , F × , X) , P) , F) , P σ , F σ.

La matrice suivante indique pour chacun d'eux, A, α, de quels symboles auxiliaires il est couple initial, o(A) qui ici existe toujours, et s'il est un couple vicinal post-initial :

	X	P	F	o(A)	post-initial
X +	.			1	
P +	.			5	.
F +	.	.		10	.
P ×	.	.		5	
F ×	.	.		10	.
X)				1	.
P)	.			5	.
F)	.	.		10	.
P σ	.			5	.
F σ	.	.		10	.

Aucun couple A, α n'est à la fois vicinal post-initial et couple initial de A ; d'autre part, X,× n'est pas un couple vicinal. Le transducteur est donc déterministe.

Pour la donnée $\alpha = (a + b) \times b$, on obtient la pile

															18		10																	
				16		10		5		1		2		3	3	3	3	3												18				
	12		13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13		14		10		5		7		10	10	10		5	
δ	δ	δ	δ	δ	δ	δ	δ	δ	δ	δ	δ	δ	δ	δ	δ	δ	δ	δ	δ	δ	δ	δ	δ	δ	δ	δ	δ	δ	δ	δ	δ	δ	δ	δ

Cette pile est transcrite par g en

															b		F																	
				α		F		P		X		+		P	P	P	P	P												b				
	(		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		)		F		P		×		F	F	F		P	
X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

Cette dernière pile est attachée à la pseudo-ramification gauche de la pseudo-ramification de $\mathcal{C}(N, T, ::=, X)$ dont α est la suite de feuilles ; ces deux pseudo-ramifications sont représentées sur la figure 6. 4.

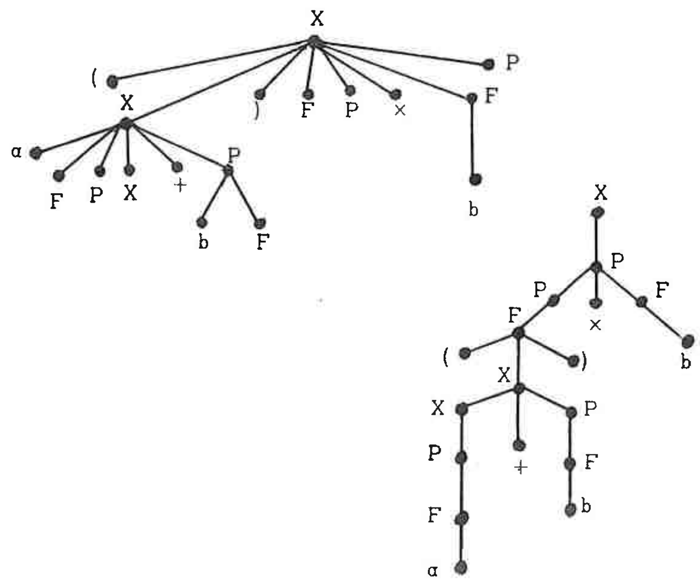


Figure 6. 4

CHAPITRE VII

ANALYSE PAR CONSTRUCTION D'UNE PILE CONJOINTE

7. 1. Grammaires pluriaxiomatiques.

En vue de transformations que nous ferons subir aux grammaires en 7. 10 et 7. 11, nous introduirons une légère généralisation de la notion de grammaire. Soit une structure de Chomsky de vocabulaire $V = T \cup N$, de relation de production $::=$ et un sous-ensemble N' de son vocabulaire non terminal N. Nous appelons *langage engendré par la grammaire pluriaxiomatique* (N, T, ::=, N') l'ensemble des propositions dérivant d'un élément de N' ; les éléments de N' sont les *axiomes* ; ce langage est la réunion des langages engendrés par les grammaires (N, T, ::=, X) où $X \in N'$. Nous noterons $\mathcal{C}(N, T, ::=, N')$ la réunion des $\mathcal{C}(N, T, ::=, X)$ pour $X \in N'$. La grammaire (N, T, ::=, N') est *ambiguë* si plusieurs pseudo-arborescences de $\mathcal{C}(N, T, ::=, N')$ ont la même suite de feuilles. Elle est *réduite* lorsque chacune des grammaires (N, T, ::=, X), $X \in N'$, engendre un langage non vide et est réduite. Les autres termes introduits au chapitre 4 ont des définitions inchangées. Il est clair que toute grammaire pluriaxiomatique est équivalente à une grammaire à un seul axiome : il suffit d'adjoindre à N un élément X' et de prolonger la relation de production par X' ::= X pour tout axiome X.

Nous ferons sur N' l'hypothèse suivante :
(P A) *si* X_1 *et* X_2 *sont deux axiomes distincts,* X_1 *ne dérive pas de* X_2.
Cette hypothèse n'est d'ailleurs pas essentielle, mais elle simplifiera un peu les générateurs qui seront introduits ; si PA n'est pas vérifiée et si la grammaire (N, T, ::=, N') est réduite, elle est ambiguë.

Nous désignerons par $\mathfrak{J}(N, T, ::=, N')$ l'ensemble des piles conjointes aux pseudo-arborescences de $\mathcal{C}(N, T, ::=, N')$. La grammaire (N, T, ::=, N') étudiée sera aussi notée G.

7. 2. Caractérisation des piles de $\mathfrak{J}(G)$.

7. 2. 1. D'après l'étude de la pile conjointe à une ramification orientée (2.10.2), la pile U conjointe à une pseudo-arborescence (I, f) de $\mathcal{C}(N, T, ::=, N')$ possède les propriétés suivantes :

(C1) La dernière sortie est une sortie d'un axiome ; elle est immédiatement précédée par une entrée.
(C2) Si i est une entrée d'un symbole non terminal, $i \neq 1$ et i-1 est une sortie.
(C3) Soit i une entrée pour laquelle i-1 est une sortie, et i' la dernière entrée inférieure à i ; si i est une entrée de A, la suite des éléments dont $i'+1, \ldots, i-1$ sont des sorties est un mot φ tel que $A ::= \tilde{\varphi}$.

Réciproquement, soit U une pile sur V qui vérifie C1, C2, C3 ; elle est conjointe à une pseudo-ramification (I, f) dans V. Les racines de (I, f) sont les éléments qui sortent de U après la dernière entrée ; d'après C1, (I, f) est donc une pseudo-arborescence dont la racine est un axiome. Les entrées des feuilles de (I, f) sont caractérisées par le fait qu'elles ne sont pas immédiatement précédées d'une sortie : d'après C2 la suite de feuilles de (I, f) est un mot sur T. Enfin, soit $\lambda = {}'Y_1 \ldots Y_p{}'$ une famille de prédécesseur A ; une entrée dans la pile U est suivie des sorties de $Y_p, \ldots, Y_1$ et de l'entrée de A ; d'après C3, $A ::= \lambda$. (I, f) appartient donc à $\mathcal{C}(N, T, ::=, N')$. De plus, la suite de feuilles de (I, f) est la trace sur T^* du mot d'entrée de la pile U.

7. 2. 2. Cette étude va nous permettre d'abord de préciser la relation entre les dérivations au sens de [12] et les piles conjointes, en démontrant une réciproque d'une particularisation (mais il serait facile d'élargir les hypothèses) d'un résultat établi en 4. 1. 3 ; d'une pile U de $\mathfrak{J}(G)$ on peut extraire une suite d'états u_{j_i} telle que, si θ_{j_i} est la suite des symboles terminaux d'entrée supérieure à j_i, les $u_{j_i}\ \theta_{j_i}$ forment une dérivation ; les j_i sont $j_0 = 0$ et les entrées de symboles non terminaux. Soit $(\delta_0, \ldots, \delta_p)$ une dérivation quelconque : pour $0 \leqslant i < p$, on peut écrire $\delta_i = \lambda_i\ A_i\ \lambda'_i$, $\delta_{i+1} = \lambda_i\ \varphi_i\ \lambda'_i$, $A_i ::= \tilde{\varphi}_i$; dans le cas de la dérivation précédente, pour tout i, $\lambda'_i \in T^*$. Nous qualifierons de *droites* les dérivations qui possèdent cette propriété. Alors, pour $0 < i < p$, $\lambda_i\ A_i$ est facteur gauche de $\lambda_{i-1}\ \varphi_{i-1}$: nous poserons $\lambda_{i-1}\ \varphi_{i-1} = \lambda_i\ A_i\ \gamma_i$, d'où $\lambda'_i = \gamma_i\ \lambda'_{i-1}$.

Lemme 1. *Pour toute X-dérivation droite* $(\delta_0 = X, \delta_1, \ldots, \delta_p)$ *où* $X \in N'$ *et* $\delta_p \in T^*$, *avec les notations précédentes,* Λ, $\lambda_{p-1}\ \varphi_{p-1}$, $\lambda_{p-1}\ A_{p-1}, \ldots, \lambda_i\ \varphi_i$, $\lambda_i\ A_i, \ldots, \lambda_0\ \varphi_0 = \varphi_0$, $\lambda_0\ A_0 = X$ *est une suite d'états, d'indices* $j_0 = 0 < j_1 < j_2 \leqslant j_3 < \ldots \leqslant j_{2p-2i-1} < j_{2p-2i} \leqslant \ldots \leqslant j_{2p-1} < j_{2p}$, *de la pile conjointe à une pseudo-arborescence de* $\mathcal{C}(G)$ *dont* δ_p *est la suite de feuilles ; pour* $0 < k \leqslant p$, *les* j_{2k} *sont les entrées de symboles auxiliaires ; pour* $0 \leqslant i < p$, *les entiers strictement compris entre* j_{2i} *et* j_{2i+1} *sont des entrées de symboles terminaux ;* λ'_i *est la suite des symboles terminaux d'entrée supérieure à* j_{2p-2i}.

Complétons la suite précédente en une pile U : les entiers jusqu'à j_1 sont les entrées des éléments de $\lambda_{p-1}\ \varphi_{p-1} = u_{j_1}$; si $u_{j_{2p-2i-1}} = \lambda_i\ \varphi_i$, $j_{2p-2i-1}$ est suivi des sorties des éléments de φ_i, puis de l'entrée de A_i : $u_{j_{2p-2i}} = \lambda_i\ A_i$; j_{2p-2i} est suivi des entrées des symboles terminaux formant γ_i,

de manière à parvenir à $u_{j_{2p-2i+1}} = \lambda_i\ A_i\ \gamma_i = \lambda_{i-1}\ \varphi_{i-1}$; $u_{j_{2p}} = X$, $u_{j_{2p}+1} = \Lambda$. Cette pile vérifie C1, C2, C3. $\lambda'_i = \gamma_i\ \lambda'_{i-1}$ permet de démontrer, par récurrence sur i, la dernière partie du lemme ; et $\delta_p = \lambda_{p-1}\ \varphi_{p-1}\ \lambda'_{p-1}$ est la suite des symboles terminaux entrant dans la pile.

7. 2. 3. Nous définirons dans la suite un certain nombre de générateurs de piles engendrant, pour toute donnée α de T^*, les piles conjointes aux pseudo-arborescences de $\mathcal{C}(G)$ qui ont α pour suite de feuilles. Nous emploierons les notations du chapitre 5 ; les entiers j_i seront les entrées de feuilles ; les générateurs seront définis par $W = V$, d'où $E = V \cup \{\sigma\}$, un entier cd et un transducteur ; on posera $T' = T \cup \{\sigma\}$. Comme l'entrée d'un symbole non terminal A intervient après les sorties d'éléments formant un mot φ tel que $A ::= \tilde{\varphi}$, nous représenterons la relation ::= par un pseudo-graphe de production $(J, \Gamma ; g)$ du troisième type possédant les propriétés PG1, PG2, PG3' (4. 6. 4). Nous pourrons donc traiter des grammaires généralisées. Nous poserons ici $J_1 = J \cup \{\sigma\}$. Nous allons d'abord définir un générateur gn^0 fort simple, qui n'est pas pratiquement utilisable, mais auquel nous ramènerons tous les autres.

7. 3. Générateur gn^0.

Il est défini par $W = V$, $cd = 1$ et un transducteur fini tr^0 :

a) $M = E \times J_1 \times T'$. Dans les définitions qui suivent, a et b sont des éléments quelconques de T', Y et Z des éléments quelconques de E, w un élément quelconque de J.
b) $q = 2$; $l(a, b) = (b, \sigma, a)$.
c) $d[(b, \sigma, a), Z] = \{(Z, o(b), a)\}$;
$d[(Y, w, a), Z] = \{(Z, w', a)\}$ si w' vérifie $w \Gamma w'$ et $g(w') = Y$: il existe au plus un tel point w' dans J, d'après PG 2 ; il n'en existe aucun si $Y = \sigma$, ou si $w \in J''$ c'est-à-dire si $g(w) \in \bar{N}$ (P G 3').
d) Si $a \neq \sigma$, $m(b, \sigma, a) = \{(s_f, {}'a{}'\}$;
$m(Y, w, a)$ contient $\{(Y, w', a), {}'A{}'\}$ pour tout point w' de J'' tel que $w \Gamma w'$ et $g(w') = \bar{A}$: il n'existe aucun tel point w' si $w \in J''$;
si $w \in J''$ et $g(w) = \bar{A}$, $m(Y, w, a)$ contient $(s_f, {}'a{}')$ lorsque $a \neq \sigma$ ou $Y = a = \sigma$ et $A \in N'$, et contient toujours $[(Y, o(A), a), {}'\sigma{}']$;
$m(Y, w, a)$ ne contient pas d'autre élément.

Si la condition GP 2 n'est pas satisfaite, il existe une suite d'éléments de M, $(Y, o(A_1), a), \ldots, (Y, o(A_n), a)$ tels que $A_1 = A_n$ et que, pour $1 \leqslant i \leqslant n-1$, $A_i\ \bar{A}_{i+1}$ soit un chemin du pseudo-graphe $(J, \Gamma ; g)$ qui représente une règle, autrement dit que $A_{i+1} ::= A_i$. La condition DA entraîne donc GP 2. La propriété établie en 6.4.3 reste vraie pour les grammaires pluriaxiomatiques réduites : DA *est l'hypothèse la plus faible qui permette à un générateur de piles d'engendrer, pour une donnée* α, *un ensemble de piles en correspondance biunivoque avec l'ensemble des pseudo-arborescences de* $\mathcal{C}(G)$ *dont* α *est la suite de feuilles.*

Etudions les calculs complets du transducteur qui ont pour entrée $\nu =$ '$\alpha\, x_1 \dots x_r\, \sigma$' avec $r \geqslant 0$ (si $r = 0$, $\nu =$ 'α'), $\alpha \in T'$, $x_1 \in T$ si $r \geqslant 1$, '$x_2 \dots x_r$' $\in V^*$; convenons que $x_{r+1} = \sigma$. Dans tous les cas, $s_0 = (x_1, \sigma, \alpha)$. Pour $r \geqslant 1$, ν est l'entrée du calcul de longueur 1 (s_f, Λ, 'α'), qui a pour résultat 'α'. Si un autre calcul complet possède ν pour entrée, le premier triplet du calcul est $[(x_2, o(x_1), \alpha)$, 'x_2', 'σ'$]$ et le $i^{\text{ème}}$ triplet, s'il n'est pas le dernier, est de la forme $[(Y_i, w_i, \alpha_i), y_i, \lambda_i]$ où $Y_i \in E$, $w_i \in J$, $\alpha_i \in T'$, y_i est un mot de E^* de longueur 0 ou 1, $\lambda_i \in E^*$; montrons, par récurrence sur i, que $\alpha_i = \alpha$ et que Y_i est le dernier élément du mot $y_1 \dots y_i$, ce qui est vrai pour $i = 1$:

- si $y_i \neq \Lambda$, $(Y_i, w_i, \alpha_i) \in d[(Y_{i-1}, w_{i-1}, \alpha), y_i]$ d'où $\alpha_i = \alpha$ et $y_i =$ 'Y_i' ;
- si $y_i = \Lambda$, $[(Y_i, w_i, \alpha_i), \lambda_i] \in m(Y_{i-1}, w_{i-1}, \alpha)$ d'où $\alpha_i = \alpha$ et $Y_i = Y_{i-1}$: Y_i est le dernier élément de $y_1 \dots y_{i-1} = y_1 \dots y_{i-1}\, y_i$.

Envisageons ceux des triplets où $w_i \in J'$: $w_i = o(Z_i)$, $Z_i \in V$; le premier est de ceux-là. Après l'un d'entre eux $[(x_k, o(Z), \alpha), y, \lambda]$ on trouve les $k' > 0$ autres triplets $[(x_{k+1}, w'_1, \alpha), x_{k+1}$, '$\sigma$'$]$, ..., $[(x_{k+k'}, w'_{k'}, \alpha), x_{k+k'}$, '$\sigma$'$]$, puis $[(x_{k+k'}, w'', \alpha), \Lambda$, '$A$'$]$, puis $[(x_{k+k'}, o(A), \alpha), \Lambda$, '$\sigma$'$]$ ou le dernier triplet du calcul (s_f, Λ, 'α'), car $g(w'') \in \overline{N}$; dans les deux cas $A ::= x_{k+k'-1} \dots x_k\, Z$ ($A ::= Z$ si $k' = 0$). Il existe donc un entier $p \geqslant 1$ et une suite de p entiers k_i, $1 \leqslant k_1 \leqslant \dots \leqslant k_p \leqslant r$ tels que le résultat du calcul soit $\sigma^{k_1} A_1\, \sigma\, \sigma^{k_2-k_1} \dots \sigma^{k_p-k_{p-1}} A_p\, \alpha$ et que $A_1 ::= x_{k_1} \dots x_1$, $A_i ::= x_{k_i} \dots x_{k_{i-1}+1} A_{i-1}$ pour $1 < i \leqslant p$.

On voit alors que, pour tout mot α de T^*, les piles de l'ensemble $gn^0(\alpha)$ vérifient les conditions C2 et C3. Pour $\alpha = \sigma$, d'après la définition de l'application m, A_p est un axiome et $k_p = r$; il en résulte que les piles de $gn^0(\alpha)$ vérifient aussi C1 et de plus que la condition GP 1 est satisfaite par le générateur. Toute pile de $gn^0(\alpha)$ est conjointe à une pseudo-arborescence de $\mathcal{C}(N, T, ::=, N')$ dont la suite de feuilles est α, trace sur T^* de son mot d'entrée.

Au paragraphe 7.7, nous définirons un générateur gn^2 contenu dans $gn^0(5.5)$ et tel que, pour tout mot α sur T, $gn^2(\alpha)$ contienne toute pile conjointe à une pseudo-arborescence de $\mathcal{C}(G)$ dont α est la suite de feuilles.

Théorème 7. 1. *Le générateur* gn^0 *associé à une grammaire pluriaxiomatique* $G = (N, T, ::=, N')$ *vérifiant les conditions* DA *et* PA *est tel que, pour tout mot* α *sur* T, $gn^0(\alpha)$ *soit l'ensemble des piles conjointes aux pseudo-arborescences de* $\mathcal{C}(G)$ *dont* α *est la suite de feuilles.*

7. 4. Complément à l'étude des piles de l'ensemble $\mathcal{J}(G)$.

Nous préciserons ici C1, C2, C3 par d'autres propriétés de la pile $U = (u_0, \dots, u_n)$ conjointe à une pseudo-arborescence (I, f) de $\mathcal{C}(G)$. Nous emploierons les notations suivantes : $\alpha^U_{ii'}$ (resp. α^U_{i-}) est le mot formé par la suite des symboles terminaux dont l'entrée dans U est strictement comprise entre i et i' (resp. supérieure à i). $U' = (u'_0, u'_1, \dots, u'_n)$ est la pile conjointe à l'arborescence orientée I, R son ordre associé ; y et z sont des points de I, $Y = f(y)$, $Z = f(z)$, x est le R-prédécesseur de y (y n'est pas la racine de I).

Supposons d'abord que, i étant l'entrée de z dans U', y est le sommet de u'_{i-1}. D'après 2.10.3, y, z est un couple de points consécutifs et x R z : x est le plus grand R-minorant commun à y et z ; d'après l'étude des points consécutifs (2.6.3), il existe un chemin de I, $(x, z_1, \dots, z_p = z)$ tel que y, z_1 soient consécutifs dans la même famille et, si $p \geqslant 2$, $z_2, \dots, z_p$ premiers de leur famille (figure 7.1). Posons $A = f(x)$ et $C = f(z_1)$; Z est une initiale de C et il existe deux mots Ψ et Ψ' de V^* tels que $A ::= \Psi\, Y\, C\, \Psi'$. D'autre part, si i' est l'entrée de x et j la sortie de z_1, y est le sommet de u'_j, $u'_{i-1} = u'_j$, $u_{i-1} = u_j = u_{i'-1} \Psi\, Y$, $u_i = u_{i'-1} \Psi\, Y\, Z$, $u_{i'} = u_{i'-1} A$. Enfin, les points de $R(z_1)$ forment une sous-arborescence complète (1.4.1) de I ; il en est de même pour les points de $R(z_1) \setminus (R(z) \setminus \{z\})$ car, avec tout point, $R(z) \setminus \{z\}$ contient ses majorants et ceux de sa famille ; les feuilles de cette sous-arborescence sont z et les feuilles de I qui entrent dans U' après z et avant z_1 (rappelons que l'ordre d'entrée de U' est l'ordre de sortie de I). Les feuilles de I qui entrent après z_1 et avant x sont celles qui majorent pour R les points suivant z_1 dans sa famille (2.3.3). Comme $(I, f) \in \mathcal{C}(G)$, $Z\, \alpha^U_{ii'}$ dérive donc de $C\, \Psi'$.

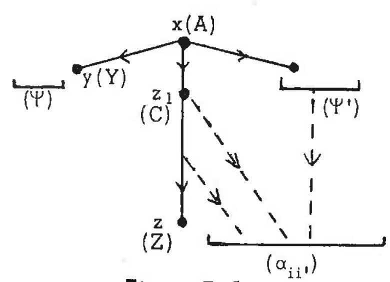


Figure 7. 1

Lemme 2. *Soit* i *une entrée dans* U *de* $Z \in V$, *et* $u_i = \eta Y Z$ $(\eta \in V^*, Y \in V)$. *Il existe deux symboles auxiliaires* A, C, *trois mots* Ψ, Ψ', η' *sur* V *et une entrée* i' *dans* U *tels que* $A ::= \Psi\ Y\ C\ \Psi'$, Z *est initiale de* C, $u_{i'} = \eta'\ A$, $\eta = \eta'\ \Psi$, $Z\ \alpha^U_{ii'}$ *dérive de* $C\ \Psi'$.

Supposons maintenant que, i étant l'entrée de z, $u'_{i-1} = \Lambda$. D'après 2.10.3, il existe un chemin de I ayant pour origine la racine, pour extrémité z et dont chaque point est le premier de sa famille : Z est une initiale d'un axiome X. Comme ci-dessus, les points de $I \setminus [R(z) \setminus \{z\}]$ forment une sous-arborescence complète de I et on en déduit que $Z\ \alpha^U_{i-}$ dérive de X.

Lemme 3. *Soit* i *une entrée dans* U *de* $Z \in V$ *telle que* $u_i = {}'\ Z\ {}'$: Z *est initiale d'un axiome* X *et* $Z\ \alpha^U_{i-}$ *dérive de* X.

Si la sortie de y est précédée de son entrée et si z est la première feuille qui entre dans U' après cette sortie, (y, z) est un couple de points consécutifs et $x\ \bar{R}\ z$ (2.10.3). On déduit alors de 2.6.3 :

Lemme 4. *Si* i *est une sortie de* Y *de la pile* U, i-1 *une entrée et si la première entrée de symbole terminal supérieure à* i *est une entrée de* Z, *il existe trois symboles auxiliaires* A, B, C *et deux mots* Ψ *et* Ψ' *de* V^* *tels que* $A ::= \Psi B C \Psi'$, Y *est finale stricte de* B, Z *est initiale de* C.

Enfin, on déduit encore de 2.10.3 :

Lemme 5. *Si* i *est une sortie de* Y *de la pile* U, i-1 *une entrée et si aucune entrée de symbole terminal n'est supérieure à* i, Y *est une finale d'un axiome*.

Ces lemmes conduisent à définir deux relations dans V :
$Y \prec Z$ *si, et seulement si, il existe deux symboles auxiliaires* A *et* C *et deux mots* Ψ *et* Ψ' *de* V^* *tels que* $A ::= \Psi\ Y\ C\ \Psi'$ *et* Z *soit initiale de* C.
$Y \succ Z$ *si, et seulement si,* $Z \in T$ *et il existe trois symboles auxiliaires* A, B, C *et deux mots* Ψ *et* Ψ' *de* V^* *tels que* $A ::= \Psi\ B\ C\ \Psi'$, Y *soit finale stricte de* B, Z *soit initiale de* C.
$Y \prec Z$ ou $Y \succ Z$ équivaut à $Z \in T$ et Y, Z est un couple vicinal (théorème 4.3).
Prolongeons ces relations dans $V \cup \{\sigma\}$ par :
$\sigma \prec Z$ *si, et seulement si,* Z *est initiale d'un axiome* ;
$Y \succ \sigma$ *si, et seulement si,* Y *est finale d'un axiome*.

7. 5. Générateur gn^1 utilisant les relations $\prec$ et $\succ$.

7.5.1. Ce générateur est défini par W = V, cd = 1 et un transducteur fini tr^1 que l'identité projette (5. 5) dans tr^0 (7. 3) :

a) $M = E \times J_1 \times T'$. Dans les définitions qui suivent, $\alpha \in T'$, $b \in T'$, $Y \in E$, $Z \in E$, $w \in J$.
b) $q = 2$; $l(\alpha, b) = (b, \sigma, \alpha)$.
c) Si $b \succ \alpha$ (d'où $b \neq \sigma$), $d[(b, \sigma, \alpha), Z] = \{(Z, o(b), \alpha)\}$;
$d[(Y, w, \alpha), Z] = \{(Z, w', \alpha)\}$ si w' vérifie $w\ \Gamma\ w'$ et $g(w') = Y$.
d) Si $b \prec \alpha$ (d'où $\alpha \neq \sigma$), $m(b, \sigma, \alpha) = \{(s_f, {}'\alpha{}')\}$;
$m(Y, w, \alpha)$ contient $[(Y, w', \alpha), {}'A{}']$ pour tout point w' de J'' tel que $w\ \Gamma\ w'$, $g(w') = \bar{A}$ et $Y \prec A$;
si $w \in J''$ et $g(w) = \bar{A}$, $m(Y, w, \alpha)$ contient : $(s_f, {}'\alpha{}')$ lorsque $A \prec \alpha$ (d'où $\alpha \neq \sigma$) ou lorsque $Y = \alpha = \sigma$ et $A \in N'$, $[(Y, o(A), \alpha), {}'\sigma{}']$ lorsque $A \succ \alpha$, sauf si $Y = \alpha = \sigma$ et $A \in N'$;
$m(Y, w, \alpha)$ ne contient pas d'autre élément.

Puisque l'identité projette tr^1 dans tr^0, tr^1 vérifie GP 2 dès que DA est satisfaite, et gn^1 vérifie GP 1. Au paragraphe 7.7, nous définirons un générateur gn^2 contenu dans gn^1 et tel que, pour $\alpha \in T^*$, $gn^2(\alpha)$ contienne toute pile conjointe à une pseudo-arborescence de $\mathcal{C}(G)$ dont α est la suite de feuilles.

Théorème 7.2. *Le générateur* gn^1 *associé à une grammaire pluriaxiomatique* $G = (N, T, ::=, N')$ *vérifiant les conditions* DA *et* PA *est tel que, pour tout mot* α *sur* T, $gn^1(\alpha)$ *soit l'ensemble des piles conjointes aux pseudo-arborescences de* $\mathcal{C}(G)$ *dont* α *est la suite de feuilles.*

La figure 7.2 schématise la construction de ces piles.

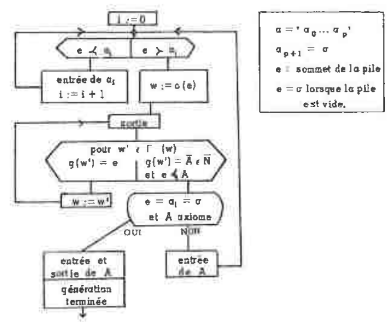


Figure 7. 2

7.5.2. Etude du déterminisme du transducteur. Pour tout couple (s, Z) de $M \times E$, $d(s, Z)$ contient un élément au plus ; s étant donné, l'existence de cet élément ne dépend pas de Z.

Si $s = (b, \sigma, a)$ où a et b appartiennent à T', $m(s)$ contient un élément au plus ; $m(s)$ et $d(s, Z)$ sont simultanément non vides si, et seulement si, $b \prec a$ et $b \succ a$ (d'où $a \neq \sigma$, $b \neq \sigma$).

Si $s = (Y, w, a)$, $a \in T'$, $Y \in E$, $w \in J''$, $g(w) = \bar{A} \in \bar{N}$, $d(s, Z)$ est vide, $m(s)$ contient deux éléments si, et seulement si, $A \prec a$ et $A \succ a$ (d'où $a \neq \sigma$).

Soit enfin $s = (Y, w, a)$, $a \in T'$, $Y \in E$, $w \in J$, $g(w) \notin \bar{N}$. $m(s)$ contient plusieurs éléments si, et seulement si, il existe deux symboles non terminaux distincts A et A' tels que $\bar{A}$ et $\bar{A}'$ soient images par g de deux points de $\Gamma(w)$, que $Y \prec A$ et $Y \prec A'$. $m(s)$ et $d(s, Z)$ sont tous deux non vides si, et seulement si, il existe $A \in N$ tel que Y et $\bar{A}$ soient images par g de deux points de $\Gamma(w)$ et que $Y \prec A$. Ces deux propriétés ne dépendent pas de a ; *nous désignerons par* Det *l'ensemble des* (Y, w), $Y \in E$, $w \in J$, $g(w) \notin \bar{N}$ *tels que, pour tout* $a \in T'$ *et tout* $Z \in E$, *la réunion des ensembles* $d[(Y, w, a), Z]$ *et* $m(Y, w, a)$ *contienne au plus un élément.*

D'après ce qui précède, et la propriété PG2' du pseudo-graphe de production :

Théorème 7. 3. *Pour que le transducteur* tr^1 *soit déterministe, il faut et il suffit qu'aucune des conditions suivantes ne soit réalisée :*

1) *il existe* $Y \in V$, $a \in T$ *tels que* $Y \prec a$ *et* $Y \succ a$;
2) *il existe* $A \in N$, $A' \in N$, $Y \in E$, $\varphi \in V^*$ *tels que* $A \neq A'$, $A ::= \varphi$, $A' ::= \varphi$, $Y \prec A$, $Y \succ A'$;
3) *il existe* $A \in N$, $Y \in V$, $\varphi \in V^*$ *tels que* $A ::= \varphi$, $Y \prec A$ *et* $Y\varphi$ *soit facteur droit du second membre d'une règle de la grammaire.*

Il est facile d'étudier, à partir des matrices des relations $\prec$ et $\succ$, et du pseudo-graphe de production, si l'une de ces conditions est réalisée. Lorsque le transducteur n'est pas déterministe, il n'en résulte pas, comme en 6.3.3, que tout générateur des piles de $\mathcal{J}(G)$ pour lequel $cd = 1$ et j_i est la $i^{\text{ème}}$ entrée de symbole terminal, soit un générateur de piles multiples : nous étudierons plus bas des cas où il n'en est rien ; on peut seulement montrer, lorsque la grammaire est réduite, qu'il existe une phrase α telle que, dans la génération de $gn^1(\alpha)$, l'un des calculs du transducteur tr^1 possède un triplet (s, y, λ) pour lequel la réunion de $m(s)$ et $d(s, Z)$ contient plus d'un élément.

7.5.3. Remarque. Supposons la condition 3 du théorème 7.3 réalisée : il existe deux règles $A ::= Z\varphi$, $A' ::= \Psi Y Z \varphi'$ et $Y \prec A$; il existe donc aussi une

règle $A'' ::= \Psi_1 Y C \Psi'_1$ où C a A pour initiale, donc Z pour initiale stricte. Pour distinguer ces deux possibilités, définissons deux nouvelles relations dans V, dont la relation $\prec$ soit réunion :

$Y \doteq Z$ *si, et seulement si, il existe* $A \in N$, $\Psi \in V^*$, $\Psi' \in V^*$ *tels que* $A ::= \Psi Y Z \Psi'$;
$Y \angle Z$ *si, et seulement si, il existe* $A \in N$, $C \in N$, $\Psi \in V^*$, $\Psi' \in V^*$ *tels que* $A ::= \Psi Y C \Psi'$ *et* Z *soit initiale stricte de* C.

Si ces deux relations sont incompatibles, la condition 3 du théorème ne peut être satisfaite. On en déduit une condition simple pour que le transducteur soit déterministe : il suffit que les trois relations $\doteq$, $\angle$, $\succ$ soient deux à deux incompatibles et que deux règles distinctes aient des seconds membres distincts.

Pour la pile conjointe à une pseudo-arborescence de $\mathcal{C}(G)$, si une sortie de Z est immédiatement suivie d'une sortie de Y, $Y \doteq Z$; si au contraire, après une sortie de Z, le sommet de la pile étant Y, se produit une entrée, il s'agit de l'entrée d'un élément B tel que $Y \prec B$ (lemme 2) et dont Z est initiale : d'après les définitions des relations $\prec$ et $\angle$, $Y \angle Z$. Lorsque les relations $\doteq$ et $\angle$ sont incompatibles, on peut conclure aux réciproques : après une sortie, on sait s'il se produit une nouvelle sortie ou une entrée en comparant le sommet au seul élément qui vient de sortir, sans tenir compte de ceux qui sont sortis avant lui. Si on traite une grammaire de Chomsky non généralisée, on peut ne pas utiliser un pseudo-graphe de production, mais seulement une liste des règles. D'autre part, pour éviter deux tests différents ($Y \prec Z$ ou $Y \succ Z$), ($Y \doteq Z$ ou $Y \angle Z$) portant sur le même couple, on peut, en s'inspirant de 2.10.2, à l'entrée d'un élément Z sur le sommet Y, noter si $Y \angle Z$ par l'entrée d'un séparateur $\ddagger$ avant Y : une suite de sorties consécutives se termine alors par une sortie de $\ddagger$. Nous ne formaliserons par cette variante, nous contentant de la schématiser sur la figure 7. 3. λ est le plus grand facteur droit de l'état courant de la pile qui ne contient pas $\ddagger$; les autres notations sont celles de la figure 7. 2 ; on prolonge la relation $\angle$ par : $\sigma \angle Z$ si, et seulement si, Z est une initiale de l'axiome X.

Pour que les trois relations $\doteq$, $\angle$, $\succ$ soient deux à deux incompatibles, il suffit qu'il existe deux symboles terminaux a et b tels que toute règle soit de la forme $A ::= a\,\varphi\,b$ ou φ ne contient aucune occurrence de a ni de b. En effet $Y \doteq Z$ entraîne alors $Y \neq b$ et $Z \neq a$, $Y \angle Z$ entraîne $Y \neq b$ et $Z = a$, $Y \succ Z$ entraîne $Y = b$. A toute grammaire $G = (N, T, ::=, N')$ on peut associer une grammaire $G' = (N, T', ::=', N')$ et une application h de T' dans T^* qui définit un homomorphisme h^* transformant le langage engendré par G' en le langage engendré par G (4.1.5), les pseudo-arborescences de $\mathcal{C}(G')$ qui ont β pour suite de feuilles et celles de $\mathcal{C}(G)$ qui ont $h^*(\beta)$ pour suite de feuilles étant en correspondance biunivoque : il suffit de prendre $T' = T \cup \{a, b\}$ où ni a, ni b n'appartient à $T \cup N$, $h(a) = h(b) = \Lambda$, $h(c) = c$ si $c \in T$, et de définir $::='$ par : $A ::=' a\varphi b$ si, et seulement si, $A ::= \varphi$. Cette remarque montre qu'il suffit de peu modifier une grammaire pour rendre très facile le problème de l'analyse.

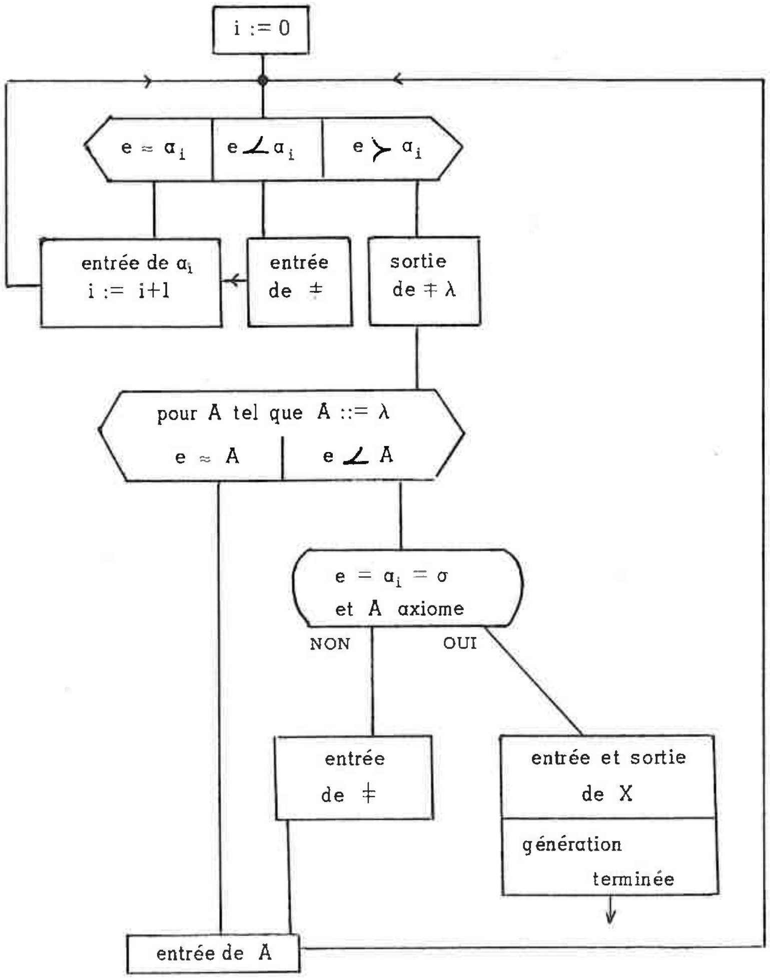


Figure 7. 3

7. 5. 4. Exemple. (cf. expressions et instructions conditionnelles d'Algol) :

$N = \{\dot{X}, P, C, A, E, D, W, I\}$; $T = \{$*SI, ALORS, SINON*, :=, $a, b, \ldots, z\}$;
$N' = \{X\}$;

X ::= A | C | C *SINON* X
C ::= P *ALORS* A
P ::= *SI* E
E ::= D *SINON* E | W
D ::= P *ALORS* W
W ::= I
I ::= a | b | ... | z
A ::= I := E

La pseudo-ramification de production du troisième type est schématisée figure 7. 4.

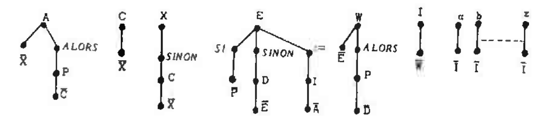


Figure 7. 4

Matrice des trois relations $\approx, \angle, \succ$:

1er elt \ 2ème elt	X	P	C	A	E	D	I	W	*SI*	*ALORS*	*SINON*	:=	a...z	σ
X														$\succ$
P										$\approx$				
C											$\approx$			$\succ$
A											$\succ$			$\succ$
E										$\succ$	$\succ$			$\succ$
D											$\approx$			
I										$\succ$	$\succ$	$\approx$		$\succ$
W										$\succ$	$\succ$			$\succ$
SI		$\angle$			$\approx$	$\angle$	$\angle$	$\angle$	$\angle$				$\angle$	
ALORS				$\approx$			$\angle$	$\approx$					$\angle$	
SINON	$\approx$	$\angle$	$\angle$	$\angle$	$\approx$	$\angle$	$\angle$	$\angle$	$\angle$				$\angle$	
:=		$\angle$			$\approx$	$\angle$	$\angle$	$\angle$	$\angle$				$\angle$	
a...z										$\succ$	$\succ$	$\succ$		$\succ$
σ	$\angle$	$\angle$	$\angle$	$\angle$			$\angle$		$\angle$				$\angle$	

Les trois relations sont deux à deux disjointes et deux règles distinctes ont des seconds membres distincts. Le transducteur tr^1 est déterministe. On peut aussi utiliser l'algorithme de 7. 5. 3. Nous représenterons ci-dessous certains états de la pile obtenue par ce dernier algorithme pour l'analyse de la phrase:

SI b *ALORS* a := *SI* c *ALORS* d *SINON* h.

Il suffit de supprimer les ǂ, c'est-à-dire de prendre la trace de cette pile sur V*, pour trouver la pile engendrée avec la même donnée par le générateur gn^1.

(*AL* désigne *ALORS*)

																c		I		W
																ǂ		ǂ		ǂ
															SI	SI	SI	SI	SI	SI
															ǂ	ǂ	ǂ	ǂ	ǂ	ǂ
														:=	:=	:=	:=	:=	:=	:=
											a		I	I	I	I	I	I	I	I
	b		I		W						ǂ		ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ
	ǂ		ǂ		ǂ		E			AL	AL	AL	AL	AL	AL	AL	AL	AL	AL	AL
SI	SI	SI	SI	SI	SI	SI	SI		P	P	P	P	P	P	P	P	P	P	P	P
ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ		ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ	ǂ

etc ...

7. 6. Contextes d'un symbole du vocabulaire.

7.6.1. Dans le cas où le transducteur tr^1 n'est pas déterministe, nous définirons un autre générateur qui tentera d'éviter les "fausses pistes" en utilisant l'état courant de la pile, au lieu de son seul sommet, ou en étendant le contexte droit (cd > 1). Pour cela, nous emploierons plus complètement les lemmes 2 et 3. $\alpha^U_{ii'}$ et α^U_{i-} ont ici le même sens qu'en 7. 4.

Définition. (η, γ) *est un* contexte *de* $Z \in V$ *s'il existe une pile* U *de* $\mathfrak{J}(G)$, *une entrée* i *de* Z *dans* U *et un entier* $r \geqslant 0$ *tels que* $u_i = \eta Z$ *et que* γ *soit facteur gauche de* $\alpha^U_{i-} \sigma^r$.

Si $\eta \neq \Lambda$ $(\eta = \eta_1 Y)$, d'après le lemme 2, il existe alors deux symboles auxiliaires A, C, trois mots φ, Ψ', η' de V* et une entrée i' dans U tels que $A ::= \varphi C \Psi'$, $\varphi \neq \Lambda$, $u_{i'} = \eta' A$, $\eta = \eta' \varphi$ et $Z \alpha^U_{ii'}$ dérive de $C \Psi'$; deux cas sont possibles : γ est facteur gauche de $\alpha^U_{ii'}$, ou bien $\gamma = \alpha^U_{ii'} \gamma'$, γ' étant un facteur gauche non vide de $\alpha^U_{i'-} \sigma^r$. Si $\eta = \Lambda$, il existe un mot ρ dérivant d'un axiome et un entier $r \geqslant 0$ tels que $Z \gamma$ soit facteur gauche de $\rho \sigma^r$ (lemme 3).

Pour éviter d'avoir à considérer des ensembles contenant le mot vide, nous remplacerons η par le mot $\sigma \eta$ sur $V \cup \{\sigma\}$. *Nous noterons* $\mathcal{C}(Z, \gamma)$ *l'ensemble des mots* $\sigma \eta$ *tels que* (η, γ) *soit un contexte de* Z. $\mathcal{C}(Z, \gamma)$ est

contenu dans la réunion des sous-ensembles suivants du monoïde libre $(V \cup \{\sigma\})^*$:

- $\mathcal{C}(A, \Lambda) \varphi$ pour toute règle $A ::= \varphi C \Psi'$ telle que $\varphi \neq \Lambda$ et $Z \gamma$ soit facteur gauche d'un mot qui dérive de $C \Psi'$;
- $\mathcal{C}(A, \gamma') \varphi$ pour toute décomposition de γ en $\gamma'' \gamma'$ $(\gamma' \neq \Lambda)$ et toute règle $A ::= \varphi C \Psi'$ telle que $\varphi \neq \Lambda$ et $Z \gamma''$ dérive de $C \Psi'$;
- $\{\sigma\}$ s'il existe un mot ρ dérivant d'un axiome et un entier $r \geqslant 0$ tels que $Z \gamma$ soit facteur gauche de $\rho \sigma^r$.

Des deux lemmes qui suivent résulte immédiatement que tout mot de l'un de ces ensembles appartient à $\mathcal{C}(Z, \gamma)$, qui est donc leur réunion, à condition que la grammaire soit réduite, ce que nous supposerons dans la suite de ce paragraphe 7. 6.

Lemme 6. *Si la grammaire* $G = (N, T, ::=, N')$ *est réduite et si* $Z \in V$, $A \in N$, $C \in V$, $\varphi \in V^*$, $\Psi' \in V^*$, $\beta \in T^*$, $\eta' \in V^*$ *sont tels que* $A ::= \varphi C \Psi'$, $Z \beta$ *dérive de* $C \Psi'$, $\eta' A$ *est un état* u'_k *d'une pile* U' *de* $\mathfrak{J}(G)$, *alors il existe une pile* U *de* $\mathfrak{J}(G)$ *et deux entrées* i, i' *dans* U *telles que* $u_i = \eta' \varphi Z$, $u_{i'} = \eta' A$ *et* $\alpha^U_{i-} = \beta \alpha^{U'}_{k-}$.

D'après 7. 2. 2, il existe un axiome X et une X-dérivation droite du mot $\eta' A \theta$ où $\theta = \alpha^{U'}_{k-}$. D'autre part, il existe une $C \Psi'$-dérivation droite de $Z \beta$ et la grammaire est réduite. On peut donc construire une X-dérivation droite d'une phrase du langage qui contienne les mots $\eta' A \theta$, $\eta' \varphi C \Psi' \theta$, $\eta' \varphi Z \beta \theta$.

Supposons d'abord que Z est un symbole auxiliaire. D'après le lemme 1, il existe deux entrées i, i' dans une pile U de $\mathfrak{J}(G)$ telles que $u_i = \eta' \varphi Z$, $u_{i'} = \eta' A$ et $\alpha^U_{i-} = \beta \theta$.

Supposons ensuite que Z appartient à T et que $Z \beta = C \Psi'$. $\eta' \varphi$ peut s'écrire $\eta'' \beta'$ où $\beta' \in T^*$ et η'' est vide ou se termine par un symbole de N. D'après le lemme 1, il existe une pile U de $\mathfrak{J}(G)$ et trois entiers $i' > j > j'$ tels que i' soit une entrée, $u_{i'} = \eta' A$, $u_j = \eta' \varphi Z \beta = \eta'' \beta' Z \beta$, $\alpha^U_{j-} = \theta$, $u_{j'} = \eta''$, j' et j étant séparés par les entrées des éléments de $\beta' Z \beta$; d'où une entrée i telle que $u_i = \eta' \varphi Z$ et $\alpha^U_{i(j+1)} = \beta$; donc $\alpha^U_{i-} = \beta \theta$.

Si $Z \in T$ et $Z \beta \neq C \Psi'$, le mot qui précède $\eta' \varphi Z \beta \theta$ dans la dérivation s'écrit $\eta' \varphi \beta_1 B \theta'$, avec $Z \beta \theta = \beta_1 \beta_2 \theta'$ et $B ::= \beta_2 : \beta_1 \beta_2 \neq \Lambda$, on peut poser $\beta_1 \beta_2 = Z \beta_3$, d'où $\beta_3 \theta' = \beta \theta$. Comme ci-dessus, on écrit $\eta' \varphi = \eta'' \beta'$. Il existe $U \in \mathfrak{J}(G)$ et trois entiers $i' > j \geqslant j'$ tels que i' soit une entrée, $u_{i'} = \eta' A$, $u_j = \eta' \varphi \beta_1 \beta_2 = \eta'' \beta' \varphi Z \beta_3$, $\alpha^U_{j-} = \theta'$, $u_{j'} = \eta''$, j' et j étant séparés par les entrées des éléments de $\beta' Z \beta_3$; d'où une entrée i telle que $u_i = \eta' \varphi Z$, $\alpha^U_{i(j+1)} = \beta_3$; donc $\alpha^U_{i-} = \beta_3 \theta' = \beta \theta$.

Lemme 7. *Si la grammaire* $G = (N, T, ::=, N')$ *est réduite, si* $Z \in V$ *et un mot* β *sur* T *sont tels que* $Z\beta$ *dérive d'un axiome, alors il existe une entrée* i *dans une pile* U *de* $\mathcal{J}(G)$ *telle que* $u_i =' Z'$ *et* $\beta = \alpha^{U}_{i-}$.

Ce lemme est vrai si $Z \in N'$ et $\beta = \Lambda$. Sinon, il existe un axiome X et une règle $X ::= C\Psi'$ tels que $Z\beta$ dérive de $C\Psi'$; il suffit alors d'appliquer le lemme 6.

$\mathcal{C}(Z, \gamma)$ est donc réunion d'ensembles $\mathcal{C}(A, \gamma')\varphi$ où $\varphi \neq \Lambda, |\gamma'| \leqslant |\gamma|$, et éventuellement de l'ensemble réduit à 'σ'. Le théorème 4. 1, dont la démonstration [54] reste valable pour une grammaire généralisée, montre qu'étant donné un entier $n \geqslant 0$, les ensembles $\mathcal{C}(Z, \gamma)$ pour $Z \in V$ et $|\gamma| \leqslant n$ sont les langages Q(D) des propositions d'une structure $\mathcal{S}(N_1 \cup V \cup \{\sigma\}, ::=_1)$ qui dérivent des divers symboles auxiliaires D de N'' ; $V \cup \{\sigma\}$ est le vocabulaire terminal de cette structure et ses règles sont du type $D ::=_1 D'\varphi$ ou $D ::=_1 \sigma$ avec $D \in N_1$, $D' \in N_1$, $\varphi \in V^*$ et $\varphi \neq \Lambda$. Si la grammaire G est une grammaire de Chomsky, il est facile de transformer cette structure en une structure gauche de Kleene dont les $\mathcal{C}(Z, \gamma)$ soient certains des langages des propositions qui dérivent d'un symbole auxiliaire.

Si G est une grammaire généralisée (4.6.5), les phrases de chacun des langages L_A qui définissent sa relation de production ::= sont les chemins d'un pseudo-graphe $(J_A, \Gamma_A ; g_A)$ dont l'origine appartient à un ensemble J'_A et l'extrémité à un ensemble J''_A. Etant donné $Z \in V$ et $\gamma_1 \in V^*$, si un chemin φ de ce pseudo-graphe, d'origine dans J'_A, peut être prolongé en un chemin $\varphi C \Psi'$, d'extrémité dans J''_A, tel que $Z\gamma_1$ dérive (resp. soit facteur gauche d'un mot qui dérive) de $C\Psi'$, il en est de même pour tout chemin d'origine dans J'_A et de même extrémité ; les φ possédant cette propriété sont donc les chemins dont l'origine appartient à J'_A et l'extrémité à un ensemble $J_{Z\gamma_1}$(1) : pour $D \in N_1$ et $D' \in N_1$, les mots φ tels que $D ::=_1 D'\varphi$ forment un langage de Kleene. Il en résulte (4.6.2) que les ensembles $\mathcal{C}(Z, \gamma)$ sont certains des langages Q(D) des propositions d'une structure gauche de Kleene qui dérivent des différents symboles auxiliaires D.

Théorème 7. 4. *Etant donné un entier* $n \geqslant 0$, *pour* $Z \in V$ *et* $|\gamma| \leqslant n$, *les ensembles* $\mathcal{C}(Z, \gamma)$ *des mots* $\sigma\eta$ *tels que* (η, γ) *soit un contexte de* Z *sont certains des langages des propositions d'une structure de Kleene qui dérivent d'un symbole auxiliaire.*

(1) Le fait pour un point de J_A d'appartenir à $J_{Z\gamma_1}$ est décidable, car si $Z\gamma_1$ est facteur gauche d'un mot dérivant de $C\Psi'$, il est aussi facteur gauche d'un mot dérivant d'un facteur gauche de $C\Psi'$ dont la longueur est au plus $|Z\gamma_1|$.

7. 6. 2. Etudions les cas particuliers n = 0 et n = 1. Pour n = 0, à tout Z de V associons, de manière bijective, un élément $<Z>$ n'appartenant pas à $V \cup \{\sigma\}$: N_1 est l'ensemble de ces $<Z>$. La relation de production $::=_1$ est définie par : $<Z> ::=_1 <A>\varphi$ si, et seulement si, $\varphi \neq \Lambda$ et il existe $C \in V$, $\Psi' \in V^*$ tels que $A ::= \varphi C \Psi'$ et Z est une initiale de C ; $<Z> ::=_1 \sigma$ si, et seulement si, Z est initiale d'un axiome.

Si $\alpha \in T$, d'après 2.10.3 et le théorème 4. 3, d, pour que $\mathcal{C}(Z, \alpha)$ ne soit pas vide, il faut et il suffit que le couple (Z, α) soit vicinal ; $\mathcal{C}(Z, \sigma)$ n'est pas vide si, et seulement si, Z est une finale d'un axiome. Ici N_1 sera obtenu en réunissant à l'ensemble N_1 du cas n = 0 les couples vicinaux d'un symbole non terminal et d'un terminal et les couples (Z, σ) où Z est finale d'un axiome. Prolongeons la relation $::=_1$ précédente par :

- $(Z, \alpha) ::=_1 <A>\varphi$ si, et seulement si, $\varphi \neq \Lambda$ et il existe $C \in V$, $\Psi' \in V^*$ tels que $A ::= \varphi C \Psi'$, Z dérive de C alors que α est initiale du premier élément de $\Psi' \neq \Lambda$ ou Z α est facteur gauche d'un mot qui dérive de C ;
- $(Z, \alpha) ::=_1 (A, \alpha)\varphi$ si, et seulement si, $\varphi \neq \Lambda$ et il existe $C \in V$ tel que $A ::= \varphi C$ et Z dérive de C ;
- $(Z, \alpha) ::=_1 \sigma$ si, et seulement si, Z α est facteur gauche d'un mot dérivant d'un axiome ou $\alpha = \sigma$ alors que Z dérive d'un axiome.

$\mathcal{C}(Z, \alpha)$ est l'ensemble des propositions de la structure $\mathcal{S}(N_1 \cup V \cup \{\sigma\}, ::=_1)$ qui dérivent de (Z, α).

7. 6. 3. Les générateurs introduits au paragraphe suivant sont fondés sur deux conséquences de la définition des contextes et de la propriété C3 des piles U de $\mathcal{J}(G)$: soit un entier $r \geqslant 0$;

(C4) si i est une entrée de Z dans U et γ un facteur gauche de $\alpha^{U}_{i-}\sigma^{r}$, σu_{i-1} appartient à $\mathcal{C}(Z, \gamma)$;

(C5) soit i une sortie autre que la dernière, i' la dernière entrée inférieure à i, γ un facteur gauche de $\alpha^{U}_{i-}\sigma^{r}$; si la suite des éléments dont les entiers strictement compris entre i' et i sont des sorties est un mot φ (éventuellement vide), il existe $A \in N$ et deux mots η, Ψ sur $V \cup \{\sigma\}$ tels que $A ::= \Psi\tilde{\varphi}$, $\Psi \neq \Lambda$, $\sigma u_{i-1} = \eta\Psi$ et $\eta \in \mathcal{C}(A, \gamma)$.

Pour démontrer C5, on applique C3 à la première entrée supérieure à i (entrée de A).

Si $\varphi \neq \Lambda$, $\tilde{\varphi}$ est un chemin du pseudo-graphe de production $(J, \Gamma ; g)$ dont l'origine est dans J' et dont l'extrémité est un point w de J, unique d'après PG2' ; A étant donné, les mots Ψ non vides pour lesquels $A ::= \Psi\tilde{\varphi}$ sont ceux tels que $\Psi\bar{A}$ soit un chemin d'origine w : ils forment un langage de Kleene, qui ne dépend que de A et w, L_{Aw}. L'ensemble, noté $\mathcal{C}'(\varphi, \gamma)$ ou $\mathcal{C}'(w, \gamma)$, des mots $\eta\Psi$ tels que $\Psi \neq \Lambda$ et qu'il existe $A \in N$ pour lequel $A ::= \Psi\tilde{\varphi}$ et $\eta \in \mathcal{C}(A, \gamma)$, est la réunion, pour $A \in N$, des ensembles $\mathcal{C}(A, \gamma) L_{Aw}$. Si $\varphi = \Lambda$, les mots Ψ tels que $A ::= \Psi$ forment un langage de Kleene L_A. L'ensemble, noté $\mathscr{C}'(\Lambda, \gamma)$ ou $\mathscr{C}'(\sigma, \gamma)$ des mots $\eta\ \Psi$ tels qu'il existe $A \in N$ pour lequel $A ::= \Psi$ et $\eta \in \mathscr{C}(A, \gamma)$ est la réunion des $\mathscr{C}(A, \gamma)\ L_A$. Que la grammaire soit ou non généralisée, on peut étendre le théorème 7. 4, d'après le théorème 4. 1 et le paragraphe 4.6.2 ; comme nous emploierons les mots réfléchis des états de la pile, nous énoncerons un résultat portant sur les ensembles $\tilde{\mathscr{C}}(Z, \gamma)$ et $\tilde{\mathscr{C}}'(w, \gamma)$ des réfléchis des mots de $\mathscr{C}(Z, \gamma)$ et $\mathscr{C}'(w, \gamma)$:

Théorème 7. 5. *Etant donné un entier* $n \geqslant 0$, *pour* $Z \in V$, $w \in J_1 = J \cup \{\sigma\}$, $\gamma \in T'^*$ *et* $|\gamma| \leqslant n$, *les ensembles* $\tilde{\mathscr{C}}(Z, \gamma)$ *et* $\tilde{\mathscr{C}}'(w, \gamma)$ *sont certains des langages des propositions d'une structure droite de Kleene, dérivant de ses divers symboles auxiliaires.*

Avec ces définitions, C5 devient :
(C'5) soit i une sortie de U autre que la dernière, i' la dernière entrée inférieure à i, γ un facteur gauche de $\alpha_{i'}^{U} \sigma^r$; si la suite des éléments dont les entiers strictement compris entre i' et i sont des sorties est un mot φ, $\sigma\ u_{i-1}$ appartient à $\mathscr{C}'(\varphi, \gamma)$.

7. 7. Générateurs utilisant des contextes.

7. 7. 1. Nous serons amenés à faire dépendre certains états d'une pile de l'état précédent, si long soit-il. Aussi ne pourrons nous nous contenter d'employer un transducteur fini ; nous utiliserons la généralisation introduite en 5. 4. La "partie droite" γ d'un contexte (η, γ) sera un mot de longueur bornée, ce qui aura pour effet de remplacer la valeur 1 de cd par une valeur quelconque. Il est inutile de considérer un contexte dans le cas où, pour le transducteur tr^1, la réunion des ensembles m(s) et d(s, Z) contient au plus un élément (cf. 7.5.2, où en particulier est défini Det).

Définissons donc des générateurs de piles gn^2 par $W = V$, un entier $cd \geqslant 1$ et le transducteur fini tr^2 donné par :

a) $M = E \times J_1 \times T'^{cd}$; dans les définitions qui suivent, α et b décrivent T', γ décrit T'^{cd-1} (on convient que $T'^0 = \{\Lambda\}$), Y et Z décrivent E, w décrit J et ν décrit E^*.
b) $q = cd + 1$; $l(\alpha \gamma b) = (b, \sigma, \alpha\gamma)$.
c) Si $b \succ \alpha$ et $\neg(b \prec \alpha)$, $d[(b, \sigma, \alpha\gamma), Z] = \{(Z, o(b), \alpha\gamma)\}$;
pour $(Y, w) \in$ Det, $d[(Y, w, \alpha\gamma), Z] = \{(Z, w', \alpha\gamma)\}$ si $w \sqsubset w'$ et $g(w') = Y$.
d) Si $b \prec \alpha$ et $\neg(b \succ \alpha)$, $m(b, \sigma, \alpha\gamma) = \{(s_f, {}'\alpha')\}$;
pour $(Y, w) \in$ Det, $m(Y, w, \alpha\gamma) = \{[(Y, w', \alpha\gamma), {}'A']\}$ si $w \sqsubset w'$, $g(w') = \bar{A}$ et $Y \prec A$ (le couple w', A est unique s'il existe puisque $(Y, w) \in$ Det) ;
pour $w \in J''$ et $g(w) = \bar{A}$, $m(Y, w, \alpha\gamma)$ a pour seul élément :
- $(s_f, {}'\alpha')$ si $A \prec \alpha$ et $\neg(A \succ \alpha)$, ou si $Y = \alpha = \sigma$ et $A \in N'$,
- $[(Y, o(A), \alpha\gamma), {}'\sigma']$ si $A \succ \alpha$ et $\neg(A \prec \alpha)$, sauf si $Y = \alpha = \sigma$ et $A \in N'$.
e) Pour $b \prec \alpha$ et $b \succ \alpha$ (d'où $\alpha \neq \sigma$), $k[(b, \sigma, \alpha\gamma), Z\ \nu]$ contient $[(Z, o(b), \alpha\gamma), 1, {}'\sigma']$ si $b\ Z\ \nu \in \tilde{\mathscr{C}}'(\sigma, \alpha\gamma)$, $[s_f, 0, {}'\alpha']$ si $b\ Z\ \nu \in \tilde{\mathscr{C}}(\alpha, \gamma)$, et pas d'autre élément ;
pour $(Y, w) \notin$ Det et $w \notin J''$, $k[(Y, w, \alpha\gamma), Z\ \nu]$ contient les seuls éléments :
- $[(Z, w', \alpha\gamma), 1, {}'\sigma']$ si $Y\ Z\ \nu \in \tilde{\mathscr{C}}'(w, \alpha\gamma)$, w' étant le point de $\Gamma(w)$, s'il existe, tel que $g(w') = Y$,
- $[(Y, w', \alpha\gamma), 0, {}'A']$ pour tous les $w' \in \Gamma(w)$ et tous les $A \in N$ tels que $g(w') = \bar{A}$ et $Y\ Z\ \nu \in \tilde{\mathscr{C}}(A, \alpha\gamma)$;
pour $w \in J''$, $g(w) = \bar{A}$, $A \prec \alpha$ et $A \succ \alpha$ (d'où $\alpha \neq \sigma$), $k[(Y, w, \alpha\gamma), Z\ \nu]$ contient les seuls éléments :
- $[s_f, 0, {}'\alpha']$ si $A\ Y\ Z\ \nu \in \tilde{\mathscr{C}}(\alpha, \gamma)$ (d'où $\alpha \neq \sigma$),
- $[(Y, o(A), \alpha\gamma), 0, {}'\sigma']$ si $A\ Y\ Z\ \nu \in \tilde{\mathscr{C}}'(\sigma, \alpha\gamma)$.

7. 7. 2. Vérifions que l'application k est un automate fini. Envisageons les ensembles $\tilde{\mathscr{C}}(Z, \gamma)$ et $\tilde{\mathscr{C}}'(w, \gamma)$ pour $Z \in V$, $w \in J_1$, $\gamma \in E^*$ et $|\gamma| \leqslant cd$: d'après les théorèmes 7.5 et 5.1, il existe une fonction d'automate fini k', d'alphabet E, dont l'ensemble des états S' contient s_f et les couples (Z, γ), (w, γ) et telle que, pour que η appartienne à $\tilde{\mathscr{C}}(Z, \gamma)$ (resp. $\tilde{\mathscr{C}}'(w, \gamma)$), il soit nécessaire et suffisant que s_f appartienne à $k[(Z, \gamma), \eta]$ (resp. $k'[(w, \gamma), \eta]$). Pour engendrer l'état initial (Z, γ) ou (w, γ), nous effectuerons une composition à gauche (théorème 5.2) ; pour engendrer l'état final, nous effectuerons le produit de k' par d'autres fonctions d'automate et une composition à droite (théorème 5.3).

Il est facile de construire une fonction d'automate k'' qui à l'un de ses états s_0 et à chaque mot ν sur E de longueur 2 au moins, associe le second élément de ν : on prend pour ensemble d'états $S'' = E \cup \{s_0, s_1\}$ où s_0, s_1 sont distincts et n'appartiennent pas à E, et pour $k''(s, \nu)$ l'ensemble ayant pour seul élément : - le deuxième élément de ν si $s = s_0$ et $|\nu| \geqslant 2$;
- s_1 si $s = s_0$ et $|\nu| = 1$;
- le premier élément de ν si $s = s_1$ et $\nu \neq \Lambda$;
- s si $s \in E$ ou $\nu = \Lambda$.

Soit k''' la fonction d'automate d'alphabet E, d'ensemble d'états $J_1 \times T'^{cd}$ telle que pour tout couple (s, ν), $k'''(s, \nu) = \{s\}$.

Définissons une application h_1 de M dans l'ensemble des parties finies de $(S' \times S'' \times J_1 \times T'^{cd} \times E^*$; α, b, γ, Y, w ont la même signification qu'en 7. 7. 1 :
- pour $b \prec \alpha$ et $b \succ \alpha$, $h_1(b, \sigma, \alpha\gamma)$ est formé des deux éléments $[((\sigma, \alpha\gamma), s_0, o(b), \alpha\gamma), {}'b']$ et $[((\alpha, \gamma), \sigma, \sigma, \alpha\gamma), {}'b']$;
- pour $(Y, w) \notin$ Det et $w \notin J''$, $h_1(Y, w, \alpha\gamma)$ contient les seuls éléments : $[((w, \alpha\gamma), s_0, w', \alpha\gamma), {}'Y']$ si un point w' de $\Gamma(w)$ vérifie $g(w') = Y$, $[((A, \alpha\gamma), Y, w', \alpha\gamma), {}'Y']$ pour tous les $w' \in \Gamma(w)$, $A \in N$ tels que $g(w) = \bar{A}$;

- pour $w \in J''$, $g(w) = \bar{A}$, $A \prec \alpha$ et $A \succ \alpha$, $h_1(Y, w, \alpha\gamma)$ est formé des deux éléments $[((\alpha, \gamma), \sigma, \sigma, \alpha\gamma), {}'AY']$ et $[((\sigma, \alpha\gamma), Y, o(A), \alpha\gamma), {}'AY']$;
- pour les autres $s \in M$, $h_1(s) = \phi$.

Enfin, définissons une application h_2 de $S' \times S'' \times J_1 \times T'^{cd}$ dans l'ensemble des parties de $M \times \{0,1\} \times E$ (E est identifié à l'ensemble des mots sur E de longueur 1) : $h_2(s, Z, w_1, \alpha\gamma)$ est vide si $s \neq s_f$ ou $Z \notin E$; si $s = s_f$ et $Z \in E$, il contient un seul élément :

$[(Z, w_1, \alpha\gamma), 0, {}'A']$ si $w_1 \in J''$ et $g(w_1) = \bar{A}$,
$[(Z, w_1, \alpha\gamma), 0, {}'\sigma']$ si $w_1 = o(A) \in J'$,
$[(Z, w_1, \alpha\gamma), 1, {}'\sigma']$ dans les autres cas où $w_1 \in J$,
$[s_f, 0, {}'\alpha']$ si $w_1 = \sigma$.

On vérifie facilement que k est l'automate fini d'ensemble initial M, d'ensemble final $M \times \{0,1\} \times E$, obtenu par composition de la fonction d'automate $k' \times k'' \times k'''$ à gauche par h_1 et à droite par h_2.

7. 7. 3. Les transducteurs tr^2 et les générateurs gn^2 dépendent du paramètre cd. Si $cd' < cd$, le générateur associé au paramètre cd est contenu dans le générateur associé à cd', car l'application ϖ de $(E \times J_1 \times T'^{cd}) \times E^*$ dans $(E \times J_1 \times T'^{cd'}) \times E^*$: $\varpi[(Y, w, \gamma_1\gamma_2), v] = [(Y, w, \gamma_1), v]$ projette le transducteur associé à cd dans le transducteur associé à cd'. D'autre part, si $cd = 1$, l'identité projette tr^2 dans le transducteur fini tr^1 car $Y Z v \in \mathscr{C}(A, \alpha)$ entraîne $Y \prec A$ (lemme 2).

Par conséquent, quel que soit cd, les transducteurs tr^2 vérifient GP2 dès que DA est satisfaite ; les générateurs gn^2 vérifient GP1 et les piles de l'ensemble $gn^2(\alpha)$ sont conjointes à des pseudo-arborescences de $\mathscr{C}(G)$ dont α est la suite de feuilles. Nous allons démontrer que réciproquement les piles conjointes à toutes ces pseudo-arborescences appartiennent à $gn^2(\alpha)$. Il en résulte, comme nous l'avions annoncé, que les générateurs gn^0 et gn^1 engendrent aussi les piles cherchées.

Soit U la pile conjointe à une pseudo-arborescence de $\mathscr{C}(G)$, dont $\alpha = {}'\alpha_0 \ldots \alpha_{p-1}'$ est la suite de feuilles. Les entrées dans U de symboles terminaux forment, avec 0 et le nombre d'états de U, un p-partage ; nous utilisons les notations de 5. 2 : $\alpha_i = {}'\alpha_i \ldots \alpha_{i+cd-1}'$ avec $\alpha_{p+i'} = \sigma$ pour $i' \geqslant 0$; si $1 \leqslant i \leqslant p$, α_{i-1} est le sommet de v_i. Lorsque $j_{i+1} = j_i + 1$, $i \neq p$ (C1): $i = 0$ et $\sigma \prec \alpha_0$, ou bien $i \neq 0$ et $\alpha_{i-1} \prec \alpha_i$ (lemmes 3,2) ; dans les deux cas, $\sigma v_i \in \mathscr{C}(\alpha_i, {}'\alpha_{i+1} \ldots \alpha_{i+cd-1}')$, d'après C4 ; $\mu_i = {}'\alpha_i'$ est le résultat d'un calcul complet d'entrée $\alpha_i \tilde{v}_i \sigma$, de longueur 1. Supposons maintenant $j_{i+1} \neq j_i + 1$, d'où $i \neq 0$, avec $v_i = {}'x_r \ldots x_1'$ (on posera $x_{r+i'} = \sigma$ pour $i' > 0$). Entre j_i et j_{i+1} se trouvent p' entrées de symboles non terminaux (par exemple $j_{i+1}-1$), $A_1, \ldots, A_{p'}$. D'après C2 et C3, il existe p' entiers r_i, pour lesquels : $1 \leqslant r_1 \leqslant \ldots \leqslant r_{p'} \leqslant r$; j_i est suivi de sorties de $x_1, \ldots, x_{r_1}$ et d'une entrée de A_1 ; pour $1 \leqslant i' < p'$, l'entrée de $A_{i'}$ est suivie d'une sortie de $A_{i'}$, de sorties de $x_{r_{i'}+1}, \ldots, x_{r_{i'+1}}$ si $r_{i'+1} \neq r_{i'}+1$ et d'une entrée de $A_{i'+1}$; l'entrée de $A_{p'}$ est suivie de celle de $\alpha_i \neq \sigma$ ou d'une sortie si $\alpha_i = \sigma$:

$$\mu_i = \sigma^{r_1} A_1 \ \sigma \ \sigma^{r_2 - r_1} \ \ldots \ \sigma \ \sigma^{r_{p'} - r_{p'-1}} \ A_{p'} \ \alpha_i.$$

$l(\alpha_i x_1) = (x_1, \sigma, \alpha_i)$; $x_1 \succ \alpha_i$ (lemmes 4,5) et $\sigma v_i \in \mathscr{C}'(\sigma, \alpha_i)$ (C'5) : selon que $x_1 \prec \alpha_i$ ou non, $k[(x_1, \sigma, \alpha_i), {}'x_2 \ldots x_r \sigma']$ contient $[(x_2, o(x_1), \alpha_i), 1, {}'\sigma']$, ou $d[(x_1, \sigma, \alpha_i), x_2]$ contient $(x_2, o(x_1), \alpha_i)$. Pour montrer que μ_i est le résultat d'un calcul complet d'entrée $\alpha_i \tilde{v}_i \sigma$, il suffit de montrer que sous les cinq hypothèses :

- j' est une entrée de $A \in N$ et j'' la dernière entrée précédant j',
- $j''+1, \ldots j'-1$ sont des sorties de $z_1, \ldots, z_{j'-j''-1}$,
- z' est le dernier élément de $\sigma u_{j'-1}$,
- $\sigma u_{j''+1} = v z''$ ($z'' = z'$ si $j''+1 = j'-1$, $z'' = z_2$ sinon),
- $\alpha \in T'$, $\gamma \in T'^{cd-1}$, $\alpha\gamma$ est facteur gauche de $\alpha_{j''}^{U} \sigma^{cd}$,

il existe un calcul d'entrée $\tilde{v}$, de premier état $(z'', o(z_1), \alpha\gamma)$, de dernier état $(z', o(A), \alpha\gamma)$ et de résultat $\sigma^{j'-j''-2} A \sigma$ lorsque $j'+1$ est une sortie autre que la dernière, de dernier état s_f et de résultat $\sigma^{j'-j''-2} A \alpha$ lorsque $j'+1$ est une entrée ou la dernière sortie . Cela résulte immédiatement des propriétés suivantes :

- $z_1 \ldots z_{j'-j''-1} \bar{A}$ est un chemin du pseudo-graphe de production (C3) ;
- pour $1 \leqslant i'' \leqslant j'-j''-2$, $\sigma u_{j''+i''} \in \mathscr{C}({}'z_1 \ldots z_{i''}', \alpha\gamma)$ (C'5) ;
- $z' \prec A$ (lemmes 2,3) et $\sigma u_{j'-1} \in \mathscr{C}(A, \alpha\gamma)$ (C4) ;
- si $j'+1$ est une sortie, $A \succ \alpha$ (lemmes 4,5) ;
- si $j'+1$ est une sortie autre que la dernière et si $\alpha = \sigma$ et $u_{j'-1} = \Lambda$, $j'+2$ est une entrée, $j'+3$ une sortie, et ainsi de suite jusqu'à la dernière sortie : A dérive de la racine (C3) ; d'après DA et PA, A n'est pas un axiome;
- si $j'+1$ est une sortie autre que la dernière, $\sigma u_{j'} \in \mathscr{C}(\sigma, \alpha\gamma)$ (C'5) ;
- si $j'+1$ est une entrée, c'est une entrée de α (C2), $A \prec \alpha$ (lemme 2), $\sigma u_{j'} \in \mathscr{C}(\alpha, \gamma)$ (C4) ;
- si $j'+1$ est la dernière sortie, $\alpha = \sigma$, $u_{j'-1} = \Lambda$ et $A \in N'$.

Théorème 7. 6. *Les générateurs gn^2 associés à une grammaire pluriaxiomatique $G = (N, T, ::=, N')$ vérifiant les conditions DA et PA sont tels que, pour tout mot α sur T, $gn^2(\alpha)$ soit l'ensemble des piles conjointes aux pseudo-arborescences de $\mathscr{C}(G)$ dont α est la suite de feuilles.*

La figure 7. 5 schématise le fonctionnement des générateurs gn^2.

7. 7. 4. Déterminisme du transducteur tr^2. Pour que tr^2 soit déterministe, il faut et il suffit que, pour tout $s \in M$ et tout $v \in E^*$, l'ensemble $k(s, v)$ contienne au plus un élément ; pour cela, il suffit que simultanément :

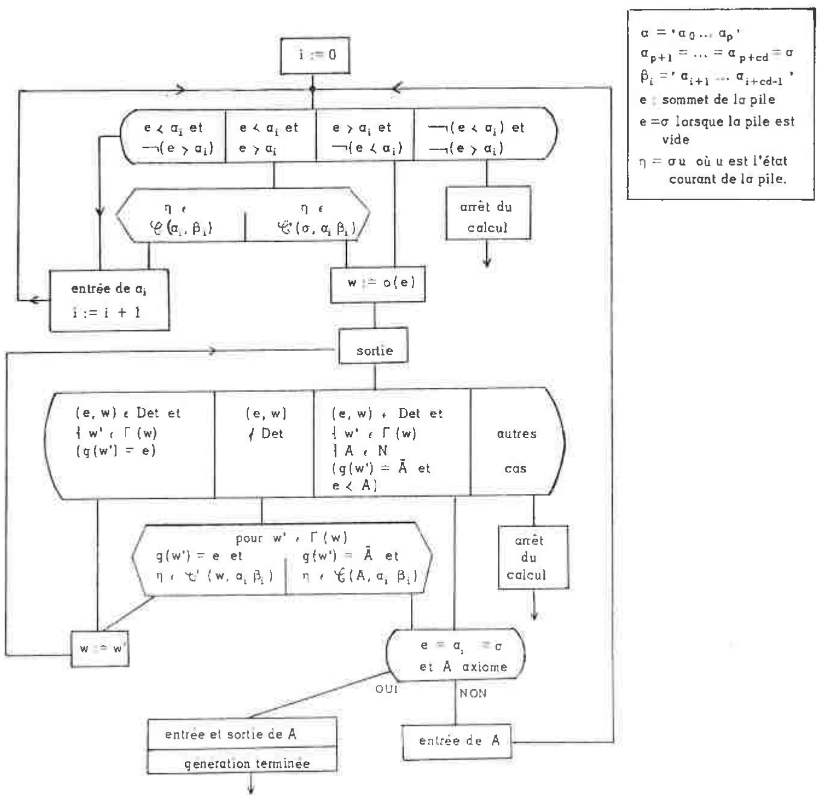


Figure 7. 5

(CD 1) pour $\alpha \in T$, $\gamma \in T'^{cd-1}$, $\mathscr{C}(\alpha, \gamma)$ et $\mathscr{C}'(\sigma, \alpha\gamma)$ soient disjoints ;

(CD 2) pour $\gamma' \in T'^{cd}$, $w \in J$ et $A \in N$ tels qu'un point w' de $\Gamma(w)$ vérifie $g(w') = \bar{A}$, $\mathscr{C}(A, \gamma')$ et $\mathscr{C}'(w, \gamma')$ soient disjoints ;

(CD 3) pour $\gamma' \in T'^{cd}$, $w \in J$, $A \in N$ et $A' \in N$ tels que deux points distincts w' et w'' de $\Gamma(w)$ vérifient $g(w') = \bar{A}$, $g(w'') = \bar{A}'$, $\mathscr{C}(A, \gamma')$ et $\mathscr{C}(A', \gamma')$ soient disjoints.

Ces conditions sont décidables [2]. On peut aussi les écrire, d'après 7. 6. 2 :

(CD 1) pour $\alpha \in T$, $\gamma' \in T'^{cd-1}$ et toute règle $A ::= \Psi$, les ensembles $\mathscr{C}(\alpha, \gamma)$ et $\mathscr{C}(A, \alpha\gamma)\Psi$ sont disjoints ;

(CD 2 et 3) pour $\gamma' \in T'^{cd}$, $A \in N$, $A' \in N$ et $\Psi \in V^*$ vérifiant $A \neq A'$ ou $\Psi \neq \Lambda$, tels qu'il existe un mot φ de V^* pour lequel $A ::= \varphi$ et $A' ::= \Psi\varphi$, $\mathscr{C}(A, \gamma')$ et $\mathscr{C}(A', \gamma')\Psi$ sont disjoints (CD 3 correspond à $\Psi = \Lambda$).

Intuitivement, la première condition assure qu'on peut choisir entre une sortie et l'entrée d'un symbole terminal, la seconde entre une sortie (précédée d'un certain nombre d'autres sorties) et l'entrée d'un symbole non terminal, et la troisième entre deux symboles non terminaux susceptibles d'entrer (cf. figure 7. 5).

Nous allons montrer qu'inversement, si la grammaire est réduite et si l'une de ces conditions n'est pas réalisée, on peut trouver deux mots α et α' de T^* tels que $\alpha\sigma^{cd}$ et $\alpha'\sigma^{cd}$ aient un facteur gauche commun $\beta\gamma'$ où $|\gamma'| = cd$ et qu'il existe une pile U de $gn^2(\alpha)$, une pile U' de $gn^2(\alpha')$ et un entier i pour lesquels $u_i \neq u'_i$ et $\beta = \alpha^U_{0i} = \alpha^{U'}_{0i}$. Par conséquent :

Théorème 7. 7. *Si une grammaire G réduite et un entier cd > 1 vérifient les conditions CD 1, CD 2, CD 3, ils définissent un transducteur tr^2 déterministe. Dans le cas contraire, aucun générateur de piles dans la définition duquel entre l'entier cd, qui engendre les piles de $\mathscr{J}(G)$ avec comme donnée la suite des symboles terminaux y entrant et comme p-partage la suite de leurs entrées, n'est à pile unique.*

Supposons d'abord $A ::= \varphi$, $A' ::= \Psi\varphi$, $A \neq A'$ ou $\Psi \neq \Lambda$, $\sigma\eta \in \mathscr{C}(A, \gamma')$, $\eta = \eta'\Psi$, $\sigma\eta' \in \mathscr{C}(A', \gamma')$. ηA et $\eta' A'$ sont des états de deux piles de $\mathscr{J}(G)$: d'après 7. 2. 2, il existe deux axiomes X et X', une X-dérivation d'un mot $\eta A \gamma_1$ et une X-dérivation d'un mot $\eta' A' \gamma'_1$, γ' étant facteur gauche de $\gamma_1 \sigma^{cd}$ et $\gamma'_1 \sigma^{cd}$. Comme la grammaire est réduite, il existe aussi une dérivation $(\delta_0 = \eta\varphi, \delta_1, \dots, \delta_p = \beta)$ où $\beta \in T^*$, d'où deux dérivations $(X, \dots, \eta A \gamma_1, \eta\varphi\gamma_1, \delta_1\gamma_1, \dots, \beta\gamma_1)$ et $(X', \dots, \eta' A' \gamma'_1, \eta\varphi\gamma'_1, \delta_1\gamma'_1, \dots, \beta\gamma'_1)$. On en déduit (lemme 1) deux piles U et U' conjointes à deux pseudo-arborescences de $\mathscr{E}(G)$ dont $\beta\gamma_1$ et $\beta\gamma'_1$ sont les suites de feuilles, et un entier i, tels que $u_i = \eta A$, $u'_i = \eta' A'$, $\alpha^U_{0i} = \alpha^{U'}_{0i} = \beta$.

Supposons maintenant $A ::= \Psi$, $\sigma\eta \in \mathscr{C}(\alpha, \gamma)$, $\eta = \eta'\Psi$, $\sigma\eta' \in \mathscr{C}(A, \alpha\gamma)$. $\eta\alpha$ est un état u''_j d'une pile U'' de $\mathscr{F}(G)$ et γ est facteur gauche de $\alpha^{U''}_{j-}\sigma^{cd}$ soit B le premier symbole auxiliaire dont une entrée dans U'' est supérieure à j : il existe un axiome X et une X-dérivation dont les deux derniers mots sont $\eta''B\gamma_1$ et $\eta''\varphi\gamma_1 = \eta\alpha\gamma_2\gamma_1$ $(B ::= \varphi)$, γ étant facteur gauche de $\gamma_2\gamma_1\sigma^{cd}$. Comme $\sigma\eta' \in \mathscr{C}(A, \alpha\gamma)$ et $A ::= \Psi$, il existe un axiome X' et une X'-dérivation dont les deux derniers mots sont $\eta'A\alpha\gamma'_1$ et $\eta'\Psi\alpha\gamma'_1 = \eta\alpha\gamma'_1$, γ étant facteur gauche de $\gamma'_1\sigma^{cd}$. D'où deux dérivations $(X, \ldots, \eta''B\gamma_1, \eta\alpha\gamma_2\gamma_1, \delta_1\alpha\gamma_2\gamma_1, \delta_2\alpha\gamma_2\gamma_1, \ldots, \beta\alpha\gamma_2\gamma_1)$ et $(X', \ldots, \eta'A\alpha\gamma'_1, \eta\alpha\gamma'_1, \delta_1\alpha\gamma'_1, \delta_2\alpha\gamma'_1, \ldots, \beta\alpha\gamma'_1)$, où $\beta \in T^*$. De la première, on déduit (lemme 1) une pile U qui contient un état $u_{j'} = \eta\varphi = \eta\alpha\gamma_2$, avec $\alpha^{U}_{j'-} = \gamma_1$, j' étant précédé des entrées des symboles terminaux qui forment $\alpha\gamma_2$; cette pile possède donc deux états $u_{i-1} = \eta$, $u_i = \eta\alpha$, avec $\alpha^{U}_{i-} = \gamma_2\gamma_1$ et donc $\alpha^{U}_{0i} = \beta$. De la seconde, on déduit une pile U' telle que $u'_{i-1} = \eta$, i soit une sortie et $\alpha^{U'}_{i-} = \alpha\gamma'_1$, d'où $\alpha^{U'}_{0i} = \beta$.

7. 7. 5. Remarque. Une variante du générateur gn^2, pour $cd = 1$, n'utilise les ensembles $\mathscr{C}(Z, \gamma)$ et $\mathscr{C}'(w, \gamma)$ que pour $\gamma = \Lambda$: dans la partie (e) de la définition de tr^2, il suffit de remplacer $\widetilde{\mathscr{C}}'(\sigma, \alpha\gamma)$ par $\widetilde{\mathscr{C}}'(\sigma, \Lambda)$, $\widetilde{\mathscr{C}}'(w, \alpha\gamma)$ par $\widetilde{\mathscr{C}}'(w, \Lambda)$, $\widetilde{\mathscr{C}}(A, \alpha\gamma)$ par $\widetilde{\mathscr{C}}(A, \Lambda)$. L'identité projette dans tr^1 le transducteur tr'^2 ainsi obtenu et projette tr^2, avec $cd = 1$, dans tr'^2. Le générateur engendre encore les mêmes piles. Mais ses conditions de déterminisme sont plus strictes que celles obtenues ci-dessus : si $(Y, w) \notin$ Det et si un point w' de $\Gamma(w)$ vérifie $g(w') = \bar{A}$, il faut que $\mathscr{C}(A, \Lambda)$ et $\mathscr{C}'(w, \Lambda)$ soient disjoints, ce qui est plus strict que : pour tout $\alpha \in T'$, $\mathscr{C}(A, \alpha)$ et $\mathscr{C}'(w, \alpha)$ sont disjoints.

7. 8. Générateurs utilisant des contextes bornés.

7. 8. 1. Les transducteurs tr^2 décident comme poursuivre un calcul en cours, alors qu'il y avait plusieurs possibilités pour le tranducteur tr^1, en fonction de l'appartenance d'un mot à un langage de Kleene parmi plusieurs. Les états demandant un choix ont été caractérisés au paragraphe 7. 5. 2 : on tentera ici de remédier aussi "économiquement" que possible à chacun de ces cas d'indéterminisme ; plus précisément, on choisira, chaque fois qu'on le pourra, grâce à un facteur gauche du mot considéré ayant une longueur bornée.

Les générateurs de piles gn^3 qui vont être définis seront donnés par $W = V$, un entier $cd \geq 1$ et des transducteurs tr^3 qui dépendent d'un certain nombre de sous-ensembles finis de E^* séparant un langage de Kleene d'un autre (4. 5. 6). L'ensemble des phrases d'un langage de Kleene K qui commencent par un élément Y donné est un langage de Kleene (cela résulte immédiatement, par exemple, du théorème 4. 4) : nous noterons K_Y ce langage.

1) Pour tout couple $Y \in V$, $\alpha \in T$ tel que $Y \nless \alpha$ et $Y \gtrdot \alpha$ et tout mot γ de T'^{cd-1}, soit $Q(Y, \alpha\gamma)$ un ensemble séparant $\widetilde{\mathscr{C}}(\alpha, \gamma)_Y$ et $\widetilde{\mathscr{C}}'(\sigma, \alpha\gamma)_Y$ si ces langages ont une borne d'intersection.

2) Envisageons un mot γ' de T'^{cd} et un couple (Y, w) de $E \times J$ tel que $g(w) \notin \bar{N}$, n'appartenant pas à l'ensemble Det (7. 5. 2) : il existe $p \geq 1$ symboles auxiliaires A pour lesquels A est l'image par g d'un point de $\Gamma(w)$ et $Y \lessdot A$; on les ordonne : $A_1, \ldots, A_p$. On construit p ensembles $Q(Y, w, A_i, \gamma')$, pour $i = 1, \ldots p$, séparant $\widetilde{\mathscr{C}}(A_i, \gamma')_Y$ de la réunion de $\widetilde{\mathscr{C}}'(w, \gamma')_Y$ et des $\widetilde{\mathscr{C}}(A_j, \gamma')_Y$ pour $j > i$, et un ensemble $Q(Y, w, \gamma')$ séparant la réunion des $\widetilde{\mathscr{C}}(A_i, \gamma')_Y$ de $\widetilde{\mathscr{C}}'(w, \gamma')_Y$, lorsque $\widetilde{\mathscr{C}}'(w, \gamma')_Y$ et les p langages $\widetilde{\mathscr{C}}(A_i, \gamma')_Y$ ont une borne d'intersection : on peut prendre pour $Q(Y, w, \gamma')$ la réunion des $Q(Y, w, A_i, \gamma')$.

La recherche des ensembles $Q(Y, \alpha\gamma)$ peut être facilitée par la remarque suivante : si γ_1 est facteur gauche de γ, $\mathscr{C}(\alpha, \gamma_1)_Y$ contient $\widetilde{\mathscr{C}}(\alpha, \gamma)_Y$, $\widetilde{\mathscr{C}}'(\sigma, \alpha\gamma_1)_Y$ contient $\mathscr{C}'(\sigma, \alpha\gamma)_Y$, donc tout ensemble séparant $\widetilde{\mathscr{C}}(\alpha, \gamma_1)_Y$ de $\widetilde{\mathscr{C}}'(\sigma, \alpha\gamma_1)_Y$ est un ensemble séparant $\widetilde{\mathscr{C}}(\alpha, \gamma)_Y$ de $\widetilde{\mathscr{C}}'(\sigma, \alpha\gamma)_Y$. On peut utiliser une remarque analogue pour déterminer les $Q(Y, w, A_i, \gamma')$ et $Q(Y, w, \gamma')$.

Soit cg le maximum des longueurs des mots appartenant à l'un des ensembles $Q(Y, \alpha\gamma)$, $Q(Y, w, A_i, \gamma')$, $Q(Y, w, \gamma')$. Le transducteur tr^3 associé à ces ensembles est défini de la manière suivante :

a) $M = E^{cg} \times J_1 \times T'^{cd}$. Dans les définitions qui suivent, α et b décrivent T', η décrit E^{cg-1}, γ décrit T'^{cd-1}, Y et Z décrivent E, w décrit J et ν décrit E^*. Nous désignerons par Det' l'ensemble formé des éléments :

- $(b, \sigma, \alpha\gamma)$ si $b \lessdot \alpha$, $b \gtrdot \alpha$, $\widetilde{\mathscr{C}}(\alpha, \gamma)_b$ et $\widetilde{\mathscr{C}}'(\sigma, \alpha\gamma)_b$ ont une borne d'intersection,
- $(Y, w, \alpha\gamma)$ si $g(w) \in \bar{N}$, le symbole auxiliaire A tel que $g(w) = \bar{A}$ vérifie $A \lessdot \alpha$ et $A \gtrdot \alpha$, $\widetilde{\mathscr{C}}(\alpha, \gamma)_A$ et $\widetilde{\mathscr{C}}'(\sigma, \alpha\gamma)_A$ ont une borne d'intersection,
- $(Y, w, \alpha\gamma)$ si $g(w) \notin \bar{N}$, $(Y, w) \notin$ Det et si pour les points w'_i $(i = 1, \ldots, p)$ de $\Gamma(w)$ tels que $g(w'_i) = \bar{A}_i \in \bar{N}$, les $p + 1$ langages $\widetilde{\mathscr{C}}(A_i, \alpha\gamma)_Y$ et $\widetilde{\mathscr{C}}'(w, \alpha\gamma)_Y$ ont une borne d'intersection.

b) $q = cd + cg$; $l(\alpha\gamma b\eta) = (b\eta, \sigma, \alpha\gamma)$.

c) Si $b \gtrdot \alpha$ et $\neg(b \lessdot \alpha)$ ou si $(b, \sigma, \alpha\gamma) \in$ Det' et aucun facteur gauche de $b\eta$ n'appartient à $Q(b, \alpha\gamma)$, $d[(b\eta, \sigma, \alpha\gamma), Z] = \{(\eta Z, o(b), \alpha\gamma)\}$;
si $(Y, w) \in$ Det ou si $(Y, w, \alpha\gamma) \in$ Det' et aucun facteur gauche de $Y\eta$ n'appartient à $Q(Y, w, \alpha\gamma)$, $d[(Y\eta, w, \alpha\gamma), Z] = \{(\eta Z, w', \alpha\gamma)\}$ si le point w' vérifie $w \,\Gamma\, w'$ et $g(w') = Y$.

d) Si $b \lessdot \alpha$ et $\neg(b \gtrdot \alpha)$ ou si $(b, \sigma, \alpha\gamma) \in$ Det' et un facteur gauche de $b\eta$ appartient à $Q(b, \alpha\gamma)$, $m(b\eta, \sigma, \alpha\gamma) = \{(s_f, \text{'}\alpha\text{'})\}$;
si $(Y, w) \in$ Det et $Y \lessdot A$ ou si $(Y, w, \alpha\gamma) \in$ Det' et A est le A_i d'indice i minimum tel qu'un facteur gauche de $Y\eta$ appartienne à $Q(Y, w, A_i, \alpha\gamma)$, $m(Y\eta, w, \alpha\gamma) = \{[(Y\eta, w', \alpha\gamma), \text{'}A\text{'}]\}$ si le point w' vérifie $w \,\Gamma\, w'$ et $g(w') = \bar{A}$;
pour $w \in J''$ et $g(w) = \bar{A}$:
- si $A \lessdot \alpha$ et $\neg(A \gtrdot \alpha)$, ou si $(Y, w, \alpha\gamma) \in$ Det' et un facteur gauche de $AY\eta$ appartient à $Q(A, \alpha\gamma)$, ou encore si $Y = \alpha = \sigma$ et $A \in N'$, $m(Y\eta, w, \alpha\gamma) = \{s_f, \text{'}\alpha\text{'}\}$,
- si $A \gtrdot \alpha$ et $\neg(A \lessdot \alpha)$, sauf pour $Y = \alpha = \sigma$ et $A \in N'$, ou si $(Y, w, \alpha\gamma) \in$ Det' et aucun facteur gauche de $AY\eta$ n'appartient à $Q(A, \alpha\gamma)$, $m(Y\eta, w, \alpha\gamma) = \{[(Y\eta, o(A), \alpha\gamma), \text{'}\sigma\text{'}]\}$.

e) Si $b \prec \alpha$, $b \succ \alpha$ et $(b, \sigma, \alpha\gamma) \notin$ Det', $k[(b\eta, \sigma, \alpha\gamma), Z\nu]$ contient les seuls éléments :

- $[(\eta Z, o(b), \alpha\gamma), 1, '\sigma']$ lorsque $b\eta Z\nu \in \widetilde{\mathscr{C}}'(\sigma, \alpha\gamma)\sigma^{cg-1}$,
- $[s_f, 0, '\alpha']$ lorsque $b\eta Z\nu \in \widetilde{\mathscr{C}}(\alpha, \gamma)\sigma^{cg-1}$;

si $g(w) \notin \overline{N}$, $(Y, w) \notin$ Det et $(Y, w, \alpha\gamma) \notin$ Det', $k[(Y\eta, w, \alpha\gamma), Z\nu]$ contient les seuls éléments :

- $[(\eta Z, w', \alpha\gamma), 1, '\sigma']$ si $Y\eta Z\nu \in \widetilde{\mathscr{C}}'(w, \alpha\gamma)\sigma^{cg-1}$, w' étant le point de $\Gamma(w)$, s'il existe, tel que $g(w') = Y$,
- $[(Y\eta, w', \alpha\gamma), 0, 'A']$ pour tous les $w' \in \Gamma(w)$ et tous les $A \in N$ tels que $g(w') = \overline{A}$ et $Y\eta Z\nu \in \widetilde{\mathscr{C}}(A, \alpha\gamma)\sigma^{cg-1}$;

si $w \in J''$, $g(w) = \overline{A}$, $A \prec \alpha$, $A \succ \alpha$ et $(Y, w, \alpha\gamma) \notin$ Det', $k[(Y\eta, w, \alpha\gamma), Z\nu]$ contient les seuls éléments :

- $[s_f, 0, '\alpha']$ si $AY\eta Z\nu \in \widetilde{\mathscr{C}}(\alpha, \gamma)\sigma^{cg-1}$,
- $[(Y\eta, o(A), \alpha\gamma), 0, '\sigma']$ si $AY\eta Z\nu \in \widetilde{\mathscr{C}}'(\sigma, \alpha\gamma)\sigma^{cg-1}$.

On démontre comme en 7. 7. 2 que k est un automate fini.

7. 8. 2. Comparons tr^3, pour cd quelconque, au transducteur tr^0 (7. 3), puis tr^3 et le transducteur tr^2 (7. 7. 1) associé au même cd. Nous désignerons par χ^0, χ^2, χ^3 les applications χ attachées (5. 4. 2) à tr^0, tr^2, tr^3 ; plus généralement, nous affecterons de ^{0},^{1},^{3} les applications l, d, m, k, de tr^0, tr^2, tr^3.

Définissons une application ϖ par : $\varpi[(Y\eta, w, \alpha\gamma), \nu] = [(Y, w, \alpha), \nu_0]$ où $Y \in E$, $\eta \in E^{cg-1}$, $w \in J_1$, $\alpha \in T'$, $\gamma \in T'^{cd-1}$, $\nu \in E^*$ et ν_0 est déterminé par $\eta\nu = \nu_0\nu_1$, $|\nu_1| = cg-1$. Montrons que ϖ projette tr^3 dans tr^0, en étudiant d'abord les divers cas où (s', ν', λ) appartient à $\chi^3(s, \nu)$:

1) $s = (Y\eta, w, \alpha\gamma)$, $s' = (\eta Z, w', \alpha\gamma)$, $\nu = Z\nu'$, $\lambda = '\sigma'$, lorsque $s' \in d^3(s, Z)$ ou $(s', 1, '\sigma') \in k^3(s, \nu)$. Alors $\varpi(s, \nu) = [(Y, w, \alpha), \nu_0]$, $\eta\nu = \nu_0\nu_1$, $\varpi(s', \nu') = [(Z', w', \alpha), \nu'_0]$, si $\eta Z = Z'\eta'$, $\eta'\nu' = \nu'_0\nu'_1$, $|\nu_1| = |\nu'_1| = cg-1$: $Z'\nu'_0\nu'_1 = Z'\eta'\nu' = \eta Z\nu' = \eta\nu = \nu_0\nu_1$ d'où $Z'\nu'_0 = \nu_0$; $(Z', w, \alpha) \in d^0[(Y, w, \alpha), Z']$.

2) $s = (Y\eta, w, \alpha\gamma)$, $s' = (Y\eta, w', \alpha\gamma)$, $\nu = \nu'$, lorsque $(s', \lambda) \in m^3(s)$ ou $(s', 0, \lambda) \in k^3(s, \lambda)$. Alors $\varpi(s, \nu) = [(Y, w, \alpha), \nu_0]$, $\varpi(s', \nu) = [(Y, w', \alpha), \nu_0]$; $[(Y, w', \alpha), \lambda] \in m^0(Y, w, \alpha)$.

3) $s = (Y\eta, w, \alpha\gamma)$, $s' = s_f$, $\nu = \nu'$ lorsque $(s_f, \lambda) \in m^3(s)$ ou $(s_f, 0, \lambda) \in k^3(s, \nu)$. $\varpi(s, \nu) = [(Y, w, \alpha), \nu_0]$; $(s_f, \lambda) \in m^0(Y, w, \alpha)$ car $Y \prec \alpha$ ou $A \prec \alpha$ entraîne $\alpha \neq \sigma$.

D'autre part, avec des notations analogues à celles de 5. 5, si $\gamma^2 = \alpha\gamma$ ($\alpha \in T'$, $\gamma \in E^{cd-1}$) et $\nu\sigma^{cg} = b\eta\nu'$ ($b \in T'$, $\eta \in E^{cg-1}$), $\gamma^0 = '\alpha'$, $\rho^2 = \alpha\gamma b\eta$, $\nu^2 = \nu'$, $\rho^0 = '\alpha b'$, $\nu^0\sigma^{cg-1} = \eta\nu'$, $l^0(\rho^0) = (b, \sigma, \alpha)$, $l^2(\rho^2) = (b\eta, \sigma, \alpha\gamma)$, ϖ vérifie $\varpi[l^2(\rho^2), \nu^2] = [l^0(\rho^0), \nu^0]$.

On pourrait montrer de même que ϖ projette aussi tr^3 dans tr^1.

Envisageons maintenant l'application ϖ' : $\varpi'[(Y, w, \gamma'), \nu] = [(Y\eta, w, \gamma'), \nu_2]$ où $Y \in E$, $w \in J_1$, $\gamma' \in T'^{cd}$, $\nu \in E^*$, η et ν_2 sont déterminés par $\eta\nu_2 = \nu\sigma^{cg-1}$

et $|\eta| = cg - 1$; montrons que ϖ' projette tr^2 dans tr^3 ; étudions les cas où (s', ν', λ) appartient à $\chi^2(s, \nu)$:

a) $s = (Y, w, \alpha\gamma)$, $s' = (Z, w', \alpha\gamma)$, $\nu = Z\nu'$, $\lambda = '\sigma'$, lorsque $s' \in d^2(s, Z)$ ou $(s', 1, '\sigma') \in k^2(s, \nu)$. $\varpi'(s, \nu) = [(Y\eta, w, \alpha\gamma), \nu_2]$, $\eta\nu_2 = \nu\sigma^{cg-1}$, $|\eta| = cg - 1$, $\varpi'(s', \nu') = [(Z\eta', w', \alpha\gamma), \nu'_2]$, $\eta'\nu'_2 = \nu'\sigma^{cg-1}$, $|\eta'| = cg-1$: $Z\eta'\nu'_2 = Z\nu'\sigma^{cg-1} = \nu\sigma^{cg-1} = \eta\nu_2$; comme $|\eta| = |\eta'|$, il existe $Z' \in E$ tel que $Z\eta' = \eta Z'$, d'où $\nu_2 = Z'\nu'_2$. Lorsque $s' \in d^2(s, Z)$, $(\eta Z', w', \alpha\gamma) \in d^3[(Y\eta, w, \alpha\gamma), Z']$; il en est de même lorsque $(s', 1, '\sigma') \in k^2(s, \nu)$ et $(Y\eta, w, \alpha\gamma) \in$ Det', d'après la définition des ensembles $Q(Y, \alpha\gamma)$ et $Q(Y, w, \alpha\gamma)$; enfin, lorsque $(s', 1, '\sigma') \in k^2(s, \nu)$ et $(Y\eta, w, \alpha\gamma) \notin$ Det', $[(\eta Z', w', \alpha\gamma), 1, '\sigma']$ appartient à $k^3[(Y\eta, w, \alpha\gamma), \nu_2]$.

b) $s = (Y, w, \alpha\gamma)$, $s' = (Y, w', \alpha\gamma)$, $\nu = \nu'$, lorsque $(s', \lambda) \in m^2(s)$ ou $(s', 0, \lambda) \in k^2(s, \nu)$. $\varpi'(s, \nu) = [(Y\eta, w, \alpha\gamma), \nu_2]$, $\varpi'(s', \nu) = [(Y\eta, w', \alpha\gamma), \nu_2]$. Lorsque $(s', \lambda) \in m^2(s)$, $[(Y\eta, w', \alpha\gamma), \lambda] \in m^3[(Y\eta, w, \alpha\gamma)]$; il en est de même lorsque $(s', 0, \lambda) \in k^2(s, \nu)$ et $(Y\eta, w, \alpha\gamma) \in$ Det', d'après la définition des ensembles $Q(Y, w, A_i, \alpha\gamma)$ et $Q(A, \alpha\gamma)$; lorsque $(s', 0, \lambda) \in k^2(s, \nu)$ et $(Y\eta, w, \alpha\gamma) \notin$ Det', $[(Y\eta, w', \alpha\gamma), 0, \lambda] \in k^3(s, \nu_2)$.

c) $s = (Y, w, \alpha\gamma)$, $s' = s_f$, $\nu = \nu'$ lorsque $(s_f, \lambda) \in m^2(s)$ ou $(s_f, 0, \lambda) \in k^2(s, \nu_2)$; $\varpi'(s, \nu) = [(Y\eta, w, \alpha), \nu_2]$; $(s_f, \lambda) \in m^3(s)$ ou $(s_f, 0, \lambda) \in k^3(s, \nu_2)$. Avec des notations analogues à celles de 5. 5, $\gamma^2 = \gamma^3$; si $\nu\sigma = b\nu^2$, $\rho^2 = \gamma^2 b$; si $\nu\sigma^{cg} = b\nu^2\sigma^{cg-1} = b\eta\nu^3$ ($\eta \in E^{cg-1}$), $\rho^3 = \gamma^2 b\eta$; $l^2(\rho^2) = (b, \sigma, \gamma^2)$, $l^3(\rho^3) = (b\eta, \sigma, \gamma^2)$, $\nu^2\sigma^{cg-1} = \eta\nu^3$, d'où $\varpi'[l^2(\rho^2), \nu^2] = [l^3(\rho^3), \nu^3]$.

Pour que les transducteurs tr^3 soient déterministes, il faut et il suffit que, pour tout $s \in M$ et tout $\nu \in E^*$, $k(s, \nu)$ contienne au plus un élément ; il suffit pour cela que la grammaire vérifie les conditions CD 1, CD 2, CD 3 qui assurent le déterminisme de tr^2 (7. 7. 4) ; nous avons vu (théorème 7. 7) que ces conditions sont nécessaires lorsque la grammaire est réduite.

En résumé :

Théorème 7. 8. *Les transducteurs* tr^3 *associés à une grammaire pluriaxiomatique* $G = (N, T, ::=, N')$ *vérifiant les conditions* DA *et* PA *satisfont à* GP 2. *Ils définissent des générateurs de piles* gn^3 *vérifiant* GP 1 *et tels que, pour tout mot* α *sur* T, $gn^3(\alpha)$ *soit l'ensemble des piles conjointes aux pseudo-arborescences de* $\mathscr{C}(G)$ *dont* α *est la suite de feuilles. Pour qu'un transducteur* tr^3 *soit déterministe, il faut et il suffit que* G *vérifie* CD 1, CD 2, CD 3.

7. 8. 3. Condition pour que tr^3 soit un transducteur fini. D'après la définition de l'ensemble Det', pour qu'un transducteur tr^3 soit fini (il est alors déterministe), il faut et il suffit que soient réalisées les deux conditions :

(CB 1) si $Y \in V$, $\alpha \in T'$, $\gamma \in T'^{cd-1}$, $Y \prec \alpha$ et $Y \succ \alpha$, les langages $\widetilde{\mathscr{C}}(\alpha, \gamma)_Y$ et $\widetilde{\mathscr{C}}'(\sigma, \alpha\gamma)_Y$ ont une borne d'intersection ;

(CB2) si $w \in J$, $Y \in E$, $\gamma' \in T'^{cd}$, $(Y, w) \notin$ Det, pour les points w'_i $(i = 1, \ldots, p)$ de Γ (w) tels que $g(w_i) = \overline{A}_i \in \overline{N}$, les p + 1 langages $\widetilde{\mathscr{C}}(A_i, \gamma')_Y$ et $\widetilde{\mathscr{C}}'(w, \gamma')_Y$ ont une borne d'intersection.

Si $\widetilde{\mathscr{C}}(a, \gamma)_Y \neq \emptyset$, $Y \lessdot a$ (lemme 2) ; si $\mathscr{C}'(\sigma, a\gamma)_Y \neq \emptyset$, par définition de cet ensemble, il existe $A \in N$, dont Y est finale stricte, tel que $\mathscr{C}(A, a\gamma) \neq \emptyset$, donc $\mathscr{C}(A, a) \neq \emptyset$; A, a est un couple vicinal (7. 6. 2), $A \lessdot a$ ou $A \gtrdot a$, d'où $Y \gtrdot a$ (7. 4). D'autre part, $\widetilde{\mathscr{C}}(a, \gamma)$ est la réunion des $\widetilde{\mathscr{C}}(a, \gamma)_Y$, et $\mathscr{C}'(\sigma, a\gamma)$ la réunion des $\widetilde{\mathscr{C}}'(\sigma, a\gamma)_Y$ pour $Y \in V$. On en déduit que CB 1 équivaut à :
(CB' 1) si $a \in T'$, $\gamma \in T'^{cd-1}$, les langages $\widetilde{\mathscr{C}}(a, \gamma)$ et $\widetilde{\mathscr{C}}'(\sigma, a\gamma)$ ont une borne d'intersection.
De même, (CB 2) équivaut à :
(CB' 2) si $w \in J$, $\gamma' \in T'^{cd}$, s'il existe p (> 0) points w'_i de Γ (w) et p symboles non terminaux A_i tels que $g(w'_i) = \overline{A}_i$, les p+1 langages $\widetilde{\mathscr{C}}(A_i, \gamma')$ et $\widetilde{\mathscr{C}}(w, \gamma')$ ont une borne d'intersection.

Lorsque deux règles distinctes de la grammaire ont des seconds membres différents, p = 1 s'il existe.

Les grammaires qui satisfont à CB 1 et CB 2 sont les "bounded right context grammars" de [24]. Les grammaires "fully nested" introduites par [38] peuvent être transformées en des grammaires équivalentes ne possédant plus aucune règle A ::= A', où $A' \in N$, sans modifier les sous-propositions d'une proposition (7. 10. 4) ; alors, pour tout point w de J, Γ (w) contient au plus un élément et pour $a \in T'$, $\widetilde{\mathscr{C}}(a, \Lambda)$ et $\widetilde{\mathscr{C}}'(\sigma, \Lambda)$ ont une borne d'intersection inférieure au maximum de la longueur des seconds membres des règles. La méthode d'analyse de [47] revient à la construction d'une pile conjointe, dans le cas où les seconds membres des règles sont distincts, de longueur 2 au plus, et où il existe un transducteur tr^2 fini pour lequel cd = 1, cg = 2.

7. 8. 4. Exemple. N = {X, A, B, C, D} ; T = {a, b, c, d, f} ; N' = {X} ;

X ::= A b C | d A D | B   A ::= f A | a
B ::= f B b | a   C ::= c C | c
D ::= b D | b

Matrice des relations $\lessdot$ et $\gtrdot$:

	X	A	B	C	D	a	b	c	d	f	σ
X											$\gtrdot$
A					$\lessdot$		$\lessdot\ \gtrdot$				
B							$\lessdot$				$\gtrdot$
C											$\gtrdot$
D											$\gtrdot$
a							$\gtrdot$				$\gtrdot$
b				$\lessdot$	$\lessdot$		$\lessdot\ \gtrdot$	$\lessdot$			$\gtrdot$
c				$\lessdot$				$\lessdot$			$\gtrdot$
d		$\lessdot$				$\lessdot$				$\lessdot$	
f		$\lessdot$	$\lessdot$			$\lessdot$				$\lessdot$	
σ	$\lessdot$	$\lessdot$	$\lessdot$			$\lessdot$			$\lessdot$	$\lessdot$	

1) $A \lessdot b$ et $A \gtrdot b$: étudions $\mathscr{C}(b, \Lambda)$ et $\mathscr{C}'(\sigma, b)$, avec les notations et les méthodes de 7. 6. 2 et 7. 6. 3 : b est initiale de b et D ;

$<b> ::=_1 \ <X> A \mid <X> d A \mid <B> f B \mid <D> b, \quad <X> ::=_1 \ \sigma$:

les mots de $\mathscr{C}(b, \Lambda)$ se terminant par A sont σ A et σ d A. $\mathscr{C}'(\sigma, b)$ est la réunion des ensembles $\mathscr{C}(A', b)$ Ψ pour A' ::= Ψ ; les mots de $\mathscr{C}'(\sigma, b)$ se terminant par A sont ceux de $\mathscr{C}(A, b)$ f A. L'ensemble formé par A σ et A d, ou encore l'ensemble des mots de longueur 2 autres que A f, est un ensemble séparant $\widetilde{\mathscr{C}}(b, \Lambda)_A$ de $\widetilde{\mathscr{C}}'(\sigma, b)_A$.

2) $b \lessdot b$ et $b \gtrdot b$: les mots de $\mathscr{C}(b, \Lambda)$ se terminant par b sont ceux de $\mathscr{C}(D, \Lambda)$ b : $<D> ::=_1 \ <X> d A \mid <D> b$; les mots de $\mathscr{C}'(\sigma, b)$ se terminant par b sont ceux de $\mathscr{C}(B, b)$ f B b et ceux de $\mathscr{C}(D, b)$ b, mais ce dernier ensemble est vide car D, b n'est pas un couple vicinal (ni $D \lessdot b$, ni $D \gtrdot b$). On peut prendre l'ensemble des mots de longueur 2 autres que b B comme ensemble séparant $\widetilde{\mathscr{C}}(b, \Lambda)_f$ de $\widetilde{\mathscr{C}}'(\sigma, b)_f$.

3) Les seuls mots Ψ à la fois second membre et facteur droit de second membre d'une règle sont :
- 'b' : D ::= b et B ::= f B b, mais on n'a pas $B \lessdot D$.
- 'a' : A ::= a, B ::= a et $f \lessdot A$, $f \lessdot B$, $\sigma \lessdot A$, $\sigma \lessdot B$: Γ [o (a)] contient deux points d'images $\overline{A}$ et $\overline{B}$, mais aucun autre point ; (f, o (a)) et (σ, o (a)) n'appartiennent pas à Det, mais $\mathscr{C}'$ (o (a), Λ) = $\emptyset$.

Etudions $\mathscr{C}(A, \Lambda)$ et $\mathscr{C}(B, \Lambda)$:

$<A> ::=_1 \ \sigma \mid <X> d \mid <A> f, \quad <B> ::=_1 \ \sigma \mid <B> f$;

les mots de $\mathscr{C}(A, \Lambda)$ sont les σf^n et $\sigma d f^n$, ceux de $\mathscr{C}(B, \Lambda)$ les σf^n $(n \geqslant 0)$. $\widetilde{\mathscr{C}}(A, \Lambda)_f$ et $\widetilde{\mathscr{C}}(B, \Lambda)_f$ ne sont pas disjoints, ni $\widetilde{\mathscr{C}}(A, \Lambda)_\sigma$ et $\widetilde{\mathscr{C}}(B, \Lambda)_\sigma$. b est le seul symbole terminal tel que $\mathscr{C}(A, b) \neq \emptyset$ et aussi le seul tel que $\mathscr{C}(B, b) \neq \emptyset$; de plus $\mathscr{C}(B, \sigma)$ = {'σ'}, mais $\mathscr{C}(A, \sigma) = \emptyset$:

$(A, b) ::=_1 \ \sigma \mid <X> d \mid (A, b) f, \quad (B, b) ::=_1 \ <B> f.$

$\widetilde{\mathscr{C}}(B, b)_\sigma$ est vide, mais $\widetilde{\mathscr{C}}(A, b)_f$ et $\widetilde{\mathscr{C}}(B, b)_f$ ne sont pas disjoints, car ils contiennent les mots σf^n $(n \geqslant 1)$. Etudions les $\mathscr{C}(A, bx)$ et $\mathscr{C}(B, bx)$ où $x \in T'$; d'après 7. 6. 1, $\mathscr{C}(B, bx) = \mathscr{C}(B, x)$ f, donc est vide sauf pour x = b ou x = σ : $\mathscr{C}(B, bb)$ est l'ensemble des σf^n $(n \geqslant 2)$, $\mathscr{C}(B, b\sigma)$ = {'σ f'}. $\mathscr{C}(A, bb) = \mathscr{C}(X, \Lambda)$ d $\cup$ $\mathscr{C}(A, bb)$ f : il s'agit de l'ensemble des $\sigma d f^n$ $(n \geqslant 0)$. $\mathscr{C}(A, b\sigma) = \mathscr{C}(X, \sigma)$ d $\cup$ $\mathscr{C}(A, b\sigma)$f : il s'agit du même ensemble.

On choisit cd = 2, on ordonne A, B en plaçant B avant A et on définit : Q (σ, o (a), B, γ') = $\emptyset$ si $\gamma' \neq \sigma\sigma$; Q (σ, o (a), B, $\sigma\sigma$) = {Λ} ; Q (σ, o (a), A, γ') = {Λ} pour tout $\gamma' \in E^2$; Q (f, o (a), B, γ') = $\emptyset$ si γ' diffère de $\sigma\sigma$, b σ et bb ; Q (f, o (a), B, $\sigma\sigma$) = {Λ} ; Q (f, o (a), B, b σ) = {f σ} ; Q (f, o (a), A, γ') = {Λ} pour tout γ'. $\widetilde{\mathscr{C}}(A, bb)_f$ et $\widetilde{\mathscr{C}}(B, bb)_f$ sont disjoints, mais n'ont pas de borne d'intersection. On obtient un transducteur pour lequel cg = 2, qui est déterministe mais n'est pas un transducteur fini. Il est d'ailleurs facile de voir que quel que soit cd, $\widetilde{\mathscr{C}}(A, b^{cd})_f$ et $\widetilde{\mathscr{C}}(B, b^{cd})_f$ n'ont pas de borne d'intersection.

7. 9. Générateurs contrôlés par des couples de symboles terminaux.

7. 9. 1. Dans les paragraphes précédents, on a limité l'indéterminisme qui rendait peu utilisable le générateur gn^0 en contrôlant les entrées et les sorties des piles par les relations $<, >$ et par les contextes ; le pseudo-graphe de production participe à ce contrôle et joue aussi un rôle, théoriquement plus essentiel, de vérification, assurant la condition C 3. Nous chercherons ici des générateurs plus simples à mettre en œuvre, où le pseudo-graphe de production ne conservera que ce rôle de vérification et où les entrées et les sorties seront contrôlées par des couples d'éléments de T' : on ne retiendra du contexte, à droite qu'un élément (cd = 1), à gauche que le dernier symbole terminal et le nombre de symboles auxiliaires qui le suivent.

7. 9. 2. Définition des ensembles Δ_1 (b, c) et Δ_2 (b, c). Soit U une pile de l'ensemble $\mathcal{Y}(G)$. Pour chacun de ses états u_i, nous nommerons ω_i le plus grand facteur droit de u_i appartenant à N^*, b_i l'élément de $T' = T \cup \{\sigma\}$ tel que $b_i\, \omega_i$ soit facteur droit de $\sigma\, u_i$, n_i le nombre de sorties supérieures à i, inférieures ou égales au plus petit entier suivant i qui est une entrée, une sortie de symbole terminal, ou l'indice du dernier état de la pile ; nous utiliserons notamment les couples $(|b_i\, \omega_i|, n_i)$.

Associons à tout couple (b, c) d'éléments de T' deux ensembles Δ_1 (b, c) et Δ_2 (b, c) de couples d'entiers (n", n') vérifiant respectivement $0 < n' \leqslant n''$ et $0 \leqslant n' \leqslant n''$, de manière que :
(C 6) si i est une entrée dans U telle que i + 1 soit une sortie autre que la dernière et si c est le premier signe de $\alpha^U_{i-}\, \sigma$, $(|b_i\, \omega_i|, n_i) \in \Delta_1\, (b_i, c)$;
(C 7) si i est une sortie de $c \in T$, $(|b_i\, \omega_i|, n_i) \in \Delta_2\, (b_i, c)$.
Nous imposerons de plus que, pour tout $b \in T'$, $\Delta_2\, (b, \sigma)$ soit vide. Soit alors i un entier pour lequel i + 1 est une entrée d'un symbole terminal c. Il existe une sortie i' de c telle que $u_i = u_{i'}$, (théorème 1. 2). En appliquant C 7 à i' :
(C 8) si i + 1 est une entrée d'un symbole terminal c, il existe dans $\Delta_2\, (b_i, c)$ un couple de premier élément $|b_i\, \omega_i|$:

Etudions la détermination des ensembles Δ_1 (b, c) et Δ_2 (b, c), à l'aide de C 3 et C 4. Soit un état u_i de U, i' la première entrée supérieure à i. Pour n_i, deux cas se présentent :
a) une sortie de symbole terminal est strictement comprise entre i et i' : $n_i = |b_i\, \omega_i|$;
b) aucune sortie de symbole terminal n'est strictement comprise entre i et i' : $n_i < |b_i\, \omega_i|$.
Δ_1(b, c)(resp. Δ_2(b, c)) est réunion des deux ensembles Δ'_1 (b, c) et Δ''_1(b, c) (resp. Δ'_2 (b, c) et Δ''_2 (b, c)) formés des couples(n", n') vérifiant n" = n' d'une part, n" < n' d'autre part.

Si i est une entrée pour laquelle i + 1 est une sortie autre que la dernière et si c est le premier élément de $\alpha^U_{i-}\sigma = \alpha^U_{i'-}\sigma$, dans le cas a il existe $A \in N$, $\Psi \in V^*$ tels que $A ::= \Psi\, b_i\, \omega_i$ et $\mathscr{C}\, (A, c) \neq \emptyset$; dans le cas b, il existe $A \in N$, $\omega' \in N^*$, $\omega'' \in N^*$ tels que $A ::= \omega'$, $\omega_i = \omega''\, \omega'$, et un mot de $\mathscr{C}$ (A, c) dont $b_i\, \omega''$ est facteur droit. Pour qu'un ensemble de couples d'entiers positifs (n', n') soit un Δ'_1 (b, c), il suffit qu'il contienne l'ensemble ${}^0\Delta'_1$ (b, c) des couples $(|\omega| + 1, |\omega| + 1)$ associés aux mots ω sur N tels qu'il existe $A \in N$, $\Psi \in V^*$ vérifiant $A ::= \Psi\, b\, \omega$ et $\mathscr{C}\, (A, c) \neq \emptyset$. Pour qu'un ensemble de couples d'entiers (n", n') tels que $0 < n' < n''$ soit un Δ''_1 (b, c), il suffit qu'il contienne l'ensemble ${}^0\Delta''_1$ (b, c) des couples $(|\omega'| + |\omega''| + 1, |\omega'|)$ associés aux mots ω', ω'' sur N tels qu'il existe $A \in N$ pour lequel $A ::= \omega'$ et $b\, \omega''$ est facteur droit d'un mot de $\mathscr{C}$ (A, c).

Si i est une sortie de $c \in T$, dans le cas a il existe $A \in N$, $\Psi \in V^*$, $\Psi' \in V^*$ tels que $A ::= \Psi\, b_i\, \omega_i\, c\, \Psi'$; dans le cas b, il existe $A \in N$, $\Psi \in V^*$, $\omega' \in N^*$, $\omega'' \in N^*$ tels que $A ::= \omega'\, c\, \Psi$, $\omega_i = \omega''\, \omega'$, $b_i\, \omega''$ soit facteur droit d'un mot de $\mathscr{C}$ (A, Λ).Pour qu'un ensemble de couples d'entiers positifs (n', n') soit un Δ'_2 (b, c), il suffit qu'il contienne l'ensemble ${}^0\Delta'_2$ (b, c) des couples $(|\omega| + 1, |\omega| + 1)$ associés aux mots ω sur N tels qu'il existe $A \in N$, $\Psi \in V^*$, $\Psi' \in V^*$ vérifiant $A ::= \Psi\, b\, \omega\, c\, \Psi'$. Pour qu'un ensemble de couples d'entiers (n", n') vérifiant $0 \leqslant n' < n''$ soit un Δ''_2 (b, c), il suffit qu'il contienne l'ensemble ${}^0\Delta''_2$ (b, c) des couples $(|\omega'| + |\omega''| + 1, |\omega'|)$ associés aux mots ω', ω'' sur N tels qu'il existe $A \in N$, $\Psi \in V^*$ pour lesquels $A ::= \omega'\, c\, \Psi$ et $b\, \omega''$ est facteur droit d'un mot de $\mathscr{C}$ (A, Λ).

On peut montrer, par exemple à l'aide du lemme 1, que réciproquement, lorsque la grammaire est réduite, les ensembles Δ'_1 (b,c), Δ''_1 (b, c), Δ'_2 (b, c) , Δ''_2 (b, c) doivent contenir respectivement ${}^0\Delta'_1$ (b, c), ${}^0\Delta''_1$ (b, c), ${}^0\Delta'_2$ (b, c) , ${}^0\Delta''_2$ (b, c).

Exemple : N = { K, B, C, D, F } ; T = { d, f, g, h } ; N' = { K } ;
K ::= f F B   C ::= h | h C   F ::= g | F g
B ::= C D   D ::= d F
F ::= g ; F, h est un couple vicinal ; donc (1, 1) $\in$ ${}^0\Delta'_1$ (g, h).
B ::= C D ; σ f F $\in$ $\mathscr{C}$ (B, σ) ; donc (4, 2) $\in$ ${}^0\Delta''_1$(f, σ).
D ::= d F ; σ f F C $\in$ $\mathscr{C}$ (D, Λ) ; donc (3, 0) $\in$ ${}^0\Delta''_2$(f, d).
Pour tout couple b, c, ${}^0\Delta'_2$ (b, c) est vide.

La génération des piles cherchées, qui va être étudiée, est résumée par l'organigramme de la figure 7. 6. Nous définirons un générateur de piles gn^4 dans le cas où, pour tout couple b, c, les ensembles Δ_1 (b, c) et Δ_2 (b, c) sont finis. D'après C 6 et C 7, pour qu'il existe de tels ensembles, il faut et suffit que soit vérifiée la condition :
(C T) *les facteurs droits des mots des langages* $\mathscr{C}$ (A, Λ) (A $\in$ N), *qui appartiennent à* N^* *ont une longueur bornée.*

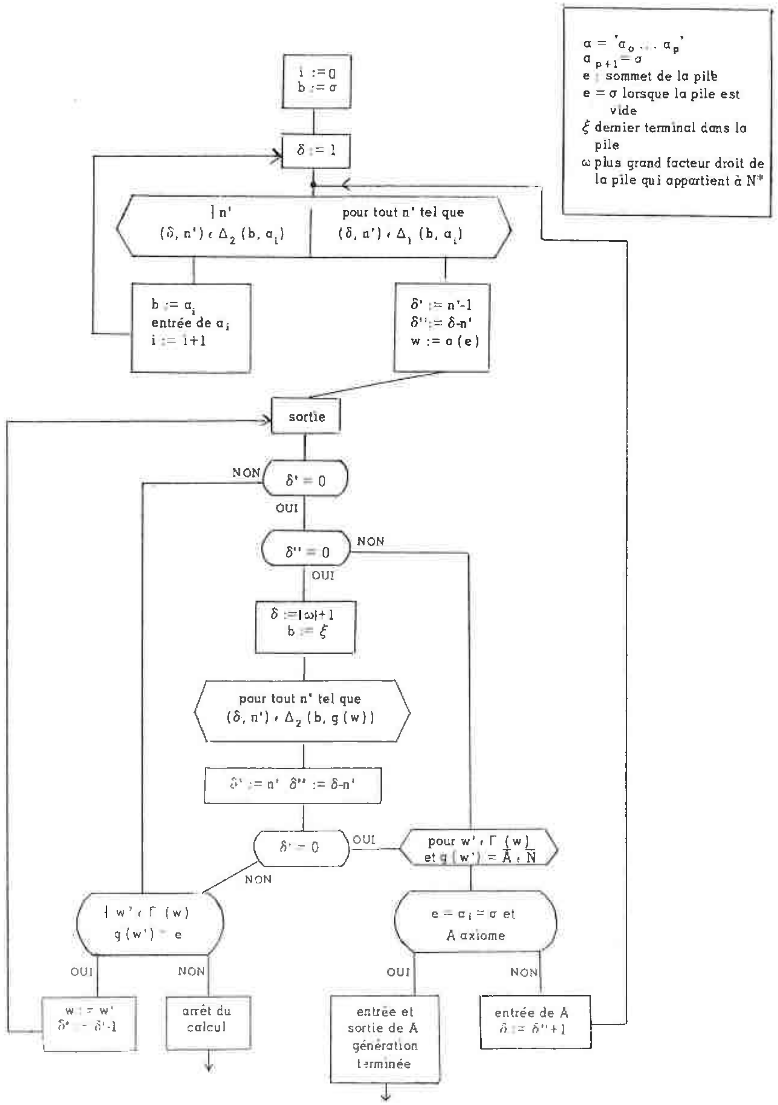


Figure 7. 6

C T signifie que, pour tout A ϵ N, les deux langages de Kleene $\mathcal{C}$ (A, Λ) et N*\{Λ} ont une borne d'intersection : C T est décidable. Nous désignerons par cg le plus grand des premiers entiers des couples appartenant aux différents ensembles Δ_1 (b, c) et Δ_2 (b, c) , par [cg] l'ensemble des entiers compris au sens large entre 0 et cg.

7. 9. 3. Générateur gn ^{4}. Il est défini par W = V, cd = 1 et un transducteur fini tr ^{4} :

a) M = $E^{cg} \times J_1 \times T' \times [cg] \times [cg]$. Dans les définitions qui suivent, a, b et c décrivent T', B décrit N, η décrit E^{cg-1}, Y et Z décrivent E, w décrit J, δ' et δ'' décrivent [cg]. Lorsque η' est un mot de $E^{cg} \setminus N^*$ et ω son plus long facteur droit appartenant à N*, on pose δ (η') = |ω| + 1 et on nomme ξ (η') l'élément de T' tel que ξ (η') ω soit facteur droit de η'.
b) q = cg + 1 ; l (ab η) = (b η, σ, a, 0, 0).
c) Si η Z ∉ N* et (1, 1) ∈ Δ_1 (b, a), d [(b η, σ, a, 0, 0), Z] est formé des (η Z, o (b), a, n', δ (η Z) - n') pour tout entier n' tel que (δ (η Z), n') ∈ Δ_2 (ξ(η Z), b) ;
si δ' ≠ 0, (δ', δ'') ≠ (1, 0) et si un point w' de Γ (w) vérifie g (w') = B, d [(B η, w, a, δ', δ''), Z] = {(η Z, w', a, δ' - 1, δ'')} ;
si η Z ∉ N* et si un point w' de Γ (w) vérifie g (w') = c, d[(c η, w, a, 1, 0),Z] est formé des (η Z, w', a, n', δ (η Z) - n') pour tout entier n' tel que (δ (η Z), n') ∈ Δ_2 (ξ(η Z), c).
d) Si (1, 0) ou (1, 1) appartient à Δ_2 (b, a) (d'où a ≠ σ), m (b η, σ, a, 0, 0) = {(s_f, 'a')} ;
m (Y η, w, a, 0, δ'') contient [(Y η, w', a, 0, δ''), 'A'] pour tout point w' de J'' tel que w Γ w' et g (w') = $\overline{A}$, et pas d'autre élément ;
si w ∈ J'', g (w) = $\overline{A}$ et Y η ∉ N*, m (Y η, w, a, 0, δ'') contient les seuls éléments :
- (s_f, 'a') lorsque dans Δ_2 (ξ(Y η), a) se trouve un couple de premier entier δ'' + 1 (d'où a ≠ σ) ou lorsque Y = a = σ et A ∈ N',
- [(Y η, o (A), a, n' - 1, δ'' + 1 - n'), 'σ'] pour tout n' tel que (δ'' + 1, n') ∈ Δ_1 (ξ(Y η), a) sauf lorsque Y = a = σ et A ∈ N'.

On utilisera aussi le transducteur tr $^{4'}$ qui ne diffère de tr ^{4} que par la suppression des mots soulignés en pointillé à la fin de la définition, et le générateur gn $^{4'}$ correspondant : l'identité projette tr $^{4'}$ dans tr ^{4}, gn $^{4'}$ est contenu dans gn ^{4}.

Soit ϖ' la projection de M sur $E^{cg} \times J_1 \times T'$: ϖ' (η', w_1, a, δ', δ'') = (η', w_1, a). ϖ étant l'application définie en 7. 8. 2, qui projette le transducteur tr ^{3} associé à cd = 1 dans tr ^{0}, une démonstration analogue à celle de 7. 8. 2 montre que ϖ o ϖ' projette tr ^{4} dans tr ^{0}. tr ^{4} vérifie GP 2 dès que AG est satisfaite, gn ^{4} vérifie GP 1 et engendre des piles conjointes à des pseudo-arborescences de $\mathcal{C}$ (G).

Il reste à montrer que gn^4 engendre toutes ces piles. Soit U l'une d'elles et $'a_0 \dots a_{p-1}'$ la suite de feuilles de la pseudo-arborescence à laquelle elle est conjointe. Les entrées des a_i, avec 0 et le nombre d'états de U, forment un p-partage de U. Nous utiliserons les notations de 5. 2 ; en particulier $a_p = \sigma$. A tout état u_k de U, nous associerons un triplet (s_k, y_k, Z_k), $s_k = (\eta_k, w_k, c_k, \delta'_k, \delta''_k) \in M$ ou $s_k = s_f$, y_k est un mot sur E de longueur 0 ou 1 ; $Z_k \in E$; puis nous montrerons que les triplets associés aux états u_k tels que $j_i < k \leqslant j_{i+1}$ forment un calcul complet du transducteur tr^4, d'entrée $a_i\, u_{j_i}\, \sigma^{cg}$, de résultat μ_i. Dans tous les cas, c_k est le premier signe de $\alpha_k^U \sigma$.

a) Si k est une sortie autre que la dernière, η_k est le facteur gauche de $\tilde{u}_k \sigma^{cg}$ ayant cg pour longueur ; si k' est la dernière entrée inférieure à k, les entiers k' + 1, ..., k sont des sorties de $z_1, \dots z_{k-k'}$, donc $z_1 \dots z_{k-k'}$ est un chemin du pseudo-graphe de production, d'origine $o(z_1)$: w_k est son extrémité ; $\delta'_k = n_k$ (défini au début de 7. 9. 2) ; $\delta''_k = \delta(\eta_k) - n_k$; $Z_k = \sigma$; si k - 1 est une entrée de symbole non terminal, $y_k = \Lambda$; sinon y_k est le dernier élément de η_k.

b) Si k est une entrée de $A \in N$, k - 1 est une sortie ; $\eta_k = \eta_{k-1}$; si k' est la dernière entrée inférieure à k, les entiers k + 1, ... k - 1 sont des sorties de $z_1, \dots, z_{k-k'-1}$, donc $z_1 \dots z_{k-k'-1} \bar{A}$ est un chemin du pseudo-graphe de production d'origine $o(z_1)$: w_k est son extrémité ; $\delta'_k = 0$; $\delta''_k = \delta(\eta_k)$; $y_k = \Lambda$; $Z_k = A$.

c) Si k est une entrée de $a \in T$, $s_k = s_f$, $y_k = \Lambda$, $Z_k = a$.

d) Si k est la dernière sortie, $s_k = s_f$, $y_k = \Lambda$, $Z_k = \sigma$.

Si les triplets associés aux u_k tels que $j_i < k \leqslant j_{i+1}$ forment un calcul complet, le résultat de ce calcul est μ_i. Pour tous ces k, $c_k = a_i$. Posons $a_i\, \tilde{u}_{j_i}\, \sigma^{cg} = 'a_i x_1 \dots x_r' = a_i x_1 \eta x_{cg+1} \dots x_r$; $x_1 = a_{i-1}$ si $i \neq 0$, $x_1 = \sigma$ si $i = 0$. $l(a_i x_1 \eta) = (x_1 \eta, \sigma, a_i, 0, 0)$. Si $j_{i+1} = j_i + 1$, $j_i + 1$ est une entrée de a_i, (1, 0) ou (1, 1) appartient à $\Delta_2(x_1, a_i)$ (C 8) : $(s_f, 'a_i') \in m(x_1 \eta, \sigma, a_i, 0, 0)$.
Si $j_{i+1} \neq j_i + 1$, $k = j_i + 1$ est une sortie de $x_1 \in T$; $\eta_k = \eta x_{cg+1}$, $w_k = o(x_1)$; $d[(x_1 \eta, \sigma, a_i, 0, 0), x_{cg+1}]$ contient $(\eta_k, w_k, a_i, n_k, \delta(\eta_k) - n_k)$ car $(\delta(\eta_k), n_k) \in \Delta_2(x_1, a_i)$ (C 7). Supposons maintenant $k > j_i + 1$.

a) k et k - 1 sont des sorties : $w_{k-1} \Gamma w_k$ et k est une sortie de $g(w_k)$; $\delta'_{k-1} = n_{k-1} \neq 0$. Si k est une sortie de symbole auxiliaire, $\delta'_{k-1} = n_{k-1} > 1$ ou $\delta''_{k-1} = \delta(\eta_{k-1}) - n_{k-1} \neq 0$, $n_k = n_{k-1} - 1$, $\delta(\eta_k) = \delta(\eta_{k-1}) - 1$: $d(s_{k-1}, y_k)$ contient s_k. Si k est une sortie d'un symbole terminal c, $n_{k-1} = \delta(\eta_{k-1}) = 1$ et $(\delta(\eta_k), n_k)$ appartient à $\Delta_2(\xi(\eta_k), c)$ (C 7) : $d(s_{k-1}, y_k)$ contient s_k.

b) $k \neq j_{i+1}$ est une sortie et k - 1 une entrée de $A \in N$: si $a_i = \sigma$ et $u_k = \Lambda$, k est suivi d'une entrée, d'une sortie, d'une entrée, ... jusqu'à j_{i+1}, dernier indice ; A dérive de la racine, donc, d'après DA et PA, n'est pas un axiome ; ce cas ne se présente jamais si aucune règle n'a son second membre réduit à un seul symbole auxiliaire ; dans tous les cas, $\eta_k = \eta_{k-2} = \eta_{k-1}$, $g(w_{k-1}) = \bar{A}$, $w_k = o(A)$, $\delta'_{k-1} = 0$, $\delta'_k = n_k = n_{k-1} - 1$, $\delta''_{k-1} = \delta(\eta_{k-1})$, $(\delta''_{k-1} + 1, n_{k-1}) \in \Delta_1(\xi(\eta_k), a_i)$ (C 6), $\delta''_k = \delta(\eta_k) - n_k = \delta''_{k-1} - n_{k-1} + 1$: $m(s_{k-1})$ contient $(s_k, '\sigma')$, $y_k = \Lambda$, $Z_k = \sigma$.

c) k est une entrée de $A \in N$: k - 1 est une sortie (C 2) ; $\eta_k = \eta_{k-1}$; $w_{k-1} \Gamma w_k$ et $g(w_k) = \bar{A}$; $\delta'_k = \delta'_{k-1} = 0$; $\delta''_k = \delta''_{k-1}$: $m(s_{k-1})$ contient $(s_k, 'A')$, $y_k = \Lambda$, $Z_k = A$.

d) $k = j_{i+1}$ est entrée de a_i ; k - 1 est une entrée (C3) d'un symbole auxiliaire A et k - 2 une sortie ; $g(w_{k-1}) = \bar{A}$; $\delta'_{k-1} = 0$; d'après C8, $\Delta_2(\xi(\eta_{k-1}), a_i)$ contient un couple de premier élément $\delta(\eta_{k-2}) + 1 = \delta(\eta_{k-1}) + 1 = \delta''_{k-1} + 1$; $s_k = s_f$, $y_k = \Lambda$, $Z_k = a_i$, $m(s_{k-1})$ contient $(s_f, 'a_i')$.

e) $k = j_{i+1}$ est sortie d'un axiome A : $a_i = \sigma$, $u_k = u_{k-2} = \Lambda$; k - 1 est entrée de A, $g(w_{k-1}) = \bar{A}$, $\delta'_{k-1} = 0$; $s_k = s_f$, $y_k = \Lambda$, $Z_k = \sigma$, $m(s_{k-1})$ contient $(s_f, '\sigma')$.

Théorème 7. 9. *Les transducteurs* tr^4 *associés à une grammaire pluriaxiomatique* G = (N, T, ::=, N') *vérifiant les conditions* DA, PA, CT *satisfont à* GP 2. *Ils définissent des générateurs de piles* gn^4 *vérifiant* GP 1 *et tels que, pour tout mot* a *sur* T, $gn^4(a)$ *soit l'ensemble des piles conjointes aux pseudo-arborescences de* $\mathcal{C}$ (G) *dont* a *est la suite de feuilles.*

Sous les mêmes hypothèses, $tr^{4'}$ vérifie GP 2 et $gn^{4'}$ vérifie GP 1 ; l'étude précédente montre que $gn^{4'}(a) = gn^4(a)$ si de plus aucune règle de G n'a son second membre réduit à un symbole auxiliaire.

7. 9. 4. Déterminisme de tr4. On vérifie facilement que, pour que tr4 ou tr4' soit déterministe, il faut et il suffit que la grammaire vérifie :

(CT 1) *Pour tout couple* (b, c) *d'éléments de* T', $\Delta_1(b, c)$ *et* $\Delta_2(b, c)$ *sont disjoints et, pour tout entier* n'', *leur réunion contient au plus un couple de premier entier* n'' ;

(CT 2) *la relation de production ne possède pas deux règles de même second membre.*

La condition CT 1 est fort restrictive : nous allons voir que si la grammaire est réduite, elle ne peut être satisfaite dès que, pour trois symboles non terminaux A, A', B et deux mots φ et Ψ sur V, A ::= B, A' ::= $\varphi A \Psi$, $|\varphi| + |\Psi| \neq 0$. La raison en est que le remplacement de B par A dans la pile laisse subsister le dernier symbole terminal présent dans la pile et le premier à entrer. Plus précisément :

1) Si $\Psi = \Lambda$, puisque la grammaire est réduite, il existe un axiome X et une X-dérivation droite d'un mot de T* qui contient des mots consécutifs $\eta A' \theta$, $\eta \varphi A \theta$, $\eta \varphi B \theta$, où $\theta \in T^*$: d'après la démonstration du lemme 1, il existe une entrée i dans une pile U de $\mathcal{J}$ (G) telle que $u_i = \eta \varphi B$, $u_{i+1} = \eta \varphi$, $u_{i+2} = \eta \varphi A$, $u_{i+3} = \eta \varphi$: i + 3 est une sortie autre que la dernière, $n_i = 1$, $n_{i+2} = 1 + |\varphi| > 1$, $|\omega_i| = |\omega_{i+2}|$, $b_i = b_{i+2}$: $(|b_i \omega_i|, n_i)$ et $(|b_i \omega_i|, n_{i+2})$ appartiennent à $\Delta_1(b_i, c)$ où c est le premier signe de $\alpha_i^U \sigma$.

2) Si $\Psi \neq \Lambda$, son premier élément Z a une initiale $c \in T$; posons $\Psi = Z \Psi'$; il existe un axiome X et une X-dérivation droite d'un mot de T^* qui contient $\eta A' \theta$, $\eta \varphi A Z \Psi' \theta$, $\eta \varphi A c \theta' \theta$, $\eta \varphi B c \theta' \theta$, où $\theta \in T^*$, $\theta' \in T^*$; d'où une pile U de $\mathcal{J}(G)$, une entrée i telle que $u_i = \eta \varphi B$, i + 1 soit une sortie et c soit le premier élément de α^U_{i-1}, et d'autre part (que Z = c ou $Z \neq c$) une sortie i' de c telle que $u_{i'} = \eta \varphi A$: $|\omega_i| = |\omega_{i'}|$, $b_i = b_{i'}$, $(|b_i \omega_i|, n_i)$ appartient à $\Delta_1 (b_{i'}, c)$ et $(|b_i \omega_i|, n_{i'})$ à $\Delta_2 (b_{i'}, c)$.

7. 10. Transformations de la grammaire.

7. 10. 1. Nous allons transformer la grammaire $G = (N, T, ::=, N')$ donnée en une grammaire qui ne possède pas de règle dont le second membre se réduit à un symbole non terminal (elle vérifie donc DA et PA), qui engendre le même langage que G et où toute phrase a les mêmes pseudo-arborescences des sous-propositions (4. 1. 2). L'analyse relative à la seconde grammaire (qui peut être effectuée grâce à $gn^{4'}$) ne résout pas complètement le problème de l'analyse relatif à la première. Mais elle résout le problème de la reconnaissance et permet de trouver les pseudo-arborescences des sous-propositions d'un mot donné ; si la grammaire transformée n'est pas ambiguë, G peut l'être, mais elle n'est pas fortement ambiguë. Nous proposerons ici deux transformations de G ; la première ne modifie que la relation de production ; elle est très proche de celle indiquée par [2] ; la seconde modifie relation de production et ensemble des axiomes ; toutes deux remplacent une pseudo-arborescence (I, f) par une pseudo- arborescence, image d'une arborescence sous-réduite de I (1. 5. 3). $A \xrightarrow{*} B$ signifie que B dérive de A pour la grammaire G.

7. 10. 2. Première transformation. Définissons une nouvelle relation de production : pour $A \in N$, $\Psi \in V^*$,
$A ::=_1 \Psi \iff (|\Psi| \geq 2$ ou $\Psi \in T^*)$ et $(\exists B \in N)$ $(A \xrightarrow{*} B$ et $B ::= \Psi)$.
Montrons que la grammaire G_1 $(N, T, ::=_1, N')$ possède les propriétés exigées.

Soit $(I_1, f_1) \in \mathcal{C}(G_1)$: I_1 est une ramification orientée (F_1, γ_1). Envisageons un nœud y de I_1 ; f_1 le transforme en A et transcrit en Ψ la famille dont il est prédécesseur : $A ::=_1 \Psi$, donc il existe $p \geq 0$ symboles non terminaux A_i tels que $A_{i-1} ::= A_i$ pour $1 \leq i \leq p$ et $A_p ::= \Psi$ (en posant $A_0 = A$) ; associons à y un ensemble F_y de p éléments x_i. On choisit les ensembles F_y disjoints, et disjoints de F_1 ; soit F la réunion de F_1 et des F_y. Définissons une ramification orientée (F, γ) de même racine et de mêmes feuilles que I_1 en posant pour chaque nœud y : $\gamma(x_{i-1}) = {}'x_i{}'$ pour $1 \leq i \leq p$ et $\gamma(x_p) = \gamma_1(y)$ $(x_0 = y)$. Définissons d'autre part une application f de F dans V : pour chaque nœud y, $f(x_i) = A_i$ pour $1 \leq i \leq p$; si $z \in F_1$, $f(z) = f_1(z)$. L'image de (F, γ) par f est une pseudo-arborescence qui a même racine et mêmes feuilles que (I_1, f_1)

et qui appartient à $\mathcal{C}(G)$. (I_1, f_1) en est une sous-pseudo-arborescence, les points de I_1 sont la racine de (F, γ), ses feuilles et ceux de ses nœuds qui ne sont pas seuls dans leur famille.

Réciproquement, soit $(I, f) \in \mathcal{C}(G)$; la racine de I, ses feuilles et ceux de ses nœuds qui ne sont pas seuls dans leur famille forment une sous-arborescence I' de I, qui est une arborescence sous-réduite de I. La restriction f' de f à I' transforme I' en une pseudo-arborescence qui a même racine et même suite de feuilles que (I, f). Soit Ψ une de ses familles, de prédécesseur A : $A = f(y)$, $y \in I'$; I' est transformée de I par un quasi-isomorphisme Θ : d'après 1. 5. 2, l'ensemble $\Theta^{-1}(y)$ des points de I transformés par Θ en y possède un plus grand élément z (pour l'ordre associé à I) et $A \xrightarrow{*} f(z)$; dans I, z est le prédécesseur d'une famille β dont les points appartiennent à I' ; dans I' le prédécesseur de β est y ; f transcrit donc β en Ψ : $f(z) ::= \Psi$, d'où $A ::=_1 \Psi$. $((I', f')) \in \mathcal{C}(G_1)$.

L'application qui à (I, f) associe ((I', f')) est une surjection de $\mathcal{C}(G)$ dans $\mathcal{C}(G_1)$; I' est sa propre ramification sous-réduite. Par suite, *les grammaires G et G_1 engendrent le même langage et chaque phrase y possède les mêmes pseudo-arborescences des sous-propositions.*

7. 10. 3. Détermination des ensembles Δ_1 (b, c) et Δ_2 (b, c) relatifs à G_1. Nous les déduirons des ensembles analogues relatifs à la grammaire G donnée, grâce à l'étude précédente dont nous conservons les notations. D'après le théorème 2. 9, dans les piles $\bar{U}$ et $\bar{U}'$ conjointes à I et I', les états de même sommet y sont égaux. Par suite :
1) Si, pour $\bar{U}'$, i' est une entrée et i' + 1 une sortie autre que la dernière, il existe une entrée i dans $\bar{U}$ telle que $\bar{u}_i = \bar{u}'_{i'}$, i + 1 soit une sortie (sinon i + 1 serait l'entrée d'un élément z et les états de sommet z dans les deux piles seraient différents), autre que la dernière (car I et I' ont même racine) ; i + 2 est alors une sortie, ou le sommet de $\bar{u}_i$ une feuille, puisque ce sommet est un point de I'.
2) Réciproquement, si, pour $\bar{U}$, i est une entrée, i + 1 une sortie autre que la dernière, i + 2 est une sortie ou le sommet de $\bar{u}_i$ une feuille, alors ce sommet est un point de I', il existe une entrée i' dans $\bar{U}'$ telle que $\bar{u}_i = \bar{u}'_{i'}$, et i' + 1 soit une sortie.
3) Si i' est la sortie d'une feuille de $\bar{U}'$, il existe une sortie i de $\bar{U}$ telle que $\bar{u}_{i-1} = \bar{u}'_{i'-1}$.
4) Si i est la sortie d'une feuille de $\bar{U}$, il existe une sortie i' de $\bar{U}'$ telle que $\bar{u}_{i-1} = \bar{u}'_{i'-1}$.

D'autre part, puisque I' est une sous-arborescence orientée de I, la première feuille entrant après un point de I, si elle existe, est la même pour $\bar{U}$ et $\bar{U}'$; si elle n'existe pas pour l'une des piles, elle n'existe pas pour l'autre.

Enfin, d'après le théorème 2. 9, le nombre de sorties qui séparent la sortie i d'un point de I' du premier entier supérieur à i qui est une entrée, la sortie d'une feuille ou l'indice du dernier état de la pile est le même pour $\overline{U}$ et $\overline{U}'$.

Les piles conjointes à (I, f) et ((I', f')) sont transcrites de $\overline{U}$ et $\overline{U}'$ par f. *Les ensembles* Δ'_1 (b, c), Δ'_2 (b, c), Δ''_2 (b, c), d'après leurs définitions et l'étude précédente, *sont les mêmes pour les grammaires* G *et* G_1 ; *on obtient les ensembles* Δ''_1 (b, c) *relatifs à* G_1 (nous les noterons Δ'''_1(b, c) pour éviter toute confusion) *en enlevant de ceux relatifs à* G *les couples, ou certains des couples,* (n'', 1). Autrement dit, pour qu'un ensemble de couples d'entiers (n'', n') tels que $0 < n' < n''$ soit un Δ''^{1}_1 (b, c), il suffit qu'il contienne l'ensemble ${}^0\Delta''^{1}_1$ (b, c) des couples ($|\omega'| + |\omega''| + 1, |\omega'|$) associés aux mots ω', ω'' sur N tels que $|\omega'| \geqslant 2$ et qu'il existe $A \in N$ pour lequel A ::= ω' et bω'' est facteur droit d'un mot de $\mathscr{C}$(A, c). Nous noterons Δ^1_1 (b, c) la réunion de Δ'_1 (b, c) et Δ''^{1}_1 (b, c).

7. 10. 4. Deuxième transformation de la grammaire G. Ψ et Ψ' étant des mots sur V, nous noterons $\Psi' \xrightarrow[=]{*} \Psi$ lorsque Ψ dérive de Ψ' et $|\Psi| = |\Psi'|$: si $\Psi = Z_1 \dots Z_n$ et $\Psi' = Z'_1 \dots Z'_n$, $Z'_i \xrightarrow{*} Z_i$ pour $1 \leqslant i \leqslant n$. Définissons une relation de production $::=_2$: pour $A \in N$, $\Psi \in V^*$:
$A ::=_2 \Psi \Longleftrightarrow (|\Psi| \geqslant 2$ ou $\Psi \in T^*)$ et $(\exists\, \Psi' \in V^*)$ (A ::= Ψ' et $\Psi' \xrightarrow[=]{*} \Psi$).
Soit N'_2 l'ensemble des symboles non terminaux dérivant d'un axiome de la grammaire G. Par une démonstration analogue à celle donnée en 7. 10. 2, on voit qu'on définit une surjection de $\mathscr{C}$(G) dans $\mathscr{C}$(N, T, $::=_2$, N'_2) en associant à (I, f) ∈ $\mathscr{C}$(G) une de ses sous-pseudo-arborescences ((I'', f'')) où les points de I'' sont les nœuds de I prédécesseurs de plusieurs points ou d'une feuille, et les feuilles de I : I'' est une arborescence sous-réduite de I. *Les deux grammaires* G *et* $G_2 = (N, T, ::=_2, N'_2)$ *engendrent le même langage, et chaque phrase y possède les mêmes pseudo-arborescences des sous-propositions.*

En choisissant un ensemble N' réduit à un seul élément, on peut dire aussi que si, pour la structure $\mathscr{S}$(N ∪ T, ::=) , un mot λ dérive de A ∈ N, il existe un A' ∈ N dérivant de A tel que, pour $\mathscr{S}$(N ∪ T, $::=_2$), λ dérive de A'.

A une pseudo-arborescence (I, f) de $\mathscr{C}$(G) on a associé ((I', f')) ∈ $\mathscr{C}(G_1)$(7. 10. 2) et ((I'', f'')) ∈ $\mathscr{C}(G_2)$, qui possèdent la même suite de feuilles ; I' et I'' sont des arborescences orientées isomorphes ayant la même suite de feuilles ; leurs piles conjointes $\overline{U}'$ et $\overline{U}''$ sont isomorphes et f les transcrit en les piles conjointes U' et U'' à ((I', f')) et ((I'', f'')) : les entrées de U' sont les entrées de U'', les sorties de U' sont les sorties de U'', $\alpha^{U'}_{i-} = \alpha^{U''}_{i-}$ pour tout i. *Les ensembles* Δ_1 (b, c) *et* Δ_2 (b, c) *relatifs à la grammaire* G_2 *sont les mêmes que ceux relatifs à* G_1, d'après leur définition. On sait donc les déduire de la grammaire G (7. 10. 3).

L'isomorphisme Θ qui transforme I' en I'' est tel que, pour tout point x de I, f [Θ (x)] dérive de f (x). Soit η'' Z un état u''_i de U'' : η'' Z est transcrit par f d'un état $\overline{u}''_i = \overline{\eta}''z$ de $\overline{U}''$; Θ^{-1} transcrit $\overline{u}''_i$ en $\overline{u}'_i = \overline{\eta}'y$; d'après le théorème 2. 9, $\overline{u}'_i$ est aussi un état $\overline{u}_j$ de la pile $\overline{U}$ conjointe à I, et l'état $\overline{u}_k$ de $\overline{U}$ dont z est le sommet est $\overline{\eta}'z$; f transcrit $\overline{\eta}'z$ en η Z, état de la pile conjointe à (I, f) ; et $\eta = f^*(\overline{\eta}') \xrightarrow[=]{*} \eta'' = (f \circ \Theta)^*(\overline{\eta}')$. Désignons par $\mathscr{C}$(Z, γ) et $\mathscr{C}_2$(Z, γ) les ensembles des mots σ η tels que (η, γ) soit un contexte de Z pour les grammaires G et G_2.

Lemme 8. *Pour tout mot σ η'' de* $\mathscr{C}_2$ (Z, γ) *il existe un mot σ η de* $\mathscr{C}$ (Z, γ) *tel que* $\eta \xrightarrow[=]{*} \eta''$

7. 10. 5. La première transformation de G en G_1 introduit des règles ayant le même second membre, de sorte que même si elle permet de vérifier la condition CT 1, CT 2 n'est pas vérifiée : nous verrons en 7.11 comment on peut se ramener à une grammaire qui réalise CT 2. La transformation de G en G_2 n'est guère directement utilisable, car elle risque d'augmenter considérablement le nombre des règles. Au paragraphe 7. 12, nous en déduirons une méthode pratique.

7. 11. Transformation de la grammaire permettant de réaliser CT 2.

7. 11. 1. *A toute grammaire* G = (N, T, ::=, N') , *nous allons associer une grammaire* $G_3 = (N_3, T, ::=_3, N'_3)$ *de même vocabulaire terminal* T, *vérifiant* CT 2, *de manière que pour toute pseudo-arborescence* (I, f) *de l'un des ensembles* $\mathscr{C}$(G), $\mathscr{C}(G_3)$, *il existe une pseudo-arborescence* (I, f') *de l'autre, image de la même arborescence orientée* I *et possédant même suite de feuilles.* Les deux grammaires engendrent donc le même langage. Une phrase n'est pas nécessairement suite de feuilles du même nombre d'arborescences dans $\mathscr{C}$ (G) et $\mathscr{C}(G_3)$ car, (I, f) étant donnée, (I, f') n'est pas nécessairement unique ; mais toute phrase possède les mêmes pseudo-arborescences des sous-propositions, car la pseudo-arborescence des sous-propositions définie par une pseudo-arborescence (I, f) ne dépend que de I et de la suite de feuilles de (I, f).

N_3 est l'ensemble des parties non vides de N. Posons V = N ∪ T, $V_3 = T \cup N_3$: il sera commode, pour unifier les expressions, de confondre tout élément a de T avec l'ensemble {a}. La relation de production est définie de la manière suivante : F, $F_1, \dots, F_n$ appartenant à V_3, $F ::=_3 F_1 \dots F_n$ si, et seulement si, F est l'ensemble des A ∈ N tels qu'il existe $Z_1 \in F_1, \dots, Z_n \in F_n$ vérifiant $A ::= Z_1 \dots Z_n$. D'après cette définition, $F_1, \dots F_n$ étant donnés, il existe au plus un F tel que $F ::= F_1 \dots F_n$. La grammaire G_3 vérifie la condition CT 2. L'ensemble des axiomes N'_3 est formé des parties de N contenant un symbole de N'.

Soit $(I, f) \in \mathscr{C}(G)$. Définissons $f'(x) \in V_3$ pour les points x de I, étudiés dans l'ordre S de sortie de l'arborescence orientée I, de manière que $f(x)$ appartienne à $f'(x)$:
a) si x est une feuille de I, $f'(x) = f(x)$;
b) si x est un nœud et '$y_1 \dots y_n$' la famille dont il est le prédécesseur, $f(x) ::= f(y_1) \dots f(y_n)$; et $y_i \; S \; x$ pour $1 \leqslant i \leqslant n$ (théorème 1. 3) ; $f'(x)$ est défini comme l'ensemble des $A \in N$ tels qu'il existe $Z_1 \in f'(y_1), \dots, Z_n \in f'(y_n)$ vérifiant $A ::= Z_1 \dots Z_n$; $f(x)$ appartient à $f'(x)$.
(I, f') appartient à $\mathscr{C}(G_3)$ et a même suite de feuilles que (I, f).

Réciproquement, soit $(I, f) \in \mathscr{C}(G_3)$. Définissons une application f' de I dans V, de manière que $f'(x)$ appartienne à $f(x)$:
a) si x est la racine de I, $f'(x) \in N' \cap f(x)$ (qui n'est pas vide d'après la définition de N'_3) ;
b) on définit $f'(x)$ pour les x dont un nœud y est prédécesseur, les y étant rangés dans l'ordre P d'entrée de I : si '$x_1 \dots x_n$' est la famille de prédécesseur y, y est la racine de I ou $pd(y) \; P y$ (théorème 1. 3) ; $f(y) ::=_3 f(x_1) \dots f(x_n)$; on choisit $f'(x_1) = Z_1 \in f(x_1), \dots, f'(x_n) = Z_n \in f(x_n)$ tels que $f'(y) ::= f'(x_1) \dots f'(x_n)$. (I, f') appartient à $\mathscr{C}(G)$ et a même suite de feuilles que (I, f) puisque, d'après la convention faite plus haut, $f'(x) \in f(x)$ signifie $f'(x) = f(x)$ lorsque $f(x) \in T$.

A toute pile de l'un des ensembles $\mathscr{J}(G)$ ou $\mathscr{J}(G_3)$, on peut donc associer une pile de l'autre, transcrite de la même pile simple, les mots d'entrée des deux piles ayant même trace sur T^*. Il résulte alors de la définition des $\Delta_1(b, c)$ et $\Delta_2(b, c)$ que ces ensembles sont les mêmes pour les deux grammaires.

7. 11. 2. Nous allons maintenant définir un pseudo-graphe de production du troisième type $(J_3, \Gamma_3 \,;\, g_3)$ pour la grammaire G_3, qui nous permette dans la pratique de nous ramener à un pseudo-graphe de production $(J, \Gamma \,;\, g)$ de la grammaire donnée G. Rappelons que pour définir $(J, \Gamma \,;\, g)$ on associe à tout $A \in N$ un $\overline{A}$ appartenant à un ensemble $\overline{N}$; $(J, \Gamma \,;\, g)$ est un pseudo-graphe dans l'ensemble $V \cup \overline{N}$ et les mots $\widetilde{\Psi}\,\overline{A}$ associés aux règles $A ::= \Psi$ sont ceux de ses chemins dont l'origine appartient à un sous-ensemble J' de J et l'extrémité à un sous-ensemble J'' ; w et w' étant deux points de J, nous dirons que w est *lié* à w' lorsque $w \; \Gamma \; w'$. Ici, à tout élément F de N_3, c'est-à-dire à tout sous-ensemble de N, associons l'ensemble $\overline{F}$ des $\overline{A} \in \overline{N}$ tels que $A \in N$: soit $\overline{N}_3$ l'ensemble de ces $\overline{F}$; $(J_3, \Gamma_3 \,;\, g_3)$ sera un pseudo-graphe dans $V_3 \cup \overline{N}_3$. J_3 est le sous-ensemble du produit cartésien $(V_3 \cup \overline{N}_3) \times \mathfrak{P}(J)$ formé des couples (F, W) tels que $W \neq \emptyset$ et $g(W)$ (ensemble des $g(w)$ pour $w \in W$) soit contenu dans F. La relation Γ_3 est définie dans J_3 par :
- pour $\overline{F}' \in V_3$, $(F, W) \; \Gamma_3 \; (F', W')$ si, et seulement si, W' est l'ensemble des points d'image dans F' liés à un point de W ;

- pour $F' \in \overline{N}_3$, $(F, W) \; \Gamma_3 \; (F', W')$ si, et seulement si, W' est l'ensemble des points de J'' liés à un point de W et $g(W') = F'$.(1)
Enfin, $g_3(F, W) = \overline{F}$.

Soit J'_3 l'ensemble des couples (F, W) de J_3 tels que W soit l'ensemble des points de J' d'image dans F, et J''_3 l'ensemble des couples $(\overline{F}, W)$ tels que $\overline{F} \in \overline{N}_3$. Montrons que, pour que $F ::=_3 F_1 \dots F_n$, il faut et il suffit que $F_n \dots F_1 \, \overline{F}$ soit un chemin du pseudo-graphe qui vient d'être défini ayant son origine dans J'_3 et son extrémité dans J''_3 ; autrement dit, qu'il existe $W_1, \dots, W_n$, W tels que $(F_n, W_n) \; \Gamma_3 \dots \Gamma_3 (F_1, W_1) \; \Gamma_3 (\overline{F}, W)$ et $(F_n, W_n) \in J'_3$.

Envisageons un tel chemin $(F_n, W_n) \; \Gamma_3 \dots \Gamma_3 \; (F_1, W_1) \; \Gamma_3 \; (\overline{F}, W)$ avec $(F_n, W_n) \in J'_3$. $g(W) = \overline{F}$ est l'ensemble des images par g des points w de J'' tels qu'il existe $w_1 \in W_1$ pour lequel $w_1 \; \Gamma \; w$, donc l'ensemble des images des points w de J'' tels qu'il existe $w_1 \in J$, $w_2 \in W_2$ pour lesquels $g(w_1) \in F_1$ et $w_2 \; \Gamma \; w_1 \; \Gamma \; w$, etc ... ; $\overline{F}$ est l'ensemble des images des w de J'' tels qu'il existe des points $w_1, \dots, w_{n-1}$ de J, w_n de W_n pour lesquels $g(w_1) \in F_1, \dots g(w_{n-1}) \in F_{n-1}$ et $w_n \; \Gamma \; w_{n-1} \dots \Gamma \; w_1 \; \Gamma \; w$, donc l'ensemble des images des w de J'' tels qu'il existe $w_1, \dots w_n$ de J pour lesquels $g(w_1) \in F_1, \dots g(w_n) \in F_n$, $w_n \in J'$ et $w_n \; \Gamma \; w_{n-1} \dots \Gamma \; w_1 \; \Gamma \; w$; enfin $\overline{F}$ est l'ensemble des $\overline{A} \in \overline{N}$ tels qu'il existe $A_1 \in F_1, \dots, A_n \in F_n$ pour lesquels $A ::= A_1 \dots A_n$. Par suite, $F ::=_3 F_1 \dots F_n$.

Réciproquement, supposons $F ::=_3 F_1 \dots F_n$, et construisons un chemin $(F_n, W_n) \; \Gamma_3 \dots \Gamma_3 \; (F_1, W_1) \; \Gamma_3 \; (\overline{F}', W)$, d'origine dans J'_3, d'extrémité dans J''_3 :
- W_n est l'ensemble des points de J' d'image dans F_n ;
- pour $1 \leqslant i < n$, W_i est l'ensemble des points liés à un point de W_{i+1} et d'image dans F_i ;
- W est l'ensemble des points de J'' liés à un point de W_1 et $g(W) = \overline{F}'$.
Il suffit de montrer qu'aucun des ensembles W_i, W n'est vide. F n'est pas vide, et si $A \in F$, il existe $A_1 \in F_1, \dots, A_n \in F_n$ tel que $A ::= A_1 \dots A_n$; donc il existe un chemin $w_n \dots w_1 \, w$ de (J, Γ) tel que $w_n \in J'$, $g(w_1) = A_1, \dots, g(w_n) = A_n$, $g(w) = \overline{A}$; de proche en proche, pour $i = n, \dots, 1$, $w_i \in W_i$, puis $w \in W$. Alors, d'après l'étude directe $F' ::=_3 F_1 \dots F_n$; comme il n'existe pas deux règles de même second membre $F = F'$.

D'après les définitions de J'_3, Γ_3, J''_3, le pseudo-graphe possède les propriétés PG 1, PG 2, PG 3' (4. 6. 4).

7. 11. 3. Dans la pratique, on effectue d'abord sur la grammaire G donnée la transformation définie en 7. 10. 2, puis on fait subir à la grammaire G_1 obtenue

(1) (J_3, Γ_3^{-1}) est un multigraphe d'ensemble de points $\mathfrak{P}(J)$, d'ensemble de noms d'arêtes $V_3 \cup \overline{N}_3$, au sens de 4. 5. 4.

la transformation qui vient d'être définie. La grammaire $G_3 = (N_3, T, ::=_3, N'_3)$ ainsi trouvée satisfait à CT 2, ne possède pas de règle dont le second membre se réduit à un symbole non terminal, vérifie donc DA et PA et engendre le même langage que la grammaire donnée, chaque phrase conservant les mêmes pseudo-arborescences des sous-propositions.

Pour définir le transducteur tr^4 ou $tr^{4'}$ servant à l'analyse (7. 9), on doit remplacer chaque ensemble Δ_1 (b, c) par un Δ_1^1 (b, c) dont on a étudié la définition à partir de la grammaire G en 7. 10. 3. On ne représente évidemment pas par une matrice d'enchaînement le pseudo-graphe de production de la grammaire G_3, défini en 7. 11. 2 : on se contente du pseudo-graphe $(J, \Gamma ; g)$ de G. La définition de $(J_3, \Gamma_3 ; g_3)$ a la signification suivante : lors de sorties consécutives de la pile d'ensembles $F_n, \dots, F_1$, encadrées de deux entrées, on détermine l'ensemble W_n des points o (A_n) de J tels que $A_n \in F_n$, puis l'ensemble W_{n-1} des points liés à un point de W_n et ayant leur image dans F_{n-1} etc ... (si l'un des ensembles W_i est vide, le mot analysé n'appartient pas au langage) ; on dispose enfin de l'ensemble W_1 des extrémités des chemins de $(J, \Gamma ; g)$, $A_n \dots A_1$ où $A_i \in F_i$, qui sont seconds membres de règles ; on détermine alors l'ensemble F des premiers membres de ces règles et F entre dans la pile (s'il n'est pas vide). Pour que tr^4 ou $tr^{4'}$ soit déterministe, il faut et il suffit que les ensembles Δ_1^1 (b, c) et Δ_2 (b, c) vérifient CT 1.

On peut montrer directement que, dans le cas général, si une grammaire G vérifie DA, il en est de même pour la grammaire G_3 qui en est déduite par la transformation étudiée dans le paragraphe 7. 11.

7. 12. Simplication du pseudo-graphe de production par transcription des règles.

7. 12. 1. Introduction. Pour le générateur gn^4, le pseudo-graphe de production a comme seul rôle, nous l'avons déjà dit, de vérifier que les piles engendrées satisfont à la condition C 3. En réalité, ce rôle est aussi joué en partie par les tests utilisant les ensembles Δ_1 (b, c) et Δ_2 (b, c), qui permettent parfois de reconnaître que le mot étudié n'appartient pas au langage, si bien que certaines vérifications sont en quelque sorte effectuées à deux reprises. Nous tenterons ici de tenir compte de ces tests pour pouvoir remplacer le pseudo-graphe de production par un pseudo-graphe plus simple. Cette simplication sera appliquée au pseudo-graphe de la grammaire G_2 associée en 7. 10. 4 à une grammaidonnée G : comme nous l'avons signalé, les règles de G_2 peuvent être beaucoup plus nombreuses que celles de G, ce qui rend une telle simplication pratiquement indispensable.

Soit donc une grammaire $G = (N, T, ::=, N')$, transformée en $G_2 = (N, T, ::=_2, N'_2)$, un ensemble N_5 et une application de N dans N_5, prolongée par l'identité en une application τ de $V = N \cup T$ dans $V_5 = N_5 \cup T$.

τ définit une transcription τ^* de V^* dans V^*_5, et une relation de production $::=_5$, par les seules règles : $\tau(A) ::=_5 \tau^*(\Psi)$ lorsque $A ::=_2 \Psi$. Plusieurs règles peuvent être ainsi remplacées par une seule. La relation $::=_5$ est représentée par un pseudo-graphe de production du troisième type $(J_5, \Gamma_5 ; g_5)$, possédant les propriétés PG 1, PG 2, PG 3', qui peut être bien plus simple qu'un pseudo-graphe analogue représentant $::=_2$; pour définir $(J_5, \Gamma_5 ; g_5)$, on associe à tout $B \in N_5$ un élément $\bar{B}$ d'un ensemble $\bar{N}_5$; τ sera prolongée par $\tau(\bar{A}) = \tau(A)$ en une application de $V \cup \bar{N}$ dans $V_5 \cup \bar{N}_5$.

Nous voulons maintenant définir un générateur qui, pour toute donnée $\alpha \in T^*$ engendre les piles transcrites par τ des piles conjointes aux pseudo-arborescences de $\mathscr{C}(G_2)$ dont α est la suite de feuilles, et aucune autre. Comme τ laisse invariants les symboles terminaux, on connaîtra ainsi les pseudo-arborescences des sous-propositions de α relatives à la grammaire G donnée.

7. 12. 2. Générateur gn^5. Il est défini par $W = V_5$ (nous noterons $E_5 = V_5 \cup \{\sigma\}$ et τ sera encore prolongée par $\tau(\sigma) = \sigma$), cd = 1 et un transducteur tr^5 : tr^5 utilise des ensembles Δ_1^1 (b, c) et Δ_2 (b, c) relatifs à la *grammaire* G *donnée*, le pseudo-graphe $(J_5, \Gamma_5 ; g_5)$ et l'ensemble $N'_5 = \tau(N')$. Nous introduirons aussi une application qui τ-projette (5. 5) dans tr^5 le transducteur $tr^{4'}$ d'un générateur $gn^{4'}$ (7. 9. 3) engendrant les piles de $\mathscr{J}(G_2)$: par suite, $gn^5(\alpha)$ contient les piles transcrites par τ des piles conjointes aux pseudo-arborescences de $\mathscr{C}(G_2)$ dont α est la suite de feuilles, mais peut en contenir d'autres. Ce transducteur $tr^{4'}$ est défini grâce aux mêmes ensembles Δ_1^1 (b, c) et Δ_2 (b, c) que tr^5, et à un pseudo-graphe de production $(J, \Gamma ; g)$ de G_2, que nous déduirons de $(J_5, \Gamma_5 ; g_5)$ et d'un autre pseudo-graphe de production de G_2, $(J_2, \Gamma_2 ; g_2)$, du troisième type et vérifiant PG 1, PG 2, PG 3'.

Introduisons d'abord ce pseudo-graphe $(J, \Gamma ; g)$: J est l'ensemble des $(w', w'') \in J_2 \times J_5$ tels que $\tau[g_2(w')] = g_5(w'')$; $(w'_1, w''_1) \Gamma (w'_2, w''_2)$ signifie $w'_1 \Gamma_2 w'_2$ et $w''_1 \Gamma_5 w''_2$; $g(w', w'') = g_2(w')$. Si $A ::=_2 \Psi$, $\Psi\bar{A}$ est l'image par g_2 d'un chemin $w'_1 \dots w'_n$ de (J_2, Γ_2) dont l'origine appartient à un ensemble J'_2 et l'extrémité à un ensemble J''_2 ; $\tau(A) ::=_5 \tau^*(\Psi)$, donc $\tau^*(\Psi \bar{A})$ est l'image par g_5 d'un chemin $w''_1 \dots w''_n$ de (J_5, Γ_5) dont l'origine et l'extrémité appartiennent à deux ensembles J'_5 et J''_5 ; $\Psi \bar{A}$ est donc l'image par g d'un chemin $(w'_1, w''_1) \dots (w'_n, w''_n)$ de (J, Γ) dont l'origine appartient à $J'_2 \times J'_5 = J'$ et l'extrémité à $J''_2 \times J''_5 = J''$. Réciproquement tout chemin de $(J, \Gamma ; g)$ d'origine dans J', d'extrémité dans J'' est un chemin de $(J_2, \Gamma_2 ; g_2)$ d'origine dans J'_2, d'extrémité dans J''_2, donc de la forme $\Psi \bar{A}$ avec $A ::=_2 \Psi$. Comme $(J_2, \Gamma_2 ; g_2)$ et $(J_5, \Gamma_5 ; g_5)$ vérifient les propriétés PG 1, PG 2, PG 3', on voit facilement qu'il en est de même pour $(J, \Gamma ; g)$. La définition de $(J, \Gamma ; g)$ nous permet de définir une application τ' de J dans J_5, qui à $w = (w', w'')$ associe $\tau'(w) = w''$: $\tau \circ g = g_5 \circ \tau'$; τ' transforme J' en J'_5 , J'' en J''_5 et $w_1 \Gamma w_2$ entraîne $\tau'(w_1) \Gamma_5 \tau'(w_2)$.

Venons en aux définitions de tr^5 et de l'application qui τ-projette $tr^{4'}$ dans tr^5 :

a) $M^5 = E_5^{cg} \times J_{15} \times T' \times [cg] \times [cg]$: cg et $[cg]$ sont définis comme en 7. 9, ainsi que $\delta(\eta')$ et $\xi(\eta')$ pour $\eta' \in E_5^{cg}$; $J_{15} = J_5 \cup \{\sigma\}$.

b) $q = cg + 1$. Définissons ϖ, application de $M' \times E^*$ dans $M'^5 \times E_5^*$ ($E = V \cup \{\sigma\}$ est l'alphabet de $tr^{4'}$, $M = E^{cg} \times J_1 \times T' \times [cg] \times [cg]$ entre dans la définition de $tr^{4'}$, $M' = M \cup \{s_f\}$, $M'^5 = M^5 \cup \{s_f\}$) :

$\varpi(s, v) = [\varpi_1(s), \tau^*(v)]$ avec $\varpi_1(\eta', w, \alpha, \delta', \delta'') = (\tau^*(\eta'), \tau'(w), \alpha, \delta', \delta'')$, τ' étant prolongée en une application de J_1 dans J_{15} par $\tau'(\sigma) = \sigma$, $\varpi_1(s_f) = s_f$.

Pour que la restriction de ϖ à $M \times E^*$ τ-projette $tr^{4'}$ dans tr^5, il suffit que les applications l, d, m définissant $tr^{4'}$ et les applications l^5, d^5, m^5 définissant tr^5 vérifient :

- $l^5[\tau^*(\rho)] = \varpi_1[l(\rho)]$ dès que $l(\rho)$ est défini (avec les notations de 5. 5, $\rho^0 = \tau^*(\rho^1)$ et $v^0 = \tau^*(v^1)$),
- $d^5[\varpi_1(s), \tau(x)]$ contient l'ensemble $\varpi_1[d(s, x)]$,
- $m^5[\varpi_1(s)]$ contient $\varpi[m(s)]$.

Ces propriétés sont faciles à vérifier sur les définitions qui suivent, où a, b, c décrivent T', B décrit N_5, η décrit E_5^{cg-1}, Y et Z décrivent E, w décrit J_{15}, δ' et δ'' décrivent $[cg]$.

$l^5(a\, b\, \eta) = (b\, \eta, \sigma, a, 0, 0)$.

c) Si $\eta\, Z \notin N_5^*$ et $(1, 1) \in \Delta_1^1(b, a)$, $d^5[(b\, \eta, \sigma, a, 0, 0), Z]$ est formé des $(\eta\, Z, o(b), a, n', \delta(\eta\, Z) - n')$ pour tout entier n' tel que $(\delta(\eta\, Z), n') \in \Delta_2(\xi(\eta\, Z), b)$;

si $\delta' \neq 0$, $(\delta', \delta'') \neq (1, 0)$ et si un point w' de $\Gamma_5(w)$ vérifie $g_5(w') = B$, $d^5[(B\, \eta, w, a, \delta', \delta''), Z] = \{(\eta\, Z, w', a, \delta' - 1, \delta'')\}$;

si $\eta\, Z \notin N_5^*$ et si un point w' de $\Gamma_5(w)$ vérifie $g_5(w') = c$, $d^5[(c\eta, w, a, 1, 0), Z]$ est formé des $(\eta\, Z, w', a, n', \delta(\eta\, Z) - n')$ pour tout entier n' tel que $(\delta(\eta\, Z), n') \in \Delta_2(\xi(\eta\, Z), c)$.

d) Si $(1, 0)$ ou $(1, 1)$ appartient à $\Delta_2(b, a)$, $m^5(b\, \eta, \sigma, a, 0, 0) = \{(s_f, 'a')\}$;

$m^5(Y\eta, w, a, 0, \delta'')$ contient $[(Y\eta, w', a, 0, \delta''), 'A']$ pour tout w' de J''_5 tel que $w \Gamma_5 w'$ et $g_5(w') = \bar{A}$, et pas d'autre élément ;

si $w \in J''_5$, $g_5(w) = \bar{A}$ et $Y\eta \notin N_5^*$, $m^5(Y\eta, w, a, 0, \delta'')$ contient les seuls éléments :

- $(s_f, 'a')$ lorsque dans $\Delta_2(\xi(Y\eta), a)$ se trouve un couple de premier entier $\delta'' + 1$ ou lorsque $Y = a = \sigma$ et $A \in N'_5$,
- $[(Y\eta, o(A), a, n' - 1, \delta'' + 1 - n'), '\sigma']$ pour tout n' tel que $(\delta'' + 1, n') \in \Delta_1^1(\xi(Y\eta), a)$, sauf lorsque $Y = a = \sigma$.

On montre, comme pour tr^4, que le transducteur tr^5 se projette dans le transducteur tr^0 relatif à la grammaire $G_5 = (N_5, T, ::=_5, N'_5)$. Cette grammaire vérifie DA et PA puisque le second membre d'une de ses règles ne se réduit jamais à un symbole auxiliaire. tr^5 satisfait à GP 2 ; gn^5 satisfait à GP 1

et les piles qu'il engendre appartiennent à $\mathcal{J}(G_5)$. Pour que tr^5 soit déterministe, il faut et il suffit que les ensembles $\Delta_1^1(b, c)$ et $\Delta_2(b, c)$ vérifient la condition CT 1 et que la relation $::=_5$ vérifie CT 2.

7. 12. 3. *Montrons que les piles engendrées par* gn^5 *possèdent les propriétés* C 6 *et* C 7 (7. 9. 2), *pour les ensembles* $\Delta_1^1(b, c)$ *et* $\Delta_2(b, c)$ *utilisés par* tr^5, qui sont définis à partir de la grammaire G. Soit $U = (u_0, \ldots, u_n)$ une pile engendrée pour une donnée $\alpha = 'a_0 \ldots a_{p-1}'$. D'après 5. 3. 1, il existe un p-partage $(j_0, \ldots, j_{p+1})$ de U et une suite de triplets (s_k, y_k, λ_k), $1 \leqslant k \leqslant n'$, tels que les triplets pour lesquels $j_i < k \leqslant j_{i+1}$ forment un calcul complet d'entrée $a_i\, \tilde{u}_{j_i}\, \sigma^{cg}$ (en posant $a_p = \sigma$), de résultat μ_i (notations de 5. 3. 1). Le transducteur est tel que, pour tout k, $|\lambda_k| = 1$, de sorte que $n' = n$ et $u_k = u_{k-1} \boxtimes \lambda_k$. Lorsque k n'est pas un j_i, nous poserons $s_k = (Y_k\, \eta_k, w_k, a'_k, \delta'_k, \delta''_k)$: $Y_k \in E_5$, $\eta_k \in E^{cg-1}$, $w_k \in J_{15}$, $a'_k \in T'$, $\delta'_k \in [cg]$, $\delta''_k \in [cg]$. $a_i\, \tilde{u}_{j_i}\, \sigma^{cg}$ s'écrit $a_i\, b_{j_i}\, \eta\, v'_i$ où $b_{j_i} = a_{i-1}$ si $i \neq 0$, $b_{j_0} = \sigma$ et $|\eta| = cg - 1$; nous définirons aussi u'_k par $v'_i = y_{j_i+1} \ldots y_k\, u'_k$.

Outre d'autres propriétés, nous allons montrer, par récurrence sur k vérifiant $j_i < k < j_{i+1}$, que $a'_k = a_i$ (qui est le premier signe de $\alpha \overset{U}{\underset{k}{-}} \sigma$), $\delta'_k + \delta''_k = \delta(Y_k\, \eta_k)$ et que $\tilde{u}_k\, \sigma^{cg} = Y_k\, \eta_k\, u'_k$ si k est une sortie, $\tilde{u}_k\, \sigma^{cg} = A\, Y_k\, \eta_k\, u'_k$ si k est une entrée de $A \in N$: d'où $|b_k\, \omega_k| = \delta(Y_k\, \eta_k)$ dans le premier cas, $|b_k\, \omega_k| = \delta(Y_k\, \eta_k) + 1$ dans le second, $b_k = \xi(Y_k\, \eta_k)$, si b_k et ω_k sont définis pour la pile U comme en 7. 9. 2.

1) $k = j_i + 1$: $l^5(a_i\, b_{j_i}\, \eta) = (b_{j_i}\, \eta, \sigma, a_i, 0, 0)$; k est une sortie :

$s_k \in d^5[(b_{j_i}\, \eta, \sigma, a_i, 0, 0), y_k]$: $a'_k = a_i$, $\tilde{u}_k\, \sigma^{cg} = \eta y_k\, u'_k = Y_k\, \eta_k\, u'_k$; $\delta'_k + \delta''_k = \delta(Y_k\, \eta_k)$. D'autre part, $(1, 1) \in \Delta_1^1(b_{j_i}, a_i)$ alors que $n_{j_i} = 1$ et $|b_{j_i}\, \omega_{j_i}| = 1$. Enfin $(|b_k\, \omega_k|, \delta'_k) \in \Delta_2(b_k, b_{j_i})$.

Dans ce qui suit, on démontre les hypothèses de récurrence pour k + 1 en les supposant vraies pour k.

2) k et k + 1 sont des sorties : $\lambda_k = \sigma$, d'où $w_k \neq \sigma$ et $w_k \notin J''_5$; $\lambda_{k+1} = \sigma$ d'où $s_{k+1} \in d^5(s_k, y_{k+1})$: $a'_{k+1} = a'_k = a_i$, $Y_{k+1}\, \eta_{k+1} = \eta_k\, y_{k+1}$, $\tilde{u}_k\, \sigma^{cg} = Y_k\, \eta_k\, y_{k+1}\, u'_{k+1}$, donc $\tilde{u}_{k+1}\, \sigma^{cg} = Y_{k+1}\, \eta_{k+1}\, u'_{k+1}$; k + 1 est une sortie de Y_k. D'autre part :

a) si $Y_k \in N_5$, $\delta'_{k+1} = \delta'_k - 1$; $\delta''_{k+1} = \delta''_k$; $\delta'_{k+1} + \delta''_{k+1} = \delta(Y_k\, \eta_k) - 1 = \delta(Y_{k+1}\eta_{k+1})$;

b) si $Y_k \in T$, comme $w_k \neq \sigma$, $\delta'_k = 1$, $\delta''_k = 0$ et $(|b_{k+1}\, \omega_{k+1}|, \delta'_{k+1}) \in \Delta_2(b_{k+1}, Y_k)$; $\delta'_{k+1} + \delta''_{k+1} = \delta(Y_{k+1}\, \eta_{k+1})$.

3) k + 1 est une sortie et k une entrée de $A \in N$: $\lambda_k = A$, d'où $w_k \in J''_5$, donc $(s_{k+1}, \sigma) \in m^5(s_k)$: $a'_{k+1} = a'_k = a_i$, $\tilde{u}_k\, \sigma^{cg} = A\, Y_k\, \eta_k\, u'_k$. $\tilde{u}_{k+1}\, \sigma^{cg} = Y_k\, \eta_k\, u'_k = Y_{k+1}\, \eta_{k+1}\, u'_{k+1}$; $\delta'_{k+1} + \delta''_{k+1} = \delta'_k + \delta''_k = \delta(Y_k\, \eta_k) = \delta(Y_{k+1}\, \eta_{k+1})$. D'autre part $(\delta''_k + 1, \delta'_{k+1} + 1) \in \Delta_1^1(b_k, a_i)$ et $\delta''_k + 1 = \delta(Y_k\, \eta_k) + 1 = |b_k\, \omega_k|$.

4) k + 1 est une entrée de $A \in N$: $(s_{k+1}, A) \in m^5(s_k)$: $a'_{k+1} = a'_k = a_i$; $\tilde{u}_{k+1}\, \sigma^{cg} = A\, \tilde{u}_k\, \sigma^{cg} = A\, Y_k\, \eta_k\, u'_k = A\, Y_{k+1}\, \eta_{k+1}\, u'_{k+1}$; $\delta'_k = \delta'_{k+1} = 0$;

$\delta''_{k+1} = \delta''_k = \delta(Y_k \eta_k) = \delta(Y_{k+1} \eta_{k+1})$.

Etudions maintenant n_k (défini en 7. 9. 2) :
- si $k + n_k$ est sortie d'un symbole terminal et $n_k > 1$, $k + n_k - 1$ est une sortie et $\delta'_{k+n_k-1} = 1$ (2, b) : d'après 2, a, $\delta'_{k+1}, \ldots, \delta'_{k+n_k-1}$ est une progression arithmétique de raison -1 : $n_k = \delta'_{k+1} + 1$;
- si $k + n_k + 1$ est entrée d'un symbole auxiliaire, $\delta'_{k+n_k} = 0$ (4) ; $\delta'_{k+1}, \ldots, \delta'_{k+n_k}$ est une progression arithmétique de raison -1 : ici encore, $n_k = \delta'_{k+1} + 1$.

Alors, lorsque k est une entrée et k + 1 une sortie autre que la dernière, $(|b_k \omega_k|, n_k) \in \Delta_1^1(b_k, a_i)$ d'après 1 et 3. Lorsque k est une sortie d'un symbole terminal Y_{k-1} et que k - 1 est aussi une sortie, $(|b_k \omega_k|, \delta'_k) \in \Delta_2(b_k, Y_{k-1})$ (2, b) ; lorsque k est sortie d'un symbole terminal et k - 1 une entrée (de ce symbole), $k = j_i + 1$, le symbole qui sort est b_{j_i} : $(|b_k \omega_k|, \delta'_k) \in \Delta_2(b_k, b_{j_i})$ (1) ; si alors k+1 est sortie d'un symbole auxiliaire, $\delta'_k = \delta'_{k+1} + 1 = n_k$ (2, a) ; si k + 1 est sortie d'un symbole terminal, $n_k = 1$ et $\delta'_k = 1$ (2, b) ; si k + 1 est une entrée, $n_k = 0$ et $\delta'_k = 0$ (4). La pile U vérifie C 6 et C 7.

7. 12. 4. Condition suffisante pour que gn^5 engendre les seules piles cherchées.
Définitions. 1) $A \in N$ est *incompatible* avec un couple (b, c) d'éléments de T' lorsque, quels que soient b', c' dans T, les mots ω et ω' sur N, le mot Ψ sur V :
- $A ::= \omega$ et $|\omega| \geqslant 2$ entraîne qu'aucun couple de $\Delta''^1_1(b, c)$ n'a $|\omega|$ pour second entier ;
- $A ::= \omega c' \omega'$ entraîne : $(|c' \omega'|, |c' \omega'|) \notin \Delta'_1(c', c)$ ou aucun couple de $\Delta''_2(b, c')$ n'a $|\omega|$ pour second entier ;
- $A ::= \omega c' \Psi b' \omega'$ entraîne : $(|b' \omega'|, |b' \omega'|) \notin \Delta'_1(b', c)$ ou aucun couple de $\Delta''_2(b, c')$ n'a $|\omega|$ pour second entier.

2) A est *partiellement incompatible* avec le couple (σ, σ) lorsque, avec les mêmes notations :
- $A ::= \omega$ entraîne $(|\omega| + 1, |\omega|) \notin \Delta''^1_1(\sigma, \sigma)$;
- $A ::= \omega c' \omega'$ entraîne : $(|c' \omega'|, |c' \omega'|) \notin \Delta'_1(c', \sigma)$ ou $(|\omega| + 1, |\omega|) \notin \Delta''_2(\sigma, c')$;
- $A ::= \omega c' \Psi b' \omega'$ entraîne $(|b' \omega'|, |b' \omega'|) \notin \Delta'_1(b', \sigma)$ ou $(|\omega| + 1, |\omega|) \notin \Delta''_2(\sigma, c')$.

Si A est incompatible avec (σ, σ), il est aussi partiellement incompatible avec lui.

3) Un couple (b, c) d'éléments de T' est un *τ-encadrement* de $A' \in N$ lorsqu'il existe $B \in N$, $C \in N$, $\Psi \in V^*$ tels que : $C \xrightarrow{*} A'$, $\tau(C) \neq \tau(A')$ ou $C = A'$, et
- soit $B ::= \Psi C$, $\Psi \neq \Lambda$ et b est le dernier élément de T' dans un mot de $\mathscr{C}(B, c)\Psi$,
- soit il existe $Z \in V$, $\Psi' \in V^*$ pour lesquels $B ::= \Psi C Z \Psi'$, b est le dernier élément de T' dans un mot de $\mathscr{C}(B, \Lambda)\Psi$ et c est initiale de Z.

(On peut remarquer que dans les deux cas, b est le dernier élément de T' dans un mot de $\mathscr{C}(A', c)$).

4) L'application τ est séparante si, lorsque $\tau(A) = \tau(A')$, d'une part $A' \xrightarrow{*} A$ ou A est incompatible avec tout τ-encadrement de A', et d'autre part, quand A' dérive d'un axiome, A dérive d'un axiome ou est partiellement incompatible avec (σ, σ).

Théorème 7. 10. *Si la grammaire G est réduite et l'application τ séparante, pour tout mot α de T*, les piles de gn$^5(\alpha)$ sont les transcrites par τ des piles conjointes aux pseudo-arborescences de $\mathscr{C}(G_2)$ qui ont α pour suite de feuilles.*

Il reste à montrer que la pseudo-arborescence (I, f') de $\mathscr{C}(G_5)$ à laquelle une pile U engendrée par gn^5 est conjointe est la transformée par τ d'une pseudo-arborescence (I, f) de $\mathscr{C}(G_2)$; nous avons pour cela à définir une application f telle que $f' = \tau \circ f$.

Si y est une feuille de I, f (y) = f' (y). Les nœuds seront considérés dans l'ordre S^{-1} inverse de l'ordre de sortie de I. Pour chacun d'eux, y, prédécesseur d'une famille '$z_1 \ldots z_r$', on va déterminer f (y), r symboles Z_j et r symboles Z_j^0 $(1 \leqslant j \leqslant r)$ de manière que $f'(y) = \tau[f(y)]$, $f'(z_j) = \tau(Z_j)$, $Z_j^0 = Z_j$ ou $\tau(Z_j^0) \neq \tau(Z_j)$, $Z_j^0 \xrightarrow{*} Z_j$ et $f(y) ::= Z_1^0 \ldots Z_r^0$ (d'où $f(y) ::=_2 Z_1 \ldots Z_r$) ; si y est la racine, on impose $f(y) \in N'_2$; sinon y possède un prédécesseur x, prédécesseur d'une famille '$y_1 \ldots y_{r'}$', avec $y = y_p$: y S x, de sorte qu'on a déjà associé à x r' symboles $Y_1, \ldots, Y_{r'}$ tels que $f'(y_j) = \tau((Y_j)$ $(1 \leqslant j \leqslant r')$ et $f(x) ::=_2 Y_1 \ldots Y_{r'}$ et r' symboles $Y_1^0, \ldots, Y_{r'}^0$ tels que $Y_j^0 = Y_j$ ou $\tau(Y_j^0) \neq \tau(Y_j)$, $Y_j^0 \xrightarrow{*} Y_j$ et $f(x) ::= Y_1^0 \ldots Y_{r'}^0$; on choisira f (y) tel que $Y_p^0 \xrightarrow{*} f(y)$ et que $f(x) ::=_2 Y_1 \ldots Y_{p-1} f(y_p) \ldots f(y_{r'})$ si $p \neq 1$, $f(x) ::=_2 f(y_1) \ldots f(y_{r'})$ si p = 1. Alors $(I, f) \in \mathscr{C}(G_2)$. Montrons que cette construction est possible.

Puisque $f'(y) ::=_5 f'(z_1) \ldots f'(z_r)$, il existe Y, $Z_1^1, \ldots, Z_r^1$ tels que $\tau(Y) = f'(y)$, $\tau(Z_j^1) = f'(z_j)$, $Y ::=_2 Z_1^1 \ldots Z_r^1$, donc il existe $Z_1^0, \ldots, Z_r^0$ tels que $Y ::= Z_1^0 \ldots Z_r^0$, $Z_j^0 \xrightarrow{*} Z_j^1$; si $\tau(Z_j^1) = \tau(Z_j^0)$ on prend $Z_j = Z_j^0$, sinon $Z_j = Z_j^1$. Nous allons montrer qu'on peut choisir f (y) = Y.

Si y est la racine de I, $f'(y) \in N'_5$ et il existe $Y' \in N'_2$ tel que $f'(y) = \tau(Y') = \tau(Y)$. Il suffit donc de montrer que Y n'est pas partiellement incompatible avec (σ, σ). '$f'(z_1) \ldots f'(z_r)$' est un état u_i de U, i étant une entrée suivie de r sorties, d'une entrée de f' (y) et de l'indice du dernier état ; comme la pile U satisfait à C 6, $(|b_i \omega_i|, n_i) \in \Delta_1^1(b_i, \sigma)$. Dans le cas où tous les Z_i sont non terminaux, $b_i = \sigma$, $|\omega_i| = n_i = r$. Dans le cas où Z_k et $Z_{k'}$ sont le premier et le dernier symbole terminal de '$Z_1 \ldots Z_r$', $b_i = Z_{k'}$, $|\omega_i| = r - k'$, $n_i = r - k' + 1$; $i' = i + r - k + 1$ est une sortie de $Z_{k'}$, $b_{i'} = \sigma$, $|\omega_{i'}| = k - 1$, $n_{i'} = k - 1$ et $(k, k - 1) \in \Delta''_2(\sigma, Z_k)$ (C 7).

Dans la suite, y est un nœud de I autre que la racine ; avec les notations précédentes, $f(x) ::=_2 Y_1 \ldots Y_{p-1}\ Y_p\ f(y_{p+1}) \ldots f(y_{r'})$, $\tau(Y_p) = f'(y) = \tau(Y)$. Si $Y_p \xrightarrow{*} Y$, $f(x) ::=_2 Y_1 \ldots Y_{p-1}\ Y\ f(y_{p+1}) \ldots f(y_{r'})$. Une entrée i dans U et un mot λ sur V sont tels que $u_i = \lambda f'(z_1) \ldots f'(z_r)$, $i+1$ soit une sortie, $u_{i+r+1} = \lambda f'(y)$; soit b le dernier élément de T' dans $\sigma\lambda$ et c le premier élément de $\alpha \overset{U}{\underset{i-}{}} \sigma = \alpha \overset{U}{\underset{i+r-1-}{}} \sigma$. Montrons que Y n'est pas incompatible avec (b, c). $(|b_i\,\omega_i|, n_i) \in \Delta_1^1(b_i, c)$ (C 6). Si aucun Z_j n'est terminal, $n_i = r$, $b_i = b$, $|\omega_i| \geqslant r$. Dans le cas où Z_k et $Z_{k'}$ sont le premier et le dernier symbole terminal de ' $Z_1 \ldots Z_r$ ', $b_i = Z_{k'}$, $|\omega_i| = r - k'$, $n_i = r - k' + 1$; $i' = i + r - k + 1$ est une sortie de Z_k, $b_{i'} = b$, $n_{i'} = k - 1 \leqslant |\omega_{i'}|$, $(|b\,\omega_i|, k-1) \in \Delta_2''(b, Z_k)$ (C 7). Il suffit maintenant de montrer que (b, c) est un τ-encadrement de Y_p.

Soit I'' la sous-arborescence de I contenant, pour tout point v tel que y S v, la famille de v. I'' contient la racine de I et est une sous-arborescence complète (1. 4. 1) de I : en effet, soit R l'ordre associé à I ; si v' est un point de I'', il existe un point v de même prédécesseur z' tel que y S v ; lorsque v'' R v' et v'' ≠ v', v'' R z' d'où v'' R v, v S v'' et y S v''. Définissons une application f'' de manière que l'image de I'' par f'' appartienne à $\mathcal{S}(V, ::=_2)$: si y S v et y ≠ v, f (v) a été défini avant f (y) et on pose f'' (v) = f (v) ; soit une famille ' $y'_1 \ldots y'_k$ ' de prédécesseur z' pour laquelle il existe p' $(1 \leqslant p' < k)$ tel que $y'_{p'}$ S y S $y'_{p'+1}$ et $y \neq y'_{p'+1}$ (il en est ainsi en particulier pour la famille de y et p' = p) : y S z', de sorte qu'à z' on a déjà associé $Y'_1, \ldots, Y'_{p'}$ vérifiant $f(z') ::=_2 Y'_1 \ldots Y'_{p'}\ f(y'_{p'+1}) \ldots f(y'_k)$; la grammaire G étant réduite, il existe un mot sur T qui, pour G, dérive de chaque Y'_j $(1 \leqslant j \leqslant p')$; pour G_2, ce mot dérive d'un Y''_j tel que $Y'_j \xrightarrow{*} Y''_j$ (7. 10. 4), et $f(z') ::=_2 Y''_1 \ldots Y''_{p'}\ f(y'_{p'+1}) \ldots f(y'_k)$: on pose $f''(y'_j) = Y''_j$. (I'', f'') est une pseudo-arborescence de $\mathcal{S}(V, ::=_2)$, qu'on peut compléter par réunion verticale en une pseudo-arborescence (I''_1, f''_1) de $\mathcal{C}(G_2)$. Les suites de feuilles de (I, f'), (I'', f''), (I''_1, f''_1) ont un facteur droit commum, formé des images des feuilles v de I telles que y S v : ce facteur droit commence par c. I'' est une sous-arborescence complète de I et I''_1 (2. 6. 2) : d'après le théorème 2. 8, les piles conjointes à I, I'' et I''_1 ont les mêmes états de sommets x et y. Ces états sont respectivement transcrits par f' en $\eta'\ f'(x)$ et $\eta'\ f'(y_1) \ldots f'(y_p) = \lambda f'(y)$, par f''_1 en $\eta''\ f(x)$ et $\eta''\ Y''_1 \ldots Y''_p$. Si v est un point de I'' et si f' (v) est terminal, $f'(v) = f''(v) = f''_1(v)$; si f' (v) est non terminal, il en est de même pour $f''_1(v)$: b, dernier élément de T' dans $\sigma\,\eta'\ f'(y_1) \ldots f'(y_{p-1})$ l'est également dans $\sigma\,\eta''\ Y''_1 \ldots Y''_{p-1}$. Avec les notations du lemme 8 (7. 10. 4), $\sigma\,\eta''$ appartient à $\mathcal{C}_2(f(x), \Lambda)$, et si $p = r'$ à $\mathcal{C}_2(f(x), c)$; il existe donc un mot $\sigma\,\eta$ appartenant à $\mathcal{C}(f(x), \Lambda)$, et à $\mathcal{C}(f(x), c)$ si $p = r'$, tel que $\eta \xrightarrow[=]{*} \eta''$. b est le dernier élément de T' dans $\sigma\,\eta\ Y^0_1 \ldots Y^0_{p-1}$; $Y^0_p \xrightarrow{*} Y_p$; si $p < r'$, c est une initiale de $f(y_{p+1})$ (d'après 2. 3. 3), donc de Y^0_{p+1} dont dérive $f(y_{p+1})$. (b, c) est bien un τ-encadrement de Y_p.

L'application τ est particulièrement commode si $A \xrightarrow{*} A'$ entraîne $\tau(A) = \tau(A')$. Alors, les règles de la relation $::=_5$ sont obtenues à partir de celles de la relation de production ::= donnée : ce sont les $\tau(A) ::=_5 \tau^*(\varphi)$ pour $A ::= \varphi$ et ($|\varphi| \geqslant 2$ ou $\varphi \in T^*$). D'autre part, la définition d'un τ-encadrement se simplifie : nécessairement, C = A'.

7. 13. Simplification de l'analyse par le choix des ensembles Δ'_1 (b, c), Δ''^1_1 (b, c), Δ'_2 (b, c), Δ''_2 (b, c).

Lors de la génération d'une pile par gn^4, $gn^{4'}$ ou gn^5, on a à déterminer si certains couples d'entiers appartiennent à un ensemble $\Delta_1(b, c)$ ou $\Delta_2(b, c)$. Si la simulation du générateur sur calculatrice oblige à conserver ces ensembles en mémoire, elle risque d'être coûteuse en place. Une grammaire G étant donnée, on a vu (7. 9. 2, 7. 10. 3) comment en déduire des ensembles ${}^0\Delta'_1(b, c)$, ${}^0\Delta''^1_1(b, c)$, ${}^0\Delta'_2(b, c)$, ${}^0\Delta''_2(b, c)$ tels que tout ensemble contenant ${}^0\Delta'_1(b, c)$ (resp. ${}^0\Delta''^1_1(b, c)$, ${}^0\Delta'_2(b, c)$, ${}^0\Delta''_2(b, c)$) soit un $\Delta'_1(b, c)$ (resp. $\Delta''^1_1(b, c)$, $\Delta'_2(b, c)$, $\Delta''_2(b, c)$).

Un cas simple est celui où tous les ensembles $\Delta'_1(b, c)$ (resp. $\Delta''^1_1(b, c)$, $\Delta'_2(b, c)$, $\Delta''_2(b, c)$) non vides, relatifs à une grammaire donnée, ou même à une classe de grammaires, sont choisis égaux à un ensemble $\overline{\Delta}'_1$ (resp. $\overline{\Delta}''_1$, $\overline{\Delta}'_2$, $\overline{\Delta}''_2$) contenant la réunion des ${}^0\Delta'_1(b, c)$ (resp. ${}^0\Delta''^1_1(b, c)$, ${}^0\Delta'_2(b, c)$, ${}^0\Delta''_2(b, c)$) pour les divers couples b, c et éventuellement les diverses grammaires de la classe ; on supposera de plus que $\Delta'_1(b, c)$ (resp. $\Delta''^1_1(b, c)$, $\Delta'_2(b, c)$, $\Delta''_2(b, c)$ est choisi vide lorsque ${}^0\Delta'_1(b, c)$ (resp. ${}^0\Delta''^1_1(b, c)$, ${}^0\Delta'_2(b, c)$, ${}^0\Delta''_2(b, c)$) l'est. Il suffit alors de connaître quatre relations binaires dans T' : ${}^0\Delta'_1(b, c)$ (resp. ${}^0\Delta''^1_1(b, c)$, ${}^0\Delta'_2(b, c)$, ${}^0\Delta''_2(b, c)$) n'est pas vide ; nous noterons ces relations $\gtrdot$ (resp. $\dot{\neq}$, $\doteq$, $\lessdot$) et nous les appellerons *relations de priorité* :

$b \gtrdot c$ *si, et seulement si, il existe* $A \in N$, $\Psi \in V^*$, $\omega \in N^*$ *tels que* $A ::= \Psi\, b\, \omega$ *et* $\mathcal{C}(A, c) \neq \emptyset$ *(c'est-à-dire* A, c *couple vicinal lorsque* $c \in T$, A *finale d'un axiome lorsque* $c = \sigma$, *si la grammaire* G *est supposée réduite) ;*

$b \dot{\neq} c$ *si, et seulement si, il existe* $A \in N$, $\omega' \in N^*$, $\omega'' \in N^*$ *tels que* $|\omega'| \geqslant 2$, $A ::= \omega'$ *et* $b\omega''$ *est facteur droit d'un mot de* $\mathcal{C}(A, c)$;

$b \doteq c$ *si, et seulement si, il existe* $A \in N$, $\Psi \in V^*$, $\Psi' \in V^*$, $\omega \in N$ *tels que* $A ::= \Psi\, b\, \omega\, c\, \Psi'$;

$b \lessdot c$ *si, et seulement si, il existe* $A \in N$, $\Psi \in V^*$, $\omega' \in N^*$, $\omega'' \in N^*$ *tels que* $A ::= \omega'\, c\, \Psi$ *et* $b\,\omega''$ *soit facteur droit d'un mot de* $\mathcal{C}(A, \Lambda)$.

Alors, pour que (n'', n') appartienne à $\Delta'_1(b, c)$ (resp. $\Delta''^1_1(b, c)$, $\Delta'_2(b, c)$, $\Delta''_2(b, c)$), il faut et il suffit que $b \gtrdot c$ (resp. $b \dot{\neq} c$, $b \doteq c$, $b \lessdot c$) et que (n'', n') appartienne à $\overline{\Delta}'_1$ (resp. $\overline{\Delta}''_1$, $\overline{\Delta}'_2$, $\overline{\Delta}''_2$).

Remplacer ainsi ${}^0\Delta'_1(b, c)$, ${}^0\Delta''^1_1(b, c)$, ${}^0\Delta'_2(b, c)$, ${}^0\Delta''_2(b, c)$ par des ensembles les contenant strictement peut cependant avoir pour inconvénient de remplacer un transducteur déterministe par un transducteur qui ne l'est plus, parce qu'il ne vérifie plus CT 1. Pour que les ensembles Δ'_1 (b, c), Δ''^1_1 (b, c), Δ'_2 (b, c), Δ''_2 (b, c) qui viennent d'être définis vérifient CT 1, il faut et il suffit que :

1) l'ensemble $\overline{\Delta}''_1$ (resp. $\overline{\Delta}''_2$) contienne au plus un couple (n'', n') *de premier entier* n'' *donné ou la relation $\not\doteq$ (resp. $\lessdot$) soit vide ;*

2) si deux des ensembles $\overline{\Delta}'_1$, $\overline{\Delta}''_1$, $\overline{\Delta}'_2$, $\overline{\Delta}''_2$ contienne respectivement deux couples ayant le même premier entier, les deux relations de priorité correspondantes soient incompatibles.

Les relations de priorité permettent aussi de donner une forme simple à la définition de l'incompatibilité d'un symbole non terminal avec un couple d'éléments de T' (7. 12. 4). Si la grammaire G donnée est réduite, pour tout A ϵ N il existe b'' ϵ T' et c'' ϵ T' tels que b'' soit le dernier élément de T' dans un mot de $\mathscr{C}$(A, c''), donc (7. 10. 3) si un mot ω sur N vérifie $|\omega| \geqslant 2$ et A :: = ω, il existe un entier n'' tel que (n'', $|\omega|$) appartienne à Δ''^1_1 (b'', c''), d'où à $\overline{\Delta}''_1$. De même, si A :: = ω c' Ψ' où $\omega \,\epsilon$ N*, c' ϵ T, Ψ' ϵ V*, il existe un entier n'' tel que (n'', $|\omega|$) appartienne à $\overline{\Delta}''_2$; si A ::= Ψ' c' ω, ($|\omega|$+ 1, $|\omega|$+ 1) appartient à $\overline{\Delta}'_1$. Par suite, *pour que* A ϵ N *soit incompatible avec un couple* (b, c) *d'éléments de* T', *il faut et il suffit que, quels que soient* b' *et* c' *dans* T, *les mots ω et ω' sur* N, *le mot Ψ sur* V:

- A ::= ω et $|\omega| \geqslant 2$ $\Longrightarrow$ non (b $\not\doteq$ c)
- A ::= ω c' ω' $\Longrightarrow$ non (b $\lessdot$ c' et c' $\gtrdot$ c)
- A ::= ω c' Ψ b' ω' $\Longrightarrow$ non (b $\lessdot$ c' et b' $\gtrdot$ c)

Dans le cas où la grammaire ne serait pas réduite, ces conditions resteraient suffisantes.

7. 14. Cas particuliers.

Nous donnerons trois exemples d'application de l'étude faite aux paragraphes 7. 9 à 7.13 à une classe de grammaires qui, pour simplifier, seront supposées réduites. Les deux premiers correspondent aux cas où les arborescences sous-réduites des arborescences orientées I dont une image (I, f) appartient à $\mathscr{C}$(G) satisfont aux hypothèses A et A_1, ou B de 2. 11.

7. 14. 1. Le premier cas est celui où, dans le second membre des règles de la grammaire G, deux occurences de symboles auxiliaires ne sont jamais consécutives : il s'agit du cas étudié par [23].

Pour tout A ϵ N, les mots de $\mathscr{C}$(A, Λ) se terminent tous par un élément de T' ; l'hypothèse CT est toujours satisfaite. Les ensembles Δ'_1 (b, c) ne peuvent contenir que (1, 1) et (2, 2) : nous prenons $\overline{\Delta}'_1$ = {(1, 1), (2, 2)} ; les Δ''^1_1 (b, c) sont vides : $\overline{\Delta}''_1 = \emptyset$; les Δ'_2 (b, c) ne peuvent contenir que (1, 1) et (2, 2) , les Δ''_2 (b, c) que (1, 0) et (2, 1) : $\overline{\Delta}'_2$ ={(1, 1), (2, 2)}, $\overline{\Delta}''_2$ = {(1, 0), (2, 1)} ; cg = 2. Les ensembles $\overline{\Delta}'_1$ et $\overline{\Delta}'_2$ contiennent tous les couples possibles pour cg = 2, $\overline{\Delta}''_2$ contient un couple (n'', n') commençant par tout n'' possible (n'' = 1, 2) et n' = n'' - 1 ; les transducteurs tr^4, $tr^{4'}$, tr^5 se simplifient notablement : en utilisant les relations de priorité, on évite toute référence explicite aux ensembles Δ_1 (b, c), Δ_2 (b, c) et même $\overline{\Delta}'_1$, $\overline{\Delta}''_1$, $\overline{\Delta}'_2$, $\overline{\Delta}''_2$.

La relation $\not\doteq$ est vide ; les trois autres relations de priorité (1) peuvent être définies de la manière suivante, en tenant compte de l'étude de $\mathscr{C}$(A, Λ) (7. 6. 2) et de la définition d'un couple vicinal (4. 3. 2) : pour b ϵ T, c ϵ T :

b $\gtrdot$ c si, et seulement si, il existe quatre symboles auxiliaires A, A', B, D et trois mots φ, φ', Ψ sur V tels que B ::= φ A' c φ', A est finale de A', A ::= Ψ b D ou A ::= Ψ b ;

b $\doteq$ c si, et seulement si, il existe deux symboles auxiliaires A, D et deux mots Ψ et Ψ' sur V tels que A ::= Ψ b D c Ψ' ou A ::= Ψ b c Ψ' ;

b $\lessdot$ c si, et seulement si, il existe quatre symboles auxiliaires A, A', B, D et trois mots φ, φ', Ψ sur V tels que B ::= φ b A' φ', A est initiale de A', A ::= D c Ψ ou A ::= c Ψ ;

b $\gtrdot$ σ si, et seulement si, il existe deux symboles auxiliaires A, D et un mot Ψ sur V tels que A soit finale d'un axiome, A ::= Ψ b D ou A ::= Ψ b ;

σ $\lessdot$ c si, et seulement si, il existe deux symboles auxiliaires A, D et un mot Ψ sur V tels que A soit initiale d'un axiome, A ::= D c Ψ ou A ::= c Ψ.

Le transducteur tr^4 (ou $tr^{4'}$) relatif à la grammaire G_3 obtenue en faisant subir successivement à G les transformations étudiées en 7. 10. 2 et 7.11 est déterministe si, et seulement si, les relations de priorité sont deux à deux incompatibles. La méthode a été mise en œuvre par [16].

Il est particulièrement facile ici de voir si une application r est séparante. La condition d'incompatibilité d'un symbole non terminal avec un couple d'éléments de T', donnée en 7. 13, se simplifie puisqu'il n'existe aucune règle A ::= ω où $\omega \,\epsilon$ N*. Surtout, les cas où (b, c) est un r-encadrement de A' sont ici les suivants : C étant un symbole auxiliaire tel que C $\xrightarrow{*}$ A', r(C) $\neq$ r(A') ou C = A', il existe des symboles auxiliaires B, B', D et des mots φ, φ', Ψ, Ψ_1 sur V tels que :

- soit B ::= Ψ_1 b C, D ::= φ B' c φ', B est finale de B' (car si c ϵ T, pour que $\mathscr{C}$(B, c) $\neq \emptyset$, il faut et il suffit que le couple B, c soit vicinal) ;
- soit B ::= Ψ_1 b C, c = σ, B est finale d'un axiome ;
- soit B ::= C c Ψ, D ::= φb B' φ' , B est initiale de B' (d'après l'étude de $\mathscr{C}$(B, Λ) en 7. 6. 2) ;
- soit B ::= C c Ψ, b = σ, B est initiale d'un axiome ;
- soit B ::= Ψ_1 b C c Ψ.

(1) $\gtrdot$, $\doteq$, $\lessdot$ sont des notations de [23].

Exemple (cf. instructions d'affectation et expressions arithmétiques Algol) :

N={A,E,P,F,B,G} ; T={:=,+,×,(,),0,1,2,3,4,5,6,7,8,9,α,b} ; N'={A};
A ::= G := E   E ::= E + P | P   P ::= P × F | F   F ::= G | B | (E)
G ::= G α | G b | α | b   B ::= B0 | B1 | ... | B9 | 0 | 1 | ... | 9

Matrice des relations de priorité : elles sont deux à deux incompatibles.

	:=	+	×	(	)	α, b	0 ⋮ 9	σ
:=		$\lessdot$	$\lessdot$	$\lessdot$		$\lessdot$	$\lessdot$	$\gtrdot$
+		$\gtrdot$	$\lessdot$	$\lessdot$	$\gtrdot$	$\lessdot$	$\lessdot$	$\gtrdot$
×		$\gtrdot$	$\gtrdot$	$\lessdot$	$\gtrdot$	$\lessdot$	$\lessdot$	$\gtrdot$
(		$\lessdot$	$\lessdot$	$\lessdot$	$\doteq$	$\lessdot$	$\lessdot$	
)		$\gtrdot$	$\gtrdot$		$\gtrdot$			$\gtrdot$
α, b	$\gtrdot$	$\gtrdot$	$\gtrdot$		$\gtrdot$	$\gtrdot$		$\gtrdot$
0...9		$\gtrdot$	$\gtrdot$		$\gtrdot$		$\gtrdot$	$\gtrdot$
σ	$\lessdot$					$\lessdot$		

Définissons τ : $\tau(A) = $ <i>, $\tau(E) = \tau(P) = \tau(F) = \tau(B) = $ <e>, $\tau(G) = $ <v> .
N_5 = { <i>, <e>, <v> } ; N'_5 = { <i> } ;
<i> $::=_5$ <v> := <e>
<e> $::=_5$ <e> + <e> | <e> + <v> | <v> + <e> | <v> + <v> | <e> × <e> |
<e> × <v> | <v> × <e> | <v> × <v> | (<e>) | (<v>) |
<e> 0 | <e> 1 | ... | <e> 9 | 0 | 1 | ... | 9
<v> $::=_5$ <v> α | <v> b | α | b.

τ-encadrements de E : := σ ; := + ; (+ ; () ;
de P : + σ ; + + ; +) ; + × ; := × ; (× ;
de F ; × σ ; × + ; ×) ; × × ;
de B : × c ; + c ; := c ; (c : c désigne l'un des chiffres 0, 1, ..., 9.

E est incompatible avec tout τ-encadrement de P car ni + $\lessdot$ +, ni + $\gtrdot$ × ne sont vrais, avec tout τ-encadrement de F car × $\lessdot$ + est faux, avec tout τ-encadrement de B car + $\gtrdot$ c est faux. P est incompatible avec tout τ-encadrement de F car × $\lessdot$ × est faux, avec tout τ-encadrement de B car × $\gtrdot$ c est faux. Enfin F est incompatible avec tout τ-encadrement de B car) $\gtrdot$ c est faux. L'application τ est séparante. Le générateur gn^5 permet l'analyse.

Envisageons la génération d'une pile U par gn^4, $gn^{4'}$ ou gn^5. Soit i + 1 une entrée dans U de $c \in T$; d'après C 8 et avec ses notations, il existe un entier n' tel que $(| b_i \omega_i |, n') \in \Delta_2 (b_i, c)$: $b_i \doteq c$ ou $b_i \lessdot c$. Le premier $j > i$ tel

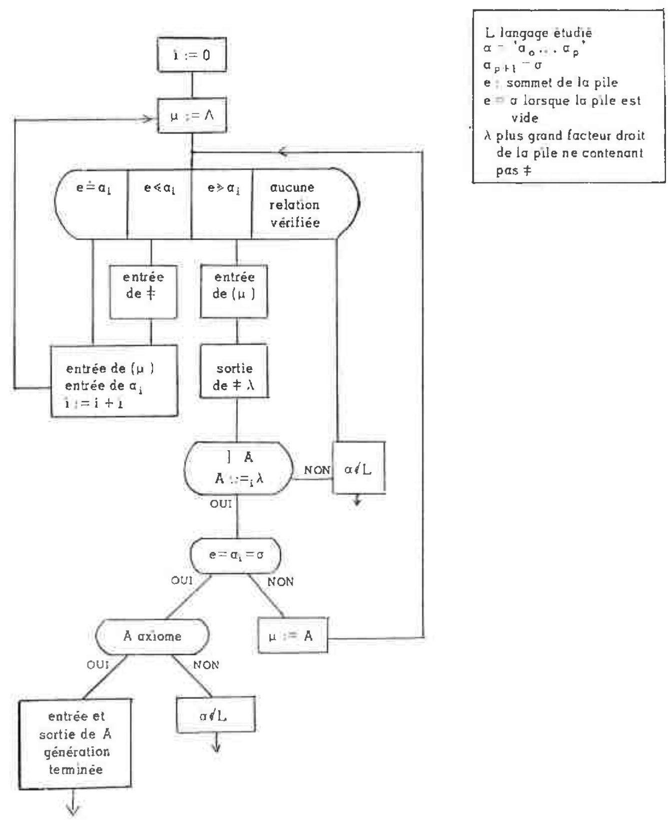


Figure 7. 7

que $u_j = u_i$ est une sortie de c (théorème 1. 2) ; si cette sortie est suivie, après des sorties de symboles auxiliaires, d'une sortie de b_i, $b_i \doteq c$; sinon, $b_i \lessdot c$ et j est suivi de la sortie de l'élément ω_i si $\omega_i \neq \Lambda$. Si les relations $\doteq$ et $\lessdot$ sont incompatibles, on peut déterminer comme en 7. 5. 3, dès l'entrée i + 1 de c, dans lequel de ces deux cas on se trouve et, en s'inspirant de l'algorithme de 2. 11. 2, limiter les sorties consécutives en plaçant dans la pile des séparateurs $\ddagger$. Nous nous contenterons de schématiser l'algorithme ainsi obtenu, en supposant les trois relations de priorité deux à deux incompatibles (figure 7. 7). On pourrait le justifier directement grâce à l'étude du paragraphe 2. 11. 2. Il s'applique avec les relations de priorité définies ci-dessus, pour relation de production $::=_i$ et pour ensemble des axiomes soit ceux de la grammaire G_3 obtenue en faisant subir successivement à G les transformations de 7. 10. 2 et 7. 11, soit ceux d'une grammaire G_5 définie par une application r séparante (7. 12).

7. 14. 2. Le second cas particulier que nous allons étudier est celui où tout second membre d'une règle de G, qui n'est pas réduit à un symbole auxiliaire, commence par un symbole terminal. Nous supposons de plus que la condition CT est satisfaite, ce qui est toujours le cas pour une grammaire non généralisée de ce type.

On peut alors définir pour chaque grammaire G un entier cg et prendre pour $\overline{\Delta}'_1$ et $\overline{\Delta}'_2$ l'ensemble des couples (n', n') tels que $1 \leqslant n' \leqslant cg$; les ${}^0\Delta''_1$ (b, c) sont vides : $\overline{\Delta}''_1 = \emptyset$; les ${}^0\Delta''_2$ (b, c) sont formés de couples (n'', 0) : prenons pour $\overline{\Delta}''_2$ l'ensemble des couples (n'', 0) tels que $1 \leqslant n'' \leqslant cg$. Ici encore, les transducteurs tr^4, $tr^{4'}$, tr^5 peuvent être définis grâce aux relations de priorité, sans faire appel à ces ensembles. La condition CT 1 est vérifiée si, et seulement si, les relations de priorité sont deux à deux incompatibles.

La relation $\ddagger$ est vide et les trois autres relations de priorité sont définies par : pour $b \in T$, $c \in T$:
$b \gtrdot c$ si, et seulement si, il existe $Z \in V$, trois symboles auxiliaires A, A', B, trois mots φ, φ', Ψ sur V et un mot ω sur N tels que $B ::= \varphi A' Z \varphi'$, A soit finale de A', $A ::= \Psi\, b\, \omega$, c soit initiale de Z (pour que c soit initiale de Z, il faut et il suffit qu'un $Z' \in V$ et un mot Ψ' sur V vérifient $Z \xrightarrow{*} Z'$, $Z' = c$ ou $Z' ::= c\Psi'$) ;
$b \doteq c$ si, et seulement si, il existe un symbole auxiliaire A, deux mots Ψ et Ψ' sur V et un mot ω sur N tels que $A ::= \Psi\, b\, \omega\, c\, \Psi'$;
$b \lessdot c$ si, et seulement si, il existe trois symboles auxiliaires A, A', B, trois mots φ, φ', Ψ sur V et un mot ω sur N tels que $B ::= \varphi b \omega A' \varphi'$, $A' \xrightarrow{*} A$ et $A ::= c\Psi$;
$b \gtrdot \sigma$ si, et seulement si, il existe un symbole auxiliaire A, un mot Ψ sur V et un mot ω sur N tels que A soit finale d'un axiome et $A ::= \Psi\, b\, \omega$;

$\sigma \lessdot c$ si, et seulement si, il existe un symbole auxiliaire A et un mot Ψ sur V tels que A dérive d'un axiome et $A ::= c\,\Psi$.

Lorsque ces relations ne sont pas incompatibles, on peut espérer obtenir un transducteur tr^4, $tr^{4'}$ ou tr^5 déterministe en utilisant directement les ensembles ${}^0\Delta'_1$ (b, c), ${}^0\Delta'_2$ (b, c), ${}^0\Delta''_2$ (b, c). Un cas particulièrement commode est celui où ceux de ces ensembles qui ne sont pas vides contiennent un seul élément (n'_1 (b, c), n'_1 (b, c)), (n'_2 (b, c), n'_2 (b, c)), (n'' (b, c), 0). Pour que la condition CT 1 soit satisfaite, il faut et il suffit que : si $b \gtrdot c$ et $b \doteq c$, $n'_1 (b, c) \neq n'_2 (b, c)$; si $b \gtrdot c$ et $b \lessdot c$, $n'_1 (b, c) \neq n'' (b, c)$; si $b \doteq c$ et $b \lessdot c$, $n'_2 (b, c) \neq n'' (b, c)$.

Dans tous les cas, on peut, comme au paragraphe précédent, en s'inspirant de 2. 11. 3, limiter les sorties consécutives grâce à un séparateur $\ddagger$: si i+1 est une entree de c dans une pile U engendrée et si $(|b_i\, \omega_i|, 0) \in \Delta''_2 (b, c)$, on fait entrer $\ddagger$ avant c ; toute suite de sorties consécutives se termine par une sortie de

Un couple (b, c) est ici un r-encadrement de A' lorsqu'il existe des symboles auxiliaires B, C, D, B', un $Z \in V$, des mots Ψ, Ψ', φ sur V tels que $C \xrightarrow{*} A'$, $r(C) \neq r(A')$ ou $C = A'$, b soit le dernier élément de T dans $\Psi \neq \Lambda$ et que l'une des conditions suivantes soit remplie :
- $B ::= \Psi\, C$, $D ::= \varphi B' Z \Psi'$, B est finale de B' et c initiale de Z ;
- $B ::= \Psi\, C$, $c = \sigma$, B est finale d'un axiome ;
- $B ::= \Psi\, C\, Z\, \Psi'$, c est initiale de Z.

Dans le cas particulier où le second membre de toute règle de G ne contient aucune occurence de symbole terminal autre que la première les ensembles ${}^0\Delta'_2$ (b, c) sont vides, ainsi que la relation $\doteq$; les transformations des paragraphes 7. 10. 2 et 7. 10. 4 remplacent G par une grammaire normale au sens de [30].

On peut généraliser le cas étudié dans ce paragraphe au cas où tout second membre de regle non réduit à un symbole auxiliaire a un facteur gauche formé de r (fixe) symboles auxiliaires suivis d'un symbole terminal : les couples des ensembles ${}^0\Delta''_2$ (b, c) sont tous de la forme (n'', r).

7. 14. 3. Nous étendons ici le cas etudié en 7. 14. 1 en nous donnant un entier $r \geqslant 2$ et en envisageant les grammaires dont toute règle est de l'un des types : $A ::= A'$, $A ::= \omega$, $A ::= \varphi$ où $A \in N$, $A' \in N$, $\omega \in N^*$, $|\omega| = r$, $\varphi \in V^*$, φ ne contient pas deux occurences consécutives de symboles auxiliaires. Nous supposons de plus que ces grammaires vérifient CT.

Pour chaque grammaire, étudions les ensembles :

- ${}^0\Delta'_1(b, c)$: ses seuls éléments possibles sont (1, 1) et (2, 2) ;
- ${}^0\Delta''_1(b, c)$: ses éléments sont de la forme (n'', r) avec $n'' > r \geqslant 2$;
- ${}^0\Delta'_2(b, c)$: ses seuls éléments possibles sont (1, 1) et (2, 2) ;
- ${}^0\Delta''_2(b, c)$: ses éléments sont de la forme $(n'', 1)$ ou $(n'', 0)$.

Soit cg le plus grand des premiers entiers des couples des ${}^0\Delta''^{\,1}_1(b, c)$ et ${}^0\Delta''_2(b, c)$: $cg > r \geqslant 2$ s'il existe effectivement une règle $A ::= \omega$, ce que nous supposerons. On peut prendre $\overline{\Delta}'_1 = \overline{\Delta}'_2 = \{(1, 1), (2, 2)\}$, et pour $\overline{\Delta}''_1$ l'ensemble des couples (n'', r) pour lesquels $r < n'' \leqslant cg$. Plutôt que de définir de même un ensemble $\overline{\Delta}''_2$, nous définirons $\overline{\Delta}''_{20}$, ensemble des $(n'', 0)$ tels que $1 \leqslant n'' \leqslant cg$, et $\overline{\Delta}''_{21}$ ensemble des $(n'', 1)$ tels que $2 \leqslant n'' \leqslant cg$. Nous remplacerons la relation de priorité $\lessdot$ par deux relations dont elle est réunion :

$b <_0 c$ si, et seulement si, ${}^0\Delta''_2(b, c)$ contient un couple de second entier 0 ;
$b <_1 c$ si, et seulement si, ${}^0\Delta''_2(b, c)$ contient un couple de second entier 1.
Autrement dit :
$b <_0 c$ si, et seulement si, il existe $A \in N$, $\Psi \in V^*$, $\omega'' \in N^*$ tels que $A ::= c\,\Psi$ et $b\,\omega''$ soit facteur droit d'un mot de $\mathscr{C}(A, \Lambda)$;
$b <_1 c$ si, et seulement si, il existe $A \in N$, $B \in N$, $\Psi \in V^*$, $\omega'' \in N^*$ tels que $A ::= B\,c\,\Psi$ et $b\,\omega''$ soit facteur droit d'un mot de $\mathscr{C}(A, \Lambda)$.
Choisissons $\Delta'_1(b, c)$, $\Delta''^{\,1}_1(b, c)$, $\Delta'_2(b, c)$ comme en 7. 13 et : $\Delta''_{20}(b, c) = \overline{\Delta}''_{20}$ si $b <_0 c$, $\Delta''_{20}(b, c) = \emptyset$ sinon ; $\Delta''_{21}(b, c) = \overline{\Delta}''_{21}$ si $b <_1 c$, $\Delta''_{21}(b, c) = \emptyset$ sinon ; $\Delta''_2(b, c) = \Delta''_{20}(b, c) \cup \Delta''_{21}(b, c)$. Pour que la condition CT 1soit satisfaite par les ensembles $\Delta^1_1(b, c)$ et $\Delta_2(b, c)$, il faut et il suffit que les relations $\gtrdot$, $\doteq$, $<_0$, $<_1$ soient deux à deux incompatibles et que $\not\doteq$ soit incompatible avec $<_0$ et avec $<_1$.

BIBLIOGRAPHIE

[1] J.W. BACKUS
Report on the algorithmic language Algol 60 - Nümerische Mathematik 2 (1960), 106-136.

[2] Y. BAR-HILLEL, M. PERLES, E. SHAMIR
On formal properties of simple phrase structure grammars - Zeitschrift für Phonetik, Schprachwiss, Kommunik. 14 (1961) 143-172.

[3] C. BERGE
Théorie des graphes et ses applications - Dunod, Paris (1958).

[4] L. BOLLIET, N. GASTINEL, P.J. LAURENT
Un nouveau langage scientifique, Algol - Hermann, Paris (1964).

[5] N. BOURBAKI
Eléments de mathématique. Théorie des ensembles, Chapitre 3 - Hermann, Paris (1963).

[6] N. BOURBAKI
Eléments de mathématique. Algèbre, Chapitre 1 - Hermann, Paris (1964).

[7] M. BRASSEUR, J. COHEN
Algorithmes d'analyse syntaxique pour langages "context free" - Institut de Mathématiques Appliquées, Grenoble (1964).

[8] R.A. BROOKER, D. MORRIS
A general translation program for phrase structure languages - Journal of the ACM 9 (1962), 1-10.

[9] W. H. BURGE
Sorting, trees, and measures of order - Information and Control 1 (1958), 181-197.

[10] T. CHEATHAM, K. SATTLEY
Syntax directed compiling - Proceedings Spring Joint Computer Conference 25 (1964), 31-57.

[11] C. CHEVALLEY
Fundamental concepts of algebra - Academic Press, New-York (1956).

[12] N. CHOMSKY
On certain formal properties of grammars - Information and Control 2 (1959), 137-167.

[13] N. CHOMSKY
Formal properties of grammars - Handbook of mathematical psychology (Luce, Bush, Galanter editors), Wiley and Sons, New-York (1963), 323-418.

[14] N. CHOMSKY et G.A. MILLER
Finite state languages - Information and Control (1958), 91-112.

[15] N. CHOMSKY et M.P. SCHÜTZENBERGER
The algebric theory of context-free languages - Computer programming and formal systems (Braffort, Hirschberg editors) - North-Holland, Amsterdam (1963), 118-161.

[16] A. COLMERAUER
Recherche d'erreurs syntaxiques - Institut de Mathématiques Appliquées, Grenoble (1965).

[17] M. DAVIS
Computability and unsolvability - Mc Graw-Hill, New-York (1958).

[18] P. DUBREIL
Algèbre, tome 1 - Gauthiers-Villars, Paris (1954).

[19] M.L. DUBREIL-JACOTIN, L. LESIEUR, R. CROISOT
Leçons sur la théorie des treillis, des structures algéques ordonnées et des treillis géométriques - Gauthier - Villars, Paris (1953).

[20] J. EICKEL, M. PAUL
The parsing and ambiguity problem for Chomsky-languages - I.F.I.P. Working Conference on formal languages description languages Vienne (1964).

[21] J. EICKEL, M. PAUL, F.L. BAUER, K. SAMELSON
A syntax controlled generator of formal language processors - Communications of the ACM 6 (1963), 451-455.

[22] A. EMOND
Application de la notion de pile à des problèmes portant sur les chemins des graphes - Thèse de troisième cycle, Faculté des Sciences de Nancy (1965).

[23] R.W. FLOYD
Syntactic analysis and operator precedence - Journal of the ACM 10 (1963), 316-333.

[24] R.W. FLOYD
Bounded context syntactic analysis - Communications of the ACM (1964), 62-67.

[25] F. GENUYS
Commentaires sur le langage Algol - Chiffres 5 (1962), 29-53.

[26] S. GORN
Processors for infinite codes of the Shannon-Fano type. Proceedings of the symposium on mathematical theory of automata - Polytechnic institute of Brooklyn (1962), 223-240.

[27] S. GORN
An axiomatic approach to prefix languages - Symbolic Languages in data processing - Gordon and Breach, New-York - Londres (1962), 1-21.

[28] A.A. GRAU
Recursive processes and Algol translation - Communications of the ACM 4 (1961), 10-15.

[29] S. GREIBACH
Formal parsing systems - Communications of the ACM 7 (1964), 499-504.

[30] S. GREIBACH
A new normal-form theorem for context-free phrase structure grammars - Journal of the ACM 12 (1965), 42-52.

[31] T.N. HIBBARD
Some combinatorial properties of certain trees with applications to searching and sorting - Journal of the ACM 9 (1962), 13-28.

[32] P. HILDEBRANDT, H. ISBITZ
Radix exchange, an internal sorting procedure for digital computers - Journal of the ACM 6 (1959), 156-163.

[33] D.A HUFFMAN
A method for the construction of minimum redudancy codes - Proc. I.R.E. 40 (1952), 1098-1101.

[34] P.Z. INGERMAN
A translation technique for languages whose syntax is expressible in extended Backus normal form - Symbolic languages in data processing - Gordon and Breach. New-York (1962), 23-64.

[35] E.T. IRONS
A syntax directed compiler for ALGOL 60 - Communications of the ACM 4 (1961), 51-55.

[36] E.T. IRONS
An error-correcting parse algorithm - Communications of the ACM 6 (1963), 669-673.

[37] K.E. IVERSON
A programming language - Wiley and Sons, New-York (1962).

[38] J.B. JOHNSTON
A class of unambiguous computer languages - Communications of the ACM 8 (1965), 147-149.

[39] C. JORDAN
Calculus of finite differences - Chelsea Publishing Company, New-York (1950).

[40] S.C. KLEENE
Representation of events in nerve sets and finite automata - Automata studies (Shannon, Mac Carthy) - Princeton University Press (1956), 3-41.

[41] S. KUNO, A. ŒTTINGER
Multiple-path syntactic analyzer - Information Processing 62 (Popplewell editor) - North Holland, Amsterdam (1963), 306-312.

[42] M. LIETZKE
A method of syntax checking Algol 60 - Communications of the ACM 7 (1964), 475-478.

[43] P. NAUR (editor)
Revised report on the algorithmic language Algol 60 - Communications of the ACM 6 (1963), 1-17.

[44] A. NEWELL, J.C. SHAW
Programming the logic theory machine - Proc. Western Joint Computer Conference (1957), 230-240.

[45] M. NIVAT, L. NOLIN
Sur un procédé de définition de la syntaxe d'Algol - Institut Blaise Pascal, Paris (1963).

[46] C. PAIR
Séminaire de programmation du Centre de Calcul Automatique de l'Université de Nancy (1964).

[47] M. PAUL
Zur Struktur formaler Sprachen - Université J. Gutenberg, Mayence (1962).

[48] C. PICARD
Théorie des questionnaires - Institut Blaise Pascal, Paris (1963).

[49] M.O. RABIN, D. SCOTT
Finite automata and their decision problems - IBM Journal of Research and Development 3 (1959), 114-125.

[50] R. ROMAC
Etude des méthodes de tri - Thèse de troisième cycle, Faculté des Sciences de Nancy (1964).

[51] G. SALTON
Manipulation of trees in information retrieval - Communications of the ACM 5 (1962), 103-114.

[52] K. SAMELSON, F.L. BAUER
Sequentielle Formelübersetzung - Elektron. Rechenanlagen 1 (1959), 176-182.

[53] M.P. SCHÜTZENBERGER
A remark on finite transducers - Information and Control 4 (1961), 185-196.

[54] M.P. SCHÜTZENBERGER
On context-free languages and push-down automata - Information and Control 6 (1963), 246-264.

[55] A.K.SCIDMORE, B.L. WEINBERG
Storage and search properties of a tree-organised memory system - Communications of the ACM 6 (1963), 28-31.

[56] E.H. SUSSENGUTH
Use of tree structures for processing files - Communications of the ACM 6 (1963), 272-279.

[57] S.H. UNGER
On syntax directed translators - R.C.A. Laboratories (1963).

[58] P.F. WINDLEY
Trees, forests and rearranging - Computer Journal 3 (1960), 84-88.

INDEX DES NOTATIONS

(les munéros renvoient aux pages)

INDEX DES CONDITIONS

(les numéros renvoient aux pages)

INDEX TERMINOLOGIQUE

(les numéros renvoient aux pages)

DEUXIEME THESE

Propositions données par la Faculté

ALGEBRE DIFFERENTIELLE

Vu et approuvé,

Le Doyen de la Faculté
des Sciences de Nancy,

J. AUBRY

Vu et permis d'imprimer

Le Recteur,
Président du Conseil de
l'Université,

P. IMBS

Cet exemplaire est imprimé
à des fins particulières et ne
peut en aucun cas être vendu.

Achevé d'imprimer le 31 octobre 1966
à l'atelier de Documentation Universitaire
3, Rue Sellier, 3 NANCY

Dépot légal N° 219 Faculté des Sciences Editeur
4ème trimestre 1966