50376
1975
113

N° d'ordre : 534

50376
1975
113

THESE

présentée à

L'UNIVERSITE DES SCIENCES ET TECHNIQUES DE LILLE

pour obtenir le titre de

DOCTEUR de SPECIALITE

(INFORMATIQUE THEORIQUE)

par

Max DAUCHET

BIBLIOTHEQUE UNIVERSITAIRE
Section de SCIENCES
LILLE

TRANSDUCTIONS INVERSIBLES DE FORETS

Soutenue le 19 Juin 1975, devant la COMMISSION D'EXAMEN

Membres du Jury : MM.	P. BACCHUS	Président
	L. BOASSON	Examinateur
	C. CARREZ	Examinateur
	V. CORDONNIER	Examinateur
	M. NIVAT	Rapporteur

UNIVERSITE DES SCIENCES
ET TECHNIQUES DE LILLE

DOYENS HONORAIRES DE L'Ancienne Faculté des Sciences

MM. R. DEFRETIN, H. LEFEBVRE, M. PARREAU.

PROFESSEURS HONORAIRES Des Anciennes Facultés de Droit
Et Sciences Economiques, des Sciences et des Lettres

M. ARNOULT, Mme BEAUJEU, MM. BROCHARD, CHAPPELON, CHAUDRON, CORDONNIER, DEHEUVELS, DEHORS, DION, FAUVEL, FLEURY, P. GERMAIN, HEIM DE BALSAC, HOCQUETTE, KAMPE DE FERIET, KOUGANOFF, LAMOTTE, LASSERRE, LELONG, Mme LELONG, MM. LHOMME, LIEBAERT, MARTINOT-LAGARDE, MAZET, MICHEL, NORMANT, PEREZ, ROIG, ROSEAU, ROUBINE, ROUELLE, SAVART, WATERLOT, WIEMAN, ZAMANSKI.

PRESIDENT HONORAIRE DE L'UNIVERSITE DES
SCIENCES ET TECHNIQUES DE LILLE

R. DEFRETIN

PRESIDENT DE L'UNIVERSITE DES
SCIENCES ET TECHNIQUES DE LILLE

M. PARREAU.

PROFESSEURS TITULAIRES

M. BACCHUS Pierre	Astronomie
M. BEAUFILS Jean-Pierre	Chimie Physique
M. BECART Maurice	Physique Atomique et Moléculaire
M. BLAYS Pierre	Géographie
M. BONNEMAN Pierre	Chimie Appliquée
M. BONTE Antoine	Géologie Appliquée
M. BOUGHON Pierre	Algèbre
M. BOUISSET Simon	Physiologie Animale
M. BOURIQUET Robert	Biologie Végétale
M. CELET Paul	Géologie Générale
M. CONSTANT Eugène	Electronique
M. CORSIN Pierre	Paléontologie
M. DECUYPER Marcel	Géométrie
M. DELATTRE Charles	Géologie Générale
M. DELHAYE Michel	Chimie Physique
M. DERCOURT Michel	Géologie Générale
M. DURCHON Maurice	Biologie Expérimentale
M. FAURE Robert	Mécanique
M. FOURET René	Physique du Solide
M. GABILLARD Robert	Electronique
M. GLACET Charles	Chimie Organique
M. GONTIER Gérard	Mécanique
M. GRUSON Laurent	Algèbre
M. GUILLAUME Jean	Microbiologie
M. HEUBEL Joseph	Chimie Minérale
M. LANSRAUX Guy	Physique Atomique et Moléculaire
M. LEBRUN André	Electronique
M. LEHMANN Daniel	Géométrie
Mme LENOBLE Jacqueline	Physique Atomique et Moléculaire
M. LINDER Robert	Biologie et Physiologie Végétales
M. LOMBARD Jacques	Sociologie

.../...

M. LUCQUIN Michel	Chimie Physique
M. MAILLET Pierre	Sciences Economiques
M. MONTARIOL Frédéric	Chimie Appliquée
M. MONTREUIL Jean	Biochimie
M. PARREAU Michel	Analyse
M. POUZET Pierre	Analyse Numérique
M. PROUVOST Jean	Minéralogie
M. SCHILTZ René	Physique Atomique et Moléculaire
Mme SCHWARTZ Marie-Hélène	Géométrie
M. TILLIEU Jacques	Physique Théorique
M. TRIDOT Gabriel	Chimie Appliquée
M. VAILLANT Jean	Analyse
M. VIDAL Pierre	Automatique
M. VIVIER Emile	Biologie Cellulaire
M. WERTHEIMER Raymond	Physique Atomique et Moléculaire
M. ZEYTOUNIAN Radyadour	Mécanique

PROFESSEURS SANS CHAIRE

M. BELLET Jean	Physique Atomique et Moléculaire
M. BILLARD Jean	Physique du Solide
M. BODARD Marcel	Biologie Végétale
M. BOILLET Pierre	Physique Atomique et Moléculaire
M. BONNOT Ernest	Biologie Végétale
M. BRIDOUX Michel	Chimie Physique
M. CAPURON Alfred	Biologie Animale
M. DEPREZ Gilbert	Physique Théorique
M. DEVRAINNE Pierre	Chimie Minérale
M. GOUDMAND Pierre	Chimie Physique
M. GUILBAULT Pierre	Physiologie Animale
M. LABLACHE-COMBIER Alain	Chimie Organique
M. LACOSTE Louis	Biologie Végétale
Mme LEHMANN Josiane	Analyse
M. LOUCHEUX Claude	Chimie Physique
M. MAES Serge	Physique Atomique et Moléculaire
Melle MARQUET Simone	Probabilités
M. MIGEON Michel	Chimie Physique
M. MONTEL Marc	Physique du Solide
M. PANET Marius	Electrotechnique
M. RACZY Ladislas	Electronique
M. ROUSSEAU Jean-Paul	Physiologie Animale
M. SALMER Georges	Electronique
M. SEGUIER Guy	Electrotechnique

MAITRES DE CONFERENCES (et chargés d'Enseignement)

M. ADAM Michel	Sciences Economiques
M. ANDRE Charles	Sciences Economiques
M. ANGRAND Jean-Pierre	Géographie
M. ANTOINE Philippe	Analyse
M. BART André	Biologie Animale
M. BEGUIN Paul	Mécanique
M. BKOUCHE Rudolphe	Algèbre
M. BOILLY Bénoni	Biologie Animale
M. BONNEMAIN Jean-Louis	Biologie Végétale
M. BOSCQ Denis	Probabilités
M. BREZINSKI Claude	Analyse Numérique
M. BRUYELLE Pierre	Géographie
M. CARREZ Christian	Informatique
M. CORDONNIER Vincent	Informatique
M. CORTOIS Jean	Physique Nucléaire et Corpusculaire

.../...

M. COQUERY Jean-Marie	Psycho-Physiologie
M. COULON Jean	Electrotechnique
Mlle DACCHARI Monique	Géographie
M. DEBOURSE Jean-Pierre	Gestion des Entreprises
M. DEBRABANT Pierre	Géologie Appliquée
M. DHAINAUT André	Biologie Animale
M. DELAUNAY Jean-Claude	Sciences Economiques
M. DERIEUX Jean-Claude	Microbiologie
M. DOUKHAN Jean-Claude	Physique du Solide
M. DRIEUX Baudouin	Informatique
M. DUEE Gérard	Géologie Appliquée
M. DYMENT Arthur	Mécanique
M. ESCAIG Bertrand	Physique du Solide
Mme EVRARD Micheline	Chimie Appliquée
M. FONTAINE Jacques-Marie	Electronique
M. FOURNET Bernard	Biochimie
M. FROELICH Daniel	Chimie Physique
M. GAMBLIN André	Géographie
M. GOBLOT Rémi	Algèbre
M. GOSSELIN Gabriel	Sociologie
M. GRANELLE Jean Jacques	Sciences Economiques
M. GUIGOU Jean-Louis	Sciences Economiques
M. GUILLAUME Henri	Sciences Economiques
M. HECTOR Joseph	Géométrie
M. HERMAN Maurice	Physique Spatiale
M. JOURNEL Gérard	Physique Atomique et Moléculaire
Mlle KOSMANN Yvette	Géométrie
M. KREMBEL Jean	Biochimie
M. LANGRAND Claude	Probabilités
M. LAURENT François	Automatique
Mlle LEGRAND Denise	Algèbre
Mlle LEGRAND Solange	Algèbre
M. LENTACKER Firmin	Géographie
M. LEROY Jean-Marie	Chimie Appliquée
M. LEROY Yves	Electronique
M. LHENAFF René	Géographie
M. LOCQUENEUX Robert	Physique Théorique
M. LOUAGE Francis	Electronique
M. MAHIEU Jean-Marie	Physique Atomique et Moléculaire
Mme N'GUYEN VAN CHI Régine	Géographie
M. MAIZIERES Christian	Automatique
M. MALAUSSENA Jean-Louis	Sciences Economiques
M. MESSELYN Jean	Physique Atomique et Moléculaire
M. MONTUELLE Bernard	Biologie Appliquée
M. NICOLE Jacques	Chimie Appliquée
M. PAQUET Jacques	Géologie Générale
M. PARSY Fernand	Mécanique
M. PECQUE Marcel	Chimie Physique
M. PERROT Pierre	Chimie Appliquée
M. PERTUZON Emile	Physiologie Animale
M. PONSOLLE Louis	Chimie Physique
M. POVY Lucien	Automatique
M. ROGALSKI Marc	Analyse
M. ROY Jean-Claude	Psycho-Physiologie
M. SIMON Michel	Sociologie
M. SLIWA Henri	Chimie Organique
M. SOMME Jean	Géographie
Mlle SPIK Geneviève	Biochimie
M. STANKIEWICZ François	Sciences Economiques
M. THERY Pierre	Electronique
M. TOULOTTE Jean-Marc	Automatique
M. TREANTON Jean-René	Sociologie

.../...

M. VANDORPE Bernard	Chimie Minérale
M. VILLETTE Michel	Mécanique
M. WERNIER Georges	Informatique
M. WATERLOT Michel	Géologie Générale
M. YVON Jean-Pierre	Analyse Numérique
Mme ZINN JUSTIN Nicole	Algèbre

Je remercie Monsieur Bacchus, Professeur à l'Université des Sciences et Techniques de Lille I, qui me fait l'honneur de présider le Jury et qui m'a conseillé dans la rédaction de cette thèse.

Je remercie Monsieur Boasson, Maître de Conférences à l'Université d'Amiens, qui s'est intéressé à mon travail et a accepté de le juger.

Je remercie Monsieur Carrez et Monsieur Cordonnier, Maîtres de Conférences à l'Université des Sciences et Techniques de Lille I, qui ont bien voulu faire partie du Jury.

Je remercie particulièrement Monsieur Nivat, Professeur à l'Université de Paris VII, qui me fait l'important honneur d'orienter ma recherche et de rapporter cette thèse.

Je remercie André Arnold : l'essentiel de ce travail a été réalisé dans le cadre de notre collaboration et sera par ailleurs publié en commun.

Françoise Desquiens et Michèle Driessens ont agrémenté de leur charme la mise en page qu'elles ont effectuée avec compréhension, rapidité et gentillesse. Monsieur et Madame Debock ont fait preuve des mêmes qualités dans la réalisation matérielle de cette thèse. Je les remercie tous.

TABLE DES MATIÈRES

Les deux parties sont numérotées séparément

INTRODUCTION

PREMIÈRE PARTIE

PRESENTATION DES RESULTATS
ET DISCUSSION

page

CHAPITRE I : RAPPELS I.1

1.DEFINITIONS ET RAPPELS

A - Définitions élémentaires I.1
B - Grammaires d'arbres I.15
C - Reconnaisseurs d'états finis I.21
D - Rappel de résultats I.23

2. LES TRANSDUCTEURS

A - Définitions I.24
B - Exemples complémentaires I.30
C - Propriétés I.31

CHAPITRE II : TRANSDUCTEURS GÉNÉRALISÉS, BIMORPHISMES ET BITRANSDUCTIONS I.33

1. DEFINITIONS I.33

2. RESULTATS PRINCIPAUX I.35

CHAPITRE III : COMMENTAIRES ET ILLUSTRATIONS DES RÉSULTATS OBTENUS I.37

CHAPITRE IV : EXAMEN DES DIFFERENTES NOTIONS DE TRANSDUCTIONS AVANTAGES DES BITRANSDUCTIONS I.48

1. CAS DES TRANSDUCTIONS NON GENERALISEES I.48

2. AUTRES NOTIONS DE BIMORPHISMES I.49

A - Restriction des transductions inversibles I.49

B - Généralisations des transductions inversibles I.50

CONCLUSION I.51

DEUXIÈME PARTIE

DEMONSTRATION

CHAPITRE I : CARACTÉRISATION DES TRANSDUCTEURS INVERSIBLES PAR BIMORPHISMES II.1

CHAPITRE II : CARACTÉRISATION DES BITRANSDUCTIONS DU TYPE ASCENDANT II.5

CHAPITRE III : UN RÉSULTAT INTERMÉDIAIRE SUR LES BITRANSDUCTIONS QUASI-GÉNÉRALISÉES II.17

CHAPITRE IV : CARACTÉRISATION DES BITRANSDUCTIONS DU TYPE "DESCENDANT-DESCENDANT" II.40

CHAPITRE V : RÉSULTATS COMPLÉMENTAIRES

1. BITRANSDUCTIONS ET FORETS REGULIERES II.56

2. BITRANSDUCTIONS ET FORETS ALGEBRIQUES II.58

3. PROPRIETES ENSEMBLISTES DES BITRANSDUCTIONS II.59

4. UN CAS PARTICULIER II.60

5. UNE HIERARCHIE INFINIE II.62

BIBLIOGRAPHIE

INTRODUCTION

Dans cette introduction, nous situons d'abord notre sujet. Nous présentons ensuite les deux parties de notre travail. Nous concluons par les perspectives ouvertes par nos résultats.

Des études linguistiques (N. Chomsky) puis informatiques (langages évolués, automates) sont à l'origine du développement de la théorie des langages formels. Deux classes de langages sont dès le début apparues comme très importantes ([CH]) : la classe des langages algébriques (solutions de systèmes d'équations algébriques) et une sous-classe, celle des langages reconnaissables (reconnus par un automate d'états finis).

Les transductions rationnelles, essentiellement introduites et étudiées par M. NIVAT [NI 1], jouent un rôle essentiel en théorie des langages, où elles sont d'un usage constant. Il est d'ailleurs remarquable que ces transductions permettent de clarifier (L. BOASSON, [BO]) la théorie des AFL développée indépendamment par S. GINSBURG et S. GREIBACH. Les transductions rationnelles présentent l'intérêt d'être stables par composition et inversion, de conserver la reconnaissabilité et l'algébricité et de se caractériser en termes de bimorphismes, ce qui les rend très manipulables.

Dans un monoïde, chaque lettre a zéro ou un successeur immédiat, selon qu'elle est ou non en fin de mot. Il est naturel de généraliser en attribuant à chaque lettre un nombre quelconque de successeurs immédiats, ce nombre étant appelé arité de la lettre. La notion de monoïde libre se généralise alors en celle de magma - ou algèbre - libre. Nous ne considérons que des magmas libres de type fini, c'est-à-dire engendrés par des alphabets finis. Un élément d'un magma est un arbre, une partie est une forêt.

La classe des forêts reconnaissables se définit alors aisément et a déjà été beaucoup étudiée (W. BRAINERD [BR 1 et 2], J. DONER [DO], S. EILENBERG et J. WRIGHT [EI], MEZEI et WRIGHT [MW], W. ROUNDS [RO 1], J. THATCHER [TH 1 et 2] ...). La classe des langages algébriques est l'ensemble des feuillages des forêts reconnaissables, ce qui fait de ces dernières un moyen d'étude des

langages algébriques ([DO],[MW],[RO],...). Les reconnaisseurs généralisés d'états finis sont également appliqués à des problèmes de décision du second ordre (J. BUCHI et C. ELGOT [BU], [DO], [TH 2]).

L'algébricité se généralise aussi naturellement et présente un grand intérêt en Informatique Théorique : un schéma de programme récursif est un système d'équations algébriques sur un magma ([NI 2]).

Ainsi une généralisation mathématique a conduit à des notions d'un grand intérêt pratique. Il est dès lors tentant de développer une théorie des forêts semblable à celle des langages. Malheureusement, de grosses difficultés apparaissent lorsqu'on introduit les transducteurs d'arbres. THATCHER a fait récemment dans [AH] un tour d'horizon de nos connaissances dans le domaine : il met en évidence que celles-ci sont maigres et peu profondes. Une seule classe de transducteurs d'arbres a pratiquement été étudiée : celle des transducteurs d'états finis. Nous parlerons donc simplement de transducteurs pour désigner ceux-ci. Les transducteurs sont obtenus, comme dans le cas des langages, en "mettant une sortie" aux reconnaisseurs. Historiquement, cette démarche a pour support pratique des préoccupations linguistiques (Rounds) et informatiques (compilation, "syntax directed transducers" par BJORNERD [BJ], MARTIN-VERE [MA]).

Les transducteurs d'arbres ont d'abord été étudiés par ENGELFRIET, ROUNDS [RO 2] et THATCHER [TH 3 et 4]. Rappelons qu'à tout transducteur T, on associe sa relation (c'est-à-dire son graphe) notée $\hat{T}$ et appelée transduction associée. Deux transducteurs sont équivalents ssi ils ont même transduction associée. On montre que les forêts reconnues "du sommet vers les feuilles" (par reconnaisseurs descendants) et les forêts reconnues "des feuilles vers le sommet" (par reconnaisseurs ascendants) sont les mêmes. Par contre, la classe des transductions ascendantes et celle des transductions descendantes sont incomparables. Les transductions ne conservent en général ni la reconnaissabilité, ni l'algébricité (A. ARNOLD et M. DAUCHET [AR 3]) ; elles ne sont stables ni par composition ni par inversion. En transformant la classe des forêts reconnaissables par transductions itérées, des hiérarchies infinies de classes de forêts sont conjecturées (B. BAKER [BA], OGDEN et ROUNDS [OC]). La générali-

sation des langages aux forêts a conduit à des transductions dont les propriétés n'ont rien de commun avec celles des transductions rationnelles de langages. De plus, la forme des règles des transducteurs d'arbres rend ceux-ci difficiles à manipuler. Le fait que la quasi-totalité des résultats un peu profonds ne soit pas conjecturée est là pour le rappeler.

Dans ces conditions, il nous paraissait téméraire de vouloir définir d'emblée une "bonne" classe de transductions de forêts, ayant les "bonnes" propriétés des transductions rationnelles. Notre but initial était d'étudier la seule stabilité par inversion. Nous allons voir comment nos résultats ont largement dépassé nos espérances.

Notre travail comprend deux parties. Dans la première partie, nous présentons, situons et commentons nos résultats. Dans la deuxième partie, nous établissons les démonstrations. Les résultats sont faciles à énoncer mais les démonstrations en sont toujours techniques, souvent difficiles. Aussi, souhaitons-nous cette présentation plus agréable en évitant de disperser l'essentiel.

Dans la première partie, le premier chapitre est consacré aux rappels. Il débute au niveau le plus élémentaire et est illustré de nombreux exemples afin de familiariser avec le sujet le lecteur non spécialiste.

Dans le chapitre II, nous définissons de nouvelles notions puis énonçons nos principaux résultats.

Nous commençons par définir les notions de transducteur généralisé (définition 9), bimorphisme (définition 12), transducteur inversible (définition 13) et bitransduction (définition 14).

Dans les transducteurs d'arbres classiques, les arbres indexés apparaissant en partie gauche des règles sont réduits à une lettre (d'arité quelconque). Nous généralisons en autorisant à cette place des arbres indexés quelconques mais contenant au moins une lettre, de façon à rendre la forme des règles plus symétrique.

Nous notons TGA (respectivement TGD) la classe des transducteurs généralisés ascendants (respectivement descendants).

Nos bimorphismes s'inspirent de ceux caractérisant les transductions rationnelles de langages et sont comme eux d'un usage facile.

Nous appelons transducteur inversible (TI) tout transducteur généralisé "inversible par règles" c'est-à-dire tel qu'en "inversant le sens des flèches de ses règles" on définisse un autre transducteur (ce qui n'est pas possible en général). La notion de bitransduction formalise "l'inversibilité au sens des relations" comme suit : une relation entre deux forêts est une bitransduction de type TG_{XY} [où X et Y désignent les types A (ascendant) ou D (descendant)] ssi elle est une transduction généralisée de type TG_X et son inverse une transduction généralisée de type TG_Y.

Nos résultats se présentent comme suit.

La classe des transductions inversibles et celle des relations associées aux bimorphismes sont identiques (nous les confondons par la suite). Ce résultat, facile à établir, concrètise la symétrie des règles. Il est évident intuitivement que l' "inversibilité au sens des règles" entraîne l' "inversibilité au sens des relations", (c'est-à-dire que toute transduction inversible est une bitransduction). Nous montrons même que $\widehat{TI} = \widehat{TG_{AA}} = \widehat{TG_{DA}} = \widehat{TG_{AD}}$ ce qui signifie que pour toute bitransduction de l'un de ces trois types, il existe un transducteur inversible associé. Cette équivalence des deux notions d'inversibilité (par règles et par relations) est le résultat le plus simple que l'on pouvait espérer. Pourtant, les résultats ci-dessus sont loin d'être évidents.

Nous avons d'ailleurs, pour le quatrième type de bitransductions, $\widehat{TG_{DD}} \neq \widehat{TI}$. D'autre part, la classe des transductions inversibles n'est pas stable par composition ($\widehat{TI} \neq \widehat{TI} \circ \widehat{TI}$). Mais nous montrons que $\widehat{TG_{DD}} = \widehat{TI}^2 = \widehat{TI}^3$. Notre but est atteint et nous obtenons de surcroit la stabilité par composition des bitransductions. Nous reviendrons en conclusion sur cet important résultat.

Nous commentons les résultats ci-dessus dans le troisième paragraphe où nous donnons surtout des exemples montrant que $\widehat{TG}_{DD} \neq \widehat{TI}$ et $\widehat{TI} \neq \widehat{TI}.\widehat{TI}$. Nous illustrons par la même occasion deux phénomènes essentiels : la "linéarité à délai borné" et les "chevauchements de découpages" de forêts. Il est intuitivement compréhensible que la seule hypothèse d'inversibilité ne puisse pas faire passer facilement des transducteurs descendants, au comportement fort complexe, aux transducteurs inversibles plus pratiques. L'inversibilité apparaît constamment à un niveau très technique au cours des démonstrations, pour supprimer les comportements "pathologiques" des transducteurs. Nous montrons au paragraphe IV que nous ne pourrions sans généraliser obtenir des résultats comparables. Nous étudions ensuite d'autres généralisations en montrant qu'elles s'éloignent souvent beaucoup de la notion originelle et qu'elles perdent de "bonnes propriétés". Nous remarquons que les "primitives transformations" de TAKAHASHI sont un cas très particulier de transduction inversible.

Nous résumons maintenant le cheminement de la seconde partie, consacrée aux démonstrations.

Dans le premier chapitre, nous caractérisons facilement les transductions inversibles par bimorphismes.

Dans le second chapitre, nous montrons d'abord deux lemmes (lemme II.1 de fidélité et lemme II.2 de linéarité globale). Il est fait dans la suite un usage constant de ces deux lemmes, ainsi que du caractère "reconnaissable" de beaucoup de phénomènes rencontrés. Nous montrons ainsi que si une bitransduction est réalisée, au moins dans un sens, par un transducteur généralisé ascendant, alors il existe un transducteur inversible équivalent. Nous raisonnons sur la forme des règles.

Nous considérons au chapitre III une classe intermédiaire de transducteurs descendants. Un transducteur descendant (généralisé ou non) a la particularité de pouvoir transducter de deux façons différentes une même branche. Dans le cas généralisé, un problème de "chevauchement des découpages" de la branche lue se pose. Nous considérons donc dans ce chapitre une classe intermédiaire (les transduc-

teurs quasi-généralisés descendants TGQ_n) pour laquelle nous réduisons plus facilement ce problème de chevauchement. Nous appelons "chaîne" toute suite de règles appliquées en un même noeud d'un arbre lu. Les "blocs" sont les suites de sous-dérivations obtenues par composition de chaînes. Une chaîne est "homogène" si toutes ses règles ont même partie gauche. Par composition, on obtient alors les blocs homogènes. Nous montrons que dans le cas quasi-généralisé, on peut toujours se ramener à des dérivations homogènes, ce qui supprime les problèmes de chevauchement. L'hypothèse d'inversibilité permet alors d'établir la proposition III.7 dont le point (i) concrètise le phénomène de "linéarité à délai borné". Ce phénomène signifie qu'une bitransduction ne peut transducter plusieurs fois une même branche que sur une profondeur bornée. Nous avons à considérer des chaînes infinies mais le point (ii) montre que dans une chaîne (ou un bloc) il n'y a qu'un nombre fini de variables (dite "essentielles") pouvant se dériver en des arbres arbitrairement longs.

Nous montrons ensuite qu'on peut "anticiper" les dérivations des variables non essentielles puis vérifier le bien fondé des anticipations par des mémorisations sur les dérivations essentielles. On montre par ailleurs que les bitransductions sont "finiement réductrices". Nous aboutissons alors à une classe plus particulière de transducteurs (les TGQ_nDF) dont les blocs sont des suites de taille bornée et vérifient le lemme III.8. (Toutes ces propriétés seraient fausse sans l'hypothèse d'inversibilité). Nous décomposons alors toute transduction ainsi obtenue en la composée de deux transductions inversibles ce qui aboutit au théorème III.13. Ce théorème fondamental montre qu'on peut remplacer le phénomène de linéarité à délai borné par un phénomène de "chevauchement des découpages" dans la forêt intermédiaire apparaissant en codomaine de la première transduction inversible et domaine de la seconde.

Nous abordons le chapitre IV en montrant que la classe des transductions généralisées finiment réductrices ($\widehat{TG}_D$) est stable par composition avec homomorphisme (proposition IV.2) et homomorphisme inverse (IV.1). Nous devons à nouveau pour ce dernier point résoudre un problème de chevauchement de découpages. Nous en déduisons facilement que la classe $\widehat{TG}_D$ est stable par composition par transduction inversible (proposition IV.3) et donc qu'elle contient la clôture par composition des transductions inversibles (proposition IV.4).

Nous montrons ensuite en IV.5 l'inclusion de la classe des bitransductions du type $\widehat{TG}_{DD}$ dans la classe $\widehat{TI}^2$. Le point délicat consiste à montrer que l'inversibilité permet de se ramener au cas quasi-généralisé en anticipant de façon homogène les dérivations. Nous concluons alors aisément, grâce au théorème 13 du chapitre précédent, par les théorèmes IV.6 et IV.7 qui achèvent la caractérisation annoncée des bitransductions et montrent la stabilité par composition.

Le chapitre V est consacré à des résultats complémentaires, de démonstration facile. Nous y montrons en particulier que les bitransductions conservent la reconnaissabilité.

Rappelons que les bitransductions sont les transductions généralisées stables par inversion. Sans cette seule hypothèse, la classe des bitransductions a des propriétés intéressantes : elle est stable par composition, elle se caractérise en termes de bimorphismes et conserve la reconnaissabilité (les cas ascendant et descendant deviennent en outre facilement comparables). Ces propriétés nous font espérer que les bitransductions joueront un rôle aussi fécond dans l'étude des forêts que celui des transductions rationnelles de langages. De fait, les premières applications confirment cette impression (ARNOLD - DAUCHET [AR 2]). Le travail présenté ici n'est donc qu'un début, consacré à l' "amont" des notions de bitransduction et transduction inversible.

En comparaison aux transductions rationnelles, deux questions se posent immédiatement.

- Les bitransductions conservent-elles l'algébricité des forêts ? Ce problème se ramène à celui de la stabilité de la classe des forêts algébriques par homomorphisme inverse, qui est un important problème ouvert. Nous conjecturons une réponse négative.

- Pourquoi $\widehat{TI} \neq \widehat{TI}^2 = \widehat{TI}^3$, c'est-à-dire pourquoi est-ce la classe des "bi-bimorphismes" qui est stable par composition ?

Pour commencer à répondre à ces deux questions, remarquons que nous avons rencontré au cours de notre travail un phénomène important : la "linéarité à délai borné". Ce phénomène était jusqu' alors insoupçonné parce qu'inexistant dans le cas monadique. De ce fait, le formalisme actuel devient inédéquat. Ceci explique nos difficultés techniques à dégager et à démontrer les propriétés essentielles.

Nous pressentons un "nouveau formalisme" rendant compte de la linéarité à délai borné . La reconnaissabilité y serait inchangée, les bitransductions aussi mais apparaîtraient de façon plus canonique. Elles conserveraient la "nouvelle" algébricité (très proche de la notion classique) et chaque bitransduction serait équivalente à un seul "nouveau" bimorphisme. C'est là l'objet de nos travaux en cours. Nous disposerions alors d'un bon outil pour étudier profondément les forêts algébriques et les magmas.

PREMIERE PARTIE

PRESENTATION DES RESULTATS ET DISCUSSION

Les termes employés dans la littérature anglo-saxonne figurent entre guillemets.

Nous ne les précisons que lorsque leur traduction française est ambigüe ou non évidente.

CHAPITRE I

RAPPELS

Le lecteur familier des notions de grammaires, reconnaisseurs et transducteurs d'arbres pourra se dispenser de lire ce chapitre.

Thatcher a fait récemment dans [AH] un excellent tour d'horizon informel des connaissances dans le domaine des arbres.

Le lecteur pourra s'y reporter en complèment à ce chapitre.

1 . DÉFINITIONS ET RAPPELS

A] Définitions élémentaires.

Introduisons les définitions de façon "naïve".

Un alphabet est un ensemble quelconque dont les éléments sont appelés lettres.

Nous ne considérons que des alphabets finis non vides.

Exemples :

α) l'alphabet latin $\Sigma_1 = \{a,b,c...,z\}$

β) $\Sigma_2 = \{0,1,2,3,4,5,6,7,8,9,+,-,\times,/,(,)\}$

γ) $\Sigma_3 = \{$symboles de base d'Algol 60$\}$.

Nous remarquons à ce propos qu'une lettre peut être une succession de symboles : par exemple, "begin" et "go to" sont deux lettres de Σ_3 au même titre que 0,1 ou A.

Un mot sur un alphabet est une séquence finie, totalement ordonnée, d'occurences de lettres de cet alphabet.

Par exemple, les mots de la langue française sont des mots sur l'alphabet latin.

Si les guillements n'appartiennent pas à l'alphabet considéré, nous noterons souvent entre guillemets chaque mot. Par exemple, le mot français papa s'écrira "papa". Le mot contient deux occurences de "a" et deux occurences de "p".

Il contient quatre occurences de lettres de l'alphabet. Nous dirons que le mot est de longueur quatre.

On introduit souvent le mot vide, de longueur zéro, que nous noterons " " ou d'un symbole distingué (par exemple Λ) si le caractère blanc " " appartient à l'alphabet.

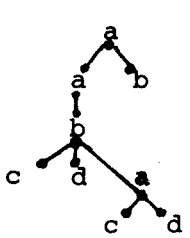
On appelle langage sur un alphabet tout ensemble fini ou infini de mots surcet alphabet. Par exemple, l'ensemble des mots du "Petit Larousse" est un langage fini sur Σ_1, l'ensemble des programmes en Algol 60 est un langage infini sur Σ_3.

Le mot "papa" est construit en concaténant "p" avec "a" pour obtenir "pa", puis "pa" avec lui-même. La concaténation est évidemment associative.

Formellement, si nous considérons un alphabet fini Σ, muni d'une opération linéaire associative partout définie (appelée concaténation), l'ensemble des mots sur Σ sera noté Σ^* et appelé monoïde libre engendré par Σ. Le mot vide est élément neutre pour la concaténation.

Dans un mot, une lettre quelconque a un seul successeur immédiat (la lettre qui la suit dans le mot) ou aucun successeur (si elle est la dernière du mot). Par exemple, dans le mot "papa", le premier "a" a un successeur ("p") et le second "a" n'a pas de successeur.

Nous allons à partir de cette remarque généraliser la notion d'alphabet en celle d'alphabet gradué et la notion de mot en celle d'arbre.



Le schéma ci-contre est un arbre t sur l'alphabet {a,b,c,d} en ce sens que les noeuds du graphe sont labellés des symboles a,b,c et d.

Les successeurs immédiats d'une occurence d'une lettre sont les occurences de lettres figurant dans l'arbre immédiatement en dessous de l'occurence considérée. Ici, l'occurence la plus en haut de "a" a deux successeurs immédiats ("a" et "d") et dans cet arbre, les diverses occurences de "a" ont tantôt deux, tantôt un successeur immédiat. "b" a trois successeurs immédiats ("c","d" et "a"), "c" et "d" n'en ont aucun.

Nous dirons que "a" est d'arité 1 et 2, "b" d'arité 3 et "c" et "d" d'arité 0.

Un alphabet gradué est donc un alphabet dans lequel chaque lettre est affectée d'un nombre fini d'entiers, ces entiers déterminant le nombre de successeurs immédiats possibles.

Formellement, un alphabet gradué Σ ("ranked") est la donnée d'un ensemble (fini) $|\Sigma|$ de labels et d'une partie finie de $|\Sigma| \times N$ appelée relation de rang ou d'arité sur Σ. Chaque lettre graduée est donc un couple ordonné dont la première composante est le label de la lettre et la seconde composante l'arité de la lettre et la seconde composante l'arité de la lettre.

Dans notre exemple, $|\Sigma| = \{a,b,c,d\}$ et la relation de rang est $\{(a,1),(a,2),(b,3),(c,0),(d,0)\}$.

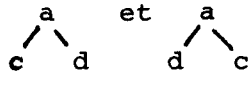
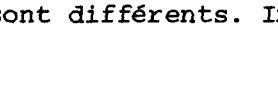
On note Σ_i l'ensemble des labels des lettres d'arité i

$$\Sigma_i = \{a \in |\Sigma| \mid (a,i) \in \Sigma\}$$

ici,

$$\Sigma_0 = \{c,d\},\ \Sigma_1 = \{a\},\ \Sigma_2 = \{a\} \text{ et } \Sigma_3 = \{b\}$$

La relation de rang sera fréquemment fonctionnelle, c'est-à-dire que chaque label aura une seule arité.

Remarquons qu'un arbre est orienté, c'est-à-dire que par exemple les arbres  et  sont différents. Intuitivement, deux arbres sont égaux si et seulement si ils sont superposables.

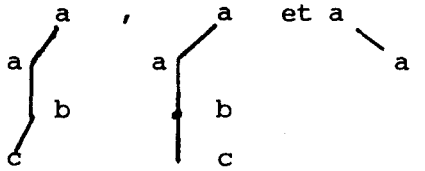
Le terme noeud désignera dans la suite un noeud accompagné de son label.

Le sommet d'un arbre est le noeud le plus en haut. Une feuille est un sommet n'ayant pas de successeur sur l'arbre. Remarquons que par définition, une feuille a un label d'arité nulle.

On appelle feuillage d'un arbre le mot obtenu en concaténant les feuilles de l'arbres, lues de gauche à droite. Le feuillage de notre exemple d'arbre t est "cdcdd".

Remarquons que pour définir le feuillage nous considérons les lettres d'arité nulle comme des lettres d'un monoïde et non pas comme des lettres d'un alphabet gradué. Nous considérons donc le feuillage d'un arbre comme un mot du monoïde Σ_o^*.

Une branche d'un arbre est une séquence de successeurs immédiats dans cet arbre, dont le premier élément est le sommet et le dernier une feuille.

Par exemple, 

sont des branches de t.

On appelle longueur d'une branche le nombre de noeuds de cette branche, diminué d'un.

Les branches ci-dessus ont respectivement les longueurs 3, 3 et 1. Une lettre d'arité nulle est une branche de longueur zéro.

La profondeur d'un arbre est la longueur de ses plus longues branches. La profondeur de t est 4.

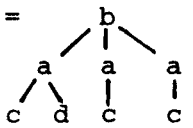
Soit Σ un alphabet gradué. Nous noterons Σ^* l'ensemble des arbres définis sur Σ. Un ensemble d'arbres sera appelé forêt. Le feuillage d'une forêt est l'ensemble des feuillages des arbres de la forêt.

Nous avons défini sur un alphabet non gradué l'opération de concaténation qui permet d'engendrer les mots à partir des lettres, et de nouveaux mots à partir de mots. Nous allons faire de même sur un alphabet gradué.

Considérons toujours l'alphabet $\Sigma = \{(a,1),(a,2),(b,3),(c,0),(d,0)\}$ à titre d'exemple.

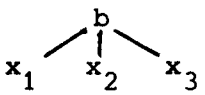
$t_1 =$ et $t_2 =$ sont des arbres sur Σ.

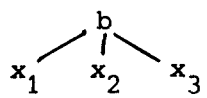
en "ajoutant b en sommet" on construit par exemple

$t_3 =$ 

en donnant dans t_3 pour successeurs immédiats de b, de gauche à droite, le sommet de t_1, le sommet de t_2 et encore le sommet de t_2. Il est commode de représenter par des symboles distingués, appelés variables ou indexes, les "noeuds virtuels", successeurs immédiats d'une lettre, auxquels se substitueront des arbres. Les variables permettent de préciser ces substitutions.

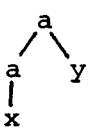
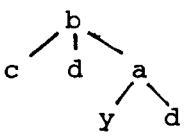
Par exemple, soient les variables x_1, x_2, x_3. Nous pourrons écrire

 au lieu de (b,3) et dirons que t_3 est obtenu en substituant

dans  l'arbre t_1 à la variable x_1, t_2 à x_2 et t_3 à x_3.

Mieux, nous pouvons considérer les variables x,y et dire que t_3 est obtenu en substituant dans l'arbre t_1 à la variable x et l'arbre t_2 à la variable y. Ici, t_2 est substitué à y en deux occurences.

Nous pouvons étendre cette notion aux arbres. Soit par exemple

$t_4 =$  , $t_5 =$ 

Des arbres tels que t_4 et t_5, qui contiennent des variables en haut de branches seront appelés arbres indexés ou dérivés. Ce dernier terme provient de l'utilisation fréquente de tels arbres dans les dérivations par transducteurs.

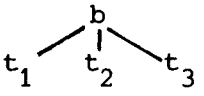
Notons t' l'arbre dérivé obtenu en substituant t_5 à x dans t_4. Cette opération se notera $t' = t_{4_x}(t_5)$ ou encore $t' = Sub_{t_4}(t_{5/x})$

Soit de même $t'' = t'_y(x)$. Nous noterons encore $t'' = t_{4x,y}(t_5, c)$. Il est facile de vérifier que $t = t''_y(x)$. Nous avons d'abord substitué les variables de t_4, puis celle de t_5.

Si nous substituons d'abord sur t_5, puis sur t_4, nous obtenons encore t. D'une façon générale, on montre l'associativité de la substition. L'ensemble des arbres sur Σ indexés par un ensemble X se note $\Sigma^*(X)$.

Faisons une importante remarque. Plusieurs occurences d'une même variable dans un même arbre impliquent des substitutions simultanées. C'est le cas de l'exemple

. Par contre, si une même variable apparait sur sur deux arbres différents, il ne s'en suit aucun lien entre les substitutions sur ces deux arbres. Par exemple, y apparait dans t_4 et t_5, mais doit être considéré comme une coïncidence fortuite de notations : si on remplace dans un des deux arbres y par z, par exemple, on ne change rien aux substitutions.

Soit une lettre d'un alphabet (par exemple (b,3)). Nous l'écrivons encore couramment $b(x_1,x_2,x_3)$ et tous les arbres de sommet b sont de la forme $b_{x_1,x_2,x_3}(t_1,t_2,t_3)$ où t_1,t_2,t_3 sont des arbres quelconques sur l'alphabet considéré. Dans ce cas, nous écrivons simplement $b(t_1,t_2,t_3)$. Nous avons ainsi une définition récursive de l'ensemble des arbres sur un alphabet. Remarquons que nous avons ici adopté une notation parenthèsée : nous avons noté $b(t_1,t_2,t_3)$ pour 

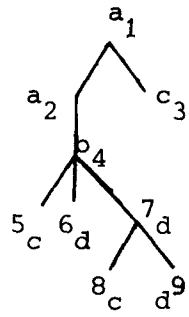
en notation arborescente.

Les deux notations sont également consacrées.

t s'écrit sous forme parenthèsée comme suit

$$t = a(a(b(c,d,a(c,d))),c)$$

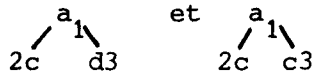
Afin de distinguer les diverses occurences d'une même lettre dans un arbre, on étiquette souvent les noeuds de cet arbre en attribuant à deux noeuds différents deux symboles évidemment différents. Par exemple, l'arbre considéré précédemment peut être étiqueté par des entiers comme suit :



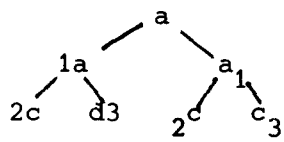
Ce procédé "naif" semble commode mais présente des inconvénients illustrés par ce qui suit : soient

$u_1 =$ et $u_2 :$

Nous pourrons les étiqueter à priori indépendamment par le même procédé.

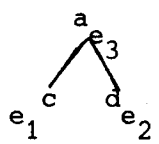


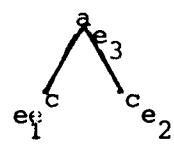
Considérons $a(u_1,u_2)$: nous obtenons en conservant les étiquetages l'arbre suivant :



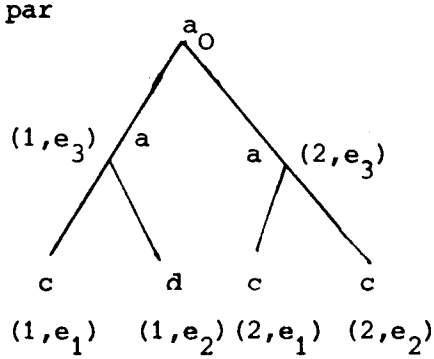
Plusieurs sommets sont affectés d'une même étiquette, il faut donc palier cet inconvénient. Un procédé classique consiste à étiqueter chaque noeud d'une suite finie d'entiers, définie par la structure de l'arbre et modifiée en ajoutant une composante chaque fois que l'on substitue. On étiquette -par exemple- le sommet par O.

Soient u_1 et u_2 étiquetés comme suit :

u_1 = 

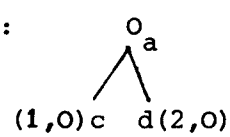
u_2 = 

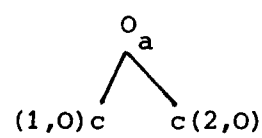
$a(u_1,u_2)$ sera étiqueté par



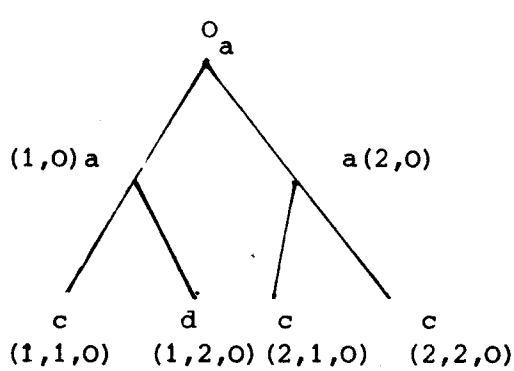
En commençant le procédé dès le feuillage nous obtenons l'étiquetage suivant :

O pour c, d, c, c étiquetés indépendamment comme sommets

u_1 : 



et $a(u_1,u_2)$:

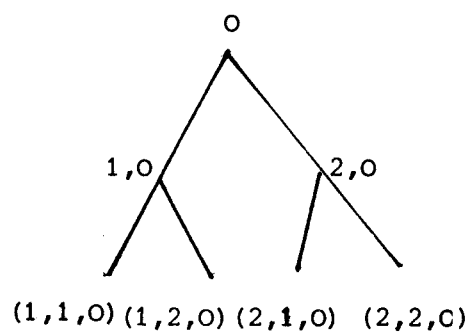


L'ensemble des étiquettes d'un arbre s'appelle alors son domaine. Ce terme se justifie par le fait que la donnée du domaine permet de reconstituer l'arbre sans ses labels.

Par exemple, soit le domaine {O,(1,O),(2,O),(1,1O),(2,1,O),(1,2,O (1,2,O),(2,2,O)}

O est l'étiquette du sommet.

Les successeurs du sommet ont une étiquette de deux chiffres : ici, (1,O) et (2,O). Le sommet est donc binaire. Les successeurs immédiats de (1,O) sont de la forme (x,1,O), d'où ici (1,1,O) et (2,1,O) etc...
On obtient



Donnons maintenant les définitions précises concernant les arbres :

Définition 0-1

Un alphabet fini Σ *est* gradué *("ranked") par une relation finie* r *de* $\Sigma \times \mathbf{N}$.

a $\in \Sigma$ *admet* i *pour arité ssi* (a,i) $\in$ r *posons pour tout* i, $\Sigma_i = r^{-1}(i)$.

Dans toute la suite, nous considérons r *fonctionnelle.*

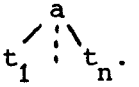
Définissons l'algèbre libre Σ^* engendrée par Σ, encore appelée magma libre.

a) $\Sigma_o \subset \Sigma$

b) si $n \geq 1$, $\forall a \in \Sigma_n$, $\forall t_1,\ldots t_n \in \Sigma^*$, alors $a(t_1,\ldots,t_n) \in \Sigma^*$

c) Σ^* est le plus petit ensemble satisfaisant à a) et b)

La définition ci-dessus nous donne la notation dite parenthèsée de Σ^*. La notation arborescente est obtenue en remplaçant dans b) l'écriture $a(t_1,\ldots,t_n)$ par



Soit S un ensemble (fini ou infini) d'indexes.

Définissons l'algèbre $\Sigma^*(S)$ des arbres sur Σ indexés par S.

a) $\Sigma_o \cup S \subset \Sigma^*(S)$

b) si $n \geq 1$, $\forall a \in \Sigma_n$, $\forall t_1,\ldots t_n \in \Sigma^*(S)$ alors $a(t_1,\ldots,t_n) \in \Sigma^*(S)$

c) tous les éléments de $\Sigma^*(S)$ sont obtenus ainsi.

En particulier, si $X = \{x_1,x_2,x_n,\ldots\}$ est un ensemble infini de variables indexées, nous appelerons encore arbre dérivé tout arbre de $\Sigma^*(X)$.

Il est évident que $\Sigma^* \subset \Sigma^*(S)$

Nous considérerons maintenant $a \in \Sigma_n$ comme élément de $\Sigma^*(X)$ et l'écrivons $a(x_1,\ldots,x_n)$. De même, tant $t \in \Sigma^*(X)$ s'écrira souvent $t(x_{i_1},\ldots,x_{i_p})$ en explicitant ses variables.

Si $u = u'(t_2) \in \Sigma^*(X)$, alors $\text{Sub}_{t_1/t_2}(u) = u'(t_1)$

Nous définissons maintenant trois relations sur $\Sigma^*(X)$ qui sont évidemment des relations d'ordre

$t_1 < t_2$ (nous dirons t_1 sous-arbre de t_2) ssi

$\exists t', t'_1,\ldots t'_n \in \Sigma^*(X)$ tq $t_2 = t'(t_1(t'_1,\ldots,t'_n))$

$t_1 \underset{I}{<} t_2$ (nous dirons t_1 sous arbre initial de t_2)

ssi $\exists t'_1,\ldots t'_n \in \Sigma^*(X)$ tels que $t_2 = t_1(t'_1,\ldots,t'_n)$

$t_1 \underset{F}{\leq} t_2$ (nous dirons t_1 sous arbre final de t_2)

ssi $\exists t' \in \Sigma^*(X)$ tel que $t_2 = t'(t_1)$

On appelle noeud de $t \in \Sigma^*(X)$ toute occurence d'un élément de $\Sigma \cup X$ danst,cet élément de $\Sigma \cup X$ est le label du noeud. Il y a souvent lieu de distinguer les diverses occurences d'une même lettre dans un même arbre. Cette distinction se fait systématiquement en affectant à chaque noeud une étiquette. Les procédés sont nombreux ([JO]) mais d'un usage lourd. Bien que dans la suite les distinctions d'occurences se feront empiriquement, nous donnons ci-après un moyen d'étiquetage en même temps que nous définissons les notions de branche, feuille, feuillage, profondeur et sommet d'un arbre.

(On pourrait également distinguer les occurences d'un même sous-arbre. Nous ne le ferons pas ici).

Soit E l'ensemble des suites finies d'entiers. Nous allons étiqueter chaque noeud a d'un arbre par $e(a) \in E$.

Nous noterons $\pi(t)$ la profondeur d'un arbre t.

Une branche b sera une suite finie de $E \times (\Sigma \cup X)$ la longueur de la branche le sera notée $l(b)$. L'application feuillage Φ applique Σ^* dans le monoïde libre Σ_o^* engendré par Σ_o.

a) $\forall x_i \in X$, $\pi(x_i) = 0$
x_i est le seul noeud, il est étiqueté $e(x_i) = 0$
$(0,x_i)$ est la branche x_i, de longueur nulle.
Σ_o est l'ensemble des feuilles de Σ^*

b) Soient $a(x_1,\ldots,x_n) \in \Sigma$ et $t_1,\ldots t_n \in \Sigma^*(X)$ alors le sommet de $u = a(t_1,\ldots,t_n)$ est a . $\pi(a(t_1,\ldots,t_n)) = \sup_{i=1,\ldots n} (\pi(t_i)) + 1$
Le sommet est étiqueté 0.

Soit e(n) l'étiquette d'un noeud de t_i. L'étiquette de ce noeud dans $a(t_1,\ldots,t_i,t_n)$ devient $(i,e(n))$.

Soit B(u) l'ensemble des branches de u.

$$B(u) = \{(0,a),b \mid \exists i \leq n,\ b \in B(t_i)\}$$

$$l((0,a),b) = l(b) + 1$$

Si $t_1,\ldots t_n \in \Sigma''$, $\Phi(a(t_1,\ldots,t_n)) = \Phi(t_1),\ldots \Phi(t_n)$.

Soit F une forêt de Σ^* ; on pose $\Phi(F) = \{\Phi(t) \mid t \in F\}$

Le domaine de t, noté D(t), est l'ensemble des étiquettes de ses noeuds.

fin de la définition 0.1.

Définition 0.2.

Soient Σ^* et Δ^* les forêts engendrées par les alphabets gradués Σ et Δ.

Définissons la notion d'homomorphisme de Σ^* dans Δ^*.

Un homomorphisme de Σ^* dans Δ^* et la donnée d'une application ϕ de Σ dans $\Delta^*(X)$ telle que :

$$\forall a(x_1,\ldots x_n) \in \Sigma,\ \Phi(a(x_1,\ldots x_n)) = \Phi(a)(x_{i_1},\ldots,x_{i_p}) \in \Delta^*(X)$$

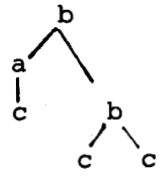
avec, pour tout $j \leq p$, $i_j \leq n$.

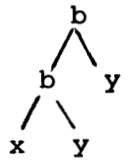
On définit alors l'homomorphisme de Σ^* dans Δ^* par :

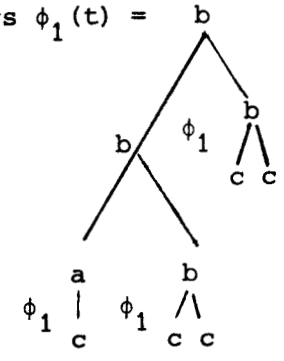
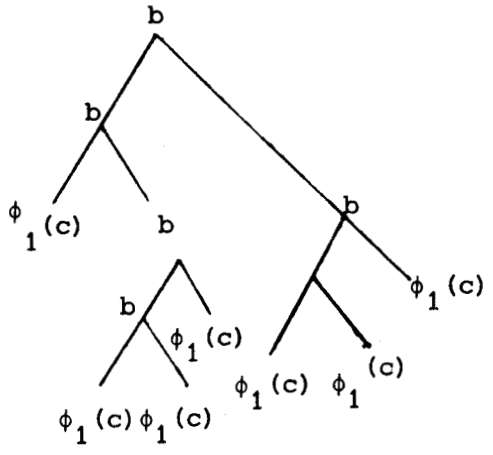
$$\phi(a(t_1,\ldots,t_n)) = \phi(a)\ (\phi(t_{i_2}),\ldots,\phi(t_{i_p}))$$

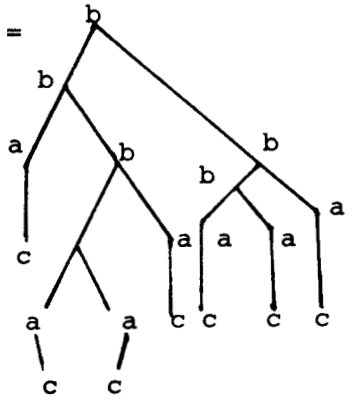
Exemples :

$$\Sigma = \Delta = \{(a,1),(b,2),(c,0)\}$$

Soit t = 

1) $\phi_1 \begin{pmatrix} a \\ | \\ x \end{pmatrix} = x$, $\phi_1 \begin{pmatrix} b \\ / \quad \backslash \\ x \quad y \end{pmatrix} =$  $\phi_1(c) = \begin{matrix} a \\ | \\ c \end{matrix}$

alors $\phi_1(t) =$  $=$ 

$=$ 

2) $\phi_2 \begin{pmatrix} a \\ | \\ x \end{pmatrix} = x$, $\phi_2 \begin{pmatrix} b \\ / \quad \backslash \\ x \quad y \end{pmatrix} = \begin{matrix} a \\ | \\ x \end{matrix}$, $\phi_2(c) = c$

BUS LILLE

alors $\phi_2(t) = \begin{array}{c} a \\ | \\ \phi_2\left(\begin{array}{c}a\\|\\c\end{array}\right)\end{array} = \begin{array}{c}a\\|\\\phi_2(c)\end{array} = \begin{array}{c}a\\|\\c\end{array}$

3) $\phi_3\left(\begin{array}{c}a\\|\\x\end{array}\right) = \begin{array}{c}a\\|\\x\end{array}$, $\phi_3\left(\begin{array}{ccc} & b & \\ / & & \backslash \\ x & & y\end{array}\right) = \begin{array}{ccc} & b & \\ / & & \backslash \\ y & & a \\ & & | \\ & & x\end{array}$ $\quad \phi_3(c) = c$

4) $\phi_4\left(\begin{array}{c}a\\|\\x\end{array}\right) = \begin{array}{c}a\\|\\x\end{array}$, $\phi_4\left(\begin{array}{ccc} & b & \\ / & & \backslash \\ x & & y\end{array}\right) = \begin{array}{ccc} & b & \\ / & & \backslash \\ z & & x\end{array}$, $\phi_4(c) = c$

<u>Cas particuliers</u> :

Si $\forall a(x_1,\ldots x_n) \in \Sigma$:

1) $\forall k = 1,\ldots n\ \exists j_k \leq p$ tel que $i_{j_k} = k$, est dit <u>complet</u> ("rank-preserving" ou "non-erasing").
Ceci signifie que toute variable de a se retrouve dans $\phi(a)$.
C'est le cas de ϕ_1, ϕ_3, ϕ_4.

2) $\forall k = 1,\ldots n$ il existe au plus un $j_k \leq p$ tel que $i_{j_k} = k$, ϕ est dit <u>linéaire</u> ("linear").
Ceci signifie qu'aucune variable de a n'apparait deux fois dans $\phi(a)$.
C'est le cas de ϕ_2, ϕ_3, ϕ_4.

3) $\forall k = 1,\ldots n$ il existe un et un seul $j_k \leq p$ tel que $i_{j_k} = k$, ϕ est dit <u>linéaire complet</u>.
Ceci signifie que les variables de $\phi(a)$ sont une permutation de celles de a. C'est le cas de ϕ_3, ϕ_4.

4) ϕ est linéaire complet <u>ordonné</u> ("distinguished") si la permutation est l'identité.
C'est le cas de l'homomorphisme identique.

5) $\pi(\phi(a)) > o$ ϕ est dit <u>continu</u> ou <u>strict</u>.
C'est le cas de ϕ_3, ϕ_4.

6) ϕ linéaire complet tel que $\phi(a) \in \Delta$ est appelé projection.
C'est le cas de ϕ_4.

B] Grammaires d'arbres.

Considérons la grammaire algébrique de mots donnée par les règles

$$\sigma \to a\ \xi\ b,\quad \xi \to a\ \sigma \text{ et } \xi \to a$$

où σ et ξ sont les variables, a et b les terminaux et σ l'axiome.

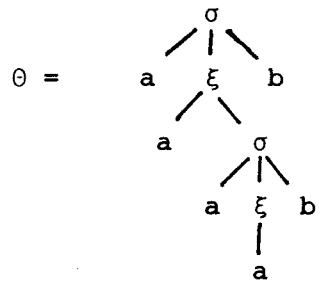
Il existe différentes notations consacrées, variant avec le domaine d'étude. Par exemple, en langages de programmation, on se donne fréquemment les grammaires sous la forme

$$\sigma ::= a\ \xi\ b$$
$$\xi ::= a\ \sigma \mid a$$

Du point de vue des séries formelles, on a un système d'équations

$$\sigma = a\ \xi\ b$$
$$\xi = a\sigma + a$$

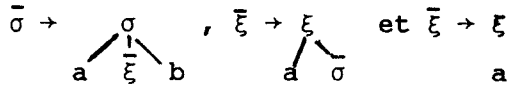
La dérivation $\sigma \Longrightarrow a\ \xi\ b \Longrightarrow aa\sigma b \Longrightarrow aaa\ \xi\ bb \Longrightarrow aaaabb$ admet pour arbre de dérivation



le mot engendré est le feuillage de cet arbre.

Nous allons maintenant nous donner une grammaire G_1 régulière d'arbres par un nombre fini de règles

Exemple 1.

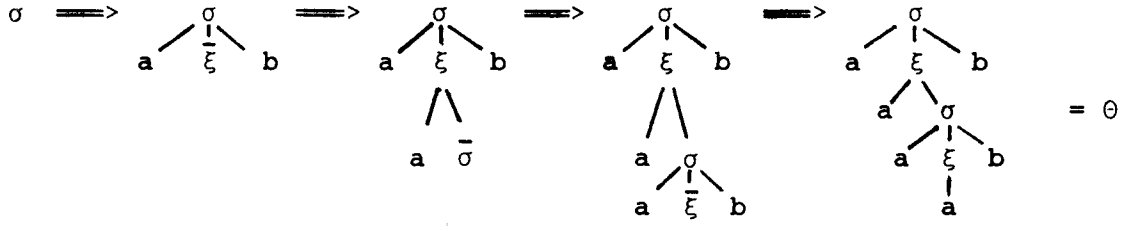


$\bar{\sigma}$ et $\bar{\xi}$ sont les variables, a,b,σ,ξ les terminaux et $\bar{\sigma}$ l'axiome. Les règles régulières auront toujours cette forme : 1 variable en partie gauche, des variables seulement en feuillage de la partie droite. Une dérivation s'effectue à partir de l'axiome, en remplaçant les variables apparaissant sur l'arbre par des parties droites de règles correspondantes.

La forêt engendrée par la grammaire est l'ensemble des arbres construits ainsi et qui ne contiennent aucune variable.

Si G est une grammaire, nous noterons F(G) la forêt engendrée.

Une forêt F est régulière si et seulement si il existe une grammaire régulière G telle que F = F(G). Donnons un exemple de dérivation



nous écrirons $\bar{\sigma} \overset{*}{\Longrightarrow} \theta$

Le lecteur vérifiera que la forêt engendrée par la grammaire d'arbres ci-dessus est l'ensemble des arbres de dérivations de la grammaire de mots considérée précédemment. On démontre que toute forêt de dérivations de tout langage algébrique est une forêt régulière. Notre exemple inspire une démonstration de ce résultat.

Remarquons que dans une dérivation d'une grammaire algébrique de mots, l'arbre se complète au fur et à mesure que l'on dérive, sans rien substituer, contrairement à ce qui se passe dans une grammaire d'arbres.

Enfin, le terme "variable" désigne selon les contextes deux notions différentes : nous avons précédemment défini les arbres indexés par des variables, et nous considérons ici des variables de grammaires d'arbres.
Le terme étant consacré dans les deux cas, nous l'utiliserons sans précision autre, sauf si une confusion est possible.

Dans une grammaire d'arbres, les variables doivent être considérés, au même titre que les terminaux, comme des lettres graduées. Dans le cas d'une grammaire régulière, toutes les variables sont d'arité nulle : elles ne peuvent donc figurer qu'en feuillage, tant dans les règles que dans les dérivations.

Les grammaires régulières d'arbres sont donc bien une généralisation des grammaires régulières de mots, dans les dérivations desquelles il n'apparait une variable qu'en "bout de mot".

Donnons un autre exemple de grammaire G_2 régulière d'arbres.

Exemple 2.

$G_2 = (\{S,A,A'\},\{(b,2),(a,1),(\bar{a},o),(\bar{a}',O)\}, R,S)$

$\{S,A,A'\}$ est l'ensemble des variables

$\{(b,2),\ldots(\bar{a}',O)\}$ l'ensemble des terminaux

S l'axiome

R l'ensemble des règles écrites ci-dessous en notations parenthèsées

$R = \{S \to b(A,A'),\ A \to a(A),\ A \to \bar{a},\ A' \to a(A'),\ A' \to \bar{a}'\}$

Le lecteur vérifiera que $F(G_2) = \{b(a^n(\bar{a}), a^p(\bar{a}')) \mid n \text{ et } p \in N\}$

Soit une occurrence de a dans un arbre t de $F(G_2)$.

Si cette occurrence est sur "la branche gauche de t", il pourra lui succéder a ou $\bar{a}$. Si a est sur la "branche droite de t", il lui succédera a ou $\bar{a}'$.

Les successeurs possibles de a dépendent donc des occurrences de a.

Par contre, dans tout arbre de $F(G_1)$, les successeurs d'un label sont ceux indiqués par les règles de la grammaire algébrique associée, et sont donc indépendants de l'occurrence. C'est pourquoi par analogie avec les langages locaux, les forêts de dérivations des langages algébriques sont dites locales.

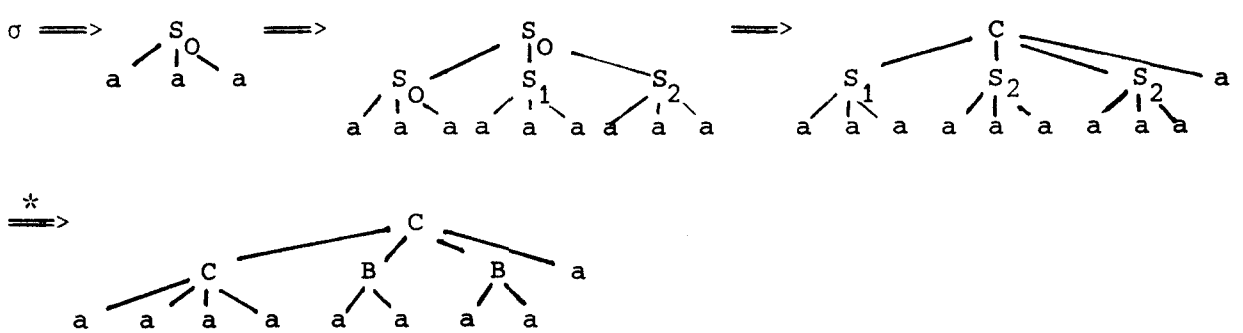
Nous allons maintenant donner un exemple de grammaire d'arbres plus générale : les variables seront d'arité quelconque. Néanmoins notre grammaire sera algébrique (ou créative, ou à contexte libre) car, par analogie au cas des langages les parties gauches de règles seront réduites à une variable.

Exemple 3.

$$G_3 = (\{(\sigma,0),(S_0,3),(S_1,3),(S_2,3)\},\ (a,0),(B,2),(C,4)\},\ R,\sigma)$$

$$R = \{\sigma \to S_0(a,a,a),\ S_0(x,y,z) \to S_0(S_0(x,y,z),\ S_1(x,y,z),\ S_2(x,y\ z))$$
$$S_0(x,y,z) \to C(y,z,z,a),\ S_1(x,y,z) \to C(y,z,z,a),$$
$$S_2(x,y,z) \to B(a,z)\}$$

Donnons un exemple de dérivation :



On montre que $\Phi(F(G_3)) = \{a^{n^2} \mid n \geqslant 2\}$ qui n'est pas algébrique.

Aho a défini une classe de langages plus large que celle des algébriques pour tenter de rendre compte de certaines contraintes "non libres" dans les langages de programmation.

Il a ainsi défini la classe des langages indexés. Nous n'utiliserons plus par la suite que des grammaires régulières, et la définition des langages indexés est complexe. Aussi nous bornerons nous à énoncer que "les feuillages des forêts algébriques sont les langages indexés" (résultat du à Fischer, cité dans [RO1])

Nous donnons maintenant les définitions précises des notions introduites précédemment, et citons quelques résultats importants déjà suggérés. Le lecteur trouvera dans [AH], [AR4], [BR1,2 et 3], [DO], [RO1 et 2], [TH1 et 2] ... des compléments sur le sujet.

Définition 1 : *Une grammaire d'arbres est la donnée d'un quadruplet*

$$G = (V, \Sigma, R, S)$$

V est un ensemble fini de variables graduées
Σ est un alphabet fini terminal gradué
R est un ensemble fini de règles
$S \in V_0$ est l'axiome.

Une grammaire est algébrique ("créative", "context-free") ssi toutes ses règles sont de la forme

$$A(x_1,\ldots,x_n) \to \theta(x_{i_1},\ldots,x_{i_p}) \text{ où } \forall j \leqslant p,\ i_p \leqslant n \text{ et } \theta \in (\Sigma \cup V)^*(X)$$

Une grammaire est régulière ssi elle est algébrique et toutes ses variables sont de degré 0. Les règles sont alors de la forme

$$A \to t(A_{i_1},\ldots A_{i_n}) \text{ où } A, A_{i_1},\ldots A_{i_n} \in V \text{ et } t(x_1,\ldots,x_n) \in \Sigma^*(X)$$

$$\text{si } n = 0,\ A \to t \text{ avec } t \in \Sigma^*$$

Définition 2 : *Soit* G *une grammaire algébrique*

$$t_1 \underset{G}{\Longrightarrow} t_2 \text{ ssi } \begin{cases} \text{il existe une règle } A(x_1,\ldots x_n) \to \Theta(x_{i_1},\ldots x_{i_p}) \text{ telle que} \\ t_1 = t'_1(A(t''_1,\ldots,t''_n)) \\ t_2 = t'_1(\Theta(t''_{i_1},\ldots,t''_{i_p})) \end{cases}$$

Remarquons que dans le cas d'une grammaire régulière,

$$t_1 \underset{G}{\Longrightarrow} t_2 \text{ ssi } \begin{cases} \text{il existe une règle } A(x_1,\ldots x_n) \to \Theta(x_{i_1},\ldots x_{i_p}) \text{ telle que} \\ t_1 = t'(A) \text{ et} \\ t_2 = t'_1(t(A_{i_1},\ldots,A_{i_n})) \end{cases}$$

$\overset{*}{\underset{G}{\Longrightarrow}}$ *est la clôture transitive de* $\underset{G}{\Longrightarrow}$

Définition 3 : *La forêt engendrée par* G *se note* F(G) *et se définit par*

$$F(G) = \{t \in \Sigma^* \mid S \overset{*}{\underset{G}{\Longrightarrow}} t\}$$

Deux grammaires sont équivalentes ssi elles engendrent la même forêt.

Une forêt est régulière (respectivement algébrique) ssi il existe une grammaire régulière (respectivement algébrique) qui l'engendre.

Théorème 1 : [DO], [TH1],...

a) *La classe* D *des forêts de dérivations des grammaires algébriques de mots est une sous-classe de celle des forêts régulières.* D *est dite classe des forêts locales.*

b) *Toute forêt régulière est la projection d'une forêt locale.*

c) *Le feuillage d'une forêt régulière est un langage algébrique.*

Ce théorème précise l'intérêt des forêts régulières pour l'étude des langages algébriques.

C] Reconnaisseurs d'états finis.

Nous avons défini les forêts régulières de façon générative à partir des grammaires. Nous allons maintenant les définir par reconnaissance. Les reconnaisseurs de forêts régulières sont étudiés dans [BR2], [DO], [TH2]

Nous allons encore retrouver une généralisation du cas des langages : un reconnaisseur d'états finis lit les arbres de bas en haut, lettre par lettre, en prenant à chaque lecture un état déterminé par la lettre lue et les états pris après lecture des successeurs immédiats de la lettre considérée.

On se donne des états finaux et un arbre est reconnu, ssi il existe un mouvement lisant l'arbre des feuilles à la racine et se terminant dans un état final.

Définition 4 : *Un reconnaisseur* M *(ou machine) d'états finis est un quadruplet*

$$M = (\Sigma, Q, Q_F, \alpha)$$

Σ est un quadruplet gradué
Q est un ensemble fini d'états
Q_F l'ensemble des états finaux
α est une relation entre l'ensemble

$$\{(a(x_1,\ldots x_n), q_{i_1},\ldots,q_{i_n}) \mid a \in \Sigma,\ q_{i_j} \in Q\} \text{ et } Q$$

le reconnaisseur est déterministe si α est fonctionnelle.

La relation $||\ ||_M$ de réponse associe à tout $t \in \Sigma^*$ une partie $||t||_M$ de Q définie comme suit

$\forall a \in \Sigma_o$, $q \in ||a||_M$ ssi $(a,q) \in \alpha$

$\forall a \in \Sigma_n$, $\forall t_1,\ldots t_n \in \Sigma^*$, $q \in ||a(t_1,\ldots,t_n)||_M$ ssi il existe $q_{i_1},\ldots q_{i_n}$ tels que

i) $((a(x_1,\ldots,x_n), q_{i_1},\ldots q_{i_n}),q) \in \alpha$

ii) $\forall j = 1\ldots n,\ q_{i_j} \in ||t_j||_M$

$t \in \Sigma^*$ est <u>reconnu</u> par M ssi $||t||_M \, Q_F \neq \emptyset$.

On pose $F(M) = \{t \in \Sigma^* \mid ||t||_M \cap Q_F \neq \emptyset\}$

Deux reconnaisseurs sont équivalents ssi ils reconnaissent la même forêt.

<u>On montre que tout reconnaisseur est équivalent à un reconnaisseur déterministe</u>.

$F \subset \Sigma^*$ est <u>reconnaissable</u> ssi $\exists M$ telle que $F = F(M)$

<u>Exemple 4</u>.

Nous allons nous donner un reconnaisseur M de la forêt régulière générée dans l'exemple 2.

Σ est l'alphabet de l'exemple 2

$Q = \{q_1, q_1', q_0\}$

$Q_F = \{q_0\}$

Nous construisons M déterministe. α est donc fonctionnelle

$\alpha(\bar{a}) = q_1$, $\alpha(\bar{a}') = q_1'$, $\alpha(a(x),q_1) = q_1$, $\alpha(a(x),q_1') = q_1'$, $\alpha(b(x_1,x_2),q_1,q_1') = q_0$

alors $||b(\bar{a}',\bar{a})||_M = b(||\bar{a}'||_M, ||\bar{a}||_M)$

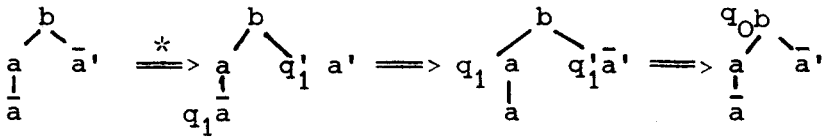
on a $||\bar{a}'||_M = q_1'$ et $||\bar{a}||_M = q_1$.

$\alpha(b(x_1,x_2),q_1',q_1)$ n'est pas défini. Le mouvement ne peut pas continuer.

$b(\bar{a}',\bar{a})$ n'est pas reconnu.

Par contre, $\left\| \begin{array}{c} b \\ / \quad \backslash \\ a \quad \bar{a}' \\ | \\ \bar{a} \end{array} \right\| M = q_0$

il sera commode d'écrire



En faisant figurer en chaque noeud la réponse de M pour le sous-arbre de sommet ce noeud.

Nous formaliserons cette écriture dans la définition 8.

D] Rappel de résultats.

Théorème 2 : On peut caractériser la classe des forêts régulières en terme de clôture de Kleene ([TH2]). *Cette classe est la même que celle des forêts reconnaissables, et que celle des forêts solutions d'un système d'équations linéaires*[BR2], [EH], [MW], [SH]. *La classe est close par intersection et complémentation, par homomorphisme linéaire ; par projection et projection inverse* (TH2).

Ces résultats sont rappelés informellement car nous les utiliserons peu par la suite. Ils étendent simplement à une algèbre graduée les résultats sur les monoïdes.

Mais les homomorphismes non linéaires ne conservent pas les forêts régulières. Nous allons maintenant étudier les problèmes de transformations de forêts. Comme le remarque Thatcher dans [AH], les problèmes de transformations sont différents dans la forêts et les langages (contrairement à la reconnaissance).

2 . Les transducteurs

Seuls les transducteurs d'arbres d'états finis (ou régulier) ont été profondément étudiés (voir [BA], [BJ], [MA], [OG], [RO1], [RO5], [TH3], [TH4].)

Comme annoncé en introduction, nous nous limitons également à l'étude de ceux-ci, que nous désignerons sans le simple terme de transducteurs.

A] Définitions.

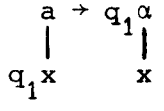
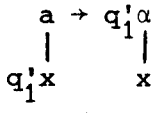
Comme dans le cas des langages un transducteur ascendant est un "reconnaisseur avec sortie" : à la lecture de chaque lettre graduée, on sort un arbre dérivé.

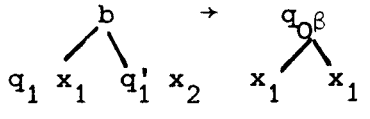
Dans le cas des langages, on concatène sans problème le mot sorti à chaque pas avec le mot sorti au cours des pas antérieurs. Ici, nous allons substituer aux variables de l'arbre dérivé (sorti au pas considéré), des arbres sortis antérieurement, selon une correspondance indiquée par les variables de chaque règle de dérivation.

Exemple.

Donnons nous un transducteur ascendant T par les règles suivantes :

$r_1 : \bar{a} \to q_1 \bar{\alpha}$ $\qquad$ $r_1' : \bar{a}' \to q_1' \bar{\alpha}'$

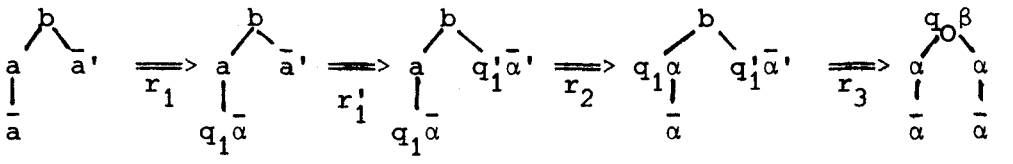
$r_2 :$  $\qquad$ $r_2' :$ 

$r_3 :$  $\qquad$ q_0 est l'état initial.

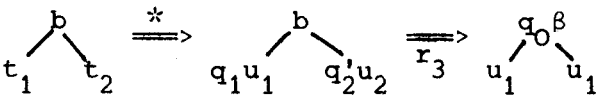
Un arbre est transducté si et seulement si il existe un mouvement le "parcourant de bas en haut" et aboutissant à un état final au sommet.

Le domaine est l'ensemble des arbres "lus". Ici, le domaine est la forêt considéré dans les exemples 2 et 4.

Exemple de mouvement (la règle utilisée est indiquée sous la flèche).



Ainsi, plus généralement, soit t_1 se transformant en q_1u_1 (nous noterons $t_1 \overset{*}{\Longrightarrow} q_1u_1$) et $t_2 \overset{*}{\Longrightarrow} q_1'u_2$.

Alors 

r_3 s'est appliquée en substituant, à gauche et à droite, u_1 à x_1 et u_2 à x_2.

Le transducteur donné ci-dessous est déterministe car on a jamais le choix de la règle à utiliser en une occurrence donnée.

Donnons une notation parenthèsée des règles et des mouvements : au lieu d'écrire l'état près du sommet de l'arbre écrit en notation arborescente, nous écrirons l'état puis entre crochets l'arbre en notation parenthèsée ainsi

$r_1 : \quad \bar{a} \to q_1\,[\bar{\alpha}] \qquad\qquad r_2 : \quad a(q_1[x]) \to q_1[\alpha(x)]$

$r_3 : \quad b(q_1[x_1],\ q_2[x_2]) \to q_0[\beta(x_1,x_2)]$ et $b(a(\bar{a}),\bar{a}') \overset{*}{\Longrightarrow} q_0[\beta(\alpha(\bar{\alpha}),\alpha(\bar{\alpha}))]$

Donnons maintenant une définition précise des transducteurs ascendants ("bottom-up", "frontier-to-root")

Définition 5 : *Un transducteur ascendant* (en abrégé TA) *est un* 5-uple $T = (Q,\Sigma,\Delta,R,P)$ *où*

- Q *est un ensemble fini d'états*
- P Q *l'ensemble des états finaux*
- Σ *l'alphabet gradué initial*
- Δ *l'alphabet gradué final.*

R est un ensemble fini de règles. Chaque règle est de la forme

$b \to q[u]$ *où* $b \in \Sigma_0$, $q \in Q$ *et* $u \in \Delta^*$

ou de la forme :

$b(q_{i_1}[x_1],\ldots,q_{i_n}[x_n]) \to q[u(x_{i_1},\ldots x_{i_p})]$ *ou* $n > 0$,
$b(x_1,\ldots,x_n) \in \Sigma_n$, $q, q_{i_1} \ldots q_{i_n} \in Q$, $u \in \Delta^*(X)$
et où $\forall j \leqslant p$, $i_j \leqslant n$.

Définissons la relation $\underset{T}{\Longrightarrow}$ *de dérivation directe par*

$t(b(q_{i_1}[u_1],\ldots,q_{i_n}[u_n])) \underset{T}{\Longrightarrow} t(q[u(u_{i_1},\ldots,u_{i_p})])$ *ssi*
$b(q_{i_1}[x_1],\ldots q_{i_n}[x_n]) \to q[u(x_{i_1},\ldots,x_{i_p})] \in R$
et $b \Longrightarrow q[u]$ *ssi* $b \to q[u] \in R$.

$\overset{*}{\underset{T}{\Longrightarrow}}$ *est la clôture transitive de* $\underset{T}{\Longrightarrow}$

On appelle <u>transduction associée à T</u> *la relation* $\hat{T}$ *de* $\Sigma^* \times \Delta^*$ *définie par*

$$\hat{T} = \{(t,u) \mid \exists q \in P,\ t \overset{*}{\underset{T}{\Longrightarrow}} q[u]\}$$

Dans notre exemple, $\hat{T} = \{(b(a^n\bar{a},a^p\bar{a}'),\ \beta(\alpha^n\bar{\alpha},\alpha^n\bar{\alpha}))$ | n et p $\in$ N$\}$

Nous allons donner un exemple de <u>transducteur descendant</u>. Le principe d'application des règles est le même que dans le cas ascendant, mais on transforme du sommet vers les feuilles.

Soient les règles de T'

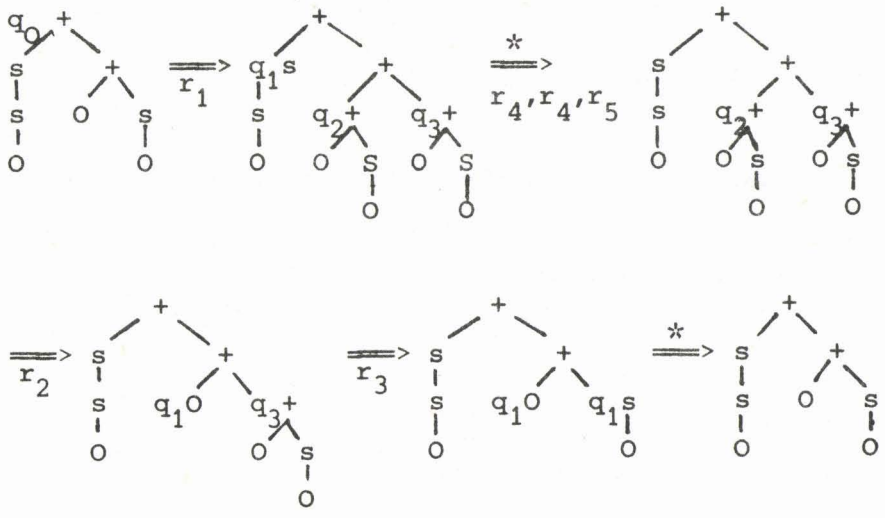
r_1 : $q_0[+(x_1,x_2)] \to +(q_1[x_1], +(q_2[x_2], q_3[x_2]))$

r_2 : $q_2[+(x_1,x_2)] \to q_1[x_1]$

r_3 : $q_3[+(x_1,x_2)] \to q_1[x_2]$

r_4 : $q_1[s(x)] \to s(q_1[x])$ $\qquad$ r_5 : $q_1[0] \to 0$

alors, en notation arborescente :



En interprétant + par le "plus" arithmétique, s par la fonction successeur et O par zéro, on voit que T' réalise l'associativité de l'addition.

Donnons des définitions précises.

Définition 6 : *Un transducteur descendant* ("top-down", "root-to-frontier", *TD en abrégé*) *est un 5-uple* $T = (Q,\Sigma,\Delta,R,P)$ *où* Q,Σ,Δ *ont la même signification que dans le cas ascendant,* P *est l'ensemble des états initiaux*
R *est un ensemble fini de règles de la forme*

$$q[b] \to u \text{ où } q \in Q,\ b \in \Sigma_0 \text{ et } u \in \Delta^*$$

ou de la forme

$$q[b(x_1,\ldots,x_n)] \to u(q_{j_1}[x_{i_1}],\ldots,q_{j_p}[x_{i_p}])$$

si $n \geq 1$, $q \in Q$, $u \in \Delta^*$, $q_{j_1},\ldots,q_{j_p} \in Qb(x_1,\ldots,x_n) \in \Sigma_n$ *et* $\forall k \leq p,\ i_p \leq n$.

Définissons la dérivation directe descendante $\underset{T}{\Longrightarrow}$ *par*

$$q[b] \underset{T}{\Longrightarrow} u \text{ ssi } q[b] \to u \in R$$

$$u_1(q[b(t_1,\ldots t_n)]) \underset{T}{\Longrightarrow} u_1(u(q_{j_1}[t_{i_1}],\ldots q_{j_p}[t_{i_p}])) \text{ ssi } q[b(\alpha_1,\ldots\alpha_n)] \to$$

$$u(q_{j_1}[x_{i_1}],\ldots q_{j_p}[x_{i_p}]) \in R$$

$\overset{*}{\underset{T}{\Longrightarrow}}$ *est la clôture transitive de* $\underset{T}{\Longrightarrow}$.

On appelle transduction associée à T *la relation* $\hat{T}$ *de* $\Sigma^* \times \Delta^*$ *définie par*

$$\hat{T} = \{(t,u) \mid \exists q \in P,\ q\ [t] \xRightarrow[T]{*} u\}$$

Définition 7 : Soit TP *la classe des transducteurs ayant la propriété* P.

Nous allégerons souvent "soit un transducteur de la classe TP" *en "soit un* TP"

Une relation R $\Sigma^* \times \Delta^*$ *est une transduction (de type)* P (*en abrégé* TP) *ssi il existe* $T \in$ TP *tel que* $R = \hat{T}$.

Deux transducteurs T *et* T' *sont équivalents ssi* $\hat{T} = \hat{T}'$.

Un transducteur (TA *ou* TD) *est déterministe ssi il ne contient pas 2 règles de même partie gauche.*

Il est à un état ("one-state", "pure") ssi Q *est réduit à un seul élément.*

Les adjectifs complet, linéaire, ordonné, continu ou strict imposent aux règles les mêmes contraintes qu'aux homomorphismes (voir 2A).

Par exemple, dans l'exemple T, toutes les règles sont linéaires complètes strictes sauf r_3 qui n'est ni linéaire (x_1 a deux occurrences droites), ni complète (x_2 n'a pas d'occurrence droite).

Dans T', r_1 est complète mais non linéaire, r_2 et r_3 ne sont ni strictes ni complètes etc...

Le domaine d'un transducteur T *est l'ensemble*

$$\text{Dom } T = \{t \in \Sigma^* \mid \exists u \in \Delta^*,\ (t,u) \in \hat{T}\}$$

et le codomaine est défini par

$$\text{Codom } T = \{u \in \Delta^* \mid \exists t \in \Sigma^*,\ (t,u) \subset \hat{T}\}.$$

Nous allons donner une nouvelle définition de reconnaisseur :

un reconnaisseur sera un transducteur qui recopie ce qu'il lit.

Par exemple, le reconnaisseur M de l'exemple 4 aura pour règles :

$$\bar{a} \to q_1[\bar{a}],\ \bar{a}' \to q_1'[\bar{a}'],\ a(q_1[x]) \to q_1[a(x)]$$

$$a(q_1'[x]) \to q_1'[a(x)] \text{ et } b(q_1[x_1],\ q_1'[x_2]) \to q_o[b(x_1,x_2)]$$

Définition 8 : *Une nouvelle définition des reconnaisseurs* :

a) *un reconnaisseur d'états finis ascendant (en abrégé* R.A*) est un* TA *tel que* $\Sigma = \Delta$ *et où toutes les règles sont de la forme* $b \to q[b]$ *ou* $b(q_{i_1}[x_1],\ldots,q_{i_n}[x_n]) \to q[b(x_1,\ldots,x_n)]$ *avec les notations de la définition 5.*

F *est reconnaissable ssi* $\exists t \in$ RA *tel que* Dom(T) = F.

b) *un reconnaisseur d'états finis descendant (en abrégé* RD*) est un* TD *tel que* $\Sigma = \Delta$ *et où toutes les règles sont de la forme* $q[b] \to b$ *ou* $q[b(x_1,\ldots,x_n)] \to b(q_{i_1}[x_1],.\ q_{i_n}[x_n])$ *avec les notations de la définition 6.*

F *est reconnaissable ssi* $\exists T \in$ RD *tel que* Dom(T) = F.

Il est facile de voir que les trois notions de reconnaissabilité coïncident. Notons que tout reconnaisseur est équivalent à un RA déterministe. mais pas à un RD déterministe.

Remarque 4.

Nous ne considérons plus dans toute la suite que des tranducteurs d'états finis à un seul état initial (au final) q_o. On montre en effet que tout transducteur est équivalent à un transducteur du même type à un seul état initial (au final). Nous nous donnerons un transducteur par l'ensemble de de ses règles, les autres éléments étant implicites.

B] Exemples complémentaires.

Les exemples seront cités dans la suite .

Exemple 5.

Soit $T_1 \in TD$ défini par les règles

$$q_0[a(x)] \to b(q_1[x], q_2[x])$$
$$q_1[a(x)] \to a[q_1(x)],\ q_1[a] \to \bar{a}$$
$$q_2[a(x)] \to a'[q_2(x)],\ q_2[\bar{a}] \to \bar{a}'$$

Il est facile de voir que $\hat{T}_1 = \{(a^n\bar{a}, b(a^{n-1}\bar{a}, a'^{n-1}\bar{a}')) \mid n \geq 1\}$

Soit $T_2 \in TA$ défini par les règles :

$$\bar{a} \to q_1[\bar{a}],\ a(q_1[x]) \to q_1[a(x)]$$
$$a(q_1[x]) \to q_1[a'(x)],\ a(q_1[x]) \to q_0[b(x,x)]$$
$$\hat{T}_2 = \{(a^n\bar{a}, b(\alpha_1\alpha_2\ldots\alpha_{n-1}\bar{a}, \alpha_1\alpha_2\ldots\alpha_{n-1}\bar{a}) \mid n \geq 1, \alpha_i \in \{a,a'\}\}$$

Soit $T_3 \in TA$ défini par les règles :

$$\bar{a} \to q_1[\bar{a}],\ q_1[a(x)] \leftarrow a(q_1[x]),\ q_1[a'(x)] \leftarrow a'(q_1[x])$$
$$\bar{a} \to q_2[\bar{a}],\ q_2[a(x)] \leftarrow b(q_2[x]),\ q_2[a'(x)] \leftarrow b'(q_2[x])$$
$$b(q_1[x], q_2[y]) \to q_0[b(x,y)]$$
$$\hat{T}_3 = \{b(\alpha_1\ldots\alpha_n\bar{a}, \alpha'_1\ldots\alpha'_p\bar{a}),\ b(\alpha_1\ldots\alpha_n\bar{a}, \beta_1\ldots\beta_p\bar{a}) \mid n \in N,\}$$

où

$\alpha_i \in \{a,a'\}$, $\alpha'_i \in \{a,a'\}$, $\beta_i = b$ ssi $\alpha'_i = a$ et $\beta_i = b'$ ssi $\alpha'_i = a'$.

Exemple 6.

Les homomorphismes sont des transducteurs à la fois ascendants et descendants, déterministes à un seul état.

Par exemple ϕ de $\{c,a,\bar{a}\}^*$ dans $\{b,a,\bar{a}\}^*$ défini par

$$\phi(c(x,y,z)) = b(b(x,y),z)$$

$\phi(a(x)) = a(x)$ et $\phi(\bar{a}) = \bar{a}$ est réalisé par T_4 comme suit

$$q_0[c(x,y,z)] \to b(b(q_0[x], q_0[y]), q_0[z])$$
$$q_0[a(x)] \to a(q_0[x]) \text{ et } q_0[\bar{a}] \to \bar{a}$$

ϕ^{-1} est réalisé par le transducteur descendant T_5

$$q_0[b(x,y)] \to c(q_1[x], q_2[x], q_0[y])$$
$$q_1[b(x,y)] \to q_0[x], \; q_2[b(x,y)] \to q_0[y]$$
$$q_0[a(x)] \to a(q_0[x]) \text{ et } q_0[\bar{a}] \to \bar{a}.$$

Par contre si nous définissons ψ par $\psi(a) = \Phi(a)$, $\psi(\bar{a}) = \Phi(\bar{a})$ et $\psi(c(x,y,z)) = b(b(x,y), b(z,c))$, on peut montrer que $\hat{\psi}^{-1}$ n'est pas une transduction. Nous reviendrons au chapitre III sur ce fait. Nous en sommes amené à étudier les propriétés des transducteurs.

C] Propriétés des transducteurs d'états finis.

Théorème 4 :

a) *Le domaine d'un transducteur est une forêt régulière.*

b) *Les transducteurs linéaires conservent les forêts régulières et les forêts algébriques.*

c) *L'image réciproque d'une forêt régulière par un transducteur est régulière.*

Ce sont les rares résultats "positifs" connus, en ce sens que la plupart des questions que l'on peut se poser raisonnablement reçoivent une réponse négative comme nous allons le voir.

Théorème 5 :

a) *Les classes* $\widehat{TD}$ *et* $\widehat{TA}$ *sont incomparables.*

exemple $\widehat{T_1} \notin \widehat{TA}$ et $\widehat{T_2} \notin \widehat{TD}$.

b) *Les classes* $\widehat{TD}$, $\widehat{TA}$ *et* $\widehat{TD} \cup \widehat{TA}$ *ne sont pas closes par composition (au sens des relations)*

exemple il est facile de voir que $\widehat{T_1} \in \widehat{TA}^2$, $\widehat{T_2} \in \widehat{TD}^2$ et $\widehat{T_3} \circ \widehat{T_2} \notin \widehat{TD} \cup \widehat{TA}$

c) *Les classes* $\widehat{TD}$, $\widehat{TA}$ *et* $\widehat{TD} \cup \widehat{TA}$ *ne sont pas closes par inversion (au sens des relations)*

exemple T_1, T_2, T_3 dans l'exmple 5, ψ^{-1} dans l'exemple 6.

Comme le remarque B. Baker [BA], ces résultats proviennent de ce que les TD peuvent engendrer deux copies différentes déterministes d'une même branche mais pas deux copies identiques non déterministes, à l'inverse des TA. Nos exemples traduisent bien ces faits.

Il est conjecturé ([BA],[OG]) un résultat beaucoup plus fort que le théorème 5 concernant la non stabilité par composition. Posons F la classe des forêts régulières, $D_n = TD^n(F)$ la classe des forêts images d'une forêt régulière par la composition de n TD, et $A_n = TA^n(F)$ définie de même.

Conjecture. Les hiérarchies $(A_n)_{n\in N}$ et $(D_n)_{n\in N}$ sont strictes.

CHAPITRE II

TRANSDUCTEURS GENERALISES, BIMORPHISMES ET BITRANSDUCTIONS

Nous avons vu que les notions de transductions ne sont ni symétriques, ni composables du sens des relations. Notre premier but a été de caractériser les relations qui étaient des transductions ainsi que leurs inverses. Comme nous l'expliquerons au chapitre suivant, nous avons été amené à généraliser la notion de transducteur d'états finis, d'où les définitions suivantes :

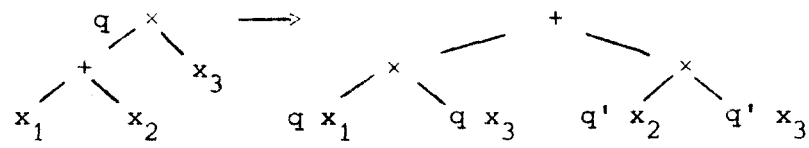
1 . DÉFINITIONS.

Définition 9 : Un transducteur généralisé ascendant (respectivement descendant) est défini comme un TA *dans la définition 6 (respectivement* TD *dans la définition 7) en y remplaçant* "$b \in \Sigma_o$" *par* "$b \in \Sigma^*$ *et* $\pi(b) > o$" *et* "$b(x_1,\ldots,x_n) \in \Sigma_n$" *par* "$b(x_1,\ldots,x_n) \in \Sigma^*(x)$ *et* $\pi(b) > o$"

Autrement dit, on généralise la notion de transducteur d'états finis en prenant pour partie gauche de règles des arbres quelconques de profondeur non nulle.

On notera TG_A et TG_D les 2 classes ainsi définies

par exemple,



est une règle généralisée.

Définition 10 : On dira qu'une règle d'un transducteur généralisé descendant est réductrice si elle est de la forme

$$q[t(x_1,\ldots,x_n)] \to q'[x_i]$$

On dira qu'un transducteur $\tau \subset TG_D$ *est finiment réducteur ("bounded erasing") si dans toute dérivation, les nombre de règles réductrices utilisées successivement est borné.*

Les règles $r_1 = q[t(x_1,\ldots,x_n)] \to t'(\ldots,q'[x_i],\ldots)$
et $r_2 = q'[t''(x_1,\ldots,x_p)] \to t'''(\ldots)$

sont dites appliquées successivement dans la dérivation

$D = q_o[U_o] \overset{*}{\Longrightarrow} u_1$ ssi

$u_o = u(t(\ldots,t"(\ldots),\ldots))$ et

$D = q_o[U_o] \overset{*}{\Longrightarrow} u'(\ldots,q[t(\ldots,t"(\ldots),\ldots)],\ldots) \underset{r_1}{\Longrightarrow}$

$u'(\ldots,t'(\ldots,q'[t"(\ldots)],\ldots),\ldots) \overset{*}{\Longrightarrow}$

$u"(\ldots,q'[t"(\ldots)],\ldots) \underset{r_2}{\Longrightarrow} u"(\ldots,t"'(\ldots),\ldots) \overset{*}{\Longrightarrow} u$

Pour tout $n \geq 1$, on notera $TG_{\overline{D}}$ la classe des transducteurs finiment réducteurs de TG_D.

Définition 11 : *Nous n'allons plus considérer par la suite que des homomorphismes linéaires, complets, strictes (voir [1] A) Le terme homomorphisme leur sera réservé. Nous nous référerons au cas général par le terme homomorphisme général.*

Définition 12 : *Nous appelerons bimorphisme tout sextuplet* $\langle A,\Sigma,\Delta,F,\phi,\psi\rangle$
où A,Σ,Δ *sont des alphabets gradués*
F *est une forêt régulière de* A^*
ϕ *est un homomorphisme de* A^* *dans* Σ^*
ψ *est un homomorphisme de* A^* *dans* Δ^*

Pour simplifier, nous écrivons $\langle F,\phi,\psi\rangle$ ce bimorphisme.
La relation $\hat{\mu}$ associée à $\mu = \langle F,\phi,\psi\rangle$ est définie par :

$$\hat{\mu} = \{(\phi(t), \psi(t)) \mid t \in F\}$$

Soit $\hat{B}$ la classe des relations associées aux bimorphismes.

Définition 13 : *Toute règle d'un* TG_A *ou d'un* TG_D *qui est linéaire complète est dite inversible.*
Un TG_A *(respectivement* TG_D*) dont toutes les règles sont inversibles est appelé transducteur généralisé inversible ascendant (respectivement descendant).*
Soient respectivement TI_A *et* TI_D *les classes ainsi définies.*
Nous appelerons transducteur généralisé inversible tout transducteur de $TI = TI_A \cup TI_D$.

Définition 14 : *Nous appelerons* bitransduction de type X,Y *toute relation* R *telle que* :

- *il existe* $T \in TG_X$, $\hat{T} = R$ *avec* X = A *ou* D
- *il existe* $T' \in TG_Y$, $\hat{T'} = R^{-1}$ avec X = A *ou* D

Nous noterons $\widehat{TG}_{XY}$ *la classe des relations ainsi définies, et* TG_{XY} *la classe des transducteurs de* TG_X *tels qu'il existe* $T' \in TG_Y$, *vérifiant* $\hat{T'} = \hat{T}^{-1}$.

2 . RÉSULTATS PRINCIPAUX

a) $\widehat{TI} = \widehat{TI}_A = \widehat{TI}_D = \hat{B}$

b) $\widehat{TG}_{AD} = \widehat{TG}_{DA} = \widehat{TG}_{AA} = \widehat{TI}$

c) $\widehat{TG}_{DD} = \widehat{TI}^2$

d) $\widehat{TI} \neq \widehat{TI}^2$

e) $\widehat{TI}^2 = \widehat{TI}^3$

Ces résultats sont démontrés en seconde partie dans le chapitre I pour le point a), dans II pour b) et dans III et IV pour c) et e).
d) est démontré en première partie dans III.

Ils s'énoncent encore ainsi :

a) Tout transducteur généralisé inversible ascendant est équivalent à un transducteur généralisé inversible descendant et réciproquement.
Tout transducteur généralisé inversible est équivalent à un bimorphisme et inversement.
Autrement dit, les transduction inversibles se caractérisent en termes de bimorphismes de façon tout à fait comparable aux K- transductions de M. NIVAT dans les langages [NI]

b) Soit T un transducteur généralisé descendant. Il existe un transducteur généralisé T' réalisant la transformation inverse si et seulement si T est équivalent à un TI.

c) Soit T un transducteur généralisé descendant. Il existe un transducteur généralisé T' ascendant (respectivement descendant) réalisant la transformation inverse si et seulement si T est équivalent à un TI (respectivement à un TI^2)

d) les transductions inversibles ne se composent pas

e) le composé de 3 TI est équivalent au composé de 2 TI.

Corollaires :

$$\widehat{TG}_{DD} = \widehat{TG}_{AA}^{\,2} = \widehat{TG}_{AD}^{\,2} = \widehat{TG}_{DA}^{\,2}$$

Les classes $\widehat{TG}_{DD}$ *et* $\widehat{TI}^2$ *sont stables par composition.*

Nous commentons au chapitre suivant ces résultats.

Remarque.

Nous omettons fréquemment par la suite le terme "généralisé". Le cas non généralisé sera toujours explicitement signalé.

CHAPITRE III

COMMENTAIRES ET ILLUSTRATIONS DES RESULTATS OBTENUS

Un TI_A a toutes ses règles de la forme

$$t(q_{i_1}[x_1],\ldots,q_{i_n}[x_n]) \rightarrow q[u(x_{\sigma(1)},\ldots,x_{\sigma(n)})]$$

où t et u sont des arbres dérivés de profondeur non nulle. En renversant les flèches, on obtient un TI_D réalisant la transduction inverse. Nous pouvons dire qu'un TI est "inversible par règles". (Dans le cas de règles non linéaires complètes strictes, parler de règle inverse est impossible au sens des transducteurs comme nous le verrons en (III,3)).

Le chapitre I de la seconde partie est facile, la caractérisation des TI en termes de bimorphismes etant intuitive, vu la symétrie des règles.

Notre but était de caractériser les transductions dont la relation inverse est également une transduction.

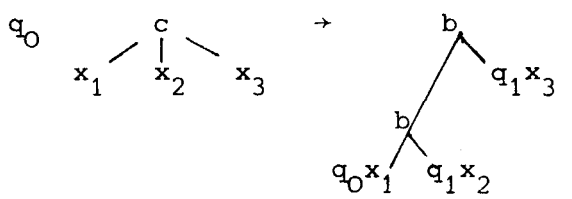
Nous obtenons en b) que "l'inversibilité du sens des relations entraîne l'inversibilité au sens des règles" dans le cas où un au moins des 2 transducteurs est ascendant, et ceci bien entendu à une équivalence de transducteurs près (c'est-à-dire que si $T \in TD_A$ et $\bar{T}$ est une bitransduction, $\exists T' \sim T$ tel que $T' \in TI$)

L'inversibilité du sens des règles entrainant de façon évidente, l'inversibilité au sens des relations, le résultats.
b) est le plus simple que l'on saurait espérer mais n'est plus contrairement à d) intuitivement prévisible (voir partie 2, chapitre II).

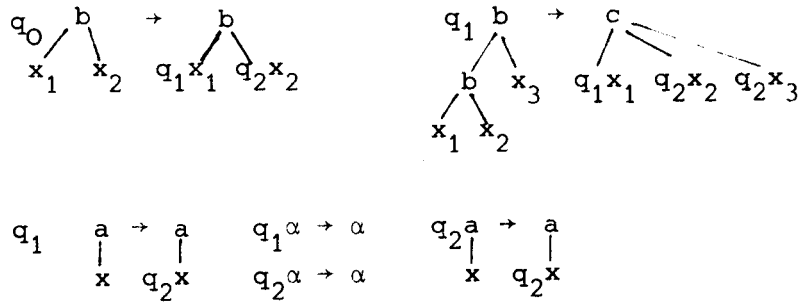
Nous aurions pu espérer le même résultat pour TG_{DD}.
Il n'en est rien comme le montre l'exemple suivant :

Théorème : $\widehat{TG_{DD}} \neq \widehat{TI}$ *et* $\widehat{TI}^2 \neq \widehat{TI}$
Afin de soutenir l'intuition du lecteur, nous adoptons une notation arborescente.

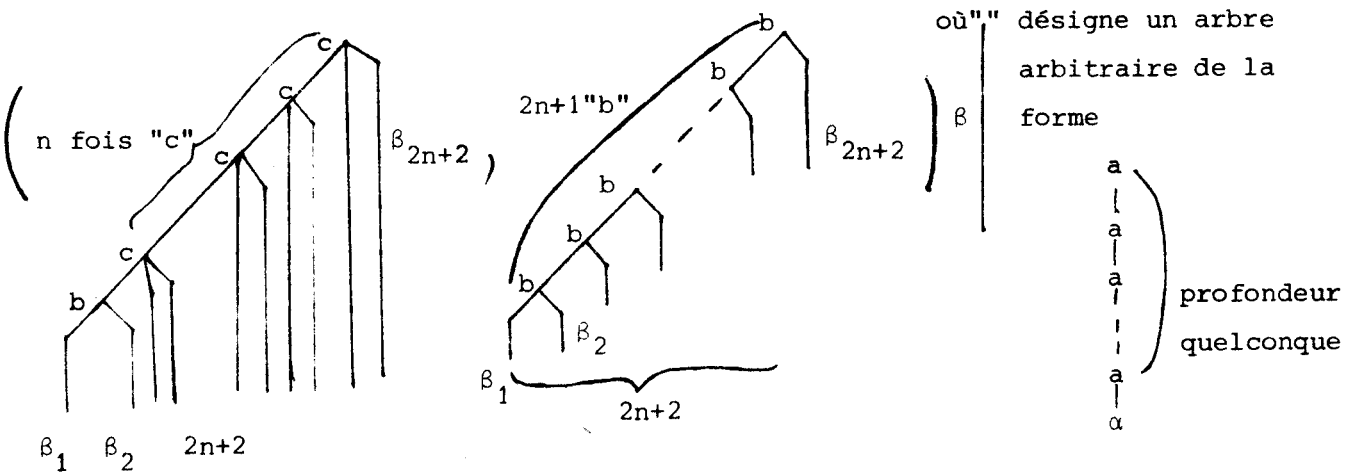
Soit $\Sigma = \Delta = \{(c,3),(b,2)(a,1)(\alpha,0)\}$
TI_1 est le TI défini par les règles :



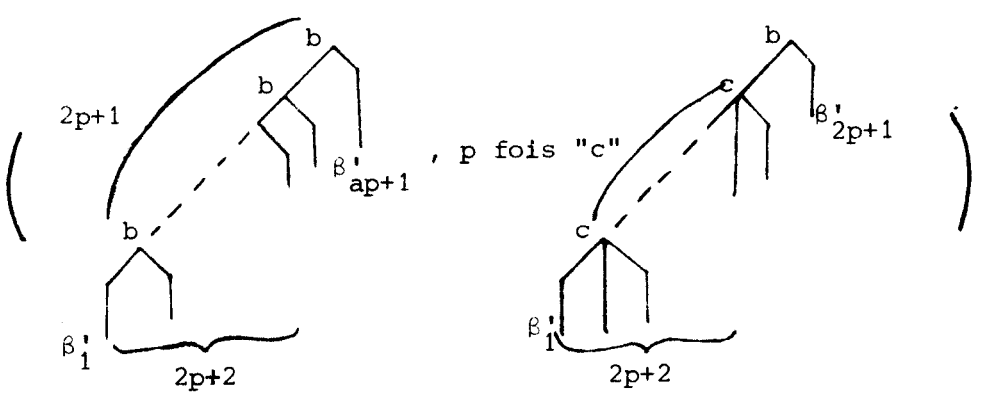
TI_2 est le TI défini par les règles :



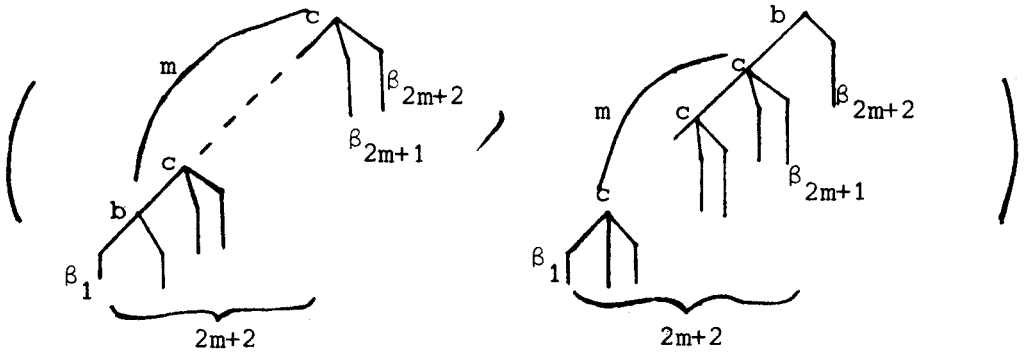
Il est facile de montrer que $\widehat{TI_1}$ est l'ensemble des couples d'arbres de la forme suivante, pour n € N.



$\widehat{TI_2}$ est l'ensemble des couples



$\widehat{TI_1} \circ \widehat{TI_2}$ est l'ensemble des couples



Supposons qu'il existe un transducteur inversible TI_3 tel que $\widehat{TI_3} = \widehat{TI_1} \circ \widehat{TI_2}$.

Pour n assez grand et les chaînes monadiques assez longues, on obtient une contradiction car on aurait une règle initial avec 2i+3 variables à gauche et 2j+4 à droite. ce qui prouve que $\widehat{TI}^2 \neq \widehat{TI}$

La non-composabilité provient ici de ce que les arbres intermédiaire sont engendrés par T_1 et transductés par T_2 selon des découpages qui ne coïncident jamais.

Soit U ∈ T DD défini par les règles :

$r_1 : q_o\ c(x\ y\ z) \rightarrow b(c(q_2x\ q_3x\ q_1y)\ q_1z)$

$r_2 : q_1\ a(x) \rightarrow d(q_1x)$

$r_3 : q_1\alpha \rightarrow \alpha$

$r_4 : q_2\ c(x\ y\ z) \rightarrow c(q_2x\ q_3x\ q_2y)$

r_5 : $q_3\,c(x,y,z) \to q_1 z$

r_6 : $q_2\,b(x,y) \to q_1(x)$ et r_7 : $q_3\,b(x,y) \to q_1(y)$

r_8 : $q_o\,b(x,y) \to b(q_1 x, q_1 y)$

Soit V $\in$ T DD défini par les règles :

s_1 : $q_o\,b(x,y) \to c(q_2 x, q_3 x, q_1 y)$

s_2 : $q_1\,a(x) \to a(q_1 x)$

s_3 : $q_1 \alpha \to \alpha$

s_4 : $q_2\,c(x,y,z) \to c(q_2 x, q_3 x, q_2 y)$

s_5 : $q_3\,c(x,y,z) \to q_1 z$

s_6 : $q_2\,b(x,y) \to q_1(x)$

s_7 : $q_3\,b(x,y) \to q_1(y)$

s_8 : $q_o\,b(x,y) \to b(q_1 x, q_1 y)$

Le lecteur vérifiera que $\hat{U} = \widehat{TI}_1 \circ \widehat{TI}_2$ et $\hat{V} = (\widehat{TI}_1 \circ \widehat{TI}_2)^{-1}$ d'où $\widehat{TDD} \neq \widehat{TI}$ et à fortiori $\widehat{TG}_{DD} \neq \widehat{TI}$

Dès lors, les résultats $\widehat{TG}_{DD} = \widehat{TI}^2$ et $\widehat{TI}^2 = \widehat{TI}^3$ sont les meilleurs que l'on pouvait espérer.

Ces résultats étaient difficiles à deviner, encore plus à démontrer. Les exemples U et V illustrent un premier fait : on ne transducte pas arbitrairement loin plusieurs fois une même branche : par exemple, r_4 a pour variables droites x et y, r_5 a pour variables droite z, donc r_4 et r_5 "séparent" la branche recopiée deux fois par le x de r_1.
Nous appelerons ce phénomène "linéarité à délai un". A l'égard de ce seul phénomène, notre exemple est simple car déterministe et non généralisé.

On démontre au chapitre III que tout $\widetilde{TGQ}_n D$ est "linéaire à délai borné" : en prenant des règles composées, (donc généralisées) on se ramène au délai un.

Dans les $\widehat{TGQ}_n D$, disons grosso modo qu'il n'y a pas de "problème de découpage". Au chapitre IV, nous montrons qu'on peut pour un $\widehat{TG}_{DD}$ supprimer les problèmes de découpage en anticipant l'arbre lu sur une profondeur borné. Ceci pose de nombreux problèmes car il faut pouvoir ensuite explorer toutes les branches, de façon déterministe, pour vérifier si l'anticipation était justifiée.

La grosse différence entre les $\widehat{TG}_{AD}$ et les $\widehat{TG}_{DD}$ est qu'en transduction descendante, on peut recopier différemment deux fois une même branche, ce qui est impossible en ascendante.

L'exemple précédent illustre le point i) de la proposition III 7). Nous allons modifier cet exemple de façon à illustrer le point ii) de la même proposition et à faire dans ce cas particulier la construction aboutissant au lemme III-9.

Faisons quelques remarques préliminaires.

Soit W_1 le TG_D donné par les règles

$q_o[a(x)] \to b(q_1[x], q_2[x])$

$q_1[a(x)] \to a(q_1[x])$

$q_2[a(x)] \to q_2[x]$

$q_1[a'(x)] \to a'(q_1'[x])$

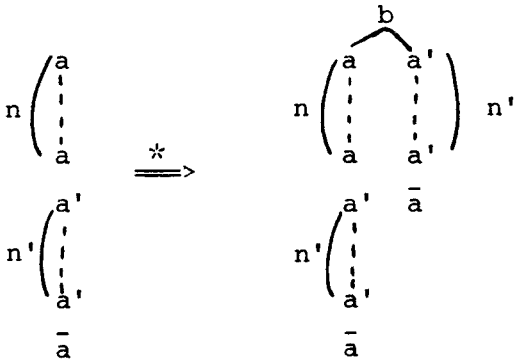
$q_2[a'(x)] \to a'(q_2'[x])$

$q_1'[a'(x)] \to a'(q_1'[x])$

$q_2'[a'(x)] \to a'(q_2'[x])$

$q_1'[\bar{a}] \to \bar{a}$ et $q_2'[\bar{a}] \to \bar{a}$.

alors



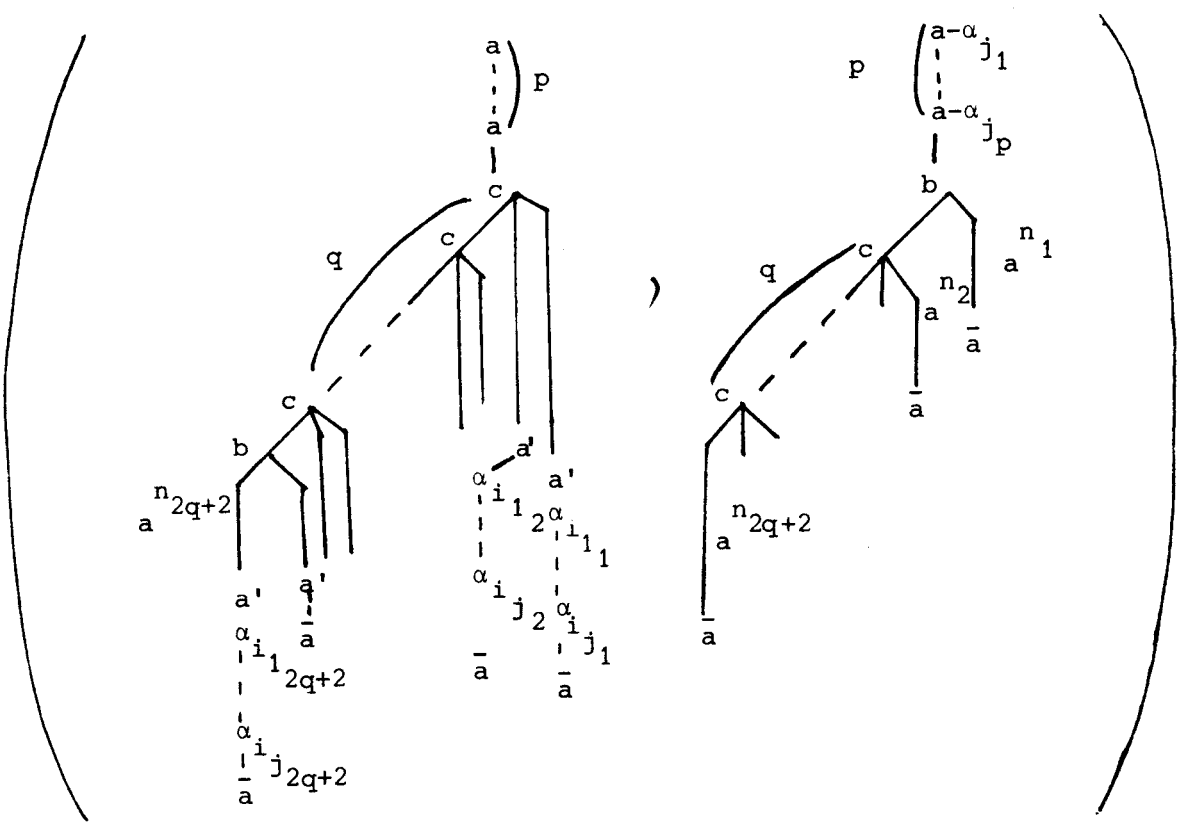
W_1 n'est pas finiment réducteur et il est facile de montrer par le lemme II-2 que $\widehat{W_1}$ n'est pas une bitransduction.

Mais il existe des TG_{DD} qui ne sont pas finiment réducteurs. On montre néanmoins en III-7 ii) qu'après avoir réduit arbitrairement loin, on ne peut plus recopier que des arbres bornés".

Donnons l'exemple annoncé, qui illustre ce fait.

Soit $\Sigma = \Delta = \{(a,1),(b,2),(c,3),(a',1),(\alpha_1,1),\ldots(\alpha_k,1),(\bar{a},0)\}$

Posons $A = \{\alpha_1,\ldots\alpha_k\}$



P est l'ensemble des couples de la forme ci-dessus où

$p,q,n_1,\ldots,n_{2q+2} \in N$ et $\forall\ l = 1,\ldots 2q+2$, $\alpha_{i_{1_l}},\ldots,\alpha_{i_{j_l}}$ sont différents 2 à 2 et

$$\{\alpha_{j_1},\ldots,\alpha_{j_p}\} \subset \bigcup_{l=1}^{2q+2} \{\alpha_{i_{1_l}},\ldots\alpha_{i_{j_l}}\}$$

P est une bitransduction.

En effet, donnons U_1 tel que $\hat{U}_1 = P$. Posons $S = \left\{ \begin{matrix}\alpha_{i_1}\\ \vdots \\ \alpha_{i_j}\end{matrix} \mid j_1 \neq j_2 \Longrightarrow \alpha_{i_{j_1}} \neq \alpha_{i_{j_1}} \right\}$

Remarquons que S est fini.

r_1 : $q_o\ a \to a$ (fils : x ; $q_o x$, $\bar{\alpha}_i x$) $\quad i = 1,\ldots k$

r_2 : $q_o\ c \to b$ (fils : c, $q_1 z$; c : $q_3 x$, $q_2 y$, $q_1 y$)

r_3 : $q_3\ c \to c$ (fils : $q_3 x$, $q_2 x$, $q_1 y$)

r_4 : $q_2\ c \to q_1 z$

r_5 : $q_1\ a \to a$ (fils : $q_1 x$)

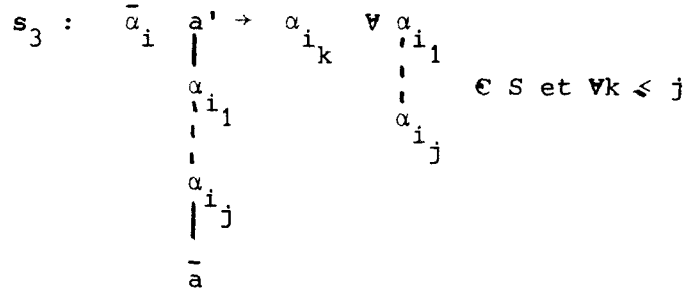
r_6 : $q_1\ a' \to \bar{a}\ \forall \begin{matrix}\alpha_{i_1}\\ \vdots \\ \alpha_{i_j}\end{matrix} \in S.$

$$\begin{matrix}\alpha_{i_1}\\ \vdots \\ \alpha_{i_j}\end{matrix}$$
$$\bar{a}$$

s_1 : $\bar{\alpha}_i\ a \to \bar{\alpha}_i x$

s_2 : $\bar{\alpha}_i\ c \to \bar{\alpha}_i x$ ou $\bar{\alpha}_i y$ ou $\bar{\alpha}_i z$

$\bar{\alpha}_i\ b \to \bar{\alpha}_i x$ ou $\bar{\alpha}_i y$

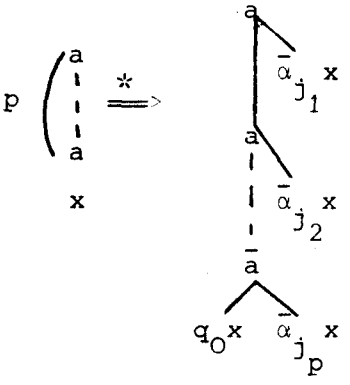
s_3 : 

Les régles du type r jouent le même rôle que dans l'exemple précédent.

U_1 n'est pas finiment réducteur car les régles s_1 et s_2 descendent au hasard "l'arbre gauche", sans rien recopier, et s_3 prend au hasard en bas de l'arbre lu, un symbole α qu'elle recopie.

Grosso modo, disons que nous appelons (au chapitre III) chaîne toute suite de régles s'appliquant en un même sommet (nous n'avons pas dans notre exemple de problème d'homogénéité). Un bloc est une chaîne dérivée. Le "délai de linéarité" étant un comme dans l'exemple précédent, il est inutile ici de faire intervenir les blocs.

La longueur d'une chaîne est le nombre d'occurences de régles qu'elle contient. Remarquons qu'une chaîne n'est pas nécessairement de longueur bornée. En effet, prenons un mouvement initial de U_1 :



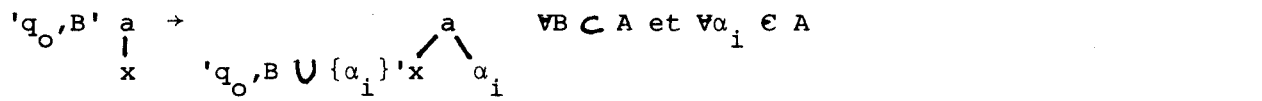
et la chaîne s'appliquant en x contiendra p+1 règles.

Nous appelons au chapitre III occurence principale (de x) toute occurence susceptible "de se dériver en un arbre arbitrairement long". Les autres occurences sont non principales. Nous démontrons en III que le nombre d'occurence principales est borné. Ici, seule l'occurence associée à q_o est principale, les autres ne se dériveront qu'en un arbre de profondeur un.

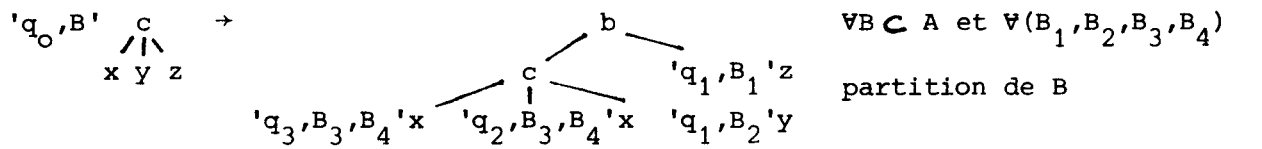
Nous allons anticiper les arbres se dérivant des occurences non principales (puisque le nombre d'arbres est fini) en mémorisant notre choix dans les états associés aux occurences principales. Nous montrons qu'on peut toujours "vérifier si l'anticipation était bonne".

Puisque les occurences non principales de variables disparaissent, les "bloc anticipés" utilisés sont de longueur bornée.

Illustrons la construction par notre exemple : nous construisons de nouveaux états que nous écrivons entre guillements.

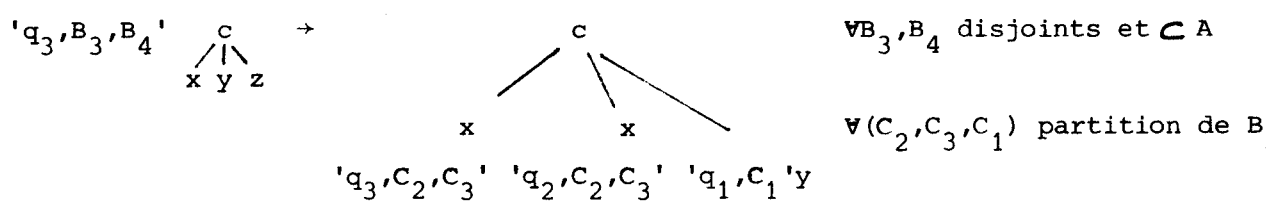


Nous avons anticipé le choix de α_i et l'avons mémorisé dans q_o.



Nous "partons vérifier en bout de branches, au hasard, que tout α de B figure bien". Le point important consiste à remarquer qu'il est indispensable de "dériver B_3, B_4 avant de les séparer". Plus précisément, en **x** va s'appliquer une chaîne de la forme (r_3,r_4) mais on ne peut dériver indépendamment les B_i dans r_3 et r_4, d'où la nécessité dans la partie III de "deviner les blocs et non "leurs règles séparément". La linéarité à délai borné nous assure que "tout bloc appliqué en **x** apparait bien en partie droite d'une même règle".

Ainsi, ici,



$'q_2,B_3,B_4'$ c → $'q_1,B_4'$z $\forall B_4 \subset A$

x y z

On termine évidemment par

$'q_1,B'$ a → a $\forall B \subset A$

x $'q_1,B'$ x

$'q_1,B'$ a' → $\bar{a}$ ssi $B \subset \{\alpha_{i_1}, \ldots \alpha_{i_j}\}$

α_{i_1}

α_{i_j}

$\bar{a}$

vérifiant ainsi que tout α anticipé est bien enbout d'une branche.

Bien entendu, ce phénomène d'anticipation, qui se produit ici en début de dérivation, peut se produire "n'importe où" et interférer avec d'autres phénomènes (problèmes de découpage, de déterminisme).
Dans la section III, nous nous débarassons délibérément du "problème de découpage" et justifions cet acte en IV.

Remarque.

Nos notations peuvent paraître redondantes (ici, on aurait pu écrire $'q_3,B_3'$, et $'q_2,B_4'$ seulement).

Elles le sont en effet parfois pour faciliter les constructions et les démonstrations. Le phénomène général d'anticipation est en effet plus complexe du fait que les arbres anticipés sont non réduits à une lettre, d'où on doit, encore de façon liée dans un bloc, deviner les dérivations intermédiaires. On a intérêt à "stocker" trop d'informations mais de façon homogène plutôt que tout juste ce qu'il faut de façon disparate".

Donnons enfin, sans autre commentaire, V_1 tel que $\hat{V}_1 = P^{-1}$.

$$\begin{array}{l} 'q_o,B' \quad \begin{array}{c} a \\ / \backslash \\ x \ \alpha_i \end{array} \quad \rightarrow \quad \begin{array}{c} a \\ | \\ 'q_o,B \cup \{\alpha_1\}' \ x \end{array} \quad \forall B \subset A \text{ et } \alpha_i \in A \end{array}$$

$$\begin{array}{l} 'q_o,B' \quad \begin{array}{c} b \\ / \backslash \\ x \ y \end{array} \quad \rightarrow \quad \begin{array}{c} c \\ \begin{array}{ccc} x & x & \\ 'q_3,B_3' & 'q_2,B_2' & 'q_1,B_1' \ y \end{array} \end{array} \quad \begin{array}{l} \forall B \subset A \\ \forall (B_1,B_2,B_3) \text{ partition de } B \end{array}$$

$$\begin{array}{l} 'q_1,B' \quad \begin{array}{c} a \\ | \\ x \end{array} \quad \rightarrow \quad \begin{array}{c} a \\ | \\ 'q_1,B' \ x \end{array} \quad \forall B \subset A \end{array}$$

$$\begin{array}{l} 'q_1,B' \quad \bar{a} \rightarrow a' \quad \forall B \subset A, \ \forall \begin{array}{c} \alpha_{i_1} \\ \vdots \\ \alpha_{i_j} \end{array} \in S \text{ et } B \subset \{\alpha_{i_1},\ldots,\alpha_{i_j}\} \\ \qquad\qquad\quad \begin{array}{c} \alpha_{i_1} \\ \vdots \\ \alpha_{i_j} \\ | \\ \bar{a} \end{array} \end{array}$$

$$\begin{array}{l} 'q_2,B' \rightarrow \begin{array}{c} c \\ / | \backslash \\ x \ y \ z \end{array} \quad 'q_1,B'z \quad \forall B \subset A \end{array}$$

$$\begin{array}{l} 'q_3,B' \quad \begin{array}{c} c \\ / | \backslash \\ x \ y \ z \end{array} \quad \rightarrow \quad \begin{array}{c} c \\ \begin{array}{ccc} x & x & y \\ 'q_3,B_3' & 'q_2,B_2' & 'q_1,B_1' \end{array} \end{array} \quad \forall B \subset A, \ \forall (B_1,B_2,B_3) \text{ partition de } B \end{array}$$

$$\begin{array}{l} 'q_3,B' \quad \begin{array}{c} c \\ / | \backslash \\ x \ y \ z \end{array} \quad \rightarrow \quad \begin{array}{c} b \\ \begin{array}{cc} x & y \\ 'q_1,B_1' & 'q_1,B_2' \end{array} \end{array} \quad \forall B \subset A, \ \forall (B_1,B_2) \text{ partition de } B \end{array}$$

<u>Remarque</u>.

On pourrait être tenté de croire que l'hypothèse d'inversibilité est superflue pour l'anticipation. Il n'en est rien.

Soit $\Sigma = \Delta = \{(a,1),(a',1),(\bar{a},o),(\bar{b},o)\}$ et Θ de règles $q_o \begin{array}{c} a \\ | \\ x \end{array} \rightarrow q_o x$, $q_o\bar{a} \rightarrow \bar{a}$ et $q_o\bar{b} \rightarrow \bar{b}$.

Alors $\hat{\Theta} = \{(a^n\bar{a},\bar{a}) \mid n \in N\} \cup \{(a^n\bar{b},\bar{b}) \mid n \in N\}$.

Codom $\hat{\Theta}$ est fini donc régulier. Pourtant, si on anticipe, on s'interdit de vérifier que a' n'apparait pas.
C'est l'hypothèse du lemme III.9 qui est ici en défaut.

CHAPITRE IV

EXAMEN DE DIFFERENTES NOTIONS DE TRANSDUCTIONS.

AVANTAGES DES BITRANSDUCTIONS.

1 CAS DES TRANSDUCTEURS NON GÉNÉRALISÉS.

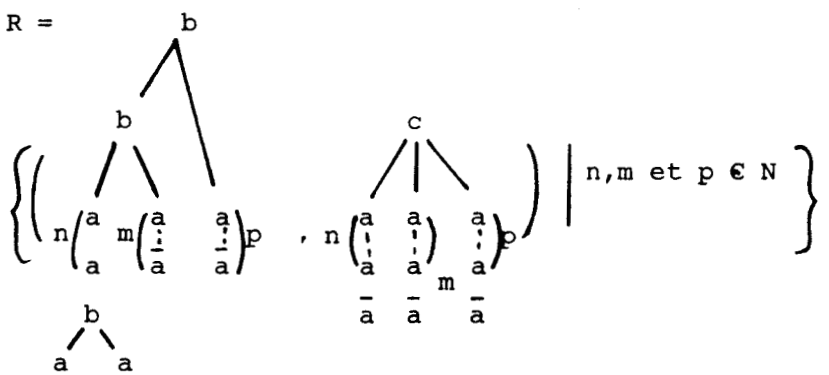
Soit $\Sigma = \{(b,2),(\bar{a},o)\}$. La relation $\left(\begin{array}{c} b \\ \bar{a}\ \bar{a} \end{array} , \bar{a} \right)$ n'est pas une $\widehat{T_D}$!

En effet, toute règle initiale serait de la forme $q_o \begin{array}{c} b \\ x\ y \end{array} \rightarrow \bar{a}$, x ou y donc non complète. Une infinité d'arbres se transformerait donc en $\bar{a}$.

Un raisonnement analogue différencie ϕ^{-1} de ψ^{-1} dans l'exemple 6 de (I 3^{eme}_B).

Notre généralisation fait disparaître ce genre de problème en autorisant des parties gauches de règles de profondeur quelconque, et toute relation finie appartient alors à $\widehat{TG_A} \cap \widehat{TG_D}$.

De plus, soit R =

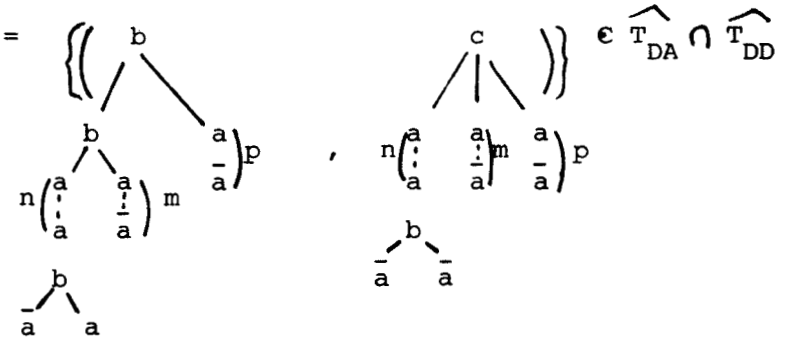


Le lecteur vérifiera que $R \notin \widehat{T_D} \cup \widehat{T_A}$.

Le fait que $R \notin \widehat{T_D}$ provient de la remarque précédent.

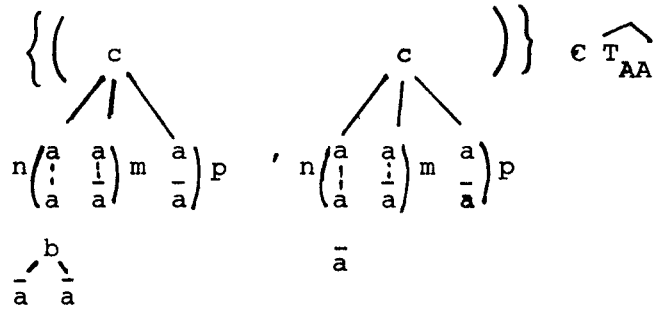
Le fait que $R \notin \widehat{T_A}$ Provient de l'incapacité des T_A de donner deux copies différentes d'une même branche (considérer la forme des dernières règles utilisées).

Mais $R_1 =$



$\in \widehat{T_{DA}} \cap \widehat{T_{DD}}$

et

$R_2 =$ 

(le lecteur construira aisément les transducteurs non généralisés associés).

Comme $R_1 \circ R_2 = R$, nous avons

$$\widehat{T}_{AA} \circ \widehat{T}_{DA} \notin \widehat{T}_D \cup \widehat{T}_A$$

ce qui ne laisse espérer aucun résultat comparable à ceux du cas généralisé.

De plus, nos bitransductions sont donc une généralisation stricte des classes usuelles. Il ne pourrait alors être possible de caractériser en termes de bi-morphismes les bitransductions non généralisées, qu'avec d'autres morphismes que ceux des TI. Nous allons voir qu'il n'en est rien.

2 AUTRES NOTIONS DE BIMORPHISMES.

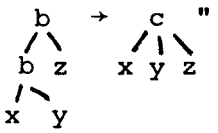
Soit M une classe de morphismes. Nous dirons qu'un bimorphisme B est de type M si B = $\langle K, \phi, \psi \rangle$ où K est une forêt régulière, ϕ et ψ deux morphismes de la classe M.

Nous étudions rapidement ici les autres types que celui des T.I, dans un ordre de généralité croissant.

A] Restriction des T. I : bimorphismes du type projection.

Ce type a été étudié par Takahashi dans (TA) sans le nom de "primitive transformation".
Ce type est stable par composition, mais ne donne que des résultats intuitivement évidents.

Les transductions associées ne font que changer régulièrement les labels, sans déformer les arbres".

Une transformation de type "  "

étant impossible, $\hat{U}$ n'est par exemple équivalente à aucune transduction de ce type.

B] Généralisation des TI.

a) Définissons TJ comme la classe associée aux homomorphismes linéaires complets comme ceux des T I, mais non nécessairement stricts, cette généralisation apparemment faible bouleverse le comportement car nous montrons en deuxième partie (chapitre V) qu'au lieu de $\widehat{TI}^2 = \widehat{TI}^3$, la suite $\{\widehat{TJ}^n \mid n \in N\}$ est une hierarchie infinie de relations.

b) Définissons TK (respectivement TK') comme la classe de bimorphismes associée aux homomorphismes linéaires (respectivement linéaires strictes).

Nous n'avons pas étudié ces classes, mais conjecturons qu'elles ont un comportement identique à celui des TI pour la composition.

Néanmoins, les transformations associées sont toutes différentes de celles associées aux transducteurs, même généralisés. En effet, un arbre peut ici avoir une infinité de transformés, comme le montre l'exemple ci-dessous :

Soit $r = (F, \phi, \psi)$

$\Sigma = \{(a,2), (\bar{a},o)\}$

$\Delta = \{(\alpha,1), (\bar{\alpha},o)\}$

$\phi(a(x,y)) = \alpha(x)$

$\phi(\bar{a}) = \bar{\alpha}$

$\psi(a(x,y)) = \alpha(y)$

$\psi(\bar{a}) = \bar{\alpha}$

$F = \Sigma^*$

Mais $\hat{F} = \{(\alpha^n\bar{\alpha}, \alpha^p\bar{\alpha}) \mid n,p \in N\}$ donc tout arbre se transforme en Δ^* par $\hat{F}$ ou $\hat{F}^{-1}$!

c) Définissons la classe TM des bimorphismes associés aux homomorphismes quelconques.

Soit alors par exemple $r = (F,\phi,id)$ avec

$\Sigma = \{(\alpha,1)\ (b,2),\ (a,1),(\bar{a},o)\}$

$F = \alpha(\{b,a,\bar{a}\}^*)$

$\phi(\alpha(x)) = b(x,x)$

$\phi(b(x,y)) = b(x,y)$

$\phi(a(x)) = a(x)$ et $\phi(\bar{a}) = \bar{a}$

id est l'identité.

alors

$\hat{r} = \{(b(t,t),\ \alpha(t)) \mid t \in \{b,a,\bar{a}\}^*\}$

Supposons qu'il existe un transducteur Θ équivalent à $\hat{r}$. Θ doit vérifier l'égalité d'arbres arbitrairement longs tout en recopiant ; il a un domaine ni régulier, ni algébrique ([AR3]). Les TM n'ont ainsi plus aucun rapport avec les transducteurs d'états finis.

En conclusion, cette critique fait apparaître le rôle privilégié que jouent les bitransductions du point de vue des transducteurs d'états finis d'arbres. Les bitransductions sont stables par composition et inversion. Elles conservent les forêts régulières (partie 2, chapitre V). Nous conjecturons que dans un "bon cadre" chacune sera équivalente à un seul bimorphisme et conserveront les forêts algébriques. Nous pouvons ainsi espérer leur voir jouer dans les forêts un rôle aussi fécond que celui des transductions rationnelles de M. NIVAT.

DEUXIEME PARTIE

DEMONSTRATIONS

CHAPITRE I

CARACTERISATION DES TRANSDUCTEURS INVERSIBLES PAR BIMORPHISMES

Définition : *On appelle branche pendante toute branche dont la feuille est une variable.*

Définissons les transducteurs quasi-généralisés de profondeur i (TG_{Q_i}) comme étant les TG_D dont chaque règle a en partie gauche toutes les branches pendantes de profondeur i.

TG_Q désignera la classe des TG_{Q_i} tels que i=1.

Posons $TG_{Q_n} = \bigcup_{i=1}^{\infty} TG_{Q_i}$.

Par définition on a :

Proposition 1.1 :

$$TA \subset TG_A$$

$$TD \subset TG_Q \subset TG_{Q_n} \subset TG_D.$$

Proposition 1.2 :

Si τ *est un transducteur généralisé, et* F *une forêt régulière, il existe un transducteur généralisé* τ' *de la même classe que* τ *tel que* $\hat{\tau}' = \{(t,t') \in \hat{\tau} \mid t' \in F\}$.

En effet. On fait la démonstration pour un TG_D ; on procèderait exactement de la même façon dans le cas d'un TG_A, TG_{Q_n}.

Soit $f = \langle Q, Q_f, \alpha \rangle$ un automate déterministe reconnaissant F. Soit Q' l'ensemble des états de τ et Q'_o l'ensemble de ses états initiaux.

L'ensemble des états de τ' est $Q' \times Q$; l'ensemble de ses états initiaux est $Q'_o \times Q_f$ et ses règles sont définies par :

$$\langle q',q\rangle\ [t(x_1,\ldots,x_n)] \to t'(\langle q'_{i_1},q_{i_1}\rangle\ [x_{i''_1}],\ \ldots,\ \langle q'_{i_p},q_{i_p}\rangle\ [x_{i''_p}])$$

est une règle de τ' ssi

$$q'\ [t(x_1,\ldots,x_n)] \to t'(q'_{i_1}\ [x_{i''_1}],\ \ldots,\ q'_{i_p}\ [x_{i''_p}])$$

est une règle de τ et $q = \alpha_{t'}(q_{i_1},\ldots,q_{i_p})$.

Il est clair par construction que $(t,t') \in \hat{\tau}'$ implique $(t,t') \in \hat{\tau}$ et $t' \in F$. On démontre la réciproque en montrant par induction sur les dérivations dans τ que si $q'[t] \overset{*}{\underset{\tau}{\Longrightarrow}} t'$ alors $\langle q', \alpha_{t'} \rangle [t] \overset{*}{\underset{\tau'}{\Longrightarrow}} t'$.

De plus, par construction, si τ est finiment réducteur, τ' l'est aussi.

cqfd.

Proposition I.3 :

Si τ est un transducteur inversible

$$\hat{\tau}^{-1} = \hat{\tau^{-1}}.$$

En effet. On démontre facilement par induction sur les dérivations que $q[t] \overset{*}{\underset{\tau}{\Longrightarrow}} t'$ ssi $t' \overset{*}{\underset{\tau^{-1}}{\Longrightarrow}} q[t]$.

cqfd.

Si τ est un transducteur inversible descendant (resp. ascendant) on note $\tau^{\times}$ le transducteur inversible descendant (resp. ascendant) obtenu en remplaçant chaque règle

$$q[t(x_1,\ldots,x_n)] \to t'(q_{i_1}[x_{\nu(1)}], \ldots, q_{i_n}[x_{\nu(n)}])$$

par $\quad q[t'(x_1,\ldots,x_n)] \to t(q_{i'_1}[x_{\nu^{-1}(1)}],\ldots,q_{i'_n}[x_{\nu^{-1}(n)}])$

avec $i'_j = i_{\nu^{-1}(\nu^{-1}(j))}$

respectivement : $t(q_{i_1}[x_1],\ldots,q_{i_n}[x_n]) \to q[t'(x_{\nu(1)},\ldots,x_{\nu(n)})]$

par $\quad t'(q_{i_{\nu(1)}}[x_1],\ldots,q_{i_{\nu(n)}}[x_n]) \to q[t(x_{\nu^{-1}(1)},\ldots,x_{\nu^{-1}(n)})]$

Par construction $(\tau^{\times})^{\times} = \tau$.

Proposition I.4 :

Si τ est un transducteur inversible

$$\hat{\tau^{\times}} = \hat{\tau}^{-1}.$$

En effet. On démontre par induction sur les dérivations que $q[t] \overset{*}{\underset{\tau}{\Longrightarrow}} t'$ ssi $q[t'] \overset{*}{\underset{\tau^{\times}}{\Longrightarrow}} t$ (resp. $t \overset{*}{\underset{\tau}{\Longrightarrow}} q[t']$ ssi $t' \overset{*}{\underset{\tau^{\times}}{\Longrightarrow}} q[t]$)

cqfd.

Corollaire I.5 : *Si* $\hat{\tau} \in \widehat{TI}_A$ *(resp.* $\widehat{TI}_D$*) alors*

$$\hat{\tau}^{-1} \in \widehat{TI}_D \text{ (resp. } \widehat{TI}_A).$$

Immédiat d'après la proposition I.3

Corollaire I.6 : *Si* $\hat{\tau} \in \widehat{TI}_A$ *(resp.* $\widehat{TI}_D$*) alors*

$$\hat{\tau}^{-1} \in \widehat{TI}_A \text{ (resp. } \widehat{TI}_D).$$

Immédiat d'après la proposition I.4.

Corollaire I.7 : $\widehat{TI}_A = \widehat{TI}_D$.

En effet. Si $\hat{\tau} \in \widehat{TI}_A$ (resp. $\widehat{TI}_D$) alors d'après le corollaire I.6 $\hat{\tau}^{-1} \in \widehat{TI}_D$ (resp. $\widehat{TI}_A$) et d'après le corollaire I.7 $\hat{\tau} = (\hat{\tau}^{-1})^{-1} \in \widehat{TI}_D$ (resp. $\widehat{TI}_A$).

cqfd.

On notera $TI = TI_A \cup TI_D$. D'après le corollaire I.7, $\widehat{TI} = \widehat{TI}_A = \widehat{TI}_D$ et d'après le corollaire I.7, $\hat{\tau} \in \widehat{TI} \Longrightarrow \hat{\tau}^{-1} \in \widehat{TI}$.

Pour $X, Y \in \{A, D, \bar{D}, Q_n\}$, on définit $\widehat{TG}_{XY} = \{\hat{\tau} \mid \hat{\tau} \in \widehat{TG}_X \text{ et } \hat{\tau}^{-1} \in \widehat{TG}_Y\}$.

Proposition I.8 : $\widehat{TI} \subset \widehat{TG}_{XY}$ *pour* $X = A, \bar{D}$; $Y = A, \bar{D}$.

Soit $\hat{B}$ la classe des relations associées aux bitransductions.

Théorème I.9 : $\hat{B} = \widehat{TI}$.

En effet.

A. Soit $\mu = \langle F, \phi, \psi \rangle$ un bimorphisme. Soit $A = \langle Q, Q_f, \alpha \rangle$ un automate déterministe reconnaissant F. On construit un transducteur inversible descendant τ dont l'ensemble des états est Q, l'ensemble des états initiaux est Q_f et dont l'ensemble des règles est défini par

$$q[t(x_1,\ldots,x_n)] \to t'(q_{i_1}[x_{\nu(1)}],\ldots,q_{i_n}[x_{\nu(n)}])$$

est une règle de τ ssi il existe un symbole σ de degré n tel que

- $\phi(\sigma(x_1,\ldots,x_n)) = t(x_{\nu_\sigma(1)},\ldots,x_{\nu_\sigma(n)})$
- $\psi(\sigma(x_1,\ldots,x_n)) = t'(x_{\nu'_\sigma(1)},\ldots,x_{\nu'_\sigma(n)})$
- $\alpha_\sigma(q_{i_{\nu'^{-1}_\sigma(1)}},\ldots,q_{i_{\nu'^{-1}_\sigma(n)}}) = q$
- $\nu(i) = \nu'_\sigma(\nu'^{-1}_\sigma(i))$ pour $i=1,\ldots,n$.

Il est facile de démontrer par induction que $q[t] \overset{*}{\underset{\tau}{\Longrightarrow}} t'$ ssi $\exists$ u tel que $\alpha_u = q$, $t = \phi(u)$; $t' = \psi(u)$.

On en déduit que $\hat{\tau} = \hat{\mu}$ et que $\hat{B} \subset \widehat{TI}_D = \widehat{TI}$.

B. Soit τ un transducteur inversible descendant. A chaque règle $r = q\,[t(x_1,\ldots,x_n)] \to t'(q_{i_1}[x_{\nu(1)}],\ldots,q_{i_n}[x_{\nu(n)}])$ de τ

on associe (i) le symbole $\bar{r}$ de degré n

(ii) la relation $\alpha_{\bar{r}}(q_{\nu^{-1}(1)},\ldots,q_{\nu^{-1}(n)}) = q$

(iii) $\phi(\bar{r}(x_1,\ldots,x_n)) = t(x_1,\ldots,x_n)$

(iv) $\psi(\bar{r}(x_1,\ldots,x_n)) = t'(x_{\nu(1)},\ldots,x_{\nu(n)})$

L'ensemble des relations (ii) définit une automate sur l'alphabet $\{\bar{r} \mid r \text{ est une règle de } \tau\}$. Soit F la forêt régulière reconnue par cet automate dont les états terminaux sont les états initiaux de τ. De la même façon que précédemment, on démontre que $q[t] \overset{*}{\underset{\tau}{\Longrightarrow}} t'$ ssi $\exists$ u $\in$ F tel que $\phi(u) = t$ et $\psi(u) = t'$ ce qui entraîne que $\widehat{TI}_D = \widehat{TI} \subset \hat{B}$.

cqfd.

Remarque. L'homomorphisme ϕ construit dans B. est ordonné. On en déduit que si une relation appartient à $\widehat{TI} = \hat{B}$, elle est associée à un bimorphisme $\langle F,\phi,\psi\rangle$ où l'un des homomorphismes est ordonné.

CHAPITRE II

CARACTERISATION DES BITRANSDUCTIONS DU TYPE ASCENDANT

Nous établissons deux lemmes qui nous seront utiles dans toute la suite et aboutissons au théorème II.4.

NOTATIONS :

Nous noterons $\pi(t)$ la profondeur d'un arbre t.

Soient T un transducteur, $R(T)$ l'ensemble de ses règles.

Pour toute $r \in R(T)$ nous noterons $g(r)$ (resp. $d(r)$) l'arbre apparaissant en partie gauche (resp. droite) de r.

Nous posons $\pi_g(T) = \sup_{r \in R} \pi(g(r))$

et $\pi_d(T) = \sup_{r \in R} \pi(d(r))$.

Définition : *Une transduction* $\hat{T} \subset \Sigma^* \times \Delta^*$ *est fidèle ssi* $\forall u \in \Delta^*$, *Card* $\{t \mid (t,u) \in \hat{T}\} < +\infty$.

Un transducteur T *est fidèle ssi* $\hat{T}$ *est fidèle.*

Lemme II.1 : *Lemme de fidélité.*

Toute bitransduction est fidèle.

En effet : Soit T un transducteur. Par définition, toute règle r vérifie $\pi_g(r) > 0$. On en déduit que :

$$u \overset{*}{\underset{T}{\Longrightarrow}} t \text{ entraîne } \pi(t) \leq \pi_d(T) \times \pi(u)$$

donc pour tout u, Card $\{t \mid (u,t) \in \hat{T}\} < +\infty$.

Soit P une bitransduction :

$\exists$ T $\in$ TG tel que $P^{-1} = \hat{T}$

d'après ce qui précède, nous avons pour tout u,

Card $\{t \mid (u,t) \in P^{-1}\} < +\infty$

donc $\forall u \in \Delta^*$, Card $\{t \mid (t,u) \in P\} < +\infty$.

cqfd.

Définition : *Un transducteur* T *a la propriété* L *ssi*

" $\exists$ $t(x)$, $t'(x)$, t'' $\in \Sigma^*(X)$

$\exists$ $u(x_1,\ldots,x_m) \in \Delta^*(X)$ avec $m \geq 2$

$\forall$ $n \in N$, $\exists$ $u_n^1,\ldots,u_n^m \in \Delta^*$

tels que :

$$\forall\ i = 1,\ldots,m,\ \lim_{n\to\infty} \pi(u_n^i) = +\infty$$

et $\forall\ n \in N,\ (u_n,t_n) \in \hat{T}$

où $u_n = u(u_n^1,\ldots,u_n^m)$

et $t_n = t(t'^n(t''))$ "

Définition : *Soient* T *un* TG_A, $\delta : t \overset{*}{\underset{T}{\Longrightarrow}} q_o[u]$ *une dérivation de* T.

δ' *est une dérivation initiale de* δ

ssi $\exists$ $t'(x_1,\ldots,x_n) \in \Sigma^*(X)$, $q_{i_1},\ldots,q_{i_n} \in Q$,

$u_1,\ldots,u_n \in \Delta^*$ *tels que*

$$\delta' : t \overset{*}{\underset{T}{\Longrightarrow}} t'(q_{i_1}[u_1],\ldots,q_{i_n}[u_n])$$

et $$\delta : t \overset{*}{\underset{T}{\Longrightarrow}} t'(q_{i_1}[u_1],\ldots,q_{i_n}[u_n]) \overset{*}{\underset{T}{\Longrightarrow}} q_o[u]$$

Nous noterons $\delta' \underset{I}{\leq} \delta$.

Lemme II.2 : *Lemme de linéarité globale.*

Si $T \in TG_{YX}$, *alors* T *n'a pas la propriété* L.

A : CAS DE $T \in TG_{AX}$:

En effet, supposons que T ait la propriété L. $\forall$ $n \in N$, il existe $\delta_n : u_n \overset{*}{\underset{T}{\Longrightarrow}} q_o[t_n$

Il est facile de voir que pour tout n, il existe $\delta'_n \underset{I}{\subseteq} \delta_n$ telle que :

$$\forall\ i = 1,\ldots,m \quad \exists\ \bar{u}^i_n(x_1,\ldots,x_{m'_i}) \lesssim_I u^i_n$$

$$\exists\ t^i_{n,1},\ldots,t^i_{n,m'_i} \in \Sigma^*$$

$$\exists\ q^i_{i_{n,1}},\ldots,q^i_{i_{n,m'_i}} \in Q$$

tels que δ'_n s'écrira comme suit

$$\delta'_n\ u(u^1_n,\ldots,u^m_n) \underset{T}{\overset{*}{\Longrightarrow}} u(\bar{u}^1_n(q^1_{i_{n,1}}[t^1_{n,1}],\ldots,q^1_{i_{n,m'_1}}[t^1_{n,m'_1}]),\ldots$$

$$\ldots,\ \bar{u}^m_n(q^m_{i_{n,1}}[t^m_{n,1}],\ldots,q_{i_{n,m'_m}}[t^m_{n,m'_n}]))$$

avec $\forall\ n,i,\ \pi(\bar{u}^i_n) \lesssim \pi_g(T)$.

Intuitivement, δ'_n est obtenue en commençant à dériver comme dans δ_n et en s' "arrêtant juste avant de dériver $u(x_1,\ldots,x_m)$"

Notons $\delta_n - \delta'_n$ la dérivation complétant δ'_n en δ_n.
$\pi(u(\bar{u}^1_n(x_1,\ldots,x_{m'_1}),\ldots,\bar{u}^m_n(x_p,\ldots,x_q))) \lesssim \pi(u) + \pi_g(T)$.

Cette inégalité, jointe à la remarque que $\forall\ r \in R(T)\ \pi_g(r) > 0$, nous permet de majorer la longueur de δ_n-δ'_n (id. le nombre de règles utilisées dans δ_n-δ'_n) par un entier M indépendant de n.

Donc, pour une infinité $(n_i)_{i\in N}$ d'indices n, δ'_{n_i} se termine par la même δ_{n_i}-δ'_{n_i}. Plus précisément, il existe $\bar{t}(x_1,\ldots,x_{p'}) \in \Sigma^*(X)$, $\exists\ u^1(x_1,\ldots,x_{m'_1}),\ldots,\bar{u}^m(x_1,\ldots,x_{m'_m}) \in \Delta^*(X)$ $\exists\ q^1_{i_1},\ldots,q^1_{i_{m'_1}},\ldots,q^m_{i_1},\ldots,q^m_{i_{m'_m}} \in Q$
il existe δ : $u(\bar{u}^1(q^1_{i_1}[x_1],\ldots,q^1_{i_{m'_1}}[x_{m'_1}]),\ldots,\bar{u}^m(q^m_{i_1}[x_p]\ldots))$

$$\underset{T}{\overset{*}{\Longrightarrow}} q_o[t(x_{j_1}\ \ldots\ x_{j_{p'}})]$$

tels que, pour tout i $\in$ N,

δ_{n_i} - δ'_{n_i} soit obtenue à partir de δ

en substituant $t^m_{n,1}$ à $x_1,\ldots,$ $t^m_{n_i,m'_{n_i}}$ à x_p.

On a évidemment

$$\forall\, i \in N,\ \bar{t}(x_{j_1},\ldots,x_{j_{p'}}) \underset{I}{\leq} t_{n_i} = t(t'^{n_i}(t")).$$

On peut toujours supposer $\pi(t'(x)) > 0$, sinon le lemme de fidélité serait contredit.

Alors, $\exists\, i_o \in N$ tel que pour $i \geq i_o$, $\bar{t}(x_1,\ldots,x_{p'}) \underset{I}{\leq} t(t'^{n_{i_o}}(x))$.

La restriction $\bar{b}$ à $\bar{t}(x_1,\ldots,x_p)$ de la branche de $t(t'^{n_i}(x))$ ayant x pour feuille est indépendante de i. Soit donc x_k la feuille de $\bar{b}$ et soit ℓ l'indice tel que x_k soit feuille de $\bar{u}^{\ell}(x_i,\ldots,x_j)$.

Comme en terminant δ_{n_i}, les arbres $t^{n"}_{n_i,n'}$ se substituent aux x_j dans $\bar{t}$, pour tout $i \in N$, pour tout $n" \neq \ell$, $t^{n"}_{n_i,n'}$ est un sous-arbre final de t ou de t'.

Il est facile de vérifier alors que : $\forall\, i \in N, \exists\, \hat{t}_{n_i} \in \Sigma^*$

$$\hat{u}_{n_i} = u(u^1_{n_i}, u^{\ell-1}_{n_i}, u^{\ell}_{n_{i_o}}, u^{\ell+1}_{n_i}, \ldots, u^m_{n_i}) \overset{*}{\Longrightarrow} q_o[\bar{t}_{n_i}]$$

avec $\pi(t_{n_i}) \leq \pi(t_{n_{i_o}}) + \pi(t(x)) + \pi(t'(x))$.

Comme $m \geq 2$, les hypothèses de L sur $\pi(u^i_n)$ montrent que

$$\lim_{i \to +\infty} \pi(\hat{u}_{n_i}) = +\infty.$$

Donc une infinité d'arbres se dérivent en un nombre fini, ce qui contredit la fidélité.

cqfd.

B : CAS DE TG_{DX} :

En vue de la section III, nous allons montrer un résultat légèrement plus fort :

Définition : *T a la propriété* L' *ssi*

" $\exists\, G \subset \Sigma^*(X)$, G fini, $\forall\, t' \in G$, t *a au plus 1 variable*

$\exists\, u(x_1,\ldots,x_m) \in \Delta^*(X)$ avec $m \geq 2$

$\forall\, n \in N, \exists\, u^1_n, \ldots, u^m_n \in \Delta^*$

tels que

$$\forall\ i = 1,\ldots,m \quad \lim_{n\to\infty} \pi(u_n^i) = +\infty$$

et $\forall\ n \in N,\ (u_n, t_n) \in \hat{T}$

où $u_n = u(u_n^1,\ldots,u_n^m)$

et $t_n = t'_{n_1}(\ldots(t'_{n_{\ell_n}})\ldots)$ *avec* $\forall\ i=1\ldots\ell_n,\ t'_{n_i} \in G''$

il est évident qu'en posant $G = \{t(x), t'(x), t''\}$, L implique L'. Donc $\neg L' \Longrightarrow \neg L$.

Montrons que si $T \in TG_{DX}$, T n'a pas la propriété L'.

En effet :

Les étapes de la démonstration sont parallèles à celles du cas A.

On définit de façon évidente les notions de dérivation utile et de dérivation initiale dans le cas descendant.

Supposons que T ait la propriété L'.

Pour tout n, $\exists\ \delta_n : q_o[u_n] \overset{*}{\underset{T}{\Longrightarrow}} t_n$.

Il est facile de voir que :

$\forall_n$, $\exists\ \delta'_n \leq \delta_n$ telle que :

$$\delta'_n : q_o[u(u_n^1,\ldots,u_n^m)] \overset{*}{\Longrightarrow} \tilde{t}_n(qn_{i_1}^{j}[\Theta_{i_1}^{j_1} n],\ldots\ qn_{i_{p_{j_1}}}^{j_1}[\Theta_{i_{p_{j_1}}}^{j_1} n],\ \ldots$$
$$qn_{i_1}^{j_v}[\Theta_{i_1}^{j_v} n],\ldots,qn_{i_{p_{j_v}}}^{j_v}[\Theta_{i_{p_{j_v}}}^{j_v} n])$$

où $\forall\ k=1,\ldots,v.\ j_k \in \{1,\ldots,m\}$ et $\exists\ \bar{\Theta}_{j_k}^n$ tel que :

$$u_n^{j_k} = \bar{\Theta}_{j_k} n(\Theta_{i_1}^{j_k} n,\ldots,\Theta_{i_{p_{j_k}}}^{j_k} n)$$

avec $\pi(\bar{\Theta}_{j_k} n) \leq \pi_g(T)$.

Ceci signifie que δ'_n "commence comme δ_n et s'arrête dès qu'on a transducté u".

Le nombre de règles appliquées dans δ'_n est borné indépendamment de n et il existe une infinité d'indices pour lesquels δ'_n est une même dérivation initiale δ'. En rénumérotant cette sous-suite, on obtient :

$\exists\ \delta'$ telle que $\forall\ n$, $\delta' \underset{I}{\leq} \delta_n$

$$\text{avec } \delta'\ q_o[u(u_n^1,\ldots,u_n^m)] \overset{*}{\Longrightarrow} \bar{t}(q_{i_1}^{j_1}[\theta_{i_1}^{j_1} n],\ldots,q_{i_{p_{j_v}}}^{j_v}[\theta_{i_{p_{j_v}}}^{j_v} n])$$

et $\forall\ k = 1,\ldots v$

$$u_n^{j_k} = \bar{\theta}_{j_k}(\theta_{i_1}^{j_k} n,\ldots,\theta_{i_{p_{j_k}}}^{j_k} n)$$

Remarquons que les $\bar{\theta}_j$ sont, comme δ' et $\bar{t}$, indépendants de n mais les $\theta_j^k n$ dépendent des u_n^k.

$\forall\ n$, $\bar{t}(x_1,\ldots,x_v) \underset{I}{\leq} t_n$.

Si $\overline{\lim_{n\to\infty}}\ \pi(t_n) < +\infty$, la fidélité est, comme dans A. contredite.

Le lecteur vérifiera facilement que pour tout n, pour tout k=1,...,v sauf au plus un, noté k_n, δ_n se continue en transductant $\theta_{i_\alpha}^{j_k} n$ en un sous-arbre d'un arbre de G (ce résultat provient de la forme de t_n).

Si il n'existe pas une infinité de k_n, $\pi(t_n)$ est bornée, ce qui contredit la fidélité.

Sinon, il existe $k_o = k_n$ pour une infinité d'indices. On termine alors comme dans A., en se fixant $u_{n_o}^{k_o}$, ce qui borne les arbres dérivés indépendamment de $u_n^{k'}$, $k' \neq k_o$ et contredit la fidélité.

cqfd.

Définition : *Soit* $T = (Q,\Sigma,\Delta,R,q_o)$ *un* TG_A.

Une dérivation initiale δ' *est dite utile ssi il existe* $\delta : t \underset{T}{\Longrightarrow} q_o[u]$ *telle que* $\delta' \underset{I}{\leq} \delta$, $t \in \Sigma^*$ *et* $u \in \Delta^*$.

Notons $\Delta(T)$ *l'ensemble des dérivations initiales de* T.

Une règle r *est utile ssi* $\exists\ \delta \in \Delta(T)$ *utilisant* r.

Un transducteur est réduit si toutes ses règles sont utiles. Il est évident que la réduction d'un transducteur ne change pas la relation associée.

Lemme II.3 : *Si* T *est un* TG_{AX}, $\hat{T} \in \widehat{TI}$.

En effet, nous allons construire $T' \sim T$ tel que toutes les règles de T' soient complètes, linéaires et non réductrices, T' sera bien, par définition, un TI.

Remarquons que les lemmes II.1 et II.2 nous assurent que $\hat{T}$ et $\hat{T}^{-1}$ sont fidèles, que $\hat{T}$ et $\hat{T}^{-1}$ n'ont pas la propriété L.

A. REDUCTION FORTE

$\forall\ q \in Q$, posons $F_{(q)} = \{(t,u) \mid t \overset{*}{\underset{T}{\Longrightarrow}} q\ [u],\ t \in \Sigma^*$ et $u \in \Delta^*\}$

et $F_{1(q)} = \{t \mid (t,u) \in F_q\}$.

Nous dirons que T est fortement réduit ssi T est réduit et $\forall\ q \neq q_o \in Q$, Card $(F_1(q)) = +\infty$.

Montrons qu'on peut toujours supposer un TG_A fortement réduit : ceci équivaut à trouver pour tout T, un T' équivalent ayant cette propriété.

Soit $T = (Q,\Sigma,\Delta,R,q_o)$.

Construisons $T' = (Q',\Sigma,\Delta,R',q_o)$.

A toute règle r : $t(q_{i_1}[x_1],\ldots,q_{i_n}[x_n]) \to q[u(x_{j_1},\ldots,x_{j_p})]$ de R, associons l'ensemble R'(r) de règles de T'.

$R'(r) = \{(t(\alpha_1,\ldots,\alpha_n) \to q[u(\beta_{j_1},\ldots,\beta_{j_p})]) \mid$ telles que,

$\forall\, k = 1,\ldots,n,\ (\alpha_k,\beta_k) = (q_{i_k}[x_k],x_k)$ ssi Card $F_1(q_{i_k}) = +\infty$ sinon, $(\alpha_k,\beta_k) \in F(q_{i_k})\}$.

On a Card $(R'(r)) < +\infty$.

T' est obtenu en prenant pour règles $\underset{r\in R}{U} R'(r)$ en réduisant et en prenant pour Q' l'ensemble des états de Q subsistant.

Il est facile de montrer que $T' \sim T$.

T' est fortement réduit par construction.

Nous considérons maintenant un TG_A, noté T, fortement réduit, fidèle et n'ayant pas la propriété L.

B. TOUTES LES REGLES DE T SONT COMPLETES :

Soit r une règle non complète. Supposons que x_j n'apparaisse pas en partie droite de r

$$r \begin{cases} t(q_{i_1}[x_1],-q_{i_j}[x_j],\ldots\ q_{i_n}[x_n]) \to q[u(x_{j_1},\ldots,x_{j_p})] \\ \text{avec pour tout } k=1,\ldots,p,\ x_{j_k} \neq x_j. \end{cases}$$

r étant utile, il existe $t_1,\ldots,t_{j-1},\ t_{j+1},\ \ldots,\ t_n \in \Sigma^*$

$\exists\ t'(x) \in \Sigma^*[X]$ et $\bar{u} \in \Delta^*$ tels que ; en posant

$$t'(t(t_1,\ldots,t_{j-1},q_{i_j}[x_j],\ t_{j+1},\ldots,t_n)) = \bar{t}(q_{i_j}[x_j]),$$

nous ayons :

$$\bar{t}(q_{i_j}[x_j]) \overset{*}{\underset{T}{\Longrightarrow}} q_o[\bar{u}]$$

D'autre part, $\{\bar{t}(\Theta) \mid (\bar{t}(\Theta),\bar{u}) \in \hat{T}\} \supset \{\bar{t}(\Theta) \mid \Theta \in F_1(q_{i_j})\}$.

Donc $\bar{u}$ est l'image d'une infinité d'arbres puisque Card $(F_1(q_{i_j})) = +\infty$, ce qui contredit la fidélité de T.

C. TOUTES LES REGLES DE T SONT LINEAIRES :

Soit r une règle non linéaire

$$t(q_{i_1}[x_1],\ldots,q_{i_j}[x_j],\ldots,q_{i_n}[x_n]) \to q[u(x_{j_1},\ldots,x_{j_p})]$$

avec $j_{i_1} = j_{i_r} = \ldots = j_{i_\alpha} = j$ et $\alpha \geq 2$.

Comme r est utile, il existe $t'(x) \in \Sigma''(X)$,

$t_1,\ldots,t_{j-1}, t_{j+1},\ldots t_n \in \Sigma^*$, $u_1,\ldots,u_{j-1},u_{j+1},\ldots,u_n \in \Delta^*$

$\bar{u}'(x_1,\ldots,x_p) \in \Delta^*(X)$ tels que,

en posant $\bar{t}'(q_{i_j}[x_j]) = t(t_1,\ldots,t_{j-1},q_{i_j}[x_j],t_{j+1},\ldots,t_n)$

et $\bar{u}(x_{j_{i_1}},\ldots,x_{j_{i_\alpha}}) = u(u_{j_1},x_{j_{i_1}},\ldots,x_{j_{i_\alpha}},u_{j_p})$

on ait $\bar{t}(q_{i_j}[x_j]) \overset{*}{\underset{T}{\Longrightarrow}} q[\bar{u}(x_{j_{i_1}},\ldots,x_{j_{i_\alpha}})]$

et, comme toutes les règles sont supposées complètes :

$$\bar{t}(q_{i_j}[x_j]) = t'(\bar{t}'(q_{i_j}[x_j])) \overset{*}{\underset{T}{\Longrightarrow}} q_o[\bar{u}'(\bar{u}(x_{j_{1_1}},\ldots,x_{j_{i_\alpha}}),\bar{u}(x_{j_{i_1}},\ldots,x_{j_{i_\alpha}})\ldots)]$$

où le nombre d'occurrences de $\bar{u}$ dans $\bar{u}'$ est $\ell \geq 1$.

Comme Card $(F_1(q_{i_j})) = +\infty$ il existe $\Theta_1(x)$, $\Theta_2(x) \in \Sigma^*(X)$, $\Theta_3 \in \Sigma^*$, $V_1(x)$, $V_2(x) \in \Delta^*(X)$ et $V_3 \in \Delta^*$, et $q \in Q$ tels que :

$\pi(\Theta_2(x)) > 0$ et

$\Theta_1(\Theta_2(\Theta_3)) \overset{*}{\underset{T}{\Longrightarrow}} \Theta_1(\Theta_2 q[V_3]) \overset{*}{\underset{T}{\Longrightarrow}} \Theta_1 q[V_2(V_3)])$

$\overset{*}{\underset{T}{\Longrightarrow}} q_{i_j}[V_1(V_2(V_3))]$ car toutes les règles sont complètes.

Supposons que $V_2(x) = x$.

Alors, $\forall\, n \in N$, $\Theta_1(\Theta_2^n(\Theta_3)) \overset{*}{\underset{T}{\Longrightarrow}} q_{i_j}[V_1(V_3)]$

ce qui contredirait la fidélité.

Donc $\pi(V_2(x)) > 0$.

Posons $\tilde{t}_n = \bar{t}(\Theta_1(\Theta_2^h(\Theta_3))$

$$u_n^1 = u_n^2 = \ldots u_n^\alpha = V_1(V_2^{(n)}(V_3))$$

et $\tilde{u}(x_{j_{i_1}},\ldots,x_{j_i}) = \bar{u}'(\bar{u}(x_{j_{i_1}},\ldots,x_{j_{i_\alpha}}),\bar{u}(x_{j_{i_1}}\ldots))$

alors $\forall\, n \in N$, $(\tilde{t}_n,\tilde{u}(u_n^1,\ldots,u_n^\alpha)) \in \hat{T}^{-1}$

d'autre part, $\exists\, U \in TG_{XA}$ tel que $\hat{U} = \hat{T}^{-1}$ car $T \in TG_{AX}$.

Donc U satisfait à la propriété L, ce qui contredit le lemme II.2.

D. T EST FINIEMENT REDUCTEUR :

Supposons T non finiement réducteur :

$\forall\, n \in N$, $\exists\, t_n(x_1,\ldots,x_{i_n})$ tel que

$\pi(t_n) \geq n$ et $\exists\, q_1^n,\ldots,q_{i_n}^n$ et $q_n \in Q$ tels que

$t_n(q_1^n[x_1],\ldots q_{i_n}^n[x_{i_n}]) \overset{*}{\underset{T}{\Longrightarrow}} q_n[x_{j_n}]$

Toutes les règles de T sont linéaires complètes d'après ce qui précède. On en déduit que $i_n = 1$ pour tout n.

Nous pouvons donc écrire :

$\forall\ n \in N$, $\exists\ t_n(x)$, $\exists\ q_n$ et $q'_n \in Q$ tels que

$$\pi(t_n(x)) \geq n \text{ et } t_n(q_n[x]) \overset{*}{\underset{T}{\Longrightarrow}} q'_n[x]$$

Il existe une infinité d'indices (n_p) $p \in N$ pour lesquels le couple (q_n, q'_n) est un même couple (q, q'). T étant réduit, il existe $t'_1 \in \Sigma^*$ et $u_1 \in \Delta^*$ tels que $t'_1 \overset{*}{\underset{T}{\Longrightarrow}} q[u_1]$.

De même, il existe $t'_2(x)$ et $u_2(x)$ tels que

$$t'_2(q'[x]) \overset{*}{\underset{T}{\Longrightarrow}} q_o[u_2(x)].$$

Nous en déduisons que $\forall\ p$, $(t'_2(t_{n_p}(t'_1)), u_2(u_1)) \in \hat{T}$.

Comme $\pi(t_{n_p}) \geq n_p$, la fidélité est contredite.

T est donc finiement réducteur. Soit p_o sa profondeur de réduction (voir partie 0).

Construisons T' en prenant pour règles les dérivations dans T de profondeur $\leq p_o+1$, contenant du moins une règle non réductrice. Il est facile de voir que $T' \sim T$.

T' a bien toutes ses règles inversibles.

cqfd.

Théorème II.4 :

$$\widehat{TI} = \widehat{TG_{AD}} = \widehat{TG_{DA}} = \widehat{TG_{AA}}.$$

En effet, le lemme II.3 montre que

$$\widehat{TG}_{AD} \subset \widehat{TI} \quad \text{et} \quad \widehat{TG}_{AA} \subset \widehat{TI}$$

$$\text{et} \quad \widehat{TG}_{DA} = \widehat{TG}_{AD}^{-1} \subset \widehat{TI}^{-1} = \widehat{TI}.$$

Les inclusions inverses proviennent de I.8.

cqfd.

CHAPITRE III

UN RESULTAT INTERMEDIAIRE SUR LES BITRANSDUCTIONS QUASI-GENERALISEES

De nombreux intermédiaires vont être introduits en vue de montrer $\widehat{TGQ_n D} \subset TI^2$.

Définition : *Soit* $P = \{t_1,\ldots,t_m\}$ *une partie finie de* $\Sigma^*(X)$.
P *est dite* compatible *ssi* $\exists\ t$ *tel que* $\forall\ i=1,\ldots,m$ *on ait* $t_i \leq_I t$. *Il est facile de voir que la borne supérieure de* P *existe alors. Nous la noterons* $\bigvee_{i=1}^{m} t_i$.

Remarquons que si $t_1 \in \Sigma^*$ et $t_2 \in \Sigma^*$, alors $t_1 \leq_I t_2 \Longleftrightarrow t_1 = t_2$.

Définition : *Soit* T *un* TGQ_n. *On appelle* chaîne *(de* T*) toute suite finie* C *de règles* $(s_1,\ldots,s_m)$ *telle que* $P(C) = \{g(s_i) \mid \{i=1,\ldots,m\}$ *soit compatible.*

Si q_i est l'état gauche de s_i, notons $Q(C)$ la suite $(q_1,\ldots,q_m)$.

Soit $t \in \Sigma^*(X)$. Nous notons t/ℓ le sous-arbre initial de t obtenu en "coupant" t à la profondeur ℓ.

Notons $\mathcal{C}(T)$ l'ensemble des chaînes de T.

$\forall\ C \in \mathcal{C}(T)$, $P(C)$ est compatible donc $\exists\ t_1 \in \Sigma^*(X)$ tel que $\forall\ i=1,\ldots,m$, $g(s_i)/n = t_1(x_1,\ldots,x_p)$.

n désigne la profondeur de T,

alors $\forall\ i = 1,\ldots,m,\ \ g(s_i) = t_1(\theta_1^i,\ldots,\theta_p^i)$

où θ_j^i est soit x_j, soit un arbre de Σ^*

donc $\bigvee_{i=1}^{m} g(s_i) = t_1(\theta_1,\ldots,\theta_p)$

avec
$$\begin{cases} \theta_j = x_j \text{ ssi } \forall\ i = 1,\ldots,m,\ \theta_j^i = x_j \\ \theta_j = \theta_j^i \text{ ssi } \theta_j^i \in \Sigma^* \text{ (comme } \theta_j^{i_1} \in \Sigma^* \text{ et } \theta_j^{i_2} \in \Sigma^* \text{ implique } \theta_j^{i_1} = \theta_j^{i_2}, \\ \qquad\qquad \theta_j \text{ est bien indépendant de i)} \end{cases}$$

En renumérotant les variables, posons $\theta(x_1,\ldots,x_q) = Vg(s_i)$. On remarque que θ est de profondeur homogène n.

Posons $r(\Theta) = \{q[\Theta(x_1,\ldots,x_q)] \rightarrow u(q_{i_1}[x_{j_1}],\ldots,q_{i_p}[x_{j_p}])$

$|q[\Theta(x_1,\ldots,x_q)] \overset{*}{\underset{T}{\Longrightarrow}} u(q_{i_1}[x_{j_1}],\ldots,q_{i_p}[x_{j_p}])$

par une dérivation dont la première règle est r}

Nous avons en fait défini de nouvelles règles en dérivant les Θ^i_j n'appartenant pas à g(r).

Définition : *On appelle chaîne homogène (c.h.) toute chaîne dont toutes les règles ont même partie gauche.*

Définissons comme suit l'ensemble $H(C)$ des c.h. associées à la chaîne C.

$$H(C) = \{(r'_1,\ldots,r'_m) \mid \forall\ i=1,\ldots,m,\ r'_i \in r_i(\Theta)\}$$

où Θ est associé à la chaîne C.

Définition : Soit $C = (r'_1,\ldots,r'_m)$ une c.h.

où $r'_i : q_{k_i}[\Theta(x_1,\ldots,x_q)] \rightarrow u_i(q_{k_{i_1}}[x_{j_{i_1}}],\ldots,q_{k_{i_{n_i}}}[x_{j_{i_{n_i}}}])$

$\forall\ \alpha = 1,\ldots,q$ posons $Q_\alpha(C) = \{q_{k_{i_\ell}} \mid x_{j_i} = x_\alpha\} \in P(Q)$

et $\forall\ Q' \in P(Q)$, posons $Q'(T) = \{t \in \Sigma^* \mid \forall\ q \in Q',\ \exists\ u \in \Delta^*,\ q[t] \overset{*}{\underset{T}{\Longrightarrow}} u\}$

Nous dirons qu'une chaîne homogène est essentielle (c.h.e.) ssi

$$\forall\ \alpha = 1,\ldots,q \quad \text{Card}\ (Q_\alpha(C)(T)) = +\infty$$

Soit $\{\alpha_1,\ldots,\alpha_{q'}\}$ l'ensemble des indices pour lesquels Card $(Q_{\alpha_i}(C)(T)) = +\infty$

et $\{\beta_1,\ldots,\beta_{q''}\}$ l'ensemble des indices pour lesquels Card $(Q_{\beta_i}(C)(T)) < +\infty$.

Définissons l'ensemble $H\varepsilon(C)$ des c.h.e. associées à la c.h.C

$$H\varepsilon(C) = \{(r''_1,\ldots,r''_m) \mid \forall\ i=1,\ldots,q'',\ \exists\ t_{\beta_i} \in Q_{\beta_i}(C) \text{ tq } \forall\ j=1,\ldots,m$$

$$r''_j = q_{i_j}[\Theta(x_{\alpha_1},\ldots,t_{\beta_1},\ldots,t_{\beta_{q''}}\ldots)] \rightarrow u_j(q_{j,j_1}[x_{j_1}]\ \ldots)$$

où cette expression, en remplaçant $\rightarrow$ par $\overset{*}{\Longrightarrow}$ est une dérivation dans T}.

Nous poserons : $\pi Q(C) = (q_{i_1},\ldots,q_{i_m})$ (suite des états gauches des règles).

Nous avons ainsi à nouveau prolongé les règles. Les règles nouvelles r" sont en nombre fini, car on a un nombre fini de Θ^i_j et de t_{β_j}, et elles sont bien du type des règles de TGQ_n (on conserve la profondeur des variables des règles d'origine).

Définition : *Soit* T *un* TGQ_n.

Définissons l'ensemble de ses blocs.

Nous définirons la *taille* V(B) et la *profondeur* $\pi_B(x)$ dans B d'une variable x d'un bloc B.

La *longueur* (B) *d'un bloc* B est le nombre de chaînes qu'il utilise.

1) Les blocs de longueur 1 sont les c.h.e. Leur taille est la longueur de la chaîne, et pour toute x, $\pi_B(x) = 1$.

2) Tous les blocs sont obtenus par composition ainsi définie :

Soient B et B' deux blocs. Nous les noterons

$$B(q_i[t] \overset{*}{\underset{b_i}{\Longrightarrow}} u_i)_{i\in I} \text{ et } B'(q'_j[t'] \overset{*}{\underset{b'_j}{\Longrightarrow}} u'_j)_{j\in J}.$$

Explicitons B et B'.

B est la suite $(b_1,\ldots,b_{V(B)})$

où pour i=1...V(B) ;

$$b_i = q_i[t(x_1,\ldots x_m)] \overset{*}{\Longrightarrow} u_i(q_{k_{i_1}}[x_{k'_{i_1}}],\ldots,q_{k_{i_{\xi_i}}}[x_{k'_{i_{\xi_i}}}])$$

B' est la suite $(b'_1,\ldots,b'_{V(B')})$

où pour $j=1,\ldots,V(B'))$

$$b'_j = q'_j[t'(x'_1,\ldots,x'_{m'})] \overset{*}{\Longrightarrow} u'_j(q'_{k_{j_1}}[x'_{k'_{j_1}}],\ldots,q'_{k_{j_{\xi_j}}}[x'_{k'_{j_{\xi_j}}}])$$

posons $Q(B) = (q_1,\ldots,q_{V(B)})$ et $g(B) = t(x_1,\ldots,x_m)$.

$V(B) = \{x_1,\ldots,x_m\}$ désigne l'ensemble des variables de B.

Remarquons qu'étant donnés deux blocs B et B' on peut toujours à un changement de notations près supposer $V(B) \cap V(B') = \emptyset$.

$\forall\, x_i \in V(B)$, nous associons la *suite d'états dans* B *relative à* x_i, notée $Q_{x_i}(B)$ et définie comme suit

$$Q_{x_i}(B) = (q_{k_{1_{\alpha_1}}},\ldots,q_{k_{1_{\alpha_{p_1}}}},q_{k_{2_{\alpha_1}}}\ldots q_{k_{j_{\alpha_{p_j}}}},q_{k_{(j+1)_{\alpha_1}}}\ldots,q_{k_{m_{\alpha_{p_m}}}}) \,|\, x_{k'_{j_{\alpha_\beta}}} = x_i$$

autrement dit, $Q_{x_i}(B)$ est la suite des états associés à x_i en partie droite de B, en lisant de gauche à droite dans u_1, puis u_2 ... puis $u_{V(B)}$.

Nous dirons que B' *est composable avec* B *en* x_i ssi $Q_{x_i}(B) = Q(B')$.

Le composé $B_{x_i}(B')$ est alors le bloc B" défini comme suit

$$B'' = (q_i[t''] \overset{*}{\underset{b''_i}{\Longrightarrow}} u''_i)_{i\in I} \quad \text{avec} \quad V(B'') = V(B),$$

$$t(x_1,\ldots,x_{i-1},t'(x'_1,\ldots,x'_{m'}),x_{i+1},\ldots x_m) = t''$$

et $\forall$ i = 1,...V(B")

$b''_i : q_i[t''] \overset{*}{\Longrightarrow} u_i(U_{k_{i_1}}, \ldots, U_{k_{i_{\xi_i}}}) = u''_i$

où $U_{k_{i_j}} = q_{i_j}[x_{k'_{i_j}}]$ ssi $x_{k'_{i_j}} \neq x_i$

et $U_{k_{i_j}} = u'_\gamma$ ssi $q_{k'_{i_j}}$ est le $\gamma^{\text{ème}}$ élément de $Q_{x_i}(B)$.

On a alors $V(B'') = \{x_1, \ldots, x_{i-1}, x'_1, \ldots x'_m, x_{i+1}, \ldots, x_m\}$

$\forall\, x_j \in V(B''),\ \pi_{B''}(x_j) = \pi_{B'}(x_j)$

$\forall\, x'_j \in V(B''),\ \pi_{B''}(x'_j) = \pi_B(x_i) + \pi_{B'}(x'_j)$

$$\ell(B'') = \ell(B) + \ell(B').$$

Comme pour les règles, nous numérotons les variables de manière qu'une même x_i n'apparaisse pas deux fois dans $g(B)$.

Définition : *Soient* $t \in \Sigma^*(X)$ et $u \in \Delta^*(X)$, *soit* T' *un* TGQ_n.

Nous dirons que t *se dérive en* u *par bloc dans* T', que nous noterons $q_0[t] \overset{*}{\underset{bT'}{\Longrightarrow}} u$ ssi il existe un bloc B de taille 1 de T' tel que

$$B : (q_0[t] \overset{*}{\underset{b}{\Longrightarrow}} u).$$

Nous dirons que T' *est homogène* ssi $\forall\, (t,u) \in \hat{T}'$, il existe une dérivation par bloc de t en u.

Lemme III.1 *Soit* T *un* TGQ_n. *Il existe* $T' \in TGQ_n$ *tel que* $\hat{T}' = \hat{T}$ *et* T' *homogène.*

En effet : A tout T, nous associons T' ayant pour règles les règles r" définies dans les $H\varepsilon(C)$. (voir page IL19). $\forall\, (t,u) \in \hat{T}$, par construction des r", on peut trouver dans T' une dérivation par bloc de t en u. Le lecteur vérifiera aisément dans le détail.

cqfd.

Nous ne considérerons plus dans ce chapitre que des TGQ_n homogènes (TGQ_nH) et des dérivations par bloc.

Nous poserons pour tout $T \in TGQ_n H$

$\mathcal{B}(T)$ = {blocs utiles}. L'utilité est conservée par composition (propriété élémentaire des dérivations)

$\mathcal{B}_I(T)$ = {blocs initiaux utiles}.

Ces deux notions se définissant de façon évidente.

Nous dirons qu'un bloc $B : (q_i[t] \overset{*}{\underset{b_i}{\Longrightarrow}} u_i)_{i \in I}$ est *extrait* de B' $(q'_i[t] \overset{*}{\underset{b'_i}{\Longrightarrow}} u'_i)_{i \in I'}$, ssi la suite $(b_1,\ldots,b_{V(B)})$ est une sous suite de $(b'_1,\ldots,b'_{V(B')})$.

D'une façon générale, si $(x_1,\ldots,x_n) \in V(B)$ et si on compose B_i avec B en x_i, nous noterons $B_{x_1,\ldots,x_n}(B_1,\ldots,B_n)$ le bloc $(((B_{x_1}(B_1))_{x_2} B_2\ldots)_{x_n} B_n)$ obtenu, ou même $B(B_1,\ldots,B_n)$.

Définition : *Soit* $B \in \mathcal{B}(T)$ *et* $x_i \in V(B)$.

On appelle *émondé de* B *selon* x_i le bloc $E_{x_i}(B)$ tel que

$$g(E_{x_i}(B)) = t_1(t_2(\ldots t_n(x_1,\ldots,x_n)\ldots))$$

où
- $x_i \in V(t_n)$
- $t_1,\ldots,t_n$ sont des arbres gauches de c.h.e. (c'est-à-dire de blocs de longueur 1)
- $\exists\ B_1,\ldots,B_p \in \mathcal{B}(T)$ tels que

$$B = E(B)(B_1,\ldots,B_p)$$

Il est évident que $E_{x_i}(B)$ est bien défini de façon unique.

Sous-Lemme III.1.1 $\forall\ B \in \mathcal{B}(T)$ et $x_i \in V(B)$, $E_{x_i}(B)$ *vérifie les propriétés suivantes évidentes.*

1) $\pi_{E_{x_i}(B)}(x_i) = \pi_B(x_i)$

2) $\forall\ x_j \in t_n(x_1,\ldots,x_n)$, $E_{x_j}(B) = E_{x_i}(B)$

3) $Q_{x_i}(B) = Q_{x_i}(E_{x_i}(B))$

4) posons : $G = \{g(C) \mid C \text{ est une c.h.e.}\}$.

G est *fini* et $\forall$ B et x_i, $\exists$ $t_{i_1},\ldots,t_{i_n} \in G$ tels que

$g(E_{x_i}(B)) = t_{i_1}(\ldots t_{i_n}\ldots))$

5) Si $E_1 \in \mathcal{B}(T)$ est émondé selon x_1, E_2 selon x_2, $E_{1x_1}(E_2)$ est émondé selon x_2

6) Si $B = (q_i[t] \overset{*}{\Longrightarrow} u_i)_{i\in I}$ et $E_x(B) = (q_i[t'] \overset{*}{\Longrightarrow} u'_i)_{i\in I}$

$\forall$ i $\in$ I, $\pi_{u_i}(x) = \pi_{u'_i}(x)$.

Lemme III.2 *Soit* T *un* TGQ_nDH.

L'assertion suivante est fausse

"$\forall$ p $\in$ N, $\exists$ $B_p(q_i[t_p] \overset{*}{\underset{b_i^p}{\Longrightarrow}} u_i^p)_{i\in I_p} \in \mathcal{B}(T)$, *émondé, tel que* $\exists$ j_1^p et j_2^p *différents* $\in$ I, $\pi(u^p_{j_1^p}) \geq p$ et $\pi(u^p_{j_2^p}) \geq p$ ".

(voir démonstration intuitive en commentaire après le cqfd).

En effet : Supposons l'assertion exacte.

A] Soit $S = (q_1,\ldots,q_n)$ une suite finie.

Notons [S] le support de S, défini comme suit

$$[S] = \{q \mid \exists\ i \text{ tq } q_i \in S\}.$$

$\forall$ $r = (q_{j_1},q_{j_2},Q') \in Q \times Q \times P(Q)$, posons $\mathcal{B}_r = \{B \in \mathcal{B}(T) \mid (q_{j_1},q_{j_2})$ sous suite de $Q(B)$ et $[Q(B)] = Q'\}$.

Soit $Z = \{r \mid \mathcal{B}_r \neq \emptyset\}$ alors $\forall$ r $\in$ Z, l'entier suivant est bien défini :

$$\pi_r = \inf_{B\in\mathcal{B}_r} \{\pi_{B'}(x) \mid B' \in \mathcal{B}_I(T) \text{ et } B'_x(B) \text{ défini et utile}\}$$

et posons $\pi_1 = \sup_{r\in Z} \pi_r$

comme Z est fini, $\pi_1 \in N$.

Sous-lemme III.2.1

$\forall$ B, $\forall$ B' = $(b'_1,\ldots,b'_{V(B')})$ *composable avec* B *en* x

$\forall$ B" *et* x" $\in$ V(B"), *avec* B, B' *et* B" $\in \mathcal{B}(T)$

alors $[Q_{x"}(B")] = [Q_x(B)]$ *entraîne qu'il existe* B"' = $(b"'_1, \ldots, b"'_{V(B')})$ $\in \mathcal{B}(T)$, *composable avec* B" *en* x" *tel que le support de la suite* $(b"'_1,\ldots,b"'_{V(B"')})$ *soit le même que celui de* $(B'_1,\ldots,b'_{V(B')})$.

En effet. Ce lemme est évident en remarquant que les b'_i comme les $b"'_j$ sont définis à partir de l'ensemble de dérivations de T : $\{q[t] \overset{*}{\Longrightarrow} u \mid q \in [Q_x(B)]$ et $\forall$ q' $\in [Q_x(B)]$, $\exists$ u' $\in \Delta^*$ tel que $q'[t] \overset{*}{\Longrightarrow} u'\}$ *qui ne dépend que de* $[Q_x(B)]$.

cqfd.

Il est alors facile de voir que $\forall$ p $\in$ N, à B_p (de l'énoncé de l'assertion) nous pouvons associer B'_p et $B"_p$ tels que :

$B'_{px_p}(B"_p) \in \mathcal{B}_I(T)$, $\pi_{B'_p}(x_p) \leq \pi_1$,

$B"_p = (q"_i[t_p] \overset{*}{\underset{b"^p_i}{\Longrightarrow}} u"^p_i)_{i \in I"_p}$ avec $(b"^p_1,\ldots,b"^p_{V(B"_p)})$ sous-suite

contenant $b^p_{j^p_1}$ et $b^p_{j^p_2}$ de $(b^p_1,\ldots,b^p_{V(B_p)})$. Remarquons que $g(B"_p) = g(B_p)$, donc $B"_p$ émondé.

D'après le lemme III.1.1.3) on peut supposer B'_p émondé selon x_p. Alors, d'après 5) de ce lemme, $B'_{px_p}(B"_p)$ est émondé.

B] Posons $\pi_2 = \sup_{Q' \in \mathcal{P}(Q)} (\inf_{B \in \mathcal{B}(T)} (\pi(g(B)) \mid [Q(B)] = Q'))$

pour tout p $\in$ N, posons $\bar{B}_p = B'_{px_p}(B"_p)$ le bloc émondé obtenu en A].

En utilisant le sous-lemme III.2.1, le sous-lemme III.1.1.4 et le fait que $\bar{B}_p$ est utile, il est facile de voir que $\forall$ p, $\exists$ $\bar{\bar{B}}_p$

tel que :

i) $\bar{\bar{B}}_p = \bar{B}_p(B_p^1,\ldots,B_p^{m_p})$

ii) $\bar{\bar{B}}_p(q_o[\theta_p] \overset{*}{\Longrightarrow} \bar{u}_p)$ avec $(\theta_p,\bar{u}_p) \in \Sigma^* \times \Delta^*$

(donc $(\theta_p,\bar{u}_p) \in \hat{T}$.

iii) $g(\bar{\bar{B}}_p) = \bar{t}_1^p(\ldots\bar{t}_{n_p}^p(\bar{t}_{n_{p+1}}^p)\ldots) = \theta_p$ où

$\pi(\bar{t}_i^j) \leq \pi_2 + \sup_{t_j \in G} \pi(t_j) = \pi_3$

$\bar{\bar{B}}_p$ s'obtient en terminant la dérivation associée à $\bar{B}_p$ "dès que possible".

Par construction, $\bar{\bar{B}}_p$ a la propriété suivante :

$\bar{\bar{B}}_p q_o[\theta_p] \overset{*}{\Longrightarrow} \bar{u}_p$ avec $\bar{u}_p = \bar{v}_p(\bar{v}_p',\bar{v}_p")$

où $\pi(\bar{v}_p) \leq \pi_d(T) \times (\pi_1 + \pi_2)$

et $\pi(\bar{v}_p') \geq \pi(u"_{j_1^p}^p) \geq p$ et $\pi(\bar{v}_p") \geq \pi(u"_{j_2^p}^p) \geq p$

Puisque $\pi(\bar{v}_p) \leq M$ indépendant de p d'une part, et d'après la propriété iii) de $\bar{\bar{B}}_p$ d'autre part, il existe une infinité d'indices p_q tels que, en renumérotant la suite $(\hat{B}_{p_q})$ en $(\hat{B}_q)$, on ait :

$\exists\, \bar{v}$, $\exists$ *un ensemble fini* $G' \subset \Sigma^*(X)$ *tels que :*

$\forall\, q \in N$, $\exists\, t_{q_1},\ldots,t_{q_{n_q}} \in G'$ *tels que* $\theta_p = t_{q_1}(t_{q_2}\ldots(t_{q_{n_q}})))$,

$\bar{u}_q = \bar{v}(\bar{v}_q',\bar{v}_q")$ *avec* $\pi(\bar{v}_q') \geq q$ *et* $\pi(\bar{v}_q") \geq q$

et enfin $(\bar{u}_q,\theta_q) \in \hat{T}^{-1}$.

Ce qui contredit le lemme II.2 partie B], car $T \in TGQ_nD$. Donc $\exists\, U \in TGDX$ tel que $\hat{U} = \hat{T}^{-1}$.

cqfd.

Commentaire : Dans A], nous montrons qu'on peut trouver M_1 tel que, de B_p, on déduit une dérivation dans la partie droite de laquelle deux arbres u'_p et u''_p de profondeur $\geq p$ apparaissent avec leur sommet avant la profondeur M_1.

On "émonde" cette dérivation en n'en conservant que la partie engendrant u'_p et u''_p.

Dans B] on termine dès que possible la dérivation . On obtient ainsi une famille de couples contredisant la linéarité globale car la partie gauche, émondée, est suffisamment "étroite".

Corollaire III.3 *Soit* T *un* $TGQ_n DH$

$$\exists\, p > 0 \text{ tel que } \forall\, B(q_i[t] \overset{*}{\Longrightarrow} u_i)_{i\in I} \in \mathcal{B}(T)$$

$$\forall\, x \in V(B),\ \pi_{u_{i_1}}(x) \geq p \text{ et } \pi_{u_{i_2}}(x) \geq p \text{ entraîne que } i_1 = i_2.$$

En effet. Supposons le corollaire faux.

$$\forall\, p,\ \exists\, B_p(q_i[t_p] \underset{b_i}{\overset{*}{\Longrightarrow}} u_i^p)_{i\in I_p} \in \mathcal{B}(T) \text{ tel que}$$

$$\exists\, x_p \in V(B_p), \text{ avec } \pi_{u^p_{i_1}}(x_p) \geq p,\ \pi_{u^p_{i_2}}(x_p) \geq p \text{ et } i_1 \neq i_2.$$

En émondant B_p selon x, on obtient d'après III.1.1.6),

$B'_p(q_i[t'_p] \underset{b'_i}{\overset{*}{\Longrightarrow}} u'^p_i)_{i\in I_p} \in B(T)$ émondé tel que

$$\pi_{u'_{i_1}}(x) = \pi_{u_{i_1}}(x) \geq p \text{ et } \pi_{u'_{i_2}}(x) = \pi_{u_{i_2}}(x) \geq p$$

avec $i_1 \neq i_2$, ce qui contredit le lemme III.2.

cqfd.

Corollaire III.4 *Soit* T *un* $TGQ_n DH$.

$\forall$ p, $\exists$ m, q $\in$ N *tels que* $\forall$ B : $(q_i[t] \overset{*}{\underset{b_i}{\Longrightarrow}} (u_i))_{i\in I} \in B(T)$,

$\forall$ x *tel* $\exists$ i $\in$ I, x $\in V(B) \cap V(u_i)$

si $\pi_B(x) \geq m$ *et* $\pi_{u_i}(x_o) < p$ *(pour une certaine occurrence* x_o *de* x*)* (1)

alors $[\forall B'(q_i[t'] \overset{*}{\underset{b'_i}{\Longrightarrow}} u'_i)_{i\in I'} \in B(T)$ *tel que* B' *composable*

avec B *en* x, *si* u'_{i_1} *s'applique en* x_o *à* u_i,

alors $\pi(u'_{i_1}) < q]$.

En effet. Supposons le corollaire faux.

$\exists$ p_o tel que $\forall$ p = m = n $\in$ N, $\exists$ $B_p(q_i[t_p] \overset{*}{\underset{b_i^p}{\Longrightarrow}} u_i^p)_{i\in I_p} \in \mathcal{B}(T)$

tel que $\exists$ x_p tq $\exists$ $i_p \in I_p$, $x_p \in V(B_p) \cap V(u_{i_p})$

tel que $\pi_{B_p}(x_p) \geq p$, $\pi_{u_{i_p}}(x_p) < p_o$

et $\exists$ $B'_p(q_i[t'_p] \overset{*}{\Longrightarrow} u'^p_i)_{i\in I'_p} \in \mathcal{B}(T)$ tq B'_p composable avec B_p en x_p,

u'^p_i, s'appliquant en x_p à u_i^p et $\pi(u'^p_{i'}) \geq p$.

Remarque : On définit l'émondage selon une feuille quelconque exactement comme selon une variable.

A] L'émondage selon x_p conservant les propriétés de profondeur de x_p, on peut supposer B_p émondé selon x_p. Il est facile de montrer qu'alors $\exists$ M tel que $\forall$ u_i^p, $\pi(u_i^p) \leq \pi_{u_{i_p}}(x_p) + M$.

Supposons $\exists$ p_1 tel que $\forall$ p, $\forall$ u_{i_p}, $\pi_{u_{i_p}}(x_p) < p_1$

alors $\exists$ $p_2 = p_1 + M$ tel que

---

(1) Nous ne préciserons plus dans la suite.

$\forall\, p,\ \exists\, B_p (q_i[t_p] \overset{*}{\Longrightarrow} u_1^p)_{I_p} \in \mathcal{B}(T)$ tel que

$$\pi(t_p) \geq \pi_B(x_p) \geq p \text{ et } \forall\, i \in I_p,\ \pi(u_{i_p}) < p_2$$

Les inégalités, jointes au fait que B_p est utile, contrediraient facilement la fidélité.

On composerait par un bloc initial de profondeur $\leq \pi_1$ comme page II.23, on terminerait dès que possible comme page II.25 d'où une infinité d'arbres (profondeur non bornée) se dériverait en un nombre fini d'arbres (profondeur bornée).

B] On peut donc supposer de plus que $\forall\, p_1,\ \exists\, p,\ \exists\, u_{i''_p}$ tel que $\pi_{u_{i''_p}}(x_p) > p_1$.

Il existe donc une sous-suite d'indices (p_n) telle que $\forall\, n \in N$,

$$\pi(B_{p_n}(x_{p_n})) \geq n,\quad \pi_{u_{i''_{p_n}}}(x_{p_n}) \geq n \text{ et } \pi(u'^{p_n}_{i'_{p_n}}) \geq n$$

Comme $\pi u^{p_n}_{i_{p_n}}(x_{p_n}) < p_o$, pour n assez grand nous avons $i''_{p_n} \neq i_{p_n}$.

Considérons la feuille p_n d'une plus longue branche de $u'^{p_n}_{i'_{p_n}}$ et émondons B'_{p_n} selon une feuille f'_n appartenant à la partie gauche de la chaîne dérivant f_n. Alors d'après les propriétés de l'émondation, en posant $B_n = B_{p_n x_{p_n}}(B'_{p_n})$, on a B_n émondé $\in \mathcal{B}(T)$ et $i''_{p_n} \neq i_{p_n}$ avec $\pi(u^{p_n}_{i''_{p_n}}) \geq n$ et $\pi(u^{p_n}_{i_{p_n}}(u'^{p_n}_{i'_{p_n}})) \geq n$, ce qui contredit le lemme III.2.

cqfd.

Corollaire III.5 *Soit* T *un* $TGQ_n D$.

$$\exists\, p,m,q \text{ tels que } \forall\, B : (q_i[t] \overset{*}{\underset{b_i}{\Longrightarrow}} u_i)_{i \in I} \in \mathcal{B}(T)$$

tel que $\forall\, x \in V(B)$, $\pi_B(x) = m$ *et*

- *si* $\pi_{u_{i_1}}(x) \geq p$ *et* $\pi_{u_{i_2}}(x) \geq p$, $i_1 = i_2$
- *si* $\pi_{u_i}(x) < p$, $\forall B'(q'_i[t'] \overset{*}{\underset{b'_i}{\Longrightarrow}} u'_i)_{i \in I'}$, *tel que*

  B' *composable avec* B *en* x,

  u'_i, *s'applique en l'occurrence considérée avec* u_i

  $\Longrightarrow \pi(u'_{i,}) < q$.

En effet. Ce corollaire est la conjonction logique de III.3 et III.4.

cqfd.

Remarques III.6. Soit T' dont les règles sont les $q_i[t] \overset{*}{\underset{b_i}{\Longrightarrow}} u_i$ apparaissant dans le corollaire ci-dessus : ce sont en fait les blocs b de taille 1 tels que $\pi_B(x) = m \ \forall \ x \in V(b)$. Elles sont donc bien en nombre fini, m étant fixé.

D'autre part, chaque variable est (au sens des règles) à profondeur n×m en partie gauche de règle (n est la profondeur de T). Donc T' est un TGQ_n^D (de profondeur m×n).

Remarquons enfin que $\forall$ B $\in$ B(T), $\forall$ x $\in$ V(B), on a

$$\text{Card } ([Q_x(B)]\ (T)) = +\infty$$

(rappelons que l'écriture ci-dessus désigne le cardinal de l'ensemble des $t \in \Sigma^*$ pouvant se dériver à partir de tout état associé à x dans B. Les définitions précises figurent pages II.18, II.20 et II.23).

En effet. La propriété est vraie par définition pour les c.h.e. et se conserve pour les blocs quelconques par définition de la composition (pages II.19, II.20 et II.21).

cqfd.

Supposons maintenant que, dans le corollaire III.5, il existe B tq $\forall$ x $\in$ V(B), $\pi_B(x) = m$ et $\exists$ x $\in$ V(B) tq $\forall$ u_i, $\forall$ l'occurrence de x dans u_i, $\pi_{u_i}(x) < p$. Alors, comme Card $[Q_x(B)]\ (T) = +\infty$, on procède comme dans le corollaire III.4 A] pour nier la fidélité.

Nous pouvons énoncer en vertu du corollaire III.5 et des remarques III.7

Proposition III.7 : $\forall$ T $\in$ TGQ_nD, $\exists$ T' $\in$ TGQ_nDH, $\hat{T}' = \hat{T}$ *et tq* $\exists$ p, q *tels que* $\forall$ $C(q_i[t] \to u_i)_{i\in I} \in C(T')$, $\forall$ x $\in$ V(C)

i) - *il existe un et un seul* $i\in I$ tel que $x \in V(u_i)$

avec $\pi_{u_i}(x) \geq p$

ii) - *pour toute occurrence* $\bar{x}$ *de* x *dans un* u_i *telle que* $\pi_{u_i}(\bar{x}) < p$, $\forall$ $B'(q'_i[t'] \overset{*}{\Longrightarrow} u'_i)_{i\in I}$, *tel que* B' *composable à* C *en* x, *si* u'_i, *s'applique en* $\bar{x}$, *alors* $\pi(u'_i) < q$.

C(T') désigne l'ensemble des chaînes utiles de T'.

La classe des tels TGQ_nDH sera notée TGQ_nDF et si T' $\in$ TGQ_nDF, on le dira fondamental.

Tout TGQ_nD *est donc équivalent à un* TGQ_nDF.

Dans toute chaîne, les occurrences droites de variable de profondeur $\geq$ p seront dites *principales* (cas i) les autres seront dites *non principales* (cas ii). p s'appelle le *seuil* de T', q la *borne* de T'.

La taille d'une chaîne n'étant pas majorée, le nombre d'occurrences droites d'une variable dans une chaîne n'est pas majoré . Mais les occurrences principales d'une variable n'apparaissant que dans une et une seule règle de la chaîne, le nombre de ces occurrences est majoré par un entier μ que nous appellerons *module de* T'.

Soit T un TGQ_nDF de profondeur n, seuil p, borne q et module μ

$$T = \{\Sigma^*, \Delta^*, R, Q, \{q_o\}\}$$

posons $G = \{\theta \in \Delta^* \mid \pi(\theta) \leq q\}$, $R = \{r_1, \ldots, r_\alpha\}$

avec $r_i = q_i\,[t_i(x_1,\ldots,x_m)] \to u_i(q_{s_1}[x_{s_1}], \ldots, q_{s_{m'}}[x_{s_{m'}}])$

et C_μ = {c.h.e. utiles de T de taille $\leq \mu$}.

A] A T , nous allons associer $\bar{T} \in TGQ_n DH$ tel que $\hat{T} = \hat{\bar{T}}$

posons $\bar{T} = \{\Sigma^*, \Delta^*, \bar{R}, \bar{Q}, \bar{Q}_o\}$

$$\bar{Q} \subset \{1,2,\ldots,\mu\} \times C_\mu \times \mathcal{P}(R \times G)$$

si $\bar{q} = (\bar{q}_1, \bar{q}_2, \bar{q}_3)$ nous poserons $\mathcal{P}_{r_i}(\bar{q}) = \bar{q}_i$ pour i=1,2,3.

Explicitons $\bar{q} \in \bar{Q}$:

$$\forall\ \bar{q} = (i', (r_{i_1}, \ldots, r_{i_k}), \{(r_{j_1}, \theta_{j_1'}), \ldots, (r_{j_{k'}}, \theta_{j_{k'}'})\}) \in \bar{Q},$$

nous nous imposons que :

- $\exists\ t(x_1,\ldots,x_m)$ tel que $t(x_1,\ldots,x_m) = g(r_{i_1}) = \ldots$

$$\ldots = g(r_{i_k}) = g(r_{j_1}) = \ldots = g(r_{j_k})$$

nous poserons par abus $g(\bar{q}) = t(x_1,\ldots,x_m)$.

- $\forall\ \beta = 1,\ldots,k'\ \ \exists\ \theta^1_{j_\beta'},\ldots,\theta^{m'}_{j_\beta'} \in \Delta^*$ tels que

$$\theta_{j_\beta'} = u_{j_\beta}(\theta^1_{j_\beta'},\ldots,\theta^{m'}_{j_\beta'})$$

- $i' \leq k \leq \mu$.

A toute règle r_i de R, nous allons associer un ensemble $\bar{R}(r_i)$ de règles de $\bar{R}$:

$$r_i = q_i[t_i(x_1,\ldots,x_m)] \rightarrow u_i(q_{s_1}[x_{s_1'}], \ldots, q_{s_{m'}}[x_{s_{m'}'}])$$

1) Choisissons $\bar{q} \in \bar{Q}$ tel que $r_i = r_{i_{i'}}$ (c'est-à-dire que r_i est la $i^{ème}$ règle de la chaîne $\mathcal{P}_{r_2}(\bar{q})$). On substitue dans r_i $\bar{q}$ à q_i.

2) <u>Pour</u> N=1 <u>jusque</u> m :

Soit $x_{s'_{\alpha_1}}, \ldots, x_{s'_{\alpha_\nu}}$ les occurrences principales, de gauche à droite, de x_N dans u_i.

Choisissons $C \in C_\mu$ avec $Q(C) = (q_{s_{\alpha_1}}, \ldots, q_{s_{\alpha_\nu}})$.

Pour $\nu' = 1, \ldots, \nu$, remplaçons dans u_i $q_{s_{\alpha_{\nu'}}}$ par

$\bar{q}_{s_{\alpha_{\nu'}}}$ où : $P_{r_1}(\bar{q}_{s_{\alpha_{\nu'}}}) = \nu'$, $P_{r_2}(\bar{q}_{s_{\alpha_{\nu'}}}) = c$ et $P_{r_3}(\bar{q}_{s_{\alpha_{\nu'}}}) = \emptyset$

3) Soit $x_{s'_\beta}$ une occurrence non principale dans u_i.

On choisit $\Theta \in G$.

On remplace $q_{s_\beta}[x_{s'_\beta}]$ par Θ dans u_i.

Pour toute occurrence principale $x_{s'_{\beta'}}$ de la même variable, on pose $P_{r_3}(\bar{q}_{s'_{\beta'}}) := P_{r_3}(\bar{q}_{s'_\beta}) \cup \{(r,\Theta)\}$ où r est choisie telle que $Q(r) = q_{s_{\beta'}}$ et $g(r) = g(r')$ où $r' \in P_{r_2}(\bar{q}_{s_{\beta'}})$.

On itère 3) pour toutes les occurrences non principales de u_i.

Il n'apparaît plus dans u_i que des occurrences principales de variables.

4) Pour tout $(r,\Theta) \in P_{r_3}(\bar{q})$

D'après la définition de $\bar{q}$, nous avons

$$r = q[t_i(x_1, \ldots, x_m)] \to u(q_{\alpha_1}[x_{\alpha'_1}], \ldots, q_{\alpha_m}[x_{\alpha'_m}])$$

et $\Theta = u(\Theta^1, \ldots, \Theta^{m'})$

Pour N = 1 jusque m

Soit $x_{\alpha'_{\beta_1}}, \ldots, x_{\alpha'_{\beta_\nu}}$ les occurrences dans u de x_N.

Choisissons $r_1,\ldots,r_{\nu'} \in R$ telles que pour $i=1,\ldots\nu'$,

on ait $$Q(r_i) = q_{\alpha_{\beta_i}}.$$

Pour toute occurrence x_{s_j} de x_N dans u_i, nous posons alors

$$P_{r_3}(\bar{q}_{s_j}) := P_{r_3}(\bar{q}_{s_j}) \cup \{(r_1,\Theta^{\beta_1}),\ldots(r_{\nu'},\Theta^{\beta_{\nu'}})\}$$

(il faut avoir choisi $r_1,\ldots,r_{\nu'}$ telles que les états ainsi définis soient bien dans $\bar{Q}$).

5) Toutes les règles de $\bar{R}(r_i)$ sont *obtenues* ainsi. On pose $\bar{R} = \bigcup_{r\in R} \bar{R}(r_i)$.

Enfin, $\bar{Q}_o = \{\bar{q} \mid P_{r_2}(\bar{q}) = r$ avec $Q(r) = q_o$, et $P_{r_3}(\bar{q}) = \emptyset\}$.

Commentaire : En 3) on anticipe l'arbre Θ qui sera dérivé dans T en un point qui nous assure, d'après les propriétés de T, que $\pi(\Theta) \lesssim q$ (occurrence non principale) et on mémorise afin de vérifier cette dérivation. En 4) on propage les anticipations.

Par 2), 3) et 4) on mémorise dans chaque règle "quelles seront les règles appliquées au même point", ce qui permettra d'une part de synthétiser les anticipations pour avoir $\hat{\bar{T}} = \hat{T}$, et d'autre part d'avoir le lemme III.8, ii), puis le lemme III.9.

Les parties gauches de règles de $\bar{R}$ étant celles de R, $\bar{T}$ est un TGQ_n. Le lemme suivant montre qu'il est homogène.

Lemme III.8 : *Soit* $(t,u) \in \hat{\bar{T}}$. *Il existe une dérivation de* t *en* u *par blocs. De plus, toutes les chaînes* $\bar{c}$ *sont de la forme* $(\bar{q}_i[t] \to u_i)_{i\in I}$ *avec*

i) $\forall$ x $\in$ V(t), $\exists$ 1 *et* 1 *seul* i $\in$ I *tel que* x $\in$ V(u_i) *(et alors* $\pi_{u_i}(x) \gtrsim p$)

ii) $\exists$ c $\in C_\mu$, A $\in P(R\times G)$ *tels que*

$$Q(\bar{c}) = (\bar{q}_1,\ldots,\bar{q}_{p'}) \text{ avec } p' \leq \mu$$

et pour tout $i=1,\ldots,p'$ $P_{r_1}(\bar{q}_i) = i$, $P_{r_2}(\bar{q}_i) = c$ *et* $P_{r_3}(\bar{q}_i) = A$.

En effet. Soit $\bar{\delta} : \bar{q}_o[t_1] \overset{*}{\underset{\bar{T}}{\Longrightarrow}} u$. Considérons une dérivation initiale maximale par bloc $\bar{\delta}'$ de $\bar{\delta}$, les chaînes vérifiant i) et ii).

Elle existe car la première règle de $\bar{T}$ utilisée dans $\bar{\delta}$ est initiale par bloc et vérifie évidemment i) et ii).

Supposons $\bar{\delta}' \neq \bar{\delta}$: en posant $\bar{\delta}' \; \bar{q}_o[t'] \overset{*}{\underset{b\bar{T}}{\Longrightarrow}} u'$, nous avons donc $V(u') \neq \emptyset$. Soit $x_i \in V(u')$.

Soit $x_{i_1},\ldots,x_{i_\nu}$ la suite des occurrences de x_i dans u', lues de gauche à droite, et x_i l'occurrence de x_i dans t'. Comme $\bar{\delta}'$ par bloc, il existe une et une seule chaîne $\bar{c}' \; (q'_i(t'_i) \to u'_i)_{i\in I'}$ de $\bar{\delta}'$ telle que $x_i \in V(t'_1)$.

Alors, $\forall\, j=1,\ldots,\nu$, $\exists\, j' \in I$ tel que $x_{i_j} \in V(u'_{j'})$, mais, comme $\bar{c}'$ vérifie par hypothèse i), $\exists\, j'_o$ tel que $(x_{i_1},\ldots,x_{i_\nu}) \in V(u'_{j'_o})$.

D'après la construction A] 2), 3) et 4), où $(x_{i_1},\ldots,x_{i_\nu})$ est la suite $(x_{s'_1},\ldots,x_{s'_\nu})$ et $u'_{j'_o}$ remplace u_i, la suite des règles appliquées de gauche à droite en x_i est bien une chaîne $\bar{c}$ vérifiant ii) pour c et A.

La même construction A] montre que la suite des règles de $\bar{c}$ est construite à partir de celle de c, en substituant aux variables non principales des arbres de G. Seules les variables principales restantes, le lecteur vérifiera précisément en utilisant A] et III.7.i) que i) du lemme en cours est vérifié.

cqfd.

Remarque : On a construit un $\bar{T}$ finiment réducteur (pas de règle réductrice) équivalent à un TGQ_nD donné.

Lemme III.9 : *Soit* $\bar{\delta} : \bar{q}'_o[\bar{t}(x_1,\ldots,x_m)] \overset{*}{\underset{b\bar{T}}{\Longrightarrow}} \bar{u}(\bar{q}'_{\gamma_1}[x_{\gamma_1}],\ldots,\bar{q}'_{\gamma_{p'}}[x_{\gamma_{p'}}])$

une dérivation initiale utile par bloc dans $\bar{T}$, *alors*

i) $\forall\ i = 1,\ldots,m \quad x_i \in V(\bar{u})$

ii) $x_{\gamma'_{i_1}} = x_{\gamma'_{i_2}} \Longrightarrow P_{r_2}(\bar{q}'_{\gamma_{i_2}}) = P_{r_2}(\bar{q}_{\gamma_{i_1}})$ *et* $P_{r_3}(\bar{q}'_{\gamma_{i_1}}) = P_{r_3}(\bar{q}'_{\gamma_{i_2}})$

i) et ii) permettent d'associer à x'_i, *pour* $i=1,\ldots,m$, *un élément* $c_{x_i} \times A_{x_i} \in C_\mu \times P(R\times G)$ *défini par*

$$c_{x_i} = P_{r_2}(\bar{q}'_{\gamma_{i_1}}) \text{ où } x_{\gamma'_{i_1}} = x_i$$

et $$A_{x_i} = P_{r_3}(\bar{q}_{\gamma_{i_1}}) \text{ où } x_{\gamma'_{i_1}} = x_i$$

iii) $\exists\ \bar{u}'(q'_{\Sigma_1}[x_{\Sigma'_1}], \ldots, q'_{\Sigma_{p''}}[x_{\Sigma'_{p''}}])$ *tel que*

a) $\forall\ x_i, \forall (r,\Theta) \in A_{x_i}, \exists\ x_{\Sigma'_{i_{1_\Theta}}}, \ldots, x_{\Sigma'_{i_{q'_\Theta}}}$ *tels que pour tout*

$j=1,\ldots,q'$, $q'_{\Sigma_{i_{j_\Theta}}} = g(r)$

et, en substituant dans $\bar{u}'$ *(pour tout* x_i *et tout* $(r,\Theta) \in A_{x_i}$ *)*

Θ *à* $q'_{\Sigma_{i_{j_\Theta}}}[x_{\Sigma'_{i_{j_\Theta}}}]$, *en supprimant les* q'_{Σ_h} *"restants", on obtient*

$\bar{u}(x_{\gamma'_1}, \ldots, x_{\gamma'_{p'}})$.

b) Le point a) permet alors de mettre en bijection $x_{\gamma'_1},\ldots,x_{\gamma'_{p'}}$ *avec la suite* $x_{\Sigma'_{\gamma'_1}}, \ldots, x_{\Sigma'_{\gamma'_{p'}}}$ *des* $x_{\Sigma'_k}$ *"restants".*

On a alors de plus; pour $j=1,\ldots,p'$

$$q'_{\Sigma_{\gamma'_j}} = Q(r_{P_{r_1}(\bar{q}_{\gamma'_j})})$$

c) $q_o[\bar{t}(x_1,\ldots,x_m)] \overset{*}{\underset{T}{\Longrightarrow}} \bar{u}'(q'_{\Sigma_1}[x_{\Sigma'_1}],\ldots,q'_{\Sigma_{p''}}[x_{\Sigma'_{p''}}])$.

En effet. Démontrons ce lemme par récurrence sur la longueur L du bloc.

Si L=1, le résultat est vrai par construction.

Supposons vrai jusque L, le lecteur vérifiera facilement qu'à la longueur L+1

i) est vrai d'après A] et III.7,i)

ii) et iii) a) et b) vrais par construction A] et III.8

iii) c) est vrai en utilisant i), ii) et iii) a) et b) et A]

cqfd.

Corollaire III.10 $\quad \hat{\bar{T}} \subset \hat{T}$

En effet. Soit $(\bar{t},\bar{u}) \in \hat{\bar{T}}$. Il existe une dérivation par bloc $\bar{q}_o[\bar{t}] \underset{b\bar{T}}{\overset{*}{\Longrightarrow}} \bar{u}$.

Donc d'après iii)c), $q_o[\bar{t}] \underset{\bar{T}}{\overset{*}{\Longrightarrow}} \bar{u}'$

où $\bar{u}'$ est un sous-arbre initial de $\bar{u}$ sans variable, donc $\bar{u}' = \bar{u}$ et $[\bar{t},\bar{u}] \in \hat{T}$.

cqfd.

Commentaire : Dans le lemme III.9, le point i), bien que "discret" est essentiel il assure l'existence de Ax_i, donc donne un sens à la suite.

Ce point i) provient de ce que T est une bitransduction donc "lit" tout ce qu'il écrit (toutes les variables gauches se retrouvent donc à droite).

Ceci nous permet dans le lemme de garder une occurrence de toute variable où l'on a anticipé, donc de vérifier qu'on pourrait faire la dérivation anticipée.

L'inclusion $\hat{T} \subset \hat{\bar{T}}$ étant une conséquence immédiate de la construction A], nous en laissons la démonstration au lecteur.

Nous pouvons énoncer :

Corollaire III.11 : $\quad \hat{T} = \hat{\bar{T}}$

Nous allons construire T_1 et T_2 $\in$ TI tels que $\hat{T}_1 \circ \hat{T}_2 = \hat{\bar{T}}$.

A cette fin, remarquons que d'après le lemme III.8, ii), les chaînes utiles de $\bar{T}$ sont de taille $\leq \mu$, donc sont en nombre fini. Nous poserons

$$C(\bar{T}) = \{c_1,\ldots,c_N\}$$

et $\quad c_i = (c_{i,1},\ldots,c_{i,\nu(c_i)})$ où $c_{i,j}$ est la $j^{ème}$ règle de c_i.

Construction de T_1 :

Nous allons prendre pour alphabet-image de T_1 l'ensemble des labels $c_1,\ldots,c_N$ des chaînes de $C(\bar{T})$ uni à celui des labels $c_{i,1},\ldots,c_{i,\nu(c_i)}$ des occurrences de règles dans les chaînes.

Soit C cet alphabet.

$$T_1 : \{\Sigma\ ,\ C\ ,\ P(\bar{Q}),\ R(T_1),\ \bar{Q}_o\}$$

Les chaînes initiales de $\bar{T}$ sont les chaînes commençant les dérivations par blocs. Ce sont les chaînes réduites à une règle d'état gauche initial.

$$\text{Soit } c_i = \bar{q}_o[t(x_1,\ldots,x_m)] \to u(\bar{q}_{\alpha_1}[x_{\alpha_1}],\ldots,\bar{q}_{\alpha_{m'}}[x_{\alpha_{m'}}])$$

une chaîne initiale de $C(\bar{T})$.

Nous lui associons

$$\bar{q}_o[t(x_1,\ldots,x_m)] \to c_{i,1}(\bar{Q}_1[x_1],\ldots,\bar{Q}_m[x_m]) \in R(T_1)$$

avec, pour i=1,...,m : $\bar{Q}_i$ = suite des états de $\bar{Q}$ associés de gauche à droite aux diverses occurrences de x_i dans u.

$$\text{Soit } c_j = (\bar{q}_i[t(x_1,\ldots,x_m)] \to u_i(\bar{q}^i_{\alpha_1}[x^i_{\alpha_1}],\ldots,\bar{q}^i_{\alpha_{m'_i}}\ x^i_{\alpha_{m'_i}}))_{i\in I}$$

une chaîne non initiale de $C(\bar{T})$.

Nous lui associons dans $R(T_1)$ la règle

$$Q(c_j)\ [t(x_1,\ldots,x_m)] \to c_j(c_{j,1}(\bar{Q}_{i_{1_1}}[x_{i_{1_1}}],\ldots,\bar{Q}_{i_{1_{p_1}}}[x_{i_{1_{p_1}}}]),c_{j,2}(\ldots),\ldots$$

$$\ldots,c_{j,\nu(c_i)}(\bar{Q}_{i_{\nu(c_i)_1}}[x_{i_{\nu(c_i)_1}}],\ldots,\bar{Q}_{i_{\nu(c_i)_{p_{\nu(c_i)}}}}[x_{i_{\nu(c_i)_{p_{\nu(c_i)}}}}]$$

où $V(u_k) = \{x_{i_{k_1}},\ldots,x_{i_{k_{p_k}}}\}$ pour $k=1,\ldots \nu(c_i)$

et $i_{k_1} < i_{k_2} < \ldots < i_{k_{p_k}}$.

$\bar{Q}_{i_{j_k}}$ = suite des états de $\bar{Q}$ associés aux diverses occurrences de gauche à droite de $x_{i_{j_k}}$ dans u_j.

Les règles de $R(T_1)$ sont ainsi en bijection avec les chaînes de $C(\bar{T})$ (rappelons qu'on suppose que les états initiaux ne réapparaissent pas).

Ces règles sont du type TI d'après III.8.i).

Construction de T_2 :

$$T_2 : \{C, \Delta, (q), R(T_2), (q)\}$$

T_2 aura un seul état (q) que nous omettrons dans la suite.

Les règles de $R(T_2)$ sont toutes les règles de l'une des deux formes suivantes :

1) $$c_{j,i}(c_{i_1}(x_1,\ldots x_{\beta_1}),c_{i_k}(x_{\beta+1_{k-1}},\ldots,x_{\beta_k}),\ldots,c_{i_\ell}(x_{\beta+1_{\ell-1}},\ldots,x_{\beta_\ell}))$$

$$\to u_j(x_{\sigma(1)},\ldots,x_{\sigma(\beta_\ell)})$$

explicitons $c_{j,i}$ = $i^{\text{ème}}$ règle de c_j

$$= \bar{q}[t(x_1,\ldots,x_m)] \to u_j(\bar{q}_{\alpha_1}[x_{\alpha_1'}],\ldots,\bar{q}_{\alpha_m'}[x_{\alpha_m'}])$$

donc u_j est l'arbre droit de $c_{j,i}$.

Remarquons que l'arité de $c_{j,i}$ est d'après T_1 le nombre de variables distinctes de u_j, et l'arité de c_{i_k} est la taille de la chaîne appliquée, c'est-à-dire le nombre d'occurrences d'une même variable.

On a ici β_1 occurrences de $x_{i_{j_1}}$, $\beta_2 - \beta_1$ de $x_{i_{j_2}}$ ... $\beta_\ell - \beta_{\ell-1}$ occurrences de $x_{i_{j_\ell}}$, donc on a bien $\beta_\ell = \alpha'_{m'}$.

σ est la permutation telle que : (1)

x_1 est à la place de la 1ère occurrence de $x_{i_{j_1}}$ dans $u_j(x_{\alpha'_1}, \ldots, x_{\alpha'_{m'}})$

$\vdots$

x_{β_1} - - - - - - - - - $\beta_1^{ème}$ -

$\vdots$

x_{β_1+1} - - - - - - - - 1ère - - - - - - $x_{i_{j_2}}$ - - - - - - - - - - -

$x_{\beta+1_{\ell-1}}$ - - - - - - - 1ère - - - - - - $x_{i_{j_\ell}}$ - - - - - - - - - - -

x_{β_ℓ} - - - - - - - - $\beta_\ell^{ème}$ - - - - - - $x_{i_{j_\ell}}$ - - - - - - - - - - -

2) Si $c_{j,i}$ est 0-dire, $u_j \in \Delta^*$. On construit alors la règle $c_{j,i} \to u_j$.

Le lecteur vérifiera facilement que $\hat{T}_1 \circ \hat{T}_2 = \hat{\bar{T}}$ (on a codé dans C^* les dérivations par bloc). La démonstration est sans difficulté mais d'écriture fastidieuse d'où le

Lemme III.12 : $\hat{T}_1 \circ \hat{T}_2 = \hat{\bar{T}}$

Théorème III.13 : $\widehat{TGQ_nD} \subset \widehat{TI^2}$.

En effet. Soit $T \in TGQ_nD$. Par les intermédiaires III.1, III.7 et III.11 nous construisons $\bar{T} \sim T$ tel que par III.12 $\exists$ T_1 et T_2 vérifiant $\hat{\bar{T}} = \hat{T}_1 \circ \hat{T}_2$.

cqfd.

---

(1) On a ainsi reconstitué la correspondance entre variables de $c_{j,i}$, en distinguant les occurrences de gauche à droite d'une même variable.

CHAPITRE IV

CARACTERISATION DES BITRANSDUCTIONS DU TYPE "DESCENDANT-DESCENDANT"

Si ϕ est un homomorphisme de Σ^* dans Δ^*, on note $\hat{\phi} = \{(t,\phi(t)) \mid t \in \Sigma^*\}$ qu'on écrira aussi parfois $\{(t,\phi(t))\}$.

Proposition IV.1 : *Si $\hat{\tau} \in TG_D$ et si ϕ est un homomorphisme, alors* $\hat{\tau}.\hat{\phi} = \{(t,\phi(t')) \mid (t,t') \in \hat{\tau}\} \subset \widehat{TG}_D$.

En effet. Soit τ' le transducteur dont les règles sont obtenues en remplaçant dans toute règle

$$q[t(x_1,\ldots,x_n)] \to t'(q_{i_1}[x_{i_1}],\ldots,q_{i_p}[x_{i_p}]) \text{ de } \tau$$

l'arbre en partie droite par son image par ϕ. On démontre aisément par induction que

$$q[t] \overset{*}{\underset{\tau}{\Longrightarrow}} t' \text{ ssi } q[t] \overset{*}{\underset{\tau'}{\Longrightarrow}} \phi(\tau') \text{ d'où } \hat{\tau}' = \hat{\tau}.\hat{\phi}.$$

Par construction, τ' est encore finiment réducteur.

cqfd.

Lemme IV.2 : *Si $\hat{\tau} \subset \widehat{TG}_D$ et si ϕ est un homomorphisme ordonné alors* $\hat{\tau}.\hat{\phi}^{-1} = \{(t,t') \mid (t,\phi(t')) \in \hat{\tau}\} \in \widehat{TG}_D$.

En effet. Pour pouvoir faire les constructions nécessaires à cette démonstration, il faut d'abord introduire plusieurs notions auxiliaires.

Ⓐ Soient Σ_o et Σ_1 deux alphabets gradués et ϕ un homomorphisme ordonné de Σ_o^* dans Σ_1^*. On définit un nouvel alphabet gradué Σ_1/Σ_o par

$$\begin{cases} \Sigma_1/\Sigma_o = \Sigma_1 \cup \Sigma_1 \times \Sigma_o \\ \text{si } \langle\sigma,\sigma'\rangle \in \Sigma_1 \times \Sigma_o,\ d(\langle\sigma,\sigma'\rangle) = d(\sigma) \end{cases}$$

On définit trivialement la projection π de $(\Sigma_1/\Sigma_o)^*$ sur Σ_1^*.

Enfin on définit $\bar{\phi}$ de Σ_o^* dans $(\Sigma_1/\Sigma_o)^*$ par

$$\bar{\phi}(\sigma(x_1,\ldots,x_n)) = \langle\sigma',\sigma\rangle\ (t_1,\ldots,t_p) \text{ ssi}$$

$$\phi(\sigma(x_1,\ldots,x_n)) = \sigma'(t_1,\ldots,t_p)$$

On a donc par définition

$$\phi = \pi.\bar{\phi}$$

et $\bar{\phi}$ est une bijection entre Σ_o^* et $\bar{\phi}(\Sigma_o^*)$ dont l'inverse sur $\bar{\phi}(\Sigma_o^*)$ est π. De plus $\bar{\phi}(\Sigma_o^*)$ est une forêt régulière.

On transforme d'abord τ en τ' en "marquant" de toutes les façons possibles les parties droites des règles de τ :

$$q[t(x_1,\ldots,x_n)] \to t'(q_{i_1}[x_{i_1}],\ldots,q_{i_p}[x_{i_p}]) \text{ est une règle de } \tau'$$

ssi

$$q[t(x_1,\ldots,x_n)] \to \pi(t'(q_{i_1}[x_{i_1}],\ldots,q_{i_p}[x_{i_p}]) \text{ est une règle de } \tau.$$

Il est évident que $\hat{\tau}' = \{(t,t') \mid (t,\pi(t')) \in \hat{\tau}\}$ et que si τ est finiment réducteur, τ' l'est aussi. τ' appartient donc à TG_D^-.

D'après la proposition I.3, il existe $\tau'' \in TG_D^-$ tel que $\hat{\tau}'' = \{(t,t') \in \hat{\tau}' \mid t' \in \bar{\phi}(\Sigma_o^*)\}$. On ne déduit que $\hat{\tau}'' = \{(t,t') \mid t' \in \bar{\phi}(\Sigma_o^*), (t,\pi(t')) \in \hat{\tau}\}$ et donc que $\hat{\tau}''.\hat{\bar{\phi}}^{-1} = \{(t,t') \mid (t,\bar{\phi}(t'')) \in \hat{\tau}''\} =$

$$\{(t,t'') \mid \bar{\phi}(t'') \in \bar{\phi}(\Sigma_o^*), (t,\pi(\bar{\phi}(t''))) \in \hat{\tau}\} =$$

$$\{(t,t'') \mid (t,\phi(t'')) \in \hat{\tau}\} = \hat{\tau}.\hat{\phi}^{-1}.$$

Il suffit donc de montrer que $\hat{\tau}''.\hat{\bar{\phi}}^{-1} \in TG_D^-$.

Ⓑ Puisque $\hat{\tau}'' \in \widehat{TG_D^-}$, on peut supposer que τ'' ne contient aucune règle réductrice, en remplaçant les règles de τ'' par les règles dérivées obtenues en appliquant successivement des règles de τ'' dont une seule est non-réductrice. Soit donc R l'ensemble des règles de τ'' qui sont toutes non-réductrices.

On considère maintenant l'alphabet $\Sigma_1 \cup \Sigma_1 \times \Sigma_o \cup \Sigma_1 \times R$ et la projection π' qui est l'identité sur $\Sigma_1 \cup \Sigma_1 \times \Sigma_o$ et qui vaut σ pour $\langle\sigma,r\rangle \in \Sigma_1 \times R$.

On considère le transducteur $\tau''' \in TG_D$ défini par

$q[t(x_1,\ldots,x_n)] \to t'(q_{i_1}[x_{i_1}],\ldots,q_{i_p}[x_{i_p}])$ est une règle de τ''' ssi

$$r = q[t(x_1,\ldots,x_n)] \to t''(q_{i_1}[x_{i_1}],\ldots,q_{i_p}[x_{i_p}])$$

est une règle de τ'' où t' est obtenu en marquant le sommet de t" par r si ce sommet n'est pas déjà marqué par un élément de Σ_o.

Il est clair par construction que

$(t,t') \in \hat{\tau}''$ ssi $\exists$ t" tel que $(t,t'') \in \hat{\tau}'''$ et $\pi'(t'') = t'$.

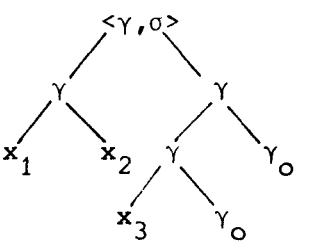
De plus, il y a bijection entre les règles de τ'' et celles de τ''' ; on identifiera donc leurs noms et la marque de R figurant au sommet de l'arbre en partie droite d'une règle de τ''' sera le nom de cette règle.

Intuitivement, ces marques ainsi introduites auront donc pour rôle de mémoriser sur tout arbre $\in$ codom $(\hat{\tau}''')$, les règles de τ''' qui ont permis de l'obtenir dans la transduction.

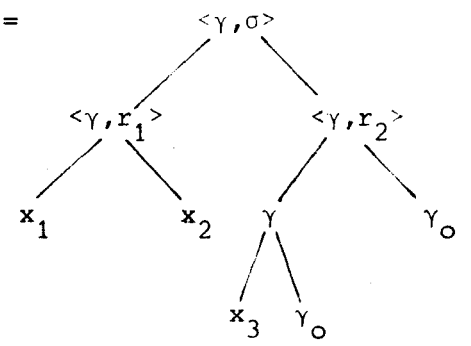
Ⓒ Soit σ un symbole de Σ_o et $\bar{\phi}(\sigma(x_1,\ldots,x_n)) = t(x_1,\ldots,x_n)$. Le sommet de t est marqué par σ. On marque les autres noeuds de t, pas nécessairement tous, par des éléments de R. En faisant ceci de toutes les façons possibles, on forme un ensemble $T_\sigma = \{u \mid \pi'(u) = \bar{\phi}(\sigma)\}$.

Soit $t \in T_\sigma$; on lui associe l'ensemble S_t des sous-arbres de t qui contiennent des variables (y compris les arbres réduits à une variable).

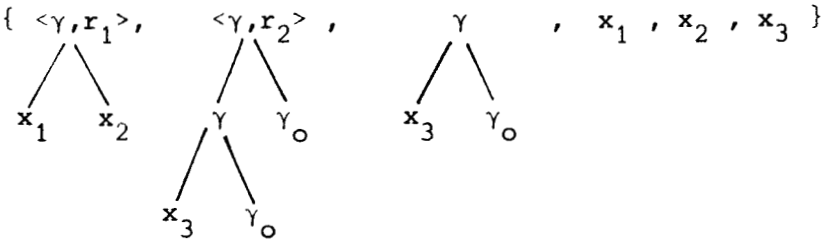
<u>Exemple</u> Soit $\{r_1,r_2\} \subset R$ et $\bar{\phi}(\sigma(x_1,x_2,x_3)) =$



Alors t =



est un arbre de T_σ et

$S_t = \{$  $\}$

fin exemple.

Pour $t \in T_\sigma$, on considère l'ensemble d'états :

$$Q_t = \{q^x_{t'} \mid t' \in S_t \text{ et } x \text{ est une variable de } t'\}$$

Si x est une variable de t, alors $x \in S_t$ et on pose alors $q^x_x = q_o$.

On peut supposer que $t, t' \in \bigcup_\sigma T_\sigma$ et $t \neq t' \Longrightarrow Q_t \cap Q_{t'} = \{q_o\}$.

On construit le transducteur μ défini sur $\bar{\phi}(\Sigma_o^*)$ dont l'ensemble des états est $Q_\mu = \bigcup_{t \in \bigcup_\sigma T_\sigma} Q_t$ dont l'état initial est q_o et dont les règles sont :

$$1) \quad q_o[\langle\gamma,\sigma\rangle(y_1,\ldots,y_p)] \to \sigma(q^{x_1}_{t_{i_1}}[y_{i_1}],\ldots,q^{x_n}_{t_{i_n}}[y_{i_n}])$$

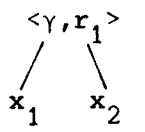
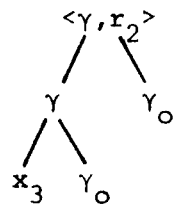
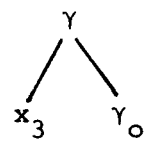
ssi $\exists\, t(x_1,\ldots,x_n) \in T_\sigma$ tel que $t = \langle\gamma,\sigma\rangle(t_1,\ldots,t_p)$ et tel que t_{i_j} est l'unique sous-arbre t_i qui contient la variable x_j. (L'unicité vient de ce que $\bar{\phi}(\sigma(x_1,\ldots,x_n))$ est égal à $t_\sigma(x_1,\ldots,x_n) = \pi'(t(x_1,\ldots,x_n))$).

$$2) \quad q^x_{t'}[\bar{\gamma}(y_1,\ldots,y_k)] \to q^x_{t''_i}[y_i]$$

ssi $t' = \bar{\gamma}(t''_1,\ldots,t''_k)$ et t''_i est l'unique arbre qui contient la variable x, où $\bar{\gamma} \in \Sigma_1 \cup \Sigma_1 \times R$.

Exemple. On reprend l'exemple ci-dessus :

$$S_t = \{t_1,t_2,t_3,x_1,x_2,x_3\}$$

où t_1 =  ; t_2 =  ; t_3 = 

Q_t est alors $\{q_{t_1}^{x_1}, q_{t_1}^{x_2}, q_{t_2}^{x_3}, q_{t_3}^{x_3}, q_o\}$

et μ contiendra les règles

$$q_o[<\gamma,\sigma>(y_1,y_2)] \to \sigma(q_{t_1}^{x_1}[y_1], q_{t_1}^{x_2}[y_1], q_{t_2}^{x_3}[y_2])$$

$$q_{t_1}^{x_1}[<\gamma,r_3>(y_1,y_2)] \to q_o[y_1]$$

$$q_{t_2}^{x_2}[<\gamma,r_3\ (y_1,y_2)] \to q_o[y_2]$$

$$q_{t_2}^{x_3}[<\gamma,r_4>(y_1,y_2)] \to q_{t_3}^{x_3}[y_1]$$

$$q_{t_3}^{x_3}[\gamma(y_1,y_2)] \to q_o[y_1]$$

<u>fin exemple</u>.

Par construction, le transducteur μ fonctionne de façon déterministe quand il est dans un état q_t^x. On en déduit facilement que

$$q_o[t(x_1,\ldots,x_n)] \overset{*}{\underset{\mu}{\Longrightarrow}} \sigma(q_o[x_1],\ldots,q_o[x_n])$$

ssi $(t(x_1,\ldots,x_n) \in T_\sigma$ et $\pi'(t(x_1,\ldots,x_n)) = \bar{\phi}(\sigma(x_1,\ldots,x_n)))$

et donc que $(t,t') \in \hat{\mu} \iff \pi'(t) = \bar{\phi}(t')$.

On a donc $(t,\bar{\phi}(u)) \in \hat{\tau}$" ssi

$\exists$ t" tel que $(t,t") \in \hat{\tau}$"' et $\pi'(t") = \bar{\phi}(u)$

ssi $\exists$ t" tel que $(t,t") \in \hat{\tau}$"' et $(t",\bar{\phi}(u)) \in \hat{\mu}$

ssi $(t,u) \in \hat{\tau}''' \circ \hat{\mu}$. Il s'ensuit que $\hat{\tau}'''.\hat{\mu} = \hat{\tau}''.\hat{\phi}^{-1}$.

(D) On va maintenant construire un transducteur $\Theta \in TG_{\bar{D}}$ tel que $\hat{\Theta} = \hat{\tau}'''.\hat{\mu}$, ce qui achèvera la démonstration.

L'ensemble des états de Θ est $Q_\mu \times Q_{\tau''}$ $(= Q_\mu \times Q_{\tau'''})$; ses états initiaux sont $\langle q_o, q\rangle$ où q est un état initial de $Q_{\tau''}$. Ses règles sont

$$\langle q', q\rangle\,[t(x_1,\ldots,x_n)] \to u(\langle q'_{i_1}, q_{i'_1}\rangle\,[x_{i''_1}],\ldots,\langle q'_{i_{p'}}, q_{i'_p}\rangle\,[x_{i''_p}])$$

ssi il existe une règle de τ''

$$r = q[t(x_1,\ldots,x_n)] \to t'(q_{j_1}[x_{j'_1}],\ldots,q_{j_K}[x_{j'_K}])$$

telle que . $q' = q_o$ ou $q' = q^{x}_{\bar{t}}$ où le sommet de $\bar{t}$ est marqué par r

$$. \; q'[t'(y_1,\ldots,y_K)] \overset{*}{\underset{\mu}{\Longrightarrow}} u(q'_{i_1}[y_{\ell_1}],\ldots,q'_{i_p}[y_{\ell_p}])$$

où en remplaçant dans la partie droite de cette dérivation $q'_m[y_{m'}]$ par $\langle q'_m, q_{jm'}\rangle[x_{j'_{m'}}]$ on retrouve la partie droite de la règle.

Par construction de Θ, on a bien $\hat{\tau}'''.\hat{\mu} \subset \hat{\Theta}$.

Pour démontrer réciproquement que $\hat{\Theta} \subset \hat{\tau}'''.\hat{\mu}$ on remarque d'abord que les seuls états figurant en partie gauche des règles de Θ sont de la forme $\langle q_o, q\rangle$ ou $\langle q^{x}_{\bar{t}}, q\rangle$ où le sommet de $\bar{t}$ est marqué par une règle de $\bar{R}$. Par ailleurs si $\langle q^{x}_{\bar{t}}, q\rangle[y]$ figure en partie droite d'une règle de Θ, tous les $\langle q^{x'}_{\bar{t}}, q\rangle[y]$ tels que x' est une variable figurant dans $\bar{t}$, y figurent aussi.

On démontre alors par induction que :

$\langle q_o, q\rangle[t] \overset{*}{\underset{\Theta}{\Longrightarrow}} u$ implique qu'il existe t' tel que

$$q[t] \overset{*}{\underset{\tau'''}{\Longrightarrow}} t' \quad \text{et} \quad q_o[t'] \overset{*}{\underset{\mu}{\Longrightarrow}} u$$

ou

si $\{x_{i_1},\ldots,x_{i_n}\}$ est l'ensemble des variables de $\bar{t} = \bar{t}(x_{i_1},\ldots,x_{i_n})$ dont le sommet est marqué par r ;

si $q_{\bar{t}}^{x_{i_j}}[t] \overset{*}{\underset{\theta}{\Longrightarrow}} u_j$ pour tout $j \in [1,n]$

alors il existe $v_1,\ldots,v_n$ tels que

$$\forall\, j \in [1,n] \quad q_{\bar{t}}^{x_{i_j}}[v_j] \overset{*}{\underset{\mu}{\Longrightarrow}} u_j$$

$$\text{et } q[t] \overset{*}{\underset{\tau'''}{\Longrightarrow}} \bar{t}(v_1,\ldots,v_n)$$

La démonstration de cette proposition s'appuie sur le fait que toute dérivation $q_{\bar{t}}^{x_{i_j}}[t] \overset{*}{\underset{\theta}{\Longrightarrow}} u_j$ commence par la règle de θ obtenue à partir de la même règle r de τ''' et que si $q_{\bar{t}}^{x}[t_1(x_1,\ldots,x_p)] \overset{*}{\underset{\mu}{\Longrightarrow}} u(\ldots,q[x_i]\ldots)$ et $q_{\bar{t}}^{x'}[t_1(x_1,\ldots,x_n)] \overset{*}{\underset{\mu}{\Longrightarrow}} u'(\ldots,q'[x_j]\ldots)$, les ensembles de variables figurant dans u et u' sont disjoints.

cqfd.

Proposition IV.3 $\quad TG_{\bar{D}} \,.\, TI \subset TG_{\bar{D}}$

En effet. Soit $\hat{\tau} \in \widehat{TG_{\bar{D}}}$ et $\mu = \langle F,\phi,\psi\rangle$ où ϕ est ordonné. D'après le lemme IV.2 $\{(t,t') \mid (t,\phi(t')) \in \hat{\tau}\} = \hat{\tau}.\{(\phi(t),t)\} \in \widehat{TG_{\bar{D}}}$ et d'après la proposition I.2 $\hat{\tau}.\{(\phi(t),t) \mid t \in F\} \in TG_{\bar{D}}$. D'après IV.1, $\hat{\tau}.\{(\phi(t),\psi(t)) \mid t \in F\} = \hat{\tau}.\hat{\mu} \in \widehat{TG_{\bar{D}}}$.

cqfd.

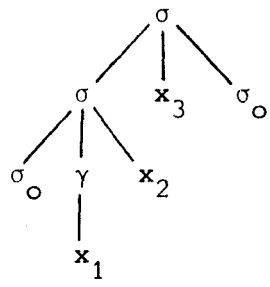
Proposition IV.4 $\quad \forall\, n \geq 1 \quad (\widehat{TI})^n \subset \widehat{TG_{\bar{D}\bar{D}}}$

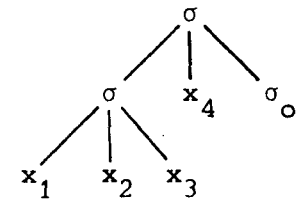
En effet. On a vu en I que $\widehat{TI} \subset TG_{\bar{D}\bar{D}}$. D'après la proposition IV.3, on en déduit immédiatement par récurrence que pour tout $n \geq 1$ $(\widehat{TI})^n \subset TG_{\bar{D}}$. Par ailleurs si $\hat{\tau} \in (\widehat{TI})^n$, il est clair que $\hat{\tau}^{-1} \in (TI)^n$ et donc $\hat{\tau}^{-1} \in TG_{\bar{D}}$. On en déduit que $\hat{\tau} \in \widehat{TG_{\bar{D}\bar{D}}}$.

cqfd.

Lemme IV.5 $\quad \widehat{TG}_{DD} \subset (TI)^2$

En effet. Soit $\hat{\tau} \in \widehat{TG}_{DD}$ et soit Σ l'alphabet gradué tel que dom $(\hat{\tau}) \subset \Sigma^*$. Soit X l'ensemble infini de variables $\{x_1,\ldots,x_n,\ldots\}$ et soit T l'ensemble des arbres de $\Sigma^*(X)$ dont les variables sont, de gauche à droite, $x_1,\ldots,x_n$. Pour tout entier ℓ, et tout arbre $t \in T$, on définit t/ℓ comme le sous-arbre initial de T obtenu en coupant les branches de t à la longueur ℓ et dont les variables sont numérotées convenablement.

Exemple. Si t = 

t/2 = 

On prend ℓ égal à la profondeur maximale des arbres figurant en partie gauche des règles de τ ($\ell = \Pi_g(\tau)$) et on pose $T/\ell = \{t/\ell \mid t \in T\}$.

Ⓐ On considère l'alphabet gradué $\Sigma' = \Sigma \times T/\ell$ où $d(<\sigma,t>) = d(\sigma)$ et la projection $\Pi : \Sigma' \to \Sigma$ qui à $<\sigma,t>$ fait correspondre σ.

On définit le transducteur non généralisé ascendant μ par :

- l'ensemble Q_μ des états est $\{q_t \mid t \in T/\ell\}$
- l'ensemble des états finaux est Q_μ
- les règles de μ sont

$$\sigma \to q_\sigma[<\sigma,\sigma>] \quad \text{si} \quad \sigma \in \Sigma_o$$

$$\sigma(q_{t_1}[x_1],\ldots,q_{t_n}[x_n]) \to q_t[<\sigma,t>(x_1,\ldots,x_n)]$$

$$\text{où } t = \sigma(t_1,\ldots,t_n)/\ell \qquad ①$$

① On peut supposer qu'un arbre quelconque de $\Sigma^*(X)$ est dans T après avoir renuméroté canoniquement ses variables. Ceci permettra en particulier d'utiliser l'identité $\sigma(t_1,\ldots,t_n)/\ell = \sigma(t_1/\ell\text{-}1,\ldots,t_n/\ell\text{-}1)$.

Par construction, le transducteur μ est déterministe et son domaine est Σ^* ; d'où pour tout $t \in \Sigma^*$ il existe un seul $t' \in \Sigma'^*$ noté $\mu(t)$ tel que $(t,t') \in \hat{\mu}$. On démontre aidément que si $t = \sigma(t_1,\ldots,t_n)$ alors $\mu(t) = \langle\sigma,t/\ell\rangle\ (\mu(t_1),\ldots,\mu(t_n))$. Comme de plus $\pi(\mu(t)) = t$ on a :

si $t' = \langle\sigma,t\rangle(t_1,\ldots,t_n) \in \mu(\Sigma^*)$ alors $t = \pi(t')/\ell$

et $t_1,\ldots,t_n \in \mu(\Sigma^*)$

$\mu(\Sigma^*)$ est une forêt régulière qui est le domaine du transducteur non-généralisé descendant déterministe ρ défini par

- L'ensemble de ses états est $Q_\rho = \{q_o\} \cup \{q_t \mid t \in T/\ell\text{-}1\}$
  et son état initial est q_o

- Ses règles sont :

1) $q_\sigma[\langle\sigma,\sigma\rangle] \to \langle\sigma,\sigma\rangle$

2) $q_{t'}[\langle\sigma,t\rangle(x_1,\ldots,x_n)] \to \langle\sigma,t\rangle(q_{t_1}[x_1],\ldots,q_{t_n}[x_n])$

ssi $t = \sigma(t_1,\ldots,t_n)$ ② et $t' = t/\ell\text{-}1$.

3) $q_o[t(x_1,\ldots,x_n)] \to t'(q_{i_1}[x_1],\ldots,q_{i_n}[x_n])$

ssi $\exists\ \bar{t}$ tel que $q_{\bar{t}}[t(x_1,\ldots,x_n)] \to t'(q_{i_1}[x_1],\ldots,q_{i_n}[x_n])$

est déjà une règle de ρ.

Ce transducteur est bien déterministe par définition :

si $q[\langle\sigma,t\rangle(x_1,\ldots,x_n)] \to \langle\sigma,t\rangle(q_{t_1}[x_1],\ldots,q_{t_n}[x_n])$

et $q[\langle\sigma,t\rangle(x_1,\ldots,x_n)] \to \langle\sigma,t\rangle(q_{t'_1}[x_1],\ldots,q_{t'_n}[x_n])$

alors $t = \sigma(t_1,\ldots,t_n) = \sigma(t'_1,\ldots,t'_n)$ d'où $t_i = t'_i \quad \forall i$.

---

② A la numérotation près des variables.

Par ailleurs il est évident que si $\rho(t)$ est défini, alors $\rho(t) = t$. Montrons maintenant que $\rho(t)$ est défini ssi $t \in \mu(\Sigma^*)$. Pour cela on montre aisément par induction que :

$$q_{\bar{t}}[t] \underset{\rho}{\overset{*}{\Longrightarrow}} t \quad \text{ssi} \quad t \in \mu(\Sigma^*) \text{ et } \bar{t} = t/\ell\text{-}1.$$

(B) On définit maintenant le transducteur généralisé descendant τ_1 défini sur Σ'^*

- les états et les états initiaux de τ_1 sont ceux de τ

- $q[t(x_1,\ldots,x_n)] \to t'(q_{i_1}[x_{i_1}],\ldots,q_{i_p}[x_{i_p}])$

  est une règle de τ_1 ssi

  $$q[\pi(t(x_1,\ldots,x_n)] \to t'(q_{i_1}[x_{i_1}],\ldots,q_{i_p}[x_{i_p}])$$

  est une règle de τ.

Par construction $(t,t') \in \hat{\tau}_1$ ssi $(\pi(t),t') \in \hat{\tau}$.

BUS LILLE

Puisque $\hat{\tau} \in TG_{DD}$, il existe $\tau' \in TG_D$ tel que $\hat{\tau}' = \hat{\tau}^{-1}$. On construit alors τ'_1 dont les états et les états initiaux sont ceux de τ' et dont les règles sont définies par :

$$q'[t'(x_1,\ldots,x_n)] \to t(q'_{i_1}[x_{i_1}],\ldots,q'_{i_p}[x_{i_p}])$$

est une règle de τ'_1 ssi

$$q'[t'(x_1,\ldots,x_n)] \to \pi(t(q'_{i_1}[x_{i_1}],\ldots,q'_{i_p}[x_{i_p}]))$$

est une règle de τ'.

Par construction $\tau'_1 \in TG_D$ et

$(t',t) \in \hat{\tau}'_1$ ssi $(t',\pi(t)) \in \hat{\tau}$. On en déduit que $\hat{\tau}'_1 = \hat{\tau}_1^{-1}$ et donc que $\hat{\tau}_1 \in \widehat{TG}_{DD}$.

© On montre maintenant qu'il existe $\hat{\tau}' \in \widehat{TG}_{DD}$ tel que $(t,t') \in \hat{\tau}'$ ssi $(t,t') \in \hat{\tau}_1$ et $t \in \mu(\Sigma^*)$.

On en déduira que $\hat{\tau} = \{(\pi(t),t)\}.\hat{\tau}'$. En effet, si $(t,t') \in \hat{\tau}$, $(\mu(t),t') \in \hat{\tau}'$ et comme $\pi(\mu(t)) = t \quad (t,\mu(t)) \in \{(\pi(t),t)\}$ alors $\hat{\tau} \subset \{(\pi(t),t)\}.\hat{\tau}'$. Réciproquement si $(t,t') \in \{(\pi(t),t)\}.\hat{\tau}'$, il existe $t'' \in \mu(\Sigma^*)$ tel que $t = \pi(t'')$ et $(t'',t') \in \hat{\tau}_1$ ce qui est équivalent à $(t,t') \in \hat{\tau}$.

Tout d'abord puisque $\hat{\tau}_1^{-1} \in TG_D$, $\hat{\tau}'^{-1} = \{(t,t') \in \hat{\tau}_1^{-1} \mid t' \in \mu(\Sigma^*)\}$ appartient à $\widehat{TG}_D$ d'après la proposition I.2. Il suffit donc de montrer que $\{(t,t') \in \hat{\tau}_1 \mid t \in \mu(\Sigma^*)\} = \hat{p}.\hat{\tau}_1 \in \widehat{TG}_D$.

On construit le transducteur τ' dont l'ensemble des états est $Q_\rho \times Q_{\tau_1}$, l'ensemble des états initiaux est $\{<q_o,q> \mid q$ est un état initial de $Q_{\tau_1}\}$ et dont les règles sont définies par :

$$<q,q'>[t(x_1,\ldots,x_n)] \to t'(<q_{i_1},q'_{i_1}>[x_{i_1''}],\ldots,<q_{i_p},q'_{i_p}>[x_{i_p''}])$$

est une règle de τ' ssi

$$q'[t(x_1,\ldots,x_n)] \to t'(q'_{i_1}[x_{i_1''}],\ldots,q'_{i_p}[x_{i_p''}])$$

est une règle de τ_1 et

$$q[t(x_1,\ldots,x_n)] \overset{*}{\underset{\rho}{\Longrightarrow}} t(q_{j_1}[x_1],\ldots,q_{j_n}[x_n])$$

avec $i_\ell = j_{i_\ell''}$.

Il est immédiat par construction que $\tau' \in TG_D \quad \hat{\rho}.\hat{\tau}_1 \subset \hat{\tau}' \subset \hat{\tau}_1$.

Supposons que $\hat{\rho}.\hat{\tau}_1 \subsetneq \hat{\tau}'$. Il existe alors $(t,t') \in \hat{\tau}'$ tel que $(t,t') \notin \hat{\rho}.\hat{\tau}_1$ ce qui entraîne que $\rho(t)$ n'est pas défini, car si $\rho(t)$ était défini, on aurait $(t,t) \in \hat{\rho}$ et comme $(t,t') \in \hat{\tau}' \subset \hat{\tau}_1$, (t,t') appartiendrait à $\hat{\rho}.\hat{\tau}_1$.

Comme $\rho(t)$ n'est pas défini, il existe nécessairement $t_1,\ldots,t_n$, $t_o(x_1,\ldots,x_n)$ tels que $t = t_o(t_1,\ldots,t_n)$ et $q_o[t] \overset{*}{\underset{\rho}{\Longrightarrow}} t_o(q_{i_1}[t_1],\ldots,q_{i_n}[t_n])$ et $\forall\, j \in [1,n]$ $t_j = \gamma(t'_1,\ldots,t'_{n_j})$ et ρ ne contient aucune règle dont la partie gauche est $q_{i_j}[\gamma(x_1,\ldots,x_{n_j})]$.

Fixons un de ces t_j qu'on notera t_{j_o}. Cet arbre ne sera jamais "lu" dans la dérivation $t \overset{*}{\Longrightarrow} t'$ de τ', car tout ce qui est "lu" par τ' est nécessairement reconnu par ρ. Si on remplace t_{j_o} dans t par n'importe quel arbre de Σ'^* on obtiendra un arbre t" tel que $t" \overset{*}{\Longrightarrow} t'$. $\hat{\tau}'$ ne sera donc pas fidèle, et comme $\hat{\tau}' < \hat{\tau}_1$, $\hat{\tau}_1$ ne sera pas non plus fidèle, ce qui est impossible, car $\hat{\tau}_1 \in \widehat{TG}_{DD}$. On a donc bien $\hat{\tau}' = \hat{\rho}.\hat{\tau}_1$.

Ⓓ Pour terminer, on montre qu'il existe $\tau''' \in TG_Q$ tel que $\hat{\tau}''' = \hat{\tau}'$. Comme $\hat{\tau}' \in \widehat{TG}_{DD}$, on en déduit que $\hat{\tau}''' \subset \widehat{TG}_{QD}$.

A chaque règle $r = q[t(x_1,\ldots,x_n)] \to t'(\ldots,q_{i_j}[x_{i_j}],\ldots)$ de τ' (où $q, q_{i_j} \in Q_\tau$, $= Q_\rho \times Q_{\tau_1}$) on associe l'ensemble E_r des sous-arbres de $t(x_1,\ldots,x_n)$ qui contiennent des variables. Chaque arbre de E_r non réduit à une variable admet une décomposition unique $\sigma(t_1,\ldots,t_p)$ où t_i est soit un arbre sans variable, soit un arbre de E_r. A chaque règle r on associe l'ensemble d'états $Q_r = \{q_t^x \mid t \in E_r$ et x est une variable de t$\}$. On peut supposer que $r \neq r' \Longrightarrow Q_r \cap Q_{r'} = \emptyset$ et on pose $Q = \bigcup_r Q_r$.

On construit d'abord $\tau" \in TG_Q$ dont l'ensemble d'états est $Q_\tau \cup Q \times Q_\tau$, et dont les états initiaux sont ceux de τ'. Les règles de ce transducteur sont :

1) $q[\langle\sigma,t'\rangle(u_1,\ldots,u_n)] \to t"(s_{i_1}[u_{i_1}],\ldots,s_{i_p}[u_{i_p}])$

où $q \in Q_\tau$,

où u_i est soit une variable, soit un arbre sans variable et où

les u_{i_j} sont tous des variables.

ssi ∃ $t(x_1,\ldots,x_k)$, $t_1,\ldots,t_n$ tels que

a) si u_i n'est pas une variable $u_i = t_i$

b) $t(x_1,\ldots,x_k) = \langle\sigma,t'\rangle(t_1,\ldots,t_n)$

c) $q[t(x_1,\ldots,x_k)] \to t''(q_{j_1}[x_{j_1}],\ldots,q_{j_p}[x_{j_p}])$
est une règle de τ'

d) $\pi(t)$ est un sous-arbre initial de t'

e) pour $\ell = 1,\ldots,p$, soit $\bar{t}_\ell$ l'unique arbre de $\{t_1,\ldots,t_n\}$ qui contient la variable $x_{j'_\ell}$; alors

$$s_{i_\ell} = \begin{cases} q_{j_\ell} \text{ si } \bar{t}_\ell = x_{j'_\ell} \\ x_{j'_\ell} \\ \langle q_{\bar{t}_\ell}, q_{j_\ell}\rangle \text{ sinon} \end{cases}$$

2) $\langle q^x_t, q\rangle\,[\langle\sigma,t'\rangle\,(u_1,\ldots,u_n)] \to s[u_i]$

où les u_j sont soit des variables, soit des arbres sans variables, et où u_i est une variable.

ssi ∃ $t'_1,\ldots,t'_n$ tels que

a) si u_j n'est pas une variable, $u_j = t'_j$

b) $t = \langle\sigma,t'\rangle(t'_1,\ldots,t'_n)$

c) $\pi(t)$ est un sous-arbre initial de t'

d) t'_i est l'unique arbre e $\{t'_1,\ldots,t'_n\}$ qui contient la variable x

$$e)\ s = \begin{cases} q \text{ si } t'_i = x \\ \langle q^x_{t'_i}, q\rangle \text{ sinon} \end{cases}$$

On montre d'abord que $\hat{\tau}' \subset \hat{\tau}''$, en associant à toute règle de τ' une règle de τ'' de type 1 et plusieurs règles de type 2. Pour vérifier les conditions 1d) et 2c), on remarque que si $q[t(t_1,\ldots,t_n)] \underset{\tau'}{\Longrightarrow} t'(\ldots q_{i_j}[t_{i_j}]\ldots) \overset{*}{\underset{\tau'}{\Longrightarrow}} u$. Alors le sommet de t est marqué par $\pi(t(t_1,\ldots,t_n))/\ell = t'$ et que, comme t a toutes ses branches de longueur inférieure à ℓ, on a bien $\pi(t)$ sous-arbre initial de t'.

Pour démontrer la réciproque, on montre d'abord, en tenant compte du fait que τ'' est déterministe quand il est dans un état q^x_t que

$$q[u(x_1,\ldots,x_k)] \overset{*}{\underset{\tau''}{\Longrightarrow}} t''(q_{j_1}[x_{j_1}],\ldots,q_{j_p}[x_{j_p}])$$

est une dérivation formée d'une règle de type 1 suivie uniquement de règles de type 2

ssi $\exists$ $t(x'_1,\ldots,x'_K)$ tel que

- $\pi(t(x'_1,\ldots,x'_K))$ est un sous-arbre initial de la marque du sommet de u
- $q[t(x'_1,\ldots,x'_K)] \to t''(q_{j_1}[x'_{j''_1}],\ldots,q_{j_p}[x'_{j''_p}])$ est une règle de τ'
- $u(x_1,\ldots,x_n)$ est un sous-arbre initial de $t(x'_1,\ldots,x'_K)$ et a comme sous-arbre initial l'arbre obtenu en remplaçant dans t les sous-arbres qui contiennent des variables autres que les $x'_{j''_i}$ par des variables. La variable $x'_{j''_i}$ de t est renumérotée x_{j_i} dans u.

On peut montrer par induction que

$$u \in \mu(\Sigma^*) \text{ et } q[u] \overset{*}{\underset{\tau''}{\Longrightarrow}} u' \text{ implique } q[u] \overset{*}{\underset{\tau'}{\Longrightarrow}} u'$$

Soit $q[u(u_1,\ldots,u_k)] \overset{*}{\underset{\tau''}{\Longrightarrow}} t''(q_{j_1}[u_{j_1}],\ldots,q_{j_p}[u_{j_p}]) \overset{*}{\underset{\tau''}{\Longrightarrow}} u'$
la dérivation qui commence par la dérivation ci-dessus. Il existe donc dans τ' une dérivation $q[t(t_1,\ldots,t_n)] \overset{}{\underset{\tau'}{\Longrightarrow}} t''(q_{j_1}[t''_{j_1}],\ldots,q_{j_p}[t_{j_p''}]) \overset{*}{\underset{\tau'}{\Longrightarrow}} u'$
avec $t''_{j_i} = u_{j_i}$. Si t' est la marque du sommet de t, $\pi(t)$ est un sous-arbre initial de t'. Comme u est un sous-arbre initial de t, la marque du sommet de u est aussi t' et comme $u(u_1,\ldots,u_K) \in \mu(\Sigma^*)$, $t' = \pi(u(u_1,\ldots,u_K))/\ell$. $\pi(t)$ est donc un sous-arbre initial de $\pi(u(u_1,\ldots,u_k))$ et donc $u(u_1,\ldots,u_K) = \bar{t}(t'_1,\ldots,t'_n)$ avec $\pi(\bar{t}) = \pi(t)$. Mais par construction de τ', $q[\bar{t}(x_1,\ldots,x_n)] \to t''(q_{j_1}[x_{j_1''}],\ldots,q_{j_p}[x_{j_p''}])$ est encore une règle de τ'. Par ailleurs, d'après la définition du sous-arbre initial de t qui est sous-arbre initial de $u(u_1,\ldots,u_K) = \bar{t}(t'_1,\ldots,t'_n)$ on a $t'_{j_i''} = t_{j_i''}$ d'où

$$q[\bar{t}(t'_1,\ldots,t'_n)] \overset{}{\underset{\tau'}{\Longrightarrow}} t''(q_{j_1}[t_{j_1''}],\ldots,q_{j_p}[t_{j_p''}]) \overset{*}{\underset{\tau'}{\Longrightarrow}} u'.$$

On a donc démontré que

$$(t,t') \in \hat{\tau}' \Longleftrightarrow (t,t') \in \hat{\tau}'' \text{ et } t \in \mu(\Sigma^*)$$

avec $\tau'' \in TG_Q$.

Il faut donc encore montrer pour obtenir le résultat annoncé plus haut, qu'il existe $\tau''' \in TG_Q$ tel que $(t,t') \in \hat{\tau}''' \Longleftrightarrow (t,t') \in \hat{\tau}''$ et $t \in \mu(\Sigma^*)$ c'est-à-dire $\hat{\tau}''' = \hat{\rho}.\hat{\tau}''$.

On procède de la même façon qu'en Ⓒ pour construire τ''' et on a bien $\hat{\rho}.\hat{\tau}'' \subset \hat{\tau}''' \subset \hat{\tau}''$; de la même façon on montre que si $\hat{\rho}.\hat{\tau}'' \subsetneq \hat{\tau}'''$, il existe $(t(t_o),t') \in \hat{\tau}''$ tel que t_o n'est pas lu dans la transduction par τ''. t_o n'est pas lu non plus si on n'utilise que les règles dérivées définies ci-dessus. En utilisant la relation entre ces règles dérivées et les règles de τ', on montre que $\hat{\tau}'$

n'est pas fidèle, ce qui est exclu puisque $\hat{\tau}' \in TG_{DD}$.

(E) On a donc $(t,t') \in \hat{\tau}''' \iff (t,t') \in \hat{\tau}' \iff (\pi(t),t') \in \hat{\tau}$.

Comme d'après le théorème III.13, $\widehat{TG_{QD}} \subset (\widehat{TI})^2$, alors $\tau''' \in (\widehat{TI})^2$ et il existe $\tau_1 = \langle F_1, \phi_1, \psi_1 \rangle$ et $\tau_2 = \langle F_2, \phi_2, \psi_2 \rangle$ tels que $\hat{\tau}''' = \hat{\tau}_1 . \hat{\tau}_2 = \{(\phi_1(t), \psi_1(t)) \mid t \in F_1\} . \hat{\tau}_2$ d'où $\hat{\tau} = \{(\pi(t),t') \mid (t,t') \in \tau'''\} = \{\pi(\phi_1(t)), \psi_1(t)) \mid t \in F_1\} . \hat{\tau}_2 \in \widehat{TI}^2$.

Théorème IV.6 :

$$\widehat{TG}_{DD} = \widehat{TG}_{\overline{DD}} = (\widehat{TI})^2$$

En effet. D'après la proposition IV.4 et le lemme IV.5

$$\widehat{TI}^2 \subset \widehat{TG}_{\overline{DD}} \subset \widehat{TG}_{DD} \subset \widehat{TI}^2.$$

cqfd.

Théorème IV.7 :

$$\forall\, n \geq 2 \quad (\widehat{TI})^2 = (\widehat{TI})^n$$

En effet. D'après la proposition IV.4

$$(\widehat{TI})^n \subset \widehat{TG}_{\overline{DD}} = (\widehat{TI})^2$$

cqfd.

Remarque. Il existe des exemples permettant d'affirmer que $\widehat{TI} \subsetneq (\widehat{TI})^2$.

Corollaire IV.8 $\quad \widehat{TG}_{DD} \cdot \widehat{TG}_{DD} \subset \widehat{TG}_{DD}$.

Corollaire IV.9 $\quad \widehat{TG}_{DD} = \widehat{TG}_{AA}^{\,2}$

CHAPITRE V

RESULTATS COMPLEMENTAIRES

1. BITRANSDUCTIONS ET FORÊTS RÉGULIÈRES

Lemme V.1 : *Les homomorphismes (linéaires complets stricts) inverses conservent les forêts régulières.*

En effet. Soit $\phi\Sigma^* \to \Delta^*$ et F une forêts régulière de Δ^*. Soit un reconnaisseur ascendant déterministe de F donné par les règles comme dans la 1ère partie (I.3.A).

Considérons alors le reconnaisseur R ayant pour règles

$\{a(q_{i_{\sigma-1(1)}}[x_1],\ldots,q_{i_{\sigma-1(n)}}[x_n]) \to q[a(x_1,\ldots,x_n)] \mid$ telles que

$a \in \Sigma, \phi(a(x_1,\ldots,x_n)) = \phi(a)(x_{\sigma(1)},\ldots,x_{\sigma(n)})$

et $\phi(a)(q_{i_1}[x_1],\ldots,q_{i_n}[x_n] \overset{*}{\Longrightarrow} q[\phi(a)(x_1,\ldots,x_n)]\}$.

Il est facile de montrer que $F(R) = \phi^{-1}(F)$.

cqfd.

Proposition V.2 :

a) Le domaine et le codomaine d'une bitransduction sont des forêts régulières. La restriction des domaines et codomaines à des forêts régulières concerne les bitransductions.

b) Les bitransductions concernent les forêts régulières.

En effet. Il est connu que la classe des forêts régulières est stable par union, intersection et homomorphismes linéaires ([TH],[RO]).

a) Soit $T \in TI$. Il existe $\langle K,\phi,\psi\rangle$ tel que $\hat{T} = \{(\phi(t),\psi(t)) \mid t \in K\}$ donc Dom $(\hat{T}) = \phi(K)$ et Codom $(\hat{T}) = \psi(K)$ sont réguliers.

Soit F une forêt régulière. Considérons

$$\hat{T}_{IF} = \{(\phi(t),\psi(t)) \mid t \in K \text{ et } \phi(t) \in F\}$$

alors $\hat{T}_{IF} = \{(\phi(t),\psi(t)) \mid t \in K \cap \phi^{-1}(F)\}$

donc $\hat{T}_{IF} \in TI$ puisque $K \cap \phi^{-1}(F)$ est régulière (lemme V.1 et rappels ci-dessus).

En restreignant le domaine de $\hat{T}^{-1}$, on restreint le codomaine de $\hat{T}$, d'où le résultat pour $\widehat{TI}$.

Soit $T \in TG_{DD}$. Alors $\hat{T} \in \widehat{TI}^2$ et $\exists\ T_1 = (K,\phi,\psi)$ et $T_2 = (K',\phi',\psi')$ tels que

$$\hat{T} = \hat{T}_1 \circ \hat{T}_2.$$

Soit $F = \psi(K) \cap \phi'(K')$, d'après a) il existe T'_1 et $T'_2 \in TI$ tels que $\hat{T}'_1 = \{(t,u) \in \hat{T}_1 \mid u \in F\}$ et $\hat{T}'_2 = \{(u,v) \in \hat{T}_2 \mid u \in F\}$ alors $\hat{T} = \hat{T}'_1 \circ \hat{T}'_2$ avec Dom $\hat{T}$ = Dom $\hat{T}'_1$ et Codom $\hat{T}$ = Codom $\hat{T}'_2$. On restreint le domaine de $\hat{T}$ à G en restreignant celui de $\hat{T}'_1$ et le codomaine de $\hat{T}$ en restreignant celui de $\hat{T}'_2$.

b) Soit $T \in TI$ et F régulière

$T(F) = \{u \mid (t,u) \in T_{IF}\}$ d'où $T(F)$ = Codom T_{IF} et $T(F)$ régulière d'après a).

c) Soit $T \in TG_{DD}$. Alors $\hat{T} = \hat{T}_1 \circ \hat{T}_2$ comme dans a) et $T(F) = T_2(T_1(F))$ est régulière.

cqfd.

Montrons maintenant que la réciproque de b) est fausse.

Proposition V.3 : Il existe des transductions concernant les forêts régulières et qui ne sont pas des bitransductions.

En effet.

Soit $\Sigma = \{(a,1),(\bar{a},0)\}$, $\Delta = \{(b,2), (a,1), (\bar{a},0)\}$.

Soit T ϵ TD défini par :

$$q_o\ [a(x)] \to b(q_1[x],q_1[x])$$
$$q_1\ [a(x)] \to a(q_1[x])$$
$$q_1\ [a(x)] \to q_1[x]$$
$$q_1\ [\bar{a}] \to \bar{a}$$

alors $\hat{T} = \{(a^n\bar{a}, b(a^p\bar{a}, a^q\bar{a})) \mid n \geq 1,\ p \text{ et } q \leq n\}$

Soit F régulière sur $a^*\bar{a}$.

Si F est finie, $\hat{T}(F)$ est finie donc régulière.

Si F est infinie, soit $b(a^p\bar{a}, a^q\bar{a})$: il existe $r \geq \sup(p,q)$ tel que $a^r\bar{a} \in F$, donc $\hat{T}(F) \ni b(a^p\bar{a}, a^q\bar{a})$. Nous avons donc alors $\hat{T}(F) = \{b(a^p\bar{a}, a^q\bar{a}) \mid p \text{ et } q \in N\}$ qui est une forêt régulière.

Mais d'autre part, $\hat{T}^{-1}(b(\bar{a},\bar{a})) = \{a^n\bar{a} \mid n \geq 1\}$ donc $\hat{T}$ n'est pas fidèle, donc n'est pas une bitransduction.

cqfd.

2. BITRANSDUCTIONS ET FORÊTS ALGÉBRIQUES

Nous ne pouvons rien dire de la stabilité des forêts algébriques par bitransduction, car la stabilité par homomorphisme inverse est un très important problème ouvert.

3. $\widehat{TI}$ ET $\widehat{TI}^2$ SONT STABLES PAR RÉUNION, PAS PAR INTERSECTION

a) Soient T et T' $\in TG_D$.

Soit Q l'alphabet des états de T, Q' celui de T'.

Considérons $\bar{Q}$ disjoint de Q' et en bijection avec Q et un nouvel état initial k_o.

Soit T" $\in TG_D$ défini par l'ensemble de règles R" suivant

i) $k_o[t(x_1,\ldots,x_n)] \to u(q_{i_1}[x_{i_1}],\ldots,q_{i_p}[x_{i_p}]) \in R"$

ssi $q_o[t(x_1,\ldots,x_n)] \to u(q_{i_1}[x_{i_1}],\ldots,q_{i_p}[x_{i_p}])$

est une règle de T et q_o état initial.

ii) $k_o[t(x_1,\ldots,x_n)] \to u(\bar{q}_{i_1}[x_{i_1}],\ldots,\bar{q}_{i_p}[x_{i_p}]) \in R"$

ssi $q'_o[t(x_1,\ldots,x_n)] \to u(q'_{i_1}[x_{i_1}],\ldots,q'_{i_p}[x_{i_p}])$

est une règle de T' et q'_o état initial.

iii) R" contient toutes les règles de T d'état gauche non initial.

iv) R" contient toutes les règles de T' d'état gauche non initial, où chaque q_i est remplacé par $\bar{q}_i$.

Il est facile de voir que $\hat{T}" = \hat{T} \cup \hat{T}'$. De plus, si T et T' $\in TI$, T" aussi, donc $\widehat{TI}$ est stable par réunion.

Soient T et T' $\in TG_{DD}$.

Alors T" $\in TG_D$. Mais il existe U et U' $\in TG_D$ tels que $\hat{U} = \hat{T}^{-1}$ et $\hat{U}' = \hat{T}'^{-1}$. D'après ce qui précède, il existe U" $\in TG_D$ tel que $\hat{U}" = \hat{U} \cup \hat{U}'$.

Mais $\hat{U} \cup \hat{U}' = \hat{T}^{-1} \cup \hat{T}'^{-1} = (\hat{T} \cup \hat{T}')^{-1} = \hat{T}"^{-1} = \hat{U}"$.

Donc T" $\in TG_{DD}$ et $\widehat{TG_{DD}} = \widehat{TI^2}$ est stable par réunion.

cqfd.

b) Soit T $\in$ TI défini par les règles

$r_1 : q_o[c(x_1,x_2,x_3)] \to c(q_1[x_1],q_2[x_2],q_3[x_3])$

$r_2 : q_1[a(x)] \to b(\alpha,q_1[x])$

$r_3 : q_2[a(x)] \to b(\beta,q_2[x])$

$r_4 : q_3[a(x)] \to b(\gamma,q_3[x])$

$r_5 : q_1[\delta] \to \alpha$

$r_6 : q_2[\delta] \to \beta$

$r_7 : q_3[\delta] \to \gamma$

Soit T' $\in$ TI défini par les règles

$r_1' : q_o[c(x_1,x_2,x_3)] \to c(q_1[x_3],q_2[x_1],q_3[x_2])$

$r_2, r_3, r_4, r_5, r_6, r_7$.

Soit $F = \{u \mid \exists\ t,\ (t,u) \in \hat{T} \cap \hat{T}'\}$.

Si $\hat{T} \cap \hat{T}'$ était une bitransduction, F serait régulière. Or le feuillage de F est $\Phi(F) = \{\alpha^n \beta^n \gamma^n \mid n \geq 1\}$. Si F était régulière, $\Phi(F)$ serait algébrique, ce qui n'est pas le cas.

cqfd.

4. UN CAS PARTICULIER OÙ LES TI SE COMPOSENT :

Définition V.4 : *Soit* T : $\langle K,\phi,\psi\rangle$.

T *est dit* 1-n (*respectivement* n-1) *ssi* ϕ (*respectivement* ψ) *est une projection.*

Proposition V.5 : *La classe des* TI_{1-n} *(respectivement* n-1*) est stable par composition.*

En effet. Il suffit d'établir le résultat dans le cas 1-n, puisque si $T = \langle K,\phi,\psi\rangle$ est 1-n, $\hat{T}^{-1}$ est réalisée par $\langle K,\psi,\phi\rangle$ qui est n-1.

Soient $T = \langle K,\phi,\psi\rangle$

et $T' = \langle K',\phi',\psi'\rangle$ où ϕ_1 et ϕ_2 sont des projections.

Soit $\bar{T} = (Q,\Sigma,\Delta,R,\{q_o\})$ le TI associé à T comme dans le chapitre I de la première partie.

Ses règles sont de la forme

$$q[a(x_1,\ldots,x_n)] \to u(q_{i_1}[x_{\gamma(1)}],\ldots,q_{i_n}[x_{\sigma(n)}])$$

avec $a \in \Sigma$ et u de profondeur ≥ 1.

De même les règles de $\bar{T}' = (Q',\Delta,\Gamma,R',\{q'\})$ sont de la forme

$$q'[b(x_1,\ldots,x_n)] \to v(q'_{i_1}[x_{\sigma(1)}],\ldots q'_{i_r}[x_{\sigma(n)}])$$

avec $b \in \Delta$ et $\pi(v) \geq 1$.

Soit $R'' = \{(q,q')[a(x_1,\ldots,x_n)] \to \Theta((q_{i_{\varepsilon(1)}},q'_{\varepsilon^{-1}\circ\sigma\circ\varepsilon(1)})[x_{\sigma\circ\varepsilon(1)}],\ldots$

$$\ldots,(q_{i_{\varepsilon(n)}},q'_{\varepsilon^{-1}\circ\sigma\circ\varepsilon(n)})[x_{\sigma\circ\varepsilon(n)}])\}$$

telles que $\exists\, q[a(x_1,\ldots,x_n)] \to u(q_{i_1}[x_{\sigma(1)}],\ldots,q_{i_n}[x_{\sigma(n)}]) \in R$

et $q'[u(x_1,\ldots,x_n)] \overset{*}{\underset{\bar{T}'}{\Longrightarrow}} \Theta(q'_{i_1}[x_{\varepsilon(1)}],\ldots,q'_{i_n}[x_{\varepsilon(n)}])$

posons $\bar{T}'' = (Q \times Q',\Sigma,\Gamma,R'',\{(q_o,q'_o)\})$

$\bar{T}''$ est associé à T" qui est 1-n vu la forme de R".

Il est facile de montrer par induction que $\hat{\bar{T}}'' = \hat{\bar{T}} \circ \hat{\bar{T}}'$.

cqfd.

5. LA SUITE $(\widehat{TJ}^n)_{n\in N}$ EST UNE HIÉRARCHIE INFINIE STRICTE DE RELATIONS

Des outils permettant une démonstration rigoureuse sont développés dans [AR2]. Mais l'idée intuitive de la démonstration étant simple, nous donnons ici des indications sur la solution.

Définissons par récurrence une famille $(F_q)_{q\in N}$ infinie de forêts régulières

F_o = {forêts finies}

$F \in F_p$, $p \geq q+1$, ssi il existe une grammaire réduite G telle que

i) $F(G) = F$

ii) $\exists$ X et Y variables de G telles que

$$X \overset{*}{\underset{G}{\Longrightarrow}} t(X,Y)$$

iii) La forêt engendrée par G avec Y comme axiome $\in F_{p'}$, avec $p' \geq q$.

Exemple : Soit G_q définie par les règles :

$X_o \to a$

pour i=1,...,q $\quad X_i \to b(X_i, X_{i-1})$

X_q est l'axiome

alors $F_q = F(G_q) \in F_q - F_{q-1}$

Lemme V.6 : $\forall T \in TJ, \forall F \in F_q$

$T(F) \in F_{q'} \Longrightarrow q' \geq q-1$

Ce lemme se démontre par récurrence sur q. On met en évidence des dérivations de F contenant un nombre arbitrairement grand de branches du type ii). Les règles de T sont inversibles et strictes sur ces dérivations qui sont non monadiques (en dérivant arbitrairement loin Y). On montre alors que T(F) a une grammaire ayant la propriété ii). L'hypothèse de récurrence appliquée à iii) permet de conclure.

L'idée de la démonstration est qu'un TJ efface des branches arbitrairement longues seulement si celles-ci sont monadiques.

Lemme V.7 : $\forall\ q \in N,\ \exists\ T_q \in TJ$ *tel que* $T_q(F_q) = F_{q-1}$.

En effet. Il suffit de considérer T_q défini par les règles :

$$q[b(x,y)] \to b(q[x],q[y])$$

$$q[b(x,a)] \to q[x]$$

$$q[a] \to a$$

cqfd.

Nous déduisons de ces deux lemmes que $\forall\ q \in N$,

$$TJ^q(F_q) \ni F_o \text{ et } TJ^{q-1}(F_q) \not\ni F_o$$

$$\text{d'où } \widehat{TJ}^q \supsetneq \widehat{TJ}^{q-1}.$$

BIBLIOGRAPHIE

[AH] AHO A.V.
Currents in the theory of computing.
Prentice Hall (1974).

[AR 1] ARNOLD A. et DAUCHET M.
Transductions inversibles de forêts régulières.
(à paraître).

[AR 2] ARNOLD A. et DAUCHET M.
Une équivalence décidable sur les forêts régulières.
(à paraître).

[AR 3] ARNOLD A. et DAUCHET M.
Un contre-exemple au théorème principal de Maibaum.
(à paraître).

[AR 4] ARNOLD A.
Langages d'arbres.
Publications du Laboratoire de Calcul de Lille n°44 (1974).

[BA] BAKER B.S.
Tree transductions and families of tree languages.
5th. A.C.M. Symposium (1973).

[BJ] BJORNER D.
The synthesis of finite-state syntax directed top-down and bottom-up transducers.
IEEE 11th. Symp. on switching and automata (1970).

[BO] BOASSON L.
Cônes rationnels et familles agréables de langages.
Thèse 3° cycle, Paris VII (1971).

[BR 1] BRAINERD W.S.
The minimalization of tree automata.
Information and Control 13 (1968).

[BR 2] BRAINERD W.S.
Tree generating regular systems.
Information and Control 14 (1969).

[BR 3] BRAINERD W.S.
Semi-thue systems and representations of tree.
IEEE 10th. Symp. on switching and automata (1969).

[BU] BÜCHI J.R. and ELGOT C.C.
Decision problems of weak second-order...
Notices Amer. Math. Soc. (1958).

[CH] CHOMSKY N., SCHÜTZENBERGER M.P.
The algebraic theory of context-free languages.
Computer Programming and Formal Systems. North. Holland Publishing Co. Amsterdam. (1963).

[CO] COURCELLE B.
Le problème de l'équivalence...
Rapport de recherche n°48 IRIA (1974).

[CR] COUSINEAU F.G. et RIFFLET J.M.
Schémas de programme : problèmes d'équivalence et complexité.
Thèse 3° cycle, Paris (1974).

[DO] DONER J.
Tree acceptors and some of their applications.
J. Computer and System Sciences 4 (1970).

[EI] EILENBERG S. et WRIGHT J.B.
Automata in general algebras.
Information and Control. 11-4. (1967).

[EL] ELGOT C.C.
Algebraic theories and program schemes.
Symposium on semantics of Algorithmic languages. Acad. Press

[JO] JOSHI, LEVY, TAKAMASHI
A tree generating system.
Automata, languages and programming. Nivat Ed. (1973).

[LE] LEVY, JOSHI
Some results in tree automata.
Mathematical Systems Theory 6-4 (1973).

[MA] MARTIN-VERE
On syntax directed transduct and tree transducers.
2nd A.C.M. Symp. (1970).

[ME] MEYERS W.J.
Linear representations of tree structure.
3rd Symp. A.C.M. (1971).

[MW] MEZEI et WRIGHT
Algebraic automata and context-free sets.
Information and Control 11 (1967).

[N 11] NIVAT M.
Transductions des langages de Chomsky.
Annales de l'Institut Fourier. Grenoble. (1968).

[N 12] NIVAT M.
On the interpretation of recursive program schemes.
Rapport de recherche n°84 IRIA (1974).

[OG] OGDEN W.F. et ROUNDS W.C.
Compositions of n tree transducers.
Proceedings A.C.M. (1972).

[RO 1] ROUNDS W.C.
Mappings and grammars on tree.
Mathematical Systems Theory. (1971)

[RO 2] ROUNDS W.C.
Tree oriented proofs...
2nd A.C.M. Symp. (1970).

[RO 5] ROSEN B.K.
Tree manipulating...
J.A.C.M. 20-1. (1973).

[SH] SHEPARD C.D.
Languages in general algebras,
1st A.C.M. Symp. (1969).

[TA] TAKAHASHI
Primitive transformations of ...
Automata, languages and programming. Nivat Ed. (1973).

[TH 1] THATCHER J.W.
Characterizing derivation trees of ...
J. Computer and System Sciences (1967).

[TH 2] THATCHER J.W. et WRIGHT J.B.
Generalized finite automata with...
Mathematical Systems Theory 2.1 (1969).

[TH 3] THATCHER J.W.
Generalized 2 sequential machine maps.
J. Computer and System Sciences 4 (1970).

[TH 4] THATCHER J.W.
Transformation 2 transl...
A.C.M. Prac. (1969).

BIBLIOTHEQUE UNIVERSITAIRE
Section de SCIENCES
LILLE