N° d'ordre : 1230

UNIVERSITE DES SCIENCES ET TECHNIQUES DE LILLE

THÈSE

présentée à

L'UNIVERSITE DES SCIENCES ET TECHNIQUES DE LILLE

pour otenir le titre de

DOCTEUR DE 3ème CYCLE

INFORMATIQUE

par

BODONIRINA RATOANDROMANANA

ORDRE ET QUOTIENT DANS LES FAMILLES DE LANGAGES ALGEBRIQUES

Thèse soutenue le 29 novembre 1984 devant la Commission d'Examen

Membres du Jury :

Président	G. COMYN
Rapporteur	M. LATTEUX
Examinateurs	J. BEAUQUIER
	M. DAUCHET
	G.JACOB

P R O F E S S E U R S CLASSE EXCEPTIONNELLE

M. CONSTANT Eugène	I.E.E.A.
M. FOURET René	Physique
M. GABILLARD Robert	I.E.E.A.
M. MONTREUIL Jean	Biologie
M. PARREAU Michel	Mathématiques
M. TRIDOT Gabriel	Chimie
M. VIVIER Emile	Biologie
M. WERTHEIMER Raymond	Physique

P R O F E S S E U R S 1ère classe

M. BACCHUS Pierre	Mathématiques
M. BEAUFILS Jean-Pierre (dét.)	Chimie
M. BIAYS Pierre	G.A.S.
M. BILLARD Jean (dét.)	Physique
M. BOILLY Bénoni	Biologie
M. BOIS Pierre	Mathématiques
M. BONNELLE Jean-Pierre	Chimie
M. BOUGHON Pierre	Mathématiques
M. BOURIQUET Robert	Biologie
M. BREZINSKI Claude	I.E.E.A.
M. CELET Paul	Sciences de la Terre
M. CHAMLEY Hervé	Biologie
M. COEURE Gérard	Mathématiques
M. CORDONNIER Vincent	I.E.E.A.
M. DEBOURSE Jean-Pierre	S.E.S.
M. DYMENT Arthur	Mathématiques

PROFESSEURS 1ère classe (suite)

M. ESCAIG Bertrand	Physique
M. FAURE Robert	Mathématiques
M. FOCT Jacques	Chimie
M. GRANELLE Jean-Jacques	S.E.S.
M. GRUSON Laurent	Mathématiques
M. GUILLAUME Jean	Biologie
M. HECTOR Joseph	Mathématiques
M. LABLACHE COMBIER Alain	Chimie
M. LACOSTE Louis	Biologie
M. LAVEINE Jean Pierre	Sciences de la Terre
M. LEHMANN Daniel	Mathématiques
Mme LENOBLE Jacqueline	Physique
M. LHOMME Jean	Chimie
M. LOMBARD Jacques	S.E.S.
M. LOUCHEUX Claude	Chimie
M. LUCQUIN Michel	Chimie
M. MIGEON Michel Recteur à Grenoble	E.U.D.I.L.
M. MIGNOT Fulbert (dét.)	Mathématiques
M. PAQUET Jacques	Sciences de la Terre
M. PROUVOST Jean	Sciences de la Terre
M. ROUSSEAU Jean-Paul	Biologie
M. SALMER Georges	I.E.E.A.
M. SEGUIER Guy	I.E.E.A.
M. SIMON Michel	S.E.S.
M. STANKIEWICZ François	S.E.S.
M. TILLIEU Jacques	Physique
M. VIDAL Pierre	I.E.E.A.
M. ZEYTOUNIAN Radyadour	Mathématiques

P R O F E S S E U R S 2ème classe

M. ANTOINE Philippe	Mathématiques (Calais)
M. BART André	Biologie
Mme BATTIAU Yvonne	Géographie
M. BEGUIN Paul	Mathématiques
M. BELLET Jean	Physique
M. BERZIN Robert	Mathématiques
M. BKOUCHE Rudolphe	Mathématiques
M. BODARD Marcel	Biologie
M. BOSQ Denis	Mathématiques
M. BRASSELET Jean-Paul	Mathématiques
M. BRUYELLE Pierre	Géographie
M. CAPURON Alfred	Biologie
M. CARREZ Christian	I.E.E.A.
M. CAYATTE Jean-Louis	S.E.S.
M. CHAPOTON Alain	C.U.E.E.P.
M. COQUERY Jean-Marie	Biologie
Mme CORSIN Paule	Sciences de la Terre
M. CORTOIS Jean	Physique
M. COUTURIER Daniel	Chimie
M. CROSNIER Yves	I.E.E.A.
M. CURGY Jean-Jacques	Biologie
Mle DACHARRY Monique	Géographie
M. DAUCHET Max	I.E.E.A.
M. DEBRABANT Pierre	E.U.D.I.L.
M. DEGAUQUE Pierre	I.E.E.A.
M. DELORME Pierre	Biologie
M. DELORME Robert	S.E.S.
M. DE MASSON D'AUTUME Antoine	S.E.S.
M. DEMUNTER Paul	C.U.E.E.P.

PROFESSEURS 2ème classe (Suite 1)

M.	DENEL Jacques	I.E.E.A.
M.	DE PARIS Jean-Claude	Mathématiques (Calais)
Mle	DESSAUX Odile	Chimie
M.	DEVRAINNE Pierre	Chimie
M.	DHAINAUT André	Biologie
Mme	DHAINAUT Nicole	Biologie
M.	DORMARD Serge	S.E.S.
M.	DOUKHAN Jean-Claude	E.U.D.I.L.
M.	DUBOIS Henri	Physique
M.	DUBRULLE Alain	Physique (Calais)
M.	DUBUS Jean-Paul	I.E.E.A.
M.	FAKIR Sabah	Mathématiques
M.	FONTAINE Hubert	Physique
M.	FOUQUART Yves	Physique
M.	FRONTIER Serge	Biologie
M.	GAMBLIN André	G.A.S.
M.	GLORIEUX Pierre	Physique
M.	GOBLOT Rémi	Mathématiques
M.	GOSSELIN Gabriel (dét.)	S.E.S.
M.	GOUDMAND Pierre	Chimie
M.	GREGORY Pierre	I.P.A.
M.	GREMY Jean-Paul	S.E.S.
M.	GREVET Patrice	S.E.S.
M.	GUILBAULT Pierre	Biologie
M.	HENRY Jean-Pierre	E.U.D.I.L.
M.	HERMAN Maurice	Physique
M.	JACOB Gérard	I.E.E.A.
M.	JACOB Pierre	Mathématiques
M.	JEAN Raymond	Biologie
M.	JOFFRE Patrick	I.P.A.

PROFESSEURS 2ème classe (suite 2)

M. JOURNEL Gérard	E.U.D.I.L.
M. KREMBEL Jean	Biologie
M. LANGRAND Claude	Mathématiques
M. LATTEUX Michel	I.E.E.A.
Mme LECLERCQ Ginette	Chimie
M. LEFEVRE Christian	Sciences de la Terre
Mle LEGRAND Denise	Mathématiques
Mle LEGRAND Solange	Mathématiques (Calais)
Mme LEHMANN Josiane	Mathématiques
M. LEMAIRE Jean	Physique
M. LHENAFF René	Géographie
M. LOCQUENEUX Robert	Physique
M. LOSFELD Joseph	C.U.E.E.P.
M. LOUAGE Francis(dét.)	E.U.D.I.L.
M. MACKE Bruno	Physique
M. MAIZIERES Christian	I.E.E.A.
M. MESSELYN Jean	Physique
M. MESSERLIN Patrick	S.E.S.
M. MONTEL Marc	Physique
Mme MOUNIER Yvonne	Biologie
M. PARSY Fernand	Mathématiques
Mle PAUPARDIN Colette	Biologie
M. PERROT Pierre	Chimie
M. PERTUZON Emile	Biologie
M. PONSOLLE Louis	Chimie
M. PORCHET Maurice	Biologie
M. POVY Lucien	E.U.D.I.L.
M. RACZY Ladislas	I.E.E.A.
M. RAOULT Jean François	Sciences de la Terre
M. RICHARD Alain	Biologie

PROFESSEURS 2ème Classe (suite 3)

M. RIETSCH François	E.U.D.I.L.
M. ROBINET Jean-Claude	E.U.D.I.L.
M. ROGALSKI Marc	Mathématiques
M. ROY Jean-Claude	Biologie
M. SCHAMPS Joël	Physique
Mme SCHWARZBACH Yvette	Mathématiques
M. SLIWA Henri	Chimie
M. SOMME Jean	G.A.S.
Mle SPIK Geneviève	Biologie
M. STAROSWIECKI Marcel	E.U.D.I.L.
M. STERBOUL François	E.U.D.I.L.
M. TAILLIEZ Roger	Institut Agricole
Mme TJOTTA Jacqueline (dét.)	Mathématiques
M. TOULOTTE Jean-Marc	I.E.E.A.
M. TURRELL Georges	Chimie
M. VANDORPE Bernard	E.U.D.I.L.
M. VAST Pierre	Chimie
M. VERBERT André	Biologie
M. VERNET Philippe	Biologie
M. WALLART Francis	Chimie
M. WARTEL Michel	Chimie
M. WATERLOT Michel	Sciences de la Terre
Mme ZINN JUSTIN Nicole	Mathématiques

CHARGES DE COURS

M. ADAM Michel	S.E.S.

CHARGES DE CONFERENCES

M.	BAFCOP Joël	I.P.A.
M.	DUVEAU Jacques	S.E.S.
M.	HOFLACK Jean	I.P.A.
M.	LATOUCHE Serge	S.E.S.
M.	MALAUSSENA DE PERNO Jean-Louis	S.E.S.
M.	NAVARRE Christian	I.P.A.
M.	OPIGEZ Philippe	S.E.S.

Il m'est agréable d'exprimer ma reconnaissance envers **Gérard COMYN** pour l'honneur et le plaisir qu'il me fait en présidant le Jury.

Je remercie **Max DAUCHET** et **Gérard JACOB** d'avoir accepté de participer au Jury .

Je suis sensible à l'honneur que me fait **Joffroy BEAUQUIER** en acceptant de venir à Lille faire partie du Jury .

Je ne saurai non plus oublier ce que je dois à **Michel LATTEUX** qui a su m'intéresser à la Théorie des Langages . Je le remercie vivement de m'avoir proposé ce travail de recherche que ses encouragements , ses conseils et la confiance qu'il m'a témoignée m'ont permis de mener à bien .

Mes remerciements vont particulièrement à toute la famille **Jean RANDRIANANTOANDRO** qui m'a toujours assuré de ses aides et de ses encouragements durant mes études .

Je remercie **Patricia CARON** qui , avec rapidité et patience s'est chargée de la mise en pages de cette thèse et **Henri GLANC** qui , avec gentillesse en a assuré la réalisation matérielle .

A mes parents ,

si loin .

TABLE DES MATIÈRES

INTRODUCTION

Le concept de langages algébriques (context-free languages) a été introduit par Chomsky [10] en 1959, dans le but de trouver un modèle mathématique pour décrire les langues naturelles. Plus précisément, il a défini d'abord les grammaires à structure de phrase (type 0) qui génèrent exactement les langages récursivement énumérables. En imposant des restrictions aux règles de production, il obtenait successivement les grammaires contextuelles (type 1) qui engendrent une partie des langages récursifs, les grammaires à contexte libre (type 2) correspondant aux langages algébriques et enfin les grammaires linéaires à droite (type 3) qui définissent la famille Rat des langages rationnels obtenue aussi à partir des langages finis par application des opérations classiques d'union, de produit et d'étoile.

Plus tard, en 1960, on a découvert que les langages "ALGOL-Like", définis par la forme normale de Backus (le métalangage utilisé pour décrire le langage de programmation ALGOL 60) ne sont autres que les langages algébriques.

En plus de leurs utilisations à des buts descriptifs, les grammaires algébriques ont beaucoup servi dans des domaines tels que l'analyse syntaxique, la compilation, la théorie des langages de programmation et la théorie des schémas de programme.

La famille des langages algébriques, notée Alg, a depuis fait l'objet de nombreuses études théoriques et pratiques, un des axes essentiels de ces recherches ayant été la définition de certaines de ses sous-classes. En effet, de nombreuses questions que l'on peut raisonnablement se poser n'ont reçu de réponses que pour ces dernières. Les plus connues ont été obtenues en ajoutant des restrictions simples aux grammaires algébriques (c'est le cas pour Lin, la famille des langages linéaires), ou aux automates à pile.

En 1965, Elgot et Mezei [13] ont défini un outil fondamental en théorie des langages, les transductions rationnelles. Nivat [38] en a donné une caractérisation en terme de bimorphismes alphabétiques qui sera utilisée tout au long de cette thèse. Ainsi est née la notion de cône rationnel, famille de langages fermée par transduction rationnelle, i.e. par homomorphisme, homomorphisme inverse et intersection avec un langage rationnel.

Simultanément, Ginsburg et Greibach [15] ont introduit un concept similaire, celui des familles agréables de langages ou FAL, cônes rationnels fermés par les opérations union, produit et étoile.

La substitution, opération très naturelle et très importante en théorie des langages, a fait l'objet de plusieurs études, en particulier dans [21], [3], [36], [37]. Le rapprochement entre cette dernière et la transduction rationnelle nous a donné une nouvelle opération, la substitution syntaxique.

En imposant certaines restrictions aux grammaires algébriques, on peut obtenir les grammaires algébriques "non expansives" qui engendrent la famille QR des langages quasirationnels, introduite par Yntema [41] en 1967. Il s'en suit que QR est une sous famille propre de Alg.

Nivat [39] a par la suite étudié cette famille et l'a caractérisée comme la clôture par substitution de la famille des langages linéaires. Ainsi QR est une famille agréable de langages. En plus, QR coïncide avec la famille des langages d'index fini [18].

Ces diverses caractérisations fournissent des moyens efficaces d'investigation qui manquent en général à l'étude des langages et expliquent le nombre de vocables différents - non expansive, finite index, superlinear - utilisés pour les désigner.

Cette famille joue avec la famille des langages à un compteur [18] un rôle fondamental dans la théorie des langages algébriques : problèmes d'ambiguité [17], théorèmes d'existence, contre-exemples à des conjectures, structure de la famille des langages algébriques etc ...

A partir de la caractérisation de Nivat, nous étudierons une nouvelle hiérarchie QR(k), ce qui nous amène à définir la notion d'ordre sur les langages quasirationnels. Nous pouvons alors calculer pour certaines opérations sur les langages, l'ordre du langage obtenu à partir des ordres des langages de départ. Ceci nous permet, en particulier de retrouver des résultats obtenus en utilisant les Edtol-systèmes ultralinéaires [30] et les séries formelles [27].

L'opération quotient est naturelle en théorie des langages, plus particulièrement pour les langages rationnels. En effet, il est bien connu que Rat, la famille des langages rationnels est fermée par quotient à droite (à gauche) par un langage arbitraire et que tout cône rationnel est clos par quotient rationnel.

En plus un langage L sur un alphabet X est rationnel si et seulement si l'ensemble $Q = \{u^{-1} L \ / \ u \in X^*\}$ des langages obtenus comme quotient de L par un mot $u \in X^*$ est un ensemble fini. Plus précisément c'est l'ensemble des états de l'automate fini déterministe minimal reconnaissant L.

En étendant l'opération quotient sur les langages non rationnels, celle-ci devient très puissante car elle regroupe les opérations intersection et effacement. En plus, il n'existe que peu d'exemples de cônes rationnels clos par quotient.

C'est ce qui nous a amené à étudier le comportement de certaines familles de langages algébriques vis à vis de cette opération.

Nous commencerons par la famille Rocl des langages à un compteur. Elle peut, en effet, être obtenue en considérant des automates dont l'alphabet de pile ne contient qu'un seul élément. Cette famille admet une seconde définition très simple : celle du cône rationnel engendré par D'^*_1, le langage de Dyck restreint défini sur une paire de parenthèses et dont une grammaire est donnée par les règles de production suivantes $S \to SS$, $S \to aSb$, $S \to \varepsilon$. L'intérêt de l'étude des langages à un compteur est renforcé par les rapports qui existent entre ces langages et les réseaux de Pétri et plus généralement les problèmes de synchronisation [28] [29]. Il n'est donc pas étonnant que de nombreux articles soient consacrés à la recherche des propriétés de cette famille (cf. [1] [4] [20] [23] [24] [25] [26] [34] [35] [41]).

Une autre famille de langages, celle des bornés (un langage L est borné s'il existe des mots $w_1, \ldots, w_k$ tels que L soit inclus dans w_1^* $w_2^* \ldots w_k^*$), occupe aussi une place importante en théorie des langages. En effet, certains problèmes indécidables dans le cas général deviennent décidables si l'on se restreint aux langages bornés. Cette famille, étudiée par Ginsburg et Spanier [14] présente l'intérêt que chacun des ses éléments n'est autre qu'un sous ensemble de $\mathbb{N}^k$. C'est ainsi que de nombreux auteurs se sont intéressés à l'étude du cône rationnel engendré par les langages algébriques bornés, noté $C(AB)$, entre autres Berstel et Boasson [6], Ginsburg [20], Ginsburg et Spanier [14] [16].

La dernière famille à laquelle nous nous intéresserons sera la famille des langages commutatifs (fermés par permutation), plus précisément AC celle des langages algébriques commutatifs. L'étude de ces langages est facilitée par le fait qu'on peut représenter un langage commutatif défini sur un alphabet X par une partie de $\mathbb{N}^k$ où k est le cardinal de X. Comme tout langage borné est l'image par transduction rationnelle d'un langage commutatif, nous essaierons d'étendre les résultats démontrés pour les algébriques bornés aux algébriques commutatifs. Remarquons enfin que AC est fermée par homomorphisme inverse.

Passons maintenant à une présentation un peu plus détaillée de notre travail.

Nous commencerons par rappeler les définitions et résultats classiques en théorie des langages utilisés tout au long de cette thèse.

Dans la première partie, nous définissons une nouvelle hiérarchie QR(k) de langages quasirationnels et nous étudions l'ordre d'un langage quasirationnel L qui est égal à k si L appartient à $QR(k) \setminus QR(k-1)$, à 1 si $L \in QR(1)$.

Nous montrons, qu'à condition d'avoir des alphabets disjoints, on peut calculer, par exemple l'ordre du produit $L_1 L_2$, de $(Ld)^*$, de $\langle LcL\rangle = \{a^n\ x\ c\ y\ b^n \ /\ n \geq 0,\ x,\ y \in L\}$ et de $\langle L\rangle = \{a^n\ x\ b^n \ /\ n \geq 0,\ x \in L\}$.

Le second paragraphe est réservé au calcul de l'ordre de la substitution marquée d'un langage quasirationnel L_2 dans un langage quasirationnel L_1, ces deux langages étant définis sur des alphabets disjoints.

Nous consacrons le paragraphe suivant à l'opération insertion. En effet, en définissant les langages quasirationnels L_1 et L_2 sur des alphabets disjoints, trois cas se présentent : le premier est que l'ordre de $i(L_1, L_2)$ est calculable (c'est le cas si L_1 ou L_2 est rationnel, si $O(L_2) = 2k_2 > k_1 = O(L_1)$ et si $O(L_2) = 2k_2+1 \geq k_1 = O(L_1) > 1$), le second est que l'ordre varie entre des bornes qui peuvent être atteintes et enfin le troisième qui reste encore un problème ouvert.

La permutation circulaire marquée, un cas particulier de conjugaison occupe le début du quatrième paragraphe que nous terminons par une proposition liant deux problèmes irrésolus pour ces deux dernières opérations.

Dans la seconde partie de cette thèse, nous étudions la fermeture par quotient de certaines familles de langages algébriques.

En effet, dans le premier chapitre, nous nous occupons de Rocl, la famille des langages à un compteur.

Le premier paragraphe est consacré à l'étude de la relation entre l'opération quotient et les cônes rationnels . Comme toute propriété vérifiée par le quotient au centre reste vraie pour les quotients à droite, à gauche, nous énonçons les résultats si possible avec le quotient au centre. Ainsi nous montrons que le quotient de deux cônes rationnels est un cône rationnel. Ensuite, nous examinons le cas où l'un de ces cônes est principal, i.e. engendré par un seul langage et nous obtenons quelques résultats utilisés dans le paragraphe suivant pour montrer que la famille des langages à un compteur est fermée par quotient à gauche par un langage algébrique. La démonstration dépend du fait que pour tout langage algébrique L défini sur l'alphabet $\{a, b\}$ il existe un langage rationnel $R \subseteq a^*$ tel que les quotients de D'^*_1 par L et par R sont égaux [Lemme II.15]. Comme conséquence, Rocl est fermée par quotient à gauche.

Dans le troisième paragraphe, nous montrons qu'il n'est pas facile d'étendre ce résultat en montrant que Ocl, la famille agréable de langages générée par Rocl, n'est pas close par quotient, que la famille des langages algébriques n'est pas fermée par quotient par un langage à un compteur et en présentant une démonstration très simple du fait que tout langage récursivement énumérable est le quotient de deux langages linéaires.

Nous abordons dans le second chapitre l'étude de la fermeture par quotient du cône engendré par les langages algébriques bornés et du cône engendré par les langages algébriques commutatifs.

Dans le premier paragraphe, nous définissons un ISA-langage comme vérifiant la propriété suivante : pour toute substitution algébrique s, l'image par s^{-1} de ce langage est algébrique. Nous étendons cette définition aux familles de langages. Ainsi nous montrons que pour tout ISA-langage, le cône rationnel qu'il engendre est une ISA-famille de langages. Nous établissons aussi que si $\mathcal{L}$ est une ISA-famille de langages, Alg $//$ $\mathcal{L} \subseteq$ Alg.

Ce qui nous permet de démontrer au paragraphe suivant que Alg est fermé par quotient au centre par un langage algébrique borné. Ainsi nous pouvons en déduire la fermeture par quotient de $\mathcal{C}$(AB).

Enfin, dans le dernier paragraphe, nous montrons que $\mathcal{C}$(AC) est aussi fermé par quotient. Puis nous démontrons que les langages algébriques commutatifs sont des ISA-langages. Nous en déduisons que la famille des langages algébriques est fermée par quotient par les langages algébriques commutatifs.

PRELIMINAIRES

A. NOTATIONS

B. LANGAGES RATIONNELS

C. TRANSDUCTIONS RATIONNELLES

D. GRAMMAIRES

E. SUBSTITUTION

A. NOTATIONS

Un *alphabet* X est un ensemble fini non vide d'éléments appelés *lettres*. Les éléments de X^*, le *monoïde libre engendré par* X, sont des *mots*. Toute partie (finie) de X^* est un *langage* (fini).

Pour tout mot u de X^*, $|u|$ est le nombre d'occurrences du mot u appelé *longueur de* u. Le *mot vide*, noté ε, est le seul mot de longueur nulle. On notera $|u|_x$ le nombre d'occurrences de la lettre x dans le mot u.

Pour tout mot u de X^*, on définit u^R le *mot miroir* de u par récurrence sur la longueur de u :

. $u^R = u$ si $|u| \leq 1$

. si $|u| > 1$ alors $u = u_1 u_2$ avec $|u_2| = 1$ et $u^R = u_2 u_1^R$.

Soient X^* et Y^* deux monoïdes libres, un *homomorphisme* h de X^* dans Y^* est une application vérifiant :

$h(w_1w_2) = h(w_1)\ h(w_2)$ pour tous $w_1, w_2 \in X^*$ et $h(\varepsilon) = \varepsilon$.

Un homomorphisme h de X^* dans Y^* sera :

. *strict* si et seulement si $h(X) \subset Y^+$

. *alphabétique* si et seulement si $h(X) \subset Y \cup \{\varepsilon\}$

. *strictement alphabétique* si et seulement si $h(X) \subset Y$.

Lemme

Si h est un homomorphisme alphabétique de X^* dans Y^* alors :

$$h^{-1}(w_1w_2) = h^{-1}(w_1)\ h^{-1}(w_2)$$
$$h^{-1}(w_1^+) = (h^{-1}(w_1))^+$$

pour tous $w_1, w_2 \in Y^*$.

La *fermeture commutative* d'un mot $u \in X^*$, notée com(u) est l'ensemble des mots obtenus à partir de u par permutation de lettres.

La *fermeture commutative* de $L \subseteq X^*$, notée com(L) est égal à $\bigcup_{u \in L} \mathrm{com}(u)$.

Un langage L est dit *commutatif* si L = com(L).

Soient L_1 et L_2 deux langages inclus dans X^*, le *shuffle* de L_1 par L_2, noté $L_1 \sqcup\!\sqcup L_2$ est

$$L_1 \sqcup\!\sqcup L_2 = \{x_1y_1x_2y_2\ldots x_ny_n / x_1x_2\ldots x_n \in L_1, y_1y_2\ldots y_n \in L_2, \forall i \in [1,n], x_i, y_i \in X^*\}$$

B. LANGAGES RATIONNELS

Définition

Un *automate d'état fini déterministe* A est un quintuple $(X, Q, q_o, *, F)$ où

. X est un alphabet d'entrée

. Q est un ensemble fini d'états

. q_o un élément distingué de Q appelé état initial

. $*$ une fonction partielle de $Q \times X \to Q$ que l'on étend à $Q \times X^*$ en posant pour tout mot f de X^* et pour toute lettre x de X

$$q * \varepsilon = q$$

$$q * fx = (q * f) * x$$

$*$ est appelée fonction de transition

. $F \subset Q$ est une partie de Q dont les éléments sont appelés états finaux.

Définition

Un mot w de X^* est reconnu par l'automate $A = (X, Q, q_o, *, F)$ si et seulement si $q_o * w \in F$.

Définition

Un langage L de X^* est *rationnel* si et seulement si il existe un automate A tel que l'ensemble des mots reconnus par A soit identique à L. On dira que A reconnait L et on note $T(A) = L$.

Définition

L'automate fini déterministe *minimal* qui reconnait L sera, parmi les automates finis déterministes qui reconnaissent L celui qui possède le moins d'états (il est unique à une numérotation des états près).

Pour tout monoïde M, la famille $Rat(M)$ de ses parties rationnelles est la fermeture rationnelle, i.e. par les opérations union, produit et étoile, de l'ensemble de ses parties finies.

Un langage est *rationnel* s'il est une partie rationnelle d'un monoïde libre.

C. TRANSDUCTIONS RATIONNELLES

Soient X^* et Y^* deux monoïdes libres. Nous appellerons *transduction* de X^* dans Y^* toute application de X^* dans l'ensemble $\mathcal{P}(Y^*)$ des parties de Y^*.

Si τ est une telle application, nous lui associons la partie $\hat{\tau}$ de $X^* \times Y^*$ définie par $\hat{\tau} = \{(f, g) \in X^* \times Y^* \; / \; g \in \tau f\}$. Ainsi la transduction τ^{-1} de Y^* dans X^* sera définie par $\forall\, g \in Y^* \; \tau^{-1} g = \{f \in X^* \; / \; (f, g) \in \hat{\tau}\}$.

Une transduction τ de X^* dans Y^* est dite *rationnelle* si et seulement si $\hat{\tau} \in \mathrm{Rat}\,(X^* \times Y^*)$.

Une synthèse des résultats concernant les transductions rationnelles et ses applications a été réalisée par Berstel [7].

Nous ne nous intéresserons par la suite qu'aux transductions rationnelles en utilisant constamment la caractérisation dite de Nivat :

Théorème [38]

Une transduction τ *de* X^* *dans* Y^* *est rationnelle si et seulement si il existe un alphabet* Z, *un langage rationnel* R *défini sur* Z, *deux homomorphismes (alphabétiques)* h *de* Z^* *dans* X^*, g *de* Z^* *dans* Y^* *tels que*

$$\hat{\tau} = \{(h(w), g(w)) \; / \; w \in R\}.$$

Ce qui nous donne en utilisant les notations d'Eilenberg [12]

$$\tau = g \circ (\cap R) \circ h^{-1}$$

Remarquons que l'on peut prendre Z tel que toutes ses lettres ne peuvent avoir des images par h et g vides simultanément.

Une transduction rationnelle τ de X^* dans Y^* sera dite

. *d'image finie* si $\forall\, w \in X^*, \tau(w)$ est un ensemble fini

. *fidèle* si $\forall\, w \in Y^* \, \tau^{-1}(w)$ est un ensemble fini

. *bifidèle* si τ est à la fois fidèle et d'image finie

. *décroissante* si $\forall\ w \in X^*,\ \forall\ w' \in \tau(w)\ |w'| \leq |w|$

Théorème d'Elgot et Mezei [13]

La composée de deux transductions rationnelles (resp. fidèle, d'image finie, bifidèle, décroissante) est une transduction rationnelle (resp. fidèle, d'image finie, bifidèle, décroissante).

Définition

Soit L une famille de langages. L est un *cône rationnel* (resp. décroissant, bifidèle) si et seulement si elle est fermée par transduction rationnelle (resp. décroissante, bifidèle). $C(L)$ sera le plus petit cône rationnel contenant L.

Définition

Un cône rationnel C sera dit *principal* s'il existe un langage L appelé *générateur* tel que $C(\{L\}) = C = C(L)$.

Un cône rationnel n'est en général pas fermé par les opérations union, produit et étoile, c'est la raison pour laquelle la notion de familles agréables de langages ou FAL a été introduite :

Définition

Une famille de langages L est une *FAL* si et seulement si L est un cône rationnel fermé par union, produit et étoile.

Lemme

Tout cône principal est fermé par union.

D. GRAMMAIRES ALGEBRIQUES

On appelle *grammaire algébrique* un quadruplet $G = (X, N, P, \sigma)$ où

. X est un alphabet dit *terminal*

. N est un alphabet disjoint de X dit *non terminal*

. P est un sous-ensemble fini de $N \times (X \cup N)^*$ dit ensemble de *règles* de la grammaire ou *production*

. $\sigma \in N$ est appelé *axiome*

On dit que g *dérive directement de* f *dans* G si et seulement si il existe $m_1, m_2 \in (X \cup N)^*$ et $\exists\ (v, m) \in P$ tels que $f = m_1\ v\ m_2$ et $g = m_1\ m\ m_2$ et on écrit $f \rightarrow g$.

La fermeture transitive et réflexive de cette relation est la relation dérive de (dans) définie par : g *dérive de* f *dans* G si et seulement si il existe une suite finie $f_1, f_2 \ldots f_{k+1}$ telle que $f_1 = f \ldots$ $f_{k+1} = g$ et $\forall\ i \in [1, k]\ f_i \rightarrow f_{i+1}$ et on note $f \overset{*}{\rightarrow} g$. k est appelé *longueur de la dérivation.*

On désignera par

$L(G, f) = \{g \in X^* \ / \ f \overset{*}{\rightarrow} g\}$

$\hat{L}(G, f) = \{g \in (X \cup N)^* \ / \ f \overset{*}{\rightarrow} g\}$.

Dans le cas où $f = \sigma$ l'ensemble $L(G, \sigma)$ est appelé *langage engendré par* G et $\hat{L}(G, \sigma)$ *langage élargi engendré par* G.

Une grammaire $G = (X, N, P, \sigma)$ est *linéaire* (respectivement linéaire à gauche, respectivement linéaire à droite) si et seulement si $P \subseteq N \times (X^* N X^* \cup X^*)$ (respectivement $P \subseteq N \times (N X^* \cup X^*)$, respectivement $P \subseteq N \times (X^* N \cup X^*)$).

Définition

Un langage L est *algébrique* (resp. *linéaire*, resp. *rationnel*) s'il est engendré par une grammaire algébrique (resp. linéaire, resp. linéaire à droite ou linéaire à gauche).

On notera respectivement *Alg*, *Lin*, *Rat* le plus petit cône engendré respectivement par les langages algébriques, les langages linéaires, les langages rationnels.

Définition

Une grammaire $G = (X, N, P, \sigma)$ est *libre* (resp. *libre à droite*, *libre à gauche*) si $\hat{L}(G, \sigma) \setminus X^*$ est inclus dans $N V^* \cup V^* N$ (resp $V^* N$, $N V^*$) où V est égal à $X \cup N$.

Remarquons que toute grammaire linéaire libre à droite (resp. libre à gauche) engendre un rationnel et réciproquement.

E. SUBSTITUTIONS

Définition

Une *substitution* est un homomorphisme de X^* dans $P(Y^*)$. Elle sera dite

- rationnelle si $s(x) \in \mathrm{Rat} \ \forall x \in X$
- algébrique si $s(x) \in \mathrm{Alg} \ \forall x \in X$

Définition

Soit s un homomorphisme de X^* dans $P(Y^*)$, soit L une famille de langages. Nous dirons que s est une *L-substitution* si et seulement si $s(x) \in L$ pour tout $x \in X$.

A cette opération sur les langages, on fait correspondre un opérateur binaire sur les familles de langages, défini de la manière suivante :

Si L_1 et L_2 sont deux familles de langages, $L_1 \square L_2 = \{s(L) / L \in L_1$, s une L_2-substitution$\}$.

Cet opérateur est fondamental en théorie des langages et se trouve à l'origine de la théorie des opérateurs syntaxiques [8] [21] [22] [11] [32] [39] [40]. Le résultat suivant est fondamental :

Théorème [18]

Si L_1 et L_2 sont des cônes (principaux) alors $L_1 \square L_2$ est un cône (principal). On a le même résultat pour les FAL.

La famille *Elm* est composée de tous les sous ensembles finis de X. La famille *Fin* est la famille de tous les langages finis.

Ainsi,

Soit L une famille de langages, Rat $\square$ L, Fin $\square$ L et Elm $\square$ L désignent respectivement la fermeture rationnelle, la fermeture par union et produit, la fermeture par union de L.

Si $\mathcal{L}$ est un cône rationnel, Rat $\square$ $\mathcal{L}$ est la FAL engendrée par $\mathcal{L}$, Fin $\square$ $\mathcal{L}$ et Elm $\square$ $\mathcal{L}$ sont des cônes.

<u>*Définition*</u>

Une famille de langages $\mathcal{L}$ est *fermée par recopie* si pour tout homomorphisme injectif strictement alphabétique h, tout langage L $\in$ $\mathcal{L}$, h(L) $\in$ $\mathcal{L}$.

<u>*Lemme*</u> [21]

Pour toutes familles de langages $\mathcal{L}_1$, $\mathcal{L}_2$, $\mathcal{L}_3$ fermées par recopie,

$$(\mathcal{L}_1 \square \mathcal{L}_2) \square \mathcal{L}_3 = \mathcal{L}_1 \square (\mathcal{L}_2 \square \mathcal{L}_3)$$

PREMIÈRE PARTIE

LES LANGAGES QUASIRATIONNELS

Introduits par Yntema [41] en 1967, les langages quasirationnels ont intéressé de nombreux auteurs, entre autres Ginsburg et Spanier [18], Crestin [11]. En particulier, Nivat a défini cette famille de langages comme étant la fermeture par substitution de la famille des langages linéaires i.e. $QR = (Lin\ \square)^* Lin$.

En posant $QR'_o = Rat$, $QR_1 = Lin$ et $QR'_k = Rat\ \square\ QR_k$, $QR_{k+1} = Lin\ \square\ QR_k$ pour tout $k \geq 1$, il est connu que $QR'_k \subsetneq QR_{k+1} \subsetneq QR'_{k+1}$ pour tout $k \in \mathbb{N}$ et par conséquent

$$QR = \bigcup_{k=1}^{\infty} QR_k = \bigcup_{k=o}^{\infty} QR'_k$$

Considérons les familles $QR(1) = Rat$ et pour tout entier $k > 0$ $QR(2k) = QR_k$, $QR(2k+1) = QR'_k$. Il est clair que $QR(k) \subsetneq QR(k+1)$ pour tout $k > 0$ et que $QR = \bigcup_{k=1}^{\infty} QR(k)$. Ainsi nous pouvons définir l'ordre d'un langage quasirationnel par :

Définition I.1

Un langage quasirationnel L *est d'ordre* k *si et seulement si* $L \in QR(k) \setminus QR(k-1)$ *pour tout* $k \geq 2$, *d'ordre 1 sinon.*

Rappelons que si l'on se donne une grammaire $G = (X, N, P, \sigma)$, l'index d'une dérivation $\sigma = w_o \to w_1 \to \ldots \to w_i \to \ldots \to w_n = w$ est le maximum des nombres de non terminaux figurant dans un des w_j, $j \in [0, n]$, celui d'un mot w le minimum des index des dérivations qui l'engendrent. Une grammaire $G = (X, N, P, \sigma)$ est d'index k ($k > 0$) si et seulement si tout mot w de $L(G, \sigma)$ est d'index au plus k. G est d'*index fini* s'il existe k tel qu'elle soit d'index k, sinon elle est d'index infini. L'index d'un langage L est le minimum des index des grammaires qui l'engendrent.

A partir de la définition des QR(2k), une caractérisation grammaticale de ces langages s'obtient facilement par :

<u>*Lemme I.2*</u> *[18]*

Soit L *un langage quasirationnel,* L *appartient à* QR(2k) *si et seulement si il est au plus d'index* k.

Remarquons que QR(2k) est un cône rationnel principal pour tout $k > 0$ et QR(2k+1) une FAL principale pour tout $k \geq 0$.

A. ORDRE D'UN LANGAGE QUASIRATIONNEL

Dans ce paragraphe, nous étudions le comportement de l'ordre vis à vis d'un certain nombre d'opérations sur les langages quasirationnels.
Commençons par le produit qui à tous langages L_1, L_2 fait correspondre $L_1 L_2 = \{xy \mid x \in L_1, y \in L_2\}$. Pour faciliter ce calcul, établissons une propriété générale.

Proposition I.3

Soient L_1 *et* L_2 *deux langages définis sur des alphabets disjoints et* $\mathcal{L}$ *un cône rationnel. Si* $L_1 L_2$ *appartient à* $Lin \,\square\, \mathcal{L}$ *alors l'un des deux langages* L_1 *ou* L_2 *appartient à* $Rat \,\square\, \mathcal{L}$.

Démonstration

Soient L_1 et L_2 deux langages définis respectivement sur X_1 et X_2 où $X_1 \cap X_2 = \emptyset$. Comme $L_1 L_2 \in \mathrm{Lin} \,\square\, \mathcal{L}$, il existe un langage linéaire L défini sur un alphabet A et une $\mathcal{L}$-substitution s de A dans X tels que $s(L) = L_1L_2$.

Pour $i = 1, 2$, posons $A_i = \{a \in A \mid s(a) \subseteq X_i^*\}$ et $\Delta = \{a \in X \mid s(a) \cap X_1^+ X_2^+ \neq \emptyset\}$. On peut supposer que tout $a \in A$ occure dans un mot de L i.e. que $A = A_1 \cup A_2 \cup \Delta$. De même on peut supposer sans nuire à la généralité de la démonstration que $\Delta = \emptyset$.

En effet, supposons Δ non vide. Posons $\Delta' = \{a' \mid a \in \Delta\}$ avec $\Delta' \cap A = \emptyset$ et soit h l'homomorphisme de A dans $A \cup \Delta'$ défini par $h(a) = a$ si $a \in A_1 \cup A_2$ et $h(a) = aa'$ si $a \in \Delta$. Lin étant un cône rationnel, $L' = h(L) \in \mathrm{Lin}$. Définissons une substitution s' de $A \cup \Delta'$ dans X par $s'(a) = s(a)$ si $a \in A_1 \cup A_2$, $s'(a) = g_1(s(a))$ si $a \in \Delta$ et $s'(a') = g_2(s(a))$ si $a' \in \Delta'$ où pour $i = 1, 2$ g_i est l'homomorphisme défini sur $X_1 \cup X_2$ par $g_i(a) = a$ si $a \in X_i$ et $g_i(a) = \varepsilon$ sinon. Comme $\mathcal{L}$ est un cône rationnel, s' est une L-substitution vérifiant $s'(a) \subseteq X_1^*$ si $a \in A_1 \cup \Delta$, $s'(a) \subseteq X_2^*$ si $a \in \Delta' \cup A_2$ et $s'(L') = s(L)$.

En effet, soit $w \in L$. Si $w \in A_1^* A_2^*$ alors $h(w) = w \in L'$, $s'(w) = s(w)$ par définition de s et $s(w) \subseteq s'(L')$ sinon il existe $w_1 \in A_1^*$, $w_2 \in A_2^*$, $a \in \Delta$ tels que $w = a_1\ a\ w_2$, $h(w) = w_1\ a\ a'\ w_2 \in L'$ et $s(w) = s(w_1\ a\ w_2) \subseteq s'(w_1\ a\ a'\ w_2) \subseteq s'(L')$.

Réciproquement, prenons un $z \in s'(L')$.

. soit $z \in s'(w_1\ w_2)$ où $w_1\ w_2 \in A_1^*\ A_2^*$ et $s'(w_1) = s(w_1)$, $s(w_2) = s(w_2)$ alors $w_1\ w_2 \in L$ et $z \in s(w_1\ w_2) \subseteq s(L)$.

. soit $z \in s'(w_1\ a\ a'\ w_2)$ où $w_1 \in A_1^*$, $w_2 \in A_2^*$, $a \in \Delta$ et $a' \in \Delta'$. Il est clair que $z = z_1\ z'_1\ z'_2\ z_2$ où $z_1 \in s'(w_1) = s(w_1)$, $z'_1 \in s'(a)$, $z'_2 \in s'(a')$ et $z_2 \in s'(w_2) = s(w_2)$. Comme $s'(a) = g_1(s(a))$, il existe $x_2 \in X_2^*$ tel que $z'_1\ x_2 \in s(a)$ et $z_1\ z'_1\ x_2\ z_2 \in s(w_1\ a\ w_2) \subseteq s(L) = L_1L_2$. Ce qui implique que $z_1\ z'_1 \in L_1$ du fait que $z_1\ z'_1 \in X_1^*$ et $x_2\ z_2 \in X_2^*$. De la même façon, $z'_2\ z_2 \in L_2$. Ainsi $z \in L_1\ L_2 = s(L)$.

Supposons $\Delta \neq \emptyset$, i.e. pour tout $a \in A$ soit $s(a) \subseteq X_1^*$, soit $s(a) \subseteq X_2^*$ ou encore $A = A_1 \cup A_2$.

Comme $L \in$ Lin, $L = L(G, \sigma)$ où $G = (X, N, P, \sigma)$ est une grammaire linéaire. Pour $i = 1, 2$ définissons les homomorphismes h_i de $A \cup N$ dans $A_i \cup N$ par $h_i(a) = a$ si $a \in A_i \cup N$, $h_i(a) = \varepsilon$ sinon et les grammaires $G_i = (A_i, N, P_i, \sigma)$ où $P_i = \{(\xi, h_i(w)) \ /\ (\xi, w) \in P\}$. Posons $H_1 = L(G_1, \sigma)$, $H_2 = L(G_2, \sigma)$, $L'_1 = L_d(G_1, \sigma)$, $L'_2 = L_g(G_2, \sigma)$ et $L' = L'_1\ H_2 \cup H_1\ L'_2$. Il est évident que $L' \in$ Lin. Montrons que $s(L) = s(L')$.

Soit $w \in s(L)$, il existe $\ell = \ell_1\ \ell_2 \in L$ tel que $w \in s(\ell)$, $\ell_1 \in A_1^*$, $\ell_2 \in A_2^*$.

. soit $\ell_1 = \varepsilon$ et est nécessairement élément de L'_1. Comme $\ell_2 \in H_2$, $\ell_1 \ell_2 \in L'_1 H_2$.

. soit $\ell_2 = \varepsilon$ et comme au cas précédent $\ell \in H_1 L'_2$

. soit $\ell_1 \neq \varepsilon$ et $\ell_2 \neq \varepsilon$. On pourra prendre la grammaire $G = (X, N, P, \sigma)$ vérifiant $P \subseteq N \times (X^* N X^* \cup \{\varepsilon\})$. Soit une dérivation $\sigma \overset{*}{\to} \ell''_1 \sigma_1 \ell'_2 \to \ell'_1 \sigma_2 \ell'_2 \overset{*}{\to} \ell_1 \ell_2$ où $(\ell'_1 = \ell_1$ et $\ell''_1 \neq \ell_1)$ ou $(\ell''_2 \neq \ell_2$ et $\ell'_2 = \ell_2)$. Alors $\ell_1 \in L'_1$ ou $\ell_2 \in L'_2$, par conséquent $w \in s(L'_1 H_2 \cup H_1 L'_2) = s(L')$.

L'inclusion inverse est évidente du fait que $s(H_1) = L_1$, $s(H_2) = L_2$, $L'_1 \subseteq H_1$ et $L'_2 \subseteq H_2$.

D'où $L_1 L_2 = s(L'_1) L_2 \cup L_1 s(L'_2)$ et soit $L_1 = s(L'_1)$ soit $L_2 = s(L'_2)$. Ainsi $L_1 \in Rat \,\square\, \mathcal{L}$ où $L_2 \in Rat \,\square\, \mathcal{L}$.

$\square$

Ce résultat a été déjà démontré par Jacob [26] dans le cadre des séries formelles et par Latteux [29] à partir des Edtol-systèmes ultralinéaires.

En prenant $\mathcal{L} = QR(2k) \,\square\, \mathcal{L}$, une extension du résultat précédent s'obtient facilement.

<u>*Corollaire I.4*</u>

Soient L_1 *et* L_2 *deux langages définis sur des alphabets disjoints et* $\mathcal{L}$ *un cône rationnel. Pour tout* $k \geq 1$, *si* $L_1 L_2$ *appartient à* $QR(2k) \,\square\, \mathcal{L}$ *alors l'un des deux langages* L_1, L_2 *appartient à* $QR(2k-1) \,\square\, \mathcal{L}$.

En particulier si $\mathcal{L} = Rat$, comme $QR(\ell) \,\square\, Rat = QR(\ell)$ pour tout $k > 0$, on en déduit :

<u>*Corollaire I.5*</u>

Pour tout $k \geq 1$, *tous langages* L_1, L_2 *définis sur des alphabets disjoints,* $L_1 L_2 \in QR(2k)$ *implique que soit* $L_1 \in QR(2k-1)$ *soit* $L_2 \in QR(2k-1)$.

A partir de ce dernier corollaire, nous pouvons déterminer l'ordre du produit de deux langages quasirationnels définis sur deux alphabets disjoints :

Proposition I.6

Soient L_1 *et* L_2 *deux langages quasirationnels définis sur des alphabets disjoints. L'ordre du produit* $L_1 L_2$, $O(L_1 L_2)$ *est égal à* $2k+1$ *si* $O(L_1) = (L_2) = 2k$, *est égal à* $\sup(O(L_1), O(L_2))$ *sinon.*

Démonstration

Posons $k' = \sup(O(L_1), O(L_2))$. Il est évident que $O(L_1 L_2) \geq k'$. Distinguons plusieurs cas :

* Si $k' = 2k-1$, comme $QR(2k-1)$ est fermé par produit, $L_1 L_2 \in QR(2k-1)$ et $O(L_1L_2) = k'$.

* Si $k' = 2k = O(L_1)$ et $O(L_2) < 2k$ i.e. $L_1 \in QR(2k)$ et $L_2 \in QR(2k-1)$. Le cas $k' = 2$ se résoud facilement à partir du fait que tout cône rationnel est fermé par produit rationnel. Supposons $k' > 2$, alors il existe deux $QR(2k-2)$ substitutions s_1 et s_2 définies sur des alphabets disjoints X et Y, deux langages $D_1 \in Lin$, $D_2 \in Rat$ tels que $D_1 \subseteq X^*$, $D_2 \subseteq Y^*$, $s_1(D_1) = L_1$ et $s_2(D_2) = L_2$. Il est clair que la $QR(2k-2)$-substitution s définie par $s/X = s_1$, $s/Y = s_2$ vérifie $s(D_1 D_2) = L_1 L_2$ et $L_1 L_2 \in QR(2k)$. Ainsi $O(L_1 L_2) = k'$.

* si $k'=2k=O(L_2)$ et $O(L_1) < 2k$, nous obtenons par un raisonnement analogue au précédent que $O(L_1 L_2) = k'$.

* si $k' = 2k = O(L_1) = O(L_2)$ alors L_1 et $L_2 \in QR(2k+1)$. Ainsi $L_1 L_2 \in QR(2k+1)$ et $2k \leq O(L_1 L_2) \leq 2k+1$. Montrons que $O(L_1 L_2) \neq 2k$. En effet, supposons que $L_1 L_2 \in QR(2k)$. D'après le corollaire I.5, soit $L_1 \in QR(2k-1)$, soit $L_2 \in QR(2k-1)$ ce qui est contraire à l'hypothèse. Par conséquent $O(L_1 L_2) = 2k+1$.

□

Les deux corollaires suivants s'en déduisent immédiatement :

Corollaire I.7

Soient L_1 *et* L_2 *deux langages de* QR(2k) *définis sur des alphabets disjoints. Alors* $L_1 L_2 \in$ QR(2k) *si et seulement si l'un des deux langages appartient à* QR(2k-1).

Corollaire I.8

Soit L *un langage défini sur un alphabet* X *et* $d \notin X$. *Alors pour tout* $k \in \mathbb{N}^*$, $(Ld)^* \in$ QR(2k) *si et seulement si* $(Ld)^* \in$ QR(2k-1).

Une des propriétés intéressantes vérifiées par l'opérateur "chevron", démontrée par Boasson et Nivat [9], a été énoncée par Boasson, Crestin et Nivat [8].

Proposition I.9

Si $\mathcal{L}$ *est un cône rationnel fermé par union et* L *un langage défini sur un alphabet ne contenant ni* a, *ni* b, $\langle L\rangle = \{a^n x b^n \ / \ x \in L\} \in \mathrm{Rat} \ \square \ \mathcal{L}$ *implique que* $L \in \mathcal{L}$.

Si nous définissons

$\langle L c L\rangle = \{a^n x c y b^n \ / \ x, y \in L \subseteq X^*, \ a, b, c \notin X, \ n \geq 0\}$ et

$\langle L\rangle \ = \{a^n x b^n \ / \ x \in L \subseteq X^*, \ a, b \notin X, \ n \geq 0\}$, il est clair que si L est un langage quasirationnel, $\langle L c L\rangle$ et $\langle L\rangle$ le sont aussi. Pour tout $k \geq 1$, les deux propriétés suivantes sont faciles à démontrer :

Lemme I.10

Pour tout $k \geq 1$, QR(2k) *est translatable i.e. pour tout* $L \in$ QR(2k), $\langle L\rangle \in$ QR(2k).

Démonstration

Comme $L \in$ QR(2k), $L \subseteq X^*$, il existe une grammaire $G = (X, N, P, \sigma)$ d'index k l'engendrant. Soit $G' = (X', N', P', \sigma')$ une nouvelle grammaire d'index k où $X' = X \cup \{a, b\}$, $N' = N \cup \{\sigma'\}$, $P' = P \cup \{(\sigma', a\,\sigma'\,b), (\sigma', \sigma)\}$. Il est évident que $L(\sigma', \sigma') = <L>$ et $<L> \in QR_k = QR(2k)$.

□

Lemme I.11

Pour tout langage rationnel L, *tout* $k \geq 1$, $<L> \in QR(2k+1)$ *si et seulement si* $<L> \in QR(2k)$.

Démonstration

En effet, QR_k étant clos par union, $<L> \in$ Rat □ $QR(2k)$ implique que $L \in QR(2k)$ [8] et $<L> \in QR(2k)$ [Lemme I.10]. La réciproque est évidente.

□

On en déduit immédiatement :

Proposition I.12

Pour tout langage quasirationnel $L \subseteq X^*$ *et* $a, b \notin X$, $O(<L>)$ *est égal à* $2([O(L) + 1] \text{ div } 2)$.

Si L est un langage quasirationnel, l'ordre de $<L\,c\,L>$ s'obtient par la proposition suivante :

Proposition I.13

Soient $L \subseteq X^*$ *un langage quasirationnel, et* a, b, c *trois nouvelles lettres. Alors* $O(<L\,c\,L>) = 2([O(L)+2] \text{ div } 2)$.

Démonstration

Distinguons les deux cas possibles :

* si $O(L) = 2k+1$ avec $k \geq 0$, il est évident que $L\,c\,L \in QR(2k+1) \subseteq QR(2k+2)$ ainsi que $<L\,c\,L>$. Supposons que $<L\,c\,L> \in QR(2k+1)$, d'après le lemme précédent, $<L\,c\,L> \in QR(2k)$ et $L\,c\,L \in QR(2k)$ ce qui est contraire à l'hypothèse.
* si $O(L) = 2k$, $O(L\,c\,L) = 2k+1$ et $O(<L\,c\,L>) = 2k+2$.

□

B. CALCUL DE L'ORDRE DE LA SUBSTITUTION MARQUEE

Soient $L_1 \subseteq X_1^*$ et $L_2 \subseteq X_2^*$, deux langages non vides définis sur des alphabets disjoints et τ, la substitution définie sur X_1 par $\tau(x) = x\, L_2$, pour tout $x \in X_1$. Alors $\tau(L_1)$, l'image de L_1 par la substitution τ, notée $s(L_1, L_2)$ est appelée la *substitution marquée de* L_2 *dans* L_1. Cette opération est très "représentative" de la substitution et permet, en particulier, de construire des générateurs pour le cône rationnel $\mathcal{L}_1 \square \mathcal{L}_2$ des images des langages de $\mathcal{L}_1$ dans une $\mathcal{L}_2$-substitution, $\mathcal{L}_1$ et $\mathcal{L}_2$ étant deux cônes rationnels principaux [20]. C'est aussi en utilisant cette opération que Greibach a montré que si $\mathcal{L}$ est un cône rationnel clos par union, $\mathcal{L}$ est clos par substitution si et seulement si $\mathcal{L} \square \mathcal{L}$ l'est aussi [22].

Les lemmes suivants sont connus sous le nom de lemme de substitution. La première version [22], due à Greibach est très utilisée et constitue un résultat de base :

Lemme I.14

Soient L_1 *et* L_2 *des langages définis sur des alphabets disjoints,* $\mathcal{L}_1$ *et* $\mathcal{L}_2$ *deux cônes rationnels fermés par union. Si* $s(L_1, L_2) \in \mathcal{L}_1 \square \mathcal{L}_2$ *alors* $L_1 \in \mathcal{L}_1$ *ou* $L_2 \in \mathcal{L}_2$.

Beauquier a montré que l'hypothèse de fermeture par union des cônes $\mathcal{L}_1$ et $\mathcal{L}_2$ n'était pas nécessaire [5]. C'est essentiellement la version la plus puissante due à Latteux [32] que nous utiliserons par la suite.

Lemme I.15

Soient $\mathcal{L}_1$ *un cône décroissant fidèle,* $\mathcal{L}_2$ *un cône décroissant,* L_1 *et* L_2 *des langages définis sur des alphabets disjoints. Alors* $s(L_1, L_2) \in \mathcal{L}_1 \square \mathcal{L}_2$ *implique* $L_1 \in \mathcal{L}_1 \square \mathrm{Fin}$ *ou* $(L_2\, c)^* \in \mathcal{L}_2$.

Nous allons montrer dans ce paragraphe, que pour deux langages quasirationnels L_1, L_2 définis sur des alphabets disjoints, l'ordre du langage $s(L_1, L_2)$ ne dépend que de $O(L_1)$ et $O(L_2)$ si $L_1 \notin$ Elm.

En effet, commençons par traiter le cas où L_2 est d'ordre pair :

<u>*Lemme I.16*</u>

Soient L_1 *et* L_2 *deux langages quasirationnels non rationnels définis sur deux alphabets disjoints d'ordres respectifs* k_1 *et* $2k_2$, $s(L_1, L_2)$ *est d'ordre* $k_1 + 2k_2$.

<u>Démonstration</u>

Soient $L_1 \subseteq X_1^*$, $L_2 \subseteq X_2^*$ où $X_1 \cap X_2 = \emptyset$. $QR(2k_2)$ étant un cône rationnel, $a\, L_2 \in QR(2k_2)$ pour $a \in X_1$. Si on définit τ par $\tau(a) = a\, L_2$ pour tout $a \in X_1$ alors τ est une $QR(2k_2)$-substitution et $s(L_1, L_2) \in QR(k_1) \square QR(2k_2) = QR(k_1 + 2k_2)$.

Supposons que $s(L_1, L_2) \in QR(k_1 + 2k_2-1) = QR(k_1-1) \square QR(2k_2)$ [$k_1 > 1$ et $\square$ associative]. D'après le lemme précédent $L_1 \in QR(k_1-1) \square \mathrm{Fin} = QR(k_1-1)$ ou $(L_2\, \#)^* \in QR(2k_2)$, $\# \notin (X_1 \cup X_2)$. Ainsi $L_1 \in QR(k_1-1)$ ou $L_2 \in QR(2k_2-1)$ [Corollaire I.8], ce qui est contraire à l'hypothèse.

D'où $s(L_1, L_2)$ est d'ordre k_1+2k_2.

$\square$

Par contre si L_1 est rationnel et $O(L_2) = 2k_2$, il est facile de voir que si $L_1 \in$ Elm alors $O(s(L_1, L_2)) = O(L_2)$ sinon $O(s(L_1, L_2)) = O(L_2)+1$.

Si L_2 est d'ordre impair, nous avons un résultat similaire :

<u>*Lemme I.17*</u>

Soient L_1 *et* L_2 *deux langages quasirationnels définis sur des alphabets disjoints d'ordres respectifs* k_1 *et* $2k_2+1$, $s(L_1,L_2)$ *est d'ordre* k_1+2k_2.

Démonstration

Soient $L_1 \subseteq X_1^*$, $L_2 \subseteq X_2^*$ où $X_1 \cap X_2 = \emptyset$. En définissant une substitution τ comme au lemme précédent, il est évident que $s(L_1, L_2) \in QR(k_1+2k_2+1)$. Comme $\square$ est associative et que pour tout $k > 0$, $QR(k) \square Rat = QR(k)$, $QR(k_1) \square QR(2k_2+1) = QR(k_1) \square (Rat \square QR(2k_2)) = (QR(k_1) \square Rat) \square QR(2k_2) = QR(k_1) \square QR(2k_2) = QR(k_1+2k_2)$ pour tout $k_1 > 0$, tout $k_2 \geq 0$ et $s(L_1, L_2) \in QR(k_1+2k_2)$.

Supposons que $s(L_1, L_2) \in QR(k_1+2k_2-1)$. Si $k_1 > 1$ alors $QR(k_1+2k_2-1) = QR(k_1-1) \square QR(2k_2)$ et $L_1 \in QR(k_1-1)$ ou $L_2 \in QR(2k_2)$ [5] ce qui est en contradiction avec les ordres des langages L_1 et L_2. Si $k = 1$, il est évident que $O(s(L_1, L_2)) = 2k_2+1 = k_1+2k_2$.

Par conséquent $O(s(L_1, L_2)) = k_1+2k_2$.

$\square$

Des lemmes I.16 et I.17, nous pouvons en déduire.

Proposition I.18

Soient L_1 *et* L_2 *deux langages quasirationnels définis sur des alphabets disjoints avec* $L_1 \notin Elm$. *Alors* $O(s(L_1, L_2)) = O(L_1) + 2[O(L_2)\ \underline{div}\ 2]$.

C. INSERTION

Soient deux langages L_1, L_2, l'*insertion de* L_2 *dans* L_1, notée $i(L_1, L_2)$ est égale à $\{x\ w\ y\ /\ xy \in L_1,\ w \in L_2\}$ [22].

En considérant des langages quasirationnels définis sur des alphabets disjoints, intéressons-nous d'abord au cas où le langage L_2 que l'on insère dans L_1 est d'ordre strictement supérieur à celui de ce dernier.

Supposons $O(L_2)$ pair, si L_1 est un langage rationnel, il est clair que $i(L_1, L_2)$ est du même ordre que L_2. Il en est de même si L_1 n'est pas régulier. Pour cela commençons par montrer la propriété suivante :

Lemme I.19

Soient L_1, L_2 *deux langages quasirationnels, définis sur des alphabets disjoints appartenant respectivement à* QR(2k-2), QR(2k) (k > 2). *Alors* $i(L_1, L_2) \in$ QR(2k).

Démonstration

Soient $L_1 \subseteq X_1^*$, $L_2 \subseteq X_2^*$ où $X_1 \cap X_2 \neq \emptyset$. Comme $L_1 \in$ QR(2k-2), il existe une grammaire $G_1 = (X_1, N_1, P_1, \sigma_1)$ d'index au plus k-1 engendrant L_1. On commence par dériver tous les non terminaux de G_1 en laissant trainer l'axiome $\sigma_2 \notin (N_1 \cup X_1)$ d'une grammaire $G_2 = (X_2, N_2, P_2, \sigma_2)$ d'index au plus k engendrant L_2, puis on termine par celle de σ_2. Il est évident que cette grammaire est au plus d'index k. D'où $i(L_1, L_2) \in$ QR(2k).

□

Ce résultat nous permet d'obtenir l'ordre de $i(L_1, L_2)$ si $O(L_2)$ est pair et $O(L_1) < O(L_2)$:

Lemme I.20

Soient L_1, L_2 *deux langages quasirationnels non rationnels, définis sur des alphabets disjoints, d'ordres respectifs* k_1 *et* $2k_2$ *où* $k_1 < 2k_2$. *Alors* $i(L_1, L_2)$ *est d'ordre* $2k_2$.

Démonstration

Soient $L_1 \subseteq X_1^*$, $L_2 \subseteq X_2^*$ où $X_1 \cap X_2 = \emptyset$. Comme $L_1 L_2 = i(L_1, L_2) \cap X_1^* X_2^*$, $O(i(L_1, L_2)) \geq O(L_1 L_2) = 2k_2$ [Proposition I.6].

Montrons que $i(L_1, L_2)$ est au plus d'ordre $2k_2$. L_1 appartenant à $QR(2k_2-1)$, $L_1 = s(R)$ où R est un rationnel défini sur un alphabet Z disjoint de $X_1 \cup X_2$, s une $QR(2k_2-2)$-substitution. Ainsi $i(L_1, L_2) = i(s(R), L_2)$.

Définissons

. pour tout $z \in Z$ le rationnel R_z obtenu à partir de R par la transduction rationnelle $\tau_z = \cap T_z \circ h_z^{-1}$ où $\bar{Z} = \{\bar{y}/y \in Z\}$, $\bar{Z} \cap Z = \emptyset$, $Z' = \{y'/y \in Z\}$, $Z' \cap (Z \cup \bar{Z}) = \emptyset$, $T_z = Z^* \bar{z} Z'^*$ et h_z l'homomorphisme alphabétique de $Z \cup \{\bar{z}\} \cup Z'$ dans Z défini par $h_z(y) = h_z(y') = y$ pour tout $y \in Z$, $y' \in Z'$ $h_z(\bar{z}) = z$.

. s' une $QR(2k_2)$-substitution par $s'(y) = s'(y') = s(y)$ pour tout $y \in Z$, $y' \in Z'$ et $s'(\bar{z}) = i(s(z), L_2)$ pour tout $\bar{z} \in \bar{Z}$ [Lemme I.19].

Il est facile de montrer que $i(L_1, L_2) = i[R \cap \{\varepsilon\}, L_2] \cup \bigcup_{z \in Z} s'(R_z)$ ce qui réduit le problème à l'appartenance de $s'(R_z)$ à $QR(2k_2)$ pour tout $z \in Z$.

Comme $R_z \subseteq Z^* \bar{z} Z'^*$, $R_z = \bigcup_{i=1}^{n} L_{z,i} L'_{z,i}$ où $L_{z,i}$, $L'_{z,i} \in \mathrm{Rat}$, $L_{z,i} \subseteq Z^*$, $L'_{z,i} \subseteq \bar{z} Z'^*$ pour tout $i \in [1, n]$ et $s'(R_z) = \bigcup_{i=1}^{n} s'(L_{z,i}.L'_{z,i})$, par conséquent il suffit de montrer que $s'(L_{z,i} L'_{z,i}) \in QR(2k_2)$.

Par définition, il existe des grammaires linéaires à droite $G_d = (Z, N_d, P_d, \sigma)$, à gauche $G_g = (Z' \cup \{\bar{z}\}, N_g, P_g, \sigma')$ vérifiant

$L_{z,i} = L(G_d, \sigma)$, $L'_{z,i} = L(G_g, \sigma')$, $P_d \subseteq N_d \times (Z\, N_d \cup \{\varepsilon\})$, $P_g \subseteq N_g \times (N_g\, Z' \cup \{\bar{z}\})$, $(Z \cup N_d) \cap (Z' \cup \{\bar{z}\} \cup N_g) = \emptyset$. Considérons la grammaire $G = (Z \cup Z' \cup \{\bar{z}\}, N_g \cup N_d, P, \sigma)$ où $P = P_g \cup \{(\xi, u\,\phi) \,/\, (\xi, u\,\phi) \in P_d\} \cup \{(\xi, \sigma')/(\xi, \varepsilon) \in P_d\}$, pour $t \in Z \cup Z'$, $G_t = (X_1, N_t, P_t, t)$ une grammaire d'index au plus k_2-1 engendrant $s'(t)$ et pour tout $\bar{y} \in \bar{Z}$ $G_{\bar{z}} = (X_1 \cup X_2, N_{\bar{z}}, P_{\bar{z}}, \bar{z})$ une grammaire d'index au plus k_2 engendrant $s'(\bar{z})$. Il est facile de montrer que $G_{z,i} = \bigcup_{t \in Z \cup Z' \cup \bar{Z}} G_t \circ G$ est d'index au plus k_2, par conséquent $s'(L_{z,i}\, L_{z',i}) = L(G_{z,i}, \sigma) \in QR(2k_2)$

D'où $i(L_1, L_2)$ est d'ordre $2k_2$

□

Examinons maintenant le cas où L_2 est d'ordre impair. Si L_1 ou L_2 est régulier, il est clair que $O(i(L_1, L_2)) = \sup(O(L_1), O(L_2))$. Par contre si L_1 et L_2 ne sont pas rationnels, rappelons un résultat dû à Greibach :

<u>*Proposition I.21*</u> [22]

Soient L_1, L_2 *deux langages définis sur des alphabets disjoints,* $\mathcal{L}$ *un cône rationnel. Alors* $i(L_1, L_2) \in \mathrm{Rat} \,\square\, \mathcal{L}$ *implique que soit* $L_1 \in \mathrm{Rat}$, *soit* $L_2 \in \mathcal{L}$.

Soient $L_1 \subseteq X_1^*$, $L_2 \subseteq X_2^*$ deux langages quasirationnels non rationnels avec $O(L_1) = k_1$, $O(L_2) = 2k_2+1$, $X_1 \cap X_2 = \emptyset$, $k_1 \leq 2k_2+1$. Comme $L_1\, L_2 = i(L_1, L_2) \cap X_1^*\, X_2^*$, $O(i(L_1, L_2)) \geq O(L_1\, L_2)$ qui est égal à $2k_2+1$ [Proposition I.6]. Ainsi :

<u>*Lemme I.22*</u>

Soient deux langages quasirationnels L_1, L_2 *non rationnels définis sur des alphabets disjoints d'ordres respectifs* k_1 , $2k_2+1$ *avec* $k_1 \leq 2k_2+1$. *Alors* $O(i(L_1, L_2)) = 2k_2+2$.

Démonstration

D'après le lemme I.19, $i(L_1, L_2) \in QR(2k_2+2)$. Supposons que $i(L_1, L_2) \in QR(2k_2+1) = Rat \,\square\, QR(2k_2)$ alors soit $L_1 \in Rat$ soit $L_2 \in QR(2k_2)$ ce qui n'est pas le cas. Par conséquent $i(L_1, L_2)$ est d'ordre $2k_2+2$.

$\square$

Par contre, pour les autres cas, on ne peut pas déterminer l'ordre de $i(L_1, L_2)$ en fonction uniquement de $O(L_1)$ et $O(L_2)$. En effet, si on prend $X = \{a, b, c, d, u, v\}$, $S_2 = \{a^n b^n \mid n \geq 0\}$, $L_2 = \{c^p u^s v^s d^p \mid p, s \geq 0\}$, il est facile de voir que $i(L_2, S_2)$ est d'ordre 4, ce qui nous permet de démontrer le résultat suivant :

Lemme I.23

Pour tout $k \geq 0$, *il existe des langages* L_{2k+4}, L'_{2k+4}, L''_{2k+4} *d'ordre* $2k+4$, S_2 *d'ordre 2 définis sur des alphabets disjoints tels que*

$$O[i(L_{2k+4}, S_2)] = 2k+4,$$
$$O[i(L'_{2k+4}, S_2)] = 2k+5,$$
$$O[i(L''_{2k+4}, S_2)] = 2k+6.$$

Démonstration

La construction des langages se fait par récurrence. Les langages employés seront tous définis sur des alphabets disjoints.

Soient $S_2 = \{c^n d^n \mid n \geq 0\}$, $\tilde{S}_2$ une recopie de S_2, $L_2 = \langle \tilde{S}_2 \rangle$, L'_2, L''_2 des recopies de L_2.

Posons pour tout $k \geq 0$,

$$L_{2k+4} = \langle L_{2k+2}\, \tilde{L}_{2k+2} \rangle$$
$$L'_{2k+4} = L''_{2k+2} . L_{2k+4}$$
$$L''_{2k+4} = \langle L'_{2k+4} \rangle$$

où $\tilde{L}_{2k+2}$ est une recopie de L_{2k+2}, L_{2k+4} et L''_{2k+2} sont définis sur des alphabets disjoints.

Il est clair que L_{2k+4}, L'_{2k+4}, L''_{2k+4} sont d'ordre 2k+4 pour tout $k \geq 0$.

La démonstration du lemme se fait par récurrence sur k

. si <u>$k = 0$</u>

- Pour L_4, il est facile de construire une grammaire d'index 2 engendrant $i(L_4, S_2)$, par conséquent $O[i(L_4, S_2)] = 4$.

- En prenant $L'_4 = L''_2 . L_4$, nous avons $i(L'_4, S_2) = i(L''_2, S_2).L_4 \cup L''_2.i(L_4,S_2)$, et $i(L'_4, S_2) \in QR(5)$. Si $L''_2 \subseteq X^*$, $S_2 \subseteq Y^*$, $L_4 \subseteq Z^*$ avec X, Y, Z disjoints deux à deux, alors $i(L'_4, S_2) \cap (X \cup Y)^* Z^* = i(L''_2, S_2)$. L_4 qui est d'ordre 5. D'où $i(L'_4, S_2)$ est d'ordre 5.

- L''_4 étant égal à $<L''_2\ L_4>$, il est clair que $i(L''_4, S_2) \in QR(6)$. En prenant les mêmes alphabets que précédemment $i(L''_4, S_2) \cap a^*(X \cup Z)^* Y^* b^* = <i(L''_2, S_2), L_4>$. D'où $i(L''_4, S_2)$ est d'ordre 6.

. Supposons la propriété vraie jusqu'à k-1 et montrons la pour k

- Pour L_{2k+4}, il suffit de construire une grammaire d'index k+2 engendrant $i(L_{2k+4}, S_2)$. Comme $O(i(L_{2k+4}, S_2)) \geq O(L_{2k+4}.S_2) = 2_{k+4}$, on obtient aisément l'ordre de $i(L_{2k+4}, S_2)$.

- Pour L'_{2k+4} et L''_{2k+4}, les démonstrations sont analogues à celles de L'_4 et L''_4 en tenant compte des hypothèses de récurrence.

□

En reprenant les mêmes langages, le résultat suivant s'obtient à partir du lemme I.21.

<u>Lemme I.24</u>

Pour tout $k \geq 0$, *tout langage* S *d'ordre* 2k+3, *il existe des langages* L_{2k+4}, L'_{2k+4}, L''_{2k+4} *d'ordre* 2k+4 *définis sur des alphabets disjoints de celui de* S *tels que* $i(L_{2k+4},S)$ *est d'ordre* 2k+4, $i(L'_{2k+4},S)$ 2k+5 *et* $i(L''_{2k+4}, S)$ 2k+6.

De ces deux derniers résultats, la généralisation suivante s'en suit facilement :

Proposition I.25

Pour tout $k \geq 0$ *et tout* $k' \in [2, 2k+3]$, *il existe des langages* L_1, L_2, L_3 *d'ordre* $2k+4$, *un langage* S *d'ordre* k' *défini sur un alphabet disjoint de ceux de* L_1, L_2, L_3 *vérifiant*

$$O(i(L_1, S)) = 2k+4$$

$$O(i(L_2, S)) = 2k+5$$

$$O(i(L_3, S)) = 2k+6.$$

Nous avons un résultat similaire si le langage où l'on insère est d'ordre impair :

Proposition I.26

Pour tout $k \geq 1$, *il existe des langages* L_1, L_2 *d'ordre* $2k+3$, *un langage* S *non rationnel d'ordre* $p \leq 2k+1$ *tels que* $i(L_1, S)$ *est d'ordre* $2k+3$ *et* $i(L_2, S)$ *d'ordre* $2k+4$.

Démonstration

Pour tout $k \geq 1$, les langages L_{2k+3}, L'_{2k+3} vont dépendre des langages L_{2k+2} et L''_{2k+2} définis précédemment.

En effet, $L_{2k+3} = L_{2k+2}.\tilde{L}_{2k+2}$

où . $\tilde{L}_{2k+2}$ est une recopie de L_{2k+2}

. L_{2k+2}, $\tilde{L}_{2k+2}$, L''_{2k+2} et S sont définis sur des alphabets disjoints.

La démonstration est identique à celle du lemme I.22.

□

Mais le problème reste ouvert si L_1 est d'ordre $2k$ ou $2k+1$ et L_2 d'ordre $2k$. En effet, il est facile de vérifier que $2k+1 \leq O[i(L_1, l_2)] \leq 2k+2$ mais on ne peut pas affirmer que la borne inférieure peut être atteinte on non. Il semble que :

Conjecture 1

Soient deux langages quasirationnels L_1 *et* L_2 *définis sur deux alphabets disjoints, d'ordres respectifs* $2k+1$ *et* $2k$ $(k > 0)$. *Alors* $i(L_1, L_2)$ *est d'ordre* $2k+2$.

Ce qui nous permettrait d'avoir.

Conjecture 2

Soient deux langages quasirationnels L_1 *et* L_2, *d'ordre* $2k$ $(k > 0)$, *définis sur deux alphabets disjoints* X_1 *et* X_2. *Alors* $i(L_1, L_2)$ *est d'ordre* $2k+2$.

En effet, soit L'_1 une recopie de L_1 sur un nouvel alphabet X'_1. Considérons le langage quasirationnel $P_1 = L_1.L'_1$ d'ordre $2k+1$. Il est facile de montrer que

$$i(P_1, L_2) = i(L_1, L_2).L'_1 \cup L_1.i(L'_1, L_2)$$

qui est d'ordre $2k+2$ d'après la première conjecture. D'où soit $i(L_1, L_2)$ L'_1, soit $L_1.i(L'_1, L_2)$ est d'ordre $2k+2$ par conséquent soit $i(L_1, L_2)$ soit $i(L'_1, L_2)$ l'est aussi.

$\square$

Ainsi nous pouvons résumer les résultats de ce paragraphe dans le tableau suivant en prenant soin de définir les langages L_1 et L_2 sur des alphabets disjoints.

$O(L_1)$	$O(L_2)$	Conditions	$O[i(L_1, L_2)]$
k_1	k_2	$k_1 = 1$ ou $k_2 = 1$	$\sup(k_1, k_2)$
k_1	$2k_2$	$1 < k_1 < 2k_2$	$2k_2$
k_1	$2k_2+1$	$1 < k_1 \leq 2k_2+1$	$2k_2+2$
$2k_1$	k_2	$2k_1 > k_2 > 1$	$2k_1$ $2k_1+1$ $2k_1+2$
$2k_1+1$	k_2	$2k_1 > k_2 > 1$	$2k_1+1$ $2k_1+2$
$2k$ $2k+1$	$2k$		$2k+1$? $2k_2$

D. PERMUTATION CIRCULAIRE MARQUEE

Définition I.27

Deux mots f *et* g *sont dits conjugués s'il existe deux mots* u *et* v *tels que* $f = uv$ *et* $g = vu$. *Etant donné un langage* L *défini sur l'alphabet* X, *on note* $conj(L)$ *le langage*

$$conj(L) = \{f \ / \ \exists\ g,\ g \in L \quad f,\ g \text{ sont conjugués}\}.$$

Un langage L *est fermé par conjugaison si* $conj(L) = L$. *Par extension, on dira qu'une famille de langages* $\mathcal{L}$ *est fermée par conjugaison si* $L \in \mathcal{L}$ *implique* $conj(L) \in \mathcal{L}$.

Il est facile de voir que Rat est fermé par conjugaison et que si L_1 et L_2 sont deux langages définis sur l'alphabet X, $\# \notin X$, $conj(L_1 \# \cap L_2\#) = conj(L_1\ \#) \cap conj(L_2\ \#)$. Nous allons montrer que :

Lemme I.28

Soient h *un homomorphisme alphabétique de* X *dans* Y, L_1 *un langage défini sur* X, L_2 *un langage défini sur* Y, $\# \notin X$, $\delta \notin Y$. *Alors il existe un homomorphisme alphabétique* g *de* $X \cup \{\#\}$ *dans* $Y \cup \{\delta\}$ *tel que*

$$conj(h(L_1)\delta) = g(conj(L_1\ \#))$$
$$conj(h^{-1}(L_2)\#) = g^{-1}(conj(L_2\ \delta)).$$

Démonstration

Considérons l'homomorphisme alphabétique g de $X \cup \{\#\}$ dans $Y \cup \{\delta\}$ défini par $g(x) = h(x)$ si $x \in X$ et $g(\#) = \delta$.

Soit $w \in conj(h(L_1)\delta)$, $w = v\ \delta\ u$ où $u\ v \in h(L_1)$. Par conséquent il existe $x_1\ x_2 \in L_1$ tel que $uv = h(x_1\ x_2)$ où $u = h(x_1)$, $v = h(x_2)$. Ainsi $w = h(x_2)\ \delta\ h(x_1) = g(x_2)\ g(\#)\ g(x_1) = g(x_2\ \#\ x_1)$ avec $x_2\ \#\ x_1 \in conj(L_1\ \#)$.

Réciproquement, soit $w \in g(conj(L_1\ \#))$. Il existe $x_2\ \#\ x_1 \in conj(L_1\ \#)$ tel que $w = g(x_2\ \#\ x_1)$. g étant un homomorphisme, $w = g(x_2)\ g(\#)\ g(x_1) =$

$h(x_2)\ \delta\ h(x_1)$ où $h(x_1\ x_2) \in h(L_1)$. D'où $w \in conj(h(L_1)\ \delta)$ et $conj(h(L_1)\delta) = g(conj(L_1\ \#))$.

De même on démontre de façon analogue que $conj(h^{-1}(L_2)\#) = g^{-1}(conj(L_2\ \delta))$.

□

Ces résultats nous permettent d'obtenir le

<u>*Lemme I.29*</u>

Soient τ *une transduction rationnelle de* X *dans* Y, L *un langage défini sur l'alphabet* X, $\# \notin X$, $\delta \notin Y$. *Il existe une transduction rationnelle* τ' *de* $X \cup \{\#\}$ *dans* $Y \cup \{\delta\}$ *telle que*

$$conj(\tau(L)\delta) = \tau'(conj(L\ \#)).$$

<u>Démonstration</u>

τ étant une transduction rationnelle peut se décomposer en $\tau = g \circ \cap R \circ h^{-1}$ où g et h sont des homomorphismes alphabétiques, R un rationnel défini sur Z et $\xi \notin Z$.

$$\begin{aligned} conj(\tau(L)\delta) &= conj(g(h^{-1}(L) \cap R)\delta) \\ &= g_1(conj[(h^{-1}(L) \cap R)\xi]) \\ &= g_1(conj[h^{-1}(L)\xi] \cap conj[R\ \xi]) \\ &= g_1[h_1^{-1}(conj(L\ \#)) \cap conj(R\ \xi)] \end{aligned}$$

où g_1 et h_1 sont des homomorphismes alphabétiques et $conj(R\ \xi) \in Rat$. D'où $\tau' = g_1 \circ \cap\ conj(R\ \xi) \circ h_1^{-1}$.

□

Soient $Sym = \{w\ \bar{w}^R : w \in \{a, b\}^*\}$, $\delta \notin \{a, b, \bar{a}, \bar{b}\}$. Comme il est facile de montrer que $conj(Sym\ \delta) \in QR(3)$, nous pouvons déduire du lemme précédent que :

<u>*Corollaire I.30*</u>

Pour tout langage linéaire L *défini sur* X, $\# \notin X$, $conj(L\ \#) \in Rat \square Lin$.

Démonstration

Comme $L \in \mathrm{Lin}$, il existe une transduction rationnelle τ telle que $L = \tau(\mathrm{Sym})$. D'après le lemme précédent, $\mathrm{conj}[\tau(\mathrm{Sym})\#] = \tau'(\mathrm{conj}(\mathrm{Sym}\ \delta))$ où τ' est une transduction rationnelle. Par conséquent $\mathrm{conj}(L\ \#) \in \mathrm{Rat} \ \square\ \mathrm{Lin}$.

$\square$

Rappelons la définition de l'opération insertion qui est une substitution particulière.

Définition I.31

Une substitution Φ *est une* ***insertion*** *sur* $L \subseteq X^*$ *s'il existe un sous ensemble* $T \subseteq X$, *tel que* $L \subseteq (X \setminus T)^* (T \cup \{\varepsilon\})(X \setminus T)^*$ *et* $\Phi(a) = \{a\}$ *pour tout* $a \in X \setminus T$. *Si pour tout* $a \in T$, $\phi(a) \in \mathcal{L}$, *on dit que* ϕ *est une* $\mathcal{L}$-***insertion***.

Définition I.32

Soient $\mathcal{L}_1$, $\mathcal{L}_2$ *deux familles de langages,* $\mathcal{L}_1 \sim \mathcal{L}_2 = \{\Phi(L) \ /\ L \in \mathcal{L}_1,\ \Phi$ *une* $\mathcal{L}_2$*-insertion sur* $L\}$.

Lemme I.33 [21]

Si L_1 *et* L_2 *sont deux langages définis sur des alphabets disjoints, alors*

$$C(L_1) \sim C(L_2) = C(i(L_1, L_2)).$$

Comme toutes les démonstrations faites jusqu'à présent sont facilitées par l'introduction du marqueur de fin de mot, il est naturel de définir une nouvelle opération, notée cir par $\mathrm{cir}(L) = \mathrm{conj}(L\ \#)$ où L est un langage défini sur un alphabet X, $\# \notin X$. Ce qui nous ramène à étudier $\mathrm{conj}(L\ \#)$ pour pouvoir énoncer des résultats sur $\mathrm{cir}(L)$. En effet étant donné un langage L défini sur un alphabet X, pour tout $\delta \notin X$, $\mathrm{conj}(L\ \delta)$ est isomorphe à $\mathrm{cir}(L)$.

Remarquons que pour tout langage quasirationnel L, cir(L) l'est aussi. Pour déterminer une borne supérieure de l'ordre de cir(L) en fonction de celui de L, essayons de trouver une expression simplifiée de cir(s(L)) où s est une L_2-substitution et L_2 un cône rationnel ou plus précisément de conj(s(L)#).

<u>*Lemme I.34*</u>

Soient $L \subseteq X_1^+$, $\delta \notin X_1$, L_2 *un cône rationnel, s une* L_2*-substitution de* X_1^* *dans* X^*, $\# \notin X$. *Alors il existe des* L_3*-insertions* $\phi_x (x \in X_1)$ *telles que* $L_1 = \mathrm{conj}(s(L)\#) = \bigcup_{x \in X_1} \phi_x (\mathrm{conj}(s(x)\#))$

où $L_3 = C(\mathrm{conj}(L\ \delta)) \,\square\, L_2$.

<u>Démonstration</u>

Pour tout $x \in X_1$, considérons les transductions rationnelles τ_x de X^* dans X^* définies par $\tau_x(w) = x^{-1} w$ pour tout $w \in X^*$ et les insertions Φ_x définies par

$$\Phi_x(y) = y \text{ si } y \neq \# \text{ et}$$

$$\phi_x(\#) = s'(\tau_x(\mathrm{conj}(L\ \delta)))$$

où $s'(x) = s(x)$ pour tout $x \in X_1$

$$s'(\delta) = \#$$

Ainsi $\mathrm{conj}(s(L)\#) = \bigcup_{x \in X_1} \phi_x(\mathrm{conj}(s(x)\#))$

En effet, soit $w \in \mathrm{conj}(s(L)\#)$, $w = v \# u$ où $u\, v \in s(L)$. Il existe $t = t_1 \dots t_i \in L$ tel que $t_j \in X_1$ pour tout $j \in [1, i]$, $u\, v \in s(t)$ et $u = u_1 u_2$, $v = v_2 v_1$ où $u_2 v_2 \in s(t_n)$, $u_1 \in s(t_1 \dots t_{n-1})$, $v_1 \in s(t_{n+1} \dots t_i)$. Par conséquent $w \in v_2\, s(t_{n+1} \dots t_i) \# s(t_1 \dots t_{n-1})\, u_2$, $v_2 \# u_2 \in \mathrm{conj}(s(t_n)\#)$ et $v_2 v_1 \# u_1 u_2 \in \Phi(v_2 \# u_2)$ et $w \in \Phi_{t_n}(\mathrm{conj}(s(t_n)\#))$ pour un $t_n \in X_1$.

Réciproquement, si $w \in \Phi_x(\mathrm{conj}[s(x)\#])$ pour un $x \in X_1$, il existe $\gamma \in \mathrm{conj}(s(x)\#)$ tel que $w \in \Phi_x(\gamma)$, $\gamma = \gamma_2 \# \gamma_1$ où $\gamma_1 \gamma_2 \in s(x)$. D'où $w \in \Phi_x(\gamma_2)\, \Phi_x(\gamma)\, \Phi_x(\gamma_1) = \gamma_2\, s'(\tau_x(\mathrm{conj}(L\ \delta)))\, \gamma_1$ i.e. que $w \in \gamma_2\, s'(z)\, \gamma_1$

pour un certain $z \in \tau_x(\mathrm{conj}(L\ \delta))$ autrement dit $z = z_2\ \delta\ z_1$ où $z_1\ x\ z_2 \in L$. Ainsi $w \in \gamma_2\ s(z_2)\ \#\ s(z_1)\ \gamma_1$ avec $s(z_1)\ \gamma_1\ \gamma_2\ s(z_2) \in s(z_1\ x\ z_2) \subseteq s(L)$ et $w \in \mathrm{conj}(s(L)\#)$.

□

Ainsi, nous avons le

Lemme I.35

Pour tout langage quasirationnel L *d'ordre* $2k(k > 0)$ *ou* $2k+1(k \geq 0)$, $\mathrm{cir}(L) \in QR(2k+2)$.

Démonstration

On peut supposer $L \subseteq X^+$ avec $\# \notin X$. La démonstration se fait par récurrence sur l'ordre de L. Considérons les deux cas :

* $O(L) = 2k+1$ avec $k \geq 0$

- si $k = 0$, $\mathrm{conj}(L\ \#) \in \mathrm{Rat} \subseteq \mathrm{Lin} = QR(2)$

- supposons la propriété vraie jusqu'à l'ordre $2k-1$ $(k > 0)$. Soit $L \in QR(2k+1)$, $L = s(R)$ où s est une Lin-substitution, $R \subseteq Z^*$, $\delta \notin Z$, $R \in QR(2k+1)$. D'après le lemme précédent, il existe des L_3-insertions Φ_z telles que $\mathrm{conj}(L\#) = \mathrm{conj}(s(R)\#) = \bigcup_{z\in Z} \Phi_z(\mathrm{conj}(s(z)\#))$. Comme $s(z) \in QR(2)$, $\mathrm{conj}(s(z)\#) \in QR(3)$ [Corollaire I.31]. Par hypothèse de récurrence $\mathrm{conj}(R\ \delta) \in QR(2k)$, τ_x étant une transduction rationnelle, $\tau_x(\mathrm{conj}(R\ \delta)) \in QR(2k)$ et s' étant une Lin-substitution $s'(\tau_x(\mathrm{conj}(R\ \delta))) \in QR(2k+2)$ et $\Phi_z(\mathrm{conj}(s(z)\#)) \in QR(3) \sim QR(2k+2) \subseteq QR(2k+2)$ [Lemme I.34, I.20] ainsi que $\mathrm{conj}(L\ \#)$.

* $O(L) = 2k$ $k > 0$

Il est clair que $L \in QR(2k+1)$, d'où $\mathrm{conj}(L\#) \in QR(2k+2)$.

□

Si L est d'ordre $2k(k > 0)$, nous pouvons donc affirmer que $2k \leq O(\mathrm{cir}(L)) \leq 2k+2$. Montrons que pour tout $k > 1$, ces bornes peuvent être atteintes :

Lemme I.36

Pour tout $k > 1$, *il existe des langages* L, L', L'' *d'ordre* $2k$ *tels que*

$O(cir(L)) = 2k,$

$O(cir(L')) = 2k+1,$

$O(cir(L'')) = 2k+2.$

Démonstration

Il suffit de considérer les langages suivants définis par récurrence

$$L_{2k+2} = \langle L_{2k}\ \tilde{L}_{2k} \rangle$$
$$L'_{2k+2} = \langle L_{2k}\ \tilde{L}_{2k} \rangle\ \tilde{\tilde{L}}_{2k} . \bar{L}_{2k}$$
$$L''_{2k+2} = s(\{a^n\ u\ x^p y^p\ v\ b^n\ /\ n,\ p \geq 0\})$$

où $s(a) = \tilde{L}_{2k}$ $\quad s(u) = L_{2k}$ $\quad s(x) = \bar{L}_{2k}$

$s(b) = b_2$ $\quad s(v) = b$ $\quad s(y) = b_1$

$L_2 = \{c^n\ d^n\ /\ n \geq 0\}$,

$\bar{L}_{2k}$, $\tilde{L}_{2k}$, $\tilde{\tilde{L}}_{2k}$ sont des recopies de L_{2k} sur des alphabets disjoints.

Ainsi, il est facile de montrer que $O[cir(L_{2+2})] = 2k+2$, $O[cir(L'_{2k+2})] = 2k+3$ et $O[cir(L''_{2k+2})] = 2k+4$.

De même, en construisant pour tout $k > 0$ les langages suivants

$$L_{2k+1} = L_{2k}\ \tilde{L}_{2k}, L'_{2k+1} = L_{2k+1}\ S_2$$

où S_2 est une recopie de L_2 sur un alphabet disjoint de celui de L_{2k+1}, $\tilde{L}_{2k}$ une recopie de L_{2k} sur un alphabet disjoint de celui de L_{2k}, # une nouvelle lettre, nous obtenons $O[conj(L_{2k+1}\ \#)] = 2k+1$ et $O[conj(L'_{2k+1}\ \#)] = 2k+2$ que nous pouvons énoncer comme suit

Lemme I.37

Pour tout $k > 0$, *il existe des langages quasirationnels* L *et* L' *d'ordre* $2k+1$ *tels que* $cir(L)$ *est d'ordre* $2k+1$ *et* $cir(L')$ *d'ordre* $2k+2$.

Mais le problème reste ouvert si $L \in$ Lin, pourtant il semble que $cir(L) \in$ Lin implique $L \in$ Rat. Ce qui est équivalent à

<u>*Conjecture 3*</u>

Tout langage linéaire fermé par conjugaison est rationnel.

Notons la dépendance de ce résultat avec la conjecture suivante :

<u>*Conjecture 4*</u>

Pour tous langages linéaires L_1 *et* L_2 *définis sur des alphabets disjoints,* $i(L_1, L_2) \in$ Rat □ Lin *implique* L_1 *ou* $L_2 \in$ Rat.

En effet, nous avons

<u>*Proposition I.38*</u>

Pour tous langages linéaires L_1 *et* L_2 *définis sur des alphabets disjoints* X_1 *et* X_2, *avec* $\# \notin (X_1 \cup X_2)$, $cir(L_1)$ appartient à Lin implique $i(L_1, L_2)$ $\in$ Rat □ Lin.

<u>Démonstration</u>

Soit δ une nouvelle lettre. Il est facile de montrer que

$i(L_1, L_2)\, \# = (h[conj(conj(L_1\ \#)\ L_2\delta]) \cap (X_1 \cup X_2)^*\{\#\})$ où h est l'homomorphisme alphabétique défini sur $X_1 \cup X_2 \cup \{\#, \delta\}$ par $h(\delta) = \varepsilon$ et $h(x) = x$ pour tout $x \neq \delta$. Comme $cir(L_1) \in$ Lin, $conj(L_1\ \#) \in$ Lin. L_2 appartenant à Lin, $conj(conj(L_1\ \#)L_1\ \delta) \in$ Rat □ Lin ainsi que $i(L_1, L_2)$

□

SECONDE PARTIE

QUOTIENT ET LES FAMILLES DE LANGAGES ALGEBRIQUES

I. LES LANGAGES A UN COMPTEUR

II. LES LANGAGES ALGEBRIQUES BORNES ET
LES LANGAGES ALGEBRIQUES COMMUTATIFS

CHAPITRE I

LES LANGAGES A UN COMPTEUR

A. QUOTIENT ET TRANSDUCTIONS RATIONNELLES

Le quotient (à gauche) d'un langage L_1 par un langage L_2, noté $L_2^{-1}\ L_1$ est égal à $\{v\ /\ uv \in L_1$ pour un certain u dans $L_2\}$ et le quotient (à gauche d'une famille de langages $\mathcal{L}_1$ par une autre $\mathcal{L}_2$ est la famille notée $\mathcal{L}_2^{-1}\ \mathcal{L}_1$, définie par $\{L_2^{-1}\ L_1\ /\ L_2 \in \mathcal{L}_2,\ L_1 \in \mathcal{L}_1\}$. $\mathcal{L}_1$ est clos par quotient par $\mathcal{L}_2$ si $\mathcal{L}_2^{-1}\ \mathcal{L}_1 \subset \mathcal{L}_1$ et $\mathcal{L}_1$ est clos par quotient si $\mathcal{L}_1^{-1}\ \mathcal{L}_1 \subseteq \mathcal{L}_1$.

Le quotient à gauche s'obtient de façon symétrique.

Si l'on définit un troisième type d'opération quotient par $L_1\ //\ L_2 = \{uw\ /\ uvw \in L_1$ pour un certain $v \in L_2\}$ où L_1 et L_2 sont des langages et $\mathcal{L}_1\ //\ \mathcal{L}_2 = \{L_1\ //\ L_2\ /\ L_1 \in \mathcal{L}_1,\ L_2 \in \mathcal{L}_2\}$, il est facile de voir que les propriétés vérifiées par //restent vraies pour les quotients à gauche, à droite.

Dans ce paragraphe, nous allons d'abord montrer que si $\mathcal{L}$ et $\mathcal{L}'$ sont des cônes rationnels alors $\mathcal{L}'\ //\ \mathcal{L}$ l'est aussi. Pour cela, nous montrerons que pour $L \in \mathcal{L}$ et $L' \in \mathcal{L}'$ l'image de $L'\ //\ L$ par une transduction rationnelle reste dans $\mathcal{L}'\ //\ \mathcal{L}$. Comme toute transduction rationnelle peut se décomposer en homomorphisme alphabétique, homomorphisme alphabétique inverse et intersection avec un langage rationnel [38], il suffit d'établir les lemmes suivants :

Lemme II.1

Soient $\mathcal{L}$ et $\mathcal{L}'$ deux cônes rationnels. Pour $L, L', R \subset X^*$ *avec* $L \in \mathcal{L}$, $L' \in \mathcal{L}'$ *et* $R \in \mathrm{Rat}$, $(L'\ //\ L) \cap R \in \mathcal{L}'\ //\ \mathcal{L}$.

Démonstration

Soient #, $\not\#$ deux nouvelles lettres. Alors $L_1 = \#\ L\ \not\# \in \mathcal{L}$ et $i(R,\ \#\ X^*\ \not\#) \in \mathrm{Rat}$. Comme le shuffle avec un langage rationnel est une transduction rationnelle

[19], $L'_1 = L' \sqcup\!\sqcup \{\#, \not\$\} = \{u \# v \not\$ w \,/\, u v w \in L'\} \in \mathcal{L}'$, d'où $L''_1 = L'_1 \cap i(R, \# X^* \not\$) \in \mathcal{L}'$. Alors il est facile de montrer l'égalité $L' // L = L''_1 // L_1$ et nous obtenons le résultat.

□

Lemme II.2

Soient $h : Y^* \to X^*$ *un homomorphisme alphabétique surjectif et* L, L' *des langages définis sur l'alphabet* X. *Alors* $h^{-1}(L' // L) = h^{-1}(L') // h^{-1}(L)$.

Démonstration

Soit $u \in h^{-1}(L' // L)$. Alors $h(u) \in L' // L$ et $u'_1 \, v \, u'_2 \in L'$ pour un certain $v \in L$ et une certaine décomposition $h(u) = u'_1 \, u'_2$. Comme h est surjectif, il existe $v' \in h^{-1}(v)$. En plus h étant alphabétique u peut se décomposer en $u_1 \, u_2$ tel que $u'_1 = h(u_1)$, $u'_2 = h(u_2)$ et $u_1 \, v' \, u_2 \in h^{-1}(u'_1 \, v \, u'_2) \subseteq h^{-1}(L')$ d'où $u \in h^{-1}(L') // h^{-1}(L)$.

L'inclusion inverse est évidente pour tout homomorphisme.

□

Lemme II.3

Soient $\mathcal{L}$, $\mathcal{L}'$ *deux cônes rationnels. Alors* $\mathcal{L}' // \mathcal{L}$ *est fermé par homomorphisme alphabétique inverse.*

Démonstration

Soeint $L \in \mathcal{L}$, $L' \in \mathcal{L}'$ où $L, L' \subseteq X^*$ et h un homomorphisme alphabétique de Y^* dans X^*. Posons $\bar{X} = \{\bar{x} \,/\, x \in X\}$ et définissons l'homomorphisme alphabétique $g : (Y \cup \bar{X})^* \to X^*$ par : $\forall\, y \in Y \;\; g(y) = h(y)$ et $\forall\, \bar{x} \in \bar{X}$ $g(\bar{x}) = x$. Ainsi g est un homomorphisme alphabétique surjectif et $h^{-1}(L'//L) = g^{-1}(L' // L) \cap Y^*$. D'où, d'après les deux lemmes précédents, nous pouvons déduire que $h^{-1}(L' // L) \in \mathcal{L}' // \mathcal{L}$.

□

Lemme II.4

Soient $\mathcal{L}$ et $\mathcal{L}'$ deux cônes rationnels. Alors $\mathcal{L}' \,//\, \mathcal{L}$ est clos par homomorphisme.

Démonstration

Soient $L \in \mathcal{L}$, $L' \in \mathcal{L}'$ où $L, L' \subseteq X^*$ et g un homomorphisme de X^* dans Y^*. On peut définir facilement une transduction rationnelle t telle que $t(L') = \{g(u_1) \# v \$ g(v_2) \;/\; u_1 v u_2 \in L'\}$ où $\#$ et $\$$ sont deux nouvelles lettres. Alors $g(L' \,//\, L) = t(L') \,//\, (\# L \$) \in \mathcal{L}' \,//\, \mathcal{L}$, d'où le résultat.

$\square$

Proposition II.5

Soient $\mathcal{L}$ et $\mathcal{L}'$ deux cônes rationnels. Alors $\mathcal{L}' \,//\, \mathcal{L}$ est un cône rationnel contenant $\mathcal{L}'$. En plus si $\mathcal{L}$ et $\mathcal{L}'$ sont clos par union, $\mathcal{L}' \,//\, \mathcal{L}$ le sera aussi.

Démonstration

Des lemmes II.1, II.3, II.4 nous pouvons affirmer que $\mathcal{L}' \,//\, \mathcal{L}$ est un cône rationnel. Comme le langage $\{\varepsilon\} \in \mathcal{L}$, pour tout $L' \in \mathcal{L}'$ $L' = L' \,//\, \{\varepsilon\} \in \mathcal{L}' \,//\, \mathcal{L}$ et $\mathcal{L}' \subseteq \mathcal{L}' \,//\, \mathcal{L}$.

Supposons maintenant que $\mathcal{L}$ et $\mathcal{L}'$ sont clos par union. Soient L, L_1 des langages de $\mathcal{L}$ et L', L'_1 des langages de $\mathcal{L}'$, nous avons à montrer que $(L' \,//\, L) \cup (L'_1 \,//\, L_1) \in \mathcal{L}' \,//\, \mathcal{L}$. Pour cela, considérons deux nouvelles lettres $\#$ et $\$$. Ainsi $L_2 = (\# L) \cup (\$ L_1) \in \mathcal{L}$ et $L'_2 = (L' \sqcup\!\sqcup \#) \cup (L'_1 \sqcup\!\sqcup \$) \in \mathcal{L}'$ et il est facile de montrer que $(L' \,//\, L) \cup (L'_1 \,//\, L_1) = L'_2 \,//\, L_2$.

$\square$

Dans le cas où $\mathcal{L}'$ est un cône rationnel principal, nous pouvons énoncer l'inclusion suivante :

Corollaire II.6

Soient L' *un langage algébrique défini sur un alphabet* X *et* $\mathcal{L}$ *un cône rationnel clos par union. Alors* $C_U(\{L'\} \,//\, \mathcal{L}) \subseteq C(L') \,//\, \mathcal{L}$.

Démonstration

Comme un cône rationnel principal est clos par union [20], nous avons, de la proposition précédente, que $C(L') \,//\, \mathcal{L}$ est un cône rationnel clos par union. Comme $\{L'\} \,//\, \mathcal{L} \subseteq C(L') \,//\, \mathcal{L}$, $C_U(L' \,//\, \mathcal{L}) \subseteq C_U(C(L') \,//\, \mathcal{L}) = C(L')//\mathcal{L}$.

□

Ces résultats tiennent pour les opérations quotient à gauche, quotient à droite. En particulier pour ces deux dernières, nous avons des résultats plus précis.

Proposition II.7

Soient L' *un langage algébrique défini sur un alphabet* X *et* $\mathcal{L}$ *un cône rationnel clos par union. Alors* $\mathcal{L}^{-1}C(L') = C_U(\mathcal{L}^{-1}\, L')$.

Démonstration

Il suffit de montrer que $\mathcal{L}^{-1}\, C(L') \subseteq C_U(\mathcal{L}^{-1}\, L')$. Pour cela considérons deux langages $L, L'_1 \subseteq Y^*$ tels que $L \in \mathcal{L}$ et $L'_1 \in C(L')$. Alors il existe un langage rationnel $R \subseteq Z^*$ et deux homomorphismes alphabétiques $h : Z^* \to X^*$ et $g : Z^* \to Y^*$ tels que $L'_1 = g(h^{-1}(L') \cap R)$. Nous avons à montrer que $L^{-1}\, L_1 \in C_U(\mathcal{L}^{-1}\, L')$.

Posons $L'_2 = h^{-1}(L') \cap R$. Comme g est alphabétique, il est facile de vérifier l'égalité suivante $L^{-1}\, L'_1 = L^{-1}\, g(L'_2) = g((g^{-1}(L))^{-1}.L'_2)$, ce qui nous ramène à montrer que $L_1^{-1}.(L'_3 \cap R) \in C_U(\mathcal{L}^{-1}\, L')$ où $L_1 = g^{-1}(L) \in \mathcal{L}$ et $L'_3 = h^{-1}(L')$.

R étant rationnel, il existe des langages rationnels $A_1, B_1, \ldots, A_n, B_n$ tels que $R = \bigcup_{i=1}^{n} A_i\, B_i$ et $\forall\, u\, v \in R$, $u \in A_i$ et $v \in B_i$ pour un $i \in [1, n]$.

Alors il est clair que $L_1^{-1}(L'_3 \cap R) = \bigcup_{i=1}^{n} (((L_1 \cap A_i)^{-1} L'_3) \cap B_i)$.

Il suffit maintenant de montrer que pour tout $i \in [1, n]$, $(L_1 \cap A_i)^{-1} L'_3 = C_i^{-1} L'_3 = C_i^{-1} h^{-1}(L') \in C_{\cup}(\mathcal{L}^{-1}.L')$. Comme $C_i \in \mathcal{L}$, le résultat vient de l'égalité suivante qui est évidente.

$$C_i^{-1} h^{-1}(L') = h^{-1}((h(C_i))^{-1} L').$$

□

Lemme II.8

Soient L *un langage défini sur un alphabet* X, $\mathcal{L}'$ *un cône rationnel. Alors* $\mathcal{L}' // C(L) = C(\mathcal{L}' // L)$.

Démonstration

D'après la proposition II.5, $\mathcal{L}' // C(L)$ est un cône rationnel et $\mathcal{L}' // L \subseteq \mathcal{L}' // C(L)$ implique que $C(\mathcal{L}' // L) \subseteq C(\mathcal{L}' // C(L)) = \mathcal{L}' // C(L)$. Pour l'inclusion inverse, considérons deux langages L_1, $L' \subseteq Y^*$ où $X \cap Y = \emptyset$, $L_1 \in C\ (L)$ et $L' \in \mathcal{L}'$. Alors il existe un langage rationnel R défini sur un alphabet Z disjoint de $X \cup Y$ et deux homomorphismes alphabétiques $h : Z^* \to X^*$ et $g : Z^* \to Y^*$ tels que $L_1 = g(h^{-1}(L) \cap R)$. Nous montrerons que $L' // L_1 \in C(\mathcal{L}' // L)$.

Définissons les homomorphismes alphabétiques $g_1 : (Z \cup Y)^* \to Y^*$ par $\forall z \in Z\ g_1(z) = g(z)$, $\forall y \in Y\ g_1(y) = y$ et $h_1 : (Z \cup Y)^* \to (X \cup Y)^*$ par $\forall z \in Z\ h_1(z) = h(z)$, $\forall y \in Y\ h_1(y) = y$.

Alors le langage $L'_1 = h_1(g_1^{-1}(L') \cap Y^* R\ Y^*) \in \mathcal{L}'$ est égal à $\{u\ h(v)\ w\ /\ v \in R,\ uw \in Y^*,\ u\ g(v)\ w \in L'\}$ et il est facile de vérifier que $L' // L_1 = (L'_1 // L) \cap Y^* \in C(\mathcal{L}' // L)$.

□

Le dernier résultat de ce paragraphe montre que l'opération quotient peut être obtenue en utilisant un homomorphisme effaçant et une intersection.

Proposition II.9

Soient $\mathcal{L}$ et $\mathcal{L}'$ deux cônes rationnels. Alors $\mathcal{L}' \mathbin{//} \mathcal{L} \subseteq \mathcal{H}_\alpha(\mathcal{L}' \wedge \mathcal{L}) =$ $\{h(L' \cap L) \;/\; L' \in \mathcal{L}', L \in \mathcal{L}$ *et* h *un homomorphisme alphabétique*$\}$.

Démonstration

Soient $L, L' \subseteq X^*$ où $L \in \mathcal{L}$ et $L' \in \mathcal{L}'$. $\bar{X} = \{\bar{x} \;/\; x \in X\}$. Définissons l'homomorphisme $h : X^* \to \bar{X}^*$ par $\forall\, x \in X\ h(x) = \bar{x}$. Alors $L_1 = X^*\, h(L)\, X^* \in \mathcal{L}$ et $L'_1 = \{u\, h(v)\, w \;/\; u\, v\, w \in L'\} \in \mathcal{L}'$. Il est facile de vérifier que $L' \mathbin{//} L_1 = \pi(L'_1 \cap L_1)$ où π est la projection de $X \cup X'$ dans X définie par $\pi(z) = z$ si $z \in X$, $\pi(z) = \varepsilon$ ailleurs.

□

Corollaire II.10

Tout cône rationnel clos par intersection est aussi fermé par quotient.

Corollaire II.11

$\mathcal{L}^{-1}\, \mathcal{L}' \subseteq \mathcal{H}_\alpha(\mathcal{L}' \wedge \mathcal{L})$.

B. QUOTIENT ET LANGAGES A UN COMPTEUR

Le but de ce paragraphe est de montrer que la famille des langages à un compteur, Rocl $= C(D'^*_1)$ est fermé par quotient à gauche (à droite). Commençons par montrer que Alg^{-1}.Rocl = Rocl. En employant la proposition II.7, il suffit de montrer que $\mathrm{Alg}^{-1}\, D'^*_1 \subseteq$ Rocl. Pour cela, considérons les langages algébriques sur deux lettres $A = \{a, b\}$ et montrons que pour tout langage algébrique $L \subseteq A^*$, il existe un langage rationnel $R \subseteq a^*$ tel que $L^{-1}D'^*_1 = R^{-1}D'^*_1$.

Nous utiliserons les notations suivantes

- Pour tout $w \in A^*$

  . $d(w) = |w|_a - |w|_b$

  . $\bar{d}(w) = \sup \{-d(w') \mid w'$ est un facteur gauche de $w\}$

  . $\mu(w) = b^i\, a^j$ où $i = \bar{d}(w)$ et $j = i+d(w)$

Ainsi $D'^*_1 = \{w \in A^* \,/\, d(w) = \bar{d}(w) = 0\} = \{w \in A^* \,/\, \mu(w) = \varepsilon\}$ et $\forall\, u, v \in A^*$, $u\, w\, v \in D'^*_1$ si et seulement si $u\, \mu(w)\, v \in D'^*_1$.

- Pour tout langage $L \subseteq A^*$, $\mu(L) = \{\mu(w) \,/\, w \in L\}$ et $\mu'(L) = \mu(L) \cap a^*$. Un langage $L \subseteq A^*$ est dit μ-*rationnel* si $\forall\, i \geq 0, \mu'(a^i\, L)$ est un langage rationnel.

- Enfin M désignera la famille $\{L \subseteq A^* \,/\, L$ est μ-rationnel$\}$.

Commençons par montrer que tout langage algébrique $L \subseteq A^*$ appartient à M. Nous avons besoin des résultats suivants :

Lemme II.12

Soient L *un langage unaire défini sur l'alphabet* $\{a\}$, i *un entier strictement positif et* R *le langage* $(a^i)^*$. *Alors* $R^{-1}L$ *et* RL *sont des langages rationnels.*

Lemme II.13

La famille M est fermée par union, produit et étoile, d'où $\mathrm{Rat} \,\square\, M = M$.

Démonstration

Comme $\mu'(a^i(L \cup L')) = \mu'(a^i L) \cup \mu'(a^i L')$, M est fermée par union.

Soient $L, L' \in M$, nous avons à montrer que $LL' \in M$ ainsi que L^*.

Soient i un entier ≥ 0. Alors, il est clair que $\mu'(a^i LL') = \mu'(\mu'(a^i L)L') = \mu'(R\, L')$ où R est un langage rationnel inclus dans a^*. Alors R est une union finie de langages de la forme $a^s(a^t)^*$. Ainsi, pour montrer que $\mu(R\, L')$ est rationnel, il suffit de montrer que $L'' = \mu'(a^s(a^t)^*L')$ l'est. Si $t = 0$ alors $L'' = \mu'(a^s L')$ est rationnel par hypothèse. Si $t > 0$, il est évident que $L'' = (a^t)^* L''$ et du lemme précédent, l'' est rationnel.

Pour l'opération étoile, considérons $L_1 = \mu'(a^i L^*)$. Si L_1 est fini, il est rationnel. Si L_1 est infini, il existe $w \in L^*$ tel que $\mu'(a^i w) = a^{i+t}$ où $t > 0$. Alors $L_1 = \mu(a^i w^* L^*) = \mu'(\mu'(a^i w^*)L^*) = \mu'(a^i(a^t)^*L^*) = (a^t)^* L_1 \in \mathrm{Rat}$ par le lemme précédent.

$\square$

Lemme II.14

Soient $i \in \mathbb{N}$ *et* $G = (A, N, P, S)$ *une grammaire algébrique satisfaisant la propriété suivante*

(P) *il existe une dérivation* $S \overset{*}{\underset{G}{\Rightarrow}} u\, S\, v$ *où* $u, v \in A^*$, $\mu(a^i u) \in a^i a^*$ *et soit* $d(u) \neq 0$ *soit* $d(v) \neq 0$.

Alors $\mu'(a^i L(G))$ *est un langage rationnel.*

Démonstration

Par hypothèse $L(G) = \{u^n w v^n \;/\; n \geq 0,\ w \in L(G)\}$, $\mu(a^i u) = a^{i+k}$ où $k = d(u) \geq 0$ et $\mu(v) = b^s a^t$ où $k' = d(v) = t-s$.

Alors $L' = \mu'(a^i\ L(G)) = L_1 \cup L_2$ où
$L_1 = \{a^j \in L' \ /\ j < s\} \in Rat$ et $L_2 = \{a^j \in L' \ /\ j \geq s\}$.

Ainsi il suffit de montrer que $L_2 \in Rat$. Nous distinguons trois cas :

Premier cas : $p = k + k' > 0$.

Alors $L_2 = (a^p)^*\ L_2 \in Rat$. En effet, si $a^j \in L_2$, alors $a^j = \mu(a^i\ w)$ pour un certain $w \in L(G)$. Comme, pour tout $n \geq 0$, $u^n\ w\ v^n \in L(G)$, il est facile de montrer que $\bar{d}(a^i\ u^n\ w\ v^n) = j + np$. D'où $(a^p)^n\ a^j \in L_2$ et $(a^p)^*\ L_2 = L_2$.

Second cas : $-p = k + k' < 0$.

Alors $k' < 0$ et $s > t$. Nous montrerons que $L_2 = ((a^p)^*)^{-1}\ L_2 \cap a^s\ a^* \in Rat$.

Pour cela, considérons $a^m \in L_2$, $w \in L(G)$ tel que $a^m = \mu(a^i\ w)$ et $n \in \mathbb{N}$ vérifiant $\ell = m-np \geq s$. Alors $u^n\ w\ v^n \in L(G)$, $d(a^i\ u^n\ w\ v^n) = a^\ell$ et comme $\bar{d}(a^i\ u) = 0$, $d(u) \geq 0$, $\ell \geq s > t$, on peut vérifier que $\bar{d}(a^i\ u^n\ w\ v^n) = 0$. Ainsi $a^\ell = \mu(a^i\ u^n\ w\ v^n) \in L_2$ d'où l'égalité.

Troisième cas : $k' + k = 0$ avec $k > 0$.

Soit $L = \{a^j \ /\ j = d(a^i\ w)$ où $w \in L(G)\}$ et montrons que $L_2 = L \cap a^s\ a^*$.

Il est clair que $L_2 \subseteq L \cap a^s\ a^*$.

Pour l'inclusion inverse, prenons $a^j \in L$ où $j \geq s$. Alors, il existe $w \in L(G)$ tel que $j = d(a^i\ w)$ et $\bar{d}(w) = r$. Ainsi $u^r\ w\ v^r \in L(G)$, $d(a^i\ u^r\ w\ v^r) = j$ et comme $d(u^r) \geq r$, $j \geq s > t$, il est facile de vérifier que $\bar{d}(a^i\ u^r\ w\ v^r) = 0$.

Ainsi $a^j = \mu(a^i\ u^r\ w\ v^r) \in L_2$ et $L_2 = L \cap a^s\ a^*$. Il revient à montrer que L est rationnel. Pour cela, considérons L' la fermeture commutative de $a^i\ L(G)$. Comme l'alphabet A ne contient que deux lettres, L' est algébrique [31] et $L = ((ab)^*)^{-1}\ L' \cap a^*$ est un langage algébrique unaire, donc rationnel.

□

Lemme II.15

Tout langage algébrique $L \subseteq A^*$ *est* μ*-rationnel.*

Démonstration

La démonstration se fait par récurrence sur le nombre de variables non terminales de la grammaire algébrique engendrant L. Plus précisément, pour une grammaire algébrique réduite $G = (A, N, P, S)$ et pour $B \in N$, $Acc(B)$ désignera l'ensemble $\{C \in N \;/\; B \overset{*}{\underset{G}{\Rightarrow}} u\,Cv$ pour $u, v \in A^*\}$.

Alors l'hypothèse de récurrence peut s'énoncer exactement comme suit :

(H_n) Si $G = (A, N, P, S)$ est une grammaire algébrique réduite telle que $Acc(S)$ contient au plus n éléments, alors $L(G)$ est μ-rationnel.

(H_o) est vérifiée car si $n = 0$, $L(G)$ est fini.

Considérons, maintenant, une grammaire algébrique réduite $G = (A, N, P, S)$ telle que $Acc(S)$ contient $n+1$ éléments avec $n \geq 0$. Distinguons les deux cas :

Premier cas : $S \in Acc(S)$

Prenons $i \in \mathbb{N}$. Si G vérifie la propriété (P) du lemme précédent alors $\mu'(a^i(L(G))) \in Rat$ et L est μ-rationnel.

Sinon prenons un entier k vérifiant $\forall B \in N$ tel que $S \in Acc(B)$, il existe une dérivation $B \overset{*}{\underset{G}{\Rightarrow}} u\,S\,v$ où $u, v \in A^*$ et $|u| \leq k$, alors il est facile de montrer la propriété suivante

(Q) Il existe un entier positif k tel que si $S \overset{*}{\underset{G}{\Rightarrow}} u\,B\,v$ où $u \in A^*$, $\bar{d}(a^i\,u) = 0$, $B \in N$ et $S \in Acc(B)$ alors $d(u) \leq k$.

Considérons, maintenant, la grammaire régulière $G' = (A', N', P', (i,s))$

où $A' = A \cup (N \setminus \{S\})$

$N' = [0, i+k] \times N$

$P' = P'_1 \cup P'_2$ où

$P'_1 = \{(q, B) \to (p, C)\ v \ /\ v \in A'^*$, il existe $u \in A^*$ vêfifiant $\mu(a^q\ u) = a^p$ et $u' \overset{*}{\underset{G}{\Rightarrow}} u$ pour un certain u' tel que $B \to u'\ C\ v \in P\}$

$P'_2 = \{(j, S) \to a^j\ v \ /\ v \in A^*,\ S \to v \in P\}$.

Nous avons à définir d'autres grammaires. Pour $B \in N'' = N \setminus \{S\}$, G_B désignera la grammaire réduite construite à partir de la grammaire (A, N'', P'', B) où $P'' = \{C \to u \in P \ /\ C \in N'',\ u \in (A \cup N'')^*\}$. Nous montrerons que $\mu'(a^i L(G)) = \mu'(s(L(G')))$ où s est la substitution définie sur A' par $s(a) = \{a\}$, $s(b) = \{b\}$ et $\forall\ B \in N''\ s(B) = L(G_B)$.

Soit $a^t \in \mu'(s(L(G')))$. Alors il existe $w' \in L(G')$ et $w \in s(w')$ tels que $a^t = \mu(w)$. Ainsi $(i, s) \overset{*}{\underset{G'}{\Rightarrow}} (j, s)\ w'_2 \underset{G'}{\Rightarrow} a^j\ w'_1\ w'_2$ et $w' \overset{*}{\underset{G}{\Rightarrow}} a^j\ v = w$. Par construction de G', nous obtenons $S \to w'_1 \in P$ et $S \overset{*}{\underset{G}{\Rightarrow}} u\ S\ w'_2$ où $u \in A^*$ et $\mu(a^i\ u) = a^j$. D'où $S \overset{*}{\underset{G}{\Rightarrow}} u\ w'_1\ w'_2 \overset{*}{\underset{G}{\Rightarrow}} uv$ avec $a^t = \mu(a^j\ v) = \mu(\mu(a^i u)v) = \mu(a^i\ uv) \in \mu'(a^i(L(G)))$.

Pour l'inclusion inverse, soit $a^t \in \mu'(a^i(L(G)))$. Alors $a^t = \mu(a^i w)$ pour un $w \in L(G)$. Considérons une dérivation gauche à partir de S jusqu'à w dans la grammaire G. En considérant dans cette dérivation le dernier mot où S apparaît, nous obtenons $S \overset{*\ell}{\underset{G}{\Rightarrow}} w_1\ S\ v \overset{\ell}{\underset{G}{\Rightarrow}} w_1\ u\ v \overset{*\ell}{\underset{G}{\Rightarrow}} w_1\ w_2 = w$ et $w_2 \in s(u\ v)$. De la propriété (Q), $\mu(a^i\ w_1) = a^j$ avec $j \leq i+k$ et par construction de G', nous avons dans G' la dérivation suivante :

$$(i, s) \overset{*}{\underset{G'}{\Rightarrow}} (j, s)\ v \underset{G'}{\Rightarrow} a^j\ u\ v \in L(G').$$

Ainsi $a^t = \mu(a^i\ w_1\ w_2) = \mu(\mu(a^i w_1)w_2) = \mu(a^j\ w_2) \in \mu(s(a^j\ u\ v)) \subseteq \mu(s(L(G')))$. De l'hypothèse de récurrence, $\forall\ z \in A'$, $s(z)$ est μ-rationnel et comme $L(G')$ est un langage rationnel, nous obtenons du lemme II.13 que $s(L(G')) \in \mathcal{M}$ d'où $\mu'(a^i(L(G))) = \mu'(s(L(G')))$ est rationnel.

<u>Second cas</u> : $S \notin Acc(S)$.

Alors $L(G) = s(F)$ où F est l'ensemble fini suivant $\{u/S \to u \in P\}$ et

s la substitution définie au premier cas. Soit $B \in N'' = N \setminus \{S\}$. Si $B \in Acc(B)$, $s(B)$ est μ-rationnel (cf. Premier cas). Sinon nous pouvons appliquer l'hypothèse de récurrence. D'où $\forall z \in A'$ $s(z)$ est μ-rationnel et $L(G) = s(F)$ est μ-rationnel.

□

Proposition II.16

$Alg^{-1}\ Rocl = Rocl$, *i.e., la famille des langages à un compteur est fermée par quotient à gauche par un langage algébrique.*

Démonstration

De la proposition II.7, $Alg^{-1}\ Rocl = Alg^{-1}\ C(D'^{*}_{1}) = C_{U}(Alg^{-1}\ D'^{*}_{1})$. Ainsi il suffit de montrer que pour un langage algébrique L, $L^{-1}\ D'^{*}_{1} \in Rocl$. Nous montrerons que $L^{-1}\ D'^{*}_{1} = R^{-1}\ D'^{*}_{1}$ où $R = \mu'(L \cap A)$.

Si $v \in L^{-1}\ D'^{*}_{1}$, alors $u\,v \in D'^{*}_{1}$ pour un certain $u \in L \cap A^{*}$. Comme $\bar{d}(uv) = 0$, $\bar{d}(u) = 0$ et $\mu(u) = a^{j} \in \mu'(L \cap A^{*}) = R$. En plus, $uv \in D'^{*}_{1}$ implique que $\mu(u)\ v \in D'^{*}_{1}$ d'où $v \in R^{-1}\ D'^{*}_{1}$.

Réciproquement, si $v \in R^{-1}\ D'^{*}_{1}$, alors $a^{j}\ v \in D'^{*}_{1}$ avec $a^{j} \in \mu(L \cap A^{*})$. Prenons $u \in A^{*} \cap L$ tel que $\mu(u) = a^{j}$. Alors $\mu(u)\ v \in D'^{*}_{1}$ implique $uv \in D'^{*}_{1}$ et $v \in L^{-1}\ D'^{*}_{1}$.

Du lemme précédent, $R = \mu'(L \cap A^{*})$ est un langage rationnel. D'où $L^{-1}\ D'^{*}_{1} = R^{-1}\ D'^{*}_{1} \in C(D'^{*}_{1}) = Rocl$.

□

Corollaire II.17

La famille des langages à un compteur est fermée par quotient à gauche.

C. EXTENSIONS

Une question qu'on peut se poser est peut-on étendre ce résultat à la FAL générée par Rocl ?

Remarquons d'abord que pour Lin, une autre sous famille classique de Alg, nous obtenons un résultat complétement différent. En effet, il est facile d'établir que tout langage récursivement énumérable est le quotient de deux langages linéaires.

Proposition II.18

$Lin^{-1}\ Lin = r.e.$

Démonstration

Comme r.e. est un cône rationnel clos par intersection, r.e. est clos par quotient [Corollaire II.7] et $Lin^{-1}\ Lin \subseteq (r.e.)^{-1}\ r.e. = r.e.$

Pour l'inclusion réciproque, considérons un langage récursivement énumérable $L \subseteq X^*$, engendré par la grammaire $G = (X, N, P, S)$ où $P \subseteq N^+ \times V^*$.

En gros, nous construirons deux langages linéaires qui permettent de simuler une dérivation dans G et d'effacer par quotient cette dérivation excepté le dernier mot.

Soient $V = X \cup N$ et $Y = V \cup \{\Rightarrow, \Leftarrow\}$.

Pour $u \in V^*$, u^R désignera le mot mirroir de u. Définissons deux langages sur l'alphabet Y :

$L_1 = \{u_n^R \Leftarrow \ldots \Leftarrow u_1^R \Leftarrow S \Rightarrow u_1 \ldots \Rightarrow u_n \Rightarrow / n \geq 1,\ \forall\, i \in [1,n]\ u_i \in V^+\}$

$L_2 = \{u_n^R \Leftarrow \ldots \Leftarrow u_1^R \Leftarrow u_o^R \Rightarrow v_o \ldots \Rightarrow v_n / n \geq 0,\ s = u_o,\ \forall\, i \in [1, n],$
$u_i, v_i \in V^*$ et $u_i \underset{G}{\Rightarrow} v_i\}$

L_1 est un langage linéaire engendré par la grammaire $G_1 = (Y, N_1, P_1, \sigma)$ où $P_1 = \{(\sigma, \sigma_1 \Rightarrow), (\sigma_1, S)\} \cup \{(\sigma_1, a_i\, \sigma_1\, a_i) / a_i \in Y\}$

On peut aussi construire une grammaire linéaire $G_2 = (Y, N_2, P_2, \sigma)$ engendrant L_2 où $P_2 = \{(\sigma_1, \Leftarrow \sigma \Rightarrow)\} \cup \{(\sigma, u^R \sigma_1 v) / (u, v) \in P\} \cup \{(\sigma, \Rightarrow)\}$.

Enfin il est facile de vérifier que $L = L_1^{-1} L_2 \cap X^*$, d'où $L \in C(L_1^{-1} L_2) \subseteq \mathrm{Lin}^{-1}\,\mathrm{Lin}$.

□

De cette proposition et de la proposition II.9, nous déduisons un résultat de Baker et Book.

Corollaire II.19 [2]

Tout langage r.e. est l'image par un homomorphisme alphabétique de l'intersection de deux linéaires.

Corollaire II.20

Lin // Lin = r.e.

Considérons maintenant Ocl = Rat □ Rocl, la FAL engendrée par Rocl. Alors les langages $L_1 = a\{b^n a^n / n \geq 1\}^*$ et $L_2 = \{a^n b^{2n} / n \geq 1\}^*$ appartiennent à Ocl. Il est clair que le langage $L' = \{b^{2^n} / n \geq 1\}$ est égal à $L_1^{-1} L_2 \cap b^*$ d'où $L' \in C(L_1^{-1} L_2) \subseteq \mathrm{Ocl}^{-1}\,\mathrm{Ocl} \setminus \mathrm{Alg}$. Ocl n'est pas clos par quotient et nous pouvons énoncer :

Proposition II.21

Ocl^{-1} Ocl *n'est pas inclus dans* Alg.

Nous pouvons nous demander si on aura le même résultat pour la famille Rocl^{-1} Ocl. La réponse est non. En fait, par la proposition II.16 et le lemme suivant, nous avons $\mathrm{Rocl}^{-1}\,\mathrm{Ocl} = \mathrm{Ocl}$.

Lemme II.22

Soient L_1 et L_2 deux cônes rationnels.
Alors $L_1^{-1}(\mathrm{Rat} \,\square\, L_2) \subseteq \mathrm{Rat} \,\square\, (((\mathrm{Rat} \,\square\, L_2)^{-1} L_1)^{-1} L_2)$.

<u>Démonstration</u>

Soient $L_1 \in \mathcal{L}_1$ et $L_2 \in \mathrm{Rat} \,\square\, \mathcal{L}_2$. Alors $L_2 = s(R)$ où R est un langage rationnel inclus dans Z^+ et s une $\mathcal{L}_2$ substitution.

Comme R est rationnel, $R = \bigcup_{i=1}^{n} A_i \, z_i \, B_i$ où $\forall\, i \in [1, n]$ A_i, $B_i \in$ Rat, $z_i \in Z$ et $\forall$ u z v $\in$ R avec $z \in Z$, il existe $i \in [1, n]$ tel que $u \in A_i$, $z = z_i$ et $v \in B_i$. Alors il est facile de vérifier l'égalité

$$L_1^{-1}\ s(R) = \bigcup_{i=1}^{n} \left((s(A_i)^{-1}\ L_1)^{-1}\ s(z_i) \right) s(B_i)$$

qui implique clairement le résultat.

$\square$

Si $(\mathrm{Rat} \,\square\, \mathcal{L}_2)^{-1}\ \mathcal{L}_2$ et $(\mathrm{Rat} \,\square\, \mathcal{L}_2)^{-1}\ \mathcal{L}_1$ sont inclus dans $\mathrm{Rat} \,\square\, \mathcal{L}_2$, alors nous obtenons $\mathcal{L}_1^{-1}(\mathrm{Rat} \,\square\, \mathcal{L}_2) = \mathrm{Rat} \,\square\, \mathcal{L}_2$. Dans le cas où $\mathcal{L}_1 = \mathcal{L}_2 = \mathrm{Rocl}$, $\mathrm{Ocl}^{-1}\ \mathrm{Rocl} \subseteq \mathrm{Alg}^{-1}\ \mathrm{Rocl} = \mathrm{Rocl} \subseteq \mathrm{Ocl}$ d'où la

<u>Proposition II.23</u>

$\mathrm{Rocl}^{-1}.\mathrm{Ocl} = \mathrm{Ocl}$.

De la proposition II.18 ou II.21, il est clair que Alg n'est pas clos par quotient. En plus, on est amené à se demander si on peut avoir un résultat similaire à celui de la proposition II.16, plus précisément $\mathrm{Rocl}^{-1}\ \mathrm{Alg} = \mathrm{Alg}$ i.e. que Alg peut être clos par quotient par un langage à un compteur.

Pour obtenir une réponse négative à cette question, il suffit de trouver un langage $L \in \mathrm{Alg}$ tel que $(D'^{*}_{1})^{-1} \notin \mathrm{Alg}$. D'après l'égalité $\mathrm{Rocl}^{-1}\ \mathrm{Ocl} = \mathrm{Ocl} \subseteq \mathrm{Alg}$, ce langage ne peut pas être dans Ocl. Prenons le donc dans $\mathrm{Ocl} \,\square\, \mathrm{Ocl}$. Plus précisément, si $\mathcal{L}_{=}$ désigne le cône rationnel engendré par le langage $\{a^n\ b^n \ /\ n \geq 0\}$, nous obtenons :

Proposition II.24

Il existe un langage $L \in \mathcal{L}_= \square \mathcal{L}_=$ *tel que* $D'^{*-1}_1 L \notin Alg$.

Démonstration

Prenons $L = s(L_1)$ où $L_1 = \{x^n\ y\ d^n\ /\ n \geq 1\} \in \mathcal{L}_=$ et s la substitution définie par $s(x) = \{b^t\ a^{2t+1}\ /\ t \geq 0\}$, $s(y) = \{b^p\ c^p\ /\ p \geq 0\}$ et $s(d) = \{d\}$. Alors $L \in \mathcal{L}_= \square \mathcal{L}_=$ et $(D'^{*}_1)^{-1} L \notin Alg$ puisqu'il est facile de vérifier que

$$(D'^{*}_1)^{-1}\ L \cap c^*\ d^* = \{c^i\ d^j\ /\ 1 \leq j \leq i < 2^j\}.$$

$\square$

Si on se restreint aux langages linéaires, on peut avoir :

Lemme II.25

$Rocl^{-1}\ Lin \subseteq Alg$.

$Rocl^{-1}\ (Rat \square Lin) \subseteq Alg$.

Démonstration

Lin est par définition le cône principal engendré par $Sym = \{w\ \bar{w}^R\ /\ w \in X = \{a, b\}\}$. Comme $Rocl^{-1}\ Lin = C_U(Rocl^{-1}\ Sym)$, il suffit de montrer que pour tout langage $L \in Rocl$, $L^{-1}\ Sym \in Alg$. A partir de l'égalité suivante qui se vérifie facilement $L^{-1}\ Sym = Sym\ \overline{(L \cap X^*)}^R \cup \overline{((L \cap X^+\ \bar{X}^+).Sym^{-1}}^R$ et de l'inclusion $Rocl\ Lin^{-1} \subseteq Rocl\ Alg^{-1} \subseteq Rocl \subseteq Alg$ nous obtenons la première inclusion. La seconde découle de la première et du lemme II.22.

$\square$

Remarquons que ces résultats restent vrais pour le quotient à droite. Mais il n'est pas facile d'étendre ces résultats pour le quotient au centre. En effet, il semble que l'on peut énoncer :

Conjecture 5

Si $L_1 \in Rocl$ et $L_2 \in Alg$, *alors* $L_1\ //\ L_2 = \{u_1\ u_2\ /\ u_1\ v\ u_2 \in L_1$ *pour un* $v \in L_2\} \in Rocl$.

CHAPITRE II

LES LANGAGES ALGEBRIQUES BORNES ET LES LANGAGES ALGEBRIQUES COMMUTATIFS

A. LES-ISA FAMILLES DE LANGAGES

Définition II.26

L est un ISA-langage si et seulement si pour toute substitution algébrique s, $s^{-1}(L) \in$ Alg.

Définition II.27

L est une ISA-famille si et seulement si pour tout $L \in \mathcal{L}$, L *est un ISA-langage.*

Rat, la famille des langages rationnels est une ISA famille. En effet, montrons une propriété plus générale.

Lemme II.28

Pour toute substitution s, *tout langage rationnel* $s^{-1}(R)$ *est rationnel.*

Démonstration

Soient s une substitution de X^* dans Y^* et $R \subseteq Y^*$ un langage rationnel. Prouvons que le langage $s^{-1}(R) = \{u \in X^* \ / \ s(u) \cap R \neq \emptyset\}$ est rationnel.

Soit $A = (Y, Q, q_o, *, F)$ l'automate d'états fini déterministe minimal qui reconnait R où

- Q est l'ensemble des états
- $q_o \in Q$, l'état initial
- $*$ la fonction de transition
- F l'ensemble des états terminaux, $F \subseteq Q$.

Nous allons construire un automate B d'états fini non déterministe, non complétement spécifié reconnaissant $s^{-1}(R)$.

Soit $B = (X, Q, q_o, \circledast, F)$ où $\circledast$ est définie par

pour tout $w \in X \cup \{\varepsilon\}$, tout $q \in Q$

$q \circledast w = \{q' \in Q \ / \ \exists \ \gamma \in s(w)$ tel que $q * \gamma = q'\}$.

Cette définition peut s'étendre facilement sur X^* par

$\forall \ (q, q') \in Q^2 \ \forall \ w \in X^* \ q' \in q \circledast w \iff \exists \ v \in s(w) \ q * v = q'$.

Ainsi $T(B) = s^{-1}(R)$.

En effet, soit $w \in T(B)$, il existe $q_n \in F$ tel que $q_n \in q_o \circledast w$, d'où $q_o * v = q_n$ où $v \in s(w)$ et $q_n \in F$ autrement dit $v \in R$ et $w \in s^{-1}(v)$ $\subseteq s^{-1}(R)$.

Réciproquement, si $w \in s^{-1}(R)$, il existe $v \in R$ tel que $w \in s^{-1}(v)$, plus précisément $q_o * v = q_n$ où $q_n \in F$, $v \in R$ et $v \in s(w)$ d'où $q_n \in q_o \circledast w$ avec $q_n \in F$ donc $w \in T(B)$.

□

Etudions d'abord le comportement d'un ISA-langage vis à vis d'une transduction rationnelle. Comme toute transduction rationnelle τ peut être décomposée en $\tau = g \circ \cap \ R \circ h^{-1}$ où g et h sont des homomorphismes alphabétiques et R un langage rationnel [38], il suffit d'établir les lemmes suivants :

<u>*Lemme II.29*</u>

Soit L *un ISA-langage,* h *un homomorphisme alphabétique,* $h^{-1}(L)$ *est un ISA-langage.*

<u>Démonstration</u>

Pour tout homomorphisme alphabétique h et toute substitution algébrique s, $(h \circ s)^{-1} = s^{-1} \circ h^{-1}$ et $s^{-1}(h^{-1}(L)) = (h \circ s)^{-1}(L)$ pour tout langage L. Il est évident que $h \circ s$ est une substitution algébrique, étant une composition de deux substitutions algébriques.

L étant un ISA$^-$langage, $(h \circ s)^{-1}(L) \in Alg$ par conséquent $h^{-1}(L)$ est un ISA$^-$langage.

□

En nous rappelant que si h est un homomorphisme alphabétique,surjectif, h^{-1} est une substitution rationnelle donc algébrique, que $s^{-1} \circ h = (h^{-1} \circ s)^{-1}$ pour toutes applications h et s, une des propriétés vérifiée par la composition de deux substitutions algébriques nous permet d'énoncer :

Lemme II.30

Soient L *un ISA-langage,* h *un homomorphisme alphabétique. Alors* h(L) *est un ISA-langage.*

Lemme II.31

Soient L *un ISA-langage défini sur un alphabet* Y, R *un langage rationnel. Alors* $L \cap R$ *est un ISA-langage.*

Démonstration

Il s'agit d'une démonstration constructive. Soient s une substitution algébrique de X^* dans Y^*, R un rationnel défini sur Y.

a - Construction

Soient $A = (Y, Q, q_o, *, F)$ l'automate d'états fini, déterministe, minimal reconnaissant R où

- . Q est l'ensemble des états
- . $q_o \in Q$, l'état initial
- . $*$ la fonction de transition
- . $F \subseteq Q$, l'ensemble des états terminaux,

et $B = (X, Q, q_o, \circledast, F)$ un automate d'états fini, non déterministe, non complétement spécifié défini au lemme II.28, reconnaissant $s^{-1}(R)$.

Soit R' le langage rationnel défini sur l'alphabet $\Delta = W \times X \times Q$ par $R' = \{(q_o, x_1, q_1)(q_1, x_2, q_2) \ldots (q_{n-1}, x_n, q_n) \ / \ q_n \in F\}$.

On définit l'homomorphisme alphabétique Φ de Δ^* dans X^* par $\Phi((q, x, q')) = x$ pour tout $(q, x, q') \in \Delta$, le rationnel $R_{qq'} = \{w \in Y^* \ / \ q * w = q'\}$ et la substitution t de Δ^* dans 2^{Y^*} par :

$$t((q, x, q')) = s(\Phi(q, x, q')) \cap R_{qq'} = s(x) \cap R_{qq'}.$$

La substitution t est bien évidemment, une substitution algébrique. Alors $s^{-1}(L \cap R) = \ (t^{-1}(L) \cap R')$.

b - Validité de la construction

En effet, soit $w \in s^{-1}(L \cap R)$, il existe $\ell \in L \cap R$ tel que $\ell \in s(w)$.

D'une part, comme $w = w_1 \ldots w_n$ où $w_i \in X$, on peut trouver une décomposition de ℓ en $\ell_1 \ldots \ell_n$ ($\ell_i \in Y^*$) vérifiant $\ell_i \in s(w_i)$ pour tout $i \in [1, n]$.

D'autre part, ℓ appartenant à R implique l'existence de $q_1, \ldots, q_n$ tels que $q_n \in F$ et $q_{i-1} * \ell = q_i$ pour tout $i \in [1, n]$ autrement dit $\ell_i \in R_{q_{i-1}q_i}$.

Ainsi $\ell_i \in s(w_i) \cap R_{q_{i-1}q_i}$ ou encore $\ell_i \in t((q_{i-1}, w_i, q_i))$. Ce qui implique que

$$\ell \in t((q_o, w_1, q_1)(q_1, w_2, q_2) \ldots (q_{n-1}, w_n, q_n)).$$

Par conséquent $w' = (q_o, w_1, q_1) \ldots (q_{n-1}, w_n, q_n) \in t^{-1}(\ell) \subseteq t^{-1}(L)$. Comme $w' = (q_o, w_1, q_1) \ldots (q_{n-1}, w_n, q_n) \in R'$, il est clair que $w \ = \Phi(w') \in \Phi(t^{-1}(L) \cap R')$.

Réciproquement, soit $w \in \Phi(t^{-1}(L) \cap R')$, $w = \Phi(v)$ où $v \in t^{-1}(L) \cap R'$.

Le mot v appartient à $t^{-1}(L)$ d'une part, i.e. qu'il existe $\ell \in L$ tel que $\ell \in t(v)$ et d'autre part peut se mettre sous la forme

$$v = (q_o, v_1, q_1)(q_1, v_2, q_2) \ldots (q_{n-1}, v_n, q_n) \text{ où } q_n \in F.$$

D'où $\ell = \ell_1 \ldots \ell_n$ avec $\ell_i \in t((q_{i-1}, v_i, q_i)) = s(v_i) \cap R_{q_{i-1}q_i}$. Par conséquent $\ell \in R_{q_o q_1} \ldots R_{q_{n-1}q_n} \subseteq R$, et comme $\ell \in s(v_1 \ldots v_n) = s(\Phi(v)) = s(w)$ alors $w \in s^{-1}(\ell) \subseteq s^{-1}(R \cap L)$.

□

Ainsi pouvons nous déduire des lemmes précédents la

Proposition II.32

Pour tout ISA-langage L, C(L) *est une ISA famille de langages.*

Montrons maintenant une propriété vérifiée par toutes les ISA-familles de langages. Pour cela établissons d'abord le lemme suivant :

Lemme II.33

Soit L *un langage agébrique inclus dans* $X^* Y^*$ *où* $X \cap Y = \emptyset$, *alors il existe*

un langage linéaire L_o *inclus dans* Z^*,

une substitution algébrique s *vérifiant* $s(z) \subseteq X^*$ *où* $s(z) \subseteq Y^*$ *pour tout* $z \in Z$,

tels que $L = s(L_o)$.

Démonstration

Il s'agit d'une démonstration constructive.

a - Construction

Soit $L \subseteq X^* Y^*$ un langage algébrique, il existe une grammaire algébrique réduite $G = (X \cup Y, N, P, A)$ engendrant L. Considérons les sous ensembles suivants :

$$N_X = \{B \in N \ / \ L(G, B) \subseteq X^*\}$$

$$N_Y = \{C \in N \ / \ L(G, C) \subseteq Y^*\}$$

$$N' = N \setminus (N_X \cup N_Y)$$

et la grammaire G' définie par $G' = (X \cup Y \cup N_X \cup N_Y, N', P', A)$ où

$$P' = P \cap [N' \times (X \cup Y \cup N)^*] \text{ et } L_o = L(G', A).$$

On définit la substitution s de $X \cup Y \cup N_X \cup N_Y$ dans $X \cup Y$ par

$$s(x) = \{x\} \qquad \text{si } x \in X \cup Y$$

$$s(B) = L(G, B) \text{ si } x \in N_X \cup N_Y.$$

b - Validité de la construction

Il est clair que s est une substitution algébrique et que pour tout $z \in Z = X \cup Y \cup N_X \cup N_Y$

$$s(z) \subseteq X^* \quad \text{si } z \in X \cup N_X$$

ou $\quad s(z) \subseteq Y^* \quad \text{si } z \in Y \cup N_Y$

Montrons que $L = s(L_o)$.

En effet, si $w \in L$, il existe une dérivation $A \overset{*}{\underset{G}{\rightarrow}} w$ où les premières étapes seront des applications des règles de G'. $A \rightarrow w_1 \rightarrow \ldots \rightarrow w_i \rightarrow \ldots \rightarrow w_n = w$
Soit w_i le premier mot de la dérivation ne contenant pas de variables non terminales de N', alors

$$A \overset{*}{\underset{G'}{\rightarrow}} w_i \text{ où } w_i \in L_o,$$

$$w_i \overset{*}{\underset{G}{\rightarrow}} w \quad \text{où } w \in X^* Y^*$$

avec $w_i = x_1 B_1 \ldots x_n B_m x_{m+1} y_1 C_1 \ldots y_p C_p y_{p+1}$

où $\quad x_i \in X^* \quad B_i \in N_X \cup \{\varepsilon\}$ pour tout $i \in [1, m]$,

$\quad y_j \in Y^* \quad C_j \in N_Y \cup \{\varepsilon\}$ pour tout $j \in [1, p]$,

$\quad x_{m+1} \in X^*, \ y_{p+1} \in Y^*.$

Par conséquent $w = x_1 \beta_1 \ldots \beta_m x_{m+1} y_1 \gamma_1 \ldots \gamma_p y_{p+1}$

où $\quad B_i \overset{*}{\underset{G}{\rightarrow}} \beta_i$ pour tout $i \in [1, n]$

$\quad C_j \overset{*}{\underset{G}{\rightarrow}} \gamma_j$ pour tout $j \in [1, p]$

$B_i = \beta_i = \varepsilon$ ou $\beta_i \in L(G, B_i) = s(B_i)\ \forall\, i \in [1, m]$

$C_j = \gamma_j = \varepsilon$ ou $\gamma_j \in L(G, C_j) = s(C_j)\ \forall\, j \in [1, p]$

$x_i \in s(x_i)$ pour tout $i \in [1, m+1]$

$y_j \in s(y_j)$ pour tout $j \in [1, p+1]$

et $w \in s(w_i) \subseteq s(L_o)$.

Réciproquement, si $\gamma \in s(L_o)$, il existe $\alpha \in L_o \subseteq (X \cup Y \cup N_X \cup N_Y)^*$ tel que $\gamma \in s(\alpha)$ i.e. que $A \overset{*}{\underset{G'}{\to}} \alpha$ donc $A \overset{*}{\underset{G}{\to}} \alpha$.

Comme $\alpha = \alpha_1 \ldots \alpha_n$, $\alpha_i \in X \cup N_X \cup N_Y \cup Y$, alors $s(\alpha) = s(\alpha_1) \ldots s(\alpha_n)$.

Par conséquent, on peut trouver une décomposition de $\gamma = \beta_1 \ldots \beta_n$ vérifiant $\beta_i \in s(\alpha_i)$ pour tout $i \in [1, n]$ et $s(\alpha_i) = \{\alpha_i\}$ si $\alpha_i \in X \cup Y$,

$s(\alpha_i) = L(G, \alpha_i)$ sinon.

D'où $\alpha_i \overset{*}{\underset{G}{\to}} \beta_i$ pour tout $i \in [1, n]$.

Ainsi pour tout $\gamma \in s(L_o)$, $A \overset{*}{\underset{G}{\to}} \alpha \overset{*}{\underset{G}{\to}} \gamma$ et $\gamma \in L$

Il nous reste à montrer que L_o est linéaire ou plus précisément $P' \subseteq N' \times ((X \cup N_X)^* (N' \cup \{\varepsilon\}) (Y \cup N_Y)^*)$.

Supposons qu'il existe une dérivation $A \overset{*}{\underset{G'}{\to}} w_1 N_1 w_2 N_2 w_3$ avec $N_1, N_2 \in N'$, on peut trouver une dérivation de N_1 en $z_1 \in X^* Y^+$ de même de N_2 en $z_2 \in X^+ Y^*$. Comme la grammaire G est réduite, $A \overset{*}{\underset{G}{\to}} w_1 z_1 w_2 z_2 w_3$ $\overset{*}{\underset{G}{\to}} w \in (X \cup Y)^* Y^+ (X \cup Y)^* X^+ (X \cup Y)^*$ ce qui est en contradiction avec le fait que $L \subseteq X^* Y^*$, d'où le résultat.

□

Lemme II.34

Pour tout ISA-langage L, Alg // {L} $\subseteq$ Alg.

Démonstration

Soient S_1, $L \subseteq X^*$ tels que $S_1 \in$ Alg.

Définissons

- un nouvel alphabet $Y = \{y_x \ / \ x \in X\}$ avec $Y \cap X = \emptyset$
- la transduction rationnelle $\tau = g \circ \cap R \circ h^{-1}$ où

  . h est un homomorphisme alphabétique de $X \cup Y$ dans X défini par

$h(y_x) = x$ pour tout $y_x \in Y$

$h(x) = x$ pour tout $x \in X$

  . g est l'identité sur $X \cup Y$

  . $R = Y^* X^* Y^*$ un langage rationnel.

Il est clair que

$$S_1 \ // \ L = h[(\tau(S_1) \ // \ L) \cap Y^*]$$

Le problème se ramène à montrer que pour tout langage algébrique $L_1 \subseteq Y^* X^* Y^*$, tout langage $L \subseteq X^*$, avec $X \cap Y = \emptyset$, $(L_1 \ // \ L) \cap Y^*$ est algébrique.

L'égalité suivante se démontre facilement

$$(L_1 \ // \ L) \cap Y^* = s_1(\mathrm{conj}([(\mathrm{conj}(L_1 \ \#) \cap Y^* \ \# \ Y^* X^*).L^{-1}]) \cap Y^* \ \#)$$

où . $\# \notin X \cup Y$

  . s_1 est un homomorphisme alphabétique de $Y \cup \{\#\}$ dans Y défini par

  $s_1(\#) = \varepsilon$

  $s_1(y) = y$ pour tout $y \in Y$

Si on pose $L'_1 = \mathrm{conj}(L_1 \ \#) \cap Y^* \{\#\} Y^* X^*$, $L'_1 \in \mathrm{Alg}$, $L'_1 \subseteq Y^* \ \# \ Y^* X^*$ où $(Y \cup \{\#\}) \cap X = \emptyset$.

En appliquant le lemme précédent,

$$L'_1 = s(L'_o)$$

où . $L'_o \in \mathrm{Lin}$

  . $L'_o \subseteq Z^*$

  . s une substitution algébrique vérifiant

$s(z) \subseteq (Y \cup \{\#\})^*$ ou $s(z) \subseteq X^*$ pour tout $z \in Z$.

Par conséquent, il existe une grammaire linéaire réduite $G' = (Z, N', P', A)$ vérifiant

. $L'_o = L(G', A)$

. $P' \subseteq N' \times (T_Y^* (N' \cup \{\varepsilon\}) T_X^*)$

où $Z = T_X \cup T_Y$

$T_X = X \cup N_X$

$T_Y = Y \cup \{\#\} \cup N_{Y\cup\{\#\}}$

Construisons une nouvelle grammaire linéaire à gauche $G'' = (Z, N', P'', A)$ engendrant un rationnel R où

$$P'' = \{(B, Cxy^R)/(B, yCx) \in P', x \in T_X^*, y \in T_Y^*, (B, C) \in N'^2\}$$
$$\cup \{(B, xy^R)/(B, yx) \in P', x \in T_X^*, y \in T_Y^*, B \in N'\}$$

Ainsi $R \subseteq (X \cup Y)^* \{\#\} (X \cup Y)^*$ et montrons que $(L_1 \,//\, L) \cap Y^* = s_1(\mathrm{conj}[(s([\pi(R \cap [s^{-1}(L) \sqcup\!\sqcup T_Y^*])]^R))^R] \cap Y^* \{\#\})$ où π est la projection sur T_Y.

En effet, soit $w \in (L_1 \,//\, L) \cap Y^*$. Il existe une décomposition de w en $w = yz \in Y^*$, $x \in L$ tels que $y\,x\,z \in L_1$ et $z \# y\,x \in L'_1 = s(L'_o)$. Ainsi $z \# yx \in s(t_o)$ où $t_o = y_1 \ldots y_n\, z_n \ldots z_1 \in L'_o$. Par conséquent $s(y_1 \ldots y_n\, z_n \ldots z_1) \subseteq L'_1$ et d'après les conditions sur la substitution s, $x \in s(z_n \ldots z_1)$, $z \# y \in s(y_1 \ldots y_n)$ d'où $z_n \ldots z_1 \in s^{-1}(x) \subseteq s^{-1}(L)$.
Par construction de R, $\gamma_o = z_n\, y_n^R \ldots z_1\, y_1^R \in R$. Or $\gamma_o \in s^{-1}(L) \sqcup\!\sqcup T_Y^*$ et $y_1 \ldots y_n \in [\pi(R \cap [s^{-1}(L) \sqcup\!\sqcup T_Y^*])]^R$ d'où $z \# y \in s([\pi(R \cap [s^{-1}(L) \sqcup\!\sqcup T_y^*])]^R)$ et $y\,z \in s_1(\mathrm{conj}([\pi(R \cap [s^{-1}(L) \sqcup\!\sqcup T_Y^*\,)]^R) \cap Y^* \#)$.

Réciproquement, si $w \in s_1(\mathrm{conj}([\pi(R \cap [s^{-1}(L) \sqcup\!\sqcup T_Y^*])]^R \cap Y^* \#))$, alors il existe $w_1 \in [\pi(R \cap [s^{-1}(L) \sqcup\!\sqcup T_Y^*])]^R$ tel que $z \# y \in s(w_1)$, $w = yz$.
Par conséquent $w_1^R = \pi(w_2)$ où $w_2 \in R \cap [s^{-1}(L) \sqcup\!\sqcup T_Y^*]$
et $\begin{cases} w_2 = \beta_n\, \alpha_n^R \ldots \beta_1\, \alpha_1^R \\ w_1 = \alpha_1 \ldots \alpha_n \end{cases}$

Comme $L \subseteq X^*$, $s^{-1}(L) \subseteq Z^*$ et $\beta_n \dots \beta_1 \in s^{-1}(L)$.
D'où il existe $\ell \in L$ tel que $\beta_n \dots \beta_1 \in s^{-1}(\ell)$ ou $\ell \in s(\beta_n \dots \beta_1)$.
w_2 appartenant à R, $w_1\ \beta_n \dots \beta_1 \in L'_o$ et $z \# y\ \ell \in s(w_1\ \beta_n \dots \beta_1) \subseteq s(L'_o) = L'_1 = \mathrm{conj}(L_1\ \#) \cap Y^* \# Y^* X^*$. Ainsi $y\ \ell\ z \in L_1$ et $w = y\ z \in L_1\ //\ L$.

Comme L est un ISA-langage, $s^{-1}(L)$ est algébrique ainsi que $(L_1\ //\ L) \cap Y^*$ et $S_1\ //\ L$ d'où le résultat.

$\square$

Corollaire II.35

Pour tout ISA-famille L de langages, $\mathrm{Alg}\ //\ L \subseteq \mathrm{Alg}$, $\mathrm{Alg}\ //\ C(L) \subseteq \mathrm{Alg}$.

B. QUOTIENT ET LANGAGES ALGEBRIQUES BORNES

Le but de ce paragraphe est de montrer que le cône engendré par les langages algébriques bornés, noté C(AB), est fermé par quotient au centre. Pour cela nous commencerons par étudier un cas plus général, la fermeture de Alg par quotient au centre par les algébriques bornés et nous en tirerons les conséquences pour C(AB).

Afin de simplifier la résolution du problème, introduisons une nouvelle famille de langages ASB_k, celle des langages algébriques bornés sur k lettres toutes différentes. Nous noterons $ASB = \bigcup_{k=o}^{\infty} ASB_k$ où $ASB_o = AB_o = \{\{\varepsilon\}\}$. Ainsi.

<u>*Lemme II.36*</u>

Pour tout $k \geq 1$, *si* $Alg \,//\, AB_{k-1} \subseteq Alg$ *alors* $Alg \,//\, AB_k \subseteq Alg$

<u>Démonstration</u>

Elle se fait en deux étapes

a - <u>Première étape</u>

$Alg \,//\, AB_{k-1} \subseteq Alg \Rightarrow$ tout langage de ASB_k est un ISA-langage, par conséquent $Alg \,//\, ASB_k \subseteq Alg$ [Lemme II.34].

Soient $L_2 \in ASB_k$ i.e. $L_2 \subseteq a_1^* \ldots a_k^*$ et s une substitution algébrique de X dans $A = \{a_1, \ldots, a_k\}$.

On peut supposer sans nuire à la généralité du problème que pour tout $x \in X$, $s(x) \subseteq a_1^* \ldots a_k^*$.

Posons $Y_x = \{x_{ij} \;/\; 1 \leq i < j \leq k,\ x \in X\}$

$Y'_x = \{x_i \;/\; x \in X,\ i \in [1, k]\}$

$Y = \bigcup_{x \in X} (Y_x \cup Y'_x)$

On définit l'homomorphisme strictement alphabétique ϕ de Y dans X par $\phi(y) = x$ pour tout $y \in Y_x \cup Y'_x$ et la substitution algébrique s' de Y dans A par $s'(x_{ij}) = s(x) \cap a_i^+ \dots a_j^+$ pour tout $x_{ij} \in Y_x$, $s(x_i) = s(x) \cap a_i^*$ si $x_i \in Y'_x$. Il est clair que $s = s' \circ \phi^{-1}$, $s^{-1} = \phi \circ s'^{-1}$ et $s^{-1}(L) = (\phi \circ s'^{-1})(L)$ pour tout langage L.

Le problème se réduit à montrer la propriété suivante "Si s' est une substitution algébrique de Y dans A vérifiant $s'(y) \subseteq a_i^+ \dots a_j^+$ $(i < j)$ ou $s'(y) \subseteq a_i^*$ et L_2 un langage algébrique inclus dans $a_1^* \dots a_k^*$ alors $s'^{-1}(L_2) \in \text{Alg}$" qui se démontre par récurrence sur le cardinal h de $Y' = \{y \,/\, s'(y) \notin \text{Rat}\}$.

Remarquons que $s'(y) \notin \text{Rat}$ implique $i < j$.

Si $h = 0$, s' est une substitution rationnelle, s'^{-1} une transduction rationnelle et $s'^{-1}(L_2) \in \text{Alg}$.

Si $h \geq 1$, soit $y \in Y'$ tel que $s'(y) \notin \text{Rat}$ et $s'(y) \subseteq a_i^+ \dots a_j^+$ où $i < j$.

Définissons les substitutions algébriques

. s_y de $A \cup \{y\}$ dans A par

$s_y(y) = s'(y)$

$s_y(a_i) = a_i$ pour tout $i \in [1, k]$

. et s'_y de Y dans $A \cup \{y\}$ par

$s'_y(y) = y$

$s'_y(y') = s'(y')$ pour tout $y' \in Y \setminus \{y\}$.

Par construction $s' = s_y \circ s'_y$, $s'^{-1} = s'^{-1}_y \circ s_y^{-1}$ et $s'^{-1} = s'^{-1}_y[s_y^{-1}(L_2)]$.

A partir de la définition de s_y, l'égalité suivante se démontre facilement :

$s_y^{-1}(L_2) = L_2 \cup ([(L_2 \,//\, s'(y)) \sqcup\!\sqcup \{y\}] \cap a_1^* \dots a_i^* \, y \, a_j^* \dots a_k^*)$.

Envisageons les deux cas possibles :

* $i = 1$ et $j = k$, alors $L_2 \,//\, s'(y) \subseteq a_1^* a_k^*$.

$\Psi(L_2 \,//\, s'(y))$ étant semilinéaire, $L_2 \,//\, s'(y) \in \text{Alg}$.

$\star$ $i \geq 2$ ou $j < k$. Comme $s'(y) \subseteq a_i^+ \ldots a_j^+$ i.e. $s'(y) \in ASB_{k-1}$, et que $L_2 \in Alg$ alors $L_2 \,//\, s'(y) \in Alg \,//\, ASB_{k-1} \subseteq Alg$.

Par conséquent $s_y^{-1}(L_2) \in Alg$.

Comme s'_y est une substitution algébrique,

$s'^{-1}(L_2) = s'^{-1}_y(L_2) \cup s'^{-1}_y([(L_2 \,//\, s'(y)) \sqcup\!\sqcup \{y\}] \cap a_1^* \ldots a_i^* \, y \, a_j^* \ldots a_k^*)$,

considérons chacun des termes de l'union

α - Reprenons les conditions sur s'_y et L_2 :

$$L_2 \subseteq a_1^* \ldots a_k^*$$

s'_y est une substitution algébrique vérifiant

$$s'_y(y') \subseteq a_i^+ \ldots a_j^+ \quad (i < j)$$

ou $s'_y(y') \subseteq a_i^*$

et $||Y'_{s'_y}|| = ||Y'|| - 1$

En appliquant l'hypothèse de récurrence, $s'^{-1}_y(L_2) \in Alg$.

β - Posons $L_1 = ([L_2 \,//\, s'(y)] \sqcup\!\sqcup \{y\}) \cap a_1^* \ldots a_i^* \{y\} a_j^* \ldots a_k^*$,

où $L_1 \subseteq a_1^* \ldots a_i^* \, y^+ \, a_j^* \ldots a_k^*$ pour faciliter les écritures.

Deux cas sont possibles :

$\star$ $j > i+1$

Comme s'_y est une substitution algébrique vérifiant $s'_y(y') \subseteq a_\alpha^+ \ldots a_\beta^+$ $(\alpha < \beta)$

ou $s'_y(y') \subseteq a_\alpha^*$ $\qquad \forall\, y' \in Y' \setminus \{y\}$

ou $s'_y(y) \subseteq \{y\}$

et que $||Y'_{s'_y}|| = ||Y'|| - 1$, le résultat s'obtient facilement en appliquant l'hypothèse de récurrence.

$\star$ $j = i+1$

Considérons l'homomorphisme alphabétique f de $A \cup \{y\}$ dans A défini par $f(a_i) = a_i$ pour tout $i \in [1, n]$ et $f(y) = \varepsilon$, $L'_1 = f(L_1) \subseteq a_1^* \ldots a_i^* a_{i+1}^* \ldots a_k^*$.

Soit $Y_1 = \{y' \mid s(y') \subseteq a_\alpha^+ \ldots a_\beta^+,\ \alpha < \beta,\ \beta \leq i \text{ ou } \alpha \geq j\}$
$\cup \{y' \mid s(y') \subseteq a_i^*,\ i \in [1, n]\}$.

$Y_1 \subseteq Y \setminus \{y\}$ et $||Y'_1|| < ||Y'||$

Or $s'^{-1}_y(L_1)$ peut se mettre sous la forme

$s'^{-1}_y(L_1) = (s''^{-1}_y(L'_1) \sqcup\!\sqcup \{y\}) \cap Y^*_{1i}\ \{y\}\ Y^*_{jk}$

où $s''_y = s'_{y/Y_1}$

$$Y_{1i} = \bigcup_{y \in X} \{y_{\alpha\beta} \mid \beta \leq i,\ \alpha < \beta\} \cup \bigcup_{\beta \leq i} \{y_\beta\}$$

$$Y_{jk} = \bigcup_{y \in X} \{y_{\alpha\beta} \mid j \leq \alpha < \beta\} \cup \bigcup_{\beta \geq j} \{y_\beta\}$$

Par récurrence $s'^{-1}_y(L'_1) \in \text{Alg}$, d'où $s^{-1}_y(L_1) \in \text{Alg}$.

Par conséquent L_2 est un ISA-langage et Alg // $ASB_k \subseteq$ Alg.

b - <u>Seconde étape</u>

Pour tout $k \geq 1$, Alg // $ASB_k \subset$ Alg implique que Alg // $AB_k \subseteq$ Alg plus précisément pour tout $L_1 \in$ Alg, $L_2 \in AB_k$, L_1 // $L_2 \in$ Alg.

Comme $L_1 \in \text{Alg} \cap X^*$, $L_2 \in \text{Alg} \cap w_1^* \ldots w_k^*$, considérons k nouvelles lettres $z_1, \ldots, z_k$ et l'homomorphisme rationnel h défini par

$h(x) = x$ pour tout $x \in X$

$h(z_i) = w_i$ pour tout $z_i \in Z = \{z_1, \ldots, z_k\}$.

Si $L_1 // L_2 = (h^{-1}(L_1) // (h^{-1}(L_2) \cap z_1^* \ldots z_k^*)) \cap X^*$, $h^{-1}(L_1)$ est algébrique, $h^{-1}(L_2) \cap z_1^* \ldots z_k^* \in ASB_k$ et $h^{-1}(L_1) // (h^{-1}(L_2) \cap z_1^* \ldots z_k^*)$ est algébrique ainsi que $L_1 // L_2$.

En effet, prenons $xy \in L_1 // L_2$, alors $x\,u\,y \in L_1$ pour un certain $u \in L_2$. Ainsi il existe $u' \in h^{-1}(u) \cap z_1^* \ldots z_k^* \subseteq h^{-1}(L_2) \cap z_1^* \ldots z_k^*$ tel que $x\,u'\,y \in h^{-1}(L_1)$. D'où $xy \in x\,u'\,y // u' \subseteq (h^{-1}(L_1) // (h^{-1}(L_2) \cap z_1^* \ldots z_k^*)) \cap X^*$.

Réciproquement, si $w \in (h^{-1}(L_1) \mathbin{//} (h^{-1}(L_2) \cap z_1^* \dots z_k^*)) \cap X^*$, alors il existe $w_1, w_2 \in X^*$, $w' \in h^{-1}(L_2) \cap z_1^* \dots z_k^*$ tels que $w_1\ w'\ w_2 \in h^{-1}(L_1)$. Comme $h(w_1\ w'\ w_2) \in L_1$, $w_1\ h(w')\ w_2 \in L_1$ avec $h(w') \in L_2$, ce qui implique que $w = w_1\ w_2 \in L_1 \mathbin{//} L_2$.

□

Comme $ASB_o = AB_o = \{\{\varepsilon\}\}$ et que $Alg \mathbin{//} AB_o \subseteq Alg$, nous pouvons en déduire que ASB ainsi que C(AB) = C(ASB) sont des ISA-familles de langages. Donc

Corollaire II.37

C(AB) *est une ISA-famille de langages.*

Proposition II.38

Alg // AB = Alg.

Démonstration

Du lemme II.36, $Alg \mathbin{//} AB \subseteq Alg$. Comme tout langage $\{\varepsilon\} \in AB$, pour tout $L \in Alg$, $L = L \mathbin{//} \{\varepsilon\}$ et $L \in Alg \mathbin{//} AB$.

□

De même, la propriété suivante devient évidente :

Corollaire II.39

AB // AB = AB.

Compte tenu du fait que C(ASB) = C(AB) et des corollaires II.37 et II.35, nous pouvons étendre ces derniers résultats en :

Corollaire II.40

Alg // C(AB) = Alg

AB // C(AB) = AB.

Le lemme suivant facilitera l'étude de la fermeture par quotient au centre de $C(AB)$. En effet

Lemme II.41

Pour tout langage algébrique L *borné sur* n *lettres différentes, tout cône rationnel* $\mathcal{L}$ *clos par union,* $C(L) \mathbin{//} \mathcal{L} = C_{\cup}(\{L\} \mathbin{//} \mathcal{L})$.

Démonstration

Soient $L \subseteq a_1^* \ldots a_n^*$ où les a_i sont des lettres toutes différentes, $L_1 \in C(L)$. Il existe deux homomorphismes alphabétiques h, g et un langage rationnel R tel que $L_1 = g[h^{-1}(L) \cap R]$. Il est facile de vérifier l'égalité suivante

$$L_1 \mathbin{//} L_2 = g[(h^{-1}[L] \cap R) \mathbin{//} g^{-1}(L_2)] \quad \text{où } L_2 \in \mathcal{L}.$$

Notons $A_i = h^{-1}(\{a_i\})$ pour tout $i \in [1, n]$

$$A_\varepsilon = h^{-1}(\varepsilon)$$

$$A = \bigcup_{i \in [1,n]} A_i \cup A_\varepsilon.$$

Il est évident que $h^{-1}(L) \subseteq [A_1^* \ldots A_n^*] \sqcup\!\sqcup A_\varepsilon^*$.

On pourra prendre $R \subseteq [A_1^* \ldots A_n^*] \sqcup\!\sqcup A_\varepsilon^*$ sans nuire à la généralité du problème.

Pour montrer l'inclusion de départ, commençons par le lemme suivant où l'on considère un automate fini déterministe $A = (X, Q, q_0, *, F)$ reconnaissant R avec pour tout $q_1, q_2 \in Q$, $R_{q_1q_2} = \{w \in X^* \mid q_1 * w = q_2\}$ et $R_{q_1F} = \bigcup_{q \in F} R_{q_1q}$

Lemme II.42

Pour tout langage algébrique $L \subseteq X^*$ *borné sur* n *lettres différentes, tout homomorphisme alphabétique* h, *tout langage* L_2, *nous avons l'égalité suivante*

$$[h^{-1}(L) \cap R] \mathbin{//} L_2 = \bigcup_{(q_1,q_2) \in Q^2} (h^{-1}(L) \mathbin{//} [L_2 \cap R_{q_1q_2}]) \cap R_{q_0q_1} R_{q_2F}.$$

Démonstration

Soit $u \in [h^{-1}(L) \cap R] // L_2$, il existe $\ell \in L_2$, une décomposition $u = u_1 u_2$ tels que $u_1 \ell u_2 \in h^{-1}(L) \cap R$ i.e. que $q_o * u_1 = q_1$, $q_1 * \ell = q_2$, $q_2 * u_2 = q$, $q \in F$ pour $q_1, q_2 \in Q$.

Par conséquent $u_1 u_2 \in R_{q_oq_1} R_{q_2F}$ et $u = u_1 u_2 \in u_1 \ell u_2 // \ell \subseteq h^{-1}(L) // [L_2 \cap R_{q_1q_2}]$.

Réciproquement, soit $u \in \bigcup_{(q_1, q_2) \in Q^2} (h^{-1}(L) // [L_2 \cap R_{q_1q_2}]) \cap R_{q_oq_1} R_{q_2F}$, $u \in (h^{-1}(L) // [L_2 \cap R_{q_1q_2}]) \cap R_{q_oq_1} R_{q_2F}$ pour un certain $(q_1, q_2) \in Q^2$.

D'une part, $u = u_1 u_2$ avec $u_1 \in R_{q_oq_1}$, $u_2 \in R_{q_2F}$, d'autre part $u = w_1 w_2$ où $w_1 s w_2 \in h^{-1}(L)$ et $s \in R_{q_1q_2} \cap L_2$.

Il est clair que $u_1 s u_2 \in R$.

Considérons le cas où w_1 est facteur gauche de u_1 i.e. qu'il existe t tel que $u_1 = w_1 t$ et $w_2 = t u_2$, le cas où w_2 est facteur droit de u_2 se traite de façon analogue.

Envisageons les divers cas possibles :

Premier cas :

$t \in A_\varepsilon^*$, $w_1 s w_2 \in h^{-1}(L)$ implique que

$$h(w_1 s w_2) = h(w_1) h(s) h(t) h(u_2)$$
$$= h(w_1) h(s) h(u_2) \in L$$

Considérons $u_1 s u_2$,

$$h(u_1 s u_2) = h(w_1) h(t) h(s) h(u_2)$$
$$= h(w_1) h(s) h(u_2) \in L.$$

Par conséquent, $u_1 s u_2 \in h^{-1}(L) \cap R$ et $u \in (h^{-1}(L) \cap R) // L_2$.

Second cas :

Il existe $i \in [1, n]$ tel que $t \in A_\varepsilon^* A_i(A_i \cup A_\varepsilon)^*$. On démontre facilement que $s \in (A_i \cup A_\varepsilon)^*$.

En effet, supposons qu'il existe $\ell \in [1, i[$ tel que $s \in A^* A_\ell A^*$, alors $u_1 s u_2 = w_1 t s u_2 \in A^* A_i A^* A_\ell A^*$ avec $\ell < i$ ce qui est contraire au fait que $R \subseteq (A_1^* \ldots A_n^*) \sqcup\!\sqcup A_\varepsilon^*$.

S'il existe $\ell \in]i, n]$ vérifiant $s \in A^* A_\ell A^*$ alors $w_1 st u_2 \in A^* A_\ell A^* A_i A^*$ où $\ell > i$ ce qui est incompatible avec le fait que $h^{-1}(L) \subseteq (A_1^* \ldots A_n^*) \sqcup\!\sqcup A_\varepsilon^*$.

Par conséquent $st \cup (A_i \cup A_\varepsilon)^* A_i (A_i \cup A_\varepsilon)^*$ et $h(w_1)\, h(st)\, h(u_2) = \ell_1\, \ell_2\, \ell_3$.

Comme $\ell_2 \in a_i^+$, il existe $p \in \mathbb{N}^+$ tel que $h(st) = a_i^p = \ell_2 = com(a_i^p) = com(h(st))$. h étant un homomorphisme alphabétique, h et com commutent et $h(ts) = a_i^p = \ell_2$.D'où $h(w_1\ ts\ u_2) = \ell_1\, \ell_2\, \ell_3 \in L$.

Ainsi $u_1\, s\, u_2 \in h^{-1}(L) \cap R$ et $u = u_1 u_2 \in (h^{-1}(L) \cap R) \,//\, L_2$.

Troisième cas :

Il existe $i < j$ tel que $t \in A^* A_i A^* A_j A^*$. Un raisonnement analogue au précédent montre que $s \in A_\varepsilon^*$ et il est facile de démontrer que $u_1 u_2 \in (h^{-1}(L) \cap R) \,//\, L_2$.

$\square$

Démonstration du lemme II.41 (fin)

Ainsi $L_1 \,//\, L_2 = \bigcup_{(q_1,q_2)\in Q^2} g[(h^{-1}(L) \,//\, [R_{q_1q_2} \cap g^{-1}(L_2)]) \cap R_{q_oq_1}\, R_{q_2F}]$.

Comme h est un homomorphisme alphabétique, l'égalité suivante se démontre facilement $L_1 \,//\, L_2 = \bigcup_{(q_1,q_2)\in Q^2} g[h^{-1}(L \,//\, h(R_{q_1q_2} \cap g^{-1}(L_2))) \cap R_{q_oq_1} R_{q_2F}]$

Ainsi L_1 // $L_2 \in C_U(\{L\}$ // $L)$. D'où l'égalité à partir du corollaire II.6.

□

La proposition suivante s'en déduit facilement

Proposition II.43

$C(AB)$ *est fermé par quotient i.e.* $C(AB)$ // $C(AB) = C(AB)$.

Démonstration

Comme $C(ASB) = C(AB)$ il suffit de montrer que $C(ASB)$ // $C(AB) \subseteq C(AB)$.

Soient $L_1 \in C(ASB)$, $L_2 \in C(AB)$ alors il existe $L \in ASB$ tel que $L_1 \in C(L)$. Ainsi L_1 // $L_2 \in C(L)$ // $C(AB) \subseteq C_U(\{L\}$ // $C(AB)) = C_U(AB) = C(AB)$.

Réciproquement, comme tout langage $\{\varepsilon\} \in C\ (AB)$ pour tout $L \in C(AB)$, $L = L$ // $\{\varepsilon\} \in C(AB)$ // $C(AB)$.

□

C. QUOTIENT ET LANGAGES ALGEBRIQUES COMMUTATIFS

Commençons d'abord par montrer que $C(AC)$ est fermé par quotient à droite (à gauche). Pour cela établissons des résultats plus généraux :

Définition II.44

Un langage L *est semilinéaire si* $com(L) = com(R)$ *pour un certain langage rationnel* R. *Un cône rationnel contenant uniquement des langages semilinéaires est dit un cône rationnel semilinéaire.*

Lemme II.45

Soient L' *un langage commutatif,* $\mathcal{L}$ *un cône rationnel semilinéaire. Alors* $\mathcal{L}^{-1}.C(L') = C(L')$.

Démonstration

De la proposition II.7, $C(L') \subseteq \mathcal{L}^{-1}\, C(L') \subseteq C_U(\mathcal{L}^{-1}\, L')$. Soit $L \in \mathcal{L}$. Comme L est un langage semilinéaire, il existe un langage rationnel R ayant même image commutative que L. L' étant commutatif, $L^{-1}\, L' = R^{-1}\, L' \in Rat^{-1}\, C(L') = C(L')$ et $C_U(\mathcal{L}^{-1}\, L') \subseteq C_U(L') = C(L')$, d'où le résultat.

En particulier, si L' est un langage commutatif semilinéaire, $C(L')$ est un cône rationnel semilinéaire [20] et nous obtenons :

Corollaire II.46

Pour tout langage commutatif semilinéaire L, $C(L)$ *est clos par quotient à gauche (à droite).*

Comme tout langage algébrique est semilinéaire, il s'en suit

Corollaire II.47

Pour tout langage algébrique commutatif L, $C(L)$ *est clos par quotient à gauche (à droite).*

Un exemple illustratif sera le langage $D_1^* = \{w \in \{a, b\}^* \ / \ |w|_a = |w|_b\}$ commutatif algébrique.

Corollaire II.48

$C(D_1^*)$ *est fermé par quotient à gauche (à droite).*

C(AC) étant un cône rationnel semilinéaire, la fermeture par quotient à gauche (à droite) de C(AC) se démontre facilement.

Lemme II.49

$[C(AC)]^{-1}\ C(AC) = C(AC)$.

Démonstration

Considérons $L_1\ L_2 \in C(AC)$, il existe $A_1 \in AC$ tel que $L_1 = \tau(A_1)$ où τ est une transduction rationnelle. $L_2^{-1}.L_1 = L_2^{-1}\ \tau(A_1) \in (C(AC))^{-1}\ C(A_1) = C_U(C(AC)^{-1}\ A_1) \subseteq C_U(C(A_1)) = C(A_1) \subseteq C(AC)$.

Peut-on étendre ce résultat au quotient au au centre ? Pour résoudre ce problème, montrons d'abord la propriété suivante

Lemme II.50

Pour tout langage commutatif L, *toute substitution algébrique* s, $s^{-1}(L) \in C(L)$.

Démonstration

s étant une substitution algébrique, $s(x) \in$ Alg pour tout $x \in X$, par conséquent pour tout $x \in X$, il existe $R_x \in$ Rat tel que $\Psi(R_x) = \Psi(s(x))$.

En posant $t(x) = R_x$, $\forall\ x \in X$, t est une substitution rationnelle et $s^{-1}(L) = t^{-1}(L)$.

En effet, soit $w \in s^{-1}(L)$, il existe $\ell \in L$ tel que $\ell \in s(w)$. Nous pouvons avoir une décomposition de ℓ en $\ell = \ell_1 \ldots \ell_n$ telle que $w = w_1 \ldots w_n$ où $w_i \in X$ pour tout $i \in [1, n]$ et $\ell_i \in s(w_i)$.

Considérons les $\gamma_i \in \mathrm{com}(\ell_i) \cap R_{w_i}$, alors $\gamma_1 \ldots \gamma_n \in L$ car L est commutatif.
De même $\gamma_1 \ldots \gamma_n \in t(w_1 \ldots w_n)$ et $w \in t^{-1}(\gamma_1 \ldots \gamma_n) \subseteq t^{-1}(L)$.

Réciproquement, soit $w \in t^{-1}(L)$. Il existe $\ell \in L$ tel que $\ell \in t(w)$.
$w = w_1 \ldots w_n$ où $w_i \in X$ par conséquent ℓ peut s'écrire $\ell = \ell_1 \ldots \ell_n$ où $\ell_i \in t(w_i)$ pour tout $i \in [1, n]$.
Comme $\Psi(R_{w_i}) = \Psi(s(w_i))$, pour tout $\ell_i \in R_{w_i}$, il existe $\gamma_i \in s(w_i)$ tel que $\gamma_i \in \mathrm{com}(\ell_i)$. ℓ appartenant à $R_{w_1} \ldots R_{w_n}$, il existe $\gamma \in \mathrm{com}(\ell_1) \ldots \mathrm{com}(\ell_n)$ tel que $\gamma \in s(w_1 \ldots w_n)$. Comme L est commutatif, $\gamma \in s(w_1 \ldots w_n) \cap L$ et $w = w_1 \ldots w_n \in s^{-1}(L)$.

□

Nous pouvons en tirer les conséquences suivantes :

Corollaire II.51

AC *est une ISA-famille de langages.*

Corollaire II.52

$$\mathrm{Alg} \ // \ \mathrm{AC} = \mathrm{Alg}$$

$$\mathrm{AC} \ // \ \mathrm{AC} = \mathrm{AC}$$

$$\mathrm{AC} \ // \ C(\mathrm{AC}) = \mathrm{AC}.$$

Etablissons une propriété vérifiée par les langages commutatifs algébriques qui permet d'obtenir la fermeture par quotient au centre de C(AC).

Lemme II.53

Pour tout langage commutatif algébrique L, *tout cône rationnel* $\mathcal{L}$ *fermé par union, nous avons*

$$C(L) \ // \ \mathcal{L} = C_U(\{L\} \ // \ \mathcal{L}).$$

Démonstration

Soient $L_1 \in C(L)$, $L_2 \in L$ i.e. $L_1 = g[h^{-1}(L) \cap R]$ où g et h sont des homomorphismes alphabétiques, R un rationnel. L'égalité suivante se démontre facilement

$$L_1 \mathbin{/\!/} L_2 = g[(h^{-1}(L) \cap R) \mathbin{/\!/} g^{-1}(L_2)].$$

Comme AC est fermé par homomorphisme inverse, $h^{-1}(L) \in AC$ et le problème revient à montrer que $(h^{-1}(L) \cap R) \mathbin{/\!/} g^{-1}(L_2) \in C_U(\{L\} \mathbin{/\!/} L)$. Pour cela, considérons le lemme suivant

Lemme II.54

Pour tout automate fini déterministe minimal $A = (X, Q, q_o, *, F)$ *reconnaissant* R, *tout langage algébrique commutatif* L_1, *tout langage* S_2, *il existe un homomorphisme alphabétique* h_1 *vérifiant*

$$(L_1 \cap R) \mathbin{/\!/} S_2 = \bigcup_{(q_1,q_2)\in Q^2} h_1[(h_1^{-1}(L_1) \mathbin{/\!/} [h_1^{-1}(S_2) \cap R_{q_1q_2}] \cap R_{q_oq_1} \; R_{q_2F}]$$

où

- $R' = \{(q_o, x_1, q_1)(q_1, x_2, q_2) \ldots (q_{n-1}, x_n, q_n) \;/\; q_n \in F\}$
- *pour tous* $q_1, q_2 \in Q$

  $R_{q_1q_2} = F(R') \cap [(\{q_1\}\times X\times Q)(Q\times X\times Q)^*(Q\times X\times\{q_2\}) \cup (\{q_1\}\times X\times\{q_2\})]$
- *et* $R_{q_2F} = \bigcup_{q\in F} R_{q_2q}$.

Démonstration

Il est évident que

R', $R_{q_1q_2}$, R_{q_2F} sont rationnels pour tous $q_1, q_2 \in Q$.

On définit l'homomorphisme alphabétique h_1 de $\Delta = Q \times X \times Q$ dans X par

$h_1((q_1, x, q_2)) = x$ pour tout $(q_1, x, q_2) \in \Delta$.

Considérons $u \in (L_1 \cap R) \mathbin{/\!/} S_2$, $u = u_1 u_2$ où $u_1 \; s \; u_2 \in L_1 \cap R$ et $s \in S_2$.

Comme $u_1 \, s \, u_2 \in R$, il existe des états q_1, q_2, q tels que $q_o * u_1 = q_1$, $q_1 * s = q_2$ et $q_2 * u_2 = q \in F$. Par conséquent $s = h_1(s')$ avec $s' \in R_{q_1q_2}$. s' appartenant à $h_1^{-1}(s)$ et $s \in L_2$ implique que $s' \in R_{q_1q_2} \cap h_1^{-1}(S_2)$. De même il existe $u'_1 \in R_{q_oq_1}$ et $u'_2 \in R_{q_2F}$ vérifiant $h_1(u'_1) = u_1$ et $h_1(u'_2) = u_2$, ainsi $u'_1 \, u'_2 \in R_{q_oq_1} \, R_{q_2F}$ et $u'_1 \, u'_2 \in h^{-1}(u_1 \, u_2) \subseteq h_1^{-1}(u_1 \, s \, u_2) \, // \, s' \subseteq h_1^{-1}(L_1) \, // \, (R_{q_1q_2} \cap h_1^{-1}(S_2))$. Et $u_1u_2 = h_1(u'_1u'_2) \in h_1((h_1^{-1}(L_1) \, // \, [R_{q_1q_2} \cap h_1^{-1}(S_2)]) \cap R_{q_oq_1} \, R_{q_2F})$.

Réciproquement, si $u \in h_1((h_1^{-1}(L_1) \, // \, [R_{q_1q_2} \cap h_1^{-1}(S_2)]) \cap R_{q_oq_1} R_{q_2F})$ pour un certain $(q_1, q_2) \in Q^2$, alors $u = h_1(v)$ où $v \in R_{q_oq_1} \, R_{q_2F}$ et $v \in h_1^{-1}(L_1) \, // \, [R_{q_1q_2} \cap h_1^{-1}(S_2)]$ autrement dit $v = v'_1 \, v'_2 \in R_{q_oq_1} \, R_{q_2F}$ et $v = v_1 \, v_2$ où $v_1 \, t \, v_2 \in h_1^{-1}(L_1)$ et $t \in R_{q_1q_2} \cap h_1^{-1}(S_2)$.

Comme $t \in R_{q_1q_2}$, $v'_1 \, s \, v'_2 \in R'$. $h_1^{-1}(L_1)$ appartenant à AC, $v'_1 \, s \, v'_2 \in \mathrm{com}(v_1 \, t \, v_2) \subseteq h_1^{-1}(L_1)$.

Ainsi $u \in h_1(v'_1 \, t \, v'_2) \, // h_1(t) \subseteq (L_1 \cap R) \, // \, L_2$.

□

<u>Démonstration du lemme 53</u> (fin)

A partir des propriétés des homomorphismes alphabétiques et des homomorphismes alphabétiques inverses, l'égalité suivante s'obtient facilement.

$$L_1 \, // \, L_2 = \bigcup_{(q_1,q_2)\in Q^2} gh_1((hh_1)^{-1}(L//(hh_1)^{-1}(hh_1)[(gh_1)^{-1}(L_2)\cap R_{q_1q_2}]\cap R_{q_oq_1} R_{q_2F})$$

et $L_1 \, // \, L_2 \in C_U(\{L\} \, // \, L)$.

L'inclusion réciproque a été établie au corollaire II.6.

□

La fermeture par quotient s'en déduit immédiatement.

<u>Lemme II.55</u>

C(AC) // C(AC) = C(AC).

Démonstration

Soient L_1, $L_2 \in C(AC)$, il existe $L \in AC$ tel que L_1 // $L_2 \in C(L)$ // $C(AC) =$ $C_\cup(\{L\}$ // $C(AC)) \subseteq C_\cup(AC)) = C(AC)$.

L'inclusion inverse est évidente.

$\square$

BIBLIOGRAPHIE

[1] J.M. AUTEBERT
Non-principalité du cylindre des langages à compteur.
Math. Syst. Theory 11, 1977, 157-167.

[2] B.S. BAKER et R.V. BOOK
Reversal-bounded Multipushdown Machines
J. Comp. Syst. Sc. 8, 1974, 315-332.

[3] J. BEAUQUIER
Substitution de langages linéaires et à compteur.
Publication LITP n° 78-5.

[4] J. BEAUQUIER
Independence of linear and one-counter generators.
Fundamentals of Computation Theory, Akademie-Verlag.
Berlin, 1979, 45-51.

[5] J. BEAUQUIER
A remark about a substitution property.
Math. Syst. Theory 14, 1981, 189-191.

[6] J. BERSTEL et L. BOASSON
Une suite décroissante de cônes rationnels.
Automata, Languages and Programming, 2nd Colloquium Saarbrücken,
Springer Verlag, 1974, 383-397.

[7] J. BERSTEL
Transductions and Context-Free Languages.
Teubner Verlag Stuttgart, 1979.

[8] L. BOASSON, M. NIVAT, J.P. CRESTIN
Familles de langages translatables et fermées par crochet.
Acta Informatica 2, 1973, 383-393.

[9] L. BOASSON, M. NIVAT
Sur diverses familles de langages fermées par transduction rationnelle.
Acta Informatica 2, 1973, 180-188.

[10] M. CHOMSKY
On certain formal properties of grammars.
Information and Control 2, 1959, 137-167.

[11] J.P. CRESTIN
Langages quasirationnels.
Langages algébriques, J.P. Crestin et M. Nivat, eds, 1973, 123-166.

[12] S. EILENBERG
Automata languages and machines.
Academic Press, Vol. A, New-York, 1974.

[13] C.C. ELGOT et J.F. MEZEI
On relations defined by generalized finite automata.
IBM. J. Res. Dev. 9, 1965, 47-68.

[14] S. GINSBURG et E.H. SPANIER
Bounded-Algol-Like Languages.
Transductions of the American Math. Soc. 113, 1964, 333-368.

[15] S. GINSBURG et S.A. GREIBACH
Studies in Abstract Families of languages.
Memoirs of the American Math. Soc. 113, 1966, 285-396.

[16] S. GINSBURG et E.H. SPANIER
Bounded Regular Sets
Proceedings of the American Math. Soc. 17, 1966, 1043-1049.

[17] S. GINSBURG et J.S. ULLIAN
Ambiguity in context-free languages.
J. Assoc. Comput. Mach. 13, 1966.

[18] S. GINSBURG et E.H. SPANIER
Derivation bounded languages.
J. Comp. Syst. Sc. 2, 1968, 228-250.

[19] S. GINSBURG et S.A. GREIBACH
Principal AFL
J. Comp. Syst. Sc. 4, 1970, 303-338.

[20] S. GINSBURG
Algebraic and Automata-Theoretic Properties of Formal Languages.
North-Holland Publishing Company, Amsterdam, 1975.

[21] S.A. GREIBACH
Chains of full AFL's.
Math. Syst. Theory 4, 1970, 231-242.

[22] S.A. GREIBACH
Syntactic Operators on Full semi-AFLs.
J. Comp. Syst. Sc. 6, 1972, 30-76.

[23] S.A. GREIBACH
One counter languages and the IRS condition.
J. Comp. Syst. Sc. 10, 1975, 237-247.

[24] S.A. GREIBACH
A note on the recognition of one-counter languages.
RAIRO Inf. Theor. R2, 1975, 5-12.

[25] S.A. GREIBACH
Remarks on blind and partially blind one-way multi-counter machines
Theoretical Comp. Sc. 7, 1978, 311-324.

[26] S.A. GREIBACH
One counter languages and the chevron operation.
RAIRO Inf. Theor. 13, 1979, 189-194.

[27] G. JACOB
Représentation et Substitutions Matricielles dans la Théorie
Algébrique des Transductions.
Thèse Sc. Math. Université Paris VII, Paris, 1975.

[28] M. JANTZEN
On the hierarchy of Petri Net languages.
RAIRO Inf. Theor. 13, 1979, 19-30.

[29] M. JANTZEN
The power of synchronizing operations on strings.
Theoretical Comp. Sc. 14, 1981, 127-154.

[30] M. LATTEUX
Edtol-systèmes ultralinéaires et opérateurs associés.
Publication du laboratoire de calcul de l'Université de Lille I, n° 100, 1977.

[31] M. LATTEUX
Cônes rationnels commutatifs.
J. Comp. Syst. Sc. 18, 1979, 307-333.

[32] M. LATTEUX
Substitutions dans les Edtol-systèmes ultralinéaires.
Information and Control 42, 1979, 194-260.

[33] M. LATTEUX
A propos du lemme de substitution.
Theoretical Comp. Sc. 14, 1981, 119-123.

[34] M. LATTEUX
Langages à un compteur.
J. Comp. Syst. Sc. 26, 1983, 14-33.

[35] M. LATTEUX et G. ROZENBERG
Commutative one-counter languages are regular.
J. Comp. Syst. Sc. 29, 1984, 54-57.

[36] J. LEGUY
Rational cone and substitution.
Fundamentals of Computation Theory, FCT' 81.
Lecture Notes in Computer Science 117, 1981, 234-243.

[37] J. LEGUY
Deux nouvelles extensions du lemme de substitution.
Colloque / Symposium Paris
Les Mathématiques de l'Informatique AFCET, 1982, 199-204.

[38] M. NIVAT
Transductions des langages de Chomsky.
Ann. Inst. Fourier <u>18</u>, Grenoble, 1967, 339-455.

[39] M. NIVAT
Opérateurs sur les familles de langages.
Rapport de recherche IRIA, n° 106, 1975.

[40] F. RODRIGUEZ
Familles de langages fermés par crochet et crochet ouvert.
Lectures Notes in Computer Science, Springer Verlag Heidelberg
<u>48</u>, 1977, 154-168.

[41] M.K. YNTEMA
Inclusion relations among families of context free languages.
Information and Control <u>10</u>, 1967, 572-597.

RESUME

Dans la première partie de cette thèse, nous introduisons la notion d'ordre d'un langage quasirationnel qui affine celle de rang dû à Yntema. L'intérêt de cette notion est, alors, mis en évidence par la preuve que pour certaines opérations telles que produit, étoile, chevron et substitution, l'ordre du langage obtenu dépend uniquement des ordres des langages de départ.

Dans la deuxième partie, nous étudions l'opération de quotient en liaison avec les principaux cônes rationnels algébriques. Nous montrons d'abord que tout langage récursivement énumérable peut être obtenu comme quotient de deux langages linéaires ce qui atteste de la puissance de cette opération. A l'opposé, nous démontrons que la famille des langages à un compteur est fermée par quotient algébrique. Enfin, grâce à l'introduction des ISA-langages, nous établissons que la famille des langages algébriques est fermée par quotient par un langage algébrique borné ou commutatif.