Table des matières

1 Notions fondamentales **19**

1 Notations et généralités 19
2 Graphes orientés 20
3 Structures algébriques 21
3.1 Monoïdes 21
3.2 Semi-anneaux 25
3.3 Polynômes et séries formelles 28
3.4 Ensembles et séries rationnels 30
3.5 Ensembles reconnaissables 31
4 Automates 31
4.1 Automates sur un semi-anneau 31
4.2 Automates sur un monoïde 33
4.3 Automates sur un alphabet 35
4.4 Automates à multiplicité 41

2 Automate universel **45**

1 Définitions et propriétés de l'automate universel 46
1.1 Automate universel dans un monoïde quelconque 46
1.2 Automate universel d'un ensemble reconnaissable 51
1.3 Automate universel et générateurs du monoïde 52
1.4 Automate universel d'un langage rationnel de A^* 54
2 Calcul effectif de l'automate universel 56
3 Écorché de l'automate universel 59
4 Développement d'un automate 64
4.1 Motivations et définitions 64
4.2 Propriétés de l'automate développé 68
4.3 Écorché du développé 70

3 Automates universels et langages réversibles **75**

1 Langages réversibles 76
2 Langages à groupe 80
3 Automate universel d'un langage à groupe 80
3.1 Structure générale de l'automate universel d'un langage à groupe . . 80
3.2 Structure des composantes fortement connexes de l'automate universel 82
4 Automate universel d'un langage réversible 86
4.1 Structure générale de l'automate universel d'un langage réversible . 86
4.2 Structure des pelotes de l'automate universel 88
5 Construction d'un automate réversible 90
5.1 Cordes 90
5.2 Automate quasi-réversible et automate universel 91

4 Hauteur d'étoile **97**

1 Hauteur d'étoile et degré d'enlacement 98
1.1 Hauteur d'étoile d'un langage rationnel 98
1.2 Enlacement d'un graphe orienté 99
1.3 Enlacement et hauteur d'étoile 104
1.4 Du calcul d'une expression au théorème d'Eggan 106
2 Hauteur d'étoile des langages à groupe 110
3 Hauteur d'étoile des langages réversibles 114
4 Automate universel et hauteur d'étoile 120

5 Déterminisation des automates (max,+) **123**

1 Le semi-anneau tropical et sa famille 124
2 Caractérisation des séries séquentielles 125
2.1 Séries translatées 125
2.2 Le problème d'une caractérisation topologique 128
3 Décidabilité de la séquentialité dans le cas des alphabets unaires 130
4 Algorithmes 135
4.1 Décidabilité 135
4.2 Déterminisation 136
5 Non-ambiguïté des séries rationnelles sur un alphabet à une lettre 140
6 Automates univoques 143
7 Le cas général 147
8 Problème de la généralisation à d'autres semi-anneaux 148

6 Dérivation d'expressions rationnelles avec multiplicité 149

1 Expressions rationnelles 150
2 Motivation de la dérivation 154
3 Dérivation et termes dérivés 156
4 L'automate des termes dérivés 165
5 Les termes dérivés fantômes 167
6 Variations 169
7 Le cas commutatif 172

— ∘ —

Chapitre 3

Automates universels et langages réversibles

Les langages réversibles forment une famille de langages rationnels qui se trouvent à la croisée de différents domaines.

Une restriction de cette famille est celle des langages réalisés par des automates réversibles déterministes. Les premières extensions des résultats sur la hauteur d'étoile aux langages réversibles se restreignent à cette famille [17, 31]. Nous aurons l'occasion d'y revenir. Dans un autre domaine, D. Angluin [2] classe ces langages selon le nombre d'états terminaux de l'automate minimal qui les reconnaît et montre qu'un langage d'une classe donnée peut être inféré par un panel de mots du langage. Les automates *inverses* sont eux aussi des automates réversibles « déterministes » sur un alphabet $A \cup \overline{A}$, où, à chaque transition (p,a,q) correspond une transition $(q,\overline{a},p)$. Ces automates se révèlent particulièrement utiles pour l'étude des semi-groupes inversifs [62, 57]. T.E. Hall [29] a établi que les langages clos par multiplication et reconnus par des automates réversibles déterministes sont des étoiles de *codes bipréfixes*. C'est encore à des automates réversibles déterministes que les automates quantiques sur un alphabet unaire ont été comparés [1]. Il a été montré que si l'automate quantique doit donner la bonne réponse avec une probabilité forte, il peut être simulé par un automate réversible.

La famille des langages reconnus par des automates réversibles déterministes n'est cependant pas stable par union. Les langages réversibles eux, forment une *variété positive* de langages. Une caractérisation algébrique en a été donnée par J.-E. Pin [51], ce qui permet de décider de la réversibilité d'un langage réversible. P.-C. Héam [36] a montré que ces langages jouent un rôle central dans l'étude topologique des langages rationnels.

Toutefois, ces résultats ne donnent pas d'algorithme raisonnable pour construire un automate réversible qui accepte un langage réversible donné. C'est ce que nous nous proposons de faire, à travers le résultat suivant :

Théorème *L'automate universel d'un langage réversible contient un automate quasi-réversible équivalent.*

Pour ce faire, il nous faut évidemment introduire les automates *quasi-réversibles* dont

on montre qu'ils sont équivalents aux automates réversibles, dans le sens où, non seulement ils reconnaissent les mêmes langages, mais aussi parce qu'on peut facilement les transformer en automates réversibles.

Avant de se lancer dans l'étude des langages réversibles, nous étudierons d'abord une sous-classe de langages, les langages à groupe. On montre que les automates universels des langages à groupe se décomposent en « étages » et que leurs composantes fortement connexes sont elles-mêmes des automates à groupe. Cette première étude nous permet d'introduire des notions qui apparaîtront plus complexes lorsqu'elles seront appliquées aux langages réversibles auxquels est consacrée la dernière partie de ce chapitre.

Nous y montrons, dans un premier temps, que les pelotes de l'automate universel d'un tel langage sont réversibles, puis que l'automate universel d'un langage réversible contient un sous-automate quasi-réversible qui accepte le langage. Nous donnons un algorithme qui permet d'obtenir un tel automate.

1 Langages réversibles

Définition 3.1 *Un automate $\mathcal{A}$ sur un alphabet A est un* ***automate réversible*** *si pour tout état q de $\mathcal{A}$, pour toute lettre a de A, il existe au plus une transition étiquetée par a qui part de q et au plus une qui y arrive. Un langage $\mathcal{L}$ est un* ***langage réversible*** *s'il existe un automate réversible qui reconnaît $\mathcal{L}$.*

Malheureusement, si un automate réversible n'est pas déterministe, c'est-à-dire s'il a plusieurs états initiaux, il se peut que l'automate minimal du langage réalisé ne soit pas réversible. Nous aurons l'occasion de le vérifier sur des exemples. P.-C. Héam [36] a d'ailleurs montré qu'il existe des langages réversibles qui ne peuvent pas être reconnus par un automate réversible déterministe.

La première question, naturelle, qui se pose donc est de savoir si, étant donné un langage rationnel, il est réversible ou non. J.-E. Pin [51] a apporté une réponse à ce problème. La caractérisation d'un langage réversible ne peut pas être faite à la seule vue de son monoïde syntaxique. Celui-ci doit en effet avoir la propriété que ses idempotents commutent, mais ce n'est pas suffisant; il faut aussi que la partie du monoïde qui reconnaît le langage respecte une propriété de clôture :

Théorème 3.1 *[51] Soit $\mathcal{L}$ un langage reconnaissable, M son monoïde syntaxique, P une partie de M et φ un morphisme de A^* dans M tels que $\mathcal{L} = P\varphi^{-1}$. Le langage $\mathcal{L}$ est réversible si et seulement si les deux propriétés suivantes sont vérifiées :*

i) *Les idempotents de M commutent.*

ii) *Quels que soient u, v et w appartenant à A^*, si $u.v^+.w$ appartient à $\mathcal{L}$, alors $u.w$ aussi.*

Les automates réversibles peuvent avoir plusieurs états initiaux ou terminaux. On va voir que ce « défaut de réversibilité » peut être étendu dans une certaine mesure, sans

altérer les propriétés de l'automate.

Définition 3.2 *Un automate $\mathcal{A} = \langle Q,A,E,I,T \rangle$ est un* ***automate quasi-réversible*** *si quel que soit le couple de transitions (e_1,e_2) de E étiquetées par la même lettre, ayant même origine ou même destination, ni e_1 ni e_2 n'appartiennent à une composante fortement connexe.*

En d'autres termes, si l'on a l'une des deux configurations suivantes,

l'état r n'appartient pas à la même composante fortement connexe que p ou q, qui eux, peuvent éventuellement faire partie de la même composante fortement connexe.

Remarque 3.1 Un automate localement réversible (*cf.* [36]), c'est-à-dire dont les composantes fortement connexes sont réversibles, n'est pas forcément quasi-réversible. On peut par exemple, dans un automate localement réversible, avoir une des deux configurations précédentes, avec p et r dans la même composante fortement connexe et q en dehors.

Il s'avère que la classe de langages ainsi reconnus n'est pas plus riche que les langages réversibles :

Lemme 3.1 *Le langage reconnu par un automate quasi-réversible est réversible.*

Démonstration. Soit $\mathcal{A}$ un automate quasi-réversible. Soit n le nombre de paires de transitions en contradiction avec l'hypothèse de réversibilité. On montre par récurrence sur n que l'on peut transformer un automate quasi-réversible en un automate réversible. Si n est nul, l'automate $\mathcal{A}$ est réversible. Supposons qu'on sait effectuer la transformation pour tout entier strictement inférieur à n. Soit $\{e_1,e_2\}$ une paire de transitions en contradiction avec l'hypothèse de réversibilité. On considère $\mathcal{A}_1$ et $\mathcal{A}_2$, deux copies de $\mathcal{A}$. On supprime e_1 de $\mathcal{A}_1$ et e_2 de $\mathcal{A}_2$. L'union de ces deux copies reconnaît le même langage que $\mathcal{A}$. Le nombre de paires de transitions qui contredisent la réversibilité est strictement inférieur à n dans chacune des deux copies, on peut donc les transformer en automates réversibles dont l'union est un automate réversible équivalent à $\mathcal{A}$. □

Il se peut que la représentation réversible soit de taille exponentielle par rapport à la représentation quasi-réversible. Ceci a été montré par P.-C. Héam [35] sur l'exemple suivant :

Exemple 13 Soit k un entier positif et $\mathcal{L}_{r_2} = (ab + ba + aa)^{2k}$. Ce langage (fini) est reconnu par l'automate minimal suivant qui est quasi-réversible : P.-C. Héam montre que la taille du plus petit automate réversible qui reconnaît $\mathcal{L}_{r_2}$ est au moins $\left(\frac{3\sqrt{2}}{4}\right)^{\frac{k}{6}}$.

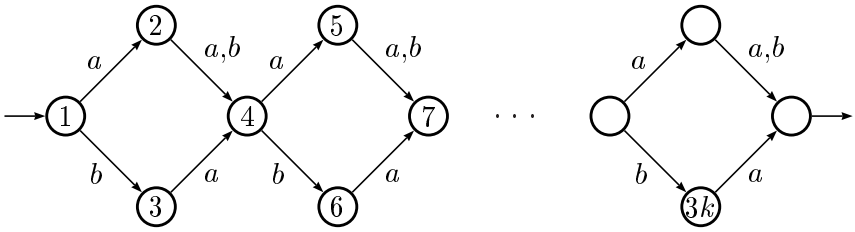


FIG. 1 – *L'automate minimal de* $\mathcal{L}_{r_2}$

On peut donc se demander si, en pratique, une bonne représentation d'un langage réversible n'est pas une représentation quasi-réversible. L'intérêt d'un automate réversible est que la lecture d'un mot, aussi bien par la gauche que par la droite, ne peut se faire que sur un nombre borné de chemins.

On peut définir une mesure de non déterminisme d'un automate de la façon suivante. On considère que l'automate est à multiplicité dans le semi-anneau $(\mathbb{N} \cup \{-\infty\}, \mathsf{max}, +)$; chaque transition a pour valeur 0, sauf celles qui sont non déterministes qui ont valeur 1. Ainsi, la valeur d'un chemin est le nombre de transitions non déterministes qu'on y a rencontré et la valeur d'un mot est le nombre maximal de choix non déterministes qu'on peut avoir à effectuer pour parcourir un calcul étiqueté par ce mot. La mesure de non-déterminisme de l'automate est la borne supérieure des coefficients de la série ainsi réalisée. Notons que cette mesure de non déterminisme est sensiblement différente de celle introduite par I. Simon [59].

On peut, de même, définir une mesure de non co-déterminisme. Il est facile de voir qu'un automate est quasi-réversible si, et seulement si, ses mesures de non déterminisme et de non co-déterminisme sont toutes deux finies.

Le petit lemme suivant est élémentaire. C'est pourtant à lui qu'on recourra dans de nombreuses preuves sur les langages réversibles.

LEMME 3.2 *Soit $\mathcal{A}$ un automate réversible. Il existe un entier strictement positif k tel que, pour tout mot u de A^*, tout chemin étiqueté par u^k est une boucle.*

Démonstration. Considérons le monoïde de transition de $\mathcal{A}$. Soit u un mot dont l'image dans ce monoïde est un idempotent. Pour tout état p de $\mathcal{A}$, soit $p \cdot u = \emptyset$, soit $p \cdot u = q$ et, comme l'image de u est un idempotent, $q \cdot u = q$, et, comme $\mathcal{A}$ est réversible, $p = q$. Un idempotent du monoïde de transition de $\mathcal{A}$ est donc une identité partielle, en d'autres termes, tout mot dont l'image dans le monoïde de transition est un idempotent, ne peut étiqueter que des boucles.

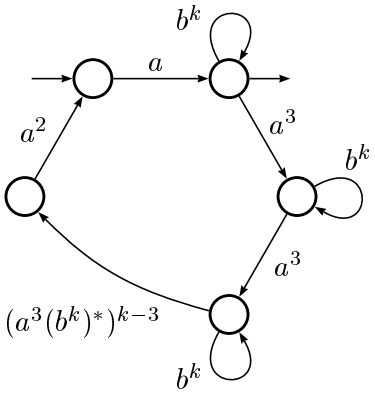
Pour tout élément α du monoïde de transition, il existe un entier strictement positif k_α tel que α^{k_α} est un idempotent. En prenant, par exemple, le plus petit multiple commun de tous les k_α, pour α dans le monoïde de transition, on obtient un entier k tel que α^k est un idempotent pour tout α. L'entier k vérifie le lemme. □

Pour clore cette présentation des langages réversibles, il faut signaler qu'il existe un

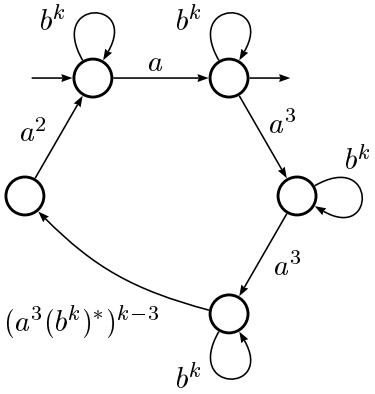
certain antagonisme entre réversibilité et déterminisme ou même non-ambiguïté :

PROPOSITION 3.3 *Il existe des langages réversibles tels qu'il n'existe aucun automate réversible non ambigu qui reconnaît l'un de ces langages.*

Démonstration. La preuve de cette proposition se présente sous forme d'exemple. En guise d'exercice, on peut montrer que tout automate réversible qui reconnaît le langage $a^*ba^* + b^*ab^*$ compte au moins deux calculs étiquetés par ab (de même pour ba). On va ici développer un autre exemple dans lequel l'ensemble des mots étiquetant plusieurs chemin est infini. Considérons le langage reconnu par l'automate réversible $\mathcal{A}_{r_1}$ présenté exemple 12.1 page 71. Les langages $a(b^*a^3)^*$ et $b^*(a^3)^*a$ sont des sous-ensembles du langage accepté par $\mathcal{A}_{r_1}$. Soit $\mathcal{B}$ un automate réversible qui accepte $\mathcal{L}_{r_1}$. Il existe un entier k tel que, quel que soit le mot u de A^*, tout chemin de $\mathcal{B}$ étiqueté par u^k est une boucle. Le mot $a(b^ka^3)^k$ est accepté par $\mathcal{B}$. La configuration suivante apparaît donc dans $\mathcal{B}$:



Si le calcul étiqueté par a^{3k+1} est unique dans $\mathcal{B}$, comme $b^k(a^{3k+1})$ est aussi accepté par $\mathcal{B}$, on obtient la configuration suivante qui accepte un langage strictement plus gros que $\mathcal{L}$:



Donc, il existe au moins deux calculs étiquetés par a^{3k+1} dans $\mathcal{B}$, et, tout mot de $a(a^{3k})^*$ étiquette au moins deux chemins. □

— ∘ —

2 Langages à groupe

Une sous-classe importante des langages réversibles est celle des langages à groupes :

Définition 3.3 *Un automate $\mathcal{A}$ est un* ***automate à groupe*** *s'il est réversible et complet. Un langage $\mathcal{L}$ est un* ***langage à groupe*** *s'il peut être accepté par un automate à groupe.*

Remarque 3.2 Dans un automate à groupe, chaque lettre induit une bijection de l'ensemble des états sur lui-même, le monoïde de transition d'un tel automate est donc un groupe.

Lemme 3.4 *L'automate minimal d'un langage à groupe est un automate à groupe.*

Démonstration. Le monoïde de transition de l'automate minimal d'un langage est le monoïde syntaxique du langage, qui est le quotient du monoïde de transition de n'importe quel automate qui accepte le langage. Un monoïde quotient d'un groupe est un groupe, donc l'automate minimal d'un langage à groupe est un automate à groupe. □

Les propriétés très particulières des langages à groupe permettent de les étudier de façon bien plus simple que les langages réversibles. Avant d'entamer l'étude des automates universels de langages réversibles, nous allons donc, en guise de hors d'œuvre, étudier les automates universels de langages à groupe, ce qui, espérons-le, facilitera la digestion du plat principal.

— ∘ —

3 Automate universel d'un langage à groupe

3.1 Structure générale de l'automate universel d'un langage à groupe

L'automate minimal d'un langage à groupe est un automate à groupe. Afin d'exploiter ses propriétés lors de l'étude de l'automate universel, nous considérons que ce dernier est construit selon la méthode exposée dans le théorème 2.1 et que ses états sont indexés par des sous-ensemble de Q (ensemble des états de l'automate minimal). Dans ce qui suit, un état p de $\mathcal{U}_\mathcal{L}$ sera donc un sous-ensemble de Q.

Définition 3.4 *Soit $\mathcal{U}_\mathcal{L}$ l'automate universel d'un langage à groupe $\mathcal{L}$. On dit qu'un état p appartient à l'****étage*** *k de $\mathcal{U}_\mathcal{L}$ si $\mathsf{Card}(p) = k$.*

Les étages ainsi définis forment une véritable hiérarchie dans l'automate universel : on peut dire qu'un étage est supérieur à un autre s'il existe une transition entre un état du premier et un état du second. La proposition suivante assure que cette relation est bien une relation d'ordre. Toutefois, cet ordre n'est généralement pas total, comme le montre l'exemple 15.

Proposition 3.5 *Soit p_1 et p_2 deux états de l'automate universel $\mathcal{U}_\mathcal{L}$ d'un langage à groupe $\mathcal{L}$ appartenant respectivement aux étages k_1 et k_2. S'il y a une transition de p_1*

à p_2 dans $\mathcal{U}_\mathcal{L}$, alors $k_1 \leqslant k_2$.
En particulier, si p_1 et p_2 sont dans la même composante fortement connexe, ce sont deux états du même étage.

Démonstration. Soit p_1 et p_2 deux états de $\mathcal{U}_\mathcal{L}$ et a une lettre de A qui étiquette une transition entre ces deux états. Par définition de l'automate universel, $p_1 \cdot a \subseteq p_2$. Comme l'automate minimal est complet et que chaque lettre agit fidèlement sur les états, $\mathsf{Card}(p_1 \cdot a) = \mathsf{Card}(p_1)$, donc $\mathsf{Card}(p_1) \leqslant \mathsf{Card}(p_2)$. □

La proposition suivante résume les propriétés de ces étages :

PROPOSITION 3.6 *Soit $\mathcal{U}_\mathcal{L}$ l'automate universel d'un langage à groupe et son automate minimal $\mathcal{A} = \langle Q, A, E, \{i\}, T \rangle$.*

i) *Le premier étage est isomorphe à $\mathcal{A}$.*

ii) *Le dernier étage est isomorphe à l'automate co-déterministe minimal.*

iii) *Le nombre d'étages est au plus égal au cardinal de T.*

iv) *Si deux états appartiennent au même étage, soit ils ne sont pas accessibles l'un-l'autre, soit ils sont dans la même composante fortement connexe.*

Démonstration. i) Il n'est en effet pas difficile de voir que le morphisme qui envoie un état p de $\mathcal{A}$ sur l'état $\{p\}$ de $\mathcal{U}_\mathcal{L}$ est bijectif sur le premier étage.

ii) Il suffit de considérer l'image miroir du langage et d'utiliser le résultat de la proposition 2.8.

iii) Puisque les représentants des factorisations sont calculés en utilisant les ensembles $u \cdot T$, pour $u \in A^*$ et que, dans un automate à groupe, chacun de ces ensembles est de même cardinal que T, le cardinal des états de l'automate universel est compris entre 1 et le cardinal de T.

iv) Soit p_1 et p_2 deux états appartenant au même étage. Ils ont donc même cardinal. S'il existe un chemin de p_1 à p_2 étiqueté par u, alors $p_1 \cdot u \subseteq p_2$; comme l'automate minimal est un automate à groupe, $p_1 \cdot u = p_2$. Il existe v dans A^* tel que l'action de $u.v$ sur Q est l'identité, donc $p_1 \cdot (u.v) = p_2 \cdot v = P$; il existe donc un chemin de l'état p_1 à p_2 étiqueté par v. □

COROLLAIRE 3.7 *Soit $\mathcal{A}$ un automate à groupe émondé ayant un état initial et un état final. Alors $\mathcal{A}$ est à la fois l'automate minimal et universel du langage qu'il reconnaît.*

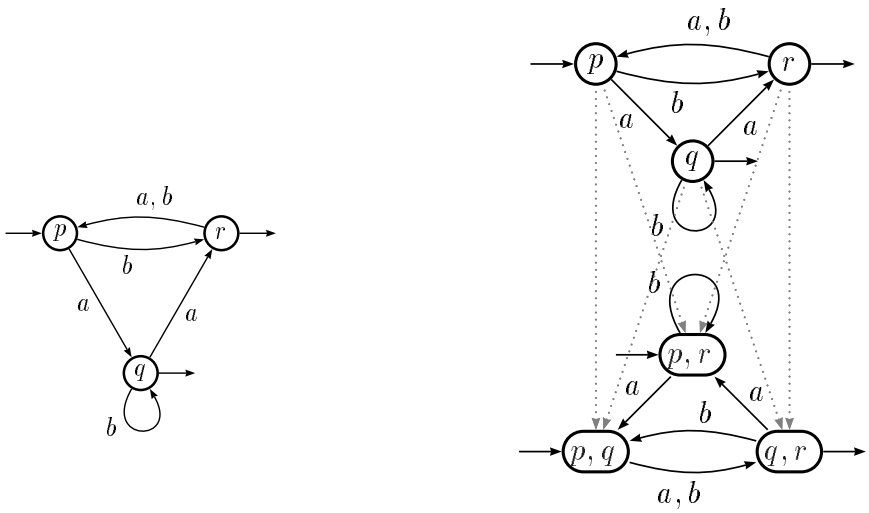
Démonstration. Comme l'automate est en même temps déterministe et co-déterministe, il est minimal. D'après la proposition précédente, l'automate universel n'a qu'un étage qui est isomorphe à l'automate minimal. □

Les automates universels présentés ci-après sont écorchés. Toutefois, par souci de clarté, on indique tous les états initiaux et terminaux.

EXEMPLE 14 Cet exemple et le suivant sont des langages dont le monoïde syntaxique est le groupe $\mathfrak{S}_3$. On verra que la structure de l'automate universel (ainsi d'ailleurs que

celle de l'automate minimal) ne dépend pas seulement du monoïde syntaxique mais aussi de la partie qui reconnaît le langage. Les états des automates universels sont étiquetés par les ensembles d'états des automates minimaux auxquels ils correspondent. Le premier exemple est le langage $\mathcal{L}_{g_1}$ reconnus par les automates de la figure 2. On peut remarquer que l'automate minimal n'est pas le graphe de Cayley du groupe syntaxique. On voit que les étages de l'automate universel dépendent bien du nombre d'états terminaux de l'automate minimal et non pas du nombre d'éléments du groupe qui forment la partie reconnaissant le langage.

Exemple 15.1 La structure des automates reconnaissants le langage $\mathcal{L}_{g_2}$ est sensiblement différente, bien que les lettres a et b représentent les mêmes générateurs de $\mathfrak{S}_3$.



(a) L'automate minimal de $\mathcal{L}_{g_1}$.

(b) L'automate universel de $\mathcal{L}_{g_1}$.

Fig. 2 – *Deux automates reconnaissant* $\mathcal{L}_{g_1}$.

L'automate universel du langage $\mathcal{L}_{g_2}$, présenté figure 4 comporte quatre étages. Les transitions spontanées entre les étages sont représentées par des arcs en pointillés gris. On voit que les étages ne sont pas totalement ordonnés : il n'y a pas de transition entre l'étage 2 (composé de trois sous-automates « en triangle ») et l'étage 3 (formé des états « *pqr* » et « *stu* »).

— ∘ —

3.2 Structure des composantes fortement connexes de l'automate universel

Les composantes fortement connexes de l'automate universel d'un langage à groupe sont très particulières.

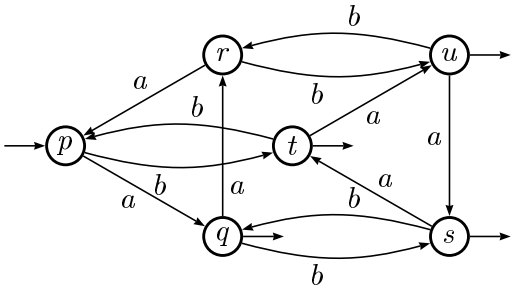


FIG. 3 – *L'automate minimal de* $\mathcal{L}_{g_2}$

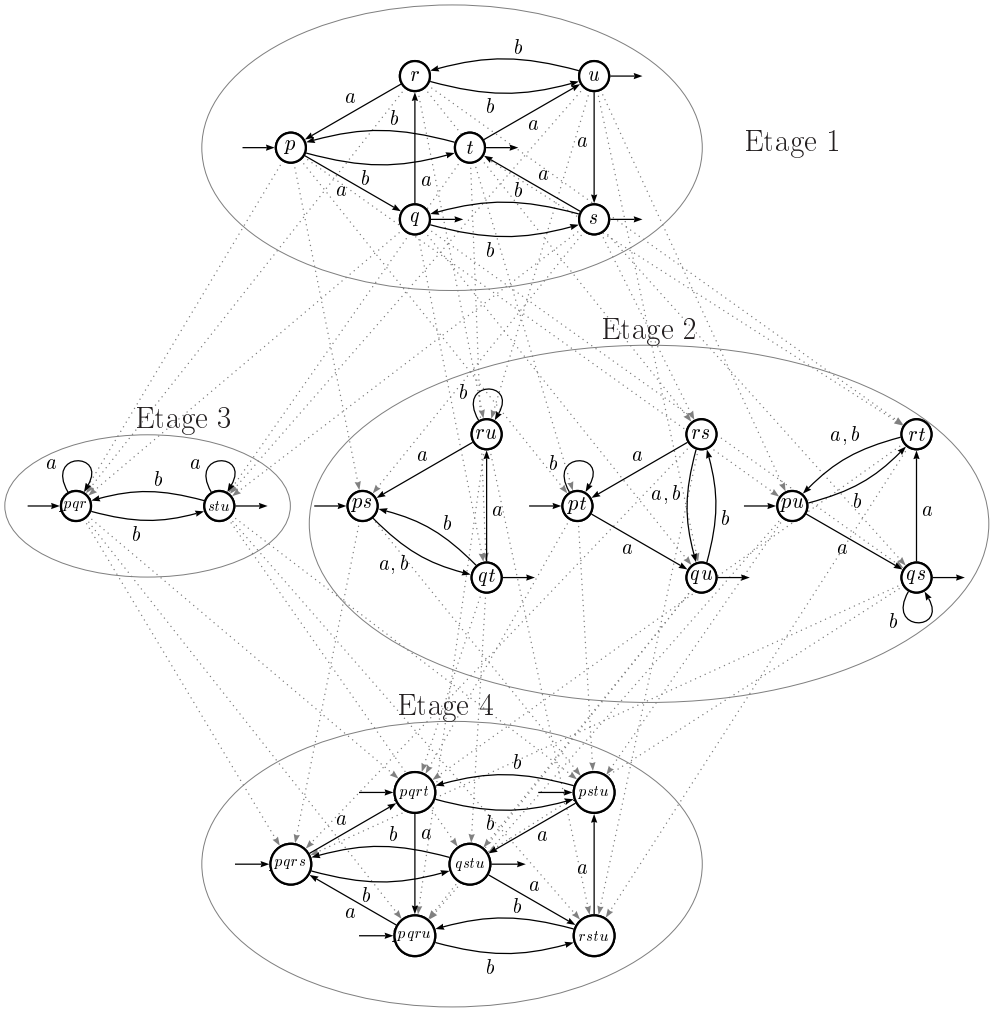


FIG. 4 – *L'automate universel de* $\mathcal{L}_{g_2}$

Chaque état d'un automate universel correspond à un sous-ensemble d'états de l'automate minimal. Comme l'automate minimal est lui-même l'image par morphisme du graphe de Cayley du monoïde syntaxique, chacun de ses états correspond à un sous-ensemble d'éléments du monoïde syntaxique.

Dans le cas des langages à groupe, l'action des lettres sur les états du graphe de Cayley est fidèle et transitive. Comme les états de l'automate minimal sont des ensembles d'états du graphe de Cayley, l'automate minimal, qui est l'orbite de l'état initial sous l'action des lettres, est, nous l'avons déjà vu, un automate à groupe (avec une seule composante fortement connexe).

Pour la même raison, chaque composante fortement connexe de l'automate universel représente l'orbite d'un sous-ensemble d'état de l'automate minimal sous l'action des lettres et est donc un automate à groupe.

THÉORÈME 3.2 *Soit $\mathcal{U}_\mathcal{L}$ l'automate universel d'un langage à groupe $\mathcal{L}$. Chaque pelote de $\mathcal{U}_\mathcal{L}$ est un sous-automate à groupe de $\mathcal{U}_\mathcal{L}$. Le monoïde de transition de chaque pelote de $\mathcal{U}_\mathcal{L}$ est un quotient du groupe syntaxique. Tout état de $\mathcal{U}_\mathcal{L}$ appartient à une pelote.*

Démonstration. Soit p un état de $\mathcal{U}_\mathcal{L}$ et u un mot de A^*. Il existe un mot v tel que l'action de $u.v$ sur les états de l'automate minimal est l'identité. Donc $p \cdot (u.v) = p$ et il existe une boucle étiquetée par $u.v$ autour de p. A partir de p, on peut donc lire n'importe quel mot u sans sortir de la composante fortement connexe de p : celle-ci est donc complète. D'autre part, puisqu'on n'a fait aucune supposition sur p, il suit que tout état de p appartient à une pelote.

L'état q d'une pelote atteint à partir d'un autre état p de la même pelote, en lisant un mot u, fait partie du même étage, donc $q = p \cdot u$, ce qui caractérise uniquement q à partir de p et u; la pelote est donc déterministe (du point de vue de ses transitions). La pelote est aussi co-déterministe, puisque l'image miroir d'un langage à groupe est un langage à groupe.

Soit $\mathcal{P}$ une pelote de l'automate universel, Q' son ensemble d'états et $G' \subseteq \mathsf{S}_{Q'}$ son groupe de transition. Si G est le groupe syntaxique, c'est-à-dire le monoïde de transition de l'automate minimal, il existe un morphisme surjectif de G sur G' :

$$\begin{array}{ccccc} G & \longrightarrow & G' & & \\ \alpha & \longmapsto & \alpha' : & \begin{array}{ccc} Q' & \to & Q' \\ p & \mapsto & \{x\alpha \mid x \in p\} \end{array} \end{array}$$

□

REMARQUE 3.3 Le fait que le groupe de transition d'une pelote soit le quotient du groupe de transition de l'automate minimal ne signifie pas que la pelote, en tant que graphe, est un quotient de l'automate minimal. Ce phénomène est illustré par l'exemple de la figure 5.

Pour conclure cet examen des automates universels des langages à groupe, on montre que l'écorché de l'automate universel contient les mêmes pelotes que l'automate universel.

PROPOSITION 3.8 *Soit $\mathcal{L}$ un langage à groupe. Les pelotes formées par les transitions non spontanées de l'écorché de l'automate universel $\check{\mathcal{U}}_\mathcal{L}$ sont exactement les pelotes de l'automate universel $\mathcal{U}_\mathcal{L}$. Si on ne conserve dans l'automate universel que les transitions des pelotes, on obtient un automate à groupe qui reconnaît $\mathcal{L}$.*

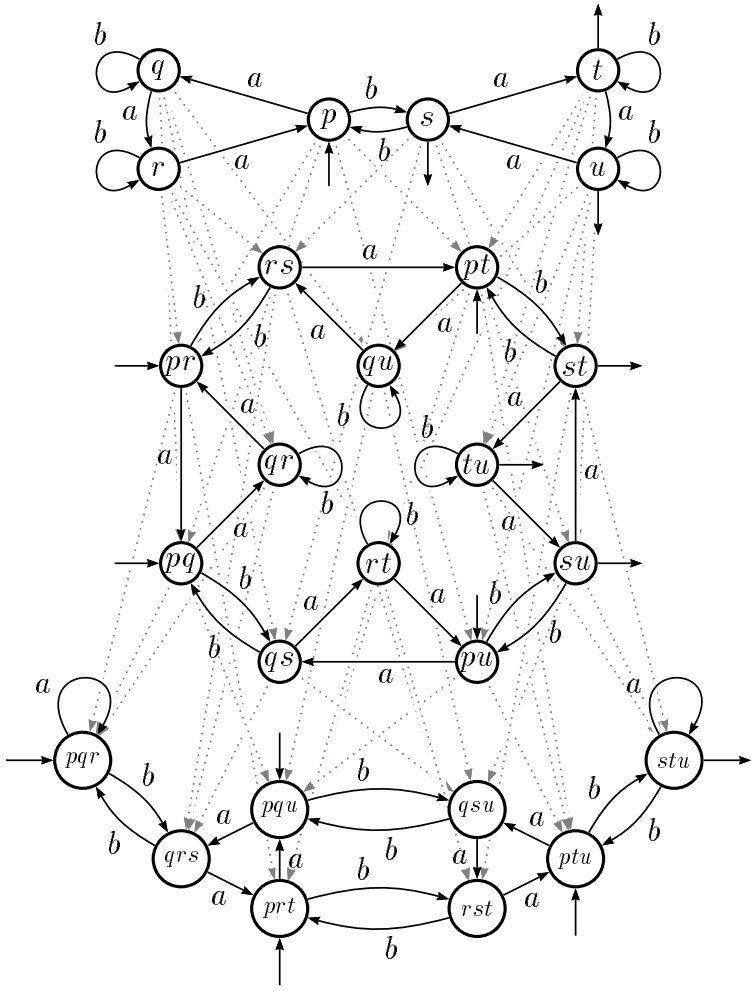


FIG. 5 – *L'automate universel d'un langage à groupe.*

Démonstration. Soit (p,a,p') une transition d'une pelote de $\mathcal{U}_\mathcal{L}$. On a vu que $p \cdot a = p'$. Soit (L,R) et (L',R') les factorisations qui correspondent respectivement à p et à q. On a :

$$R' = \bigcap_{x \in p'} \mathsf{Fut}_\mathcal{A}(x) = \bigcap_{x \in p \cdot a} \mathsf{Fut}_\mathcal{A}(x).$$

Donc (L,R') est bien un couple maximal tel que $L.a.R'$ est inclus dans $\mathcal{L}$. Donc cette transition appartient bien à $\check{\mathcal{U}}_\mathcal{L}$.

Soit (p,a,p') une transition n'appartenant pas à une pelote de $\mathcal{U}_\mathcal{L}$. Donc $p \cdot a$ est strictement inclus dans p'. Il existe $p' = p \cdot a$ qui correspond à un état dont le facteur gauche est strictement inclus dans celui de p', donc avec un facteur droit strictement plus grand que celui de p'. Le couple (L,R'') (avec des notations évidentes) est supérieur au couple (L,R'). La transition (p,a,p') n'apparaît donc pas dans $\check{\mathcal{U}}_\mathcal{L}$.

Le dernier point est trivial, on a vu que chaque pelote est un automate à groupe; leur union forme donc un automate à groupe qui contient l'automate minimal et reconnaît donc tout le langage. □

— ∘ —

4 Automate universel d'un langage réversible

4.1 Structure générale de l'automate universel d'un langage réversible

L'automate minimal d'un langage réversible n'est pas nécessairement réversible. Cette simple constatation rend les choses, certes plus intéressantes, mais aussi un tantinet plus délicates. Indexer les états de l'automate universel en fonction de l'automate minimal ne paraît en effet pas judicieux pour exploiter les propriétés de réversibilité du langage. C'est donc un automate réversible qui reconnaît le langage qui nous servira de référence.

Il faut bien souligner qu'*a priori*, on ne connaît rien de l'automate réversible de référence si ce n'est qu'il est réversible (évidemment) et qu'il reconnaît le langage. Cet automate inconnu va nous apporter des informations sur l'automate universel qui, lui, est connu et constructible à partir de l'automate minimal.

Les états de l'automate minimal d'un langage à groupe nous permettaient d'indexer ceux l'automate universel car ils étaient les images de i sous l'action des mots de A^*. Dans le cas d'un automate réversible $\mathcal{A}$, ce n'est donc pas tant les états qui nous intéresseront, que les images de l'ensemble des états initiaux (I) sous l'action des mots de A^*. Encore une fois, nous sommes ramenés aux états du déterminisé et indexer les états de l'automate universel ne consistera en rien d'autre que leur associer les états du développé $\mathcal{A}$ par la fonction donnée à la proposition 2.18.

Ce sont donc des ensembles d'ensembles qui indexent désormais les états de l'automate universel.

On ne peut donc plus décrire un étage à l'aide d'un simple entier. Toutefois, cette notion demeure.

Définition 3.5 *Soit $\mathcal{A} = \langle Q,A,E,\{i\},T \rangle$ un automate réversible reconnaissant un langage $\mathcal{L}$.*

*Le **profil** d'un état P du développé $\mathcal{V}_\mathcal{A}$ est une fonction κ_P de $[1;\mathsf{Card}(Q)]$ dans $\mathbb{N}$ qui, à tout n associe le nombre d'éléments de P de cardinal n. On dit qu'un état P appartient à l'**étage** κ de $\mathcal{V}_\mathcal{A}$ si κ est le profil de P.*

Exemple 12.2 Considérons le développé de $\mathcal{A}_{r_1}$ présenté dans l'exemple 12. Voici les profils de quelques-uns de ses états :

Etats	$\{pq,pr\}$	$\{p,qr\}$	$\{p,q,r\}$	$\{pqr\}$
Profils	[0,2,0]	[1,1,0]	[3,0,0]	[0,0,1]

On va voir que, de même que dans le cas des groupes, les étages ainsi définis sont ordonnés :

Proposition 3.9 *Soit P_1 et P_2 deux états du développé $\mathcal{V}_\mathcal{A}$ d'un automate réversible $\mathcal{A}$ appartenant respectivement aux étages κ_1 et κ_2. S'il y a une transition de P_1 à P_2 dans l'automate universel de $\mathcal{L}$, alors $\kappa_1 \leqslant_{\mathsf{lex}} \kappa_2$.*
En particulier, si p_1 et p_2 sont dans la même composante fortement connexe, ce sont deux états du même étage.

Démonstration. Soit (P,a,Q) une transition de $\mathcal{V}_\mathcal{A}$. Considérons la fonction qui à X appartenant à P associe $X \cdot a$. On peut séparer les éléments de P en deux parties. D'une part, ceux pour lesquels X et $X \cdot a$ ont le même cardinal; à cause de la réversibilité, la fonction est une bijection sur ce sous-ensemble de P. D'autre part, ceux pour lesquels le cardinal de X est strictement supérieur au cardinal de $X \cdot a$; alors, à cause de la réversibilité et parce que les éléments de P sont incomparables, il n'existe pas X' de même cardinal que $X \cdot a$ tel que $X' \cdot a = X \cdot a$ (ce qui entraînerait $X' \subseteq X$). Donc κ_Q est lexicographiquement plus grand que κ_P. □

Exemple 12.3 L'écorché de l'automate universel du langage réversible $\mathcal{L}_{r_1}$ est représenté figure 7. Là encore, on a indiqué tous les états initiaux et terminaux.

Les états sont disposés de haut en bas selont leur signature, de « pq », « qr » et « pq » dont la signature est [0,1,0], à « $\{p,q\}$ », « $\{q,r\}$ » et « $\{p,q\}$ » dont la signature est [2,0,0]. Les états « $\{pqr\}$ » et « $\{p,q,r\}$ » du développé de $\mathcal{A}_{r_1}$ ne sont respectivement pas accessible et co-accessible, ils ne peuvent donc pas faire partie de l'automate universel. On vérifie que les autres états appartiennent bien à $\mathcal{U}_{\mathcal{L}_{r_1}}$.

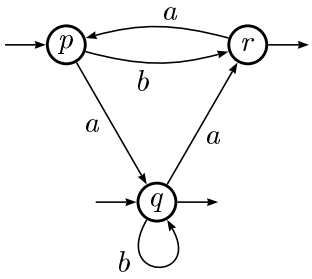


Fig. 6 – *Un automate réversible acceptant* $\mathcal{L}_{r_1}$

Dans certains cas particuliers, on peut aussi utiliser la construction de l'automate universel à partir de l'automate minimal pour retrouver certains résultats. Ainsi, le corollaire 3.7 pour les automates à groupe s'applique partiellement aux automates réversibles :

Proposition 3.10 *Soit $\mathcal{A}$ un automate réversible émondé ayant un état initial et un état final. Alors $\mathcal{A}$ est l'automate minimal et universel du langage.*

Démonstration. De même que pour les langages à groupes, l'automate est en même temps déterministe et co-déterministe, il est donc minimal. Les états de l'automate universel sont indexés par la clôture par intersection des états du co-déterminisé, c'est-dire, en l'occurrence, par les singletons. Il existe une transition $(\{p\},a,\{q\})$ dans l'automate universel si et seulement si $p \cdot a = q$, donc si et seulement s'il existe une transition (p,a,q) dans l'automate minimal. De même pour les états initiaux et terminaux. □

— ◦ —

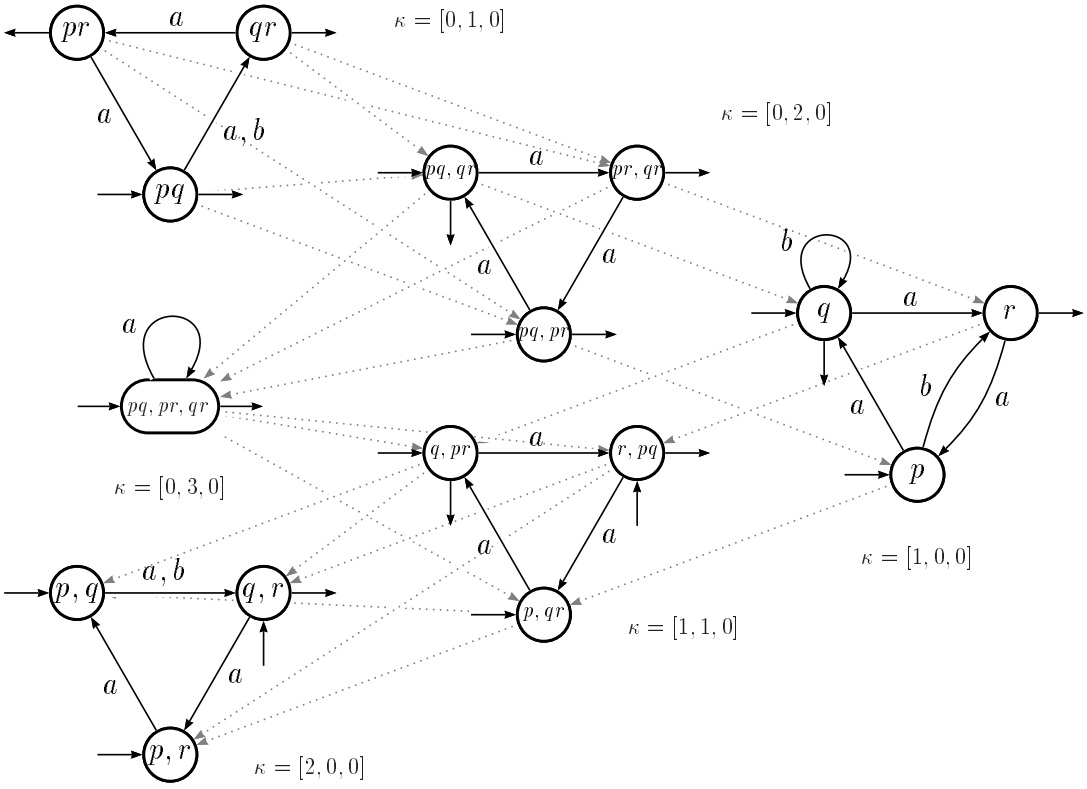


FIG. 7 – *L'automate universel de $\mathcal{L}_{r_1}$ (écorché)*

4.2 Structure des pelotes de l'automate universel

De même que chaque pelote de l'automate universel d'un langage à groupe est un automate à groupe, pour les langage réversibles, la propriété se transmet. Les deux énoncés suivants correspondent au théorème 3.2. Ils sont énoncés pour le développé d'un automate réversible. La propriété pour l'automate universel en est un simple corollaire.

PROPOSITION 3.11 *Soit $\mathcal{L}$ un langage réversible, $\mathcal{A}$ un automate réversible qui accepte $\mathcal{L}$ et $\mathcal{V}_\mathcal{A}$ son développé. Soit P et Q les sous-ensembles représentant deux états p et q d'une même pelote de l'automate $\mathcal{V}_\mathcal{A}$. Alors, pour tout mot u (resp. v) étiquetant un chemin de p à q (resp. de q à p), $P \cdot u = Q$ et $Q \cdot v = P$.*

Démonstration. D'après la proposition 3.9 les états d'une même pelote ont le même profil. On montre que u induit une bijection entre les éléments de P de cardinal i et ceux de Q de même cardinal. C'est-à-dire que quel que soit X dans P, il existe Y dans Q, de même cardinal que X tel que $X \cdot u = Y$. De même, on montre que v induit une bijection de Q sur P.
Si cette propriété est fausse, soit i minimal contredisant cette affirmation. Il existe X appartenant à P, de cardinal i et Y appartenant à Q de cardinal j tel que $Y \subset X \cdot u$. On

a donc $\mathsf{Card}(Y) = j < i$; il existe donc X_j appartenant à P tel que $X_j \cdot u = Y \subset X \cdot u$. Comme l'automate $\mathcal{A}$ est réversible, ceci entraîne $X_j \subset X$ et contredit le fait que P est une anti-chaîne. □

THÉORÈME 3.3 *Les pelotes de l'automate universel d'un langage réversible sont réversibles.*

Démonstration. La proposition 3.11 montre que l'état d'une pelote du développé d'un automate réversible atteint à partir d'un autre état de la même pelote est caractérisé par l'étiquette du chemin qui les relie. Les pelotes de l'automate développé sont donc déterministes[1]. Comme l'automate universel est un sous-automate du développé, ses pelotes sont déterministes. De plus, l'image miroir d'un langage réversible est un langage réversible; les pelotes de l'automate transposé sont donc déterministes. Par conséquence, les pelotes de l'automate universel d'un langage réversible sont aussi co-déterministes. Finalement, elles sont réversibles. □

REMARQUE 3.4 Le fait que les pelotes de l'automate universel soient réversibles n'est pas caractéristique des langages réversibles. En fait, ceci entraîne seulement que le point ii) de la caractérisation du théorème 3.1 est vérifié. Par exemple, les pelotes de l'automate universel du langage a^*b^*, dans le monoïde syntaxique duquel les idempotents ne commutent pas, sont réversibles.

Les pelotes de l'automate universel d'un langage réversible sont fortement liées à la structure du monoïde syntaxique de ce langage :

PROPOSITION 3.12 *Soit $\mathcal{U}_\mathcal{L}$ l'automate universel d'un langage réversible et p un état appartenant à une pelote $\mathcal{P}$ de $\mathcal{U}_\mathcal{L}$. Soit u un mot correspondant à un idempotent du monoïde syntaxique. Alors, s'il existe un chemin dans $\mathcal{P}$ partant de p étiqueté par u, ce chemin est une boucle autour de p étiquetée par u.*

Démonstration. Si u correspond à un idempotent, comme le futur d'un état de l'automate universel est l'image inverse d'une partie du monoïde syntaxique, s'il existe un mot v tel que $u.v$ est dans le futur de p, alors pour tout entier k, le mot $u^k.v$ est dans le futur de p. On suppose qu'il existe un chemin, interne à la pelote de p, qui part de p étiqueté par un mot u correspondant à un idempotent. Par le lemme des tiroirs, il existe deux entiers i,j ($j > 0$) et un état $q = (L,R)$ tels que u^i étiquette un chemin de p à q et que u^j étiquette une boucle autour de q. Donc $L.u^j \subseteq L$, et comme u correspond à un idempotent et que les facteurs sont reconnus par le monoïde syntaxique, on a $L.u \subseteq L$ et donc, le mot u étiquette une boucle autour de q. Les pelotes sont réversibles donc en particulier co-déterministes, le chemin de p à q étiqueté par u^i est donc i fois la boucle autour de q et $p = q$. □

— ∘ —

1. En ce qui concerne les transitions, on fait ici abstraction des états initiaux ou terminaux

5 Construction d'un automate réversible

Nous allons montrer ici que l'automate universel d'un langage réversible contient un sous-automate quasi-réversible qui reconnaît le même langage. Comme d'une part, on sait construire l'automate universel à partir de l'automate minimal et que, d'autre part, on a vu que la transformation d'un automate quasi-réversible en un automate réversible est simple, même si elle peut se révéler avoir une complexité exponentielle; ceci permet de construire un automate réversible pour un langage réversible donné.

Pour obtenir ce résultat, nous allons considérer un automate réversible (*a priori* inconnu) que l'on met sous une forme particulière, de telle façon que son image dans l'automate universel est quasi-réversible. Pour ce faire, on doit introduire des automates d'apparence singulière.

5.1 Cordes

Si on quotiente chaque composante fortement connexe d'un automate, on obtient un automate orienté acyclique. On peut alors séparer les différents chemins (réussis) de ce graphe.

Une fois ces chemins séparés, comme on est particulièrement prudent et qu'on a une bonne mémoire, on peut remplacer chaque sommet par la composante fortement connexe dont il est le quotient en prenant soin que les états extrémaux des transitions soient corrects. C'est cette manipulation que nous allons détailler dans les propositions suivantes :

Définition 3.6 *Un automate $\mathcal{A} = \langle Q,A,E,I,T \rangle$ est une* ***corde***[2] *s'il est émondé, n'a qu'un seul état initial, un seul état final et si il y a au plus une transition arrivant sur chaque composante fortement connexe*[3] *et au plus une transition partant de chaque composante fortement connexe. Un automate est un* ***quipu***[4] *si c'est une union finie disjointe de cordes.*

Proposition 3.13 *Tout automate $\mathcal{A}$ émondé est équivalent à un quipu $\mathcal{B}$ tel qu'il existe un morphisme surjectif de $\mathcal{B}$ sur $\mathcal{A}$.*

Démonstration. La preuve est la même que pour le passage d'un automate quasi-réversible à un automate réversible. Au lieu de considérer les paires de transitions qui contredisent la réversibilité, on considère les paires de transitions qui contredisent le fait d'être une corde. En dupliquant ainsi l'automate pour chacune de ces paires et en supprimant dans chaque copie l'une des deux transition, on obtient finalement un quipu qui est équivalent à $\mathcal{A}$. Cette construction est illustrée par la figure 9. □

Remarque 3.5 La construction précédente a une complexité déplorable, puisque c'est une opération encore plus brutale que celle qui permet de transformer un automate quasi-réversible en réversible, cette dernière étant déjà exponentielle. Nous ne l'utiliserons que

2. à noeuds, évidemment.

3. au sens large

4. Instrument inca formé d'une collection de cordelettes sur lesquelles une succession de nœuds permettait, selon leur nombre et leur forme, de tenir une comptabilité.

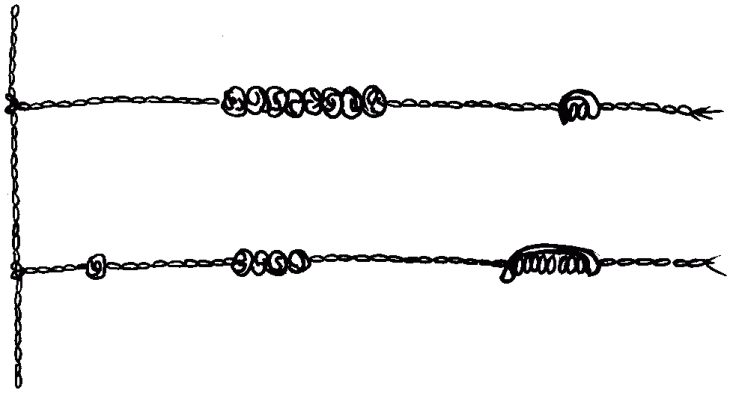


FIG. 8 – *Schéma d'un quipu inca.*

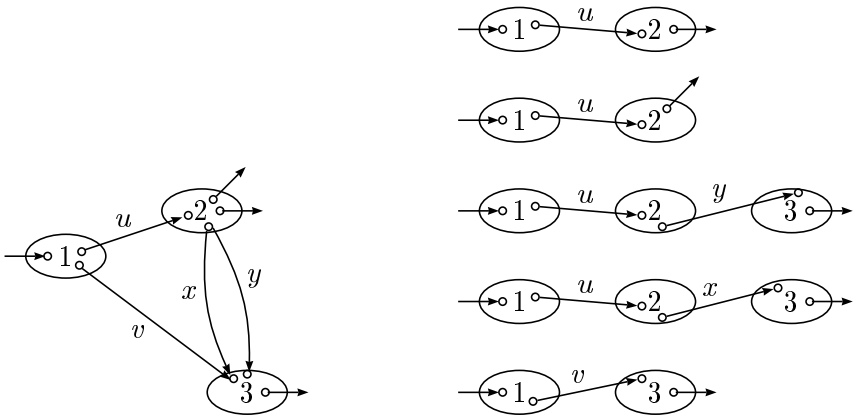


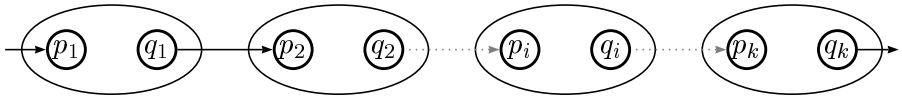
FIG. 9 – *D'un automate à un quipu.*

de façon théorique, pour obtenir un quipu ayant des propriétés similaires à l'automate de départ. Par exemple, si l'automate de départ est déterministe, chaque composante connexe du quipu l'est aussi.

— ◦ —

5.2 Automate quasi-réversible et automate universel

Nous avons montré dans le lemme 3.1 que tout automate quasi-réversible est équivalent à un quipu réversible. Ceci est vrai, en particulier, pour tout automate réversible. Nous allons utiliser dans cette partie cette forme particulière. Chaque corde d'un quipu a la forme suivante :



Si l'automate est en plus réversible, chacune de ces corde est déterministe et co-déterministe.

Noter que les états p_i et q_i jouent un rôle particulier : tout calcul de l'automate les emprunte obligatoirement.

Définition 3.7 *Un quipu est minimal si on ne peut supprimer aucune de ses cordes sans diminuer le langage reconnu.*

Proposition 3.14 *Soit $\mathcal{L}$ un langage réversible et $\mathcal{A}$ un quipu réversible minimal qui reconnaît $\mathcal{L}$. Soit φ un morphisme de $\mathcal{A}$ dans l'automate universel $\mathcal{U}_\mathcal{L}$. Alors, l'image de chaque composante fortement connexe de $\mathcal{A}$ est une composante fortement connexe de $\mathcal{U}_\mathcal{L}$.*

Démonstration. On montre que, si ce n'est pas le cas, le quipu n'est pas minimal.

S'il y a une pelote $\mathcal{P}$ de $\mathcal{A}$ dont l'image par φ n'est pas une pelote de $\mathcal{U}_\mathcal{L}$, la pelote de $\mathcal{U}_\mathcal{L}$ qui contient $\mathcal{P}\varphi$ contient une transition e qui n'appartient pas à l'image de $\mathcal{P}$.

Soit $\mathcal{B}$ la corde qui contient $\mathcal{P}$. Tout mot reconnu par $\mathcal{B}$ peut se décomposer en $u.v$, avec u et v respectivement passé et futur d'un état r de la pelote $\mathcal{P}$.

Dans $\mathcal{P}\varphi$, il existe une boucle autour de $r\varphi$, étiquetée par un mot w, qui emprunte la transition e. Comme l'automate universel reconnaît $\mathcal{L}$ et que u (*resp.* v) appartient au passé (*resp.* au futur) de $r\varphi$, pour tout entier n, le mot $u.w^n.v$ appartient à $\mathcal{L}$. Comme l'automate $\mathcal{A}$ est réversible, il existe un entier k tel que, pour tout état p de $\mathcal{A}$, soit w^k étiquette une boucle autour de $\mathcal{A}$, soit il n'y a pas de chemin étiqueté par w^k qui part de p. Le mot $u.w^k.v$ est accepté par $\mathcal{A}$. S'il était accepté par $\mathcal{B}$, w^k étiquetterait une boucle autour de r dans $\mathcal{P}$ et, puisque les pelotes de l'automate universel sont déterministes, la transition e appartiendrait à $\mathcal{P}\varphi$. Donc il existe une autre corde $\mathcal{B}'$ qui accepte $u.w^k.v$, et, puisque w^k étiquette nécessairement une boucle, qui accepte aussi $u.v$. Tout mot de $\mathcal{B}$ est donc ainsi accepté par une autre corde, ce qui contredit la minimalité du quipu $\mathcal{A}$. □

Proposition 3.15 *Soit $\mathcal{L}$ un langage réversible et $\mathcal{A}$ un quipu réversible minimal qui reconnaît $\mathcal{L}$. Soit φ un morphisme de $\mathcal{A}$ dans $\mathcal{U}_\mathcal{L}$. Alors, l'image de chaque corde de $\mathcal{A}$ dans l'automate universel est réversible.*

Démonstration. Par l'absurde, si ce n'est pas vrai, il existe deux transitions e_1 et e_2 d'une corde $\mathcal{B}$, qui ont la même étiquette, et dont les images ont soit le même état de départ soit le même état d'arrivée. On peut, sans perte de généralité, supposer que c'est la première configuration qui se présente.

Étant donnée la forme d'une corde, toute paire de transitions d'une corde est telle qu'il existe un chemin qui passe par ces deux transitions. Supposons que w est un mot qui étiquette un chemin dont la première transition est e_1 et qui se termine dans l'état de départ de e_2. L'image de ce chemin par φ est une boucle (puisque $e_1\varphi$ et $e_2\varphi$ ont le même état de départ. La transition $e_1\varphi$ appartient donc à une pelote, et comme les pelotes de $\mathcal{U}_\mathcal{L}$ sont réversibles, $e_2\varphi$ n'appartient à aucune pelote, donc e_2 non plus. Donc tout calcul de $\mathcal{B}$ emprunte e_2. Tout mot accepté par $\mathcal{B}$ se décompose en $u.v$, où v étiquette la fin d'un calcul dont la première transition est e_2.

La fin de la preuve est identique à celle de la proposition 3.14. Pour tout entier n, le mot $u.w^n.v$ appartient au langage. Pour k identique à celui de la preuve précédente, le

mot $u.w^k.v$ n'étiquette pas un chemin dans $\mathcal{B}$, sinon e_1 et e_2 auraient le même état de départ, ce qui contredirait la réversibilité. Il existe donc une corde $\mathcal{B}'$ qui accepte $u.w^n.v$ et aussi $u.v$, ce qui contredit la minimalité de $\mathcal{A}$. □

— ∘ —

Le théorème principal de cette partie apparaît comme un corollaire des deux propositions précédentes.

THÉORÈME 3.4 *Il existe un sous-automate de l'automate universel d'un langage réversible, qui est quasi-réversible et qui reconnaît tout le langage.*

Démonstration. On montre que l'image d'un quipu réversible minimal, qui reconnaît le langage $\mathcal{L}$, dans $\mathcal{U}_\mathcal{L}$ est quasi-réversible, quel que soit le morphisme.

Soit $\mathcal{A}$ un quipu réversible minimal qui reconnaît $\mathcal{L}$ et φ un morphisme de $\mathcal{A}$ dans $\mathcal{L}$. L'image de chaque corde de $\mathcal{A}$ est réversible. Supposons $\mathcal{A}\varphi$ non quasi-réversible. Il existe alors deux transitions e_1 et e_2 qui appartiennent respectivement à deux cordes $\mathcal{B}_1$ et $\mathcal{B}_2$ de $\mathcal{A}$ et telles que $e_1\varphi$ et $e_2\varphi$ contredisent la quasi-réversibilité. L'une de celles-ci (supposons qu'il s'agit de $e_1\varphi$) appartient à une pelote, et l'autre non (les pelotes sont réversibles). L'état p commun à ces deux transitions fait partie d'une pelote qui contient $e_1\varphi$. L'état p est l'image par φ d'une des extrémités q de e_2. L'état q appartient à une composante fortement connexe dont l'image est la pelote de p tout entière (proposition 3.14). Donc $e_1\varphi$ appartient à $\mathcal{B}_2\varphi$, ce qui contredit la réversibilité de $\mathcal{B}_2\varphi$ et la proposition 3.15. □

Ce théorème nous dit que l'automate universel d'un langage réversible contient un sous-automate quasi-réversible qui reconnaît le langage. En fait, la preuve est plus précise et indique qu'il existe un tel sous-automate dont les composantes fortement connexes sont exactement celles de l'automate universel. On peut donc définir un sous-automate quasi-réversible maximal de l'automate universel qui reconnaît le langage.

PROPOSITION 3.16 *Soit $\mathcal{U}_\mathcal{L}$ l'automate universel d'un langage réversible $\mathcal{L}$. Soit $\mathcal{R}_\mathcal{L}$ l'automate obtenu à partir de $\mathcal{U}_\mathcal{L}$ en appliquant l'algorithme suivant :*

1. *Calcul des composantes fortement connexes de $\mathcal{U}_\mathcal{L}$ et marquage des transitions qui appartiennent à une pelote.*
2. *Suppression des transitions qui contredisent la quasi-réversibilité et qui n'appartiennent à aucune pelote.*

Alors, $\mathcal{R}_\mathcal{L}$ est un automate quasi-réversible qui reconnaît le langage $\mathcal{L}$.

Démonstration. Comme les pelotes de $\mathcal{U}_\mathcal{L}$ sont réversibles, $\mathcal{R}_\mathcal{L}$ est quasi-réversible. De plus, d'après la preuve du théorème 3.4, l'image de tout quipu réversible minimal qui reconnaît $\mathcal{L}$ est un sous-automate de $\mathcal{R}_\mathcal{L}$ qui reconnaît donc $\mathcal{L}$. □

Cet automate quasi-réversible est en quelque sorte universel pour les automates quasi-réversibles :

PROPOSITION 3.17 *Soit $\mathcal{A}$ un automate quasi-réversible qui reconnaît un langage $\mathcal{L}$ avec*

un nombre minimal d'états. Alors $\mathcal{A}$ est un sous-automate de l'automate $\mathcal{R}_\mathcal{L}$ défini ci-dessus.

Démonstration. Soit φ un morphisme de $\mathcal{A}$ dans l'automate universel $\mathcal{U}_\mathcal{L}$. Soit $\mathcal{B}$ la décomposition en quipu réversible (rendu minimal) de $\mathcal{A}$ et μ le morphisme de $\mathcal{B}$ sur $\mathcal{A}$ qui en découle (morphisme surjectif à cause de la minimalité de $\mathcal{A}$). Alors $\mu.\varphi$ est un morphisme de $\mathcal{B}$ dans $\mathcal{U}_\mathcal{L}$ et, plus précisément, comme on l'a montré, un morphisme de $\mathcal{B}$ dans $\mathcal{R}_\mathcal{L}$. Donc μ est un morphisme de $\mathcal{A}$ dans $\mathcal{R}_\mathcal{L}$; ce morphisme est injectif à cause de la minimalité de $\mathcal{A}$. □

— ∘ —

Exemple 16 Considérons le langage $\mathcal{L}_{r_3}$ défini par l'automate minimal de la figure 10. On peut vérifier que le langage $\mathcal{L}_{r_3}$ est un langage réversible. L'automate universel de $\mathcal{L}_{r_3}$

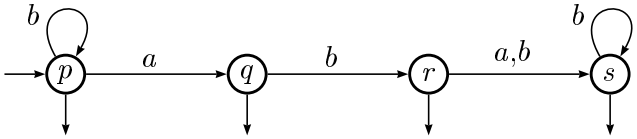


Fig. 10 – *L'automate minimal de $\mathcal{L}_{r_3}$.*

est représenté figure 11. En supprimant toutes les transitions hors des composantes forte-

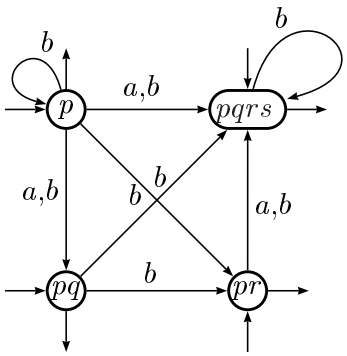


Fig. 11 – *L'automate universel de $\mathcal{L}_{r_3}$.*

ment connexes qui sont en contradiction avec la quasi réversibilité, on obtient l'automate de la figure 12.

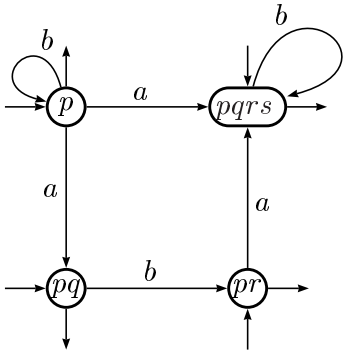


FIG. 12 – *Un automate quasi-réversible pour* $\mathcal{L}_{r_3}$.

Table des exemples

1 Le langage $\mathcal{L}_1 = A^*abA^*$ 21

1.1 Un graphe orienté . . . 21
1.2 Le monoïde fini M_1 . . . 22
1.3 Langage rationnel . . . 30
1.4 Langage reconnaissable . . . 31
1.5 Automate reconnaissant $\mathcal{L}_1$ . . . 35
1.6 Langage reconnu par un automate . . . 36
1.7 Factorisations de $\mathcal{L}_1$ . . . 46
1.8 Automate universel de $\mathcal{L}_1$ . . . 48
1.9 L'automate universel de $\mathcal{L}_1$ est reconnaissable . . . 51
1.10 Automate A-universel de $\mathcal{L}_1$ . . . 54
1.11 Écorché de l'automate universel de $\mathcal{L}_1$ . . . 61
1.12 Développé d'un automate reconnaissant $\mathcal{L}_1$ . . . 66
1.13 Ecorché du développé . . . 70

2 Distance préfixe 25

3 Le semi-anneau $\mathbb{N}$ 25

4 Le semi-anneau de Boole 26

5 Semi-anneaux principaux 26

6 Etoile de 1/2 27

7 Langages reconnus par la partie $\{1,2\}$ de $\mathbb{Z}/3\mathbb{Z}$ 31

7.1 Langage reconnaissable . . . 31
7.2 Factorisations de $\mathcal{L}_3 = \{1,2\}$ dans le monoïde $N_2 = \mathbb{Z}/3\mathbb{Z}$ . . . 46
7.3 Automate universel du langage $\mathcal{L}_3$ . . . 48
7.4 Factorisations du langage $\mathcal{L}_3' = \{u \in A^* \mid |u|_a \neq |u|_b \mod 3\}$ . . . 52

7.5 Automate {1}-universel de $\mathcal{L}_3$ 53
7.6 Automate universel du langage $\mathcal{L}_3'$ 55
7.7 Écorché de l'automate {1}-universel du langage $\mathcal{L}_3$ 61

8 Le langage $\mathcal{L}_2 = a^+$ 47
8.1 Factorisations du langage $\mathcal{L}_2 = a^+$ dans le monoïde a^* 47
8.2 Automate universel du langage $\mathcal{L}_2$ 48
8.3 Automate $\{a\}$-universel du langage $\mathcal{L}_2$ 53
8.4 Écorché de l'automate universel du langage $\mathcal{L}_2$ 61
8.5 Développé de l'automate minimal du langage $\mathcal{L}_2$ 66

9 Factorisations d'un langage non reconnaissable 47

10 Le langage $\mathcal{L}_4 = ((a+c)(b+c)+(b+(a+c)a)b^*(a+c))^*(b+(a+c)a)$ sur A^* 59
10.1 Calcul de l'automate universel 59
10.2 Ecorché de l'automate universel de $\mathcal{L}_4$ 61
10.3 Un automate d'enlacement minimal pour $\mathcal{L}_4$ 120

11 Ordre sur $\mathcal{P}(Q)$ 64

12 Le langage $\mathcal{L}_{r_1}$ sur A^* (Langage réversible) 71
12.1 Automate réversible $\mathcal{A}_{r_1}$ et développé de $\mathcal{A}_{r_1}$ 71
12.2 Profil des états du développé de $\mathcal{A}_{r_1}$ 86
12.3 Automate universel de $\mathcal{L}_{r_1}$ 87
12.4 Un automate quasi-réversible de $\mathcal{L}_{r_1}$ 119

13 Automate réversible 77

14 Automate minimal et universel de $\mathcal{L}_{g_1}$ (Langage à groupe) 81

15 Le langage $\mathcal{L}_{g_2}$ sur A^* (Langage à groupe) 81
15.1 Automate minimal et universel de $\mathcal{L}_{g_2}$ 81
15.2 Hauteur d'étoile de $\mathcal{L}_{g_2}$ 113

16 Construction d'un automate réversible 94

17 Enlacement d'un graphe 100

18 Morphismes conforme et non conforme 102

19 Indice i_{E} **105**

20 Automate sur $\mathbb{N}_{\mathsf{m}}$ avec valeur minimale éloignée **126**

21 Une série non-séquentielle uniformément bornée **129**

22 L'automate unaire avec multiplicité $\mathcal{A}_{\mathsf{m}_1}$ **130**

22.1 Les paramètres de $\mathcal{A}_{\mathsf{m}_1}$ 130

22.2 Décision de la séquentialité de α_{m_1} 135

22.3 Calcul d'un automate non ambigu pour α_{m_1} 141

23 Fréquence des circuits de poids maximum **133**

24 $\mathsf{E}_4 = (\mathbf{x}(a\cdot(\mathbf{y}\,b)^\star\cdot a)\,\mathbf{y}+\mathbf{y}\,(b\cdot(a\,\mathbf{x})^\star\cdot b)\,\mathbf{x})^\star$ **151**

24.1 Présentation de l'expression 151

24.2 Arbre de l'expression E_4 154

24.3 Automate de l'expression E_4 155

24.4 Dérivation unitaire 170

25 $\mathsf{E}_1 = (\frac{1}{6}a^* + \frac{1}{3}b^*)^*$ **153**

25.1 Terme constant de E_1 153

25.2 Dérivées de E_1 160

25.3 Termes dérivés 162

26 $\mathsf{E}_2 = (a\,b\,a + (a\,(a - b\,a)))$ **153**

26.1 Ecriture de E_2 153

26.2 Dérivées de E_2 161

26.3 Termes dérivés fantômes 163

27 $\mathsf{E}_3 = 5\,((2\,a\,b) + ((3\,b)\cdot(4\,(a\,b)^\star)))^\star$ **151**

27.1 Terme constant de E_3 151

27.2 Dérivation et dérivation unitaire 171

28 Problème sur les termes dérivés fantômes **168**

Index

action, 23
alphabet, 23
automate
    quasi-réversible, 77
    à groupe, 80
    à multiplicité, 41
    complet, 36
    déterminisé d'un automate, 38
    déterministe, co-déterministe, 36
    des termes dérivés, 166
    émondé, 36
    généralisé, 104
    non-ambigu, 34
    normalisé, 32
    réversible, 76
    sous-jacent, 42
    sur un alphabet, 35
    sur un monoïde, 33
    sur un semi-anneau, 31
    universel, 48
    univoque, 144
automates équivalents, 32

boucle, 20

cône, 28
calcul, 36
    victorieux, 125
carré d'un automate, 41
chemin, 20
circuit (élémentaire), 20
composante fortement connexe, 21
corde, 90

dérivée
    cassante, 171
    unitaire, 169
dérivée d'une expression
    par rapport à un mot, 158
    par rapport à une lettre, 157
développé d'un automate, 66
distance préfixe, 24

écorché
    de l'automate universel, 60
    du développé, 70
enlacement
    d'un automate, 104
    d'un graphe orienté, 100
ensembles rationnels, 30
étage de l'automate universel
    d'un langage à groupe, 80
    d'un langage réversible, 86
état accessible, 36
étoile, 27
expression rationnelle, 98
    à multiplicité, 150
    réduite, 154
    valide, 152
expressions équivalentes, 152

facteur, 24, 46
factorisation, 46
    initiale, terminale, 49
fonction de production, 44

graphe
    acyclique, 21
    connexe, 21
    critique, 130
    de Cayley, 22
    fortement connexe, 21
    orienté, 20
    sous-jacent à un automate, 32

hauteur d'étoile

d'un langage rationnel, 99
d'une expression rationnelle, 99

idempotent, 25
image miroir, 24
indice i_E d'un automate généralisé, 105

langage, 23
à groupe, 80
réversible, 76
rationnel, 30
reconnaissable, 31
reconnu par un automate, 36
lettres, 23
longueur d'une expression, 151

monoïde, 21
de transition, 37
libre, 23
syntaxique, 41
morphisme
conforme, 102
de graphes, 21
syntaxque, 41
mot idempotent pour un automate, 114
mots, 23

ordre
lexicographique, 24
radiciel, 24

partie propre d'une série, 29
passé/futur d'un état, 34
pelote, 21
poids d'un circuit, 130
polynôme, 29
polynômes d'expressions, 156
préfixe, 24
produit de deux automates, 41
profondeur d'une expression, 151

quipu, 90
quotient à gauche
d'une série, 30
dans un monoïde, 22

rationnel, 27
relation, 20
représentation linéaire, 37

semi-anneau, 25
idempotent, 25
positif, 167
principal, 26
semi-module, 28
séries, 28
rationnelles, 30
séquentielles, 43
sous-factorisation, 46
sous-graphe, 20
suffixe, 24
support, 20
support d'une série, 29

terme constant
d'une expression rationnelle, 152
d'une série, 29
termes dérivés
d'une expression, 162
fantômes, 163
transducteurs, 44
transitions, 32, 35
transitions spontanées, 36
translatée, 126
transposé d'un automate, 41
type fini, 28

uniformément divergente, 128

vecteur translaté, 137

Bibliographie

[1] A. Ambainis et R. Freivalds, 1-way quantum finite automata: strengths, weaknesses and generalizations. In *39th Ann. Symposium on FOCS* (1998), 332–342.

[2] D. Angluin, Inference of reversible languages. *J. of the ACM* **29** (1982), 741–765.

[3] V. Antimirov, Partial derivatives of regular expressions and finite automaton constructions. *Theoret. Comput. Sci.* **155** (1996), 291–319.

[4] A. Arnold, A. Dicky, et M. Nivat, A note about minimal non-deterministic automata. *Bull ot the EATCS* **47** (1992), 166–169.

[5] M.-P. Béal, O. Carton, C. Prieur, et J. Sakarovitch, Squaring transducers. *Proc. of Latin 2000, Lect. Notes in Comp. Sci.* **1776** (2000), 397–406.

[6] G. Berry et R. Sethi, From regular expressions to deterministic automata, *Theoret. Comput. Sci.* **48** (1986), 117–126.

[7] J. Berstel et J.-E. Pin, Local languages and the Berry-Sethi algorithm, *Theoret. Comput. Sci.* **155** (1996), 439–446.

[8] J. Berstel et Ch. Reutenauer, *Les séries rationnelles et leurs langages.* Masson, 1984. Traduction: *Rational Series and their Languages.* Springer, 1986.

[9] J. A. Brzozowski, Derivatives of regular expressions. *J. Assoc. Comput. Mach.* **11** (1964), 481–494.

[10] A. Buchsbaum, R. Giancarlo, et J. Westbrook, On the Determinization of Weighted Finite Automata. *Proc. of ICALP'98, Lect. Notes in Comp. Sci.* **1443** (1998), 482–493.

[11] P. Caron et M. Flouret, Glushkov construction for multiplicities. *Pre-Proceedings of CIAA'00*, M. Daley, M. Eramian and S. Yu, eds, Univ. of Western Ontario, (2000), 52–61.

[12] O. Carton, Factorisations et matrice des facteurs. Manuscrit (1994).

[13] J.-M. Champarnaud et D. Ziadi, New finite automaton constructions based on canonical derivatives. *Pre-Proceedings of CIAA'00*, M. Daley, M. Eramian and S. Yu, eds, Univ. of Western Ontario, (2000), 36–43.

[14] Ch. Choffrut, Une caractérisation des fonctions séquentielles et des fonctions sous-séquentielles en tant que relations rationnelles. *Theoret. Comput. Sci.* **5** (1977), 325–337.

[15] M. Chrobak, Finite Automata and Unary Languages. *Theoret. Comput. Sci.* **47** (1986), 149–158.

[16] G. Cohen, P. Moller, J.-P. Quadrat, et M. Viot, Algebraic Tools for the Performance Evaluation of Discrete Event Systems. *IEEE Proc.: Special issue on D.E.S.* **77.1** (1989).

[17] R. Cohen, Star height of certain families of regular events. *J. Computer System Sci.* **4** (1970), 281–297.

[18] R. Cohen et J. A. Brzozowski, General properties of star height of regular events. *J. Computer System Sci.* **4** (1970), 260–280.

[19] J. H. Conway, *Regular algebra and finite machines.* Chapman and Hall, 1971.

[20] B. Courcelle, D. Niwinski, A. Podelski, A geometrical view of the dererminization ond minimization of finite-state automata. *Math. Systems Theory* **24** (1991), 117–146.

[21] F. Dejean et M.-P. Schützenberger, On a question of Eggan. *Inform. and Control* **9** (1966), 23–25.

[22] L. C. Eggan, Transition graphs and the star-height of regular events. *Michigan Mathematical J.* **10** (1963), 385–397.

[23] S. Eilenberg, *Automata, Languages and Machines, volume A.* Academic Press, 1974.

[24] S. Gaubert, Rational Series over Dioids and Discrete Event Systems. *Proc. of the 11th Conf. on Anal. and Opt. of Systems, Lect. Notes in Contr. and Inf. Sci.* **199** (1994).

[25] S. Gaubert, On the Burnside problem for Semigroups of Matrices in the (max,+) Algebra. *Semigroup Forum* **52** (1996), 271–292.

[26] S. Gaubert, Methods and applications of (max,+) linear algebra. *rapport de recherche INRIA* **3088** (1997).

[27] V. Glushkov, The abstract theory of automata. *Russian Mathematical Surveys 16* (1961), 1–53.

[28] Ch. Hagenah et A. Muscholl, Computing ε-Free NFA from regular expressions in $O(n \log^2(n))$ time. *R.A.I.R.O. Inf. Théorique* **34**, (2000), 257–277.

[29] T.E. Hall, Biprefix codes, inverse semigroups and syntactic monoids of injective automata. *Theoret. Comput. Sci.* **32** (1984), 201–213.

[30] T. Harju et J. Karhumäki, The equivalence problem of multitape finite automata. *Theoret. Comput. Sci.* **78** (1991), 347–355.

[31] K. Hashiguchi et N. Honda, The star height of reset-free events and strictly locally testable events. *Inform. and Control* **40** (1979), 267–284.

[32] K. Hashiguchi, Limitedness theorem on finite automata with distance functions. *J. of Comput. Syst. Sci.* **24** (1982), 233–244.

[33] K. Hashiguchi, Algorithms for determining relative star height and star height. *Inform. and Computation* **78** (1988), 124–169.

[34] K. Hashiguchi, Improved limitedness theorem on finite automata with distance functions. *Theoret. Comput. Sci.* **72** (1990), 27–38.

[35] P.-C. Héam, A lower bound for reversible automata. *Theoret. Informatics Appl.* **34** (2000), 331–341.

[36] P.-C. Héam, Contribution à l'algorithmique des automates : complexité et aspects topologiques. *Thèse de doctorat*, Université Paris 7, 2001.

[37] J. Hromkovič, S. Seibert, et T. Wilke, Translating regular expressions into small ε-free nondeterministic finite automata. *Proc. of STACS'97, Lect. Notes in Comp. Sci.* **1200** (1997), 55–66.

[38] D. Krob, Differentiation of K-rational expressions. *Int. J. of Algebra and Computation 2* (1992), 57–87.

[39] W. Kuich et A. Salomaa, *Semirings, Automata, Languages.* Springer, 1986.

[40] G. Lallement, *Semigroups and combinatorial applications.* Wiley, 1979.

[41] H. Leung, Limitedness theorem on finite automata with distance functions: an algebraic proof *Theoret. Comput. Sci.* **81** (1991), 137–145.

[42] S. Lombardy et J. Sakarovitch, On the star height of rational languages, a new presentation for two old results *Proc. of 3rd ICLWC, Kyoto* (2000) (M. Ito, ed.), World Scientific, *à paraître.*

[43] S. Lombardy et J. Sakarovitch, Star height of reversible languages and universal automata, *accepté à Latin'02.*

[44] S. Lombardy, Sequentialization and unambiguity of (max, +) rational series over one letter. *Pre-Proc. of Workshop on Max-plus algebra, Prague* (2001) (S. Gaubert, ed.).

[45] O. Matz et A. Potthoff, Computing small nondeterministic finite automata. *proc. of TACAS'95, BRICS Notes Series* (1995), 74–88.

[46] R. McNaughton et H. Yamada, Regular expressions and state graphs for automata. *IRE Trans. on electronic computers* **9** (1960), 39–47.

[47] R. McNaughton, The loop complexity of pure-group events. *Inform. and Control* **11** (1967), 167–176.

[48] Moebius, *Sur l'étoile.* Gentiane, 1983, Rééd. Casterman, 1990.

[49] M. Mohri, Finite-State Transducers in Language and Speech Processing. *Computat. Ling.* **23.2** (1997), 269–311.

[50] C. Nicaud, Étude du comportement en moyenne des automates finis et des langages rationnels. *Thèse de doctorat*, Université Paris 7, 2000.

[51] J.-E. Pin, On reversible automata. In *Proc. 1st LATIN Conf., (I. Simon, Ed.), Lecture Notes in Comput. Sci.* **583** (1992), 401–416.

[52] J.-E. Pin, Variétés de langages formels. Masson, 1984. Traduction: *Varieties of formal languages* North Oxford Acad. Pub., 1986.

[53] G. Raney, Sequential functions. *J. Assoc. Comput. Mach.* **5** (1958) 177–180.

[54] J. Sakarovitch, A construction on automata that has remained hidden. *Theoret. Comput. Sci.* **204** (1998), 205–231.

[55] J. Sakarovitch, *Éléments de théorie des automates.* Vuibert, *à paraître.*

[56] A. Salomaa, *Jewels of formal language theory.* Computer Science Press, 1981.

[57] P.V. Silva, On free inverse monoid languages. *Theoret. Informatics and Appl.* **30** (1996), 349–378.

[58] I. Simon, Recognizable sets with multiplicities in the tropical semiring. *Proc. of MFCS'88, Lect. Notes in Comp. Sci.* **324** (1988), 107–120.

[59] I. Simon, The non deterministic complexity of a finite automaton. *Mots (M. Lothaire)*, Hermès (1990) 384–400.

[60] I. Simon, On semigroups of matrices over the tropical semiring. *R.A.I.R.O. Inf. Théorique*, (1994), 277–294.

[61] N.J.A. Sloane, *The On-Line Encyclopedia of Integer Sequences.* `http://www.research.att.com/~njas/sequences/`

[62] J.B. Stephen, Presentations of inverse monoids. *J. Pure Appl. Alg.* **63** (1990), 81–112.

[63] M. Szalay, On the maximal order in S_n and S_n^*. *Acta arithm.* **37** (1980), 321–331.