

UNIVERSITÉ PARIS 7

THÈSE

---

MOTS INFINIS,  $\omega$ -SEMIGROUPES ET TOPOLOGIE.

---

*Présentée par*

**Olivier CARTON**

*pour l'obtention du titre de*

DOCTEUR DE L'UNIVERSITÉ PARIS 7

*Spécialité : INFORMATIQUE*

*Soutenue le 3 décembre 1993,  
devant le jury composé de Messieurs :*

|                  |                   |
|------------------|-------------------|
| Serge GRIGORIEFF | <i>Président</i>  |
| Dominique PERRIN | <i>Directeur</i>  |
| André ARNOLD     | <i>Rapporteur</i> |
| Jean-Éric PIN    | <i>Rapporteur</i> |
| John RHODES      | <i>Examineur</i>  |
| Wolfgang THOMAS  | <i>Examineur</i>  |

# Remerciements

Je remercie très chaleureusement Dominique Perrin, mon directeur de thèse, pour son dévouement et ses encouragements constants.

Je remercie Jean-Éric Pin pour ses précieuses explications et pour ses judicieux conseils. Je le remercie ainsi qu'André Arnold d'avoir accepté la tâche de rapporteur.

Je remercie Serge Grigorieff pour les discussions fécondes. Je le remercie aussi de présider le jury de cette thèse.

Je remercie Wolfgang Thomas et John Rhodes d'avoir la gentillesse de participer au jury.

Je remercie Michelle Sarfati, sans laquelle je n'aurais pu m'inscrire en thèse.

Je remercie également tous ceux qui m'ont apporté leur soutien pendant ces années de thèse.



# Introduction

Ce travail s'inscrit dans le cadre général de la théorie des automates et plus particulièrement des langages reconnaissables de mots infinis. Ces langages furent considérés, il y a une trentaine d'années, par Büchi [Büc60] dont la motivation était la recherche de théories logiques décidables.

Par la suite cette théorie s'est développée dans diverses directions. D'une part, elle s'est enrichie de résultats concernant les automates utilisés sur les objets infinis avec les théorèmes de McNaughton [McN66] et Rabin [Rab69] et récemment la construction de Safra [Saf88]. D'autre part, cette théorie s'est révélée liée à d'autres domaines comme la théorie descriptive des ensembles [Mos80], la théorie des jeux [REBC82a, REBC82b] ou encore la théorie des processus [Arn92].

Pour les langages de mots finis, il est bien connu que la notion combinatoire d'automate est équivalente à la notion algébrique de semigroupe. A chaque automate correspond un semigroupe de relations. Les langages rationnels de mots finis sont caractérisés par leur semigroupe syntaxique fini. Cette approche s'est imposée pour l'étude des langages rationnels à travers la correspondance entre les variétés de langages et les variétés de semigroupes [Pin84].

L'analogie pour les mots infinis a d'abord été étudiée par J.P. Pécuchet [Péc86b, Péc86a], mais l'approche la plus satisfaisante est celle proposée par T. Wilke [Wil91] puis [PP93a]. Il s'agit de remplacer les semigroupes par des  $\omega$ -semigroupes qui sont en gros des semigroupes munis d'un produit infini. Les langages de mots infinis reconnaissables par automates sont alors caractérisés par un  $\omega$ -semigroupe syntaxique fini.

Pour les automates de mots infinis, il existe des liens étroits entre la structure d'un automate et la complexité topologique du langage reconnu par cet automate. Ainsi les langages fermés sont reconnus par des automates dont tous les états sont acceptants. Le premier résultat dans cette direction fut obtenu par L.H. Landweber qui caractérisa les langages reconnaissables  $G_\delta$  [Lan69]. La hiérarchie introduite par K. Wagner [Wag79] fut ensuite une avancée impressionnante dans ce domaine puisqu'elle généralise tous les résultats antérieurs.

Dans cette thèse, nous avons développé plus particulièrement cette approche algébrique dans le but de donner des caractérisations purement algébriques de la hiérarchie de Wagner. Plus précisément, nous avons établi les résultats suivants.

- Nous généralisons un résultat de J.P. Pécuchet [Péc86b] en montrant que  $n$  couples liés conjugués d'un semigroupe correspondent à différentes factorisations ramseyennes d'un même mot infini.
- Nous étendons aux algèbres introduites par Wilke les opérations classiques sur les semigroupes. Nous avons ainsi défini le  $\omega$ -semigroupe des parties, le produit de Schützenberger et le produit en couronne qui dans le cas des semigroupes sont les outils fondamentaux pour l'étude des variétés de semigroupes.
- Pour chacune de ces trois constructions, nous obtenons pour les langages de mots infinis l'analogie des résultats classiques pour les langages de mots finis [Pin84].
- La définition du produit en couronne et l'utilisation d'un beau résultat d'I. Simon [Eil76] nous permet de donner un algorithme construisant un  $\omega$ -semigroupe équivalent à un automate de Muller donné.
- Nous caractérisons de façon purement algébrique les classes de la hiérarchie introduite par K. Wagner. Ceci nous autorise à donner de nouvelles démonstrations plus simples de nombreuses propriétés comme la stabilité par image inverse de fonction séquentielle des classes de cette hiérarchie.
- De cette caractérisation, nous déduisons la construction effective d'un automate de Muller équivalent à un  $\omega$ -semigroupe donné.
- Lors de la caractérisation des classes de Wagner, nous avons été amené à considérer des chaînes qui sont des suites emboîtées d'ensembles. À l'aide de diagramme, nous donnons une nouvelle démonstration d'un résultat de Hausdorff [Hau57] concernant ces chaînes.
- Nous introduisons une variante des automates de Rabin, appelés *automates à chaînes* parce que leur condition d'acceptation est une chaîne. Ce mode de reconnaissance se révèle aussi puissant que celui des automates de Rabin habituels. Il a le grand avantage de faciliter de nombreuses opérations comme la complémentation ou l'intersection qui sont des opérations difficiles à effectuer sur les automates de Rabin habituels.
- Nous décrivons un algorithme qui calcule un automate à chaîne équivalent à un automate de Rabin habituel. Si l'automate de Rabin de départ a  $n$

---

états et  $p$  paires de Rabin, l'automate obtenu a  $O(n^p 4^{2p^2})$  et  $p$  paires de Rabin.

- Étant donnés deux automates à chaîne reconnaissant deux  $\omega$ -langages  $X$  et  $Y$ , nous décrivons des constructions simples pour des automates à chaîne reconnaissant le complémentaire  $X^c$ , l'intersection  $X \cap Y$  et l'union  $X \cup Y$ . Si les longueurs des chaînes des automates de départ sont  $p$  et  $q$ , les longueurs des chaînes des automates obtenus ont  $p + 1$ ,  $p + q$  et  $p + q$ .
- Étant donné un automate à chaîne, nous décrivons un algorithme qui réduit la longueur de la chaîne de la condition d'acceptation à l'indice de Rabin du langage reconnu par l'automate. La longueur de la chaîne est alors minimale. Ceci fournit une nouvelle caractérisation de l'indice de Rabin.

Cette thèse est constituée de cinq chapitres :

Dans le premier chapitre, nous introduisons les notations et les premières définitions. Les notions de langages et d'automates sont définies. Nous reprenons également quelques résultats classiques utiles pour la suite comme le théorème de Landweber

Dans le deuxième chapitre, nous développons les notions de semigroupes et de  $\omega$ -semigroupes. Nous étendons aux  $\omega$ -semigroupes le produit en couronne et le produit de Schützenberger. Ces opérations permettent, en particulier de calculer un  $\omega$ -semigroupe équivalent à un automate donné.

Dans le troisième chapitre, à l'aide de la nouvelle preuve du résultat de Hausdorff, nous explicitons l'algèbre de Boole engendrée par une semi-algèbre de Boole.

Dans le quatrième chapitre, nous donnons une présentation algébrique de la hiérarchie de Wagner. Nous caractérisons à l'aide des  $\omega$ -semigroupes les classes de complexité. Nous terminons par la construction d'un automate équivalent à un  $\omega$ -semigroupe.

Dans le dernier chapitre, nous introduisons une classe particulière d'automates de Rabin pour lesquels la construction d'un automate pour l'union, l'intersection et le passage au complémentaire est élémentaire. Nous établissons une nouvelle caractérisation de l'indice de Rabin d'un langage.

Ce travail devrait se prolonger naturellement dans deux directions. D'une part, la continuation des travaux de K. Wagner, comme les liens avec les classes de Wadge méritent d'être approfondis et d'autre part les résultats de T. Wilke [Wil91] et J.P. Pécuchet [Péc86b] concernant les variétés de langages de mots infinis invitent également à développer la théorie des variétés de  $\omega$ -semigroupes.



# Chapitre Premier

## Premières définitions

### 1.1 Notations

Nous fixons ici quelques notations générales. Pour deux ensembles  $E$  et  $F$ , le complémentaire de  $E$  est noté  $E^c$  et l'union  $E \cup F$  de  $E$  et de  $F$  est notée  $E + F$ . Cette notation sera en particulier utilisée pour les expressions rationnelles. De façon cohérente, la différence  $E \setminus F$  est notée  $E - F$ .

Les applications sont notées à droite. Pour une fonction  $f$  d'un ensemble  $E$  dans un ensemble  $F$ , l'image d'un élément  $x$  de  $E$  est notée  $xf$ . On étend cette notation à un sous-ensemble  $X$  de  $E$  en écrivant  $Xf = \{xf \mid x \in X\}$  l'ensemble des images des éléments de  $X$  par la fonction  $f$ . De la même façon on note  $Yf^{-1}$  l'image réciproque d'une partie  $Y$  de  $F$  par la fonction  $f$ . Si  $g$  est une fonction de l'ensemble  $F$  dans un ensemble  $G$ , la composée de  $f$  et de  $g$ , habituellement notée  $g \circ f$ , sera notée  $fg$  par simple juxtaposition. Ainsi l'image de l'élément  $x$  par la composée  $fg$  est écrit  $xfg$ .

Une relation d'équivalence sur un ensemble  $E$  *sature* une partie  $X$  de  $E$  si toute classe d'équivalence est incluse dans  $X$  ou disjointe de  $X$ . La partie  $X$  s'écrit alors comme une union de classes d'équivalence. Par extension, on dit qu'une fonction  $f$  de  $E$  dans  $F$  *sature* une partie  $X$  si la relation d'équivalence nucléaire associée à la fonction  $f$  *sature*  $X$ . On a alors l'égalité  $X = Xff^{-1}$ .

### 1.2 Mots et langages

Nous fixons ici le vocabulaire en donnant les définitions élémentaires. Les notions d'alphabet, de mot fini et infini et de langages sont successivement définies.

Dans toute la suite, nous allons considérer des suites finies ou infinies d'élé-

ments d'un ensemble fini  $A$  appelé *alphabet*. Les éléments de cet ensemble sont appelés les *lettres*. On suppose toujours l'alphabet fini.

Une suite finie de lettres s'appelle un *mot fini* sur l'alphabet  $A$  (ou plus simplement un *mot*). On écrit un mot  $u$  par simple juxtaposition de ses lettres

$$u = a_1 a_2 \dots a_n.$$

Le nombre  $n$  de lettres du mot  $u$  est par définition la *longueur* du mot  $u$  et elle est notée  $|u|$ . Pour une lettre  $a$ , on note  $|u|_a$  le nombre d'occurrences de la lettre  $a$  dans le mot  $u$ .

**Exemple :** Quelques exemples de mots sur l'alphabet  $A = \{a, b\}$  :  $a$ ,  $bab$  et  $abbabbabab$ .

La suite vide est notée  $\varepsilon$  ou  $1$  et elle est appelée *mot vide*. L'ensemble des mots sur l'alphabet  $A$  est noté  $A^*$  et l'ensemble des mots non vides est noté  $A^+$  ( $A^* = A^+ + \{\varepsilon\}$ ). L'ensemble des mots est muni de l'opération *produit de concaténation* qui à deux mots  $u = a_1 \dots a_m$  et  $v = b_1 \dots b_n$  associe le mot  $uv = a_1 \dots a_m b_1 \dots b_n$ . Cette opération est associative et le mot vide est un élément neutre pour cette opération.

Un *mot infini* sur l'alphabet  $A$  est une suite infinie de lettres de  $A$ . C'est donc une application de  $\mathbb{N}$  dans  $A$  que l'on note également par juxtaposition

$$u = a_0 a_1 a_2 \dots$$

L'ensemble des mots infinis sur l'alphabet  $A$  est noté  $A^\omega$  et l'on pose

$$A^\infty = A^+ + A^\omega$$

qui est donc l'ensemble des mots non vides, finis ou infinis, sur l'alphabet  $A$ . Le produit d'un mot fini  $u = a_0 \dots a_n$  et d'un mot infini  $v = b_0 b_1 b_2 \dots$  est égal au mot infini  $uv$  obtenu par simple concaténation

$$uv = a_0 \dots a_n b_0 b_1 b_2 \dots$$

Un mot infini  $u = a_0 a_1 a_2 \dots$  est dit *périodique* s'il existe un entier  $p$  tel que  $a_k = a_{k+p}$  pour tout  $k \in \mathbb{N}$ . De même un mot infini  $u = a_0 a_1 a_2 \dots$  est dit *ultimement périodique* s'il existe deux entiers  $k_0$  et  $p$  tels que  $a_k = a_{k+p}$  pour tout  $k \geq k_0$ .

**Exemple :** Les mots infinis  $abababab\dots$  et  $bbababab\dots$  sont respectivement périodique et ultimement périodique.

Un mot  $v$  fini est dit *préfixe* d'un mot  $u$  éventuellement infini si et seulement si il existe un mot  $w$  tel que  $u = vw$ .

Un ensemble de mots finis est appelé un *langage* et un ensemble de mots infinis est appelé un  $\omega$ -*langage*. Un langage est dit *préfixe* si et seulement si aucun mot n'est préfixe d'un autre mot du langage.

### 1.2.1 Langages rationnels

Nous définissons dans cette partie la notion de langage rationnel. Nous commençons par les mots finis puis nous étendons les définitions aux mots infinis.

Pour deux langages  $X_1$  et  $X_2 \subset A^*$ , on définit les opérations *rationnelles*

- (1) l'union  $X_1 \cup X_2$  notée  $X_1 + X_2$ ;
- (2) le produit  $X_1 X_2 = \{x_1 x_2 \mid x_1 \in X_1 \text{ et } x_2 \in X_2\}$ ;
- (3) l'itération  $X_1^+ = \{x_1 \dots x_n \mid n \geq 0 \text{ et } x_1, \dots, x_n \in X_1\}$ .

Pour un langage  $X \subset A^*$ , on note  $X^* = X^+ + \{\varepsilon\}$ .

L'ensemble  $\mathcal{Rat}(A^+)$  des *langages rationnels* de  $A^+$  est la plus petite famille  $\mathcal{R}$  de langages de  $A^+$  telle que

- (i)  $\mathcal{R}$  contient l'ensemble vide et les singletons  $\{a\}$  pour tout  $a \in A$ .
- (ii)  $\mathcal{R}$  est fermée par union finie, produit fini et par itération.

Les ensembles rationnels sont décrits par des *expressions rationnelles* formées avec les lettres de l'alphabet  $A$  et les symboles des opérations  $+$ ,  $*$  et  $^+$ .

**Exemple :** L'expression  $A^*a$  représente le langage constitué des mots sur l'alphabet  $A = a + b$  se terminant par la lettre  $a$ .

**Exemple :** L'expression  $(b + ab^*a)^*$  représente le langage constitué des mots sur l'alphabet  $A = a + b$  ayant un nombre pair d'occurrences de la lettre  $a$ .

Pour les langages de mots infinis, on étend de façon naturelle le produit de deux langages en posant pour  $X_1 \subset A^*$  et  $X_2 \subset A^\omega$

$$X_1 X_2 = \{x_1 x_2 \mid x_1 \in X_1 \text{ et } x_2 \in X_2\}$$

On définit également l'itération infinie d'un langage  $X \subset A^*$  par

$$X^\omega = \{x_0 x_1 x_2 \dots \mid x_0, x_1, x_2 \dots \in X\}$$

L'ensemble  $\mathcal{Rat}(A^\infty)$  des *langages rationnels* de  $A^\infty$  est la plus petite famille  $\mathcal{R}$  de langages de  $A^\infty$  telle que

- (i)  $\mathcal{R}$  contient l'ensemble vide et les singletons  $\{a\}$  pour tout  $a \in A$ .
- (ii)  $\mathcal{R}$  est fermée par union finie, produit fini, par itération et par itération infinie.

On a alors le théorème suivant qui caractérise les langages rationnels de  $A^\omega$ . Ce théorème découle directement des définitions.

**Théorème 1** *Un  $\omega$ -langage est un langage rationnel de  $A^\omega$  si et seulement si il est union finie de langages de la forme  $X_1X_2^\omega$  où  $X_1$  et  $X_2$  sont des langages rationnels de  $A^+$ .*

**Exemple :** L'ensemble des mots infinis sur l'alphabet  $A = a + b$  commençant par le mot  $aab$  est égal à  $aabA^\omega$ .

**Exemple :** L'ensemble des mots infinis sur l'alphabet  $A = a + b$  ayant un nombre fini d'occurrences de  $a$  est égal à  $A^*b^\omega$ .

**Exemple :** L'ensemble des mots infinis sur l'alphabet  $A = a + b$  ayant un nombre infini d'occurrences de  $a$  est égal à  $(A^*a)^\omega$ .

## 1.2.2 Topologie

Dans ce paragraphe, nous présentons quelques rappels de topologie. Un exposé plus exhaustif peut être trouvé dans [Bou74].

On munit l'alphabet de la topologie discrète. L'espace  $A^\omega = A^{\mathbb{N}}$  des mots infinis est alors naturellement muni de la topologie produit. Une base d'ouverts est la famille des langages de la forme  $uA^\omega$  où  $u$  est un mot fini de  $A^*$ . Cette famille constitue une base dénombrable d'ouverts fermés rationnels.

Les ouverts sont les  $\omega$ -langages de la forme  $XA^\omega$  où  $X$  est un langage quelconque de  $A^*$ . De façon traditionnelle, on note  $G$  la famille des ouverts et  $F$  la famille des fermés. Pour une famille  $C$  de langages, on note  $C_\delta$  la famille des langages qui peuvent s'écrire comme une intersection dénombrable de langages de la famille  $C$ . De manière duale, on note  $C_\sigma$  la famille des langages qui peuvent s'écrire comme une union dénombrable de langages de la famille  $C$ . On a ainsi les familles classiques  $G_\delta$ ,  $F_\sigma$ ,  $G_{\delta\sigma}$  et  $F_{\sigma\delta}$  de la hiérarchie borélienne.

La topologie discrète de  $A$  peut être définie par la distance  $\delta$  définie par  $\delta(a, a) = 0$  pour tout  $a \in A$  et  $\delta(a, b) = 1$  pour tous  $a \neq b$ . Comme produit dénombrable d'espaces métrisables, l'espace  $A^\omega$  est métrisable. En effet la

topologie de  $A^\omega$  peut être définie par la distance ultramétrique  $d$  définie par

$$d(u, v) = \begin{cases} 0 & \text{si } u = v \\ 2^{-r} & \text{sinon avec } r = \min\{n \mid u_n \neq v_n\} \end{cases}$$

De manière intuitive, deux mots sont proches s'ils coïncident sur un long préfixe. Pour deux mots  $u$  et  $v$ , on a  $d(u, v) \leq 2^{-k}$  si et seulement si les  $k$  premières lettres de  $u$  sont égales aux  $k$  premières lettres de  $v$ . En particulier un mot  $u$  appartient à la fermeture  $\overline{X}$  d'un langage  $X$  si pour tout  $k \in \mathbb{N}$ , il existe un mot  $x_k$  de  $X$  qui coïncide avec  $u$  sur un préfixe de longueur au moins  $k$ .

À titre d'illustration, nous allons démontrer la proposition suivante qui caractérise les  $\omega$ -langages ouverts et fermés.

**Proposition 1** *Un  $\omega$ -langage de  $A^\omega$  est ouvert et fermé si et seulement si il est de la forme  $XA^\omega$  où  $X$  est un langage fini de  $A^*$ .*

En particulier, les  $\omega$ -langages ouverts et fermés sont nécessairement rationnels.

**Démonstration :** Nous démontrons auparavant un lemme technique.

**Lemme 1** *Soit  $U$  un ouvert de  $A^\omega$ . Il existe un langage  $X \subset A^*$  préfixe tel que  $U = XA^\omega$ .*

**Démonstration :** L'ouvert  $U$  s'écrit  $U = YA^\omega$ . On pose  $X = Y - YA^+$ . Le langage  $X$  est préfixe et vérifie  $U = XA^\omega$ . Si de plus le langage  $Y$  est rationnel,  $X$  est également rationnel.  $\square$

Il est clair qu'un  $\omega$ -langage de la forme  $XA^\omega$  est ouvert. Si de plus le langage  $X$  est fini, le  $\omega$ -langage  $XA^\omega$  est fermé puisque

$$A^\omega - XA^\omega = (A^n - XA^*)A^\omega$$

où  $n$  est un majorant de la longueur des mots de  $X$ .

Réciproquement, soit  $XA^\omega$  un  $\omega$ -langage ouvert et fermé. Grâce au lemme précédent, on peut supposer que le langage  $X$  est préfixe. Si  $X$  n'est pas fini, on peut trouver par compacité de  $A^\omega$  une suite  $x_n$  d'éléments de  $X$  telle que la suite de mots  $x_n a^\omega$  converge vers un mot  $u$  qui appartient nécessairement à  $XA^\omega$  puisque ce  $\omega$ -langage est fermé par hypothèse. Le mot  $u$  a donc un préfixe  $x$  dans  $X$ . Pour  $n$  assez grand, le mot  $x$  est préfixe de  $x_n$ , ce qui contredit le fait que langage  $X$  soit préfixe.  $\square$

### 1.3 Automates

Nous définissons ici la notion d'automate. Pour les mots finis la condition d'acceptation est l'arrêt dans un état final. Pour les mots infinis, plusieurs modes de reconnaissance ont été proposés. Nous introduisons ici les automates de Büchi, les automates de Muller, les automates de Rabin et finalement les automates à tables de transitions.

Un *automate* sans condition d'acceptation est un quadruplet  $\mathcal{A} = (Q, A, E, I)$ , où  $Q$  est un ensemble fini d'états,  $A$  est un alphabet,  $E$  un sous-ensemble de  $Q \times A \times Q$  appelé l'ensemble des *transitions* et  $I \subset Q$  est l'ensemble des états *initiaux*. Un automate est généralement représenté par un diagramme. Les états sont alors représentés par des cercles, les transitions par des arcs étiquetés et les états initiaux sont marqués par des flèches entrantes.

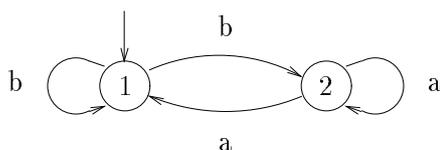


Figure 1.1 : Un automate

**Exemple :** L'automate  $\mathcal{A} = (Q, A, E, I)$  où

$$Q = \{1, 2\}, A = \{a, b\}, E = \{(1, b, 1), (1, b, 2), (2, a, 1), (2, a, 2)\} \text{ et } I = \{1\}$$

est représenté à la figure 1.1

Deux transitions  $(p_1, a_1, q_1)$  et  $(p_2, a_2, q_2)$  sont dites *consécutives* si  $q_1 = p_2$ . Un *chemin* de l'automate  $\mathcal{A}$  est une suite finie de transitions consécutives

$$f_1 = (q_0, a_1, q_1), f_2 = (q_1, a_2, q_2), \dots, f_n = (q_{n-1}, a_n, q_n)$$

notée également

$$q_0 \xrightarrow{a_1} q_1 \xrightarrow{a_2} q_2 \cdots q_{n-1} \xrightarrow{a_n} q_n \quad \text{ou plus simplement} \quad q_0 \xrightarrow{a_1 \dots a_n} q_n$$

L'état  $q_0$  est *l'origine* du chemin et l'état  $q_n$  son *extrémité* et on dit que le chemin passe par les états  $q_0, q_1, \dots, q_n$ . Le mot  $a_1 a_2 \dots a_n$  est *l'étiquette* du chemin et l'entier  $n$  sa *longueur*. Le *contenu* du chemin est l'ensemble  $\{f_1, \dots, f_n\}$  des transitions qui constituent le chemin. Par extension, deux chemins  $c_1$  et  $c_2$  sont

dits *consécutifs* si l'extrémité de  $c_1$  coïncide avec l'origine de  $c_2$ . Ils seront dits *coterminaux* s'ils ont même origine et même extrémité.

Un *chemin infini* dans l'automate  $\mathcal{A}$  est une suite infinie  $c$  de transitions consécutives

$$f_1 = (q_0, a_1, q_1), f_2 = (q_1, a_2, q_2), \dots$$

notée également

$$q_0 \xrightarrow{a_1} q_1 \xrightarrow{a_2} q_2 \dots$$

L'état  $q_0$  est *l'origine* du chemin infini et le mot infini  $a_1 a_2 \dots$  est son étiquette. On dit que le chemin  $c$  *passé infiniment souvent* par un état  $q$  (ou que  $q$  est *infiniment répété* dans  $c$ ) s'il existe une infinité d'entiers  $n$  tels que  $q_n = q$ . L'ensemble des états infiniment répétés de  $c$  est noté  $\text{Inr}(c)$ .

Un état  $q$  d'un automate  $\mathcal{A} = (Q, A, E, I)$  est dit *accessible* s'il existe un chemin fini  $c$  d'extrémité  $q$  tel que l'origine du chemin soit un état initial.

Un automate  $\mathcal{A} = (Q, A, E, I)$  est dit *déterministe* si pour tout état  $p \in Q$  et toute lettre  $a \in A$ , il existe au plus un état  $q$  tel que  $(p, a, q)$  soit une transition et si l'ensemble des états initiaux est réduit à un unique état  $q_0$ . Dans ce cas, l'automate est noté  $(Q, A, E, q_0)$ . Si l'automate  $(Q, A, E, q_0)$  est déterministe, l'ensemble des transitions est le graphe d'une fonction partielle de  $Q \times A$  dans  $Q$ , appelée la *fonction de transition*. S'il existe, l'unique état  $q$  tel que  $(p, a, q)$  soit une transition est noté  $p.a$ . Puisque la donnée de l'ensemble  $E$  est équivalente à la donnée de la fonction de transition, l'automate est souvent défini par un quadruplet  $(Q, A, \cdot, q_0)$ . La fonction de transition d'un automate déterministe se prolonge en une fonction partielle de  $Q \times A^*$  dans  $Q$  en posant  $q.1 = q$  pour  $q \in Q$  et  $q.(ua) = (q.u).a$  pour tout mot  $u \in A^*$  et toute lettre  $a \in A$  si  $q.u$  et  $(q.u).a$  sont définis.

De façon duale, un automate  $\mathcal{A} = (Q, A, E, I)$  est dit *complet* si pour tout état  $q \in Q$  et toute lettre  $a \in A$ , il existe au moins un état  $p$  tel que  $(p, a, q)$  soit une transition. Si l'automate est déterministe et complet, la fonction de transition est partout définie sur  $Q \times A^*$ . La fonction ainsi prolongée définit une action à droite de  $A^*$  sur  $Q$ .

**Exemple :** L'automate de la figure 1.1 n'est ni déterministe ni complet. En revanche, l'automate de la figure 1.2 est déterministe et complet.

## Automates de Büchi

Pour reconnaître un langage avec un automate  $\mathcal{A} = (Q, A, E, I)$ , on se fixe un ensemble  $F \subset Q$  d'états *finaux*. Un *automate de Büchi* est donc un quintuplet  $\mathcal{A} = (Q, A, E, I, F)$ . Un chemin fini de l'automate de Büchi est dit *réussi* si

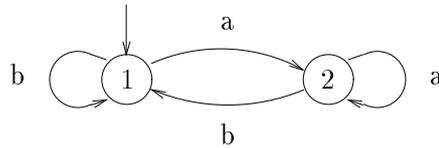


Figure 1.2 : Un automate déterministe et complet.

l'origine du chemin est un état initial et si l'extrémité du chemin est un état final. Un mot  $u$  est *reconnu* par l'automate s'il est l'étiquette d'un chemin réussi de l'automate. Le langage des mots finis reconnus par un automate de Büchi  $\mathcal{A}$  est noté  $L^*(\mathcal{A})$ . Un langage est dit *reconnaisable* s'il est l'ensemble des mots reconnus par un automate. On a alors le théorème

**Théorème 2 (Kleene)** *Un langage est reconnaissable si et seulement si il est rationnel.*

Pour les mots infinis, ce mode de reconnaissance n'a pas de sens puisque un chemin infini ne possède pas d'extrémité. On dit qu'un chemin infini  $c$  d'un automate de Büchi est *réussi* si l'origine du chemin est un état initial et si le chemin passe infiniment souvent par un état de  $F$ , *i.e.*,  $\text{Inr}(c) \cap F \neq \emptyset$ . Un mot infini  $u$  est *reconnu* par l'automate s'il est l'étiquette d'un chemin réussi de l'automate. L'ensemble des chemins réussis d'un automate de Büchi  $\mathcal{A}$  est noté  $|\mathcal{A}|$  et le  $\omega$ -langage des mots infinis reconnus est noté  $L^\omega(\mathcal{A})$ . Un  $\omega$ -langage  $X$  est dit *reconnaisable* s'il existe un automate  $\mathcal{A}$  tel que  $X = L^\omega(\mathcal{A})$ .

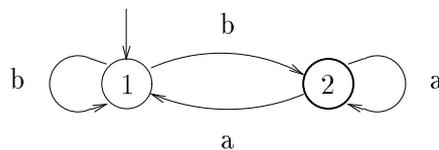


Figure 1.3 : Automate de Büchi.

**Exemple :** L'automate  $\mathcal{A}$  de la figure 1.3 avec  $F = \{2\}$  (L'état final est marqué en gras) reconnaît le langage des mots commençant par la lettre  $b$ . On a  $L^*(\mathcal{A}) = b(a+b)^*$ . Le  $\omega$ -langage reconnu par cet automate est égal au  $\omega$ -langage des mots commençant par la lettre  $b$  et comportant une infinité d'occurrences de la lettre  $a$  :  $L^\omega(\mathcal{A}) = b(b^*a)^\omega$ .

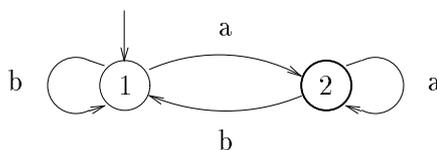


Figure 1.4 : Automate de Büchi déterministe et complet.

**Exemple :** L'automate  $\mathcal{A}$  de la figure 1.4 avec  $F = \{2\}$  (L'état final est marqué en gras) reconnaît le langage des mots se terminant par la lettre  $a$ . On a  $L^*(\mathcal{A}) = (a + b)^*a$ . Le  $\omega$ -langage reconnu par cet automate est égal au  $\omega$ -langage comportant une infinité d'occurrences de  $a$  :  $L^\omega(\mathcal{A}) = (b^*a)^\omega$ .

Nous avons l'analogie du théorème de Kleene pour les mots infinis.

**Théorème 3 (Büchi)** *Un  $\omega$ -langage est reconnaissable si et seulement si il est rationnel.*

On peut trouver une preuve de ce théorème dans [PP93a] ou [Eil72].

Nous allons ici établir un premier lien entre les propriétés topologiques et la reconnaissance par automate. Soit  $\mathcal{A} = (Q, A, E, I, F)$  un automate de Büchi. Un état  $q$  est dit *coaccessible* (pour le mode de reconnaissance des mots infinis) s'il existe un chemin infini  $c$  d'état initial  $q$  tel que  $\text{Inr}(c) \cap F \neq \emptyset$ . On note  $C^\omega(\mathcal{A})$  l'ensemble de ces états coaccessibles.

**Proposition 2** *Soit  $\mathcal{A} = (Q, A, E, I, F)$  un automate de Büchi reconnaissant un  $\omega$ -langage  $X$ . L'automate  $(Q, A, E, I, C^\omega(\mathcal{A}))$  reconnaît alors la fermeture topologique  $\overline{X}$  du langage  $X$ .*

De plus, on peut toujours supprimer les états non coaccessibles d'un automate. L'automate n'est alors plus nécessairement complet mais le  $\omega$ -langage reconnu reste le même.

**Démonstration :** Soit un mot  $u$  reconnu par l'automate  $(Q, A, E, I, C^\omega(\mathcal{A}))$ . Il existe un chemin  $q_0 \xrightarrow{u_0} q \xrightarrow{u_1} q \xrightarrow{u_2} q \dots$  où  $q_0$  est un état initial et  $q$  un état de  $C^\omega(\mathcal{A})$ . Puisque l'état  $q$  est coaccessible, il existe un chemin infini  $c$  d'étiquette un mot infini  $v$  tel que  $\text{Inr}(c) \cap F \neq \emptyset$ . Pour tout  $k \in \mathbb{N}$  le mot  $x_k = u_0 \dots u_k v$  appartient à  $X$  et coïncide avec  $u$  sur un préfixe de longueur au moins  $k$ , ce qui montre que  $u \in \overline{X}$ .

Réciproquement, soit un mot  $u \in \overline{X}$ . Il existe une suite de mots  $x_k$  tels que  $x_k$  et  $u$  coïncident sur un préfixe de longueur au moins  $k$ . Pour chaque

mot  $x_k$ , il existe un chemin réussi  $c_k$ . On note  $q_{k,0}q_{k,1}q_{k,2} \dots$  la suite des états du chemin  $c_k$ . On peut remarquer que tous les états  $q_{k,l}$  pour  $k, l \in \mathbb{N}$  sont coaccessibles. On va alors procéder par extraction diagonale pour exhiber un chemin réussi d'étiquette  $u$ . Puisque le nombre d'états de l'automate est fini, il existe une suite d'entiers  $k\varphi_0$  telle que  $q_{k\varphi_0,0}$  soit constant. On extrait alors une suite d'entiers  $k\varphi_1\varphi_0$  de la suite  $k\varphi_0$  telle que  $q_{k\varphi_1\varphi_0,1}$  soit constant. On itère le procédé et l'on pose  $q_k = q_{k\varphi_k \dots \varphi_0, k}$ . La suite  $q_k$  constitue la suite d'états d'un chemin réussi de l'automate  $(Q, A, E, I, C^\omega(\mathcal{A}))$  puisque que tous les états  $q_{k,l}$  sont coaccessibles.  $\square$

La proposition précédente a le corollaire suivant

**Corollaire 1** *Pour tout  $\omega$ -langage reconnaissable  $X$ , la fermeture  $\overline{X}$  de  $X$  est également reconnaissable.*

Il est bien connu que pour les mots finis les automates de Büchi déterministes et les automates de Büchi non déterministes sont équivalents. Tout langage reconnaissable peut être reconnu par un automate de Büchi déterministe. Ce résultat ne s'étend pas aux mots infinis. Si l'alphabet contient au moins deux lettres, il existe des  $\omega$ -langages reconnaissables qui ne sont reconnus par aucun automate de Büchi déterministe. Par abus de notation, un  $\omega$ -langage est dit *déterministe* s'il est reconnu par automate de Büchi déterministe. Afin de caractériser les  $\omega$ -langages déterministes, nous introduisons un nouvel opérateur. Pour  $X \subset A^*$ , on pose

$$\overrightarrow{X} = \{u \in A^\omega \mid u \text{ a une infinité de préfixes dans } X\}$$

Nous commençons par donner une caractérisation topologique des  $\omega$ -langages  $X$  qui peuvent s'écrire  $\overrightarrow{L}$  avec  $L \subset A^*$ .

**Théorème 4** (Landweber) *Un  $\omega$ -langage  $X$  peut s'écrire  $\overrightarrow{L}$  avec  $L \subset A^*$  si et seulement si  $X$  est un langage  $G_\delta$ .*

**Démonstration :** On suppose que  $X = \overrightarrow{L}$  et on pose  $L_n = \{v \in L \mid |v| \geq n\}$ . Tout mot infini  $u$  de  $\overrightarrow{L}$  possède un préfixe dans  $L$  de longueur supérieure à  $n$  et donc un préfixe dans  $L_n$ . Le mot  $u$  appartient donc à  $L_n A^\omega$  pour tout  $n$  et à l'intersection dénombrable d'ouverts  $\bigcap_{n \geq 0} L_n A^\omega$ . Réciproquement tout mot de cette intersection possède des préfixes dans  $L$  de longueur arbitraire et appartient par conséquent au langage  $X = \overrightarrow{L}$ . Ceci montre que tout  $\omega$ -langage de la forme  $\overrightarrow{L}$  est un  $G_\delta$ .

Réciproquement, on suppose que  $X$  s'écrit  $\bigcap_{n \geq 0} U_n$  où les  $U_n$  sont des ouverts. En posant  $V_n = \bigcap_{k=0}^n U_k$ , on a  $X = \bigcap_{n \geq 0} V_n$  et les  $V_k$  forment une suite

décroissante d'ouverts. On écrit  $V_n = L_n A^\omega$  où  $L_n$  est un langage préfixe. On pose alors  $L = \bigcup_{n \geq 0} L_n A^n$ . Vérifions que  $X = \overrightarrow{L}$ . Si  $u$  appartient à  $X$ , le mot  $u$  possède un préfixe  $p_n$  dans chacun des langages  $L_n A^n$  et donc une infinité de préfixes dans  $L$  puisque la longueur de  $p_n$  est supérieure à  $n$ . Réciproquement, si  $u$  possède une suite  $p_n$  de préfixes dans  $L$  de longueur croissante, ces préfixes appartiennent nécessairement à des langages  $L_i A^i$  distincts puisque ces langages sont préfixes. Le mot  $u$  appartient donc à un infinité d'ouverts  $V_i$  et donc à l'intersection  $\bigcap_{n \geq 0} V_n$  puisque ces ouverts forment une suite décroissante.  $\square$

Cet opérateur permet d'établir le lien entre  $L^*(\mathcal{A})$  et  $L^\omega(\mathcal{A})$  pour un automate de Büchi déterministe  $\mathcal{A}$ . On a la proposition

**Proposition 3** *Pour tout automate de Büchi déterministe  $\mathcal{A} = (Q, A, E, q_0, F)$  on a*

$$L^\omega(\mathcal{A}) = \overrightarrow{L^*(\mathcal{A})}.$$

**Démonstration :** Soit un mot infini  $u$  reconnu par l'automate  $(Q, A, E, q_0, F)$ . Il existe un chemin  $q_0 \xrightarrow{u_0} q \xrightarrow{u_1} q \xrightarrow{u_2} q \cdots$  où  $q$  appartient à  $F$ . Pour tout  $k \in \mathbb{N}$ , le mot fini  $v_k = u_0 \dots u_k$  est un préfixe de  $u$  qui appartient à  $L^*(\mathcal{A})$  ce qui montre l'inclusion  $L^\omega(\mathcal{A}) \subset \overrightarrow{L^*(\mathcal{A})}$ .

Réciproquement, si  $u$  admet une suite de préfixes  $v_k$  de longueur croissante appartenant à  $L^*(\mathcal{A})$ , ceci définit un chemin infini puisque l'automate est déterministe. Ce chemin est bien sûr réussi.  $\square$

Cette dernière proposition permet de démontrer le résultat bien connu.

**Théorème 5** *Pour un  $\omega$ -langage  $X$ , les conditions suivantes sont équivalentes*

- (1)  $X$  est un  $\omega$ -langage déterministe;
- (2) il existe un langage reconnaissable  $L \subset A^*$  tel que  $X = \overrightarrow{L}$ .

**Démonstration :** L'implication (1)  $\Rightarrow$  (2) découle directement de la proposition précédente. Réciproquement, tout langage de mots finis est reconnu par un automate de Büchi déterministe.  $\square$

## Automates de Rabin

Comme tous les  $\omega$ -langages ne sont pas reconnus par un automate de Büchi déterministe, nous introduisons une extension des automates de Büchi. Un *automate de Rabin* est un quintuplet  $\mathcal{A} = (Q, A, E, q_0, \mathcal{R})$  où  $(Q, A, E, q_0)$  est

un automate déterministe, et  $\mathcal{R} = (R_1, R_2, \dots, R_m)$  est une suite d'ensembles d'états. Afin de simplifier les notations, nous adoptons la convention de poser  $R_0 = Q$ . Pour un ensemble d'états  $R \subset Q$ , on définit  $|R|$  comme l'ensemble  $|\mathcal{A}_R|$  des chemins réussis de l'automate de Büchi associé  $\mathcal{A}_R = (Q, A, E, q_0, R)$ . En particulier,  $|R_0|$  est l'ensemble de tous les chemins de  $\mathcal{A}$ . De façon analogue, on définit  $L^\omega(R)$  comme l'ensemble des étiquettes des chemins de  $|R|$ . Si de plus l'automate  $\mathcal{A}$  est complet, le langage  $L^\omega(R_0)$  est égal à  $A^\omega$ . L'ensemble des chemins réussis de l'automate de Rabin est défini par l'expression

$$|\mathcal{A}| = (|R_1| - |R_2|) + (|R_3| - |R_4|) + \dots$$

Un automate de Büchi déterministe est un cas particulier d'un automate de Rabin. Un automate de Büchi déterministe  $\mathcal{A} = (Q, A, E, q_0, F)$  est équivalent à l'automate de Rabin  $\mathcal{A}' = (Q, A, E, q_0, \mathcal{R})$  où la suite  $\mathcal{R}$  est réduite au seul élément  $F$ .

La suite  $\mathcal{R}$  est habituellement (cf. [PP93a]) définie comme une famille de couples d'ensembles d'états  $\mathcal{R} = \{(L_j, U_j) \mid 1 \leq j \leq n\}$ . Un chemin est alors dit réussi s'il existe un indice  $j$  tel que  $1 \leq j \leq n$ ,  $p$  passe infiniment souvent par  $L_j$  et seulement un nombre fini de fois par  $U_j$ . Cela revient à dire que

$$\text{Inr}(p) \cap L_j \neq \emptyset \text{ et } \text{Inr}(p) \cap U_j = \emptyset.$$

Les deux définitions sont équivalentes. On passe naturellement de la famille de couples  $\{(L_1, U_1), (L_2, U_2), \dots\}$  à la suite  $(R_1, R_2, \dots)$  en posant  $R_1 = L_1$ ,  $R_2 = U_1$ ,  $R_3 = L_2$ ,  $R_4 = U_2, \dots$  et réciproquement en posant éventuellement  $U_m = \emptyset$  si le nombre de parties  $R_i$  est impair égal à  $2m - 1$ .

Un mot infini est reconnu par l'automate  $\mathcal{A}$  s'il est l'étiquette d'un chemin réussi de  $\mathcal{A}$ . L'ensemble de mots reconnus est égal à l'expression

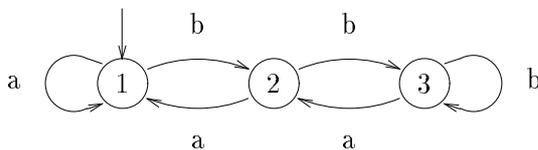
$$L^\omega(\mathcal{A}) = (L^\omega(R_1) - L^\omega(R_2)) + (L^\omega(R_3) - L^\omega(R_4)) + \dots$$

Le nombre  $m$  de parties  $R_i$  de  $\mathcal{R}$  est par définition l'*indice de Rabin* de l'automate. L'*indice de Rabin* d'un  $\omega$ -langage est le plus petit indice de Rabin d'un automate de Rabin reconnaissant ce langage.

**Exemple :** L'automate de Rabin de la figure 1.5 avec la suite  $\mathcal{R} = \{\{1, 3\}, \{2\}\}$  reconnaît le langage  $X = (a + b)^*(a^\omega + b^\omega)$ .

Le théorème suivant énonce que le mode de reconnaissance des automates de Rabin est équivalent à celui des automates de Büchi.

**Théorème 6 (McNaughton)** *Un  $\omega$ -langage est reconnaissable si et seulement si il est reconnu par un automate de Rabin.*

Figure 1.5 : Automate de Rabin :  $\mathcal{R} = \{\{1, 3\}, \{2\}\}$ 

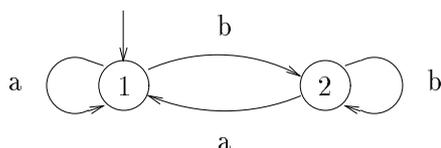
Nous ne démontrons pas ce théorème essentiellement dû à McNaughton [McN66]. La preuve de ce théorème est difficile et peut être obtenue de plusieurs façons. La plus récente et la plus directe est due à Safra [Saf88]. On peut aussi se référer à [PP93a] ou à [Eil72].

### Automates de Muller

Nous introduisons maintenant les automates de Muller ou automates à table. Un *automate de Muller* ou *automate à table* est un quintuplet  $\mathcal{A} = (Q, A, E, q_0, \mathcal{T})$  où  $(Q, A, E, q_0)$  est un automate déterministe, et  $\mathcal{T}$  un ensemble de parties de  $Q$ . Un chemin infini  $c$  est dit *réussi* si l'origine du chemin est l'état initial et si  $\text{Inr}(c) \in \mathcal{T}$ , autrement dit si l'ensemble des états infiniment répétés est un élément de la table  $\mathcal{T}$ . L'ensemble des mots reconnus par l'automate est l'ensemble des étiquettes des chemins réussis et il est noté  $L^\omega(\mathcal{A})$ .

Un automate de Rabin est un cas particulier d'un automate de Muller. Un automate de Rabin  $\mathcal{A} = (Q, A, E, q_0, \mathcal{R})$  avec  $\mathcal{R} = (R_1, \dots, R_m)$  est équivalent à l'automate de Muller  $\mathcal{A}' = (Q, A, E, q_0, \mathcal{T})$  où la table  $\mathcal{T}$  est définie par

$$\mathcal{T} = \{T \subset Q \mid \exists k \ T \cap R_{2k-1} \neq \emptyset \text{ et } T \cap R_{2k} = \emptyset\}$$

Figure 1.6 : Automate de Muller :  $\mathcal{T} = \{\{1\}, \{2\}\}$ 

**Exemple :** L'automate de Muller de la figure 1.6 avec la table  $\mathcal{T} = \{\{1\}, \{2\}\}$  reconnaît également le langage  $X = (a + b)^*(a^\omega + b^\omega)$ .

Le théorème suivant énonce que le mode de reconnaissance des automates de Muller est équivalent à celui des automates de Büchi.

**Théorème 7** *Un  $\omega$ -langage est reconnaissable si et seulement si il est reconnu par un automate de Muller.*

Ce théorème se déduit du précédent.

**Lemme 2** *Soit  $\mathcal{A} = (Q, A, E, q_0)$  un automate déterministe et  $R \subset Q$  une partie de  $Q$ . Les propositions suivantes sont équivalentes*

- (1) *il existe un chemin infini  $c$  d'origine  $q_0$  tel que  $\text{Inr}(c) = R$ ;*
- (2) *il existe un état accessible  $q \in R$  et un chemin  $q \xrightarrow{w} q$  qui passe par tous les états de  $R$  et par aucun autre état;*
- (3) *pour tout  $q \in R$ , l'état  $q$  est accessible et il existe un chemin  $q \xrightarrow{w} q$  qui passe par tous les états de  $R$  et par aucun autre état.*

*Une tel ensemble d'état est appelé un ensemble répétable ou une partie répétable et l'on dit que le chemin  $q \xrightarrow{w} q$  définit cette partie répétable.*

**Démonstration :** On suppose qu'il existe un chemin infini  $c$  d'origine  $q_0$  tel que  $\text{Inr}(c) = R$ . On note  $q_0q_1q_2 \dots$  la suite des états du chemin  $c$ . Soit  $q$  un état de  $R$ . Puisque  $q$  apparaît dans  $c$ , il existe un chemin  $q_0 \xrightarrow{w_0} q$  et  $q$  est accessible. Puisque  $\text{Inr}(c) = R$ , il existe un entier  $k_0$  tel que pour tout  $k \geq k_0$ , l'état  $q_k$  appartient à  $R$ . Il existe un entier  $k_1 \geq k_0$  tel que  $q = q_{k_1}$  puisque  $q$  apparaît infiniment souvent dans le chemin  $c$ . Pour la même raison, il existe pour chaque état  $p \in R$  un entier  $k_p \geq k_1$  tel que  $p = q_{k_p}$ . On pose  $k_2 = \max\{k_p \mid p \in R\}$ . Il existe  $k_3 \geq k_2$  tel que  $q = q_{k_3}$ . Le chemin  $q_{k_1} \xrightarrow{w} q_{k_3}$  convient. Tous les états de  $R$  apparaissent car pour  $p \in R$  il existe  $k_1 \leq k_p \leq k_3$  tel que  $p = q_{k_p}$  et tous les états apparaissant sont dans  $R$ . Ceci montre l'implication (1)  $\Rightarrow$  (3).

L'implication (3)  $\Rightarrow$  (2) est évidente.

On suppose que  $q$  est accessible. Il existe alors un chemin fini  $q_0 \dots q_n$  tel que  $q = q_n$ . Il existe d'après l'hypothèse un chemin fini  $p_1 \dots p_m$  tel que  $p_1 = p_m = q$ , et tel que tous les états de  $R$  et uniquement les états de  $R$  apparaissent dans ce chemin. Le chemin infini  $c = q_0 \dots q_n (p_1 \dots p_m)^\omega$  vérifie alors  $\text{Inr}(c) = R$  ce qui montre (2)  $\Rightarrow$  (1).  $\square$

Nous allons maintenant établir un lien entre les automates de Büchi déterministes et les automates de Muller. Un automate de Muller  $(Q, A, E, q_0, \mathcal{T})$  est

dit à *table pleine* si et seulement si pour toutes parties répétibles  $R, S \subset Q$ , on a l'implication

$$\left. \begin{array}{l} R \subset S \\ R \in \mathcal{T} \end{array} \right\} \Rightarrow S \in \mathcal{T}.$$

**Théorème 8** (Landweber) *Tout automate de Büchi déterministe est équivalent à un automate de Muller à table pleine et tout automate de Muller à table pleine est équivalent à un automate de Büchi déterministe.*

**Démonstration :** Soit  $\mathcal{B} = (Q, A, E, I, F)$  un automate de Büchi déterministe. L'automate de Muller  $\mathcal{A} = (Q, A, E, I, \mathcal{T})$  où la table  $\mathcal{T}$  est déterminée par

$$\mathcal{T} = \{R \subset Q \mid R \cap F \neq \emptyset\}$$

est à table pleine et il est équivalent à l'automate  $\mathcal{A}$ .

Réciproquement, soit  $\mathcal{A} = (Q, A, E, q_0, \mathcal{T})$  un automate de Muller à table pleine. On définit un automate de Büchi déterministe

$$\mathcal{B} = (\mathcal{P}(Q) \times Q, A, \cdot, (\emptyset, q_0), F)$$

où l'ensemble  $F$  des états finaux est donné par  $F = \{(\emptyset, q) \mid q \in Q\}$  et où pour tout  $(P, q) \in \mathcal{P}(Q) \times Q$  et toute lettre  $a \in A$ , la fonction de transition est définie par

$$(P, q).a = \begin{cases} (\emptyset, q.a) & \text{si } P + \{q.a\} \text{ contient un élément de } \mathcal{T} \\ (P + \{q.a\}, q.a) & \text{sinon.} \end{cases}$$

Soit  $u$  un mot reconnu par l'automate  $\mathcal{B}$ . Le mot  $u$  définit dans l'automate  $\mathcal{B}$  un chemin  $c$  qui passe infiniment souvent par un état final. Il existe donc un ensemble répétable  $R \in \mathcal{T}$  dont tous les états sont infiniment répétés. Par conséquent  $R \subset \text{Inr}(c)$  et puisque la table de  $\mathcal{A}$  est pleine, on a aussi  $\text{Inr}(c) \in \mathcal{T}$ . Le mot  $u$  est reconnu par  $\mathcal{A}$ .

Réciproquement, soit  $u$  un mot reconnu par l'automate  $\mathcal{A}$ . Le mot  $u$  définit dans l'automate  $\mathcal{B}$  un chemin  $c$  tel que  $\text{Inr}(c) \in \mathcal{T}$ . On note  $q_0q_1q_2\dots$  la suite des états de  $c$ . Pour tout entier  $k$  assez grand, il existe  $p$  tel que  $R \subset \{q_k \dots q_{k+p}\}$ . Le chemin de l'automate  $\mathcal{B}$  passe donc par un état final infiniment souvent et reconnaît le mot  $u$ .  $\square$

Grâce aux théorèmes 4 et 5 et au théorème précédent, on a, pour un  $\omega$ -langage reconnaissable  $X$ , l'équivalence entre les propriétés suivantes

- (1)  $X$  est reconnu par un automate de Büchi déterministe;

- (2)  $X$  est reconnu par un automate de Muller à table pleine;
- (3)  $X$  peut s'écrire  $X = \overrightarrow{L}$  avec  $L \subset A^*$ ;
- (4)  $X$  est un langage  $G_\delta$ .

Nous verrons plus loin que le théorème 8 est un cas particulier de la théorie développée par Wagner [Wag79] et que les langages  $G_\delta$  constituent une des premières classes de la hiérarchie étudiée plus loin.

### Automates à table de transitions

Nous introduisons maintenant les automates à table de transitions. Il s'agit d'une variante des automates de Muller introduite par B. Le Saec [Sae90]. Un *automate à table de transitions* est un quintuplet  $\mathcal{A} = (Q, A, E, q_0, \mathcal{T})$  où  $(Q, A, E, q_0)$  est un automate déterministe et  $\mathcal{T}$  un ensemble de parties de  $E$ . Un chemin infini  $c$  est dit *réussi* si l'origine du chemin est l'état initial et si l'ensemble des transitions apparaissant infiniment souvent dans le chemin est un élément de la table  $\mathcal{T}$ . L'ensemble des mots reconnus par l'automate est l'ensemble des étiquettes des chemins réussis et il est noté  $L^\omega(\mathcal{A})$ . Un automate de Muller est un cas particulier d'un automate à table de transitions. Réciproquement, il est aisé à partir d'un automate à table de transitions de construire un automate de Muller équivalent (cf. [SPW91b]).

## Chapitre 2

# Semigroupes et $\omega$ -semigroupes

### 2.1 Introduction

Pour les langages de mots finis, la reconnaissance par automate et la reconnaissance par morphisme de semigroupe sont deux notions équivalentes. Étant donné un automate fini reconnaissant un langage, il est aisé de construire un semigroupe en considérant le semigroupe des relations binaires sur l'ensemble des états de l'automate. Réciproquement, un semigroupe agit sur lui-même à droite par multiplication et peut donc être considéré comme un automate déterministe.

Nous allons voir qu'il existe des notions correspondantes pour les langages de mots infinis. De manière intuitive, les semigroupes sont particulièrement adaptés à l'étude des langages de mots finis parce que la structure algébrique naturelle de l'ensemble  $A^*$  des mots finis est la structure de semigroupe. De façon surprenante, l'ensemble  $A^\omega$  ne possède pas de structure algébrique naturelle. Par contre, si l'on adjoint les mots finis aux mots infinis, l'ensemble  $A^\infty = A^* + A^\omega$  admet plusieurs opérations naturelles. D'une part,  $A^*$  est muni d'une structure de semigroupe, et ce semigroupe agit à gauche sur l'ensemble  $A^\omega$  par concaténation d'un mot fini et d'un mot infini. D'autre part, le semigroupe  $A^*$  est muni d'un produit infini. Étant donné une suite  $u_0, u_1, u_2, \dots$  de mots finis, le produit  $u_0 u_1 u_2 \dots$  définit un mot infini. Ce produit infini induit également l'opération  $\omega$  de  $A^*$  dans  $A^\omega$  qui à un mot fini  $u$  associe le produit infini  $uuu \dots = u^\omega$ . Réciproquement, il n'est pas possible de définir le produit infini à partir de l'opération  $\omega$  même en utilisant la concaténation à gauche de mots finis. En effet, il existe dans  $A^\omega$  des mots non ultimement périodiques. Ceci conduit à introduire deux nouvelles structures algébriques, les  $\omega$ -semigroupes d'une part qui sont des semigroupes munis d'un produit infini et les *algèbres de Wilke* d'autre part qui sont des semigroupes munis d'une opération  $\omega$ . Les algèbres de Wilke furent introduites par T. Wilke [Wil91], les  $\omega$ -semigroupes par J.E. Pin [PP93b]. Nous verrons que ces deux

notions coïncident dans le cas fini.

## 2.2 Semigroupes

Nous commençons par quelques définitions et rappels sur les semigroupes, en particulier sur les semigroupes finis. Nous ne donnons pas les démonstrations de ces résultats classiques. Celles-ci peuvent être trouvées dans [How76] et [Lal79].

Un *semigroupe* est un ensemble appelé *support* muni d'une opération (ou loi de composition interne) associative mais ne possédant pas nécessairement d'élément neutre. On adopte le plus souvent la notation multiplicative. Le produit de deux éléments  $s$  et  $t$  est noté  $s.t$  ou plus simplement  $st$ . Le langage  $A^+$  des mots finis non vides sur un alphabet  $A$  muni de la concaténation comme produit est un semigroupe.

Un *zéro* d'un semigroupe  $S$  est un élément  $0$  tel que  $0s = s0 = 0$  pour tout  $s$  dans  $S$ . S'il existe, le zéro d'un semigroupe est unique et est noté  $0$ . Il est toujours possible d'ajouter un nouvel élément à un semigroupe en le considérant comme un zéro. Si  $S$  est un semigroupe, on note  $S^0$  le semigroupe de support  $S \cup \{0\}$ . La loi de  $S$  est prolongée à  $S \cup \{0\}$  en posant  $0s = s0 = 0$  pour tout  $s$  dans  $S^0$ .

Un *monoïde* est un semigroupe muni d'un élément neutre pour la loi de composition, noté  $1$  le plus souvent. Le langage  $A^*$  des mots sur un alphabet  $A$  est un monoïde dont l'élément neutre est le mot vide.

Si  $S$  est un semigroupe, on note  $S^1$  le monoïde égal à  $S$  si  $S$  est un monoïde et égal à  $S \cup \{1\}$  si  $S$  ne possède pas d'élément neutre, l'opération de  $S$  étant étendue en posant  $1s = s1 = s$  pour tout  $s$  dans  $S^1$ .

De façon générale, un morphisme entre deux structures algébriques est une application qui préserve les opérations. En particulier, un *morphisme de semigroupe* est une application  $\varphi$  d'un semigroupe  $S$  dans un semigroupe  $T$  telle que pour tous  $s_1, s_2 \in S$ , on ait

$$(s_1s_2)\varphi = (s_1\varphi)(s_2\varphi).$$

On adopte la convention de noter  $s_1s_2\varphi$  l'élément  $(s_1s_2)\varphi$ .

Un *idempotent* d'un semigroupe est un élément  $e$  tel que  $e^2 = e$ . L'ensemble des idempotents du semigroupe  $S$  est noté  $E(S)$ . On a les résultats classiques suivants pour les semigroupes finis.

**Proposition 4** *Tout élément d'un semigroupe fini admet une puissance idempotente.*

**Proposition 5** *Pour tout semigroupe fini  $S$ , il existe un entier noté  $\pi$  tel que pour tout élément  $s$  de  $S$ , l'élément  $s^\pi$  soit un idempotent.*

Nous définissons maintenant les relations de Green. Pour deux éléments  $s_1$  et  $s_2$  d'un semigroupe  $S$ , on définit

$$\begin{aligned}
s_1 \leq_{\mathcal{J}} s_2 &\Leftrightarrow S^1 s_1 S^1 \subset S^1 s_2 S^1 \\
s_1 \leq_{\mathcal{R}} s_2 &\Leftrightarrow s_1 S^1 \subset s_2 S^1 \\
s_1 \leq_{\mathcal{L}} s_2 &\Leftrightarrow S^1 s_1 \subset S^1 s_2 \\
s_1 \leq_{\mathcal{H}} s_2 &\Leftrightarrow s_1 S^1 \subset s_2 S^1 \text{ et } S^1 s_1 \subset S^1 s_2 \\
s_1 \mathcal{J} s_2 &\Leftrightarrow s_1 \leq_{\mathcal{J}} s_2 \text{ et } s_2 \leq_{\mathcal{J}} s_1 \\
s_1 \mathcal{R} s_2 &\Leftrightarrow s_1 \leq_{\mathcal{R}} s_2 \text{ et } s_2 \leq_{\mathcal{R}} s_1 \\
s_1 \mathcal{L} s_2 &\Leftrightarrow s_1 \leq_{\mathcal{L}} s_2 \text{ et } s_2 \leq_{\mathcal{L}} s_1 \\
s_1 \mathcal{H} s_2 &\Leftrightarrow s_1 \leq_{\mathcal{H}} s_2 \text{ et } s_2 \leq_{\mathcal{H}} s_1
\end{aligned}$$

**Proposition 6** *Les relations  $\mathcal{R}$  et  $\mathcal{L}$  commutent.*

On note  $\mathcal{D}$  la relation d'équivalence  $\mathcal{D} = \mathcal{R}\mathcal{L} = \mathcal{L}\mathcal{R}$ . Pour un élément  $s$  d'un semigroupe  $S$ , la classe d'équivalence de  $s$  pour la relation  $\mathcal{R}$ ,  $\mathcal{L}$ ,  $\mathcal{D}$ , et  $\mathcal{H}$  est respectivement appelée une  $\mathcal{R}$ -classe,  $\mathcal{L}$ -classe,  $\mathcal{D}$ -classe, et une  $\mathcal{H}$ -classe et cette classe est notée  $R(s)$ ,  $L(s)$ ,  $D(s)$  et  $H(s)$ . Pour tout  $s \in S$ , on a, par définition,  $H(s) = R(s) \cap L(s) \subset D(s)$ . La proposition suivante décrit la structure des  $\mathcal{D}$ -classes.

**Proposition 7** *Soient  $D$  une  $\mathcal{D}$ -classe d'un semigroupe  $S$  et  $s_1$  et  $s_2$  deux éléments de  $S$ .*

- (1) *Si  $s_1 \mathcal{R} s_2$ , et si  $s_1 = s_2 x$  et  $s_2 = s_1 y$ , les applications  $s \mapsto s x$  et  $s \mapsto s y$  induisent des bijections réciproques l'une de l'autre de  $L(s_1)$  sur  $L(s_2)$  et ces bijections préservent les  $\mathcal{H}$ -classes.*
- (2) *On a  $s_1 s_2 \in R(s_1) \cap L(s_2)$  si et seulement si  $R(s_2) \cap L(s_1)$  contient un idempotent.*
- (3) *Une  $\mathcal{H}$ -classes est un groupe si et seulement si elle contient le produit de deux de ses éléments.*
- (4) *Si la  $\mathcal{D}$ -classe  $D$  contient un idempotent, elle en contient au moins un dans chaque  $\mathcal{R}$ -classe et dans chaque  $\mathcal{L}$ -classe.*

**Proposition 8** Soit  $S$  un semigroupe fini

- (1) Soient  $s \in S$ ,  $s_1, s_2 \in S^1$ . Si  $s = s_1 s s_2$  alors  $s_1 s \in L(s)$  et  $s s_2 \in R(s)$ .
- (2) Soient  $s_1, s_2 \in S$ . Alors  $s_1 \mathcal{D} s_2$  si et seulement si  $s_1 \mathcal{J} s_2$ .

## 2.3 $\omega$ -semigroupes

Une *action à gauche* d'un semigroupe  $S$  sur un ensemble  $E$  est une application  $S \times E \rightarrow E$  qui à  $(s, e) \in S \times E$  associe  $s.e \in E$  telle que pour tous  $s_1, s_2 \in S$  et  $e \in E$

$$s_1.(s_2.e) = (s_1 s_2).e$$

On s'autorise à écrire  $se$  à la place  $s.e$  lorsque cela n'introduit pas une ambiguïté.

Un  $\omega$ -semigroupe est une algèbre  $S = (S_f, S_i)$  munie des opérations suivantes

- une loi de composition interne associative sur  $S_f$ ;
- une action à gauche de  $S_f$  sur  $S_i$ ;
- une application  $\pi : S_f^{\mathbb{N}} \rightarrow S_i$  appelée *produit infini*.

Le produit infini est en outre compatible avec la loi de composition interne et l'action à gauche. Pour toute suite croissante d'entiers  $(k_n)_{n \geq 1}$  et pour toute suite  $(s_n)_{n \geq 0} \in S_f^{\mathbb{N}}$  et pour tout  $s \in S_f$  on a

$$\pi((s_0 \dots s_{k_1-1}), (s_{k_1} \dots s_{k_2-1}), \dots) = \pi(s_0, s_1, s_2, \dots, s_{k_1-1}, s_{k_1}, \dots) \quad (1)$$

$$s.\pi(s_0, s_1, s_2, \dots) = \pi(s, s_0, s_1, s_2, \dots) \quad (2)$$

Ces dernières conditions montrent que l'on peut sans ambiguïté remplacer la notation  $\pi(s_0, s_1, s_2, \dots)$  par  $s_0 s_1 s_2 \dots$  puisque la valeur du produit infini ne dépend pas de la factorisation choisie. Nous utiliserons donc cette notation dans la suite. Les ensembles  $S_f$  et  $S_i$  sont respectivement appelés la partie *finitaire* et la partie *infinitaire* du  $\omega$ -semigroupe  $(S_f, S_i)$ .

**Exemple :** Le couple  $(A^*, A^\omega)$  muni des opérations habituelles est bien sûr un  $\omega$ -semigroupe. On écrira  $A^\infty$  pour désigner cet  $\omega$ -semigroupe.

**Exemple :** On considère l'intervalle  $I = [0; 1] \subset \mathbb{R}$  muni de la multiplication des réels. Pour  $(x_i)_{i \in \mathbb{N}} \in I^{\mathbb{N}}$ , une suite d'éléments de  $I$ , la suite  $X_n = \prod_{i=0}^n x_i$  est une suite décroissante de l'intervalle  $I$  et converge donc dans  $I$ . Il est possible de définir un produit infini sur  $I$  en posant  $\prod_{i \geq 0} x_i = \lim_{n \rightarrow \infty} X_n$ . Le couple  $(I, I)$  constitue alors un  $\omega$ -semigroupe.

Dans ce dernier exemple, le produit infini est la limite des produits partiels. Cela n'est pas toujours les cas comme le montre l'exemple suivant.

**Exemple :** On munit le semigroupe multiplicatif  $S = \{0, 1\}$  d'un produit infini constant égal à 0. Le couple  $(S, S)$  constitue bien un  $\omega$ -semigroupe mais  $1^\omega = 0$ .

Si  $S = (S_f, S_i)$  et  $T = (T_f, T_i)$  sont deux  $\omega$ -semigroupes, un *morphisme de  $\omega$ -semigroupe*  $\varphi$  de  $S$  dans  $T$  est un couple  $(\varphi_f, \varphi_i)$  formé d'une application  $\varphi_f$  de  $S_f$  dans  $T_f$  et d'une application  $\varphi_i$  de  $S_i$  dans  $T_i$  respectant la structure de  $\omega$ -semigroupe. L'application  $\varphi_f$  est en particulier un morphisme de semigroupe de  $S_f$  dans  $T_f$ .

## 2.4 Algèbres de Wilke

Une *algèbre de Wilke*  $S$  est un couple  $S = (S_f, S_i)$  muni des opérations suivantes

- une loi de composition interne associative sur  $S_f$ ;
- une action à gauche de  $S_f$  sur  $S_i$ ;
- une application  $\omega$  de  $S_f$  dans  $S_i$  notée  $s \mapsto s^\omega$ .

L'opération  $\omega$  vérifie en outre les propriétés suivantes. Pour tous  $s_1, s_2 \in S_f$ , on a

$$(s_1^n)^\omega = s_1^\omega \quad (3)$$

$$s_1(s_2s_1)^\omega = (s_1s_2)^\omega \quad (4)$$

Si  $S = (S_f, S_i)$  et  $T = (T_f, T_i)$  sont deux algèbres de Wilke, un *morphisme d'algèbre de Wilke*  $\varphi$  de  $S$  dans  $T$  est un couple  $(\varphi_f, \varphi_i)$  formé d'une application  $\varphi_f$  de  $S_f$  dans  $T_f$  et d'une application  $\varphi_i$  de  $S_i$  dans  $T_i$  respectant la structure d'algèbre de Wilke.

**Exemple :** Tout  $\omega$ -semigroupe  $S = (S_f, S_i)$  peut être considéré comme une algèbre de Wilke en définissant  $s^\omega$  comme le produit infini  $sss\dots$  pour tout élément  $s$  de  $S_f$ . Le  $\omega$ -semigroupe  $A^\infty$  peut en particulier être considéré comme une algèbre de Wilke.

**Exemple :** On note  $A^{\text{up}}$  l'ensemble des mots infinis ultimement périodiques sur l'alphabet  $A$ . Le couple  $(A^*, A^{\text{up}})$  muni des opérations naturelles est une algèbre de Wilke.

## 2.5 Factorisations Ramseyennes

Nous rappelons ici un cas particulier du théorème de Ramsey.

**Théorème 9** (Ramsey) *Soit  $A$  un alphabet non nécessairement fini et  $f$  une application de  $A^+$  dans un ensemble fini  $E$ . Pour tout mot  $u \in A^\omega$  il existe une factorisation  $u = u_0u_1u_2 \dots$  telle que  $u_1f = u_2f = \dots = e \in E$ .*

Une factorisation  $u = v_0v_1v_2 \dots$  est dite *extraite* d'une factorisation  $u = u_0u_1u_2 \dots$  s'il existe une suite  $(k_n)_{n \geq 1}$  d'entiers telle que

$$\begin{aligned} v_0 &= u_0 \dots u_{k_1-1} \\ v_1 &= u_{k_1} \dots u_{k_2-1} \\ &\vdots \end{aligned}$$

**Corollaire 2** *Soit  $A$  un alphabet non nécessairement fini et  $\varphi$  un morphisme de  $A^+$  dans un semigroupe fini  $S$ . Pour toute factorisation  $u = u_0u_1u_2 \dots$  d'un mot  $u \in A^\omega$  il existe une factorisation  $u = v_0v_1v_2 \dots$  extraite de la première telle que  $v_0\varphi = s$ ,  $v_1\varphi = v_2\varphi = \dots = e$ ,  $se = s$  et  $e^2 = e$ . Une telle factorisation du mot  $u$  est appelée factorisation ramseyenne extraite de la factorisation du mot  $u$  pour le morphisme  $\varphi : A^+ \rightarrow S$ .*

Dans la majorité des cas, ce corollaire est appliqué quand la factorisation  $u = u_0u_1u_2 \dots$  est la factorisation du mot  $u$  en lettres de l'alphabet. Soit  $S$  un semigroupe. Un couple  $(s, e) \in S \times E(S)$  tel que  $se = s$  est appelé un *couple lié* du semigroupe  $S$ .

Il faut remarquer que pour un mot infini  $u$  donné, on peut en général trouver plusieurs couples liés correspondant à des factorisations ramseyennes du mot  $u$ . Deux couples liés  $(s_1, e_1)$  et  $(s_2, e_2)$  sont dits *conjugués* si et seulement si il existe  $x_1, x_2 \in S^1$  tels que  $s_1x_1 = s_2$ ,  $x_1x_2 = e_1$  et  $x_2x_1 = e_2$ . Ces égalités entraînent également  $s_2x_2 = s_1x_1x_2 = s_1$ , ce qui assure la symétrie de la définition. La relation de conjugaison caractérise les couples liés qui peuvent correspondre à des factorisations ramseyennes différentes d'un même mot. Nous avons le résultat suivant. Il généralise un résultat de Pécuchet [Péc86b] qui donne la propriété pour  $n = 2$ .

**Théorème 10** *La relation de conjugaison est une relation d'équivalence. Les  $n$  couples liés  $(s_1, e_1), \dots, (s_n, e_n)$  sont conjugués si et seulement si il existe un mot  $u \in A^\omega$  tel que  $u$  admette des factorisations ramseyennes correspondant aux couples liés  $(s_1, e_1), \dots, (s_n, e_n)$ .*

Nous commençons par quelques lemmes techniques

**Lemme 3** *Soient  $(s_1, e_1)$  et  $(s_2, e_2)$  deux couples liés de  $S$  et  $x_1 \in S$  tels que  $s_1x_1 = s_2$ ,  $x_1e_2 \mathcal{R} e_1$  et  $e_1 \mathcal{D} e_2$ . Il existe alors  $x'_1$  appartenant à la  $\mathcal{H}$ -classe  $R(e_1) \cap L(e_2)$  tel que  $s_1x'_1 = s_2$*

Nous avons le diagramme

|         |         |
|---------|---------|
| $* e_1$ | $x'_1$  |
|         | $* e_2$ |

**Démonstration :** L'élément  $x'_1 = e_1x_1e_2$  vérifie  $s_1x'_1 = s_1e_1x_1e_2 = s_1x_1e_2 = s_2e_2 = s_2$ . Les égalités  $e_1x'_1 = x'_1$  et  $x'_1e_2 = x'_1$  impliquent  $x'_1 \leq_{\mathcal{R}} e_1$  et  $x'_1 \leq_{\mathcal{L}} e_2$ . Puisque  $x_1e_2 \mathcal{R} e_1$ , il existe  $u \in S^1$  tel que  $x_1e_2u = e_1$ . On a alors  $x'_1u = e_1x_1e_2u = e_1^2 = e_1$  et donc  $x'_1 \mathcal{R} e_1$ . Puisque  $x'_1 \leq_{\mathcal{L}} e_2$  et  $x'_1 \mathcal{D} e_1 \mathcal{D} e_2$ , on a également  $x'_1 \mathcal{L} e_2$ .  $\square$

**Lemme 4** *Soient  $e_1$  et  $e_2$  deux idempotents et  $x_1$  un élément d'un semigroupe  $S$  tels que  $x_1 \in R(e_1) \cap L(e_2)$ . Alors il existe  $x_2 \in R(e_2) \cap L(e_1)$  tel que  $x_1x_2 = e_1$ . De plus, tout  $x_2 \in R(e_2) \cap L(e_1)$  tel que  $x_1x_2 = e_1$  vérifie  $x_2x_1 = e_2$ .*

Nous avons alors le diagramme

|                  |                  |
|------------------|------------------|
| $* e_1 = x_1x_2$ | $x_1$            |
| $x_2$            | $* e_2 = x_2x_1$ |

**Démonstration :** Puisque  $x_1 \mathcal{R} e_1$ , il existe  $u \in S^1$  tel que  $x_1u = e_1$ . En posant  $x_2 = e_2ue_1$ ,  $x_2$  appartient à  $R(e_2) \cap L(e_1)$  et vérifie  $x_1x_2 = x_1e_2ue_1 = x_1ue_1 = e_1^2 = e_1$ .

Soit  $x_2 \in R(e_2) \cap L(e_1)$  vérifiant  $x_1x_2 = e_1$ . Puisque la  $\mathcal{H}$ -classe  $R(x_1) \cap L(x_2)$  contient l'idempotent  $e_1$ , le produit  $x_2x_1$  appartient à la  $\mathcal{H}$ -classe  $R(x_2) \cap L(x_1)$ . De plus, l'élément  $x_2x_1$  est idempotent d'après les égalités  $(x_2x_1)^2 = x_2x_1x_2x_1 = x_2e_1x_1 = x_2x_1$  et est donc égal à  $e_2$ .  $\square$

Nous sommes maintenant en mesure de démontrer une caractérisation équivalente des couples liés conjugués. On a la proposition suivante due à Pécuchet [Péc86a].

**Proposition 9** (Pécuchet) *Deux couples liés  $(s_1, e_1)$  et  $(s_2, e_2)$  d'un semigroupe  $S$  sont conjugués si et seulement si il existe  $x_1 \in S^1$  tel que  $s_1x_1 = s_2$ ,  $x_1e_2 \mathcal{R} e_1$  et  $e_1 \mathcal{D} e_2$ .*

**Démonstration :** Si les couples liés  $(s_1, e_1)$  et  $(s_2, e_2)$  sont conjugués, il existe  $x_1, x_2 \in S^1$  tels que  $s_1x_1 = s_2$ ,  $e_1 = x_1x_2$  et  $e_2 = x_2x_1$ . Nous avons alors  $x_1e_2 = x_1x_2x_1 \mathcal{R} x_1x_2 = e_1$  puisque  $(x_1x_2x_1)x_2 = e_1^2 = e_1 = x_1x_2$ . De plus, les égalités  $x_2e_1x_1 = x_2x_1x_2x_1 = e_2$  et  $x_1e_2x_2 = x_1x_2x_1x_2 = e_1$  montrent que  $e_1$  et  $e_2$  sont  $\mathcal{D}$ -équivalents.

Réciproquement, s'il existe  $x_1 \in S^1$  tel que  $s_1x_1 = s_2$ ,  $x_1e_2 \mathcal{R} e_1$  et  $e_1 \mathcal{D} e_2$ , une application du lemme 3 permet de supposer que  $x_1 \in R(e_1) \cap L(e_2)$ . Le lemme 4 assure alors l'existence de  $x_2$  tel que  $x_1x_2 = e_1$  et  $x_2x_1 = e_2$ .  $\square$

La réflexivité et la symétrie de la relation de conjugaison sont évidentes. Soient  $(s_1, e_1)$ ,  $(s_2, e_2)$  et  $(s_3, e_3)$  trois couples liés. Supposons que  $(s_1, e_1)$  et  $(s_2, e_2)$  soient conjugués et que  $(s_2, e_2)$  et  $(s_3, e_3)$  soient également conjugués. D'après la proposition 9, il existe  $x_1$  et  $x_2$  tels que  $s_1x_1 = s_2$ ,  $x_1e_2 \mathcal{R} e_1$ ,  $s_2x_2 = s_3$  et  $x_2e_3 \mathcal{R} e_2$ . Nous avons  $s_1x_1x_2 = s_2x_2 = s_3$  et  $x_1x_2e_3 \mathcal{R} x_1e_2 \mathcal{R} e_1$  et les couples  $(s_1, e_1)$  et  $(s_3, e_3)$  sont conjugués. La relation de conjugaison est donc transitive. Ceci montre que la relation de conjugaison est une relation d'équivalence.

Nous introduisons ici quelques notations. Étant donnés  $n$  éléments  $x_1, \dots, x_n$  d'un semigroupe  $S$ , nous définissons le sous-ensemble circulaire  $E_{i,j}(n)$  et le produit circulaire  $C_{i,j}(x_1, \dots, x_n)$  par

$$E_{i,j}(n) = \begin{cases} \{i, \dots, n\} & \text{si } 1 \leq i \leq n \text{ et } j = 1 \\ \{i, \dots, j-1\} & \text{si } 1 \leq i < j \leq n \\ \{i, \dots, n, 1, \dots, j-1\} & \text{si } 1 < j \leq i \leq n \end{cases}$$

$$C_{i,j}(x_1, \dots, x_n) = \begin{cases} x_i \dots x_n & \text{si } 1 \leq i \leq n \text{ et } j = 1 \\ x_i \dots x_{j-1} & \text{si } 1 \leq i < j \leq n \\ x_i \dots x_n x_1 \dots x_{j-1} & \text{si } 1 < j \leq i \leq n \end{cases}$$

Par abus de notation, nous écrirons  $C_{i,j}(x)$  pour  $C_{i,j}(x_1, \dots, x_n)$ . Dans le produit  $C_{i,j}(x)$  interviennent uniquement les  $x_i$  pour  $i$  appartenant à l'ensemble  $E_{i,j}(n)$ . Pour cette raison, on s'autorise à écrire  $C_{i,j}(x)$  même si certains  $x_i$  pour  $i \notin E_{i,j}(n)$  ne sont pas définis. On peut remarquer que l'ensemble  $E_{i,j}(n)$  est un singleton si  $j = i + 1$  pour  $1 \leq i < n$  ou si  $i = n$  et  $j = 1$ . Dans ces cas, nous avons les égalités.  $C_{i,i+1}(x) = x_i$  et  $C_{n,1}(x) = x_n$ . On peut aussi remarquer que si  $k$  appartient à l'ensemble  $E_{i,j} - \{i\}$ , la relation de Chasles  $C_{i,k}(x)C_{k,j}(x) = C_{i,j}(x)$  est vérifiée. La proposition suivante étend le lemme de Pécuchet.

**Proposition 10** Soient  $n$  couples liés conjugués  $(s_1, e_1), \dots, (s_n, e_n)$ .  
Il existe alors  $x_1, \dots, x_n \in S^1$  tels que

- Pour  $k = 1, \dots, n-1$   $s_k x_k = s_{k+1}$  et  $s_n x_n = s_1$ ;

- Pour  $k = 1, \dots, n$   $C_{k,k}(x_1, \dots, x_n) = e_k$ .

**Démonstration :** Nous commençons par démontrer un lemme.

**Lemme 5** Soient  $i, j \in \{1, \dots, n\}$ . Si pour tout  $k \in E_{i,j}(n)$  on a  $x_k \in R(e_k) \cap L(e_{k+1})$ , alors on a  $C_{i,j}(x) \in R(e_i) \cap L(e_j)$ .

**Démonstration :** Nous effectuons une récurrence sur le cardinal de l'ensemble  $E_{i,j}(n)$ . L'hypothèse  $x_k \in R(e_k) \cap L(e_{k+1})$  assure le résultat lorsque  $E_{i,j}(n)$  est un singleton. Si  $E_{i,j}(n)$  est de cardinal au moins 2, il est possible de choisir  $k$  appartenant à  $E_{i,j}(n) - \{i\}$ . Le produit  $C_{i,j}(x)$  se décompose alors  $C_{i,j}(x) = C_{i,k}(x)C_{k,j}(x)$ . La  $\mathcal{H}$ -classe  $R(C_{k,j}(x)) \cap L(C_{i,k}(x))$  contient par hypothèse de récurrence l'idempotent  $e_k$  et le produit  $C_{i,j}(x)$  appartient donc à la  $\mathcal{H}$ -classe  $R(C_{i,k}(x)) \cap L(C_{k,j}(x)) = R(e_i) \cap L(e_j)$ . Ceci termine la preuve du lemme.  $\square$

Puisque les couples liés  $(s_i, e_i)$  sont conjugués, les  $e_i$  sont  $\mathcal{D}$ -équivalents et il existe  $x_1, \dots, x_{n-1}$  tels que  $s_i x_i = s_{i+1}$  pour  $1 \leq i \leq n-1$ . Grâce au lemme 3, nous pouvons supposer que  $x_i \in R(e_i) \cap L(e_{i+1})$  pour  $1 \leq i \leq n-1$ . Le choix des  $x_i$  implique  $C_{i,j}(x) \in R(e_i) \cap L(e_j)$  pour  $1 \leq i < j \leq n$  puisque dans ce cas on a  $E_{i,j}(n) \subset \{1, \dots, n-1\}$ .

Montrons maintenant qu'il existe  $x_n \in R(e_n) \cap L(e_1)$  tel que  $\forall 1 \leq k \leq n$   $C_{k,k}(x) = e_k$ . D'après le lemme 4, il existe  $y \in R(e_2) \cap L(e_1)$  tel que  $x_1 y = e_1$  et  $y x_1 = e_2$ . Puisque  $y$  appartient à la  $\mathcal{R}$ -classe  $R(e_2) = R(C_{2,n}(x))$ , il existe  $x_n \in S^1$  tel que  $C_{2,1}(x) = C_{2,n}(x)x_n = y$ . Nous pouvons en outre supposer grâce au lemme 3 que  $x_n$  appartient à la  $\mathcal{H}$ -classe  $R(e_n) \cap L(e_1)$ . Montrons que  $x_n$  ainsi choisi vérifie  $\forall 1 \leq k \leq n$   $C_{k,k}(x) = e_k$ . On raisonne alors par récurrence sur  $k$ . Les égalités pour  $k = 1$  et  $k = 2$  proviennent directement du choix de  $x_n$ ; nous avons les égalités

$$\begin{aligned} C_{1,1}(x) &= x_1 C_{2,1}(x) = x_1 y = e_1 \\ C_{2,2}(x) &= C_{2,1}(x) x_1 = y x_1 = e_2. \end{aligned}$$

On suppose maintenant que le résultat est acquis pour  $k$ . Grâce à la relation de Chasles, on a les égalités  $C_{k,k+1}(x)C_{k+1,k}(x) = C_{k,k}(x)$  et  $C_{k+1,k}(x)C_{k,k+1}(x) = C_{k+1,k+1}(x)$ . D'après l'hypothèse de récurrence, on a  $e_k = C_{k,k}(x)$ . Par application du lemme 4, on obtient le résultat pour  $k+1$ . Ceci termine la preuve de la proposition.  $\square$

Nous avons finalement le diagramme

|                        |                        |                        |          |                        |
|------------------------|------------------------|------------------------|----------|------------------------|
| ${}^*e_1 = C_{1,1}(x)$ | $C_{1,2}(x)$           | $C_{1,3}(x)$           | $\dots$  | $C_{1,n}(x)$           |
| $C_{2,1}(x)$           | ${}^*e_2 = C_{2,2}(x)$ | $C_{2,3}(x)$           | $\dots$  | $C_{2,n}(x)$           |
| $C_{3,1}(x)$           | $C_{3,2}(x)$           | ${}^*e_3 = C_{3,3}(x)$ | $\dots$  | $C_{3,n}(x)$           |
| $\vdots$               | $\vdots$               | $\vdots$               | $\ddots$ | $\vdots$               |
| $C_{n,1}(x)$           | $C_{n,2}(x)$           | $C_{n,3}(x)$           | $\dots$  | ${}^*e_n = C_{n,n}(x)$ |

Nous terminons maintenant la démonstration du théorème 10. Soient  $(s_1, e_1)$  et  $(s_2, e_2)$  deux couples liés correspondant aux deux factorisations ramseyennes  $u = u_0u_1u_2\dots$  et  $u = v_0v_1v_2\dots$  d'un mot infini  $u$ . Nécessairement, les deux idempotents  $e_1$  et  $e_2$  sont  $\mathcal{D}$ -équivalents. En effet, pour  $n$  assez grand, le mot  $u_n$  est facteur d'un mot  $v_i\dots v_{i+j}$  et le mot  $v_n$  est facteur d'un mot  $u_k\dots u_{k+l}$ , ce qui assure  $e_1 \mathcal{D} e_2$ . Par symétrie, on peut supposer que  $u_0$  est un préfixe de  $u_1$ . Le mot  $u_1$  s'écrit alors  $u_1 = u_0w$ . Le mot  $wv_1$  est un préfixe du mot  $u_1u_2u_3\dots u_n$  pour  $n$  assez grand. En posant  $x_1 = w\varphi$ , on obtient  $s_1x_1 = s_2$  et  $x_1e_2 \mathcal{R} e_1$ . La proposition 9 permet de conclure que les couples  $(s_1, e_1)$  et  $(s_2, e_2)$  sont conjugués.

Réciproquement, soient  $n$  couples liés conjugués  $(s_1, e_1), \dots, (s_n, e_n)$ . Il existe,  $n$  éléments  $x_1, \dots, x_n$  de  $S_f$  vérifiant les conditions de la proposition précédente. On choisit  $n$  mots  $u_0$  et  $u_1, \dots, u_n \in A^*$  tels que  $u_0\varphi = s_1$  et que  $u_i\varphi = x_i$  pour  $1 \leq i \leq n$ . On constate alors que le mot infini  $u = u_0(u_1\dots u_n)^\omega$  admet des factorisations ramseyennes correspondant aux couples liés  $(s_1, e_1), \dots, (s_n, e_n)$ .

## 2.6 $\omega$ -semigroupes finis

Tout  $\omega$ -semigroupe  $S = (S_f, S_i)$  peut être considéré comme une algèbre de Wilke en définissant  $s^\omega$  égal au produit infini  $sss\dots$  pour tout élément  $s$  de  $S_f$ . Nous allons voir que les notions de  $\omega$ -semigroupe et d'algèbre de Wilke coïncident dans le cas fini. Nous avons en effet le résultat suivant.

**Théorème 11** *Soit  $S = (S_f, S_i)$  une algèbre de Wilke finie. Il existe un unique produit infini de  $S_f$  dans  $S_i$  telle que l'opération  $\omega$  soit induite par ce produit infini.*

**Démonstration :** Soit  $(s_n)_{n \in \mathbb{N}}$  une suite d'éléments de  $S_f$ . D'après le corollaire 2, il existe un couple lié  $(s, e)$  de  $S_f$  et une suite d'entiers  $(k_n)_{n \geq 1}$  tels

que

$$\begin{aligned} s_0 s_1 \cdots s_{k_1-1} &= s \\ s_{k_n} \cdots s_{k_{n+1}-1} &= e \quad \text{pour tout } n \geq 1 \end{aligned}$$

Il en résulte que si  $(S_f, S_i)$  peut être muni d'une structure de  $\omega$ -semigroupe avec un produit infini induisant l'opération  $\omega$ , alors le produit infini  $s_0 s_1 s_2 \cdots$  est nécessairement égal à  $se^\omega$ .

Pour rendre cette définition cohérente, il reste à s'assurer que cette définition ne dépend pas du choix de la factorisation ramseyenne. Si  $(s_1, e_1)$  et  $(s_2, e_2)$  sont deux couples liés correspondant à des factorisations de la suite  $(s_n)_{n \in \mathbb{N}}$ , alors ces couples sont conjugués d'après le théorème 10 il existe  $x_1, x_2 \in S_f^1$  tels que  $s_1 x_1 = s_2$ ,  $x_1 x_2 = e_1$  et  $x_2 x_1 = e_2$  d'après la proposition 9. On vérifie que

$$s_1 e_1^\omega = s_1 (x_1 x_2)^\omega = s_1 x_1 (x_2 x_1)^\omega = s_2 e_2^\omega$$

On vérifie que le produit infini ainsi défini est compatible avec le produit de  $S_f$  et l'action à gauche de  $S_f$  sur  $S_i$ .  $\square$

Dorénavant, on ne fera plus la distinction entre  $\omega$ -semigroupe et algèbre de Wilke dans le cas fini.

Nous allons montrer que l'on peut associer un  $\omega$ -semigroupe fini  $\hat{S} = (S, S_i)$  à tout semigroupe fini  $S$ . Pour cette construction, on peut aussi se référer à [Wil91] et [PP93a]. On note  $\pi$  l'exposant du semigroupe  $S$  et  $[s, e]$  la classe de conjugaison du couple lié  $(s, e)$ . On prend pour  $S_i$  l'ensemble des classes de conjugaison des couples liés. On définit l'opération  $\omega$  et l'action à gauche en posant pour  $[s, e] \in S_i$  et  $t \in S$

$$t[s, e] = [ts, e] \quad \text{et} \quad t^\omega = [t^\pi, t^\pi]$$

Ces définitions sont cohérentes car si  $(s_1, e_1)$  et  $(s_2, e_2)$  sont deux couples conjugués, les couples  $(ts_1, e_1)$  et  $(ts_2, e_2)$  sont également conjugués pour tout  $t \in S$ . On vérifie que  $\hat{S}$  ainsi défini est un  $\omega$ -semigroupe fini. Le  $\omega$ -semigroupe  $\hat{S}$  possède la propriété universelle suivante.

**Proposition 11** *Soit  $T = (T_f, T_i)$  un  $\omega$ -semigroupe fini et soit  $S$  un semigroupe fini. Pour tout morphisme de semigroupe  $\varphi$  de  $S$  dans  $T_f$ , il existe un unique morphisme de  $\omega$ -semigroupe de  $\hat{\varphi} : \hat{S} \rightarrow T$  qui prolonge  $\varphi$ . Si de plus,  $\varphi(S)$  engendre  $T$  comme  $\omega$ -semigroupe, alors  $\hat{\varphi}$  est surjectif.*

**Démonstration :** Un morphisme de  $\omega$ -semigroupe  $\hat{\varphi} : \hat{S} \rightarrow T$  vérifiant les hypothèses de l'énoncé est nécessairement défini par

$$\begin{aligned} \hat{\varphi}(t) &= t && \text{pour tout } t \in S \\ \hat{\varphi}([s, e]) &= \varphi(s)\varphi(e)^\omega && \text{pour tout } [s, e] \in S_i \end{aligned}$$

Ces définitions sont cohérentes, car si  $(s_1, e_1)$  et  $(s_2, e_2)$  sont deux couples conjugués, il existe, d'après la proposition 9,  $x_1, x_2 \in S^1$  tels que  $s_1x_1 = s_2$ ,  $x_1x_2 = e_1$  et  $x_2x_1 = e_2$ . On a alors

$$\begin{aligned} \widehat{\varphi}([s_1, e_1]) &= \varphi(s_1)\varphi(e_1)^\omega \\ &= \varphi(s_1)\varphi(x_1x_2)^\omega \\ &= \varphi(s_2)\varphi(x_2x_1)^\omega \\ &= \varphi(s_2)\varphi(e_2)^\omega \\ &= \widehat{\varphi}([s_2, e_2]) \end{aligned}$$

□

## 2.7 Reconnaissance par morphisme

Par définition, un morphisme de  $\omega$ -semigroupe  $\varphi : A^\omega \rightarrow S$  reconnaît un  $\omega$ -langage  $X \subset A^\omega$  si et seulement si le morphisme sature le langage  $X$  c'est à dire si  $X = X\varphi^{-1}$ . Par extension, on dit qu'un  $\omega$ -semigroupe  $S$  reconnaît un  $\omega$ -langage  $X$  s'il existe un morphisme de  $A^\omega$  dans  $S$  qui reconnaît le langage  $X$ . Le théorème suivant énonce que les  $\omega$ -langages reconnus par un  $\omega$ -semigroupe fini sont exactement les  $\omega$ -langages reconnaissables par automate. Par conséquent, on parlera de  $\omega$ -langage reconnaissable.

**Théorème 12 (Wilke)** *Un  $\omega$ -langage est reconnu par un  $\omega$ -semigroupe fini si et seulement si il est reconnaissable par un automate.*

Pour la démonstration de ce théorème, on peut se référer à [Wil91] ou à [PP93a].

**Exemple :** On considère l'alphabet  $A = a + b$  et le morphisme  $\varphi$  de  $A^+$  dans le semigroupe multiplicatif  $S = \{0, 1\}$  défini par  $a\varphi = 0$  et  $b\varphi = 1$ . On identifie donc  $a$  avec 0 et  $b$  avec 1. Les couples liés de  $S$  sont  $(b, b)$ ,  $(a, b)$  et  $(a, a)$ . Le couple  $(a, a)$  est seul dans sa classe de conjugaison puisque l'idempotent  $a$  n'est pas  $\mathcal{D}$ -équivalent à  $b$ . Les couples  $(b, b)$  et  $(a, b)$  ne sont pas conjugués puisque  $a$  et  $b$  ne sont pas  $\mathcal{R}$ -équivalents. La partie  $S_i$  de  $\widehat{S}$  est donc égale à  $S_i = \{b^\omega, ab^\omega, a^\omega\}$ . On a les égalités

$$\begin{aligned} b\varphi^{-1} &= b^+ && \text{(mots finis ne contenant pas d'occurrence de } a) \\ a\varphi^{-1} &= A^*aA^+ && \text{(mots finis contenant au moins une occurrence de } a) \\ b^\omega\varphi^{-1} &= b^\omega && \text{(mots infinis ne contenant pas d'occurrence de } a) \\ ab^\omega\varphi^{-1} &= A^*ab^\omega && \text{(mots infinis contenant un nombre fini non nul} \\ &&& \text{d'occurrences de } a) \\ a^\omega\varphi^{-1} &= (A^*a)^\omega && \text{(mots infinis contenant un nombre infini} \\ &&& \text{d'occurrences de } a) \end{aligned}$$

Le morphisme de  $\omega$ -semigroupe  $\varphi : A^\infty \rightarrow \widehat{S}$  reconnaît donc le  $\omega$ -langage  $X = (A^*a)^\omega = (b^*a)^\omega$ .

**Exemple :** On considère le semigroupe présenté par

$$\langle a, b; aba = a, bab = b, a^2 = b^2 = 0 \rangle$$

|        |        |
|--------|--------|
| $^*ab$ | $a$    |
| $b$    | $^*ba$ |
| $^*0$  |        |

Relations  
 $a^2 = b^2 = 0$   
 $aba = a$   
 $bab = b$

et le morphisme canonique de  $(a + b)^+$  dans ce semigroupe. Les éléments du semigroupe sont  $\{a, ab, b, ba, 0\}$ . Les classes de conjugaisons des couples liés sont décrites par (pour le couple lié  $(s, t)$ , on écrit  $st^\omega$  si  $s \neq t$  et  $s^\omega$  sinon).

|                |           |
|----------------|-----------|
| $(ab)^\omega$  | $x_1 = a$ |
| $a(ba)^\omega$ | $x_2 = b$ |
| $(ba)^\omega$  | $x_1 = b$ |
| $b(ab)^\omega$ | $x_2 = a$ |
| $0(ab)^\omega$ | $x_1 = a$ |
| $0(ba)^\omega$ | $x_2 = b$ |
| $0^\omega$     |           |

On a les égalités

$$\begin{aligned} (ab)^\omega \varphi^{-1} &= (ab)^\omega \\ (ba)^\omega \varphi^{-1} &= (ba)^\omega \\ 0(ab)^\omega \varphi^{-1} &= A^+(ab)^\omega - \{(ab)^\omega, (ba)^\omega\} \\ 0^\omega \varphi^{-1} &= A^\omega - A^+(ab)^\omega \end{aligned}$$

## 2.8 Opérations

À travers la correspondance entre les variétés de langages et les variétés de semigroupes, certaines opérations sur les langages se traduisent par des opérations sur les semigroupes. Ainsi le passage aux images inverses par substitutions correspond au passage au semigroupe des parties [Str79]. De même, l'image inverse par fonction séquentielle correspond au produit en couronne et la concaténation au produit de Schützenberger.

Nous allons maintenant définir pour les  $\omega$ -semigroupes, le  $\omega$ -semigroupe des parties, le produit en couronne et le produit de Schützenberger. Pour chacune de ces constructions, nous allons démontrer l'analogie pour les mots infinis des résultats pour les mots finis.

### 2.8.1 Semigroupe et $\omega$ -semigroupe des parties

Soit  $S$  un semigroupe. Le semigroupe des parties de  $S$  est le semigroupe de support  $\mathcal{P}(S)$  muni du produit défini pour tous  $P, Q \subset S$  par

$$PQ = \{st \mid s \in P \text{ et } t \in Q\}.$$

Soit  $\varphi : S \rightarrow T$  est un morphisme d'un semigroupe  $S$  dans un semigroupe  $T$ . Le morphisme  $\varphi$  s'étend naturellement en un morphisme de  $\mathcal{P}(S)$  dans  $\mathcal{P}(T)$  en posant  $P\varphi = \{s\varphi \mid s \in P\}$  pour toute partie  $P$  de  $S$ .

Soit  $S = (S_f, S_i)$  un  $\omega$ -semigroupe fini. Le  $\omega$ -semigroupe des parties de  $S$ , noté  $\mathcal{P}(S)$  est le  $\omega$ -semigroupe de support  $(\mathcal{P}(S_f), \mathcal{P}(S_i))$ . Pour  $P \subset S_f$  et  $Q \subset S_i$  l'action à gauche de  $P$  sur  $Q$  est définie par

$$PQ = \{st \mid s \in P \text{ et } t \in Q\}.$$

Il est immédiat de vérifier que ceci définit effectivement une action à gauche de  $\mathcal{P}(S_f)$  sur  $\mathcal{P}(S_i)$ . Pour une partie  $P$  de  $S_f$  l'opération  $\omega$  est définie par

$$P^\omega = \{s^\omega \mid s \in P^+\}$$

ou  $P^+$  dénote le sous-semigroupe engendré par la partie  $P$  dans  $S_f$ . Il reste à vérifier que pour toutes parties  $P$  et  $Q$  de  $S_f$ , on a les égalités

$$\begin{aligned} (P^n)^\omega &= P^\omega \\ P(QP)^\omega &= (PQ)^\omega. \end{aligned}$$

De l'inclusion  $P^n \subset P^+$ , on déduit les inclusions  $(P^n)^+ \subset P^+$  et  $(P^n)^\omega \subset P^\omega$ . Réciproquement, soit  $x \in P^+$ . L'élément  $x^n$  appartient à  $(P^n)^+$  et  $x^\omega = (x^n)^\omega$  appartient à  $(P^n)^\omega$  ce qui montre l'inclusion inverse.

Un élément  $x$  de  $(PQ)^+$  s'écrit  $x = s_1 t_1 \dots s_m t_m$  avec  $s_i \in P$  et  $t_i \in Q$  pour  $1 \leq i \leq m$ . On a alors l'inclusion  $(PQ)^\omega \subset P(QP)^\omega$  puisque  $x^\omega = (s_1 t_1 \dots s_m t_m)^\omega = s_1 (t_1 s_2 \dots s_m t_m s_1)^\omega$  qui appartient à  $P(QP)^\omega$ . De la même façon, on montre l'inclusion inverse  $P(QP)^\omega \subset (PQ)^\omega$ .

Si le langage  $L$  de mots finis est reconnu par le semigroupe  $S$  et si  $\sigma$  est une substitution, le langage  $L\sigma^{-1}$  est reconnu par le semigroupe  $\mathcal{P}(S)$ . Le théorème suivant étend ce résultat aux  $\omega$ -langages.

**Théorème 13** *Soit  $\sigma : A^* \rightarrow B^*$  une substitution. Si le  $\omega$ -langage  $L \subset B^\omega$  est reconnu par le  $\omega$ -semigroupe  $S = (S_f, S_i)$ , alors le  $\omega$ -langage  $L\sigma^{-1} \subset A^\omega$  est reconnu par le  $\omega$ -semigroupe  $\mathcal{P}(S)$ .*

**Démonstration :** Soit  $\varphi : B^\infty \rightarrow S$  un morphisme surjectif reconnaissant  $L$ . On définit le morphisme  $\psi : A^\infty \rightarrow \mathcal{P}(S)$  pour toute lettre  $a \in A$  par

$$a\psi = \{v\varphi \mid v \in a\sigma\}.$$

La substitution  $\sigma$  est par définition un morphisme de  $A^*$  dans  $\mathcal{P}(B^*)$ . Le morphisme  $\psi$  est donc égal à la composée  $\sigma\varphi$  et pour tout mot  $u \in A^*$ , on a l'égalité

$$u\psi = \{v\varphi \mid v \in u\sigma\}. \quad (5)$$

Nous étendons maintenant ce résultat au mots infinis.

**Lemme 6** *Pour tout mot  $u \in A^\omega$ , on a*

$$u\psi = \{v\varphi \mid v \in u\sigma\}.$$

**Démonstration :** On écrit  $u = a_0a_1a_2 \dots$  pour désigner les lettres du mot  $u$ . Soit  $v$  un mot de  $u\sigma$ . Le mot  $v$  admet donc une factorisation  $v = v_0v_1v_2 \dots$  où chacun des  $v_i$  appartient à  $a_i\sigma$ . Par application du corollaire 2 (p. 28), on obtient une factorisation ramseyenne  $v = v'_0v'_1v'_2 \dots$  du mot  $v$  extraite de la factorisation précédente pour le morphisme  $\varphi$ . On a donc  $v'_0\varphi = s$  et  $v'_1\varphi = v'_2\varphi = \dots = e$  et l'image du mot  $v$  par  $\varphi$  est  $se^\omega$ . La factorisation  $v = v'_0v'_1v'_2 \dots$  induit une factorisation  $u = u_0u_1u_2 \dots$  du mot  $u$  telle que  $v'_i$  appartienne à  $u_i\sigma$  pour tout  $i$ . De cette factorisation de  $u$ , on extrait une factorisation ramseyenne  $u = u'_0u'_1u'_2 \dots$  de  $u$  pour le morphisme  $\psi$ . On a donc  $u'_0\psi = P$  et  $u'_1\psi = u'_2\psi = \dots = Q$  et l'image du mot  $u$  par  $\psi$  est  $PQ^\omega$ . D'après la formule 5, on a  $s \in P$  et  $e \in Q$  et donc  $v\varphi = se^\omega \in PQ^\omega = u\psi$ . Ceci montre l'inclusion  $\{v\varphi \mid v \in u\sigma\} \subset u\psi$ .

Réciproquement, soit  $u = u_0u_1u_2 \dots$  une factorisation ramseyenne du mot  $u$  pour le morphisme  $\psi$ . On a donc  $u_0\psi = P$  et  $u_1\psi = u_2\psi = \dots = Q$  et l'image du mot  $u$  par  $\psi$  est  $PQ^\omega$ . Un élément de  $u\psi$  s'écrit  $s_0(s_1 \dots s_m)^\omega$  où  $s_0 \in P$  et  $s_i \in Q$  pour  $1 \leq i \leq m$ . Il existe alors des mots finis  $v_0, v_1, v_2, \dots$  tels que  $v_i \in u_i\sigma$  et  $v_0\varphi = s_0$  et  $v_i\varphi = s_i$  pour  $i \geq 1$  où  $\hat{i}$  dénote l'entier de  $\{1, 2, \dots, m\}$  congru à  $i$  modulo  $m$ . Le mot infini  $v = v_0v_1v_2 \dots$  appartient à  $u\sigma$  et vérifie  $v\varphi = s_0(s_1 \dots s_m)^\omega$ . Ceci termine la preuve du lemme.  $\square$

Un mot infini  $u$  appartient à  $L\sigma^{-1}$  si et seulement si  $u\psi$  contient un élément de  $L\varphi$ . Ceci termine la preuve du théorème.  $\square$

## 2.8.2 Produit de Schützenberger

Soient  $T$  et  $S$  deux semigroupes. Le *produit de Schützenberger* noté  $T \diamond S$  des semigroupes  $S$  et  $T$  est le semigroupe de support  $T \times \mathcal{P}(T^1 \times S^1) \times S$ . Un élément

de ce semigroupe est donc un triplet constitué d'un élément de  $T$ , d'une partie du produit cartésien  $T^1 \times S^1$  et d'un élément de  $S$ . Le produit est défini par

$$(t_1, P_1, s_1)(t_2, P_2, s_2) = (t_1 t_2, t_1 P_2 + P_1 s_2, s_1 s_2).$$

où les parties  $t_1 P_2$  et  $P_1 s_2$  sont définies, de façon naturelle, par les formules

$$\begin{aligned} t_1 P_2 &= \{(t_1 t, s) \mid (t, s) \in P\} \\ P_1 s_2 &= \{(t, s s_2) \mid (t, s) \in P\} \end{aligned}$$

L'élément  $(t, P, s)$  du semigroupe  $T \diamond S$  peut être identifié avec la matrice

$$\begin{pmatrix} t & P \\ 0 & s \end{pmatrix}$$

et le produit du semigroupe est identique au produit matriciel.

Soient deux morphismes  $\varphi : A^+ \rightarrow S$  et  $\psi : A^+ \rightarrow T$  où  $S$  et  $T$  sont deux semigroupes. Le morphisme  $\varphi \times \psi : A^+ \rightarrow S \diamond T$  est défini pour toute lettre  $a \in A$  par

$$a(\varphi \times \psi) = \begin{pmatrix} a\varphi & \{(a\varphi, 1), (1, a\psi)\} \\ 0 & a\psi \end{pmatrix}$$

Nous définissons maintenant le produit de Schützenberger d'un semigroupe  $T$  et d'un  $\omega$ -semigroupe  $S = (S_f, S_i)$ . Le produit de schützenberger  $T \diamond S$  est le  $\omega$ -semigroupe de support  $(T \diamond S_f, \mathcal{P}(T^1 \times S_i) \times S_i)$ . La partie finitaire de cet  $\omega$ -semigroupe est donc le semigroupe  $T \diamond S_f$ . Un élément  $(Q, e)$  de la partie infinitaire de  $T \diamond S$  est un couple constitué d'une partie  $Q$  du produit cartésien  $T^1 \times S_i$  et d'un élément  $e$  de  $S_i$ . L'élément  $(Q, e)$  peut être identifié avec la matrice colonne

$$\begin{pmatrix} Q \\ e \end{pmatrix}.$$

L'action à gauche du  $\omega$ -semigroupe s'identifie au produit matriciel et se définit par

$$\begin{pmatrix} t_1 & P_1 \\ 0 & s_1 \end{pmatrix} \begin{pmatrix} Q_1 \\ e_1 \end{pmatrix} = \begin{pmatrix} t_1 Q_1 + P_1 e_1 \\ s_1 e_1 \end{pmatrix}$$

où les parties  $t_1 Q_1$  et  $P_1 e_1$  de  $S_i$  sont définies par

$$\begin{aligned} t_1 Q_1 &= \{(t_1 t, e) \mid (t, e) \in Q_1\} \\ P_1 e_1 &= \{(t, s e_1) \mid (t, s) \in P_1\}. \end{aligned}$$

On vérifie que cela définit bien une action à gauche. On a en effet

$$\begin{pmatrix} t_2 & P_2 \\ 0 & s_2 \end{pmatrix} \begin{pmatrix} t_1 & P_1 \\ 0 & s_1 \end{pmatrix} \begin{pmatrix} Q_1 \\ e_1 \end{pmatrix} = \begin{pmatrix} t_2 t_1 Q_1 + t_2 P_1 e_1 + P_2 s_1 e_1 \\ s_2 s_1 e_1 \end{pmatrix}$$

L'opération  $\omega$  est définie par

$$\begin{pmatrix} t_1 & P_1 \\ 0 & s_1 \end{pmatrix}^\omega = \begin{pmatrix} \{(t_1^k t, s s_1^\omega) \mid k \in \mathbb{N} \text{ et } (t, s) \in P_1\} \\ s_1^\omega \end{pmatrix}$$

Une vérification immédiate donne

$$\begin{pmatrix} t_1 & P_1 \\ 0 & s_1 \end{pmatrix} \left( \begin{pmatrix} t_2 & P_2 \\ 0 & s_2 \end{pmatrix} \begin{pmatrix} t_1 & P_1 \\ 0 & s_1 \end{pmatrix} \right)^\omega = \left( \begin{pmatrix} t_1 & P_1 \\ 0 & s_1 \end{pmatrix} \begin{pmatrix} t_2 & P_2 \\ 0 & s_2 \end{pmatrix} \right)^\omega$$

Pour les langages de mots finis, si  $L_1$  et  $L_2$  sont deux langages reconnus respectivement par les semigroupes  $S$  et  $T$ , alors le langage  $L_1 L_2$  est reconnu par le produit de Schützenberger  $S \diamond T$  (cf. [Pin84]). Le théorème suivant étend ce résultat aux  $\omega$ -langages.

**Théorème 14** *Soient  $L \subset A^+$  un langage reconnu par un semigroupe  $T$  et  $X \subset A^\omega$  un  $\omega$ -langage reconnu par un  $\omega$ -semigroupe  $S = (S_f, S_i)$ , alors le  $\omega$ -langage  $LX \subset A^\omega$  est reconnu par le  $\omega$ -semigroupe  $T \diamond S$ .*

**Démonstration :** Soient un morphisme  $\varphi : A^+ \rightarrow T$  tel que  $L = L\varphi^{-1}$  et un morphisme de  $\psi : A^\omega \rightarrow S$  tel que  $X = X\psi^{-1}$ . On pose  $\psi = (\psi_f, \psi_i)$  où  $\psi_f$  est un morphisme de  $A^+$  dans le semigroupe  $S_f$ . On note  $\varphi \diamond \psi$  le morphisme de  $A^\omega$  dans  $T \diamond S$ . La première composante du morphisme  $\varphi \diamond \psi$  est le morphisme  $\varphi \diamond \psi_f$  de  $A^+$  dans  $T \diamond S_f$ . L'image d'un mot fini  $u$  par le morphisme  $\varphi \diamond \psi$  est donnée par

$$u(\varphi \diamond \psi) = \begin{pmatrix} u\varphi & \{(v\varphi, w\psi_f) \mid v \in A^*, w \in A^* \text{ et } u = vw\} \\ 0 & u\psi_f \end{pmatrix}$$

Montrons que l'image d'un mot infini  $u$  dans  $T \diamond S$  est donnée par

$$u(\varphi \diamond \psi) = \begin{pmatrix} \{(v\varphi, w\psi_i) \mid v \in A^*, w \in A^\omega \text{ et } u = vw\} \\ u\psi_i \end{pmatrix} \quad (6)$$

**Lemme 7** *Si le mot infini  $w$  vérifie la formule 6, alors pour tout mot fini  $v$ , le mot infini  $vw$  vérifie la formule 6.*

**Démonstration :** On pose

$$v(\varphi \diamond \psi) = \begin{pmatrix} t_0 & P_0 \\ 0 & s_0 \end{pmatrix} \quad w(\varphi \diamond \psi) = \begin{pmatrix} Q_0 \\ e_0 \end{pmatrix} \quad vw(\varphi \diamond \psi) = \begin{pmatrix} Q_1 \\ e_1 \end{pmatrix}$$

Par définition de l'action à gauche,  $e_1 = s_0 e_0 = v\psi \cdot w\psi = vw\psi$ , ce qui montre que la deuxième composante de  $vw(\varphi \diamond \psi)$  est égale à  $vw\psi$ .

Soit  $v_1w_1$  une factorisation de  $vw$ . Soit le mot  $v$  est préfixe du mot  $v_1$ , soit le mot  $v_1$  est préfixe du mot  $v$ .

Dans le premier cas, le mot  $v_1$  s'écrit  $v_1 = vv_2$  et  $v_2w_1$  est alors une factorisation du mot  $w$ . Le couple  $(v_2\varphi, w_1\psi)$  appartient à  $Q_0$  et le couple  $(v_1\varphi, w_1\psi)$  appartient à  $t_0Q_0$  puisque  $v_1\varphi = v\varphi.v_2\varphi = t_0.v_2\varphi$ . Réciproquement, si un couple  $(t_0t, e)$  appartient à  $t_0Q_0$ , le couple  $(t, e)$  est égal  $(v_2\varphi, w_1\psi)$  pour une factorisation  $v_2w_1$  de  $w$ , et le couple  $(t_0t, e)$  correspond à la factorisation  $v_1w_1$  de  $vw$  où  $v_1 = vv_2$ .

Dans le deuxième cas, le mot  $v$  s'écrit  $v_1v_2$  et  $v_2w$  est une factorisation du mot  $w_1$ . Le couple  $(v_1\varphi, w_1\psi) = (v_1\varphi, v_2\psi e_0)$  appartient donc à  $P_0e_0$ . Réciproquement, si un couple  $(t, se_0)$  appartient à  $P_0e_0$ , le couple  $(t, s)$  est égal à  $(v_1\varphi, v_2\psi)$  pour une factorisation  $v_1v_2$  de  $v$ , et le couple  $(t, se_0)$  correspond à la factorisation  $v_1w_1$  de  $vw$  où  $w_1 = v_2w$ .  $\square$

Pour montrer que tout mot  $u$  vérifie la formule (6), on considère une factorisation ramseyenne  $u_0, u_1, u_2, \dots$  du mot  $u$  pour le morphisme  $\varphi \diamond \psi$ . On suppose que  $u_0(\varphi \diamond \psi) = s$  et  $u_1(\varphi \diamond \psi) = u_2(\varphi \diamond \psi) = \dots = e$  où  $(s, e)$  est un couple lié du semigroupe  $T \diamond S_f$ . Grâce au lemme précédent, il suffit de montrer que le mot  $u' = u_1u_2u_3\dots$  vérifie la formule (6). On pose

$$u_1(\varphi \diamond \psi) = e = \begin{pmatrix} f_0 & P_0 \\ 0 & e_0 \end{pmatrix} \quad u'(\varphi \diamond \psi) = e^\omega = \begin{pmatrix} Q_0 \\ e_0^\omega \end{pmatrix}$$

Puisque  $u_1\psi = u_2\psi = \dots = e_0$ , l'image de  $u'$  est  $u'\psi = e_0^\omega$ .

Soit  $vw$  une factorisation du mot  $u'$ . Il existe un entier  $k \geq 1$  et une factorisation  $v_1w_1$  du mot  $u_k$  telle que  $v = u_1u_2\dots u_{k-1}v_1$  et  $w = w_1u_{k+1}u_{k+2}\dots$ . Le couple  $(v_1\varphi, w_1\psi)$  appartient à la partie  $P_0$  et le couple  $(v\psi, w\psi) = (f_0^k v_1\varphi, w_1\psi e_0^\omega)$  appartient à  $Q_0$ . Réciproquement, tout couple  $(t, e)$  de  $Q_0$  est de la forme  $(f_0^k t', s'e_0^\omega)$  où le couple  $(t', s')$  appartient à  $P_0$ . Il existe donc une factorisation  $v_1w_1$  de  $u_k$  telle que  $v_1\varphi = t'$  et  $w_1\psi = s'$ . Le couple  $(t, e)$  correspond alors à la factorisation  $vw$  de  $u'$  où  $v = u_1u_2\dots u_{k-1}v_1$  et  $w = w_1u_{k+1}u_{k+2}\dots$ .  $\square$

### 2.8.3 Produit en couronne

Soient  $S$  et  $T$  deux semigroupes. Le produit de  $s_1$  et  $s_2 \in S$  sera noté  $s_1.s_2$ . Pour  $t_1, \dots, t_n$  des éléments de  $T$  et  $f$  une application de  $T$  dans  $S$ , l'image du produit  $t_1\dots t_n$  par  $f$  sera notée  $t_1\dots t_n f$ . Le *produit en couronne*  $S \circ T$  est le semigroupe  $S^{T^1} \times T$  muni du produit

$$(f_1, t_1)(f_2, t_2) = (t \mapsto tf_1.tt_1f_2, t_1t_2).$$

Soient  $(S_f, S_i)$  un  $\omega$ -semigroupe. Le produit de  $S_f$  et l'action à gauche de  $S_f$  sur  $S_i$  seront notés par un point. Ainsi pour  $s_1, s_2 \in S_f$  et  $e \in S_i$ , le produit  $s_1.s_2.e$  dénotera sans ambiguïté  $(s_1.s_2).e = s_1.(s_2.e)$ .

Le lemme suivant montre que pour un  $\omega$ -semigroupe fini, il suffit de définir l'opération  $\omega$  d'un idempotent.

**Lemme 8** *Soient  $S_f$  un semigroupe et  $S_i$  un ensemble fini munis d'une action à gauche de  $S_f$  sur  $S_i$  et d'une application  $\omega$  de  $E(S_f)$  dans  $S_i$  telle que pour tout  $s_1$  et  $s_2$  tels que  $s_1s_2$  et  $s_2s_1$  soient idempotents,  $(s_1s_2)^\omega = s_1(s_2s_1)^\omega$  alors  $(S_f, S_i)$  est naturellement muni d'une structure de  $\omega$ -semigroupe.*

**Démonstration :** Afin d'assurer  $(s^n)^\omega = s^\omega$  pour tout entier  $n$ , il est nécessaire de poser  $s^\omega = (s^\pi)^\omega$  où  $\pi$  est l'exposant du semigroupe  $S_f$ . Soient  $s_1$  et  $s_2$  deux éléments de  $S_f$ . On vérifie que  $(s_1s_2)^\omega = ((s_1s_2)^\pi)^\omega = s_1((s_2s_1)^\pi)^\omega = s_1(s_2s_1)^\omega$ .  $\square$

Soit  $(S_f, S_i)$  un  $\omega$ -semigroupe fini et  $T$  un semigroupe fini. Le produit en couronne  $(S_f, S_i) \circ T$  est égal à l' $\omega$ -semigroupe  $(S_f \circ T, S_i^{T^1})$ . L'action à gauche de  $S_f \circ T$  sur  $S_i^{T^1}$  est définie par

$$(f_1, t_1)g_1 = t \mapsto tf_1.tt_1g_1.$$

Pour  $(f_1, t_1)$  un idempotent de  $S_f \circ T$ , l'opération  $\omega$  est définie pour  $t \in T^1$  par

$$(f_1, t_1)^\omega = t \mapsto tf_1.(tt_1f_1)^\omega.$$

On peut remarquer que l'expression de  $(f_1, t_1)(f_2, t_2)^\omega$  se simplifie lorsque  $((f_1, t_1), (f_2, t_2))$  est un couple lié. Dans ce cas, on a les relations  $t_1t_2 = t_1$ ,  $t_2^2 = t_2$  et  $tf_1.tt_1f_2 = tf_1$  pour tout  $t \in T$ . On a alors l'expression

$$\begin{aligned} (f_1, t_1)(f_2, t_2)^\omega &= t \mapsto tf_1.tt_1f_2(tt_1t_2f_2)^\omega \\ &= t \mapsto tf_1(tt_1f_2)^\omega \end{aligned} \quad (7)$$

Vérifions que ceci définit une action à gauche. Soient  $(f_1, t_1), (f_2, t_2) \in S_f \circ T$  et  $g_1 \in S_i^{T^1}$ .

$$\begin{aligned} ((f_1, t_1)(f_2, t_2))g_1 &= (t \mapsto tf_1.tt_1f_2, t_1t_2)g_1 \\ &= t \mapsto (tf_1.tt_1f_2).(tt_1t_2g_1) \\ &= t \mapsto tf_1.tt_1f_2.tt_1t_2g_1 \\ (f_1, t_1)((f_2, t_2)g_1) &= (f_1, t_1)(t \mapsto tf_2.tt_2g_1) \\ &= t \mapsto (tf_1).(tt_1f_2.tt_1t_2g_1) \\ &= t \mapsto tf_1.tt_1f_2.tt_1t_2g_1 \end{aligned}$$

Vérifions que  $((f_1, t_1)(f_2, t_2))^\omega = (f_1, t_1)((f_2, t_2)(f_1, t_1))^\omega$  pour deux éléments  $(f_1, t_1)$  et  $(f_2, t_2)$  de  $S_f \circ T$  tels que  $(f_1, t_1)(f_2, t_2)$  et  $(f_2, t_2)(f_1, t_1)$  soient idempotents

$$\begin{aligned}
((f_1, t_1)(f_2, t_2))^\omega &= (t \mapsto tf_1.tt_1f_2, t_1t_2)^\omega \\
&= t \mapsto (tf_1.tt_1f_2).(tt_1t_2f_1.tt_1t_2t_1f_2)^\omega \\
&= t \mapsto tf_1.tt_1f_2.(tt_1t_2f_1.tt_1t_2t_1f_2)^\omega \\
(f_1, t_1)((f_2, t_2)(f_1, t_1))^\omega &= (f_1, t_1)(t \mapsto tf_2.tt_2f_1.(tt_2t_1f_2.tt_2t_1t_2f_1)^\omega) \\
&= t \mapsto (tf_1).(tt_1f_2.tt_1t_2f_1.(tt_1t_2t_1f_2.tt_1t_2t_1t_2f_1)^\omega) \\
&= t \mapsto tf_1.tt_1f_2.tt_1t_2f_1.(tt_1t_2t_1f_2.tt_1t_2t_1t_2f_1)^\omega \\
&= t \mapsto tf_1.tt_1f_2.tt_1t_2f_1.tt_1t_2t_1f_2(tt_1t_2t_1t_2f_1.tt_1t_2t_1f_2)^\omega
\end{aligned}$$

Puisque l'élément  $(f_1, t_1)(f_2, t_2)$  de  $S_f \circ T$  est idempotent, on a les égalités

$$\begin{aligned}
t_1t_2t_1t_2 &= t_1t_2 \\
tf_1.tt_1f_2.tt_1t_2f_1.tt_1t_2t_1f_2 &= tf_1.tt_1f_2
\end{aligned}$$

ce qui termine la vérification.

**Exemple :** On considère le produit en couronne  $\widehat{S} \circ \mathbb{Z}/2\mathbb{Z}$  où  $S$  est le semigroupe multiplicatif  $S = \{a = 0, b = 1\}$  et  $\mathbb{Z}/2\mathbb{Z}$  le groupe cyclique d'ordre 2. La structure de cet  $\omega$ -semigroupe est donné à la figure 2.1. À gauche est donnée la structure en  $\mathcal{D}$ -classes de la partie finitaire, c'est à dire du semigroupe  $S \circ \mathbb{Z}/2\mathbb{Z}$ . L'élément  $(f, s)$  de  $S \circ \mathbb{Z}/2\mathbb{Z}$  où  $f$  est une fonction de  $\mathbb{Z}/2\mathbb{Z}$  dans  $S$  est noté par le triplet  $(0f, 1f, s)$ . À droite, est associé à chaque classe de conjugaison de couples liés, l'élément correspondant de la partie infinitaire de  $\widehat{S} \circ \mathbb{Z}/2\mathbb{Z}$ . La fonction  $g$  de  $\mathbb{Z}/2\mathbb{Z}$  dans la partie infinitaire de  $\widehat{S}$  est notée par le couple  $(0g, 1g)$ . On remarque en effet qu'il y a une exacte correspondance entre les classes de conjugaison de couples liés et les éléments de la partie infinitaire de  $\widehat{S} \circ \mathbb{Z}/2\mathbb{Z}$ .

De façon plus générale, on a le résultat

**Proposition 12** *Soient  $S$  un semigroupe fini et  $G$  un groupe fini. On a*

$$\widehat{S} \circ G = \widehat{S \circ G}$$

**Démonstration :** Les deux  $\omega$ -semigroupes ont la même partie finitaire, à savoir le semigroupe  $S \circ G$ . Le  $\omega$ -semigroupe  $\widehat{S} \circ G$  est donc un quotient du  $\omega$ -semigroupe  $\widehat{S \circ G}$  d'après la proposition 11. Il suffit donc de montrer que si deux couples liés  $((f_1, s_1), (g_1, e_1))$  et  $((f_2, s_2), (g_2, e_2))$  du semigroupe  $S \circ G$  vérifient  $(f_1, s_1)(g_1, e_1)^\omega = (f_2, s_2)(g_2, e_2)^\omega$  dans  $\widehat{S \circ G}$ , alors les deux couples sont conjugués. On est alors assuré de la même égalité dans  $\widehat{S} \circ G$  ce qui montre que

Structure en  $\mathcal{D}$ -classes

|   |              |             |             |              |
|---|--------------|-------------|-------------|--------------|
| $\begin{matrix} * & (b, b, 0) \\ & (b, b, 1) \end{matrix}$  |              |             |             |              |
| <table style="border-collapse: collapse; margin: 0 auto;"> <tr> <td style="border: 1px solid black; padding: 2px;"><math>*(a, b, 0)</math></td> <td style="border: 1px solid black; padding: 2px;"><math>(a, b, 1)</math></td> </tr> <tr> <td style="border: 1px solid black; padding: 2px;"><math>(b, a, 1)</math></td> <td style="border: 1px solid black; padding: 2px;"><math>*(b, a, 0)</math></td> </tr> </table> | $*(a, b, 0)$ | $(a, b, 1)$ | $(b, a, 1)$ | $*(b, a, 0)$ |
| $*(a, b, 0)$  | $(a, b, 1)$  |             |             |              |
| $(b, a, 1)$   | $*(b, a, 0)$ |             |             |              |
| $\begin{matrix} * & (a, a, 0) \\ & (a, a, 1) \end{matrix}$  |              |             |             |              |

|                             |                          |
|-----------------------------|--------------------------|
| $(b, b, 0)(b, b, 0)^\omega$ | $(b^\omega, b^\omega)$   |
| $(a, b, 0)(b, b, 0)^\omega$ | $(ab^\omega, b^\omega)$  |
| $(b, a, 0)(b, b, 0)^\omega$ | $(b^\omega, ab^\omega)$  |
| $(a, a, 0)(b, b, 0)^\omega$ | $(ab^\omega, ab^\omega)$ |
| $(a, b, 1)(b, a, 0)^\omega$ | $(a^\omega, b^\omega)$   |
| $(b, a, 1)(a, b, 0)^\omega$ | $(b^\omega, a^\omega)$   |
| $(a, a, 0)(a, b, 0)^\omega$ | $(a^\omega, ab^\omega)$  |
| $(a, a, 1)(b, a, 0)^\omega$ | $(a^\omega, ab^\omega)$  |
| $(a, a, 0)(b, a, 0)^\omega$ | $(ab^\omega, a^\omega)$  |
| $(a, a, 1)(a, b, 0)^\omega$ | $(ab^\omega, a^\omega)$  |
| $(a, a, 0)(a, a, 0)^\omega$ | $(a^\omega, a^\omega)$   |
| $(a, a, 1)(a, a, 0)^\omega$ | $(a^\omega, a^\omega)$   |

Figure 2.1 : Structure de  $\hat{S} \circ \mathbb{Z}/2\mathbb{Z}$ 

les deux  $\omega$ -semigroupes sont isomorphes. Puisque  $e_1$  et  $e_2$  sont deux idempotents du groupe  $G$ , on a nécessairement  $e_1 = e_2 = 1_G$ . On suppose que pour tout  $t \in G$

$$tf_1(ts_1g_1)^\omega = tf_2(ts_2g_2)^\omega$$

Les deux couples liés  $(tf_1, ts_1g_1)$  et  $(tf_2, ts_2g_2)$  de  $S$  sont conjugués. Il existe donc deux éléments  $tx_1$  et  $tx_2$  de  $S$  tels que  $tf_1.tx_1 = tf_2$ ,  $tx_1.tx_2 = ts_1g_1$  et  $tx_2.tx_1 = ts_2g_2$ . On définit alors les deux fonctions  $h_1$  et  $h_2$  de  $G$  dans  $S$  en posant  $th_1 = ts_1^{-1}x_1$  et  $th_2 = ts_2^{-1}x_2$  pour tout  $t \in G$ . On vérifie alors que les éléments  $(h_1, s_1^{-1}s_2)$  et  $(h_2, s_2^{-1}s_1)$  de  $S \circ G$  vérifient

$$\begin{aligned} (f_1, s_1)(h_1, s_1^{-1}s_2) &= (f_2, s_2) \\ (h_1, s_1^{-1}s_2)(h_2, s_2^{-1}s_1) &= (g_1, 1_G) \\ (h_2, s_2^{-1}s_1)(h_1, s_1^{-1}s_2) &= (g_2, 1_G) \end{aligned}$$

ce qui montre que les couples liés  $((f_1, s_1), (g_1, e_1))$  et  $((f_2, s_2), (g_2, e_2))$  sont bien conjugués  $\square$

Si le langage  $L$  de mots finis est reconnu par le semigroupe  $S$  et si  $\sigma$  est une fonction séquentielle réalisée par un transducteur  $T$ , le langage  $L\sigma^{-1}$  est reconnu

par le semigroupe  $S \circ M(\sigma)$  où  $M(\sigma)$  est le monoïde de transition du transducteur  $T$  (cf. [Pin84]). Le théorème suivant étend ce résultat aux  $\omega$ -langages. Un exposé plus complet sur les fonctions séquentielles et les transducteurs peut être trouvé dans [Ber79].

**Théorème 15** *Soit  $\sigma : A^* \rightarrow B^*$  une fonction séquentielle réalisée par un transducteur  $T = (Q, A, B, q_0, \cdot, *)$  et soit  $M(\sigma)$  le monoïde de transition de l'automate  $(Q, A, q_0, \cdot)$ . Si le  $\omega$ -langage  $L \subset B^\omega$  est reconnu par le  $\omega$ -semigroupe  $S = (S_f, S_i)$ , alors le  $\omega$ -langage  $L\sigma^{-1} \subset A^\omega$  est reconnu par le  $\omega$ -semigroupe  $S \circ M(\sigma)$ .*

**Exemple :** On considère la fonction séquentielle  $\sigma$  réalisé par le transducteur de la figure 2.2. De manière intuitive, ce transducteur ne conserve que les  $a$  en position paire (les lettres d'un mot sont numérotées à partir de zéro). Les lettres  $a$  en position impaire sont transformées en lettre  $b$ . Le langage  $X_1 = (A^*a)^\omega$  des mots ayant une infinité d'occurrences de  $a$  est reconnu par  $\omega$ -semigroupe  $\hat{S}$  où  $S$  est le semigroupe  $S = \{a = 0, b = 1\}$ . Le langage  $X_1\sigma^{-1}$  est l'ensemble des mots ayant une infinité d'occurrences de  $a$  en position paire. Le monoïde de transition du transducteur est  $\mathbb{Z}/2\mathbb{Z}$  puisque chacune des lettres  $a$  et  $b$  échange les états 1 et 2. D'après le théorème précédent, le langage  $X_1\sigma^{-1}$  est donc reconnu par le  $\omega$ -semigroupe  $\hat{S} \circ \mathbb{Z}/2\mathbb{Z}$ .

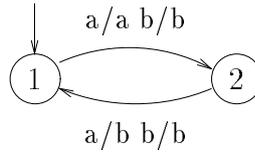


Figure 2.2 : Transducteur

**Démonstration :** Soit  $\varphi : B^\omega \rightarrow S$  un morphisme surjectif reconnaissant  $L$ . On note  $\mu : A^* \rightarrow M(\sigma)$  le morphisme induit par l'action des lettres. On définit le morphisme  $\psi : A^\omega \rightarrow S \circ M(\sigma)$  pour toute lettre  $a \in A$  par

$$a\psi = (t \mapsto (q_0t * a)\varphi, a\mu)$$

**Lemme 9** *Pour tout mot  $u \in A^*$ , on a*

$$u\psi = (t \mapsto (q_0t * u)\varphi, u\mu)$$

**Démonstration :** On montre le résultat par récurrence sur la longueur  $|u|$  du mot  $u$ . Le résultat est acquis pour les mots de longueur 1 par définition du morphisme  $\psi$ . Si  $u$  est de longueur supérieure à 1, on écrit  $u = va$  où  $a$  est une lettre et  $v$  un mot de longueur inférieure. On a alors

$$\begin{aligned}
u\psi &= v\psi.a\psi \\
&= (t \mapsto (q_0t * v)\varphi, v\mu)(t \mapsto (q_0t * a)\varphi, a\mu) \\
&= (t \mapsto (q_0t * v)\varphi(q_0tv\mu * a)\varphi, v\mu a\mu) \\
&= (t \mapsto (q_0t * v)(q_0t.v * a)\varphi, va\mu) \\
&= (t \mapsto (q_0t * va)\varphi, va\mu) \\
&= (t \mapsto (q_0t * u)\varphi, u\mu)
\end{aligned}$$

ce qui termine la preuve du lemme.  $\square$

Nous étendons maintenant ce résultat au mots infinis.

**Lemme 10** *Pour tout mot  $u \in A^\omega$ , on a*

$$u\psi = (t \mapsto (q_0t * u)\varphi).$$

**Démonstration :** On suppose l'élément  $t \in M(\sigma)$  fixé. On considère une factorisation  $u = u_0u_1u_2\dots$  telle que  $u_0\psi = s = (f_1, t_1)$  et  $u_1\psi = u_2\psi = \dots = e = (f_2, t_2)$  où  $(s, e)$  est un couple lié du semigroupe  $S_f \circ M(\sigma)$ . En particulier on a les relations  $t_1t_2 = t_1$  et  $t_2^2 = t_2$ . On définit les états  $q_1 = q_0t.u_0$  et  $q_{i+1} = q_i.u_i$  pour  $i \geq 1$ . On peut alors écrire la factorisation  $q_0t*u = v_0v_1v_2\dots$  où  $v_0 = q_0t*u_0$  et  $v_i = q_i * u_i$  pour  $i \geq 1$ . D'après le lemme précédent, on a alors les égalités

$$\begin{aligned}
v_0\varphi &= q_0t * u_0 \\
&= tf_1 \\
v_i\varphi &= q_i * v_i \\
&= q_0t.(u_0\dots u_{i-1}) * u_i \\
&= q_0t(u_0\dots u_{i-1})\mu * u_i \\
&= q_0tt_1 * u_i \\
&= tt_1f_2
\end{aligned}$$

Puisque  $(s, e)$  est un couple lié, on a la relation  $tf_1.tt_1f_2 = tf_1$ , ce qui montre que le couple  $(tf_1, tt_1f_2)$  est lié. On a donc  $(q_0t*u)\varphi = tf_1.(tt_1f_2)^\omega$  ce qui correspond à la formule 7 (p. 41) et termine ainsi la preuve du lemme.  $\square$

Pour montrer que le  $\omega$ -semigroupe  $S \circ M(\sigma)$  reconnaît le langage  $L\sigma^{-1}$ , on constate que le mot  $u \in A^\omega$  vérifie  $u\sigma \in L$  si et seulement si  $(q_0 * u)\varphi$  appartient à l'image de  $L$  dans  $S$ , c'est à dire si  $1(u\psi) \in L\varphi$ .  $\square$

Soit  $(Q, A, q_0, E, T)$  un automate de Muller ou un automate à table de transitions dont l'ensemble des transitions est  $E = \{f_1, f_2, \dots, f_n\}$ . L'automate peut être considéré comme un transducteur, en associant à la transition  $f_i$  la production  $f_i$ . L'automate considéré comme transducteur réalise une fonction séquentielle  $\sigma$  de  $A^*$  dans  $E^*$ . Soit  $X \subset E^\omega$  l'ensemble des chemins réussis de l'automate. Le  $\omega$ -langage  $L \subset A^\omega$  reconnu par l'automate est égal à  $L = X\sigma^{-1}$ . D'après le théorème précédent, il suffit donc de construire un  $\omega$ -semigroupe reconnaissant le  $\omega$ -langage  $X$  des chemins réussis pour construire un  $\omega$ -semigroupe reconnaissant le  $\omega$ -langage  $L$ .

On commence par rappeler un résultat de I. Simon [Eil76]. Étant donné un automate  $\mathcal{A}$ , une congruence de chemins est une relation d'équivalence entre les chemins finis vérifiant les deux conditions suivantes

- (1) Deux chemins équivalents sont coterminaux, *i.e.*, ont même origine et même extrémité;
- (2) Si  $p$  et  $q$  sont deux chemins équivalents et si  $r, p$  et  $s$  sont des chemins consécutifs, alors les chemins  $rps$  et  $rqs$  sont équivalents.

On a alors la proposition suivante

**Proposition 13** (Simon) *Soit  $\sim$  une congruence de chemins, telle que pour tout couple de boucles  $p$  et  $q$  autour d'un même état,  $p^2 \sim p$  et  $pq \sim qp$ . Alors deux chemins coterminaux ayant même contenu sont équivalents.*

Soit  $\mathcal{A}$  un automate d'ensemble de transitions  $E = \{f_1, f_2, \dots, f_n\}$  et soit  $\sim$  la congruence des chemins de  $\mathcal{A}$  engendrée par les relations  $p^2 \sim p$  et  $pq \sim qp$  pour tout couple de boucles  $p$  et  $q$  autour d'un même état. L'ensemble  $S = \{0\} \cup S'$  où  $S'$  est l'ensemble des classes d'équivalences de chemins modulo la congruence  $\sim$  est naturellement muni d'une structure de semigroupe. Le produit  $pq$  de deux chemins  $p$  et  $q$  est égal à la concaténation de  $p$  et  $q$  si  $p$  et  $q$  sont consécutifs. Le produit est égal à 0 sinon. On peut remarquer que le semigroupe  $S$  est égal à  $E^*/\equiv$  où  $\equiv$  est la congruence engendrée par les relations  $pq \equiv 0$  si  $p$  et  $q$  non consécutifs,  $p^2 \equiv p$  et  $pq \equiv qp$  pour tout couple de boucles  $p$  et  $q$  autour d'un même état.

Pour un automate à table transition, la proposition suivante montre qu'à partir du semigroupe des chemins considéré précédemment, il est possible de construire un  $\omega$ -semigroupe qui reconnaît le  $\omega$ -langage des chemins réussis.

**Proposition 14** *Pour un automate à table de transitions, le  $\omega$ -semigroupe  $\hat{S}$  reconnaît le  $\omega$ -langage des chemins réussis.*

**Démonstration :** Pour le semigroupe  $S$  un couple  $(s, e)$  est lié si le chemin  $e$  est une boucle autour de l'extrémité du chemin  $s$  (en particulier les chemins  $s$  et  $e$  sont consécutifs) et si le contenu de cette boucle est inclus dans le contenu du chemin  $s$ . On définit la partie  $P$  de  $\hat{S}$  comme l'ensemble des classes de conjugaison  $[s, e]$  telles que l'origine du chemin  $s$  est l'état initial et telles que le contenu de la boucle  $e$  est un élément de la table de transitions. Cette définition est cohérente. D'une part, deux chemins  $\mathcal{R}$ -équivalents ont nécessairement même origine. D'autre part, si  $c_1$  et  $c_2$  sont deux chemins consécutifs, le contenu du produit  $pq$  est égal à l'union des contenus de  $p$  et de  $q$ . Deux chemins  $\mathcal{D}$  équivalents ont donc nécessairement même contenu. Si  $(s_1, e_1)$  et  $(s_2, e_2)$  sont deux couples liés conjugués, les chemins  $s_1$  et  $s_2$  ont même origine et les deux boucles  $e_1$  et  $e_2$  ont même contenu.

Un chemin infini  $c$  d'un automate se factorise  $c = c_0c_1c_2 \dots$  où  $c_0$  est chemin de l'état initial  $q_0$  à un état  $q$  et les chemins  $c_i$  pour  $i \geq 1$  sont des boucles autour de l'état  $q$  et ont  $\text{Inr}(c)$  pour contenu. Le  $\omega$ -semigroupe  $\hat{S}$  reconnaît donc l'ensemble des chemins réussis de l'automate.  $\square$



## Chapitre 3

# Chaînes d'ensembles

### 3.1 Introduction

Nous donnons dans ce chapitre une construction de l'algèbre de Boole  $\mathcal{B}$  engendrée par une semi-algèbre de Boole  $\mathcal{S}$ . On sait que tout élément de  $\mathcal{B}$  peut s'écrire comme une union de différences d'éléments de  $\mathcal{S}$ . Nous considérons plus particulièrement les chaînes qui sont des suites de différences emboîtées. Nous donnons une nouvelle démonstration, à l'aide de diagrammes, d'un théorème de Hausdorff (cf. [Hau57]) qui affirme que l'on peut supposer que ces unions de différences sont des chaînes. Nous obtenons une construction explicite de chaînes pour l'union et l'intersection de deux parties connaissant une chaîne pour chacune de ces deux parties.

### 3.2 Notations

Dans ce chapitre uniquement, nous noterons par commodité l'intersection comme un produit et l'union comme une addition :  $XY = X \cap Y$  et  $X + Y = X \cup Y$ . Cette notation est cohérente car l'intersection est distributive par rapport à l'union, *i.e.*,  $X(Y + Z) = XY + XZ$  avec les notations adoptées. Nous noterons  $X \Delta Y = X + Y - XY$  la différence symétrique de  $X$  et de  $Y$ . Il faut remarquer que le  $-$  se comporte comme un produit en distribuant par rapport à  $+$  mais qu'il ne se comporte pas de façon associative avec  $+$ . Ainsi  $(X + Y) - Z$  est égal à  $(X - Z) + (Y - Z)$  mais est en général différent de  $X + (Y - Z)$ . On a cependant l'égalité des deux expressions dans le cas où  $X \cap Z = \emptyset$ . L'ensemble  $\mathcal{P}(E)$  des parties de  $E$  muni de la différence symétrique comme addition et de l'intersection comme produit est un anneau. En effet, celui-ci est isomorphe au produit  $\prod_{x \in E} \mathbb{Z}/2\mathbb{Z}$  en identifiant une partie de  $E$  et sa fonction caractéristique.

Soit  $E$  un ensemble et soit  $\mathcal{F}$  un ensemble de parties de  $E$ . On dit que  $\mathcal{F}$  est une *algèbre de Boole* s'il contient  $E$  et  $\emptyset$  et s'il est fermé par union, intersection et passage au complémentaire.

Si  $\mathcal{F}$  contient  $E$  et  $\emptyset$ , est fermé par union et intersection, mais pas nécessairement par passage au complémentaire, on dit que c'est une *semi-algèbre de Boole*. L'exemple le plus courant d'une telle semi-algèbre est l'ensemble des ouverts d'un espace topologique.

### 3.3 Sommes de différences et chaînes

Nous définissons une *somme de différences* de longueur  $m$  comme une expression

$$\begin{aligned} X &= (X_1 - X_2) + (X_3 - X_4) + \cdots + (X_{m-1} - X_m) && \text{si } m \text{ est pair} \\ X &= (X_1 - X_2) + (X_3 - X_4) + \cdots + X_m && \text{si } m \text{ est impair} \end{aligned}$$

Nous définissons une *chaîne de différences* (ou plus simplement une *chaîne*) comme une somme de différences où les parties  $X_i$  vérifient en outre la condition  $X_1 \supset X_2 \supset \dots \supset X_m$ . Le fait que la suite  $X_i$  soit décroissante rend le parenthésage d'une telle expression équivalent au parenthésage gauche et nous noterons

$$X = X_1 - X_2 + X_3 - \cdots \pm X_m$$

où le signe  $\pm$  devant  $X_m$  dépend de la parité de  $m$ . L'expression est alors aussi équivalente à

$$X = X_1 \Delta X_2 \Delta \dots \Delta X_m$$

Soit  $\mathcal{F}$  une semi-algèbre de Boole d'un ensemble  $E$ . Nous définissons  $C_m(\mathcal{F})$  de la façon suivante. Pour  $m = 0$  nous posons  $C_0(\mathcal{F}) = \{\emptyset\}$  et pour  $m \geq 1$  nous notons  $C_m(\mathcal{F})$  l'ensemble des parties  $X$  de  $E$  qui s'écrivent comme une chaîne de longueur  $m$  avec des éléments de  $\mathcal{F}$ , *i.e.*,

$$X = X_1 - X_2 + X_3 - \cdots \pm X_m \quad X_i \in \mathcal{F}$$

Nous omettrons dorénavant la référence à la semi-algèbre de Boole fixée  $\mathcal{F}$  et nous écrirons  $C_m$ . Pour  $m = 1$  nous avons  $C_1 = \mathcal{F}$ . Si  $m' \leq m$ , nous avons  $C_{m'} \subset C_m$  car  $\emptyset \in \mathcal{F}$  et l'on peut poser  $X_{m'+1} = \cdots = X_m = \emptyset$  pour passer d'une chaîne de longueur  $m'$  à une chaîne de longueur  $m$ . Si  $\mathcal{F}$  se trouve être une algèbre de Boole, alors  $C_m = \mathcal{F}$  pour tout  $m \geq 1$ .

**Proposition 15** *Si  $X \in C_m$  et  $Y \in C_n$ , alors*

$$XY \in \begin{cases} C_{m+n-2} & \text{si } m \text{ et } n \text{ pairs} \\ C_{m+n-1} & \text{sinon} \end{cases}$$

$$X + Y \in \begin{cases} C_{m+n-1} & \text{si } m \text{ et } n \text{ impairs} \\ C_{m+n} & \text{sinon} \end{cases}$$

**Démonstration :** Nous allons montrer qu'à partir d'une chaîne de longueur  $m$  pour  $X \in C_m$  et d'une chaîne de longueur  $n$  pour  $Y \in C_n$ , il est possible de construire des chaînes des longueurs énoncées ci-dessus pour les ensembles  $XY$  et  $X + Y$ .

Nous commençons par traiter quelques exemples pour les premières valeurs de  $m$  et  $n$ . Les cas  $m = 0$  ou  $n = 0$  sont triviaux puisque  $C_0 = \{\emptyset\}$ .

Nous considérons le cas  $n = 1$  ( $Y = Y_1 \in \mathcal{F}$ ). Le cas  $m = 1$  ( $X \in \mathcal{F}$ ) est évident car  $XY$  et  $X + Y \in \mathcal{F}$  car la semi-algèbre  $\mathcal{F}$  est stable pour l'intersection et l'union. Pour  $m = 2$ , la partie  $X$  s'écrit  $X = X_1 - X_2$ . Nous avons alors les chaînes

$$\begin{aligned} XY &= X_1Y_1 - X_2Y_1 \\ X + Y &= (X_1 + Y_1) - (X_2 + Y_1) + Y_1 && \text{Solution 1} \\ X + Y &= (X_1 + Y_1) - X_2 + X_2Y_1 && \text{Solution 2} \end{aligned}$$

Pour  $m = 3$ , la partie  $X$  s'écrit  $X = X_1 - X_2 + X_3$ . Nous avons alors les chaînes

$$\begin{aligned} XY &= X_1Y_1 - X_2Y_1 + X_3Y_1 \\ X + Y &= (X_1 + Y_1) - (X_2 + Y_1) + (X_3 + Y_1) && \text{Solution 1} \\ X + Y &= (X_1 + Y_1) - X_2 + (X_3 + X_2Y_1) && \text{Solution 2} \end{aligned}$$

Ces formules se généralisent aisément pour  $m$  quelconque. La partie  $X$  s'écrit  $X = X_1 - X_2 + X_3 - \dots \pm X_m$  et nous avons les chaînes

$$\begin{aligned} XY &= X_1Y_1 - X_2Y_1 + X_3Y_1 - \dots \pm X_mY_1 && \text{Solution 1} \\ X + Y &= (X_1 + Y_1) - (X_2 + Y_1) + \dots - (X_m + Y_1) + Y_1 && \text{si } m \text{ est pair} \\ X + Y &= (X_1 + Y_1) - (X_2 + Y_1) + \dots + (X_m + Y_1) && \text{si } m \text{ est impair} \\ X + Y &= (X_1 + Y_1) - X_2 + (X_3 + X_2Y_1) - X_4 && \text{Solution 2} \\ &\quad + (X_5 + X_4Y_1) - \dots - X_m + X_mY_1 && \text{si } m \text{ est pair} \\ X + Y &= (X_1 + Y_1) - X_2 + (X_3 + X_2Y_1) - X_4 && \text{si } m \text{ est impair} \\ &\quad + (X_5 + X_4Y_1) - \dots + (X_m + X_{m-1}Y_1) \end{aligned}$$

Nous considérons maintenant le cas  $n = 2$  ( $Y = Y_1 - Y_2$ ). Le cas  $m = 1$  est symétrique du cas  $n = 1$  et  $m = 2$ . Pour  $m = 2$  ( $X = X_1 - X_2$ ), nous avons les chaînes

$$\begin{aligned} XY &= X_1Y_1 - (X_1Y_2 + X_2Y_1) \\ X + Y &= (X_1 + Y_1) - (X_2 + Y_2) + (X_1Y_2 + X_2Y_1) - X_2Y_2 \end{aligned}$$

Pour  $m = 3$  ( $X = X_1 - X_2 + X_3$ ), nous avons les chaînes

$$\begin{aligned} XY &= X_1Y_1 - (X_1Y_2 + X_2Y_1) + X_3Y_1 - X_3Y_2 \\ X + Y &= (X_1 + Y_1) - (X_2 + Y_2) + (X_1Y_2 + X_2Y_1 + X_3) - X_2Y_2 + X_3Y_2 \end{aligned}$$

Nous revenons maintenant au cas général. Nous donnons la construction de diagrammes de représentation des chaînes. A la chaîne

$$X = X_1 - X_2 + X_3 - \cdots \pm X_m$$

nous associons un diagramme de Venn construit dans le plan. A  $E$ , correspond le demi-plan droit  $\mathcal{P}_0 = \{(x, y) \mid x \geq 0\}$  et à  $X_i$  le demi-plan droit  $\mathcal{P}_i = \{(x, y) \mid x \geq i\}$ . Nous avons bien  $\mathcal{P}_{i+1} \subset \mathcal{P}_i$  pour  $0 \leq i \leq n - 1$ . Le plan se trouve ainsi découpé en bandes verticales dont chacune représente une des différences  $X_i - X_{i+1}$  (cf. figure 3.1). La figure a été volontairement limitée à la région significative du plan. En face de chacune des droites verticales d'équation  $X = i$  est inscrite la partie  $X_i$  délimitée par celle-ci. Les bandes verticales représentant des différences incluses dans la partie  $X$  sont grisées sur la figure.

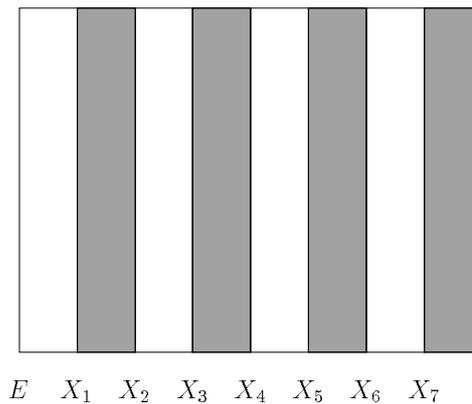


Figure 3.1 :  $X = X_1 - X_2 + X_3 - X_4 + X_5 - X_6 + X_7$

A la chaîne

$$Y = Y_1 - Y_2 + Y_3 - \cdots \pm Y_n$$

nous associons comme précédemment un diagramme de Venn dans le plan, mais en inversant les rôles des coordonnées du plan. A  $E$  correspond le demi-plan supérieur  $\{(x, y) \mid y \geq 0\}$  et à  $Y_i$  le demi-plan supérieur  $\{(x, y) \mid y \geq i\}$ . Le plan se trouve ainsi découpé en bandes horizontales dont chacune représente une des différences  $Y_i - Y_{i+1}$  (cf. figure 3.2). En face de chacune des bandes horizontales est inscrite la différence représentée.

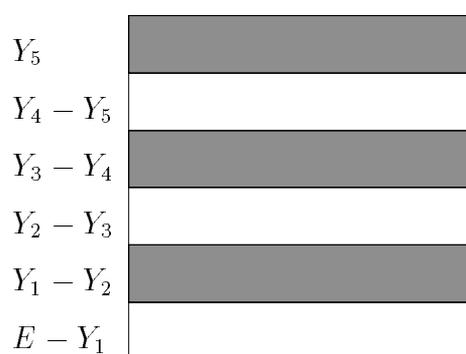


Figure 3.2 :  $Y = Y_1 - Y_2 + Y_3 - Y_4 + Y_5$

Pour obtenir un diagramme de Venn des parties  $XY$  et  $X + Y$  en fonction des parties  $X_i$  et  $Y_i$ , il suffit de superposer les deux diagrammes correspondants à  $X$  et à  $Y$  (cf. figures 3.3 et 3.4). Chacun des petits carrés élémentaires de ces diagrammes représente une intersection  $(X_i - X_{i+1})(Y_j - Y_{j+1})$ . Sur les figures, les carrés élémentaires inclus dans  $XY$  (respectivement  $X + Y$ ) sont grisés.

L'intersection élémentaire  $X_i Y_j$  est représentée sur le diagramme par la région du plan  $\{(x, y) \mid x \geq i \text{ et } y \geq j\}$  c'est à dire un quart de plan supérieur droit (cf. figure 3.5).

Une partie  $P$  égale à une union de parties du type  $X_i Y_j$  ( $P = X_{i_1} Y_{j_1} + X_{i_2} Y_{j_2} + \dots + X_{i_k} Y_{j_k}$ ) est représentée sur le diagramme par une union de différents quarts de plan supérieurs droits. Cette région du diagramme se trouve délimitée à gauche et en bas par un chemin ne comportant que des pas élémentaires Est ou de pas élémentaires Sud (si le chemin est lu de haut vers le bas). Réciproquement, une région délimitée à gauche et en bas par un tel chemin est égale à une union de parties  $X_i Y_j$  (cf. figure 3.6). Deux parties  $P_1$  et  $P_2$  étant ainsi délimitées par deux chemins  $C_1$  et  $C_2$ , il suffit que le chemin  $C_1$  soit au dessus et à droite du chemin  $C_2$  pour que  $P_1 \subset P_2$ .

Si les parties  $X_i$  et  $Y_j$  appartiennent à la semi-algèbre  $\mathcal{F}$ , toute partie  $P$  ainsi déterminée appartient également à la semi-algèbre  $\mathcal{F}$ . La donnée d'un chemin

|             |     |       |       |       |       |       |       |       |
|-------------|-----|-------|-------|-------|-------|-------|-------|-------|
| $Y_5$       |     | ■     |       | ■     |       | ■     |       | ■     |
| $Y_4 - Y_5$ |     |       |       |       |       |       |       |       |
| $Y_3 - Y_4$ |     | ■     |       | ■     |       | ■     |       | ■     |
| $Y_2 - Y_3$ |     |       |       |       |       |       |       |       |
| $Y_1 - Y_2$ |     | ■     |       | ■     |       | ■     |       | ■     |
| $E - Y_1$   |     |       |       |       |       |       |       |       |
|             | $E$ | $X_1$ | $X_2$ | $X_3$ | $X_4$ | $X_5$ | $X_6$ | $X_7$ |

Figure 3.3 : Diagramme de Venn de l'intersection

|             |     |       |       |       |       |       |       |       |
|-------------|-----|-------|-------|-------|-------|-------|-------|-------|
| $Y_5$       | ■   | ■     | ■     | ■     | ■     | ■     | ■     | ■     |
| $Y_4 - Y_5$ |     | ■     |       | ■     |       | ■     |       | ■     |
| $Y_3 - Y_4$ | ■   | ■     | ■     | ■     | ■     | ■     | ■     | ■     |
| $Y_2 - Y_3$ |     | ■     |       | ■     |       | ■     |       | ■     |
| $Y_1 - Y_2$ | ■   | ■     | ■     | ■     | ■     | ■     | ■     | ■     |
| $E - Y_1$   |     | ■     |       | ■     |       | ■     |       | ■     |
|             | $E$ | $X_1$ | $X_2$ | $X_3$ | $X_4$ | $X_5$ | $X_6$ | $X_7$ |

Figure 3.4 : Diagramme de Venn de l'union

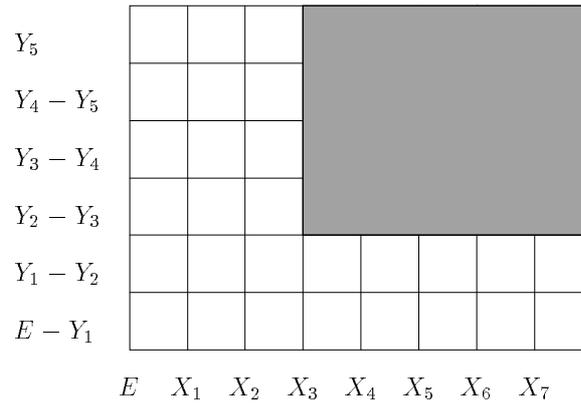


Figure 3.5 :  $X_3Y_2$

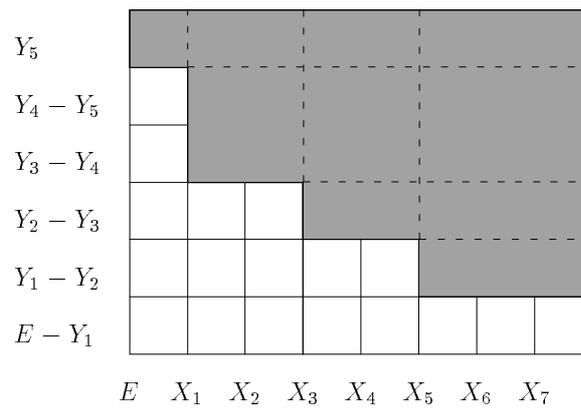


Figure 3.6 :  $X_5Y_1 + X_3Y_2 + X_1Y_3 + EY_5$

uniquement constitué de pas Est et Sud détermine ainsi une partie de la semi-algèbre  $\mathcal{F}$ . Nous allons exhiber des chaînes pour  $XY$  et  $X + Y$  constituées de telles parties. Pour ce faire, nous déterminerons les chemins délimitant ces parties. Nous allons illustrer cette construction en reprenant les exemples pour les premières valeurs de  $m$  et  $n$  (cf. figures 3.7, 3.8 et 3.9).

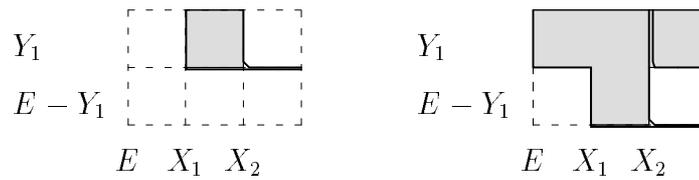


Figure 3.7 : Chaînes pour  $m = 2$  et  $n = 1$

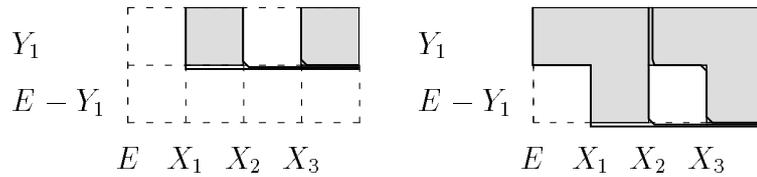


Figure 3.8 : Chaînes pour  $m = 3$  et  $n = 1$

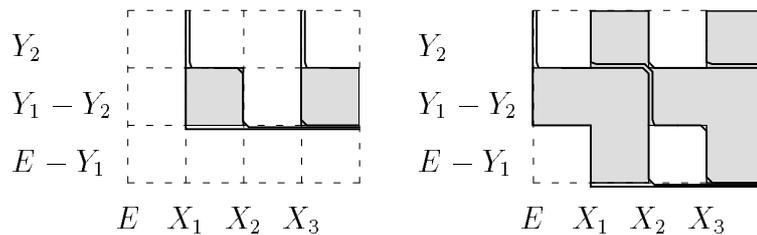


Figure 3.9 : Chaînes pour  $m = 3$  et  $n = 2$

La construction des chemins pour les chaînes de  $XY$  et de  $X + Y$  se généralise sans difficulté pour  $m$  et  $n$  quelconques (cf. figures 3.10 et 3.11).

Nous faisons remarquer que ces solutions ne sont pas uniques. Lorsque deux chemins délimitant deux parties consécutives d'une chaîne coïncident partiellement le long de deux côtés d'un petit carré élémentaire, deux configurations (cf.

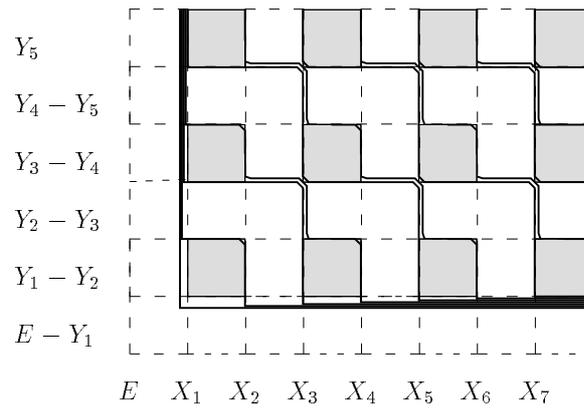


Figure 3.10 : Chaîne de  $XY$

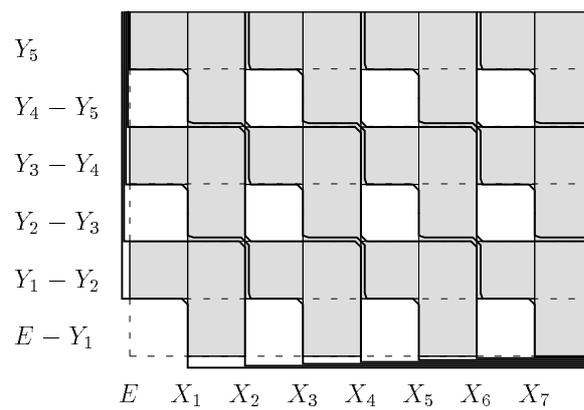


Figure 3.11 : Chaîne de  $X + Y$

figure 3.12) sont possibles. Ces deux configurations conduisent à deux chaînes égales à la même partie de  $E$ . Or pour les deux chaînes données précédemment pour l'intersection et pour l'union, ces configurations apparaissent de façon régulière, ce qui conduit à de nombreuses formules différentes.



Figure 3.12 : Configurations équivalentes

Ces deux diagrammes permettent de donner des formules explicites pour les chaînes de  $XY$  et de  $X + Y$ .

$$XY = Z_1 - Z_2 + Z_3 - \cdots \pm Z_{m+n-1}$$

où les  $Z_i$  sont définis par

$$\begin{aligned} Z_1 &= X_1 Y_1 \\ Z_2 &= X_1 Y_2 + X_2 Y_1 \\ Z_3 &= X_1 Y_3 + X_3 Y_1 \\ Z_4 &= X_1 Y_4 + X_2 Y_3 + X_3 Y_2 + X_4 Y_1 \\ &\vdots \\ Z_{m+n-2} &= X_{m-1} Y_n + X_m Y_{n-1} && \text{si } m+n \text{ pair} \\ Z_{m+n-2} &= X_{m-1} Y_n && \text{si } m \text{ pair et } n \text{ impair} \\ Z_{m+n-2} &= X_m Y_{n-1} && \text{si } m \text{ impair et } n \text{ pair} \\ Z_{m+n-1} &= X_m Y_n && \text{si } m \text{ impair ou } n \text{ impair} \\ Z_{m+n-1} &= \emptyset && \text{si } m \text{ pair et } n \text{ pair} \end{aligned}$$

ou de façon plus synthétique

$$\begin{aligned} Z_{2i} &= \sum_{k+l=2i+1} X_k Y_l \\ Z_{2i+1} &= \sum_{\substack{k+l=2i+2 \\ k \text{ impair}}} X_k Y_l \end{aligned} \tag{8}$$

$$X + Y = Z_1 - Z_2 + Z_3 - \cdots \pm Z_{m+n}$$

où les  $Z_i$  sont définis en posant  $X_0 = Y_0 = E$  par

$$\begin{aligned}
Z_1 &= Y_1 + X_1 &= X_0Y_1 + X_1Y_0 \\
Z_2 &= Y_2 + X_2 &= X_0Y_2 + X_2Y_0 \\
Z_3 &= Y_3 + X_1Y_2 + X_2Y_1 + X_3 &= X_0Y_3 + X_1Y_2 + X_2Y_1 + X_3Y_0 \\
Z_4 &= Y_4 + X_2Y_2 + X_4 &= X_0Y_4 + X_2Y_2 + X_4Y_0 \\
&\vdots \\
Z_{m+n-1} &= X_{m-1}Y_n + X_mY_{n-1} && \text{si } m+n \text{ pair} \\
Z_{m+n-1} &= X_{m-1}Y_n && \text{si } m \text{ impair et } n \text{ pair} \\
Z_{m+n-1} &= X_mY_{n-1} && \text{si } m \text{ pair et } n \text{ impair} \\
Z_{m+n} &= X_mY_n && \text{si } m \text{ pair ou } n \text{ pair} \\
Z_{m+n} &= \emptyset && \text{si } m \text{ impair et } n \text{ impair}
\end{aligned}$$

ou de façon plus synthétique

$$\begin{aligned}
Z_{2i} &= \sum_{\substack{k+l=2i \\ k \text{ pair}}} X_kY_l \\
Z_{2i+1} &= \sum_{k+l=2i+1} X_kY_l
\end{aligned} \tag{9}$$

□

**Proposition 16** *Si  $X \in C_m$  alors  $X^c \in C_{m+1}$ .*

**Démonstration :** Si la partie  $X$  s'écrit  $X = X_1 - X_2 + X_3 - \dots \pm X_m$ , le complémentaire  $X^c$  s'écrit  $X^c = E - X_1 + X_2 - X_3 + \dots \pm X_m$  et appartient à  $C_{m+1}$ . □

**Proposition 17** *Pour  $m \geq 0$ , tout  $X \in C_m$  et tout  $Y \in \mathcal{F}$ , nous avons*

- (1) *Pour  $m$  pair,  $X + Y \in C_{m+1}$  et  $X - Y \in C_m$*
- (2) *Pour  $m$  impair,  $X + Y \in C_m$  et  $X - Y \in C_{m+1}$*

**Démonstration :** Il suffit de remarquer que  $X - Y = XY^c$ . Nous appliquons d'abord la proposition 16 pour avoir  $Y^c \in C_2$  puis la proposition 15 pour conclure. □

**Proposition 18** *Toute somme de différences est égale à une chaîne de même longueur.*

**Démonstration :** Si la longueur  $m$  de la somme de différences est 1, la partie considérée est un élément de la semi-algèbre  $\mathcal{F}$  et le résultat est acquis. Si la longueur est égale à 2, il suffit d'écrire  $X = X_1 - X_2 = X_1 - X_1X_2$  ou  $X = X_1 - X_2 = X_1 = (X_1 + X_2) - X_2$  pour obtenir une chaîne de longueur 2 égale à  $X$ . Nous raisonnons ensuite par récurrence pour  $m \geq 3$ . Pour la somme de différences  $X = (X_1 - X_2) + (X_3 - X_4) + \dots$ , les parties  $X_1 - X_2$  et  $(X_3 - X_4) + \dots$  appartiennent respectivement à  $C_2$  et à  $C_{m-2}$ . Nous appliquons ensuite la proposition 15 pour conclure.  $\square$

Nous avons un résultat similaire pour les différences symétriques de longueur  $m$ .

**Proposition 19** *Toute différence symétrique  $X = X_1 \Delta \dots \Delta X_m$  d'éléments de  $\mathcal{F}$  est égale à une chaîne de même longueur.*

**Démonstration :**

$$X = \bigtriangleup_{k=1}^m Z_k \quad \text{avec} \quad Z_k = \sum_{i_1, \dots, i_k \text{ distincts}} X_{i_1} \dots X_{i_k}$$

En effet par définition de la différence symétrique,  $X$  est égal à l'ensemble des éléments de  $E$  qui appartiennent à un nombre impair de parties  $X_i$ . Si  $x \in E$  appartient à exactement  $k$  parties  $X_i$ , il appartient à  $Z_1, \dots, Z_k$  et à aucun des  $Z_i$  pour  $i > k$ . Le nombre de parties  $Z_i$  à auxquelles appartient un élément  $x$  de  $E$  est égal au nombre de parties  $X_i$  auxquelles il appartient. Nous remarquons également que les parties  $Z_i$  vérifient  $Z_1 \supset \dots \supset Z_m$  et constituent une chaîne égale à  $X$ .  $\square$

Réciproquement, nous avons pour toute chaîne

$$X_1 - X_2 + X_3 - \dots \pm X_m = X_1 \Delta X_2 \Delta \dots \Delta X_m$$

**Proposition 20** *Si  $X \in C_m$  et  $Y \in C_n$ , alors  $X \Delta Y \in C_{m+n}$ .*

**Démonstration :** Les parties  $X$  et  $Y$  s'écrivent respectivement  $X = X_1 \Delta \dots \Delta X_m$  et  $Y = Y_1 \Delta \dots \Delta Y_n$ . La différence symétrique  $X \Delta Y$  s'écrit  $X \Delta Y = X_1 \Delta \dots \Delta X_m \Delta Y_1 \Delta \dots \Delta Y_n$  et appartient à  $C_{m+n}$  d'après la proposition précédente.  $\square$

### 3.4 Algèbre de Boole

**Proposition 21** *L'ensemble  $C = \bigcup_{m \geq 0} C_m$  est l'algèbre de Boole engendrée par  $\mathcal{F}$ .*

**Démonstration :** C'est immédiat d'après les propositions 15 et 16.  $\square$

Soit  $\mathcal{F}$  une semi-algèbre de Boole. Nous notons

$$\begin{aligned} D_n(\mathcal{F}) &= \{X^c \mid X \in C_n(\mathcal{F})\} \\ \mathcal{F}^c &= \{X^c \mid X \in \mathcal{F}\} \end{aligned}$$

**Proposition 22** *On a les égalités suivantes*

$$\begin{aligned} C_n(\mathcal{F}^c) &= C_n(\mathcal{F}) \quad \text{pour } n \text{ pair} \\ C_n(\mathcal{F}^c) &= D_n(\mathcal{F}) \quad \text{pour } n \text{ impair} \end{aligned}$$

**Démonstration :** D'après la proposition 19, une partie  $X$  de  $E$  appartient à  $C_n(\mathcal{F})$  si et seulement si elle peut s'écrire  $X = X_1 \Delta \cdots \Delta X_n$  avec  $X_i \in \mathcal{F}$ . La formule  $X^c \Delta Y = (X \Delta Y)^c$  implique les deux formules

$$\begin{aligned} X_1^c \Delta \cdots \Delta X_n^c &= X_1 \Delta \cdots \Delta X_n \quad \text{pour } n \text{ pair} \\ X_1^c \Delta \cdots \Delta X_n^c &= (X_1 \Delta \cdots \Delta X_n)^c \quad \text{pour } n \text{ impair} \end{aligned}$$

Le résultat est alors immédiat.  $\square$



## Chapitre 4

# Hiérarchie de Wagner

### 4.1 Introduction

Nous présentons ici la théorie, initialement due à K. Wagner [Wag79, Kam85, Bar92], des classes de complexité de  $\omega$ -langages reconnaissables. Ces classes sont caractérisées par les longueurs des chaînes et des superchaînes. Ces deux notions peuvent, de façon indépendante, être définies pour un  $\omega$ -langage, un automate et un  $\omega$ -semigroupe. Nous montrons que ces trois définitions conduisent effectivement à la même notion. Nous étudions ensuite les liens entre ces classes de complexité et la structure topologiques des  $\omega$ -langages. Nous établissons en particulier la stabilité de ces classes de complexité par image inverse de fonction séquentielle. Ensuite nous caractérisons l'indice de Rabin d'un  $\omega$ -langage. Nous terminons par la construction d'un automate équivalent à un automate.

### 4.2 Chaînes et superchaînes

#### 4.2.1 Chaînes et superchaînes d'un $\omega$ -langage

Soit  $X \subset A^\omega$  un  $\omega$ -langage et  $m \geq 0$  un entier. Une  $X$ -chaîne de longueur  $m$  est un couple

$$C = (Y; Z)$$

formé d'un ensemble  $Y \subset A^*$  non vide et d'une suite strictement croissante

$$Z = (Z_0, Z_1, \dots, Z_m)$$

de parties de  $A^+$  telle que les conditions suivantes soient vérifiées.

- (i)  $YZ_i \subset Y$ ;

- (ii)  $Z_i^* = Z_i$ ;
- (iii)  $Z_{i-1}(Z_i - Z_{i-1}) \subset Z_i - Z_{i-1}$  et  
 $(Z_i - Z_{i-1})Z_{i-1} \subset Z_i - Z_{i-1}$  pour tout  $1 \leq i \leq m$ ;
- (iv) les ensembles  $W_0 = YZ_0^\omega$ ,  $W_1 = Y(Z_1 - Z_0)^\omega$ ,  $W_2 = Y(Z_2 - Z_1)^\omega$ , ... et  
 $W_m = Y(Z_m - Z_{m-1})^\omega$  sont alternativement inclus dans  $X$  ou disjoints de  $X$ .

La chaîne est dite *positive* quand  $W_0 \subset X$  et *negative* sinon. Ainsi, pour une chaîne positive de longueur 4, on a

$$\begin{aligned}
 YZ_0^\omega &\subset X \\
 Y(Z_1 - Z_0)^\omega &\cap X = \emptyset \\
 Y(Z_2 - Z_1)^\omega &\subset X \\
 Y(Z_3 - Z_2)^\omega &\cap X = \emptyset \\
 Y(Z_4 - Z_3)^\omega &\subset X
 \end{aligned}$$

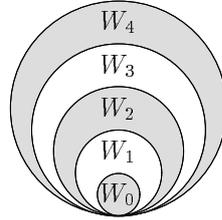


Figure 4.1 : Chaîne de longueur 4

On note  $m^+(X)$  (resp.  $m^-(X)$ ) la longueur maximale des chaînes positives (resp. négatives) de  $X$  et  $m(X) = \max(m^+(X), m^-(X))$ . On convient de poser  $m^+(X) = -1$  (resp.  $m^-(X) = -1$ ) si l'ensemble des chaînes positives (resp. négatives) est vide. Nous verrons que  $m(X)$  est toujours fini pour  $X$  reconnaissable.

**Exemple :** Soit  $X_1 = (b^*a)^\omega$  l'ensemble des mots ayant un nombre infini d'occurrences de  $a$  sur l'alphabet  $A = a + b$ . Une chaîne négative de longueur 1 est  $(A^*; b^*, A^*)$ . En effet, on a  $A^*b^\omega \cap X_1 = \emptyset$  et  $A^*(A^* - b^*)^\omega = X_1$ . On a en fait  $m^+(X_1) = 0$  et  $m(X_1) = m^-(X_1) = 1$  comme on le vérifiera plus loin.

Une  $X$ -*superchaîne* de longueur  $n$  est une suite

$$C_0, C_1, \dots, C_n$$

de chaînes  $C_i = (Y_i; Z_{i0}, Z_{i1}, \dots, Z_{im})$  toutes de longueur  $m(X)$  telle que :

- (i) chaque  $C_i$  est accessible à partir de  $C_{i-1}$ , i.e., pour  $1 \leq i \leq n$ , il existe des ensembles  $U_i \subset A^*$  non vides tels que  $Y_{i-1}U_i \subset Y_i$ ;
- (ii) les chaînes  $C_i$  sont alternativement positives et négatives.

On dit que la superchaîne est *positive* si  $C_0$  est positive et *négative* sinon. On note  $n^+(X)$  (resp.  $n^-(X)$ ) la longueur maximale des superchaînes positives (resp. négatives) et  $n(X) = \max(n^+(X), n^-(X))$ . On convient que  $n^+(X) = -1$  (resp.  $n^-(X) = -1$ ) si l'ensemble des superchaînes positives (resp. négatives) est vide. Nous verrons aussi que  $n(X)$  est fini pour  $X$  reconnaissable.

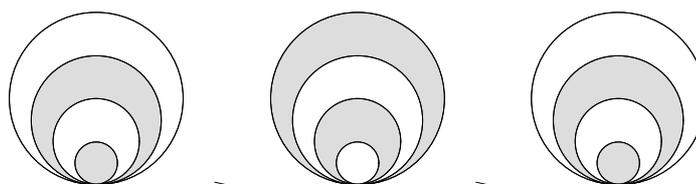


Figure 4.2 : Superchaîne de longueur 2 avec des chaînes de longueur 3

**Exemple :** Le langage  $X_2 = (b^*a)^\omega + A^*c(a + b)^\omega$  sur l'alphabet  $A = a + b + c$  est l'ensemble des mots ayant ou bien

- (1) aucune occurrence de  $c$  et un nombre infini d'occurrences de  $a$ ;
- (2) un nombre fini non nul d'occurrences de  $c$ .

Le langage  $X_2$  admet deux chaînes de longueur 1, la chaîne négative  $C_0 = ((a + b)^*; b^*, (a + b)^*)$  et la chaîne positive  $C_1 = (A^*cA^*; (a + b)^*, A^*)$ . En prenant  $U_1 = A^*cA^*$ , on vérifie que  $Y_0U_1 = A^*cA^* = Y_1$ . Le langage  $X_2$  admet donc la superchaîne positive ( $C_1$ ) de longueur 0 et la superchaîne négative ( $C_0, C_1$ ) de longueur 1. On vérifiera plus loin que  $n^+(X_2) = 0$  et  $n(X_2) = n^-(X_2) = 1$ .

### 4.2.2 Chaînes et superchaînes d'un automate

Soit  $\mathcal{A} = (Q, A, \cdot, q_0, \mathcal{T})$  un automate de Muller. Une  $\mathcal{A}$ -chaîne de longueur  $m$  de l'automate  $\mathcal{A}$  est une suite croissante

$$R_0 \subset R_1 \subset \dots \subset R_m$$

de parties répétables de  $Q$  telle que les  $R_i$  pour  $0 \leq i \leq m$  sont alternativement dans  $\mathcal{T}$  et hors de  $\mathcal{T}$ . On dit que la chaîne est *positive* ou *négative* selon que  $R_0$  appartienne ou n'appartienne pas à la table  $\mathcal{T}$ . On note  $m^+(\mathcal{A})$  (resp.  $m^-(\mathcal{A})$ ) la longueur maximale des chaînes positives (resp. négatives) de  $\mathcal{A}$  et  $m(\mathcal{A}) = \max(m^+(\mathcal{A}), m^-(\mathcal{A}))$ .

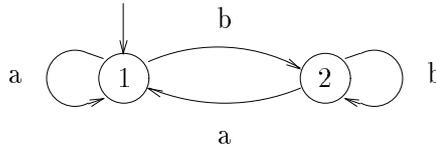


Figure 4.3 : Automate  $\mathcal{A}_1$  :  $\mathcal{T} = \{\{1\}, \{1, 2\}\}$

**Exemple :** L'automate  $\mathcal{A}_1$  de la figure 4.3 avec la table  $\mathcal{T} = \{\{1\}, \{1, 2\}\}$  reconnaît le  $\omega$ -langage  $X_1$  des mots ayant un nombre infini d'occurrences de  $a$ . On a  $m^+(\mathcal{A}_1) = 0$  et  $m(\mathcal{A}_1) = m^-(\mathcal{A}) = 1$ . En effet, la seule chaîne de longueur 1 est  $\{2\} \subset \{1, 2\}$  et cette chaîne est négative.

Une  $\mathcal{A}$ -superchaîne de longueur  $n$  de l'automate est une suite

$$(U_0, U_1, \dots, U_n)$$

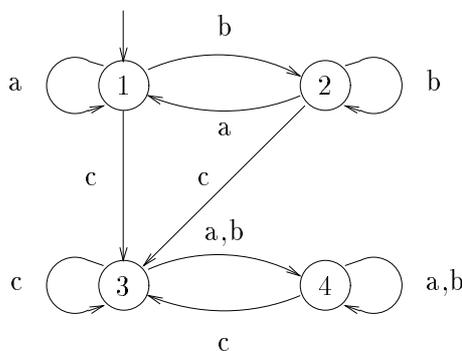
de chaînes de longueur  $m(\mathcal{A})$  telle que :

- (i) chaque  $U_i$  est accessible à partir de  $U_{i-1}$  pour  $1 \leq i \leq n$ , i.e., il existe un chemin d'un état de  $U_{i-1}$  à un état de  $U_i$ ;
- (ii) les  $\mathcal{A}$ -chaînes  $U_i$  sont alternativement positives et négatives.

On dit que la superchaîne est *positive* si  $U_0$  est une chaîne positive et *négative* sinon. On note  $n^+(X)$  (resp.  $n^-(X)$ ) la longueur maximale des superchaînes positives (resp. négatives) et  $n(X) = \max(n^+(X), n^-(X))$ . On convient que  $n^+(X) = -1$  (resp.  $n^-(X) = -1$ ) si l'ensemble des superchaînes positives (resp. négatives) est vide.

**Exemple :** L'automate  $\mathcal{A}_1$  de la figure 4.3 possède une seule chaîne de longueur maximale et cette chaîne est négative. L'automate admet donc une seule superchaîne négative de longueur 0 et aucune superchaîne positive :  $n^+(\mathcal{A}_1) = -1$  et  $n(\mathcal{A}_1) = n^-(\mathcal{A}_1) = 0$ .

**Exemple :** L'automate  $\mathcal{A}_2$  de la figure 4.4 avec la table  $\mathcal{T} = \{\{1\}, \{1, 2\}, \{4\}\}$  reconnaît le langage  $X_2$ . L'automate possède deux chaînes de longueur 1 et


 Figure 4.4 : Automate  $\mathcal{A}_2 : \mathcal{T} = \{\{1\}, \{1, 2\}, \{4\}\}$ 

aucune chaîne de longueur supérieure. La chaîne  $U_0 = (\{2\}, \{1, 2\})$  est négative et la chaîne  $U_1 = (\{4\}, \{3, 4\})$  est positive :  $m(\mathcal{A}) = m^+(\mathcal{A}) = m^-(\mathcal{A}) = 1$ . L'automate  $\mathcal{A}_2$  admet la superchaîne positive ( $U_1$ ) et la superchaîne négative ( $U_0, U_1$ ) :  $n^+(\mathcal{A}_2) = 0$  et  $n(\mathcal{A}_2) = n^-(\mathcal{A}_2) = 1$ .

### 4.2.3 Chaînes et superchaînes d'un $\omega$ -semigroupe

Soit  $S = (S_f, S_i)$  un  $\omega$ -semigroupe et  $P$  un sous-ensemble de la partie infinitaire  $S_i$  de  $S$ . Une  $(S, P)$ -chaîne de longueur  $m$  est une suite

$$s; e_0, e_1, \dots, e_m$$

de  $m + 2$  éléments de  $S$  telle que :

- (i) pour  $0 \leq i \leq m$ , le couple  $(s, e_i)$  est lié, i.e.,  $se_i = s$  et  $e_i^2 = e_i$ ;
- (ii) la suite  $e_0, \dots, e_m$  est décroissante pour l'ordre  $\mathcal{H} : e_0 \geq_{\mathcal{H}} e_1 \geq_{\mathcal{H}} \dots \geq_{\mathcal{H}} e_m$  ce qui est équivalent à  $e_i e_{i-k} = e_{i-k} e_i = e_i$  pour  $0 \leq i - k \leq i \leq m$ ;
- (iii) les éléments  $se_i^\omega$  pour  $0 \leq i \leq m$  sont alternativement dans la partie  $P$  et hors de la partie  $P$ .

On a encore les notions de chaîne positive ou négative selon que l'élément  $se_0^\omega$  appartienne ou n'appartienne pas à  $P$ . On note également  $m^+(S, P)$  (resp.  $m^-(S, P)$ ) la longueur maximale des chaînes positives (resp. négatives) dans le  $\omega$ -semigroupe  $S = (S_f, S_i)$  relativement à la partie  $P \subset S_i$  et  $m(S, P) = \max(m^+(S, P), m^-(S, P))$ .

$$\begin{array}{c} \boxed{*b = 1} \\ \boxed{*a = 0} \end{array}$$

Figure 4.5 : Structure en  $\mathcal{D}$ -classes de  $S = \{b = 1, a = 0\}$ .

**Exemple :** Le morphisme  $\varphi : A^\infty \rightarrow S_1$  dans le  $\omega$ -semigroupe  $S_1 = \widehat{S}$  où  $S$  est le semigroupe  $S = \{b = 1, a = 0\}$  reconnaît le  $\omega$ -langage  $X_1 = (b^*a)^\omega$ . La structure en  $\mathcal{D}$ -classes de  $S$  est donnée par la figure 4.5. L'image de  $X_1$  dans  $S_1$  est la partie  $P = \{a^\omega\}$ . La seule  $(S_1, P)$ -chaîne de longueur 1 est la chaîne négative  $(a; b, a)$ .

Une  $(S, P)$ -superchaîne de longueur  $n$  est une suite

$$u_0, u_1, \dots, u_n$$

de chaînes de longueur  $m = m(S, P)$  avec  $u_i = (s_i, e_{i0}, e_{i1}, \dots, e_{im})$  telle que :

- (i) la suite  $s_i$  est décroissante pour l'ordre  $\mathcal{R}$ , i.e.,  $s_0 \geq_{\mathcal{R}} s_2 \geq_{\mathcal{R}} \dots \geq_{\mathcal{R}} s_n$ ;
- (ii) les chaînes  $u_i$  sont alternativement positives et négatives.

On a encore les notions de superchaînes positives et négatives, et les entiers associés  $n^+(S, P)$ ,  $n^-(S, P)$  et  $n(S, P) = \max(n^+(S, P), n^-(S, P))$ .

$$\begin{array}{c} \boxed{*b = 1} \\ \boxed{*a} \\ \boxed{*c = 0} \end{array}$$

Figure 4.6 : Structure en  $\mathcal{D}$ -classes de  $\{b = 1, a = a^2, c = 0\}$ .

**Exemple :** Le  $\omega$ -langage  $X_2 = (b^*a)^\omega + A^*c(a+b)^\omega$  est reconnu par le morphisme  $\varphi : A^\infty \rightarrow \widehat{S}$  dans le  $\omega$ -semigroupe  $S_2 = \widehat{S}$  où  $S$  est le semigroupe idempotent  $S = \{b = 1, a = a^2, c = 0\}$ . La structure en  $\mathcal{D}$ -classes de  $S$  est donnée par la figure 4.6. L'image du  $\omega$ -langage  $X_2$  dans  $S_2$  est la partie  $P = \{cb^\omega, a^\omega, ca^\omega\}$ . Relativement à cette partie  $P$ , le  $\omega$ -semigroupe  $S_2$  admet deux chaînes de longueur 1 et aucune chaîne de longueur supérieure. La chaîne  $u_0 = (a; b, a)$  est négative et la chaîne  $u_1 = (c; a, c)$  est positive. On a donc  $m(S_2, P) = m^+(S_2, P) = m^-(S_2, P) = 1$ . Le

$\omega$ -semigroupe  $S_2$  admet la superchaîne positive ( $u_1$ ) et la superchaîne négative ( $u_0, u_1$ ) puisque  $a \geq_{\mathcal{R}} c$ . On a alors  $n^+(S_2, P) = 0$  et  $n(S_2, P) = n^-(S_2, P) = 1$ .

#### 4.2.4 Équivalences des définitions

La proposition suivante montre que les entiers  $m^+(X)$ ,  $m^-(X)$ ,  $m(X)$ ,  $n^+(X)$ ,  $n^-(X)$  et  $n(X)$  peuvent être calculés pour un  $\omega$ -langage reconnaissable  $X$  sur tout automate ou tout  $\omega$ -semigroupe reconnaissant ce langage. En particulier les entiers  $m^+(\mathcal{A})$ ,  $m^-(\mathcal{A})$ ,  $m(\mathcal{A})$ ,  $n^+(\mathcal{A})$ ,  $n^-(\mathcal{A})$ ,  $n(\mathcal{A})$  et les entiers  $m^+(S, P)$ ,  $m^-(S, P)$ ,  $m(S, P)$ ,  $n^+(S, P)$ ,  $n^-(S, P)$  et  $n(S, P)$  ne dépendent ni de l'automate ni du  $\omega$ -semigroupe considéré mais sont uniquement liés au  $\omega$ -langage reconnu.

**Proposition 23** *Soit  $X \subset A^\omega$  un  $\omega$ -langage reconnaissable, soit  $\mathcal{A} = (Q, i, T)$  un automate de Muller complet reconnaissant  $X$  et  $\varphi : A^\infty \rightarrow S$  un morphisme de  $A^\infty$  sur un  $\omega$ -semigroupe fini  $S = (S_f, S_i)$  reconnaissant le langage  $X$ . On pose  $P = X\varphi$ . On a les égalités*

$$\begin{aligned} m^+(X) &= m^+(\mathcal{A}) = m^+(S, P) \\ m^-(X) &= m^-(\mathcal{A}) = m^-(S, P) \\ n^+(X) &= n^+(\mathcal{A}) = n^+(S, P) \\ n^-(X) &= n^-(\mathcal{A}) = n^-(S, P) \end{aligned}$$

**Exemple :** Le langage  $X_1 = (b^*a)^\omega$  est reconnu par l'automate  $\mathcal{A}_1$  de la figure 4.3 et le  $\omega$ -semigroupe  $S_1$  :

$$\begin{aligned} m^+(X_1) &= m^+(\mathcal{A}_1) = m^+(S_1, P) = 0 \\ m^-(X_1) &= m^-(\mathcal{A}_1) = m^-(S_1, P) = 1 \\ n^+(X_1) &= n^+(\mathcal{A}_1) = n^+(S_1, P) = -1 \\ n^-(X_1) &= n^-(\mathcal{A}_1) = n^-(S_1, P) = 0 \end{aligned}$$

**Exemple :** Le langage  $X_2 = (b^*a)^\omega + A^*c(a+b)^\omega$  sur l'alphabet  $A = a + b + c$  est reconnu par l'automate  $\mathcal{A}_2$  de la figure 4.4 et le  $\omega$ -semigroupe  $S_2$  :

$$\begin{aligned} m^+(X_2) &= m^+(\mathcal{A}_2) = m^+(S_2, P) = 1 \\ m^-(X_2) &= m^-(\mathcal{A}_2) = m^-(S_2, P) = 1 \\ n^+(X_2) &= n^+(\mathcal{A}_2) = n^+(S_2, P) = 0 \\ n^-(X_2) &= n^-(\mathcal{A}_2) = n^-(S_2, P) = 1 \end{aligned}$$

**Démonstration :** Nous allons associer à chaque  $X$ -chaîne une  $\mathcal{A}$ -chaîne de même longueur et de même nature. Ensuite, nous allons associer à chaque  $\mathcal{A}$ -chaîne une  $(S, P)$ -chaîne de même longueur et de même nature. Enfin, nous

allons associer à chaque  $(S, P)$ -chaîne une  $X$ -chaîne de même longueur et de même nature ce qui montrera le résultat pour les chaînes. Nous allons faire de même pour les superchaînes.

### Équivalences des définitions pour les chaînes

**Lemme 11** *Soient  $\mathcal{A} = (Q, A, \cdot, q_0, \mathcal{T})$  un automate de Muller complet reconnaissant un  $\omega$ -langage  $X$ . Soient  $(Y; Z_0, Z_1, \dots, Z_m)$  une  $X$ -chaîne,  $q$  un état de l'automate et  $R \subset A^*$  un langage tel que  $q.R \subset i.Y$ . Il existe alors une  $\mathcal{A}$ -chaîne accessible de  $q$  de même longueur et de même nature que la  $X$ -chaîne  $(Y; Z_0, Z_1, \dots, Z_m)$ .*

**Démonstration :** On pose  $Q' = q.RZ_m$ , l'ensemble des états accessibles de l'état  $q$  par un chemin étiqueté d'un mot de  $RZ_m$ . L'inclusion  $Z_m Z_i \subset Z_m$  impose que  $Q'$  est stable par action de  $Z_i : Q'.Z_i = q.RZ_m Z_i \subset q.RZ_m = Q'$ . On définit le préordre  $\prec_i$  et la relation d'équivalence associée  $\equiv_i$  sur  $Q'$  par

$$q' \prec_i q \quad \Leftrightarrow \quad \exists z_i \in Z_i \quad q \xrightarrow{z_i} q'$$

Ceci définit bien un préordre puisque  $Z_i^* \subset Z_i$ . Chaque préordre  $\prec_i$  induit un ordre sur les classes d'équivalence de  $\equiv_i$ . Puisque les ensembles  $Z_i$  sont ordonnés par inclusion, il en est de même pour les préordres  $\prec_i$  et les relations d'équivalence  $\equiv_i$ .

$$\begin{aligned} q \prec_i q' &\Rightarrow q \prec_{i+k} q' \\ q \equiv_i q' &\Rightarrow q \equiv_{i+k} q' \end{aligned}$$

On choisit  $Q_m$  une classe d'équivalence minimale pour  $\prec_m$ . Puisque  $\prec_{m-1} \subset \prec_m$ , l'ensemble  $Q_m$  est saturé par  $\equiv_{m-1}$ . On choisit alors  $Q_{m-1}$  une classe d'équivalence minimale pour  $\prec_{m-1}$  incluse dans  $Q_m$ . On itère le procédé pour choisir  $Q_0 \subset Q_1 \subset \dots \subset Q_m$  des classes d'équivalence de  $\equiv_i$  minimales pour  $\prec_i$ . On choisit un état  $q'$  un état de  $Q_0$ . Puisque  $q'$  appartient à  $Q' = q.RZ_m$ , il appartient également à  $i.YZ_m \subset i.Y$ . On choisit alors un mot  $y \in Y$  tel que  $i.y = q'$ . On choisit ensuite des mots  $z_0, \dots, z_m$  appartenant respectivement à  $Z_0, Z_1 - Z_0, \dots, Z_m - Z_{m-1}$  et l'on pose  $q_i = q'.z_i$ . Puisque les ensembles  $Q_i$  sont des classes minimales, on a nécessairement  $q' \equiv_i q_i$ , ce qui assure l'existence de mots  $z'_i \in Z_i$  tels que  $q_i.z'_i = q'$ . On définit alors les mots  $w_i$  par

$$\begin{aligned} w_0 &= z_0 z'_0 \\ w_1 &= w_0 z_1 z'_1 \\ &\vdots \\ w_m &= w_{m-1} z_m z'_m \end{aligned}$$

Les mots  $w_i$  vérifient les propriétés suivantes

- (i)  $w_i \in Z_i - Z_{i-1}$ ;
- (ii)  $q'.w_i = q'$ ;
- (iii)  $yw_i^\omega \in Y(Z_i - Z_{i-1})^\omega$ .

Les chemins  $q' \xrightarrow{w_i} q'$  définissent donc des ensembles répétables  $R_0, \dots, R_m$  qui constituent une  $\mathcal{A}$ -chaîne puisque  $w_{i+1} = w_i z_{i+1} z'_{i+1}$  et  $yw_i^\omega \in Y(Z_i - Z_{i-1})^\omega$ .  $\square$

Pour montrer qu'à toute  $X$ -chaîne, on peut associer une  $\mathcal{A}$ -chaîne de même longueur et de même nature, il suffit d'appliquer le lemme avec  $q = i$  et  $R = Y$ .

Soit  $R_0 \subset \dots \subset R_m$  une  $\mathcal{A}$ -chaîne. On choisit un état  $q \in R_0$  et un mot  $y$  tel que  $i.y = q$ . On choisit également des mots  $w_0, \dots, w_m \in A^*$  tels que les chemins  $q \xrightarrow{w_i} q$  définissent les parties répétables  $R_0, \dots, R_m$ . On définit alors l'élément  $s$  et les idempotents  $e_i$  et les mots  $w'_i$  par

$$\begin{array}{ll}
 e_0 = (w_0\varphi)^\pi & \text{et } w'_0 = w_0^\pi \\
 e_1 = (e_0(w_1\varphi)e_0)^\pi & \text{et } w'_1 = (w'_0 w_1 w'_0)^\pi \\
 e_2 = (e_1(w_2\varphi)e_1)^\pi & \text{et } w'_2 = (w'_1 w_2 w'_1)^\pi \\
 \vdots & \vdots \\
 e_m = (e_{m-1}(w_m\varphi)e_{m-1})^\pi & \text{et } w'_m = (w'_{m-1} w_m w'_{m-1})^\pi \\
 s = (y\varphi)e_m & \text{et } y' = yw'_m
 \end{array}$$

On vérifie aisément que

- (1)  $e_i e_{i+1} = e_{i+1} e_i = e_{i+1}$  et donc  $e_0 \geq_{\mathcal{H}} e_1 \geq_{\mathcal{H}} \dots \geq_{\mathcal{H}} e_m$ ;
- (2)  $s e_i = s e_m e_i = s$  et donc les couples  $(s, e_i)$  sont liés;
- (3) Les mots  $y'$  et  $w'_i$  vérifient  $y'\varphi = s$ ,  $w'_i\varphi = e_i$  pour  $0 \leq i \leq m$ . Le mot infini  $u_i = y'w'_i^\omega$  qui induit un chemin infini  $c_i$  tel que  $\text{Inr}(c_i) = R_i$  vérifie  $u_i\varphi = s e_i^\omega$ .

La suite  $(s; e_0, e_1, \dots, e_m)$  constitue donc une  $(S, P)$ -chaîne de même longueur et de nature même que  $\mathcal{A}$ -chaîne  $R_0 \subset \dots \subset R_m$ .

Soit  $(s; e_0, e_1, \dots, e_m)$  une  $(S, P)$ -chaîne. On pose

$$\begin{array}{l}
 Y = s\varphi^{-1} \\
 Z_i = \{e_0, \dots, e_i\}\varphi^{-1}
 \end{array}$$

On vérifie que

- (i)  $YZ_i \subset Y$  puisque  $s e_i = s$ ;

- (ii)  $Z_i^* = Z_i$  puisque  $\{e_0, \dots, e_i\}$  est un sous-semigroupe de  $S_f$ ;
- (iii)  $Z_{i-1}(Z_i - Z_{i-1}) \subset Z_i - Z_{i-1}$  et  $(Z_i - Z_{i-1})Z_{i-1} \subset Z_i - Z_{i-1}$  puisque  $(Z_i - Z_{i-1}) = e_i\varphi^{-1}$ ;
- (iv)  $W_0 = YZ_0^\omega = se_0^\omega\varphi^{-1}$ ,  $W_1 = Y(Z_1 - Z_0)^\omega = se_1^\omega\varphi^{-1}$ , ... et  $W_m = Y(Z_m - Z_{m-1})^\omega = se_m^\omega\varphi^{-1}$ .

La suite  $(Y; Z_0, \dots, Z_m)$  constitue donc une  $X$ -chaîne de même longueur et de même nature que la  $(S, P)$ -chaîne  $(s; e_0, e_1, \dots, e_m)$ . Ceci achève de montrer que les différentes définitions des chaînes coïncident pour les  $\omega$ -langages reconnaissables.

### Équivalences des définitions pour les superchaînes

Nous montrons maintenant que les définitions des superchaînes coïncident pour les  $\omega$ -langages reconnaissables.

Soient  $(C_0, C_1, \dots, C_n)$  une  $X$ -superchaîne et  $U_i \subset A^*$  des ensembles tels que  $Y_{i-1}U_i \subset Y_i$ . On applique le lemme 11 à la  $X$ -chaîne  $C_0$  avec  $q = i$  et  $R = Y$  pour trouver une  $\mathcal{A}$ -chaîne  $U_0 = (R_{0,0}, \dots, R_{0,m})$ . On choisit un état  $q_0 \in R_{0,0}$ . On applique à nouveau le lemme 11 à la  $X$ -chaîne  $C_1$  avec  $q = q_0$  et  $R = U_1$ . On a bien  $q_0.U_1 \subset i.Y_0U_1 \subset i.Y_1$ . On trouve une  $\mathcal{A}$ -chaîne  $U_1 = (R_{1,0}, \dots, R_{1,m})$  accessible de la chaîne de l'état  $q_0$  et donc de la  $\mathcal{A}$ -chaîne  $U_0$ . On choisit un état  $q_1 \in R_{1,0}$  et l'on itère le procédé pour trouver des  $\mathcal{A}$ -chaînes  $U_0, \dots, U_n$  et donc une  $\mathcal{A}$ -superchaîne puisque la chaîne  $U_i$  est accessible de  $U_{i-1}$ . La  $\mathcal{A}$ -superchaîne trouvée est de même nature que la  $X$ -superchaîne de départ puisque chaque  $\mathcal{A}$ -chaîne  $U_i$  est de même nature que la  $X$ -chaîne  $C_i$ .

Soit  $(U_0, \dots, U_n)$  une  $\mathcal{A}$ -superchaîne où  $U_i = (R_{i,0}, \dots, R_{i,m})$  est une  $\mathcal{A}$ -chaîne. On choisit des états  $q_i \in R_{i,0}$  et des mots  $y, v_1, \dots, v_m$  tels que  $i.y = q_0$  et  $q_{i-1}.v_i = q_i$  pour  $1 \leq i \leq m$ . On choisit également des mots  $w_{i,j}$  tels que les chemins  $q_i \xrightarrow{w_{i,j}} q_i$  définissent les parties répétées  $R_{i,j}$ . On définit les éléments  $s_i$ , les idempotents  $e_{i,j}$  et les mots  $w'_{i,j}$  par

$$\begin{aligned}
e_{i,0} &= (w_{i,0}\varphi)^\pi & \text{et} & \quad w'_{i,0} = (w_{i,0})^\pi \\
e_{i,j} &= (e_{i,j-1}(w_{i,j}\varphi)^\pi e_{i,j-1})^\pi & \text{et} & \quad w'_{i,j} = (w'_{i,j-1}w_{i,j}w'_{i,j-1})^\pi \\
s_0 &= (y\varphi)e_{0,m} & \text{et} & \quad v'_0 = yw'_{0,m} \\
s_i &= s_{i-1}(v_i\varphi)e_{i,m} & \text{et} & \quad v'_i = v'_{i-1}w'_{i,m}
\end{aligned}$$

pour obtenir une  $(S, P)$ -superchaîne  $(u_0, \dots, u_n)$  en posant  $u_i = (s_i; e_{i,0}, \dots, e_{i,m})$ . La  $(S, P)$ -superchaîne ainsi obtenue est de même nature que la  $\mathcal{A}$ -superchaîne

$(U_0, \dots, U_n)$ . le mot infini  $u_{i,j} = v'_0 \dots v'_i (w'_{i,j})^\omega$  vérifie  $u_{i,j} \varphi = s_i e_{i,j}^\omega$  et induit un chemin infini  $c_{i,j}$  tel que  $\text{Inr}(c_{i,j}) = R_{i,j}$ .

Soit  $(u_0, \dots, u_n)$  une  $(S, P)$ -superchaîne et  $u_i \in S_f$  des éléments de  $S_f$  tels que  $s_{i-1} u_i = s_i$ . On pose

$$\begin{aligned} Y_i &= s_i \varphi^{-1} \\ U_i &= u_i \varphi^{-1} \\ Z_{i,j} &= \{e_{i,0}, \dots, e_{i,j}\} \varphi^{-1} \end{aligned}$$

On vérifie comme précédemment que les  $C_i = (Y_i; Z_{i,0}, \dots, Z_{i,m})$  constituent des  $X$ -chaînes de même nature que les  $(S, P)$ -chaînes  $u_i$ . De plus, l'égalité  $s_{i-1} u_i = s_i$  implique l'inclusion  $Y_{i-1} U_i \subset Y_i$ . On a donc obtenu une  $X$ -superchaîne de même nature que la  $(S, P)$ -superchaîne  $(u_0, \dots, u_n)$ . Ceci achève la démonstration de l'équivalence des différentes définitions des superchaînes pour les  $\omega$ -langages reconnaissables.  $\square$

### 4.3 Propriétés

Nous donnons maintenant quelques propriétés élémentaires des entiers  $m^+(X)$ ,  $m^-(X)$ ,  $m(X)$ ,  $n^+(X)$ ,  $n^-(X)$  et  $n(X)$ .

**Proposition 24** *Pour tout  $\omega$ -langage  $X$  reconnaissable*

- (i)  $0 \leq m(X) < \infty$  et  $0 \leq n(X) < \infty$ ;
- (ii)  $|m^+(X) - m^-(X)| \leq 1$  et  $|n^+(X) - n^-(X)| \leq 1$ ;
- (iii)  $m^+(X) = m^-(X)$  dès que  $n(X) \geq 1$ ;
- (iv)  $m^+(X) = m^-(A^\omega - X)$  et  $n^+(X) = n^-(A^\omega - X)$ .

**Démonstration :** Soit  $X$  un  $\omega$ -langage reconnaissable. Si  $X$  est vide, il admet la chaîne négative  $(A^*; A^*)$  de longueur 0. Ensuite tout  $\omega$ -langage reconnaissable non vide contient un ensemble de la forme  $YZ^\omega$  et admet donc la chaîne positive  $(YZ^*; Z^*)$ . Ceci montre que pour tout  $\omega$ -langage  $X$  reconnaissable, on a  $m(X) \geq 0$ . On peut aussi remarquer que tout automate de Muller complet admet au moins un ensemble répétable  $R$ . L'automate admet la chaîne  $(R)$ . Cette chaîne est positive si cet ensemble répétable est acceptant et négative sinon. En particulier si le  $\omega$ -langage reconnu par l'automate n'est pas vide, l'automate admet au moins un ensemble répétable acceptant et donc une chaîne positive. De même, tout  $\omega$ -semigroupe  $S = (S_f, S_i)$  admet la chaîne  $(s; s^\pi)$  pour  $s \in S_f$  quelconque.

Cette chaîne est positive si  $s^\omega \in P$  et négative sinon. En particulier si le langage reconnu par le  $\omega$ -semigroupe n'est pas vide, la partie  $P$  associée n'est pas vide et le  $\omega$ -semigroupe admet au moins une chaîne positive.

Pour une  $\mathcal{A}$ -chaîne  $R_0 \subset \dots \subset R_m$  d'un automate  $\mathcal{A} = (Q, i, \mathcal{T})$ , l'équivalence  $R_{i-1} \in \mathcal{T} \Leftrightarrow R_i \notin \mathcal{T}$  pour  $1 \leq i \leq m$  implique que les inclusions  $R_{i-1} \subset R_i$  sont nécessairement strictes. La longueur d'une chaîne est alors bornée par le nombre d'états de l'automate. De même, pour une  $(S, P)$ -chaîne  $(s; e_1, \dots, e_m)$  d'un  $\omega$ -semigroupe, l'équivalence  $se_{i-1}^\omega \in P \Leftrightarrow se_i^\omega \notin P$  impose que  $e_{i-1} \neq e_i$ . La relation  $e_{i-1} \geq_{\mathcal{H}} e_i$  implique  $e_{i-1} >_{\mathcal{D}} e_i$ . La longueur d'une  $(S, P)$ -chaîne est alors bornée par le nombre de  $\mathcal{D}$ -classes de  $S_f$ . Ceci montre que pour  $\omega$ -langage  $X$  reconnaissable, on a  $m(X) < \infty$ .

Puisque tout  $\omega$ -langage  $X$  reconnaissable admet au moins une  $X$ -chaîne, il admet une chaîne de longueur maximale puisque  $m(X) < \infty$ . Il admet donc aussi une  $X$ -superchaîne de longueur 0, ce qui prouve que  $n(X) \geq 0$ .

Afin de montrer que la longueur maximale des superchaînes d'un automate ou d'un  $\omega$ -semigroupe est bornée, nous démontrons auparavant deux lemmes qui jouent des rôles analogues, l'un pour les chaînes d'un automate et l'autre pour les chaînes d'un  $\omega$ -semigroupe.

**Lemme 12** *Soient  $U_1 = (R_{1,0}, \dots, R_{1,m})$  et  $U_2 = (R_{2,0}, \dots, R_{2,m})$  deux  $\mathcal{A}$ -chaînes de longueur maximale d'un automate  $\mathcal{A}$ . Si  $U_1$  est accessible de  $U_2$  et  $U_2$  est accessible de  $U_1$ , alors les chaînes  $U_1$  et  $U_2$  sont de même nature.*

**Démonstration :** De l'accessibilité de  $U_1$  à  $U_2$  et de  $U_2$  à  $U_1$ , on déduit l'existence de chemins  $q_1 \xrightarrow{u_1} q_2$  et  $q_2 \xrightarrow{u_2} q_1$  où  $q_1$  et  $q_2$  sont des états respectivement de  $R_{1,m}$  et  $R_{2,m}$ . Il existe également des chemins  $q_1 \xrightarrow{w_1} q_1$  et  $q_2 \xrightarrow{w_2} q_2$  qui définissent respectivement les ensembles répétables  $R_{1,m}$  et  $R_{2,m}$ . Le chemin  $q_1 \xrightarrow{w_1 u_1 w_2 v_2} q_1$  définit un ensemble répétable  $R$  contenant  $R_{1,m}$  et  $R_{2,m}$ . Les ensembles répétables  $R$  et  $R_{1,m}$  sont nécessairement de même nature (acceptants ou non acceptants). Sinon  $(R_{1,0}, \dots, R_{1,m}, R)$  est une chaîne de longueur  $m + 1$ , ce qui contredit la maximalité des chaînes  $U_1$  et  $U_2$ . De façon symétrique,  $R$  et  $R_{2,m}$  sont de même nature. Les ensembles  $R_{1,m}$  et  $R_{2,m}$  étant de même nature, les chaînes  $U_1$  et  $U_2$  le sont également.  $\square$

Une conséquence de ce lemme est que dans une  $\mathcal{A}$ -superchaîne  $(U_0, \dots, U_n)$  d'un automate  $\mathcal{A}$ , la chaîne  $U_i$  n'est pas accessible de la chaîne  $U_j$  pour  $i < j$ . Sinon la chaîne  $U_{j-1}$  est accessible de  $U_j$  puisque  $U_{j-1}$  est accessible de  $U_i$ . Ceci conduit à une contradiction puisque les chaînes  $U_{j-1}$  et  $U_j$  ne sont pas par définition de même nature. En particulier ceci montre que dans une superchaîne  $(U_0, \dots, U_n)$ , les ensembles répétables  $R_{0,m}, \dots, R_{n,m}$  sont deux à deux

disjoints. La longueur d'une superchaîne est alors bornée par le nombre d'états de l'automate.

**Lemme 13** *Soient  $u_1 = (s_1; e_{1,0}, \dots, e_{1,m})$  et  $u_2 = (s_2; e_{2,0}, \dots, e_{2,m})$  deux  $(S, P)$ -chaînes de longueur maximale d'un  $\omega$ -semigroupe  $S$ . Si  $s_1 \mathcal{R} s_2$ , alors les chaînes  $u_1$  et  $u_2$  sont de même nature.*

**Démonstration :** Puisque  $s_1 \mathcal{R} s_2$ , il existe  $x_1$  et  $x_2 \in S_f$  tels que  $s_1 x_1 = s_2$  et  $s_2 x_2 = s_1$ . On considère les idempotents  $f_1 = (e_{1,m} x_1 e_{2,m} x_2 e_{1,m})^\pi$  et  $f_2 = (e_{2,m} x_2 e_{1,m} x_1 e_{2,m})^\pi$ . On remarque que les couples  $(s_1, f_1)$  et  $(s_2, f_2)$  sont liés et que  $e_{1,m} \geq_{\mathcal{H}} f_1$  et  $e_{2,m} \geq_{\mathcal{H}} f_2$ . Un simple calcul montre que  $s_1 f_1^\omega = s_2 f_2^\omega$  :

$$\begin{aligned} s_1 f_1^\omega &= s_1 (e_{1,m} x_1 e_{2,m} x_2 e_{1,m})^\omega \\ &= s_1 e_{1,m} x_1 e_{2,m} (e_{2,m} x_2 e_{1,m} x_1 e_{2,m})^\omega \\ &= s_2 (e_{2,m} x_2 e_{1,m} x_1 e_{2,m})^\omega \\ &= s_2 f_2^\omega \end{aligned}$$

On a l'équivalence  $s_1 e_{1,m}^\omega \in P \Leftrightarrow s_1 f_1^\omega \in P$ . Sinon  $(s_1; e_{1,0}, \dots, e_{1,m}, f_1)$  constitue une chaîne de longueur  $m + 1$ , ce qui contredit la maximalité des chaînes  $u_1$  et  $u_2$ . De façon symétrique, on a aussi l'équivalence  $s_2 e_{2,m}^\omega \in P \Leftrightarrow s_2 f_2^\omega \in P$ . De l'égalité  $s_1 f_1^\omega = s_2 f_2^\omega$ , on déduit  $s_1 e_{1,m}^\omega \in P \Leftrightarrow s_2 e_{2,m}^\omega \in P$ , ce qui montre que les chaînes  $u_1$  et  $u_2$  sont de même nature.  $\square$

Une conséquence de ce lemme est que dans une  $(S, P)$ -superchaîne d'un  $\omega$ -semigroupe  $S$ , les relations  $s_i \geq_{\mathcal{R}} s_{i+1}$  sont strictes. La longueur d'une  $(S, P)$ -superchaîne est alors bornée par le nombre de  $\mathcal{D}$ -classes du semigroupe  $S_f$ .

Si  $(R_0, \dots, R_m)$  et  $(s; e_0, \dots, e_m)$  sont respectivement une  $\mathcal{A}$ -chaîne et une  $(S, P)$ -chaîne de longueur  $m$ , les chaînes  $(R_1, \dots, R_m)$  et  $(s; e_1, \dots, e_m)$  sont des chaînes de nature opposée. Ceci montre que  $|m^+(X) - m^-(X)| \leq 1$ .

De la même façon, en supprimant la première chaîne d'une superchaîne, on obtient une superchaîne de nature opposée, ce qui montre l'inégalité  $|n^+(X) - n^-(X)| \leq 1$ .

Toute superchaîne de longueur au moins égale à 1 contient une chaîne positive et une chaîne négative de longueur  $m(X)$ . Ceci implique que  $m^+(X) = m^-(X) = m(X)$ .

Toute chaîne positive (resp. négative) pour  $X$  constitue une chaîne négative (resp. positive) pour le complémentaire  $A^\omega - X$ . On a donc  $m^+(X) = m^-(A^\omega - X)$  et  $m^-(X) = m^+(A^\omega - X)$ . Il est de même pour les superchaînes et l'on a donc aussi  $n^+(X) = n^-(A^\omega - X)$  et  $n^-(X) = n^+(A^\omega - X)$ .  $\square$

la proposition suivante montre que dans la définition des  $(S, P)$ -chaînes, il est possible de remplacer l'ordre  $\mathcal{H}$  entre les idempotents par les ordres  $\mathcal{R}$  ou  $\mathcal{L}$  sans pour autant modifier la valeur des entiers  $m^+(S, P)$ ,  $m^-(S, P)$  et  $m(S, P)$ .

**Proposition 25** *Soient  $S = (S_f, S_i)$  un  $\omega$ -semigroupe, et  $t_0, \dots, t_m$  des éléments de  $S_i$ . Les propriétés suivantes sont équivalentes*

$$(i) \text{ Il existe des couples liés } (s, e_0), \dots, (s, e_m) \text{ tels que } \begin{cases} se_i^\omega = t_i \\ e_0 \geq_{\mathcal{H}} \dots \geq_{\mathcal{H}} e_m \end{cases}$$

$$(ii) \text{ Il existe des couples liés } (s, e_0), \dots, (s, e_m) \text{ tels que } \begin{cases} se_i^\omega = t_i \\ e_0 \geq_{\mathcal{R}} \dots \geq_{\mathcal{R}} e_m \end{cases}$$

$$(iii) \text{ Il existe des couples liés } (s, e_0), \dots, (s, e_m) \text{ tels que } \begin{cases} se_i^\omega = t_i \\ e_0 \geq_{\mathcal{L}} \dots \geq_{\mathcal{L}} e_m \end{cases}$$

**Démonstration :** Il est clair que (i) entraîne (ii) et (iii). Montrons que (ii) entraîne (i). Soient  $(s, e_0), \dots, (s, e_m)$  des couples liés tels que  $se_i^\omega = t_i$  et  $e_0 \geq_{\mathcal{R}} \dots \geq_{\mathcal{R}} e_m$ . Puisque  $e_i \geq_{\mathcal{R}} e_{i+k}$ , on a  $e_i e_{i+k} = e_{i+k}$  pour  $0 \leq i \leq i+k \leq m$ . On définit les idempotents  $e'_i = e_i \dots e_0$ . On vérifie que

$$\begin{aligned} se'_i &= s \\ se_i'^{\omega} &= s(e_i \dots e_0)^\omega \\ &= se_i(e_{i-1} \dots e_0 e_i)^\omega \\ &= se_i^\omega \\ e'_i e'_{i+k} &= e_i \dots e_0 e_{i+k} \dots e_0 \\ &= (e_i \dots e_0 e_{i+k}) e_{i+k-1} \dots e_0 \\ &= e'_{i+k} \\ e'_{i+k} e'_i &= e_{i+k} \dots e_0 e_i \dots e_0 \\ &= e_{i+k} \dots e_{i+1} (e_i \dots e_0 e_i) e_{i-1} \dots e_0 \\ &= e'_{i+k} \end{aligned}$$

Ceci montre que les couples liés  $(s, e'_0), \dots, (s, e'_m)$  vérifient la condition (i). La démonstration de (iii)  $\Rightarrow$  (i) est duale.  $\square$

## 4.4 Autres caractérisations

### 4.4.1 Dans les $\omega$ -semigroupes

Nous donnons maintenant une formulation équivalente des chaînes lorsque le  $\omega$ -semigroupe considéré est le  $\omega$ -semigroupe  $\hat{S} = (S, S_i)$  engendré par un semi-

groupe fini  $S$ . Nous allons montrer que dans ce cas particulier, il est possible de définir un ordre  $\geq$  sur  $S_i$  tel que les éléments  $t_0, \dots, t_m$  vérifient  $t_0 \geq \dots \geq t_m$  si et seulement si il existe des couples liés  $(s, e_0), \dots, (s, e_m)$  de  $S$  tels que  $e_0 \geq_{\mathcal{H}} \dots \geq_{\mathcal{H}} e_m$ . Nous définissons la relation  $\geq$  entre les couples liés d'un semigroupe fini  $S$  en posant

$$(s_1, e_1) \geq (s_2, e_2) \Leftrightarrow \exists x, y \in S^1 \begin{cases} s_1 x = s_2 & e_1 x \mathcal{L} e_2 & y e_1 x = e_2 \\ s_2 y = s_1 & y e_1 \mathcal{R} e_2 \end{cases}$$

Il faut remarquer que la relation  $(s_1, e_1) \geq (s_2, e_2)$  implique les relations  $s_1 \mathcal{R} s_2$  et  $e_1 \geq_{\mathcal{J}} e_2$ . Nous commençons par donner une définition équivalente de cette relation.

**Proposition 26** *Soient  $(s_1, e_1)$  et  $(s_2, e_2)$  deux couples liés d'un semigroupe fini  $S$ . On a l'équivalence*

$$(s_1, e_1) \geq (s_2, e_2) \Leftrightarrow \exists x, y \in S^1 \begin{cases} s_1 x = s_2 & y e_1 x = e_2 \\ s_2 y = s_1 \end{cases}$$

**Démonstration :** Une des implications est évidente. Réciproquement, on pose  $x' = x e_2$  et  $y' = e_2 y$ . On vérifie que

$$\begin{aligned} s_1 x' &= s_1 x e_2 = s_2 & x' &= x e_2 \text{ et } e_2 = y e_1 x' & y' e_1 x' &= e_2 y e_1 x e_2 = e_2 \\ s_2 y' &= s_2 e_2 y = s_1 & y' &= e_2 y \text{ et } e_2 = y' e_1 x \end{aligned}$$

□

la proposition suivante montre que la relation ainsi définie est compatible avec la relation de conjugaison entre les couples liés.

**Proposition 27** *Deux couples liés  $(s_1, e_1)$  et  $(s_2, e_2)$  d'un semigroupe fini  $S$  sont conjugués si et seulement si ils vérifient  $(s_1, e_1) \geq (s_2, e_2)$  et  $(s_2, e_2) \geq (s_1, e_1)$ .*

**Démonstration :** Si les couples  $(s_1, e_1)$  et  $(s_2, e_2)$  sont conjugués, il existe  $x, y \in S^1$  tels que  $s_1 x = s_2$ ,  $xy = e_1$  et  $yx = e_2$ . D'après les lemmes 3 et 4, on peut en outre supposer que  $x \in R(e_1) \cap L(e_2)$  et  $y \in R(e_2) \cap L(e_1)$ . On a alors  $e_1 x = x$  et  $y e_1 = e_1$  et donc  $(s_1, e_1) \geq (s_2, e_2)$ . Par symétrie, on a également  $(s_2, e_2) \geq (s_1, e_1)$ .

Réciproquement, l'hypothèse  $(s_1, e_1) \geq (s_2, e_2)$  et  $(s_2, e_2) \geq (s_1, e_1)$  implique  $e_1 \mathcal{D} e_2$ . □

La proposition suivante montre que la relation  $\geq$  entre les couple liés est transitive.

**Proposition 28** *Si les trois couples liés  $(s_1, e_1)$ ,  $(s_2, e_2)$  et  $(s_3, e_3)$  vérifient  $(s_1, e_1) \geq (s_2, e_2)$  et  $(s_2, e_2) \geq (s_3, e_3)$ , alors ils vérifient aussi  $(s_1, e_1) \geq (s_3, e_3)$ .*

**Démonstration :** D'après l'hypothèse, il existe  $x_1, y_1, x_2, y_2 \in S^1$  tels que

$$\begin{aligned} s_1 x_1 &= s_2 & x_1 x_1 &\mathcal{L} e_2 & y_1 e_1 x_1 &= e_2 \\ s_2 y_1 &= s_1 & y_1 e_1 &\mathcal{R} e_2 \\ s_2 x_2 &= s_3 & x_2 x_2 &\mathcal{L} e_3 & y_2 e_2 x_2 &= e_3 \\ s_3 y_2 &= s_2 & y_2 e_2 &\mathcal{R} e_3 \end{aligned}$$

On vérifie que

$$\begin{aligned} s_1 x_1 x_2 &= s_2 x_2 = s_3 & e_1 x_1 x_2 &\mathcal{L} e_2 x_2 &\mathcal{L} e_3 & y_2 y_1 e_1 x_1 x_2 &= y_2 e_2 x_2 = e_3 \\ s_3 y_2 y_1 &= s_2 y_1 = s_1 & y_2 y_1 e_1 &\mathcal{R} y_2 e_2 &\mathcal{R} e_3 \end{aligned}$$

et donc  $(s_1, e_1) \geq (s_3, e_3)$ .  $\square$

La relation  $\geq$  ainsi définie est préordre. Puisqu'elle est compatible avec la relation de conjugaison, elle définit un ordre entre les classes de conjugaison. La proposition suivante montre que dans le cas particulier d'un  $\omega$ -semigroupe du type  $\hat{S}$ , cette relation caractérise les chaînes.

**Proposition 29** *Les couples liés  $(s_1, e_1)$  et  $(s_2, e_2)$  vérifient  $(s_1, e_1) \geq (s_2, e_2)$  si et seulement si il existe deux couples liés  $(s'_1, e'_1)$  et  $(s'_2, e'_2)$  respectivement conjugués aux couples  $(s_1, e_1)$  et  $(s_2, e_2)$  tels que  $s'_1 = s'_2$  et  $e'_1 \geq_{\mathcal{H}} e'_2$ .*

**Démonstration :** Si les couples  $(s_1, e_1)$  et  $(s_2, e_2)$  vérifient  $(s_1, e_1) \geq (s_2, e_2)$ , il existe  $x, y \in S^1$  tels que  $s_1 x = s_2$ ,  $s_2 y = s_1$ ,  $y e_1 \mathcal{R} e_2$ ,  $e_1 x \mathcal{L} e_2$  et  $y e_1 x = e_2$ . On pose alors  $s'_1 = s'_2 = s_1$ ,  $e'_1 = e_1$  et  $e'_2 = e_1 x y e_1$ . On vérifie que  $(s'_2, e'_2)$  est un couple lié :  $e'^2_2 = e_1 x y e_1 x y e_1 = e_1 x e_2 y e_1 = e_1 x y e_1$ ,  $s'_2 e'_2 = s_1 e_1 x y e_1 = s_1 = s'_2$ .  $\square$

En raisonnant par récurrence, il est possible de démontrer le même résultat pour  $m$  couple liés.

**Proposition 30** *Les  $m$  couples liés  $(s_1, e_1), \dots, (s_m, e_m)$  d'un semigroupe fini  $S$  vérifient  $(s_1, e_1) \geq \dots \geq (s_m, e_m)$  si et seulement si il existe des couples liés  $(s, e'_1), \dots, (s, e'_m)$  respectivement conjugués aux couples  $(s_1, e_1), \dots, (s_m, e_m)$  tels que  $e_1 \geq_{\mathcal{H}} \dots \geq_{\mathcal{H}} e_m$ .*

#### 4.4.2 Dans les automates

Nous commençons par décrire quelques propriétés générales des chaînes pour un préordre. Nous en déduirons une caractérisation de la longueur des chaînes et des superchaînes dans les automates.

Soit  $\prec$  un préordre sur un ensemble  $E$  et  $\equiv$  la relation d'équivalence associée. On dit que le préordre *sature* une partie  $F$  si et seulement si la relation  $\equiv$  sature la partie  $F$  (et donc son complémentaire  $F^c$ ). On suppose dorénavant que le préordre  $\prec$  sature la partie  $F$ . Cette hypothèse est toujours vérifiée lorsque  $\prec$  est un ordre. Une  $F$ -chaîne de longueur  $m$  est une suite  $(e_0, \dots, e_m)$  d'éléments de  $E$  tel que

$$(i) \quad e_0 \prec e_1 \prec \dots \prec e_m;$$

(ii) les éléments  $e_i$  appartiennent alternativement à  $F$  et à  $F^c$ .

La chaîne est dite positive si  $e_0 \in F$  et négative sinon. On note  $m^+(F)$  (resp.  $m^-(F)$ ) la longueur maximale des chaînes positives (resp. négatives).

Un sous-ensemble  $I$  de  $E$  est dit un intervalle pour le préordre  $\prec$  s'il vérifie

$$\left. \begin{array}{l} x \in I \text{ et } z \in I \\ x \prec y \prec z \end{array} \right\} \Rightarrow y \in I$$

Un sous-ensemble  $I$  de  $E$  est dit un intervalle finissant pour le préordre  $\prec$  s'il vérifie

$$\left. \begin{array}{l} x \in I \\ x \prec y \end{array} \right\} \Rightarrow y \in I$$

Ces définitions sont cohérentes. En effet, un intervalle finissant est un bien un intervalle. Les classes d'équivalence de  $\equiv$  sont des exemples d'intervalles. On appelle  $F$ -recouvrement par intervalles un ensemble  $\{I_0, \dots, I_m\}$  d'intervalles tels que

$$\bigcup_{0 \leq i \leq m} I_i = F.$$

Par définition, la taille d'un tel recouvrement est  $2m$  si au moins un des intervalles est finissant et  $2m + 1$  sinon. En particulier les seuls recouvrements de taille 0 sont ceux constitués d'un unique intervalle finissant et ceux de taille 1 sont ceux constitués d'un unique intervalle non finissant.

Puisque la relation d'équivalence  $\equiv$  est supposée saturer la partie  $F$ , il existe au moins un recouvrement de  $F$ , à savoir celui formé des classes d'équivalence de  $\equiv$  incluses dans la partie  $F$ . On note  $r(F)$  la taille minimale d'un recouvrement de  $F$ . On remarquera qu'un  $F$ -recouvrement de taille minimal contient au plus un intervalle finissant puisque l'union  $I_1 + I_2$  de deux intervalles finissants  $I_1$  et  $I_2$  est un intervalle finissant. La proposition suivante montre que la taille minimale d'un  $F$ -recouvrement est lié à la longueur maximale des  $F$ -chaînes.

**Proposition 31** *On a l'égalité*

$$m^+(F) = r(F)$$

**Démonstration :** Soit  $\{I_0, \dots, I_k\}$  un  $F$ -recouvrement de taille minimale. Si  $r(F)$  est pair, on suppose que  $I_k$  est l'intervalle finissant. Soit  $(e_0, \dots, e_m)$  une  $F$ -chaîne positive. Nécessairement, les éléments  $e_0, e_2, \dots$  appartiennent à des intervalles différents.

- Si aucun des intervalles  $I_i$  n'est finissant, on a  $r(F) = 2k + 1$  et l'inégalité  $m \leq r(F)$  est immédiate.
- Si l'intervalle  $I_k$  est finissant et  $m \leq 2k$ , un des élément de la chaîne appartient à  $I_k$  et ce peut être que  $e_m$ . On alors  $m = 2k = r(F)$ .

Réciproquement, pour un élément  $e \in E$ , on pose  $m^+(e)$  la longueur maximale d'une  $F$ -chaîne positive  $(e_0, \dots, e_m)$  telle que  $e_m = e$ . On a bien sûr  $m^+(e) \leq m^+(F)$  pour tout élément  $e \in E$ . La relation  $e_0 \prec e_1$  implique que  $m^+(e_0) \leq m^+(e_1)$ . On pose alors  $I_k = \{e \in E \mid m^+(e) = 2k\}$ . On vérifie que l'on obtient un  $F$ -recouvrement de taille  $m(F)$ .  $\square$

Puisque les  $F$ -chaînes positives (resp. négatives) sont des  $F^c$ -chaînes négatives (resp. positives), la longueur maximale des  $F$ -chaînes négatives est caractérisée par la taille minimale des recouvrements de  $F^c$  par intervalles.

$$m^-(F) = m^+(F^c) = r(F^c)$$

Pour un automate de muller  $\mathcal{A} = (Q, A, \cdot, q_0, \mathcal{T})$ , les résultats précédents s'appliquent avec  $E$  l'ensemble des parties répétibles,  $F$  la table  $\mathcal{T}$  de l'automate et l'inclusion des ensembles comme préordre. Puisque l'inclusion est un ordre, la partie  $F$  est automatiquement saturée par le préordre. On obtient alors une caractérisation de la longueur maximale  $m^+(\mathcal{A})$  des  $\mathcal{A}$ -chaînes. L'entier  $m^+(\mathcal{A})$  est égal à  $m$  si et seulement si il existe un recouvrement de taille  $m$  de la table de l'automate. De même, l'entier  $m^+(\mathcal{A})$  est inférieur ou égal à  $m$  si et seulement si il existe un recouvrement de taille au plus  $m$  de la table de l'automate.

Lorsque les entiers  $m^+(\mathcal{A})$  et  $m^-(\mathcal{A})$  sont égaux, on peut obtenir de la même façon une caractérisation de l'entier  $n^+(\mathcal{A})$ . Soit  $E$  l'ensemble des parties répétibles  $R$  telles qu'il existe une  $\mathcal{A}$ -chaîne  $(R_0, \dots, R_m)$  de longueur  $m(\mathcal{A})$  telle que  $R_m = R$ . Soit  $F$  l'ensemble des parties de  $E$  qui sont acceptantes, *i.e.*,  $F = E \cap \mathcal{T}$ . L'accessibilité des parties constitue un préordre sur  $E$ . D'après le lemme 12 (p. 74), ce préordre sature l'ensemble  $F$  puisque deux  $\mathcal{A}$ -chaînes  $U_1$  et  $U_2$  telles que  $U_1$  est accessible de  $U_2$  et  $U_1$  accessible de  $U_2$  sont nécessairement de même nature. L'entier  $n^+(\mathcal{A})$  est donc égal à  $n$  si et seulement si il existe un recouvrement de taille  $n$  de  $F$  par intervalles pour l'accessibilité.

## 4.5 Classes de complexité

Nous introduisons tout d'abord quelques notations. Nous notons

- $\mathbf{R}$  l'ensemble des langages reconnaissables de  $A^\omega$ ;
- $\mathbf{G}$  l'ensemble des ouverts reconnaissables de  $A^\omega$ ;
- $\mathbf{F}$  l'ensemble des fermés reconnaissables de  $A^\omega$ .

Nous définissons un préordre sur les parties reconnaissables de  $A^\omega$  en ordonnant les quadruplets  $(m^+, m^-, n^+, n^-)$  de  $\mathbb{N}^4$  par

$$(m_1^+, m_1^-, n_1^+, n_1^-) \leq (m_2^+, m_2^-, n_2^+, n_2^-) \Leftrightarrow \begin{cases} m_1^+ < m_2^+ \text{ et } m_1^- \leq m_2^- \\ \text{ou} \\ m_1^+ \leq m_2^+ \text{ et } m_1^- < m_2^- \\ \text{ou} \\ m_1^+ = m_2^+ \text{ et } m_1^- = m_2^- \text{ et } n_1^+ \leq n_2^+ \text{ et } n_1^- \leq n_2^- \end{cases}$$

Compte tenu des contraintes liant ces quatre entiers, d'après la proposition 24 (p. 73), les classes d'équivalence associées sont les suivantes :

$$\begin{aligned} \mathbf{C}_m^n &= \{X \in \mathbf{R} \mid m(X) = m, n^+(X) = n - 1, n^-(X) = n\} \\ \mathbf{D}_m^n &= \{X \in \mathbf{R} \mid m(X) = m, n^+(X) = n, n^-(X) = n - 1\} \\ \mathbf{E}_m^n &= \{X \in \mathbf{R} \mid m(X) = m, n^+(X) = n - 1, n^-(X) = n - 1\} \end{aligned}$$

Ainsi, les éléments de  $\mathbf{D}_m^n$  sont les complémentaires de ceux de  $\mathbf{C}_m^n$ . On dit que ces classes sont *duales* l'une de l'autre. On a  $\mathbf{E}_m^n = \mathbf{C}_m^n \cap \mathbf{D}_m^n$  et les classes  $\mathbf{E}_m^n$  sont dites, elles, *autoduales* car elles contiennent les complémentaires de leurs éléments. L'ordre induit sur les classes est représenté à la figure 4.7.

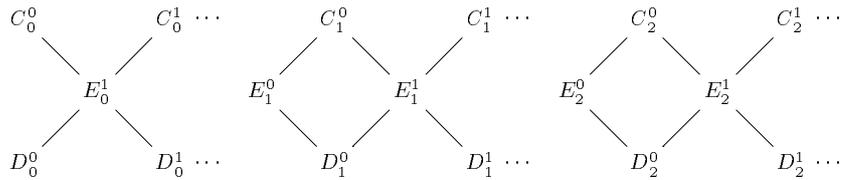


Figure 4.7 : L'ordre induit sur les classes

Nous utilisons aussi les classes 'descendantes' définies comme

$$\begin{aligned}\hat{\mathbf{C}}_m^n &= \{X \in \mathbf{R} \mid m(X) < m \text{ ou } (m(X) = m \text{ et } n^+(X) \leq n - 1)\} \\ \hat{\mathbf{D}}_m^n &= \{X \in \mathbf{R} \mid m(X) < m \text{ ou } (m(X) = m \text{ et } n^-(X) \leq n - 1)\} \\ \hat{\mathbf{E}}_m^n &= \{X \in \mathbf{R} \mid m(X) < m \text{ ou } (m(X) = m \text{ et } n(X) \leq n - 1)\}\end{aligned}$$

Nous allons voir que pour certains cas particuliers, la définition des classes se simplifie. Puisque tout  $\omega$ -langage reconnaissable  $X$  vérifie  $m(X) \geq 0$ , on a pour  $m = 0$  les égalités

$$\begin{aligned}\hat{\mathbf{C}}_0^n &= \{X \in \mathbf{R} \mid m(X) = 0 \text{ et } n^+(X) \leq n - 1\} \\ \hat{\mathbf{D}}_0^n &= \{X \in \mathbf{R} \mid m(X) = 0 \text{ et } n^-(X) \leq n - 1\} \\ \hat{\mathbf{E}}_0^n &= \{X \in \mathbf{R} \mid m(X) = 0 \text{ et } n(X) \leq n - 1\}\end{aligned}$$

De même pour  $n = 0$ , l'inégalité  $n^+(X) < -1$  signifie par exemple qu'il n'existe pas de superchaîne positive de longueur 0 et donc qu'il n'existe pas de chaîne positive de longueur maximale. La définition prend alors la forme

$$\begin{aligned}\hat{\mathbf{C}}_m^0 &= \{X \in \mathbf{R} \mid m^+(X) \leq m - 1\} \\ \hat{\mathbf{D}}_m^0 &= \{X \in \mathbf{R} \mid m^-(X) \leq m - 1\} \\ \hat{\mathbf{E}}_m^0 &= \{X \in \mathbf{R} \mid m(X) \leq m - 1\}\end{aligned}$$

On a en particulier l'égalité

$$\hat{\mathbf{E}}_m^0 = \bigcup_{n \geq 0} \hat{\mathbf{C}}_{m-1}^n = \bigcup_{n \geq 0} \hat{\mathbf{D}}_{m-1}^n \quad (10)$$

Nous allons maintenant examiner  $\mathbf{C}_m^n$  pour les premières valeurs des entiers  $m$  et  $n$ .

**Proposition 32** *Pour tout alphabet  $A$*

$$\mathbf{C}_0^0 = \hat{\mathbf{C}}_0^0 = \{\emptyset\} \quad \text{et} \quad \mathbf{D}_0^0 = \hat{\mathbf{D}}_0^0 = \{A^\omega\}$$

**Démonstration :** Il a été vu au cours de la démonstration de la proposition 24 que tout  $\omega$ -langage reconnaissable non vide admet une chaîne positive puisqu'il contient un ensemble de la forme  $YZ^\omega$ . Le langage vide est donc le seul langage à appartenir à la classe  $\mathbf{C}_0^0$ . Le résultat  $\mathbf{D}_0^0 = \{A^\omega\}$  est dual et s'obtient par passage au complémentaire.  $\square$

**Proposition 33** *Pour tout alphabet  $A$*

$$\hat{\mathbf{C}}_0^1 = \mathbf{G} \quad \text{et} \quad \hat{\mathbf{D}}_0^1 = \mathbf{F}$$

**Exemple :** Le  $\omega$ -langage  $X_3 = (b + c)^*aA^\omega$  sur l'alphabet  $A = a + b + c$  des mots ayant au moins une occurrence de  $a$  est ouvert. Un automate reconnaissant ce langage est donné à la figure 4.8. On vérifie que  $m(\mathcal{A}_3) = 0$ , que la seule  $\mathcal{A}_3$ -superchaîne de longueur 1 est  $(\{1\}, \{2\})$  et que cette superchaîne est négative.

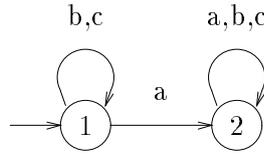


Figure 4.8 : Automate  $\mathcal{A}_3 : \mathcal{T} = \{\{2\}\}$

Nous avons montré à la proposition 23 (p. 69), que la longueur maximale des chaînes positives peut être calculée dans un  $\omega$ -semigroupe reconnaissant le  $\omega$ -langage considéré. Dans le cadre des  $\omega$ -semigroupes, les résultats de la proposition précédente s'énoncent ainsi : soit  $X \subset A^\omega$  un  $\omega$ -langage,  $S = (S_f, S_i)$  un  $\omega$ -semigroupe fini et  $\varphi : A^\omega \rightarrow S$  un morphisme surjectif reconnaissant  $X$ . Soit  $P$  l'image de  $X$  dans  $S$ .

- (1)  $X$  est fermé si et seulement si pour tous couples liés  $(s_1, e_1)$  et  $(s_2, e_2)$  de  $S_f$  on a l'implication  $s_2 \geq_{\mathcal{R}} s_1$  et  $s_1 e_1^\omega \in P \Rightarrow s_2 e_2^\omega \in P$ .
- (2)  $X$  est ouvert si et seulement si pour tous couples liés  $(s_1, e_1)$  et  $(s_2, e_2)$  de  $S_f$  on a l'implication  $s_1 \geq_{\mathcal{R}} s_2$  et  $s_1 e_1^\omega \in P \Rightarrow s_2 e_2^\omega \in P$ .

**Démonstration :** (1) Soit  $X$  un langage fermé et  $u \in A^\omega$  un mot tel que  $u\varphi = s_2 e_2^\omega$ . Le mot  $u$  se factorise  $u = u_0 u_1 u_2 \dots$  avec  $u_0 \varphi = s_2$ ,  $u_1 \varphi = u_2 \varphi = \dots = e_2$ . Puisque  $s_2 \geq_{\mathcal{R}} s_1$ , il existe un élément  $t \in S_f$  tel que  $s_2 t = s_1$ . On choisit deux mots  $v$  et  $w$  tels que  $v\varphi = t$  et  $w\varphi = e_1$ . On définit le mot  $x_k = u_0 \dots u_k v w^\omega$ . Puisque  $x_k \varphi = s_1 e_1^\omega$ , chacun des mots  $x_k$  appartient à  $X$ . Les mots  $u$  et  $x_k$  coïncident sur le préfixe  $u_0 \dots u_k$  et la suite  $x_k$  converge vers  $u$ . Le mot  $u$  appartient donc à  $X$  et  $s_2 e_2^\omega$  appartient à  $P$ .

Réciproquement, soit  $u$  un élément de l'adhérence de  $X$  et  $u = u_0 u_1 u_2 \dots$  avec  $u_0 \varphi = s_2$ ,  $u_1 \varphi = u_2 \varphi = \dots = e_2$  une factorisation du mot  $u$ . Il existe un mot  $x \in X$  tel que  $u$  et  $x$  coïncident sur un préfixe de longueur supérieure à  $|u_0|$ . On pose  $x\varphi = s_1 e_1^\omega \in P$ . Soit  $x = v_0 v_1 v_2 \dots$  une factorisation de  $x$  telle que  $v_0 \varphi = s_1$  et  $v_1 \varphi = v_2 \varphi = \dots = e_1$ . Quitte à regrouper les premiers termes de la factorisation, on peut toujours supposer  $|v_0| \geq |u_0|$  ce qui implique  $s_1 \leq_{\mathcal{R}} s_2$ . On en déduit que  $s_2 e_2^\omega \in P$  et que  $u \in X$ . L'ensemble  $X$  est donc fermé.

(2) Ce résultat est dual du précédent et se déduit naturellement en passant au complémentaire.  $\square$

**Proposition 34** *Pour tout  $n$ , on a l'égalité*

$$\hat{C}_0^n = C_n(\mathbf{G})$$

**Exemple :** Soit un entier  $n$  donné. Pour un mot  $u$ , on note  $|u|_{a,n}$  le nombre d'occurrences de  $a$  dans le mot  $u$  si ce nombre est inférieur à  $n$  ou  $n$  sinon. Soit  $X_4$  le  $\omega$ -langage des mots  $u$  tels que  $|u|_{a,n}$  soit impair. Un automate reconnaissant le langage  $X_4$  est donné à la figure 4.9. On vérifie que  $m(\mathcal{A}_4) = 0$  puisque les seuls ensembles répétables sont les singletons. La seule  $\mathcal{A}_4$ -superchaîne de longueur  $n$  est  $(\{1\}, \dots, \{n+1\})$  et cette superchaîne est négative. Le langage  $X_4$  appartient donc à  $\hat{C}_0^n$ . Le langage  $X_4$  se décompose  $X_4 = Y_1 - Y_2 + \dots \pm Y_n$  avec  $Y_k = \{u \mid |u|_{a,n} \geq k\}$  pour  $1 \leq k \leq n$ .

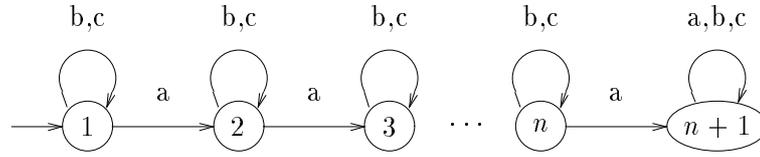


Figure 4.9 : Automate  $\mathcal{A}_4 : \mathcal{T} = \{\{2\}, \{4\}, \dots\}$

**Démonstration :** Nous établissons deux lemmes

**Lemme 14** *Soient  $X$  et  $Y$  deux  $\omega$ -langages. On a les inégalités suivantes*

$$\begin{aligned} m(X + Y) &\leq m(X) + m(Y) \\ m(XY) &\leq m(X) + m(Y) \\ m(X \Delta Y) &\leq m(X) + m(Y) \end{aligned}$$

**Démonstration :** Nous utilisons la caractérisation de l'entier  $m(X)$  à l'aide des  $\omega$ -semigroupes. Si les  $\omega$ -langages  $X$  et  $Y$  sont respectivement reconnus par les morphismes  $\varphi : A^\infty \rightarrow S$  et  $\psi : A^\infty \rightarrow T$  où  $S = (S_f, S_i)$  et  $T = (T_f, T_i)$  sont des  $\omega$ -semigroupes finis, les  $\omega$ -langages  $X + Y$ ,  $XY$  et  $X \Delta Y$  sont reconnus par le morphisme  $\varphi \times \psi : A^\infty \rightarrow S \times T$  où  $S \times T$  dénote le  $\omega$ -semigroupe produit. Nous commençons par rappeler quelques résultats élémentaires sur le  $\omega$ -semigroupe  $S \times T$

- (1)  $(e, f)$  est un idempotent de  $S_f \times T_f$  si et seulement si  $e$  et  $f$  sont respectivement des idempotents de  $S_f$  et de  $T_f$ ;
- (2)  $((s, t), (e, f))$  est un couple lié de  $S_f \times T_f$  si et seulement si  $(s, e)$  et  $(t, f)$  sont respectivement des couples liés de  $S_f$  et de  $T_f$ ;
- (3)  $(e_1, f_1) \geq_{\mathcal{H}} (e_2, f_2)$  dans  $S_f \times T_f$  si et seulement si  $e_1 \geq_{\mathcal{H}} e_2$  et  $f_1 \geq_{\mathcal{H}} f_2$ ;
- (4)  $(s_1, t_1) \geq_{\mathcal{R}} (s_2, t_2)$  dans  $S_f \times T_f$  si et seulement si  $s_1 \geq_{\mathcal{R}} s_2$  et  $t_1 \geq_{\mathcal{R}} t_2$ .

Si  $P$  et  $Q$  sont les images respectives de  $X$  et  $Y$  dans  $S$  et  $T$  par  $\varphi$  et  $\psi$ , on a

$$\begin{aligned} (X + Y)(\varphi \times \psi) &= P \times T_i + S_i \times Q \\ (XY)(\varphi \times \psi) &= P \times Q \\ (X \triangle Y)(\varphi \times \psi) &= P \times (T_i - Q) + (S_i - P) \times Q \end{aligned}$$

Nous faisons uniquement la preuve pour l'intersection  $XY$ , les deux autres étant similaires. Nous allons montrer que s'il existe une  $(S \times T, P \times Q)$ -chaîne de longueur  $m$ , il existe deux entiers  $p$  et  $q$  vérifiant  $p + q \geq m$ , une  $(S, P)$ -chaîne de longueur  $p$  et une  $(T, Q)$ -chaîne de longueur  $q$ . Soit  $((s, t); (e_0, f_0), \dots, (e_m, f_m))$  une  $(S \times T, P \times Q)$ -chaîne de longueur  $m$ . On considère les indices  $i$  où les deux éléments successifs  $se_{i-1}^\omega$  et  $se_i^\omega$  alternent c'est à dire lorsqu'ils ne sont pas simultanément dans  $P$  ou hors de  $P$ . On considère également les indices  $j$  où les éléments  $tf_{j-1}^\omega$  et  $tf_j^\omega$  alternent. De façon plus rigoureuse, on définit les ensembles d'entiers  $I$  et  $J$  par

$$\begin{aligned} I &= \{i \in [1; m] \mid se_{i-1}^\omega \in P \Leftrightarrow se_i^\omega \notin P\} \\ J &= \{j \in [1; m] \mid tf_{j-1}^\omega \in Q \Leftrightarrow tf_j^\omega \notin Q\} \end{aligned}$$

La suite  $((s, t); (e_0, f_0), \dots, (e_m, f_m))$  constitue une  $(S \times T, P \times Q)$ -chaîne. Pour tout entier  $k$ , les éléments successifs  $(se_{k-1}^\omega, tf_{k-1}^\omega)$  et  $(se_k^\omega, tf_k^\omega)$  associés à cette chaîne alternent, c'est à dire, ne sont pas simultanément dans  $P \times Q$  ou hors de  $P \times Q$ . Ceci implique qu'au moins un des couples d'éléments successifs  $(se_{k-1}^\omega, se_k^\omega)$  ou  $(tf_{k-1}^\omega, tf_k^\omega)$  est une alternance. Autrement dit, l'entier  $k$  appartient à  $I$  ou à  $J$ . On a donc l'inclusion  $[1; m] \subset I + J$ . On pose alors  $I = \{i_1 < \dots < i_p\}$  et  $J = \{j_1 < \dots < j_q\}$  où  $p$  et  $q$  sont les cardinaux respectifs de  $I$  et  $J$ . On a bien  $p + q \geq m$  et les suites  $(s; e_0, e_{i_1}, \dots, e_{i_p})$  et  $(t; f_0, f_{j_1}, \dots, f_{j_q})$  sont respectivement une  $(S, P)$ -chaîne et une  $(T, Q)$ -chaîne. En particulier si l'entier  $m$  est strictement supérieur à  $m(X) + m(Y)$ , l'entier  $p$  est supérieur à  $m(X)$  ou l'entier  $q$  est supérieur à  $m(Y)$ , ce qui conduit à une contradiction.  $\square$

**Lemme 15** Soient  $X$  et  $Y$  deux  $\omega$ -langages. On a les implications suivantes

$$\begin{aligned} m(X + Y) &= m(X) + m(Y) \Rightarrow n(X + Y) \leq n(X) + n(Y) \\ m(XY) &= m(X) + m(Y) \Rightarrow n(XY) \leq n(X) + n(Y) \\ m(X \Delta Y) &= m(X) + m(Y) \Rightarrow n(X \Delta Y) \leq n(X) + n(Y) \end{aligned}$$

**Démonstration :** Nous effectuons la preuve uniquement pour l'intersection et nous suivons un schéma de preuve identique à celui du lemme précédent. On suppose que les  $\omega$ -langages  $X$  et  $Y$  sont respectivement reconnus par les morphismes  $\varphi : A^\infty \rightarrow S$  et  $\psi : A^\infty \rightarrow T$  où  $S = (S_f, S_i)$  et  $T = (T_f, T_i)$  sont des  $\omega$ -semigroupes finis. Nous allons prouver que s'il existe une  $(S \times T, P \times Q)$ -superchaîne de longueur  $n$ , il existe deux entiers  $p$  et  $q$  vérifiant  $p + q \geq n$ , une  $(S, P)$ -superchaîne de longueur  $p$  et une  $(T, Q)$ -superchaîne de longueur  $q$ . Soit  $(u_0, \dots, u_n)$  une  $(S \times T, P \times Q)$ -superchaîne de longueur  $n$  où  $u_i$  est une  $(S \times T, P \times Q)$ -chaîne de longueur  $m = m(X) + m(Y)$  :

$$u_i = ((s_i, t_i); (e_{i,0}, f_{i,0}), \dots, (e_{i,m}, f_{i,m})).$$

On introduit à nouveau les ensembles d'alternance  $I$  et  $J$  en posant

$$\begin{aligned} I &= \{i \in [1; n] \mid s_{i-1}e_{i-1,0}^\omega \in P \Leftrightarrow s_i e_{i,0}^\omega \notin P\} \\ J &= \{j \in [1; n] \mid t_{j-1}f_{j-1,0}^\omega \in Q \Leftrightarrow t_j f_{j,0}^\omega \notin Q\} \end{aligned}$$

Puisque les chaînes successives  $u_{k-1}$  et  $u_k$  alternent pour tout  $k$ , on a l'inclusion  $[1; n] \subset I + J$ . On pose  $I = \{i_1 < \dots < i_p\}$  et  $J = \{j_1 < \dots < j_q\}$  où  $p$  et  $q$  sont les cardinaux respectifs de  $I$  et  $J$ . Puisque pour tout  $k$ , la chaîne  $u_k$  est de longueur  $m = m(X) + m(Y)$ , on peut trouver, d'après la démonstration du lemme précédent, une  $(S, P)$ -chaîne  $u_{k,S}$  de longueur  $m(X)$  et une  $(T, Q)$ -chaîne  $u_{k,T}$  de longueur  $m(Y)$  telles que  $u_{k,S}$  et  $u_{k,T}$  sont respectivement de la forme  $(s_k, e_{k,0}, \dots)$  et  $(t_k; e_{k,0}, \dots)$ . On a alors la  $(S, P)$ -superchaîne  $(u_{i_1,S}, \dots, u_{i_p,S})$  de longueur  $p$  et la  $(T, Q)$ -superchaîne  $(u_{j_1,T}, \dots, u_{j_q,T})$  de longueur  $q$ . En particulier si l'entier  $n$  est strictement supérieur à  $n(X) + n(Y)$ , l'entier  $p$  est supérieur à  $n(X)$  ou l'entier  $q$  est supérieur à  $n(Y)$ , ce qui conduit à une contradiction.  $\square$

Soit  $X \in C_n(\mathbf{G})$ . Le  $\omega$ -langage s'écrit  $X = X_1 \Delta \dots \Delta X_n$  où  $X_1, \dots, X_n \in \mathbf{G}$ . Grâce aux lemmes 14 et 15, on a  $X \in \hat{\mathbf{C}}_0^n$  ce qui démontre l'inclusion  $C_n(\mathbf{G}) \subset \hat{\mathbf{C}}_0^n$ .

Réciproquement, soit  $X \in \hat{\mathbf{C}}_0^n$  un  $\omega$ -langage,  $S = (S_f, S_i)$  un  $\omega$ -semigroupe fini et  $\varphi : A^\infty \rightarrow S$  un morphisme surjectif reconnaissant  $X$ . Soit  $P$  l'image de  $X$  dans  $S$ . Pour un élément  $s \in S_f$ , on pose  $n^+(s)$  la longueur maximale d'une  $(S, P)$ -superchaîne positive  $(u_0, \dots, u_n)$  telle que  $u_n = (s; \epsilon)$ . On a bien sûr l'inégalité  $n^+(s) \leq n^+(X) \leq n - 1$  pour tout élément  $s \in S_f$ . On pose

alors  $P_k = \{se^\omega \in S_i \mid n^+(s) \geq k \text{ pour } 0 \leq k \leq n-1\}$ . On vérifie que les  $\omega$ -langages  $X_k = P_k\varphi^{-1}$  sont des ouverts. En effet, la relation  $s_1 \geq_{\mathcal{R}} s_2$  entraîne  $n^+(s_1) \leq n^+(s_2)$  ce qui assure que l'ensemble  $P_k$  vérifie la condition nécessaire et suffisante pour que  $X_k$  soit ouvert :  $s_1 \geq_{\mathcal{R}} s_2$  et  $s_1 e_1^\omega \in P_k \Rightarrow s_2 e_2^\omega \in P_k$ . On a l'égalité

$$P = P_0 - P_1 + \cdots \pm P_n$$

qui implique immédiatement

$$X = X_0 - X_1 + \cdots \pm X_n$$

Le  $\omega$ -langage  $X$  appartient à  $C_n(\mathbf{G})$  ce qui montre l'inclusion  $\hat{C}_0^n \subset C_n(\mathbf{G})$ .  $\square$

**Proposition 35** *On a l'égalité*

$$\hat{C}_1^0 = \mathbf{G}_\delta \quad \text{et} \quad \hat{D}_1^0 = \mathbf{F}_\sigma$$

**Exemple :** Le  $\omega$ -langage  $X_1 = (b^*a)^\omega$  des mots ayant un nombre infini de  $a$  est reconnu par l'automate  $\mathcal{A}_1$  de la figure 4.3 avec  $\mathcal{T} = \{\{1\}, \{1, 2\}\}$ . Cet automate est à table pleine. Par conséquent, le langage  $X_1$  est un  $\mathbf{G}_\delta$ .

**Démonstration :** Un langage  $X$  est déterministe si et seulement si un automate  $\mathcal{A}$  reconnaissant  $X$  est à table pleine, ce qui est équivalent avec l'inégalité  $m^+(X) = m^+(\mathcal{A}) \leq 0$ .  $\square$

On obtient le corollaire suivant

**Corollaire 3** *La famille  $\hat{\mathbf{E}}_1^0 = \mathbf{F}_\sigma \cap \mathbf{G}_\delta$  coïncide avec la clôture booléenne de la famille des ouverts reconnaissables*

$$\mathbf{F}_\sigma \cap \mathbf{G}_\delta = \bigcup_{m \geq 0} C_m(\mathbf{G})$$

**Démonstration :** La proposition précédente donne l'égalité

$$\hat{\mathbf{E}}_1^0 = \hat{C}_1^0 \cap \hat{D}_1^0 = \mathbf{F}_\sigma \cap \mathbf{G}_\delta$$

D'après la formule 10, on a aussi l'égalité

$$\hat{\mathbf{E}}_1^0 = \bigcup_{n \geq 0} \hat{C}_0^n = \bigcup_{n \geq 0} C_n(\mathbf{G})$$

ce qui donne le résultat.  $\square$

**Proposition 36** *Pour tout  $m$ , on a l'égalité*

$$\hat{C}_m^0 = C_m(\mathbf{G}_\delta)$$

**Exemple :** Le  $\omega$ -langage reconnu par l'automate  $\mathcal{A}_5$  de la figure 4.10 avec  $\mathcal{T} = \{\{0, 1\}, \{0, 1, 2, 3\}, \dots\}$ . appartient à  $\hat{C}_m^0 = C_m(\mathbf{G}_\delta)$  puisque l'automate admet une seule  $\mathcal{A}_5$ -chaîne de longueur  $m : (\{0\}, \{0, 1\}, \{0, 1, 2\}, \dots, \{0, \dots, m\})$  et cette chaîne est négative.

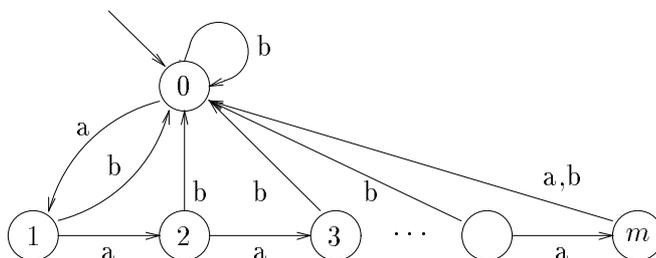


Figure 4.10 : Automate  $\mathcal{A}_5 : \mathcal{T} = \{\{0, 1\}, \{0, 1, 2, 3\}, \dots\}$

**Démonstration :** Soit  $X \in C_m(\mathbf{G}_\delta)$ . Le  $\omega$ -langage s'écrit  $X = X_1 \Delta \dots \Delta X_n$  où  $X_1, \dots, X_n \in \mathbf{G}_\delta$ . Grâce aux lemmes 14 et 15, on a  $X \in \hat{C}_m^0$  ce qui démontre l'inclusion  $C_m(\mathbf{G}_\delta) \subset \hat{C}_m^0$ .

Réciproquement, soit  $X \in \hat{C}_m^0$  un  $\omega$ -langage,  $S = (S_f, S_i)$  un  $\omega$ -semigroupe fini et  $\varphi : A^\infty \rightarrow S$  un morphisme surjectif reconnaissant  $X$ . Soit  $P$  l'image de  $X$  dans  $S$ . Pour un élément  $t \in S_i$ , on pose  $m^+(s)$  la longueur maximale d'une  $(S, P)$ -chaîne positive  $(s; e_0, \dots, e_m)$  telle que  $se_m^\omega = t$ . On a bien sur l'inégalité  $m^+(t) \leq m^+(X) \leq m - 1$  pour tout élément  $t \in S_i$ . On pose alors  $P_k = \{t \in S_i \mid m^+(t) \geq k \text{ pour } 0 \leq k \leq m - 1\}$ . On vérifie que les  $\omega$ -langages  $X_k = P_k \varphi^{-1}$  sont des ensembles  $\mathbf{G}_\delta$ . En effet si  $(s, e_0)$  et  $(s, e_1)$  sont deux couples liés tels que  $e_0 \geq_{\mathcal{H}} e_1$ , alors  $m^+(se_0^\omega) \leq m^+(se_1^\omega)$ . Il n'existe donc pas de  $(S, P_k)$ -chaîne positive. On a l'égalité

$$P = P_0 - P_1 + \dots \pm P_m$$

qui implique immédiatement

$$X = X_0 - X_1 + \dots \pm X_m$$

Le  $\omega$ -langage  $X$  appartient à  $C_m(\mathbf{G}_\delta)$  ce qui montre l'inclusion  $\hat{C}_m^0 \subset C_m(\mathbf{G}_\delta)$ .  $\square$

### 4.5.1 Indice de Rabin

Nous sommes maintenant en mesure de caractériser l'indice de Rabin d'un  $\omega$ -langage. On a le théorème

**Théorème 16** *Pour tout  $\omega$ -langage  $X$ , on a l'équivalence*

$$\text{ind}(X) \leq m \quad \Leftrightarrow \quad X \in \hat{\mathbf{C}}_m^0.$$

**Démonstration :** Si le  $\omega$ -langage vérifie  $\text{ind}(X) = m$ , il existe un automate de Rabin  $\mathcal{A} = (Q, A, E, q_0, \mathcal{R})$  avec  $\mathcal{R} = (R_1, R_2, \dots, R_m)$  tel que  $L^\omega(\mathcal{A}) = X$ . Par définition, le  $\omega$ -langage s'écrit comme une somme de différence de langages déterministes.

$$L^\omega(\mathcal{A}) = (L^\omega(R_1) - L^\omega(R_2)) + (L^\omega(R_3) - L^\omega(R_4)) + \dots$$

La longueur de la somme de différence étant  $m$ , le langage  $X$  appartient à  $C_m(\mathbf{G}_\delta) = \hat{\mathbf{C}}_m^0$ .

Réciproquement, si le  $\omega$ -langage  $X$  appartient à  $C_m(\mathbf{G}_\delta)$ , il peut s'écrire comme une chaîne de longueur  $m$  de  $\omega$ -langages déterministes reconnaissables.

$$X = X_1 - X_1 + X_3 - \dots \pm X_m$$

Chacun des  $\omega$ -langages  $X_i$  est reconnu par un automate de Büchi  $\mathcal{A}_i$ . En considérant le produit de ces automates, il est aisé de construire un automate de Rabin d'indice  $m$  reconnaissant le  $\omega$ -langage  $X$ .  $\square$

Cette caractérisation a le corollaire suivant.

**Corollaire 4** *Pour tous  $\omega$ -langages  $X$  et  $Y$  reconnaissables, on a les inégalités qui correspondent aux formules de la proposition 15 (p. 50).*

$$\begin{aligned} \text{ind}(XY) &\leq \begin{cases} \text{ind}(X) + \text{ind}(Y) - 2 & \text{si } m \text{ et } n \text{ pairs} \\ \text{ind}(X) + \text{ind}(Y) - 1 & \text{sinon} \end{cases} \\ \text{ind}(X + Y) &\leq \begin{cases} \text{ind}(X) + \text{ind}(Y) - 1 & \text{si } m \text{ et } n \text{ impairs} \\ \text{ind}(X) + \text{ind}(Y) & \text{sinon} \end{cases} \\ \text{ind}(X \Delta Y) &\leq \text{ind}(X) + \text{ind}(Y) \end{aligned}$$

### 4.5.2 Stabilité par fonction séquentielle inverse

Nous allons montrer que les classes  $\hat{\mathbf{C}}_m^n$  et  $\hat{\mathbf{D}}_m^n$  sont stables par image inverse de fonction séquentielle.

**Théorème 17** *Soit  $\sigma : A^* \rightarrow B^*$  une fonction séquentielle, et soit  $X \subset B^*$  un  $\omega$ -langage reconnaissable. On a alors les implications*

$$\begin{aligned} X \in \hat{\mathbf{C}}_m^n &\Rightarrow X\sigma^{-1} \in \hat{\mathbf{C}}_m^n \\ X \in \hat{\mathbf{D}}_m^n &\Rightarrow X\sigma^{-1} \in \hat{\mathbf{D}}_m^n \end{aligned}$$

**Démonstration :** On suppose la fonction séquentielle  $\sigma : A^* \rightarrow B^*$  réalisée par le transducteur  $T = (Q, A, B, q_0, \cdot, *)$  de monoïde de transition  $M(\sigma)$ . D'après théorème 15 (p. 44), si le  $\omega$ -langage  $X$  est reconnu par le morphisme  $\varphi : B^\infty \rightarrow S$  où  $S = (S_f, S_i)$  est un  $\omega$ -semigroupe fini, le  $\omega$ -langage  $X\sigma^{-1}$  est reconnu par le morphisme  $\psi : A^\infty \rightarrow S \circ M$ . Si l'image de  $X$  par  $\varphi$  dans  $S$  est égale à  $P$ , l'image du langage  $X\sigma^{-1}$  par  $\psi$  dans le produit en couronne  $S \circ M$  est égale à la partie  $Q$  définie par

$$Q = \{g \in S_i^M \mid 1g \in P\}$$

où on a identifié  $S_i^M$  et la partie infinitaire du  $\omega$ -semigroupe  $S \circ M$ . Soit un couple lié  $((f, s), (g, e))$  de  $S_f \circ M$ . Pour tout  $t \in M$ , le couple  $(tf, tsg)$  est un couple lié de  $S_f$  et d'après la formule 7 (p. 41), on a l'égalité

$$(f, s)(g, e)^\omega = t \mapsto tf(tsg)^\omega$$

En particulier pour  $t = 1$ , on a

$$1(f, s)(g, e)^\omega = 1f.(sg)^\omega$$

Nous allons montrer que s'il existe une  $(S \circ M, Q)$ -chaîne de longueur  $m$ , alors il existe une  $(S, P)$ -chaîne de longueur  $m$  et de même nature (positive ou négative). Nous montrerons également que s'il existe une  $(S \circ M, Q)$ -superchaîne de longueur  $n$ , alors il existe une  $(S, P)$ -superchaîne de longueur  $n$  et de même nature. Nous démontrons auparavant deux lemmes.

**Lemme 16** *Soient  $((f, s), (g_1, e_1))$  et  $((f, s), (g_2, e_2))$  deux couples liés de  $S_f \circ M$ . La relation  $(g_1, e_1) \geq_{\mathcal{R}} (g_2, e_2)$  implique la relation  $sg_1 \geq_{\mathcal{R}} sg_2$ .*

**Démonstration :** D'après l'hypothèse, il existe un élément  $(h, x)$  de  $S_f \circ M$  tel que  $(g_1, e_1)(h, x) = (g_2, e_2)$ . Pour tout  $t \in M$ , on a  $tg_1.te_1h = tg_2$ . En prenant  $t = s$ , on obtient  $sg_1.se_1h = sg_2$  ce qui termine la preuve du premier lemme.  $\square$

**Lemme 17** *Soient  $((f_1, s_1), (g_1, e_1))$  et  $((f_2, s_2), (g_2, e_2))$  deux couples liés de  $S_f \circ M$ . La relation  $(f_1, s_1) \geq_{\mathcal{R}} (f_2, s_2)$  implique la relation  $1f_1 \geq_{\mathcal{R}} 1f_2$ .*

**Démonstration :** D'après l'hypothèse, il existe un élément  $(h, x)$  de  $S_f \circ M$  tel que  $(f_1, s_1)(h, x) = (f_2, s_2)$ . Pour tout  $t \in M$ , on a  $tf_1.ts_1h = tf_2$ . En prenant  $t = 1$ , on obtient  $1f_1.s_1h = 1f_2$  ce qui termine la preuve du deuxième lemme.  $\square$

Soit  $((f, s); (g_0, e_0), \dots, (g_m, e_m))$  une  $(S \circ M, Q)$ -chaîne de longueur  $m$ . D'après le premier lemme, les idempotents  $sg_i \in S_f$  forment une suite décroissante pour l'ordre  $\mathcal{R} : sg_0 \geq_{\mathcal{R}} \dots \geq_{\mathcal{R}} sg_m$ . D'après la proposition 25 (p. 76), on peut donc obtenir une  $(S, P)$ -chaîne de longueur  $m$  puisque les éléments  $1f(sg_i)^\omega$  sont alternativement dans  $P$  et hors de  $P$ . On obtient en particulier les deux inégalités

$$\begin{aligned} m^+(X\sigma^{-1}) &\leq m^+(X) \\ m^-(X\sigma^{-1}) &\leq m^-(X) \end{aligned}$$

Si une des deux inégalités précédentes est stricte, on est assuré du résultat. On suppose donc maintenant que  $m(X\sigma^{-1}) = m(X) = m$ . Soit  $(u_0, \dots, u_n)$  une  $(S \circ M, Q)$ -superchaîne de longueur  $n$  où chaque  $u_i$  est une  $(S \circ M, Q)$ -chaîne de longueur  $m$ . On applique le résultat précédent à chaque  $u_i$  pour obtenir des  $(S, P)$ -chaînes  $v_i = (1f_i; e_{i,0}, \dots, e_{i,m})$  de longueur  $m$ . D'après le deuxième lemme, les éléments  $1f_i \in S_f$  forme une suite décroissantes pour l'ordre  $\mathcal{R}$ . La suite  $(v_0, \dots, v_n)$  constitue alors une  $(S, P)$ -superchaînes de longueur  $n$  et de même nature que la  $(S \circ M, Q)$ -superchaîne de départ. On a donc obtenu les inégalités

$$\begin{aligned} n^+(X\sigma^{-1}) &\leq n^+(X) \\ n^-(X\sigma^{-1}) &\leq n^-(X) \end{aligned}$$

ce qui termine la démonstration du théorème.  $\square$

## 4.6 Application

Nous donnons ici la construction d'un automate reconnaissant un  $\omega$ -langage  $X$  à partir d'un  $\omega$ -semigroupe reconnaissant ce langage. Dans un premier temps, nous traitons le cas où le langage  $X$  est déterministe. Nous explicitons alors un automate de Büchi déterministe reconnaissant le langage  $X$ . De ce cas particulier, nous déduisons ensuite le cas général pour lequel nous donnons un automate de Muller reconnaissant  $X$ .

Pour un semigroupe  $S$  et un élément  $s$  de  $S$ , on note  $\text{stab}(s)$  le stabilisateur de  $s$ , c'est à dire le sous-semigroupe formé des éléments  $t$  de  $S$  vérifiant  $st = s$ .

### 4.6.1 Construction de l'automate

Soit  $X \subset A^\omega$  un  $\omega$ -langage déterministe reconnu par le morphisme  $\varphi$  de  $A^\infty$  dans le  $\omega$ -semigroupe fini  $S = (S_f, S_i)$ , i.e.,  $X = X\varphi\varphi^{-1}$ . On note  $P = X\varphi$  l'image de  $X$  par  $\varphi$  dans  $S$ . Puisque le  $\omega$ -langage  $X$  est déterministe, il n'existe pas de  $(S, P)$ -chaîne positive de longueur 1. D'après la proposition 25 (p. 76), il est possible de remplacer l'ordre  $\mathcal{H}$  par l'ordre  $\mathcal{R}$  dans la définition des  $(S, P)$ -chaînes. On a donc pour tous couples liés  $(s, e_1)$  et  $(s, e_2)$  de  $S_f$

$$\left. \begin{array}{l} se_1^\omega \in P \\ e_1 \geq_{\mathcal{R}} e_2 \end{array} \right\} \Rightarrow se_2^\omega \in P.$$

On adopte les notations suivantes :

- $R = \{s \in S_f \mid \exists e \in S_f \ se^\omega \in P\}$ ;
- $\psi = \varphi \diamond \varphi$  le morphisme naturel de  $A^*$  dans le produit de Schützenberger  $S_f \diamond S_f$ ;
- $T = (A^*\psi)^0$  c'est à dire le sous-semigroupe  $A^*\psi$  de  $S_f \diamond S_f$  auquel on a adjoint un zéro;
- pour  $s \in R$ ,

$$P_s = \left\{ \begin{pmatrix} s & Q \\ 0 & s \end{pmatrix} \in T \mid \exists (x, e) \in Q \begin{array}{ll} x \in \text{stab}(s) & e^2 = e \\ e \in \text{stab}(s) & se^\omega \in P \end{array} \right\}$$

#### États de l'automate

L'ensemble des états de l'automate est  $Q = T^R \times S_f^1$ . Les états sont des couples formés d'une fonction de  $R$  dans  $T$  et d'un élément de  $S_f^1$ . L'état initial est le couple  $q_0 = (f_0, 1)$  où la fonction  $f_0$  est la fonction constante égale au zéro de  $T$  et où 1 dénote l'unité de  $S_f^1$ . L'ensemble des états finaux est  $F = \bigcup_{s \in R} F_s$  où  $F_s = \{(f, s) \in Q \mid sf \in P_s\}$ .

#### Action de l'automate

Soient un état  $q = (f, s) \in Q$  et une lettre  $a \in A$ . L'action à droite de  $a$  sur  $q$  est déterminée par  $q.a = (g, s.a\varphi)$  où la fonction  $g$  est définie par

$$tg = \begin{cases} tf.a\psi & \text{si } t \neq s \\ a\psi & \text{si } sf = 0 \\ a\psi & \text{si } sf \in P_s \\ sf.a\psi & \text{sinon} \end{cases}$$

Lorsque  $sf = 0$  ou  $sf \in P_s$  ( $q$  est alors un état final), la valeur de  $sg$  ne dépend pas de  $sf$ . On dira alors que l'état  $q$  est un *état d'oubli* pour  $s$ .

Soit  $q' = (h, t)$  l'état égal à  $q.u$  où  $u$  est un mot fini. On constate que  $t = s.u\varphi$ . En particulier, si  $q' = q$ , alors  $u\varphi$  appartient au stabilisateur de  $s$ . Dans le cas où  $q$  est égal à l'état initial  $q_0$  de l'automate,  $s$  est l'unité de  $S_f$  et  $t = u\varphi$ . On peut remarquer que dans ce cas,  $sf \neq 0$  si et seulement si le mot  $u$  possède un préfixe propre  $v$  tel que  $v\varphi = s$ . En effet, pour l'état initial,  $sf_0 = 0$  et  $sf$  demeure égal à 0 tant que l'on ne passe pas dans un état d'oubli pour  $s$  et reste ensuite non nul.

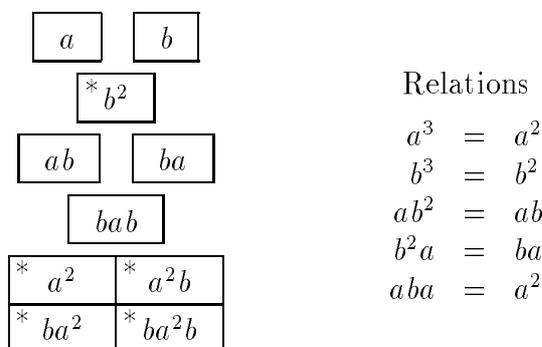


Figure 4.11 : Structure en  $\mathcal{D}$ -classes de  $T$ .

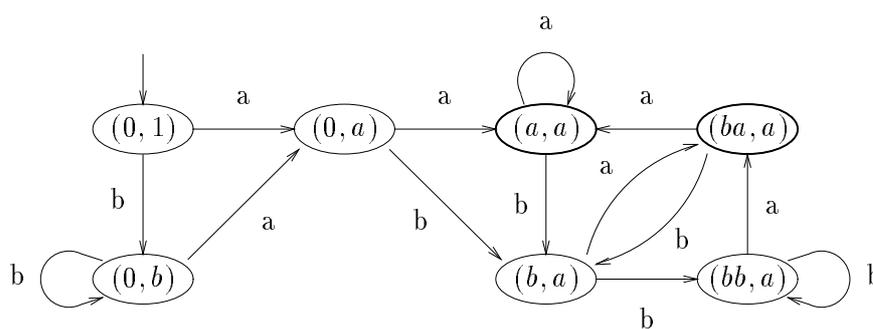


Figure 4.12 : Automate  $\mathcal{A}_6$

**Exemple :** On considère le langage déterministe  $X_1 = (b^*a)^\omega$  des mots ayant un nombre infini d'occurrences de la lettre  $a$  sur l'alphabet  $A = a + b$ . Ce  $\omega$ -langage est reconnu par le  $\omega$ -semigroupe  $S_1 = (\{b = 1, a = 0\}, \{b^\omega, ab^\omega, a^\omega\})$ . L'image  $P$  de  $X_1$  dans le  $\omega$ -semigroupe  $S_1$  est égale à  $\{a^\omega\}$ . la partie  $R$  est donc réduite au

singleton  $\{a\}$ . Le semigroupe  $T$  admet la présentation

$$\langle a, b; a^3 = a^2, b^3 = b^2, ab^2 = ab, b^2a = ba, aba = a^2 \rangle .$$

La structure en  $\mathcal{D}$ -classes de ce semigroupe est donnée à la figure 4.11. On obtient l'automate de la figure 4.12. Le couple  $(f, s)$  où  $f$  est une fonction de  $R = \{a\}$  dans  $T$  est noté par  $(af, s)$ .

### 4.6.2 Preuve de l'automate

Montrons que l'automate ainsi construit reconnaît effectivement le  $\omega$ -langage  $X$ . Soit  $u \in A^\omega$  un mot accepté par l'automate et  $q = (f, s) \in F_s$  un état infiniment répété par le mot  $u$ . Le mot  $u$  se factorise  $u = u_0u_1u_2\dots$  où  $q_0.u_0 = q$  et  $q.u_i = q$  pour  $i \geq 1$ .

**Lemme 18** *Soit  $q$  un état appartenant à  $F_s$ . Il existe un idempotent  $e \in \text{stab}(s)$  tel que  $se^\omega \in P$  et tel que tout mot  $u$  vérifiant  $q.u = q$  se factorise  $u = vw$  avec  $v\varphi \in \text{stab}(s)$  et  $w\varphi = e$ .*

**Démonstration :** Puisque l'état  $q = (f, s)$  appartient à  $F_s$ , on a  $sf \in P_s$ . Il existe donc  $x, e$  associé à  $sf$  tels que  $x \in \text{stab}(s)$ ,  $e^2 = e$ ,  $e \in \text{stab}(s)$ ,  $se^\omega \in P$ . Parmi les états visités par le chemin  $q \xrightarrow{u} q$ , on considère le dernier état  $q'$  de  $F_s$  apparaissant avant l'état  $q$  de l'extrémité. Un tel état existe toujours, puisque l'origine  $q$  du chemin appartient à  $F_s$ . Le chemin se décompose  $q \xrightarrow{u_1} q' \xrightarrow{u_3} q$  avec  $u_1\varphi \in \text{stab}(s)$ . Puisque  $q'$  est le dernier de  $F_s$  apparaissant dans le chemin, il n'existe pas d'état d'oubli entre  $q'$  et  $q$  et donc  $u_3\psi = sf$ . Le mot  $u_3$  se décompose alors  $u_3 = u_2w$  avec  $u_2\varphi = x \in \text{stab}(s)$  et  $w\varphi = e$ . Le mot  $u$  se factorise maintenant  $u = vw$  avec  $v = u_1u_2$ . On a  $v\varphi \in \text{stab}(s)$  puisque  $u_1\varphi, u_2\varphi \in \text{stab}(s)$ .  $\square$

En appliquant le lemme, chaque mots  $u_i$  pour  $i \geq 1$  se factorise  $u_i = v_iw_i$  avec  $v_i\varphi \in \text{stab}(s)$  et  $w_i\varphi = e$ . On pose  $y_0 = u_0v_1$  de sorte que  $y_0\varphi = s$  car  $u_0\varphi = s$  et  $v_1\varphi \in \text{stab}(s)$ . On pose également  $y_i = w_i v_{i+1}$  pour  $i \geq 1$  de sorte que  $y_i\varphi \in \text{stab}(s)$  et  $y_i\varphi \in eS_f^1$  pour  $i \geq 1$ . De la factorisation  $u = y_0y_1y_2\dots$ , on déduit que  $u\varphi = se'^\omega$  où  $e \geq_{\mathcal{R}} e'$ . Le mot  $u$  appartient donc à  $X$ .

Réciproquement, soit  $u \in A^\omega$  un mot de  $X$ . L'image du mot  $u$  par  $\varphi$  est  $u\varphi = se^\omega \in P$ . Le mot  $u$  se factorise  $u = u_0u_1u_2\dots$  où  $u_0\varphi = s$  et  $u_i\varphi = e$  pour  $i \geq 1$ . On pose  $q_{i+1} = q_i.u_i$ .

Puisque  $u_0\varphi = s$ , le mot  $u_0$  possède un plus petit préfixe  $v$  tel que  $v\varphi = s$ . L'état  $q = q_0.v$  est un état d'oubli pour  $s$ . Le chemin  $q_0 \xrightarrow{u_0} q_1$  passe donc par un état d'oubli.

Montrons maintenant que si le chemin  $q_i \xrightarrow{u_i} q_{i+1}$  visite un état d'oubli pour  $s$ , alors le chemin  $q_{i+1} \xrightarrow{u_{i+1}} q_{i+2}$  visite un état de  $F_s$ . Soit  $q'$  le dernier état d'oubli pour  $s$  du chemin  $q_i \xrightarrow{u_i} q_{i+1}$ . Si le chemin  $q_{i+1} \xrightarrow{u_{i+1}} q_{i+2}$  ne visite pas d'état de  $F_s$  avant  $q_{i+2}$ , le chemin  $q' \xrightarrow{v} q_{i+1} \xrightarrow{u_{i+1}} q_{i+2}$  ne contient pas d'état d'oubli pour  $s$  exceptées les extrémités. Le mot  $v$  éventuellement vide dans le cas où  $q' = q_{i+1}$  vérifie  $v\varphi \in \text{stab}(s)$  car  $q'$  est un état d'oubli pour  $s$  et  $u_0 \dots u_{i+1}\varphi = s$ . Puisque  $u_{i+1}\varphi = e$ , l'état  $q_{i+2}$  appartient à  $F_s$ .

Puisque le chemin  $q_0 \xrightarrow{u_0} q_1$  passe par un état d'oubli, chacun des chemins  $q_i \xrightarrow{u_i} q_{i+1}$  visite un état de  $F_s$  et le mot  $u$  est reconnu par l'automate.

### 4.6.3 Cas général

A partir d'un  $\omega$ -semigroupe reconnaissant un  $\omega$ -langage, il est aisé de décomposer ce langage comme combinaison booléenne de langages déterministes. Lors de la démonstration de la proposition 36 (p. 88), on montre que tout  $\omega$ -langage peut s'écrire comme une chaîne de  $\omega$ -langage  $G_\delta$ . On applique la construction précédente pour chacun de ces langages déterministes. Chaque automate de Büchi obtenu est transformé en un automate de Muller. On applique alors les opérations booléennes correspondantes sur les automates pour obtenir un automate de Muller reconnaissant le  $\omega$ -langage considéré.



## Chapitre 5

# Automates de Rabin à chaînes

### 5.1 Introduction

Nous définissons, dans ce chapitre, les automates de Rabin à chaîne qui sont des automates de Rabin pour lesquels les parties reconnaissantes sont emboîtées. Nous montrerons qu'à partir de deux automates de Rabin à chaîne reconnaissant deux  $\omega$ -langages  $X$  et  $Y$ , il est possible de construire des automates de Rabin à chaîne pour le complément  $X^c$ , l'intersection  $X \cap Y$  et l'union  $X + Y$ . Nous donnons ensuite une nouvelle caractérisation de l'indice de Rabin d'un  $\omega$ -langage.

### 5.2 Notations et définitions

Un *automate de Rabin à chaîne* (ou plus simplement un *automate à chaîne*) est un automate de Rabin dont les parties  $R_i$  de  $\mathcal{R}$  vérifient en outre la condition

$$R_1 \supset R_2 \supset R_3 \supset \cdots \supset R_m$$

Dans le cas d'un automate à chaîne, la somme de différences définissant les chemins réussis est une chaîne

$$|\mathcal{A}| = |R_1| - |R_2| + |R_3| - \cdots \pm |R_m|$$

Il est de même de la somme de différences définissant le  $\omega$ -langage reconnu

$$L^\omega(\mathcal{A}) = L^\omega(R_1) - L^\omega(R_2) + L^\omega(R_3) - \cdots \pm L^\omega(R_m)$$

### 5.3 Constructions

A partir d'automates à chaîne reconnaissant des  $\omega$ -langages  $X$  et  $Y$ , nous allons construire des automates à chaîne reconnaissant les  $\omega$ -langages  $X^c$ ,  $XY$  et  $X+Y$ .

**Lemme 19** *Toute  $\omega$ -langage reconnu par un automate de Rabin (respectivement à chaîne) est reconnu par un automate de Rabin (respectivement à chaîne) complet de même indice de Rabin.*

**Démonstration :** Soit  $\mathcal{A} = (Q, A, E, q_0, \mathcal{R})$  un automate de Rabin. Si  $\mathcal{A}$  n'est pas complet, il suffit de poser  $\mathcal{A}' = (Q', A, E', q_0, \mathcal{R})$  où  $Q'$  et  $E'$  sont définis par  $Q' = Q + \{\square\}$  et  $E' = E + \{(q, a, \square) \mid \forall q' \in Q (q, a, q') \notin E\} + \{(\square, a, \square) \mid a \in A\}$  pour obtenir un automate  $\mathcal{A}'$  équivalent à  $\mathcal{A}$  et complet. Si de plus l'automate  $\mathcal{A}$  est un automate à chaîne, l'automate construit  $\mathcal{A}'$  est aussi un automate à chaîne.  $\square$

**Proposition 37** *Soit  $\mathcal{A} = (Q, A, E, q_0, \mathcal{R})$  avec  $\mathcal{R} = (R_1, \dots, R_m)$  un automate à chaîne complet reconnaissant un  $\omega$ -langage  $X$ . Le nouvel automate  $\mathcal{A}' = (Q, A, E, q_0, \mathcal{R}')$  où  $\mathcal{R}'$  est définie par  $\mathcal{R}' = (R_0, R_1, \dots, R_m)$  avec la convention  $R_0 = Q$  reconnaît le complémentaire  $X^c$  de  $X$ .*

**Démonstration :** Les  $\omega$ -langages reconnus par les automates  $\mathcal{A}$  et  $\mathcal{A}'$  sont respectivement égaux aux chaînes

$$\begin{aligned} L^\omega(\mathcal{A}) &= L^\omega(R_1) - L^\omega(R_2) + \dots \pm L^\omega(R_m) \\ L^\omega(\mathcal{A}') &= L^\omega(R_0) - L^\omega(R_1) + L^\omega(R_2) - \dots \mp L^\omega(R_m) \end{aligned}$$

Puisque l'automate  $\mathcal{A}$  est complet,  $L^\omega(R_0) = A^\omega$  et les  $\omega$ -langages reconnus par  $\mathcal{A}$  et  $\mathcal{A}'$  sont complémentaires.  $\square$

Étant donnés deux automates à chaîne,  $\mathcal{A} = (P, A, E, p_0, \mathcal{S})$  d'une part et  $\mathcal{B} = (Q, A, F, q_0, \mathcal{T})$  d'autre part où  $\mathcal{S} = (S_1, S_2, \dots, S_m)$  et  $\mathcal{T} = (T_1, T_2, \dots, T_n)$  reconnaissant respectivement les  $\omega$ -langages  $X$  et  $Y$ , nous construisons deux automates  $\mathcal{C}_1 = (R, A, G, r_0, \mathcal{U})$  et  $\mathcal{C}_2 = (R, A, G, r_0, \mathcal{V})$  reconnaissant respectivement les  $\omega$ -langages  $XY$  et  $X+Y$ . Nous commençons par décrire la partie commune : l'automate  $\mathcal{C} = (R, A, G, r_0)$ . Nous posons  $\mathbb{B} = \{0, 1\}$  muni des opérations booléennes habituelles et nous identifions une proposition avec sa valeur dans  $\mathbb{B}$ .

L'ensemble  $R$  des états est égal à  $P \times Q \times \mathcal{M}_{m,n}(\mathbb{B} \times \mathbb{B})$ . Chaque état de  $R$  est un triplet formé d'un état de  $\mathcal{A}$ , d'un état de  $\mathcal{B}$  et d'une matrice de dimensions  $m$  et  $n$  de couples de booléens que nous écrirons  $(p, q, M)$  ou  $(p, q, e_{ij}, f_{ij})$  pour

donner de façon explicite la matrice  $M$  par  $M = ((e_{ij}, f_{ij}))$  pour  $1 \leq i \leq m$  et  $1 \leq j \leq n$ .

L'état initial  $r_0$  de l'automate  $\mathcal{C}$  est défini par  $r_0 = (p_0, q_0, e_{ij}, f_{ij})$  où les valeurs booléennes sont données par  $(e_{ij}, f_{ij}) = (p_0 \in S_i, q_0 \in T_j)$ .

L'ensemble  $G$  des transitions de l'automate  $\mathcal{C}$  est l'ensemble des transitions  $(p, q, e_{ij}, f_{ij}) \xrightarrow{a} (p', q', e'_{ij}, f'_{ij})$  pour lesquels  $p \xrightarrow{a} p'$  et  $q \xrightarrow{a} q'$  sont respectivement des transitions de  $\mathcal{A}$  et  $\mathcal{B}$  et où les  $e'_{ij}, f'_{ij}$  sont définis par

$$\begin{aligned}
 e'_{ij} &= \begin{cases} p' \in S_i & \text{si } e_{ij} \wedge f_{ij} = 1 \\ p' \in S_i \vee e_{ij} & \text{sinon} \end{cases} \\
 \text{et} & \\
 f'_{ij} &= \begin{cases} q' \in T_j & \text{si } e_{ij} \wedge f_{ij} = 1 \\ q' \in T_j \vee f_{ij} & \text{sinon} \end{cases}
 \end{aligned} \tag{11}$$

Il est possible de restreindre l'automate ainsi construit aux états accessibles à partir de l'état initial  $r_0$  par une suite de transitions de  $G$ .

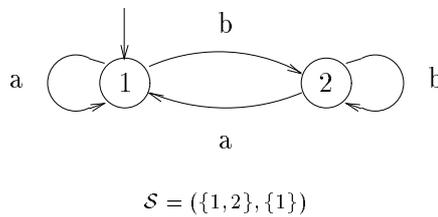


Figure 5.1 : Automate  $\mathcal{A}$

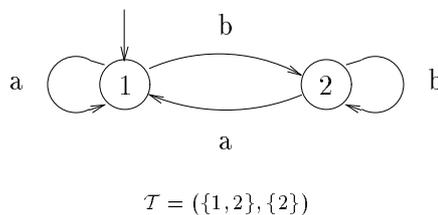
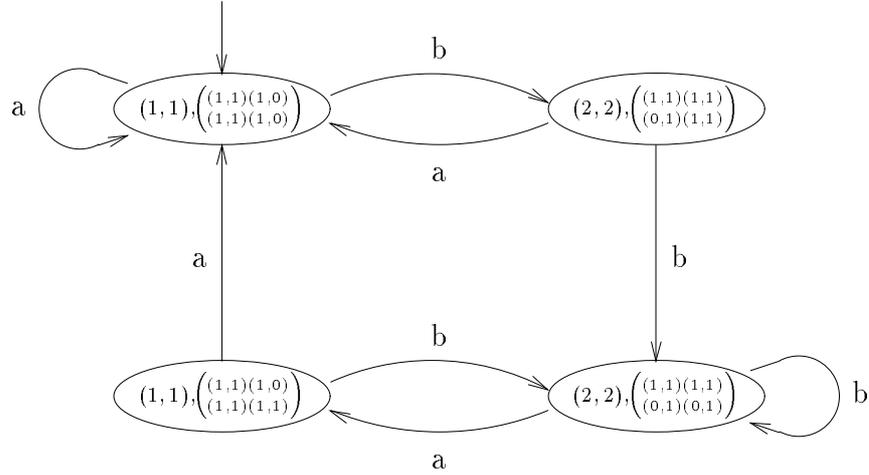


Figure 5.2 : Automate  $\mathcal{B}$

**Exemple :** Nous illustrons cette construction par un exemple sur l'alphabet  $A = a + b$ . Nous considérons les automates  $\mathcal{A} = (\{1, 2\}, A, E, 1, S)$  et  $\mathcal{B} =$

Figure 5.3 : Automate  $\mathcal{C}$ 

$(\{1, 2\}, A, E, 1, T)$  avec  $\mathcal{S} = (S_1 = \{1, 2\}, S_2 = \{1\})$  et  $\mathcal{T} = (T_1 = \{1, 2\}, T_2 = \{2\})$  (cf. figures 5.1 et 5.2). Les indices de Rabin  $m$  et  $n$  des automates  $\mathcal{A}$  et  $\mathcal{B}$  sont respectivement  $m = 2$  et  $n = 2$ . Les états de l'automate construit  $\mathcal{C}$  sont donc des triplets formés d'un état de  $\mathcal{A}$  (1 ou 2), d'un état de  $\mathcal{B}$  (1 ou 2) et d'une matrice  $2 \times 2$  de couples de booléens (cf. figure 5.3).

Avant de décrire les suites  $\mathcal{U}$  et  $\mathcal{V}$  d'états, nous énonçons quelques propriétés de l'automate  $\mathcal{C}$  ainsi construit.

**Lemme 20** Soient  $(p, q, M)$  un état de l'automate  $\mathcal{C}$ ,  $p \xrightarrow{a} p'$  une transition de  $\mathcal{A}$  et  $q \xrightarrow{a} q'$  une transition de  $\mathcal{B}$ . Il existe un unique état  $(p', q', M')$  de  $\mathcal{C}$  tel que  $(p, q, M) \xrightarrow{a} (p', q', M')$  soit une transition de  $\mathcal{C}$ .

**Démonstration :** Pour  $p, q, p', q'$  et  $M$  fixés il existe une unique matrice  $M'$  calculée suivant les formules (11) telle que  $(p, q, M) \xrightarrow{a} (p', q', M')$  soit une transition de  $\mathcal{C}$ .  $\square$

**Lemme 21** Si les deux automates  $\mathcal{A}$  et  $\mathcal{B}$  sont complets, alors l'automate  $\mathcal{C}$  est complet.

**Démonstration :** Pour un état  $(p, q, M)$  de  $\mathcal{C}$  et une lettre  $a$  de  $A$ , il existe  $p \xrightarrow{a} p'$  une transition de  $\mathcal{A}$  et  $q \xrightarrow{a} q'$  une transition de  $\mathcal{B}$  puisque  $\mathcal{A}$  et  $\mathcal{B}$  sont complets. D'après le lemme précédent, il existe une transition  $(p, q, M) \xrightarrow{a} (p', q', M')$  de  $\mathcal{C}$ .  $\square$

**Lemme 22** Si  $(p, q, M) \xrightarrow{u} (p', q', M')$  est un chemin de  $\mathcal{C}$  d'étiquette  $u \in A^*$  alors  $p \xrightarrow{u} p'$  et  $q \xrightarrow{u} q'$  sont respectivement des chemins de  $\mathcal{A}$  et de  $\mathcal{B}$ .

**Démonstration :** La démonstration est immédiate en raisonnant par récurrence sur la longueur du mot  $u$  puisque par construction de l'automate, l'existence d'une transition  $(p, q, M) \xrightarrow{a} (p', q', M')$  dans  $\mathcal{C}$  implique l'existence des transitions  $p \xrightarrow{a} p'$  et  $q \xrightarrow{a} q'$  dans  $\mathcal{A}$  et  $\mathcal{B}$ .  $\square$

Nous définissons les ensembles d'états

$$\begin{aligned} H_{00} &= R \\ H_{k0} &= \{(p, q, e_{ij}, f_{ij}) \in R \mid p \in S_k\} && \text{pour } 1 \leq k \leq m \\ H_{0l} &= \{(p, q, e_{ij}, f_{ij}) \in R \mid q \in T_l\} && \text{pour } 1 \leq l \leq n \\ H_{kl} &= \{(p, q, e_{ij}, f_{ij}) \in R \mid e_{kl} \wedge f_{kl} = 1\} && \text{pour } 1 \leq k \leq m \text{ et } 1 \leq l \leq n \end{aligned}$$

**Exemple :** Pour l'exemple de la figure 5.3, nous résumons par un tableau l'appartenance des états de l'automate  $\mathcal{C}$  aux différents ensembles  $H_{ij}$  définis.

|   | $H_{00}$ | $H_{10}$ | $H_{01}$ | $H_{20}$ | $H_{02}$ | $H_{11}$ | $H_{21}$ | $H_{12}$ | $H_{22}$ |
|---|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| $(1, 1), \left( \begin{smallmatrix} (1,1)(1,0) \\ (1,1)(1,0) \end{smallmatrix} \right)$ | •        | •        | •        | •        |          | •        | •        |          |          |
| $(1, 1), \left( \begin{smallmatrix} (1,1)(1,0) \\ (1,1)(1,1) \end{smallmatrix} \right)$ | •        | •        | •        | •        |          | •        | •        |          | •        |
| $(2, 2), \left( \begin{smallmatrix} (1,1)(1,1) \\ (0,1)(0,1) \end{smallmatrix} \right)$ | •        | •        | •        |          | •        | •        |          | •        |          |
| $(2, 2), \left( \begin{smallmatrix} (1,1)(1,1) \\ (0,1)(1,1) \end{smallmatrix} \right)$ | •        | •        | •        |          | •        | •        |          | •        | •        |

**Lemme 23** Soit  $(p, q, M) \xrightarrow{u} (p', q', M')$  un chemin de  $\mathcal{C}$  d'étiquette  $u \in A^*$ . Si les chemins  $p \xrightarrow{u} p'$  de  $\mathcal{A}$  et  $q \xrightarrow{u} q'$  de  $\mathcal{B}$  passent respectivement par un état de  $S_k$  pour  $0 \leq k \leq m$  et un état de  $T_l$  pour  $0 \leq l \leq n$ , alors le chemin  $(p, q, M) \xrightarrow{u} (p', q', M')$  passe par un état de  $H_{kl}$ .

**Démonstration :** Nous écrivons le chemin  $(p, q, M) \xrightarrow{u} (p', q', M')$  sous la forme

$$(p_0, q_0, M_0) \xrightarrow{u_1} (p_1, q_1, M_1) \xrightarrow{u_2} \dots \xrightarrow{u_{r-1}} (p_{r-1}, q_{r-1}, M_{r-1}) \xrightarrow{u_r} (p_r, q_r, M_r)$$

où  $u_1 \dots u_r = u$ ,  $(p_0, q_0, M_0) = (p, q, M)$  et  $(p_r, q_r, M_r) = (p', q', M')$ . Les cas  $m = 0$  ou  $n = 0$  sont immédiats. Nous supposons  $m \geq 1$  et  $n \geq 1$  et nous noterons par  $((e_{ij}^m, f_{ij}^m))$ , pour  $1 \leq m \leq n$ , les éléments de la matrice  $M_m$  de l'état  $(p_m, q_m, M_m)$ . Par hypothèse, il existe deux entiers  $m_2$  et  $m_3$  tels que

$p_{m_2} \in S_k$  et  $q_{m_3} \in T_l$ . Nous avons alors  $e_{kl}^{m_2} = 1$  et  $f_{kl}^{m_3} = 1$  et nous appelons  $m_0$  le plus petit entier  $m \geq 1$  tel que  $e_{kl}^m = 1$  et  $m_1$  le plus petit entier  $m \geq 1$  tel que  $f_{kl}^m = 1$ . Si  $m_0 = m_1$ , l'état  $(p_{m_1}, q_{m_1}, M_{m_1})$  appartient à  $H_{kl}$ . Sinon, nous pouvons supposer par symétrie que  $m_0 < m_1$ . Par définition de  $m_1$ , nous avons  $f_{kl}^m = 0$  pour  $m_0 \leq m < m_1$ . La règle de calcul des  $e_{kl}^m$  et  $e_{kl}^{m_0} = 1$  impliquent  $e_{kl}^m = 1$  pour  $m_0 \leq m \leq m_1$  et par conséquent  $(p_{m_1}, q_{m_1}, M_{m_1}) \in H_{kl}$ .  $\square$

**Lemme 24** Soit  $(p, q, M) \xrightarrow{u} (p', q', M')$  un chemin de  $\mathcal{C}$  d'étiquette  $u \in A^*$ . Si les états  $(p, q, M)$  et  $(p', q', M')$  appartiennent à  $H_{kl}$ , alors les chemins  $p \xrightarrow{u} p'$  de  $\mathcal{A}$  et  $q \xrightarrow{u} q'$  de  $\mathcal{B}$  passent respectivement par un état de  $S_k$  et un état de  $T_l$ .

**Démonstration :** Les cas  $m = 0$  ou  $n = 0$  sont à nouveau immédiats. Nous reprenons les notations du lemme précédent et nous appelons  $m_0$  le plus petit entier  $m \geq 1$  tel que  $e_{kl}^m = 1$  et  $m_1$  le plus petit entier  $m \geq 1$  tel que  $f_{kl}^m = 1$ . Les entiers  $m_0$  et  $m_1$  existent puisque  $e_{kl}^r = 1$  et  $f_{kl}^r = 1$ . Si  $m_0 = 1$ , le calcul de  $e_{kl}^1$  implique que  $p_1 \in S_k$  puisque  $e_{kl} = 1$  et  $f_{kl} = 1$ . Si  $m_0 > 1$ , nous avons  $e_{kl}^{m_0-1} = 0$  et  $e_{kl}^{m_0} = 1$  par définition de  $m_0$ . Le calcul de  $e_{kl}^{m_0}$  implique à nouveau que  $p_{m_0} \in S_k$ . Par un raisonnement similaire, nous obtenons  $q_{m_1} \in T_l$ .  $\square$

Nous sommes maintenant en mesure d'énoncer la propriété fondamentale de l'automate  $\mathcal{C}$  construit.

**Lemme 25** Un chemin infini  $c$  de l'automate  $\mathcal{C}$  passe infiniment souvent par  $H_{kl}$  si et seulement si les chemins dans les automates  $\mathcal{A}$  et  $\mathcal{B}$  de même étiquette que  $c$  passent infiniment souvent par  $S_k$  dans  $\mathcal{A}$  et infiniment souvent par  $T_l$  dans  $\mathcal{B}$ . Avec les notations adoptées :

$$L^\omega(H_{kl}) = L^\omega(S_k)L^\omega(T_l).$$

où  $L^\omega(S_k)$ ,  $L^\omega(T_l)$  et  $L^\omega(H_{kl})$  sont respectivement les ensembles des étiquettes des chemins passant infiniment souvent par  $S_k$ ,  $T_l$  et  $H_{kl}$  dans les automates  $\mathcal{A}$ ,  $\mathcal{B}$  et  $\mathcal{C}$ .

**Démonstration :** Tout chemin infini  $c$  de l'automate  $\mathcal{C}$  passant infiniment souvent par  $H_{kl}$  se décompose en une suite de chemins  $(p, q, M) \xrightarrow{u} (p', q', M')$  où les états  $(p, q, M)$  et  $(p', q', M')$  appartiennent à  $H_{kl}$ . Par application du lemme 24 les chemins de même étiquette que  $c$  passe infiniment souvent par  $S_k$  dans  $\mathcal{A}$  et par  $T_l$  dans  $\mathcal{B}$ .

Réciproquement, si les chemins de même étiquette que  $c$  passe infiniment souvent par  $S_k$  dans  $\mathcal{A}$  et par  $T_l$  dans  $\mathcal{B}$ , le chemin  $c$  se décompose en une suite de chemins  $(p, q, M) \xrightarrow{u} (p', q', M')$  où les chemins  $p \xrightarrow{u} p'$  de  $\mathcal{A}$  et  $q \xrightarrow{u} q'$  de  $\mathcal{B}$

passent respectivement par un état de  $S_k$  et un état de  $T_l$ . Par application du lemme 23, nous obtenons le résultat.  $\square$

Nous introduisons les ensembles d'états  $K_{kl}$  définis par

$$K_{kl} = \bigcup_{k \leq i \text{ et } l \leq j} H_{ij} \quad \text{pour } 0 \leq k \leq m \text{ et } 0 \leq l \leq n$$

**Lemme 26** *Pour  $R$  et  $R'$  deux ensembles d'états d'un automate  $(Q, A, E, I)$ , on a les égalités*

$$|R + R'| = |R| + |R'| \quad \text{et} \quad L^\omega(R + R') = L^\omega(R) + L^\omega(R')$$

**Démonstration :** Puisque le nombre d'états de l'automate est fini, tout chemin passant infiniment souvent par  $R + R'$  passe infiniment souvent par  $R$  ou par  $R'$ .  $\square$

Il est aisé de calculer les étiquettes des chemins passant infiniment souvent par  $K_{kl}$  pour  $0 \leq k \leq m$  et  $0 \leq l \leq n$

$$L^\omega(K_{kl}) = \bigcup_{k \leq i \text{ et } l \leq j} L^\omega(S_i)L^\omega(T_j) = L^\omega(S_k)L^\omega(T_l)$$

Les ensembles d'états vérifient en outre la propriété

$$k \leq k' \text{ et } l \leq l' \quad \Rightarrow \quad K_{k'l'} \subset K_{kl}$$

Nous terminons maintenant la construction des automates  $\mathcal{C}_1$  et  $\mathcal{C}_2$  en donnant explicitement les suites d'états  $\mathcal{U}$  et  $\mathcal{V}$ .

$$\begin{aligned} U_{2i} &= \sum_{k+l=2i+1} K_{kl} \\ U_{2i+1} &= \sum_{\substack{k+l=2i+2 \\ k \text{ impair}}} K_{kl} \\ V_{2i} &= \sum_{\substack{k+l=2i \\ k \text{ pair}}} K_{kl} \\ V_{2i+1} &= \sum_{k+l=2i+1} K_{kl} \end{aligned}$$

Les langages  $X$  et  $Y$  sont égaux aux chaînes

$$\begin{aligned} X &= L^\omega(S_1) - L^\omega(S_2) + \cdots \pm L^\omega(S_m) \\ Y &= L^\omega(T_1) - L^\omega(T_2) + \cdots \pm L^\omega(S_n) \end{aligned}$$

et le langage  $XY$  est égal à la chaîne

$$XY = Z_1 - Z_2 + Z_3 - \cdots \pm Z_{m+n-1}$$

où les  $Z_i$  sont définis par (cf. formules (8) page 58)

$$\begin{aligned} Z_{2i} &= \sum_{k+l=2i+1} L^\omega(S_k)L^\omega(T_l) \\ Z_{2i+1} &= \sum_{\substack{k+l=2i+2 \\ k \text{ impair}}} L^\omega(S_k)L^\omega(T_l) \end{aligned}$$

Il est alors aisé de vérifier que  $\forall k \quad L^\omega(U_k) = Z_k$ . Nous avons pour  $k$  pair

$$\begin{aligned} L^\omega(U_{2i}) &= L^\omega\left(\sum_{k+l=2i+1} K_{kl}\right) \\ &= \sum_{k+l=2i+1} L^\omega(K_{kl}) \\ &= \sum_{k+l=2i+1} L^\omega(S_k)L^\omega(T_l) \\ &= Z_{2i} \end{aligned}$$

Le cas  $k$  impair est similaire. Le langage reconnu par l'automate  $\mathcal{C}_1$  est égal à la chaîne

$$\begin{aligned} L^\omega(\mathcal{C}_1) &= L^\omega(U_1) - L^\omega(U_2) + \cdots \pm L^\omega(U_{m+n-1}) \\ &= Z_1 - Z_2 + \cdots \pm Z_{m+n-1} \\ &= XY \end{aligned}$$

Nous montrerions de la même façon que

$$L^\omega(\mathcal{C}_2) = X + Y$$

**Proposition 38** *Tout automate de Rabin est équivalent à un automate à chaîne de même indice de Rabin.*

**Démonstration :** Le résultat est acquis si l'indice de Rabin  $m$  est égal à 1. Si  $m = 2$ , l'automate  $(Q, A, E, q_0, \{R_1, R_2\})$  est équivalent à l'automate à chaîne  $(Q, A, E, q_0, \{R_1 + R_2, R_2\})$ . On raisonne ensuite par récurrence.  $\square$

## 5.4 Réduction

Soit  $\mathcal{A} = (Q, A, E, q_0, \mathcal{R})$  un automate à chaîne reconnaissant un  $\omega$ -langage  $X$ . Nous allons montrer qu'il existe une autre suite décroissante d'ensembles d'états  $\mathcal{R}'$  de longueur égale à l'indice de Rabin  $\text{ind}(X)$  du  $\omega$ -langage  $X$  telle que l'automate  $\mathcal{A}' = (Q, A, E, q_0, \mathcal{R}')$  reconnaisse également le  $\omega$ -langage  $X$ . La suite  $\mathcal{R}'$  est donc minimale.

Soit  $\mathcal{R}$  une suite  $R_1 \supset R_2 \supset \dots \supset R_m$  décroissante de parties d'un ensemble  $Q$ . On adopte la convention de poser  $R_0 = Q$ . On définit la suite  $\mathcal{D}(\mathcal{R})$  des différences par

$$\begin{aligned} D_0 &= R_0 - R_1 \\ D_1 &= R_1 - R_2 \\ D_2 &= R_2 - R_3 \\ &\vdots \\ D_{m-1} &= R_{m-1} - R_m \\ D_m &= R_m \end{aligned}$$

La donnée d'une suite  $\mathcal{R}$  décroissante est équivalente à la donnée d'une suite  $\mathcal{D}$  d'ensembles disjoints. On ne fera plus dorénavant la distinction entre la suite  $\mathcal{R}$  et la suite  $\mathcal{D}$  associée. On définit le *poids* d'une suite  $\mathcal{R}$  décroissante par

$$\text{poids}(\mathcal{R}) = m + \text{card}(D_1) + 2 \text{card}(D_2) + 3 \text{card}(D_3) + \dots + m \text{card}(D_m)$$

Soit  $\mathcal{A} = (Q, A, E, q_0, \mathcal{R})$  un automate à chaîne reconnaissant un  $\omega$ -langage  $X$ . Pour une partie  $R$  de  $Q$  on définit l'entier  $\text{Iac}(\mathcal{R}, R)$  et l'ensemble  $\text{Eac}(\mathcal{R}, R)$  par

$$\begin{aligned} \text{Iac}(\mathcal{R}, R) &= \max\{k \mid R \cap R_k \neq \emptyset\} \\ &= \max\{k \mid R \cap D_k \neq \emptyset\} \\ \text{Eac}(\mathcal{R}, R) &= R \cap R_{\text{Iac}(\mathcal{R}, R)}. \end{aligned}$$

Par abus de notation on omet la référence à la suite  $\mathcal{R}$  lorsqu'il n'y a pas d'ambiguïté. Pour un état  $q$ , on écrit  $\text{Iac}(q)$  pour  $\text{Iac}(\{q\})$ . Pour une partie répétable  $R$ , l'indice  $\text{Iac}(R)$  détermine si cette partie est acceptante ou non acceptante. En effet, un chemin infini  $c$  de l'automate d'origine  $q_0$  tel que  $\text{Inr}(c) = R$  est réussi si et seulement si l'indice  $\text{Iac}(R)$  est impair puisque par définition, l'ensemble des chemin réussis est défini par la chaîne

$$|\mathcal{A}| = |R_1| - |R_2| + |R_3| - \dots \pm |R_m|.$$

L'automate à chaîne  $\mathcal{A} = (Q, A, E, q_0, \mathcal{R})$  est d'ailleurs équivalent à l'automate de Muller  $\mathcal{B} = (Q, A, E, q_0, \mathcal{T})$  où la table  $\mathcal{T}$  est définie par

$$\mathcal{T} = \{R \subset Q \mid \text{Iac}(R) \text{ est impair}\}$$

Nous allons maintenant décrire plusieurs transformations qui permettent de diminuer le poids de la suite  $\mathcal{R}$  sans modifier le  $\omega$ -langage reconnu par l'automate à chaîne. Deux suites  $\mathcal{R}$  et  $\mathcal{R}'$  sont dites équivalentes si pour toute partie répétable  $R$  de l'automate, on a  $\text{Iac}(\mathcal{R}, R) \equiv \text{Iac}(\mathcal{R}', R) \pmod{2}$ .

### Transformation 1

On suppose que  $R_m = \emptyset$ . La suite  $\mathcal{R}' = (R_0, \dots, R_{m-1})$  vérifie

$$\text{poids}(\mathcal{R}') = \text{poids}(\mathcal{R}) - 1$$

Les suites  $\mathcal{R}$  et  $\mathcal{R}'$  sont manifestement équivalentes.

### Transformation 2

On suppose qu'il existe un état  $q \in D_k$  avec  $k \geq 1$  tel que qu'il n'existe pas de partie répétable  $R$  vérifiant  $q \in \text{Eac}(R)$ . La suite  $\mathcal{R}'$  définie par  $D'_0 = D_0 + \{q\}$ ,  $D'_k = D_k - \{q\}$  et  $D'_i = D_i$  pour tout  $i$  différent de 0 et de  $k$  vérifie

$$\text{poids}(\mathcal{R}') = \text{poids}(\mathcal{R}) - k$$

De plus, pour toute partie répétable  $R$ , on a  $\text{Iac}(\mathcal{R}', R) = \text{Iac}(\mathcal{R}, R)$  et les suites  $\mathcal{R}$  et  $\mathcal{R}'$  sont équivalentes.

### Transformation 3

On suppose qu'il existe un état  $q \in D_k$  avec  $k \geq 2$  tel que pour toute partie répétable  $R$ , on ait l'implication

$$q \in \text{Eac}(R) \quad \Rightarrow \quad R \cap D_{k-1} = \emptyset$$

La suite  $\mathcal{R}'$  définie par  $D'_{k-2} = D_{k-2} + \{q\}$ ,  $D'_k = D_k - \{q\}$  et  $D'_i = D_i$  pour tout  $i$  différent de  $k-2$  et de  $k$  vérifie

$$\text{poids}(\mathcal{R}') = \text{poids}(\mathcal{R}) - 2$$

Si la partie répétable  $R$  vérifie  $\text{Eac}(R) = \{q\}$ , on a  $\text{Iac}(\mathcal{R}', R) = \text{Iac}(\mathcal{R}, R) - 2$  et on a  $\text{Iac}(\mathcal{R}', R) = \text{Iac}(\mathcal{R}, R)$  dans le cas contraire. Les suites  $\mathcal{R}$  et  $\mathcal{R}'$  sont donc équivalentes.

Soit  $\mathcal{A} = (Q, A, E, q_0, \mathcal{R})$  un automate à chaîne où la suite  $\mathcal{R}$  est de poids minimal. Une telle suite peut être obtenue en appliquant les transformations 1, 2 et 3 tant que cela est possible. Le processus doit nécessairement converger puisque le poids est un entier et diminue à chaque étape. Nous allons montrer

que la longueur de la suite  $\mathcal{R}$  est égale à l'indice de Rabin du  $\omega$ -langage reconnu par l'automate. Puisque la suite  $\mathcal{R} = (R_0, \dots, R_m)$  est de poids minimal, aucune des transformations ne s'applique. On en déduit que  $R_m \neq \emptyset$  et que pour tout état  $q \in D_k$  avec  $k \geq 1$ , il y a une partie répétable  $R$  telle que  $q \in \text{Eac}(R)$ . Si de plus  $k \geq 2$ , il existe une partie répétable  $R$  qui vérifie  $q \in \text{Eac}(R)$  et  $R \cap D_{k-1} \neq \emptyset$ . Nous allons montrer que l'automate admet une chaîne de longueur  $m - 1$ . On choisit un état  $q_m \in R_m$  et une partie répétable  $S_m$  telle que  $q_m \in \text{Eac}(S_m)$  et  $S_m \cap D_{m-1} \neq \emptyset$ . On choisit alors un état  $q_{m-1} \in S_m \cap D_{m-1}$  et une partie répétable  $S_{m-1}$  telle que  $q_{m-1} \in \text{Eac}(S_{m-1})$  et  $S_{m-1} \cap D_{m-2} \neq \emptyset$ . On itère le procédé et l'on obtient des parties répétibles  $S_1, \dots, S_m$  telles que  $\text{Iac}(S_i) = i$  pour  $1 \leq i \leq m$  et  $S_i \cap S_{i+1} \neq \emptyset$  pour  $1 \leq i \leq m - 1$ . Cette dernière propriété assure que la partie  $T_i = S_1 + \dots + S_i$  est une partie répétable. De plus on a l'égalité  $\text{Iac}(T_i) = i$  pour  $1 \leq i \leq m$ . La suite  $(T_1, \dots, T_m)$  constitue donc une  $\mathcal{A}$ -chaîne positive de longueur  $m - 1$ . L'indice de Rabin du  $\omega$ -langage reconnu par l'automate est donc au moins égal à  $m$ . Sinon, on aurait  $X \in \hat{\mathbf{C}}_{m-1}^0$  ce qui contredit l'existence d'une chaîne positive de longueur  $m - 1$ .



# Références bibliographiques

- [Arn83a] A. Arnold. Rational  $\omega$ -languages are non-ambiguous. *Theoretical Computer Science*, 26:221–223, 1983.
- [Arn83b] A. Arnold. Topological characterizations of infinite behaviours of transition systems. In *Automata, Languages and Programming*, volume 153 of *Lecture Notes in Computer Science*, pages 28–38, Berlin, 1983. Springer-Verlag.
- [Arn85a] A. Arnold. Deterministic and non-ambiguous rational  $\omega$ -languages. In M. Nivat, editor, *Automata on Infinite Words*, volume 192 of *Lecture Notes in Computer Science*, pages 18–27, Berlin, 1985. Springer-Verlag.
- [Arn85b] A. Arnold. A syntactic congruence for rational  $\omega$ -languages. *Theoretical Computer Science*, 39:333–335, 1985.
- [Arn92] A. Arnold. *Système de transitions finis et sémantique des processus communicants*. Masson, 1992.
- [Bar92] R. Barua. The hausdorff-kuratowski hierarchy of  $\omega$ -regular languages and a hierarchy of muller automata. *Theoretical Computer Science*, 96:345–360, 1992.
- [Ber79] J. Berstel. *Transduction and context-free languages*. B.G. Teubner, 1979.
- [Bou74] N. Bourbaki. *Éléments de mathématiques, Topologie générale*. CCLS, Paris, 1974.
- [Büc60] J.R. Büchi. Weak second-order arithmetic and finite automata. *Z. Math. Logik und grundl. Math.*, 6:66–92, 1960.
- [Eil72] S. Eilenberg. *Automata, languages and machines*, volume A. Academic Press, New York, 1972.

- 
- [Eil76] S. Eilenberg. *Automata, languages and machines*, volume B. Academic Press, New York, 1976.
- [Hal74] P. Halmos. *Measure Theory*. Springer-Verlag, 1974.
- [Hau57] F. Hausdorff. *Set Theory*. Chelsea, New York, 1957.
- [How76] J. Howie. *An introduction to semigroup theory*. Academic Press, New York, 1976.
- [Kam85] M. Kaminski. A classification of  $\omega$ -regular languages. *Theoretical Computer Science*, 36:217–229, 1985.
- [Lal79] G. Lallement. *Semigroups and combinatorial applications*. Wiley, 1979.
- [Lan69] L.H. Landweber. Decision problems for  $\omega$ -automata. *Mathematical Systems Theory*, 3:376–384, 1969.
- [McN66] R. McNaughton. Testing and generating infinite sequences by a finite automaton. *Information and Control*, 9:521–530, 1966.
- [Mos80] Y.N. Moschovakis. *Descriptive Set Theory*. North Holland, 1980.
- [Péc86a] J.P. Pécuchet. Étude syntaxique des parties reconnaissables de mots infinis. volume 226 of *Lecture Notes in Computer Science*, pages 294–303, 1986.
- [Péc86b] J.P. Pécuchet. Variétés de semigroupes et mots infinis. In B. Monien and G. Vidal-Naquet, editors, *STACS '86*, volume 210 of *Lecture Notes in Computer Science*, pages 180–191, 1986.
- [Per90] D. Perrin. Finite automata. In J. van Leeuwen, editor, *Handbook of theoretical computer science*, volume B, chapter 1. Elsevier, 1990.
- [Pin84] J.E. Pin. *Variétés de langages formels*. Masson, 1984.
- [PP93a] D. Perrin and J.E. Pin. *Infinite Words*. Princeton University Press, 1993. (to appear, Version 1.1:LITP report 93.40).
- [PP93b] D. Perrin and J.E. Pin. Semigroups and automata on infinite words. In *In York Conference on semigroups*, 1993. À paraître.
- [Rab69] M.O. Rabin. Decidability of second-order theories and automata on infinite trees. *Trans. Amer. Math. Soc.*, 141:1–35, 1969.

- 
- [Rab72] M.O. Rabin. Automata on infinite objects and Church's problem. *Amer. Math. Soc. Regional Conference Series in Mathematics*, 13:1–23, 1972.
- [REBC82a] K. Richard, G. Elwin, R. Berlekamp, and J.H. Conway. *Winning Ways*, volume 1 : Games in General. Academic Press, 1982.
- [REBC82b] K. Richard, G. Elwin, R. Berlekamp, and J.H. Conway. *Winning Ways*, volume 2 : Games in Particular. Academic Press, 1982.
- [Sae90] B. Le Saec. Saturating right congruences for rational  $\omega$ -languages. *R.A.I.R.O.-Informatique Théorique et Applications*, 24:545–560, 1990.
- [Saf88] S. Safra. On the complexity of  $\omega$ -automata. In *29th Annual Symposium on Foundations of computer sciences*, pages 24–29, 1988.
- [Sim92] P. Simonnet. *Automates et Théorie Descriptive*. Thèse, Université Paris 7, 1992.
- [SPW91a] B. Le Saec, J.E. Pin, and P. Weil. Finite semigroup whose stabilizers are idempotents. *International Journal of Algebra and Computation*, 1(3):291–314, 1991.
- [SPW91b] B. Le Saec, J.E. Pin, and P. Weil. A purely algebraic proof of McNaughton's theorem on infinite words. *Foundations of Software Technology and Theoretical Computer Science*, 560:141–151, 1991.
- [Str79] H. Straubing. Recognizable sets and power sets of finite semigroups. *Semigroup Forum*, 18:331–340, 1979.
- [Tho79] W. Thomas. Star-free regular sets of  $\omega$ -sequences. *Information and Control*, 42:148–156, 1979.
- [Tho81] W. Thomas. A combinatorial approach to the theory of  $\omega$ -automata. *Information and Control*, 48:261–283, 1981.
- [Tho82] W. Thomas. Classifying regular events in symbolic logic. *J. Comput. System Sci.*, 25:360–375, 1982.
- [Tho90] W. Thomas. Automata on infinite objects. In J. van Leeuwen, editor, *Handbook of theoretical computer science*, volume B, chapter 4. Elsevier, 1990.
- [Wag79] K. Wagner. On  $\omega$ -regular sets. *Information and Control*, 43:123–177, 1979.

- [Wil91] T. Wilke. An Eilenberg theorem for  $\infty$ -languages. In *ICALP '91*, volume 510 of *Lecture Notes in Computer Science*, pages 588–599, Berlin, 1991. Springer-Verlag.

# Table des Figures

|      |   |    |
|------|---|----|
| 1.1  | Un automate . . . . .   | 12 |
| 1.2  | Un automate déterministe et complet. . . . .                      | 14 |
| 1.3  | Automate de Büchi. . . . .  | 14 |
| 1.4  | Automate de Büchi déterministe et complet. . . . .                | 15 |
| 1.5  | Automate de Rabin : $\mathcal{R} = \{\{1, 3\}, \{2\}\}$ . . . . . | 19 |
| 1.6  | Automate de Muller : $\mathcal{T} = \{\{1\}, \{2\}\}$ . . . . .   | 19 |
| 2.1  | Structure de $\hat{S} \circ \mathbb{Z}/2\mathbb{Z}$ . . . . .     | 43 |
| 2.2  | Transducteur . . . . .  | 44 |
| 3.1  | $X = X_1 - X_2 + X_3 - X_4 + X_5 - X_6 + X_7$ . . . . .           | 52 |
| 3.2  | $Y = Y_1 - Y_2 + Y_3 - Y_4 + Y_5$ . . . . .                       | 53 |
| 3.3  | Diagramme de Venn de l'intersection . . . . .                     | 54 |
| 3.4  | Diagramme de Venn de l'union . . . . .                            | 54 |
| 3.5  | $X_3Y_2$ . . . . .  | 55 |
| 3.6  | $X_5Y_1 + X_3Y_2 + X_1Y_3 + EY_5$ . . . . .                       | 55 |
| 3.7  | Chaînes pour $m = 2$ et $n = 1$ . . . . .                         | 56 |
| 3.8  | Chaînes pour $m = 3$ et $n = 1$ . . . . .                         | 56 |
| 3.9  | Chaînes pour $m = 3$ et $n = 2$ . . . . .                         | 56 |
| 3.10 | Chaîne de $XY$ . . . . .  | 57 |
| 3.11 | Chaîne de $X + Y$ . . . . .                                       | 57 |
| 3.12 | Configurations équivalentes . . . . .                             | 58 |
| 4.1  | Chaîne de longueur 4 . . . . .                                    | 64 |

---

|      |  |     |
|------|--|-----|
| 4.2  | Superchaîne de longueur 2 avec des chaînes de longueur 3 . . . . .                     | 65  |
| 4.3  | Automate $\mathcal{A}_1 : \mathcal{T} = \{\{1\}, \{1, 2\}\}$ . . . . .                 | 66  |
| 4.4  | Automate $\mathcal{A}_2 : \mathcal{T} = \{\{1\}, \{1, 2\}, \{4\}\}$ . . . . .          | 67  |
| 4.5  | Structure en $\mathcal{D}$ -classes de $S = \{b = 1, a = 0\}$ . . . . .                | 68  |
| 4.6  | Structure en $\mathcal{D}$ -classes de $\{b = 1, a = a^2, c = 0\}$ . . . . .           | 68  |
| 4.7  | L'ordre induit sur les classes . . . . .   | 81  |
| 4.8  | Automate $\mathcal{A}_3 : \mathcal{T} = \{\{2\}\}$ . . . . .                           | 83  |
| 4.9  | Automate $\mathcal{A}_4 : \mathcal{T} = \{\{2\}, \{4\}, \dots\}$ . . . . .             | 84  |
| 4.10 | Automate $\mathcal{A}_5 : \mathcal{T} = \{\{0, 1\}, \{0, 1, 2, 3\}, \dots\}$ . . . . . | 88  |
| 4.11 | Structure en $\mathcal{D}$ -classes de $T$ . . . . .                                   | 93  |
| 4.12 | Automate $\mathcal{A}_6$ . . . . .   | 93  |
| 5.1  | Automate $\mathcal{A}$ . . . . .   | 99  |
| 5.2  | Automate $\mathcal{B}$ . . . . .   | 99  |
| 5.3  | Automate $\mathcal{C}$ . . . . .   | 100 |

# Table des Matières

|          |  |           |
|----------|--|-----------|
| <b>1</b> | <b>Premières définitions</b>                             | <b>7</b>  |
| 1.1      | Notations . . . . .                                      | 7         |
| 1.2      | Mots et langages . . . . .                               | 7         |
| 1.2.1    | Langages rationnels . . . . .                            | 9         |
| 1.2.2    | Topologie . . . . .                                      | 10        |
| 1.3      | Automates . . . . .                                      | 12        |
| <b>2</b> | <b>Semigroupes et <math>\omega</math>-semigroupes</b>    | <b>23</b> |
| 2.1      | Introduction . . . . .                                   | 23        |
| 2.2      | Semigroupes . . . . .                                    | 24        |
| 2.3      | $\omega$ -semigroupes . . . . .                          | 26        |
| 2.4      | Algèbres de Wilke . . . . .                              | 27        |
| 2.5      | Factorisations Ramseyennes . . . . .                     | 28        |
| 2.6      | $\omega$ -semigroupes finis . . . . .                    | 32        |
| 2.7      | Reconnaissance par morphisme . . . . .                   | 34        |
| 2.8      | Opérations . . . . .                                     | 35        |
| 2.8.1    | Semigroupe et $\omega$ -semigroupe des parties . . . . . | 36        |
| 2.8.2    | Produit de Schützenberger . . . . .                      | 37        |
| 2.8.3    | Produit en couronne . . . . .                            | 40        |
| <b>3</b> | <b>Chaînes d'ensembles</b>                               | <b>49</b> |
| 3.1      | Introduction . . . . .                                   | 49        |
| 3.2      | Notations . . . . .                                      | 49        |

|          |   |            |
|----------|---|------------|
| 3.3      | Sommes de différences et chaînes . . . . .                  | 50         |
| 3.4      | Algèbre de Boole . . . . .                                  | 61         |
| <b>4</b> | <b>Hiérarchie de Wagner</b>                                 | <b>63</b>  |
| 4.1      | Introduction . . . . .                                      | 63         |
| 4.2      | Chaînes et superchaînes . . . . .                           | 63         |
| 4.2.1    | Chaînes et superchaînes d'un $\omega$ -langage . . . . .    | 63         |
| 4.2.2    | Chaînes et superchaînes d'un automate . . . . .             | 65         |
| 4.2.3    | Chaînes et superchaînes d'un $\omega$ -semigroupe . . . . . | 67         |
| 4.2.4    | Équivalences des définitions . . . . .                      | 69         |
| 4.3      | Propriétés . . . . .  | 73         |
| 4.4      | Autres caractérisations . . . . .                           | 76         |
| 4.4.1    | Dans les $\omega$ -semigroupes . . . . .                    | 76         |
| 4.4.2    | Dans les automates . . . . .                                | 78         |
| 4.5      | Classes de complexité . . . . .                             | 81         |
| 4.5.1    | Indice de Rabin . . . . .                                   | 89         |
| 4.5.2    | Stabilité par fonction séquentielle inverse . . . . .       | 89         |
| 4.6      | Application . . . . .                                       | 91         |
| 4.6.1    | Construction de l'automate . . . . .                        | 92         |
| 4.6.2    | Preuve de l'automate . . . . .                              | 94         |
| 4.6.3    | Cas général . . . . .                                       | 95         |
| <b>5</b> | <b>Automates de Rabin à chaînes</b>                         | <b>97</b>  |
| 5.1      | Introduction . . . . .                                      | 97         |
| 5.2      | Notations et définitions . . . . .                          | 97         |
| 5.3      | Constructions . . . . .                                     | 98         |
| 5.4      | Réduction . . . . .   | 105        |
|          | <b>Références bibliographiques</b>                          | <b>109</b> |