THESE DE DOCTORAT

en Informatique Fondamentale

Automates sur les ordres linéaires : Complémentation

Chloé RISPAL

Soutenue le 7 décembre 2004 devant le jury composé de

Jean Berstel
Véronique Bruyère
Olivier Carton
Didier Caucal
Dominique Perrin
Jean-Eric Pin

Institut Gaspard Monge,
Université de Marne-la-Vallée

Remerciements

Je remercie Olivier Carton, mon directeur de thèse, pour m'avoir laissé choisir un sujet si intéressant et pour m'avoir soutenue dans mon travail.
Je remercie Didier Caucal, mon maître de stage de DEA, pour m'avoir initiée à l'informatique théorique et m'avoir transmis sa passion de la recherche. Je le remercie également d'avoir accepté d'être rapporteur et d'être toujours disponible pour de nouveaux conseils.
Un grand merci aussi à Jean-Eric Pin d'avoir accepté d'être rapporteur malgré ses nombreuses sollicitations et aux membres de mon jury : Jean Berstel, Véronique Bruyère et Dominique Perrin.
Je tiens également à remercier Nicolas Bedon pour son aide précieuse dans la relecture de mon manuscrit et son soutien moral tout au long de ma thèse ainsi que Christophe Morvan pour m'avoir maintenue dans l'univers des graphes infinis.
Je remercie aussi tous les membres de l'IGM, en particulier Frédérique Bassino, Julien Clément, Cyril Nicaud, Marie-Pierre Béal, Jung Eun-jin, Isabelle Fagnot, Julien Allali, Jérome Petazzoni, Pirro Bracka, Anne Micheli, Eric Incerti, Giuseppina Rindone, Patrice Hérault, Dominique Revuz, Marc Zipstein, Etienne Duris, Gilles Roussel, Pierre Peterlongo, Rémi Forax, Benoît Olivieri, Serge Midonnet, Theresa Gomez-Diaz, Line Fonfrède, Madeleine Eklou, Julien Cervelle, Florent Hivert, Jean-Christophe Novelli, Jacques Désarménien, Maxime Crochemore, Claude Martineau, Dominique Rossin, Wolgang Thomas et l'équipe Galion de L'IRISA.
Pour finir je tiens à remercier mes parents, Adrien, Amélie, Nina, Arthur et Emma, George et Véro, ma famille, mes amis : Laurence, Béa, Marie, Fleur, Elsa, Sophie, Sonia, Céline, Athina, Miguel, Perez, Benj, Yoyo, Mathilde, Mathieu, Jonathan, Tarik et Christine, les insaliens, tous ceux que j'ai oublié. Je remercie surtout Flo pour avoir été à mes cotés pendant ces années de thèse.

Résumé

Cette thèse traite des ensembles rationnels de mots indexés par des ordres linéaires et en particulier du problème de la fermeture par complémentation.

Dans un papier fondateur de 1956, Kleene initie la théorie des langages en montrant que les automates sur les mots finis et les expressions rationnelles ont le même pouvoir d'expression. Depuis, ce résultat a été étendu à de nombreuses structures telles que les mots infinis (Büchi, Muller), bi-infinis (Beauquier, Nivat, Perrin), les mots indexés par des ordinaux (Büchi, Bedon), les traces, les arbres... Plus récemment, Bruyère et Carton ont introduit des automates acceptant des mots indexés par des ordres linéaires et des expressions rationnelles correspondantes. Ces structures linéaires comprennent les mots infinis, les mots indexés par des ordinaux et leurs miroirs. Le théorème de Kleene a été généralisé aux mots indexés par les ordres linéaires dénombrables et dispersés, c'est-à-dire les ordres ne contenant pas de sous-ordre isomorphe à $\mathbb{Q}$.

Pour la plupart des structures, la classe des ensembles rationnels forme une algèbre de Boole. Cette propriété est nécessaire pour traduire une logique en automates. La fermeture par complémentation restait un problème ouvert. Dans cette thèse, on résout ce problème de façon positive : on montre que le complément d'un ensemble rationnel de mots indexés par des ordres linéaires dispersés est rationnel. La méthode classique pour obtenir un automate acceptant le complémentaire d'un ensemble rationnel se fait par déterminisation. Nous montrons que cette méthode ne peut-être appliquée dans notre cas : tout automate n'est pas nécessairement équivalent à un automate déterministe. Nous avons utilisé d'autres approches. Dans un premier temps, nous généralisons la preuve de Büchi, basée sur une congruence de mots, et obtenons ainsi la fermeture par complémentation dans le cas des ordres linéaires de rang fini. Pour obtenir le résultat dans le cas général, nous utilisons l'approche algébrique. Nous développons une structure algébrique qui étend la reconnaissance classique par semigroupes finis : les semigroupes sont remplacés par les $\diamond$-semigroupes qui possèdent un produit généralisé. Nous prouvons qu'un ensemble est rationnel si et seulement s'il est reconnu par un $\diamond$-semigroupe fini. Nous montrons aussi qu'un $\diamond$-semigroupe canonique, appelé $\diamond$-semigroupe syntaxique, peut être associé à chaque ensemble rationnel. Notre preuve de la complémentation est effective.

Le théorème de Schützenberger établit qu'un ensemble de mots finis est sans étoile si et seulement si son semigroupe syntaxique est fini et apériodique. Pour finir, nous étendons partiellement ce résultat au cas des ordres de rang fini.

Abstract

This thesis treats of rational sets of words indexed by linear orderings and particularly of the problem of the closure under complementation.

In a seminal paper of 1956, Kleene started the theory of languages establishing that automata on finite words and rational expressions have the same expressive power. Since then, this result has been extended to many structures such as infinite words (Büchi, Muller), bi-infinite words (Beauquier, Nivat, Perrin), ordinal words (Büchi, Bedon), traces, trees... . More recently, Bruyère and Carton have introduced automata accepting words indexed by linear orderings and the corresponding rational expressions. Those linear structures include infinite words, ordinal words and their mirrors. Kleene's theorem has been generalized to words indexed by countable scattered linear orderings, that is orderings without any sub-ordering isomorphic to $\mathbb{Q}$.

For many structures, the class of rational sets forms a boolean algebra. This property is necessary to translate logic into automata. The closure under complementation was left as an open problem. In this thesis, we solve it in a positive way : we prove that the complement of a rational set of words indexed by scattered linear orderings is rational. The classical method to get an automaton accepting the complement of a rational set is trough determinization. We show that this method can not be applied in our case : An automaton is not necessary equivalent to a deterministic one. We have used other approaches. First, we generalize the proof of Büchi, based on congruence of words, to obtain the closure under complementation in the case of linear orderings of finite ranks. To get the whole result in the general case, we use the algebraic approach. We develop an algebraic structure extending the classical recognition by finite semigroups : semigroups are replaced by $\diamond$-semigroups equipped with a generalized product. We prove that a set is rational iff it is recognized by a finite $\diamond$-semigroup. We also prove that a canonical $\diamond$-semigroup can be associated to each rational set : the syntactic $\diamond$-semigroup. Our proof of the closure under complementation is effective.

The theorem of Schützenberger establishes that a set of finite words is star-free if and only if its syntatic semigroup is finite and aperiodic. To finish, we partially extend this result to linear orderings of finite ranks.

Table des matières

Introduction

Cette thèse s'inscrit dans le cadre général de la théorie des automates et des langages formels. Cette théorie est utilisée dans de nombreux domaines scientifiques pour spécifier et vérifier des propriétés sur des suites de symboles. En informatique, les automates ont de nombreuses applications telles que l'analyse lexicale ou la vérification de programmes.

Kleene initia la théorie des automates en 1956, établissant que les ensembles de mots finis décrits par des expressions rationnelles étaient exactement les langages acceptés par des automates finis [21]. Depuis lors, la notion d'ensemble rationnel s'est progressivement généralisée. Les ensembles rationnels de mots de longueurs ω ont tout d'abord été considérés par Büchi [11, 12], Muller [25] et McNaughton [23] qui généralisa le théorème de Kleene à ces mots infinis. Büchi [13] a ensuite introduit les automates acceptant des mots indexés par des ordinaux. Ceux-ci sont des automates finis équipés de transitions limites de la forme $P \longrightarrow p$ où p est un état de l'automate et P un ensemble d'états. Ces transitions sont utilisées à la fin de chemins dont la longueur est un multiple de ω. Les expressions rationnelles équivalentes ont été introduites plus tard par Wojciechowski [45]. Les mots bi-infinis ont ensuite été étudiés par Nivat et Perrin [26] puis par Beauquier [19]. Enfin, les mots indexés par les ordres linéaires ont été considérés. Ces mots contiennent en particulier les mots finis, infinis, transfinis et leurs miroirs. Bruyère et Carton ont étendu le théorème de Kleene aux mots indexés par des ordres linéaires dénombrables dispersés [8]. Rappelons qu'un ordre linéaire est dispersé s'il ne contient pas de sous-ordre dense isomorphe à $\mathbb{Q}$. Les automates possèdent des transitions limites de la forme $P \longrightarrow p$ ou $p \longrightarrow P$ où p est un état de l'automate et P un ensemble d'états. Les opérateurs rationnels correspondant sont l'union $+$, le produit de concaténation fini $\cdot$, l'itération finie $*$, l'itération infinie à droite ω et à gauche $-\omega$, l'itération ordinale $\natural$, l'itération ordinale inverse $-\natural$ ainsi que l'opérateur binaire $\diamond$ qui est l'équivalent d'une itération sur tous les ordres linéaires dispersés.

Pour la plupart des structures, la classe des ensembles rationnels est fermée par de nombreux opérateurs tels que les substitutions, les substitutions inverses et les opérations booléennes. La clôture par union et par intersection est généralement facile à démontrer. La fermeture par complémentation est souvent beaucoup plus difficile à obtenir. Cette propriété est importante à la fois d'un point de vue théorique et pratique. Elle signifie que la classe des ensembles rationnels forme une algèbre de Boole et est utilisée chaque fois qu'une logique est traduite en automates. Par exemple, dans les preuves de décidabilité de la théorie du second ordre monadique des entiers par Büchi [12] et celle de la décidabilité de la théorie du second ordre monadique de l'arbre binaire infini par Rabin [34], la fermeture par complémentation est une propriété essentielle. Dans [8], la clôture par complémentation restait une question ouverte pour les mots sur les ordres linéaires. Nous résolvons le problème dans cette thèse en montrant que la classe des ensembles rationnels de mots indexés par des ordres linéaires dénombrables et dispersés est fermée par complémentation.

Étant donné un automate $\mathcal{A}$ acceptant un ensemble de mots finis, la méthode classique pour construire un autre automate $\mathcal{B}$ acceptant exactement les mots qui ne sont pas acceptés par $\mathcal{A}$ se fait par déterminisation. Cette méthode est déjà plus difficile dans le cas des mots infinis. Büchi [12] a utilisé une autre approche, basée sur une congruence de mots finis. McNaughton [23] a étendu la méthode de déterminisation aux mots infinis prouvant que tout automate de Büchi est équivalent à un automate de Muller déterministe. Cette méthode a été généralisée par Büchi [13] aux mots transfinis mais elle devient alors très complexe. Contrairement aux automates sur les ordinaux [2], les automates sur les ordres linéaires ne sont pas déterminisables. La fermeture par complémentation des ensembles rationnels ne peut donc pas s'obtenir directement à partir des automates. Dans le cas des ordres linéaires dispersés de rang fini, c'est-à-dire lorsque le nombre d'imbrication de ω et $-\omega$ est fini, nous prouvons la fermeture par complémentation [16] en généralisant la preuve de Büchi [12] avec une induction supplémentaire sur le rang. Parce que les rangs des ordres linéaires dénombrables dispersés prennent valeur parmi tous les ordinaux [20], cette approche n'est pas adaptée pour les mots indexés par tous ces ordres. C'est pourquoi nous utilisons l'approche algébrique.

Dans le cas des mots finis, on peut de façon algébrique classifier les ensembles rationnels en fonction des propriétés de leur semigroupe syntaxique. Les ensembles rationnels de mots finis sont définis algébriquement comme les ensembles reconnus par un semigroupe fini. Cette importante caractérisation a pour conséquence immédiate la propriété de fermeture par complémentation. Elle se généralise de façon naturelle aux mots infinis [42, 31] reconnus par des ω-semigroupes finis et aux mots sur les ordinaux [4, 7] reconnus par des ω_1-semigroupes finis. Nous définissons ici une structure algébrique adaptée aux mots indexés par les ordres linéaires : les $\diamond$-semigroupes. Nous prouvons que, lorsqu'ils sont finis, ces $\diamond$-semigroupes sont équivalents aux automates sur les ordres linéaires et obtenons ainsi la fermeture par complémentation de la classe des ensembles rationnels de mots indexés par les ordres linéaires dénombrables et dispersés. Par analogie avec le cas des mots finis, nous prouvons également qu'un $\diamond$-semigroupe minimal peut-être associé à chaque ensemble rationnel : le $\diamond$-semigroupe syntaxique. Notre preuve de la fermeture par complémentation est effective. Elle donne également une autre preuve de la décidabilité de l'équivalence de ces automates [10]. Notons qu'il est nécessaire que les ordres linéaires soient dénombrables et dispersés. Büchi avait déjà mentionné que la classe des ensembles rationnels de mots transfinis de longueur supérieure à ω_1 (le plus petit ordinal non dénombrable) n'est pas fermée par complémentation. On peut également prouver que l'ensemble des mots indexés par des ordres linéaires dispersés n'est pas rationnel en tant que sous-ensemble des mots sur tous les ordres linéaires alors que son complémentaire l'est.

Dans la classification algébrique des ensembles rationnels, le fameux théorème de Schützenberger [38] établit que les ensembles sans étoile de mots finis sont exactement les ensembles dont le semigroupe syntaxique est fini et apériodique. Ce résultat a été étendu aux mots infinis [22, 41, 29] et même aux mots transfinis [5]. Nous le généralisons ici aux ensembles sans étoile de mots sur les ordres linéaires dispersés de rang fini.

Cette thèse est organisée en chapitres de la façon suivante :

– Dans le premier chapitre, nous rappelons certaines définitions élémentaires et fixons les notations. Les notions d'ordre linéaire et de mot sont définis. Nous rappelons également certains résultats généraux largement utilisés dans les chapitres suivants, notamment la caractérisation inductive des ordres linéaires dispersés due à Hausdorff ainsi que certaines définitions algébriques.

– Le chapitre 2 est consacré aux automates et aux expressions rationnelles. Ceux-ci sont successivement détaillés dans le cas des mots finis, infinis, bi-infinis, des mots sur les ordinaux et enfin des mots indexés par les ordres linéaires dispersés. Dans chacun des cas, les deux approches sont équivalentes et le théorème de Kleene est énoncé.

– Dans le chapitre 3, on prouve qu'un ensemble rationnel de mots sur les ordres linéaires dispersés n'est pas nécessairement accepté par un automate déterministe. On se restreint alors au cas des ordres de rang fini. Nous généralisons la preuve de Büchi et montrons que la classe des ensembles rationnels de mots indexés par des ordres linéaires dénombrables dispersés de rang fini est fermée par complémentation. La preuve utilise une congruence sur les mots et une induction sur le rang.

– Le chapitre 4 généralise les résultats du chapitre précédent en utilisant une caractérisation algébrique des ensembles rationnels. Nous rappelons les structures algébriques déjà existantes adaptées aux mots finis, infinis et transfinis ainsi que certaines propriétés algébriques des semigroupes finis et les relations de Green. Nous définissons ensuite les $\diamond$-semigroupes, dont le produit de toute suite d'éléments indexée par un ordre linéaire dénombrable et dispersé est défini. Nous montrons que lorsqu'un $\diamond$-semigroupe est fini, il peut être décrit de façon finie par un semigroupe et des fonctions compatibles à droite et à gauche avec ce semigroupe. Nous prouvons ensuite qu'un ensemble de mots sur les ordres linéaires dénombrables et dispersés est rationnel si et seulement si il est reconnu par un $\diamond$-semigroupe fini. Nous en déduisons que la classe des ensembles rationnels est fermée par complémentation. Enfin, pour tout ensemble reconnaissable X, nous définissons le $\diamond$-semigroupe syntaxique de X et prouvons que c'est le plus petit $\diamond$-semigroupe reconnaissant X au sens de la division.

– Dans le chapitre 5, nous étendons la définition de langages sans étoile aux mots indexés par des ordres linéaires de rang fini et généralisons le théorème de Schützenberger : un ensemble rationnel de mots indexés par des ordres linéaires dénombrables dispersés de rang fini est sans étoile si et seulement si son semigroupe syntaxique est fini et apériodique.

Chapitre 1

Définitions et notations

Ce chapitre est consacré aux définitions et notations utilisées dans la suite. Nous rappelons les propriétés élémentaires sur les ensembles, relations et fonctions, puis nous définissons les ordinaux et les ordres linéaires. Pour les ordres linéaires dispersés, nous énonçons la caractérisation d'Hausdorff qui sera largement utilisée dans les chapitres suivants. Les mots indexés par des ordres linéaires sont ensuite présentés, ainsi que le théorème de Ramsey dont nous verrons différentes versions par la suite. Enfin nous redonnons certaines définitions algébriques de base.

1.1 Ensembles, relations et fonctions

Pour tout ensemble E, on note $\mathcal{P}(E)$ l'ensemble des parties de E. Cet ensemble contient notamment l'ensemble vide noté $\emptyset$. On parlera souvent d'une *classe* d'ensembles pour définir une collection d'ensembles C. Les notations $\bigcup\limits_{E \in C} E$ et $\bigcap\limits_{E \in C} E$ désignent respectivement l'union et l'intersection des éléments de C. Pour deux ensembles E et F, la différence $E \setminus F$ est l'ensemble des éléments de E n'appartenant pas à F. Enfin, si E' est un sous-ensemble de E noté $E' \subseteq E$, le complémentaire de E' dans E, à savoir l'ensemble $E \setminus E'$ sera noté $\overline{E'}$.

Pour tout entier n, on note $E_1 \times E_2 \times \ldots \times E_n$ le produit cartésien des ensembles $E_1, \ldots E_n$ défini comme l'ensemble $\{(e_1, e_2, \ldots, e_n) | e_1 \in E_1, \ldots, e_n \in E_n\}$. La notation E^n dénote l'ensemble des n-uplets d'éléments de E.

Une *relation* sur des ensembles E et F est une partie de $E \times F$. Etant donnée une relation $\mathcal{R} \subseteq E \times F$, on note aussi $e\mathcal{R}f$ pour signifier que $(e, f) \in \mathcal{R}$. Une relation $\mathcal{R} \subseteq E \times E$ est dite *réflexive* si quel que soit $e \in E$, $e\mathcal{R}e$. Elle est *symétrique* si quels que soient $e_1, e_2 \in E$, $e_1\mathcal{R}e_2$ implique $e_2\mathcal{R}e_1$ et *antisymétrique* si $e_1\mathcal{R}e_2$ et $e_2\mathcal{R}e_1$ implique $e_1 = e_2$. Enfin $\mathcal{R}$ est *transitive* si quels que soient $e_1, e_2, e_3 \in E$, $e_1\mathcal{R}e_2$ et $e_2\mathcal{R}e_3$ implique $e_1\mathcal{R}e_3$.

Une relation antisymétrique et transitive est une *relation d'ordre.* Lorsqu'une relation d'ordre est réflexive, elle est généralement noté $\leq$. Une relation d'ordre $\mathcal{R}$ sur E est *stricte* si pour tout $e \in E$, non $e\mathcal{R}e$. Une relation d'ordre stricte est souvent notée $<$. Un ensemble E muni d'une relation d'ordre $<$ admet un minimum s'il existe $m \in E$ tel que quel que soit $e \in E$, $m \leq e$. Une relation d'ordre sur un ensemble E est *bien ordonnée* si et seulement si toute partie non vide de E admet un élément minimum.

Une relation réflexive, symétrique et transitive est une *relation d'équivalence*. La *classe d'équivalence* d'un élément $e \in E$ est l'ensemble des éléments e' tels que $e\mathcal{R}e'$. Lorsqu'une relation d'équivalence possède un nombre fini de classes d'équivalence, ce nombre est appelé l'*indice* de la relation. Les relations sont ordonnées par inclusions. Une relation $\mathcal{R}_1$ est *plus fine* que $\mathcal{R}_2$ si $e_1\mathcal{R}_1e_2$ implique $e_1\mathcal{R}_2e_2$. L'égalité est la relation d'équivalence la plus fine et la relation universelle est la plus grossière.

Une *application* (respectivement *fonction*) $\varphi : E \to F$ est une relation telle que quel que soit $x \in E$, il existe exactement (respectivement au plus) un $y \in F$ tel que $(x, y) \in \varphi$. On note $\varphi(x) = y$ l'*image* de l'élément x et plus généralement pour un ensemble $E' \subseteq E$, $\varphi(E') = \bigcup_{x \in E'} \varphi(x)$. Une application $\varphi : E \to F$ est *injective* si quels que soient $x, y \in E$, $\varphi(x) = \varphi(x') \Longrightarrow x = x'$ et *surjective* si quel que soit $y \in F$, il existe $x \in E$ tel que $\varphi(x) = y$. Une application injective et surjective est dite *bijective*. La *composée* de deux fonctions $f : E \to F$ et $g : F \to G$ est la fonction $g \circ f : E \to G$ définie par $g \circ f(x) = z$ s'il existe $y \in F$ tel que $f(x) = y$ et $g(y) = z$.

Soient deux ensembles E et E' munis d'ordres respectivement notés $<$ et $<'$. Une application $\varphi : E \to E'$ est un *morphisme d'ordre* si quels que soient les éléments $e_1, e_2 \in E$, $e_1 < e_2$ si et seulement si $\varphi(e_1) <' \varphi(e_2)$.

1.2 Cardinaux

Un entier naturel peut être utilisé pour décrire la taille d'un ensemble ou bien pour définir la position d'un élément dans une suite. Lorsque les ensembles sont finis, ces deux notions coïncident. Dans le cas d'ensembles infinis, il faut distinguer ces deux approches. Les tailles des ensembles sont décrits par les nombres cardinaux, découverts par Cantor, tandis que l'aspect position d'un élément dans une suite est généralisé par les ordinaux.

Le cardinal d'un ensemble fini à n éléments est n. Le cardinal de l'ensemble des nombres entiers naturels $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$ est noté $\aleph_0$. C'est le plus petit nombre cardinal transfini. Deux ensembles ont même cardinal lorsqu'ils sont en bijection. Un ensemble est dit *dénombrable* s'il est de cardinal $\aleph_0$ c'est-à-dire s'il est en bijection avec $\mathbb{N}$. Dans la suite on travaillera uniquement sur des ensembles dénombrables. Notons que l'union dénombrable d'ensembles dénombrables est un ensemble dénombrable.

1.3 Ordinaux

Nous rappelons ici brièvement les nombres ordinaux qui sont plus largement détaillés dans [37] ou [39, 40]. Les ordinaux étant en particulier des ordres linéaires, les opérations sur les ordinaux seront définies plus précisément dans le paragraphe suivant.

En théorie des ensembles, les entiers naturels peuvent être définis comme des ensembles bien ordonnés : $0 = \emptyset$, $1 = \{\emptyset\}$, $2 = \{0, 1\} = \{\emptyset, \{\emptyset\}\}$, $3 = \{0, 1, 2\} = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$, etc. L'ensemble 3, par exemple, possède les éléments 0,1,2 et est ordonné par la relation d'appartenance $0 < 1 < 2$ qui est un bon ordre. Un entier est plus petit qu'un autre si et seulement s'il est l'un de ses éléments. En identifiant les ensembles ordonnés isomorphes, tout ensemble fini bien ordonné définit un entier naturel. Cette approche se généralise aux ensembles non nécessairement finis. Un *ordinal* est une classe d'équivalence d'ensembles bien ordonnés pour la relation d'isomorphisme. Un *ordinal* β est représenté par l'ensemble des ordinaux plus petits que β sur lequel la relation d'appartenance est un bon ordre et qui vérifie que tout élément de β est aussi un sous-ensemble de β. Dans la suite, les ordinaux seront notés par des lettres grecques minuscules α, β, γ,... Le plus petit ordinal est donc $\emptyset$ noté $\mathbf{0}$. Par définition, tout entier n est l'ordinal associé à l'ensemble fini $\{0, 1, \ldots, n - 1\}$. Le premier nombre ordinal non fini est ω. Il correspond à l'ensemble des entiers naturels $\mathbb{N}$ muni de son ordre habituel.

L'addition des entiers, traduite en termes d'ensembles, permet de généraliser l'addition aux nombres ordinaux transfinis. Soient α et β deux ordinaux. L'ordinal $\alpha + \beta$ est l'ensemble $\alpha \cup \beta$ tel que tout élément de α soit plus petit que n'importe quel élément de β. Les ensembles α et β préservant leurs bons ordres respectifs. Intuitivement, l'opération $\alpha + \beta$ revient à juxtaposer β à droite de α. Par exemple, l'ordinal $\omega + \omega$ est composée de deux copies de l'ensemble des entiers naturels $0 < 1 < 2 < \ldots < 0' < 1' < 2' < \ldots$. Notons que $3 + \omega = \omega$ puisque $0 < 1 < 2 < 0' < 1' < 2' < \ldots$ est isomorphe à ω. Plus généralement, pour tout entier n, $n + \omega = \omega$. Par contre $\omega + 3$ représenté par $0 < 1 < 2 < \ldots < 0' < 1' < 2'$ n'est pas isomorphe à ω. L'addition n'est donc pas commutative.

Pour multiplier deux ordinaux α et β, on remplace chaque élément de β par une copie de α. Par exemple l'ordinal 2ω représenté par $0_0 < 1_0 < 0_1 < 1_1 < 0_2 < 1_2 < 0_3 < 1_3 < \ldots$ est isomorphe à ω. Par contre $\omega 2$ est égal à $\omega + \omega$ donc cette multiplication n'est pas non plus commutative.

Un ordinal α est *successeur* s'il possède un plus grand élément c'est-à-dire s'il existe un ordinal β tel que $\alpha = \beta + 1$. Un ordinal différent de $\mathbf{0}$ et qui n'est pas successeur est *limite*. Par exemple ω est un ordinal limite.

1.4 Ordres linéaires

Ce paragraphe est consacré aux ordres linéaires et le lecteur interessé pourra consulter l'ouvrage de Rosenstein [37] pour une description plus complète. Après les définitions et opérations sur les ordres linéaires, la caractérisation de Hausdorff des ordres linéaires dispersés et la notion de coupure sont décrites.

1.4.1 Définitions et opérations

Une relation d'ordre $<$ sur un ensemble J est *totale* si tous les éléments sont comparables : quels que soient j et k dans J, on a $j < k$ ou $k < j$. Cette propriété induit une représentation linéaire des ordres c'est pourquoi un ensemble muni d'un ordre total est appelé un *ordre linéaire*. Les ordres linéaires seront aussi définis à isomorphisme près et souvent notés I, J, K... Tout ensemble fini totalement ordonné de cardinal n est isomorphe à l'ensemble $\{0, 1, 2, \ldots, n-1\}$ et noté n. L'ordre des entiers naturels et celui des nombres rationnels sont des ordres linéaires. Contrairement aux ordinaux, une partie non vide d'un ordre linéaire n'admet pas nécessairement de minimum. Les ordres linéaires peuvent donc contenir des sous-ordres infinis à gauche. Par exemple, l'ensemble des entiers relatifs $\{\ldots, -2, -1, 0, 1, 2, \ldots\}$ muni de l'ordre habituel est un ordre linéaire noté ζ. Les ordinaux sont aussi des ordres linéaires.

Pour tout ordre linéaire J, on note $-J$ l'ordre inverse appelé aussi ordre miroir de J. Si la relation d'ordre sur J est $<$, l'ordre $-J$ est l'ensemble J muni de $<^*$ défini pour tous les éléments i,j de J par $i <^* j$ si et seulement si $j < i$. Par exemple, $-\omega$ correspond à l'ensemble des entiers négatifs avec l'ordre habituel : $\ldots < -2 < -1 < 0$.

La somme $J+K$ de deux ordres linéaires est définie comme l'ensemble $J \cup K$ muni de l'ordre $<$ étendant les ordres de J et K de la façon suivante : $j < k$ pour tout $j \in J$ et $k \in K$. Par exemple, $\zeta = -\omega + \omega$. Plus généralement, pour tous les ordres linéaires J et K_j où $j \in J$, la *somme* $\sum\limits_{j \in J} K_j$ est l'ensemble $\bigcup\limits_{j \in J} K_j$ muni de l'ordre étendu par J. On parlera d'une somme indexée par J ou d'une J-somme. Formellement, $\sum\limits_{j \in J} K_j$ est l'ensemble de toutes les paires (k, j) telles que $k \in K_j$ équipé de l'ordre lexicographique défini par $(k, j) < (k', j')$ si et seulement si $j < j'$ ou $j = j'$ et $k < k'$ dans K_j.

Exemple 1 L'ordre linéaire $\sum\limits_{\omega} \zeta$ correspond à juxtaposer ω copies de l'ordre ζ. Cet ordre est représenté à la Figure 1.1 où chaque $\bullet$ correspond à un élément de l'ordre.

$$\cdots\bullet\;\;\bullet\;\;\bullet\;\;\bullet\;\;\bullet\cdots\qquad\cdots\bullet\;\;\bullet\;\;\bullet\;\;\bullet\;\;\bullet\cdots\qquad\cdots\bullet\;\;\bullet\;\;\bullet\;\;\bullet\;\;\bullet\cdots\quad\cdots$$

FIG. 1.1 – Représentation de l'ordre $\sum\limits_{\omega} \zeta$.

Soient J et K deux ordres linéaires. La somme $\sum\limits_{J} K$ correspond à la multiplication $K \cdot J$ définie pour les ordinaux. On utilisera les puissances dans les notations. Par exemple $\sum\limits_{\zeta} \zeta = \zeta^2$ et pour tout ordinal α, $\alpha^\omega = \underbrace{\alpha \cdot \alpha \cdot \alpha \ldots}_{\omega fois}$

Deux éléments i et j d'un ordre linéaire J sont dits *consécutifs* s'il n'existe aucun $k \in J$ tel que $i < k < j$. L'élément i est alors appelé le *prédécesseur* de

j et j est le *successeur* de i. Notons qu'un élément $j \in J$ a un prédécesseur si l'ensemble $\{i \mid i < j\}$ admet un plus grand élément.

Lorsqu'un ordre linéaire J admet un plus grand élément (respectivement un plus petit élément), il est noté $max(J)$ (respectivement $min(J)$). Sinon on dit que J est un ordre *limite à droite* (*limite à gauche*).

Une *suite* d'éléments d'un ensemble E indexée par un ordre linéaire J est une application de J dans E et se note souvent $(e_j)_{j\in J}$.

1.4.2 Ordres linéaires dispersés

Un ordre linéaire $(J, <)$ est *dense* si pour tous les éléments j et k de J tels que $j < k$, il existe un élément i dans J tel que $j < i < k$. Par exemple, l'ensemble des nombres rationnels $\mathbb{Q}$ muni de l'ordre habituel $<$ est un ordre linéaire dense. Un ordre est dit *dispersé* s'il ne contient pas de sous-ordre dense. Par exemple les ordres linéaires ω et ζ sont dispersés. Cette notion correspond à celle d'ensemble discret, adaptée aux ordres. Les ordres linéaires dénombrables dispersés généralisent de façon naturelle les ordres finis et les ordinaux. De plus, ces ordres possèdent une jolie caractérisation due à Hausdorff, que nous utiliserons dans toute la suite. Ils sont construits par induction à partir des ordres finis en utilisant les ω-sommes et $-\omega$-sommes.

Notons $\mathcal{N}$ la classe des ordres linéaires finis, $\mathcal{O}$ la classe des ordinaux dénombrables et $\mathcal{S}$ celle des ordres linéaires dénombrables et dispersés.

Théorème 1 *(Hausdorff [20]) Un ordre linéaire dénombrable J est dispersé si et seulement si J appartient à $\bigcup_{\alpha\in\mathcal{O}} V_\alpha$ où les classes V_α sont définies de façon inductive par :*

1. $V_0 = \{\mathbf{0}, \mathbf{1}\}$
2. $V_\alpha = \left\{ \sum_{j\in J} K_j \mid J \in \mathcal{N} \cup \{\omega, -\omega, \zeta\} \text{ et } K_j \in \bigcup_{\beta<\alpha} V_\beta \right\}$.

où $\mathbf{0}$ et $\mathbf{1}$ sont respectivement les ordres ayant zéro et un élément.

La classe V_1 contient les ordres linéaires finis et les ordres $\omega = \sum_\omega \mathbf{1}$, $-\omega$ et ζ. Par

$$\bullet\ \bullet\ \bullet\ \bullet\ \cdots\quad \bullet\ \bullet\ \bullet\ \bullet\ \cdots\quad \bullet\ \bullet\ \bullet\ \bullet\ \cdots\quad \cdots$$

FIG. 1.2 – Représentation de l'ordre ω^2.

$$\cdots\bullet\ \bullet\ \bullet\ \bullet\quad \cdots\bullet\ \bullet\ \bullet\ \bullet\quad \cdots\bullet\ \bullet\ \bullet\ \bullet\quad \cdots$$

FIG. 1.3 – Représentation de l'ordre $\sum_\omega -\omega$.

exemple ω^2 et $\sum_{\omega} -\omega$ appartiennent à V_2 et pour tout entier n, ω^n appartient à V_n. Donc $\omega^\omega = \sum_{n\in\omega} \omega^n$ appartient à V_ω. Chaque ordre linéaire dispersé s'écrit à partir de l'ordre à un élément en utilisant les sommes finies et les sommes indexées par ω, $-\omega$ et ζ ; ces sommes étant imbriquées un nombre ordinal de fois. Les ordinaux sont en particulier des ordres linéaires dispersés. La classe des ordres linéaires dénombrables dispersés est donc fermée par les opérations d'inversion d'ordres et de somme indexée par un ordre linéaire dénombrable dispersé.

Intuitivement, le rang d'un ordre linéaire est le nombre maximum d'imbrications de ω-sommes et de $-\omega$-sommes utilisées pour le définir à partir des ordres finis. Ce nombre est lié à la classe V de Hausdorff de l'ordre, mais pour un ordinal α, la classe V_α ne contient pas exactement les ordres de rang α car le produit fini ne modifie pas le rang. Par exemple, l'ordre $\omega + \omega$ de rang 1 appartient à la classe V_2. Pour tout ordinal α, on définit la classe W_α par :

$$W_\alpha = \left\{ \sum_{j\in J} K_j \mid J \in \mathcal{N} \text{ et } K_j \in V_\alpha \right\}.$$

Ces classes sont strictement intermédiaires à celles de Hausdorff : les inclusions $V_\alpha \subset W_\alpha \subset V_{\alpha+1}$ sont vérifiées pour tout ordinal α. Par exemple, l'ordre $\omega^\alpha + \omega^\alpha$ appartient à W_α mais n'appartient pas à V_α et l'ordre $\omega^{\alpha+1}$ appartient à $V_{\alpha+1}$ mais pas à W_α. Formellement, le *rang* d'un ordre linéaire J est le plus petit ordinal α tel que $J \in W_\alpha$. Les ordres de rang 0 sont les ordres finis.

Dans le but de simplifier les preuves, on définit également les classes U_α très proches des classes V_α pour tout ordinal α. Le fait est que l'ordre ζ n'est pas nécessaire dans la définition inductive des classes puisque $\zeta = -\omega + \omega$. Pour tout ordinal α, on définit

$$U_\alpha = \left\{ \sum_{j\in J} K_j \mid J \in \mathcal{N} \cup \{\omega, -\omega\} \text{ et } K_j \in \bigcup_{\beta<\alpha} W_\beta \right\}.$$

Pour tout ordinal α, $U_\alpha \subset V_\alpha \subset W_\alpha$. D'après le théorème de Hausdorff, les rangs des ordres linéaires dénombrables dispersés sont les ordinaux : $\mathcal{S} = \bigcup_{\alpha\in\mathcal{O}} W_\alpha$. Pour montrer une propriété sur l'ensemble des ordres linéaires dénombrables dispersés, on utilisera souvent une preuve par induction transfinie sur le rang. On montre que la propriété est vraie sur l'ensemble des ordres finis W_0. Ensuite, pour tout ordinal α, on suppose la propriété pour tout ordre de rang $\beta < \alpha$. On prouve alors que la propriété est stable par sommes indexées par ω et $-\omega$ c'est-à-dire qu'elle est vérifiée pour les ordres de U_α. Enfin on vérifie avec le produit fini le cas des ordres de W_α. En effet, la classe W_α contient les sommes finies des ordres appartenant à U_α.

$$W_\alpha = \left\{ \sum_{j\in J} K_j \mid J \in \mathcal{N} \text{ et } K_j \in U_\alpha \right\}$$

Pour tout $\alpha \in \mathcal{O}$, on notera $W'_\alpha = W_\alpha \setminus \{\mathbf{0}\}$ l'ensemble W_α privé de l'ordre contenant 0 élément. Par extension, le rang d'un ensemble d'ordres linéaires est la borne supérieure des rangs de ses éléments.

1.4.3 Coupures

Un *intervalle* d'un ordre linéaire J est un sous-ensemble K tel que quels que soient $j, k \in K$ et $i \in J$, si $j < i < k$ alors $i \in K$. Cette notion correspond à celle de convexité adaptée aux ordres. Une *coupure* de J est un couple (K, L) d'intervalles disjoints de J tels que $J = K \cup L$ et quels que soient $k \in K$ et $l \in L$, $k < l$. L'ensemble de toutes les coupures de J est noté $\hat{J}$. Notons que l'ensemble des coupures d'un ordre linéaire dispersé est toujours un ordre dénombrable.

Exemple 2 L'ordre linéaire $J = \{0, 1, 2\}$ a pour ensemble de coupures $\hat{J} = \{(\emptyset, \{0, 1, 2\}), (\{0\}, \{1, 2\}), (\{0, 1\}, \{2\}), (\{0, 1, 2\}, \emptyset)\}$.

Plus généralement, si J est l'ordre linéaire fini à n éléments, l'ensemble $\hat{J}$ possède $n + 1$ coupures.

L'ensemble $\hat{J}$ des coupures d'un ordre J est naturellement muni d'un ordre total. La relation $<$ est définie pour toutes les coupures $c_1 = (K_1, L_1)$ et $c_2 = (K_2, L_2)$ de $\hat{J}$ par $c_1 < c_2$ si $K_1 \subset K_2$. Cette inclusion implique $L_2 \subset L_1$ donc la définition est symétrique. Notons que l'ordre $\hat{J}$ possède toujours un plus petit élément $(\emptyset, J)$ et un plus grand élément $(J, \emptyset)$. L'ensemble $\hat{J} \setminus \{(\emptyset, J), (J, \emptyset)\}$ est noté $\hat{J}^*$.

Exemple 3 L'ordre $\widehat{\omega}$ contient la première coupure $(\emptyset, \omega)$, les coupures $(\{0, 1, \ldots, n\}, \{n+1, n+2, \ldots\})$ pour tout entier n ainsi que la dernière coupure $(\omega, \emptyset)$. C'est l'ordinal $\omega + 1$.

Plus généralement, si l'ordre J est un ordinal α, l'ordre $\hat{J}$ est l'ordinal $\alpha + 1$. En effet, quelle que soit la coupure $c = (K, L)$ différente de la coupure maximale, l'intervalle L possède un plus petit élément j_c et l'application qui associe l'élément j_c à la coupure c est bijective et préserve l'ordre.

Les ordres linéaires J et $\hat{J}$ s'étendent naturellement en un ordre linéaire sur l'ensemble $J \cup \hat{J}$ de la façon suivante. Pour tout élément $j \in J$ et toute coupure $c = (K, L) \in \hat{J}$, l'ordre $<$ est défini par $j < c$ si $j \in K$ et $c < j$ si $j \in L$. Comme les coupures forment une partition de J, une des deux relations $j < c$ ou $c < j$ est vérifiée. Deux éléments de J ou deux coupures de $\hat{J}$ sont ordonnés par les ordres définis respectivement sur J et $\hat{J}$. Cette relation induit une structure d'ordre linéaire sur l'ensemble $J \cup \hat{J}$. Notons que quels que soient les éléments $j_1 < j_2$ de J, il existe au moins une coupure $c \in \hat{J}$ telle que $j_1 < c < j_2$. De plus si j_1 et j_2 sont deux éléments consécutifs, cette coupure est unique. De façon analogue, quelles que soient les coupures $c_1 < c_2$ de $\hat{J}$, il existe toujours au moins un élément j de J tel que $c_1 < j < c_2$ et celui-ci est unique lorsque les coupures sont consécutives. Ainsi les éléments de J et les coupures de $\hat{J}$ s'intercalent. Notons que l'ordre linéaire $J \cup \hat{J}$ possède aussi un plus petit élément $(\emptyset, J)$ et un plus grand élément $(J, \emptyset)$. L'ordre $J \cup \hat{J}$ privé de ces éléments est noté $J \cup \hat{J}^*$.

Exemple 4 L'ordre $J \cup \hat{J}$ pour $J = \zeta + \zeta$ est illustré à la Figure 4 où chaque élément de $\zeta + \zeta$ est représenté par un point et chaque coupure par une barre verticale. Il existe une coupure entre deux éléments consécutifs de J et entre les deux copies de ζ en plus des coupures minimales et maximales.

$$\Big| \;\cdots \bullet \,|\, \bullet \,|\, \bullet \,|\, \bullet \,|\, \bullet \cdots \;\Big|\; \cdots \bullet \,|\, \bullet \,|\, \bullet \,|\, \bullet \,|\, \bullet \cdots \;\Big|$$

FIG. 1.4 – L'ordre linéaire $J \cup \hat{J}$ pour $J = \zeta + \zeta$.

Pour tout élément d'un ordre linéaire J, il existe deux coupures consécutives c_j^- et c_j^+ de $\hat{J}$ telles que $c_j^- < j < c_j^+$. Ces coupures sont définies par $c_j^- = (K, \{j\} \cup L)$ et $c_j^+ = (K \cup \{j\}, L)$ où $K = \{k \in J \mid k < j\}$ et $L = \{l \in J \mid j < l\}$. Notons également qu'une coupure $c = (K, L) \in \hat{J}$ admet un prédécesseur si K possède un plus grand élément k. Dans ce cas $c = c_k^+$ et son prédécesseur est c_k^-. De même, c possède un successeur si L admet un plus petit élément l. Dans ce cas $c = c_l^-$ et son successeur est c_l^+.

1.5 Mots indexés par des ordres linéaires

Un *alphabet* est un ensemble fini de symboles appelés lettres. Un *mot* x sur un alphabet A indexé par un ordre linéaire J est une application de J dans A qui associe une lettre à chaque élément de J. Il peut être vu comme l'ordre J dont chaque élément est étiqueté par une lettre de l'alphabet. L'ordre linéaire J est appelé la *longueur* de x et noté $|x|$. Le mot indexé par l'ordre à 0 élément est appelé mot vide et noté ϵ. Un mot est souvent noté comme une suite $x = (a_j)_{j \in J}$ où chaque lettre a_j est l'image de l'élément j. Cette notion de mot généralise les définitions de mots déjà considérées dans la littérature. Si l'ordre J est fini de cardinal n, un mot de longueur J est une suite finie $x = a_0 a_1 a_2 \dots a_{n-1}$ de lettres. Un mot $a_0 a_1 a_2 \dots$ de longueur ω est un mot *infini (à droite)* et, symétriquement, un mot $\dots a_{-2} a_{-1} a_0$ indexé par $-\omega$ est *infini à gauche*. Un mot *bi-infini* $\dots a_{-2} a_{-1} a_0 a_1 a_2 \dots$ est un mot de longueur ζ.

Soient $x = (a_i)_{i \in I}$ et $y = (b_j)_{j \in J}$ deux mots indexés respectivement par les ordres linéaires I et J. Le *produit*, ou la *concaténation* de x et y est le mot $(c_k)_{k \in K}$ de longueur $K = I + J$ défini par $c_k = a_k$ si $k \in I$ et $c_k = b_k$ si $k \in J$. Il est noté xy. Plus généralement, pour tout ordre linéaire J, le *J-produit* d'une suite de mots $(x_j)_{j \in J}$ indexée par J est le mot $x = \prod_{j \in J} x_j$ obtenu par concaténation des mots x_j en respectant l'ordre J. Pour tout $j \in J$, notons $K_j = |x_j|$ et $x_j = (a_{k,j})_{k \in K_j}$. Le mot x est de longueur $K = \sum_{j \in J} K_j = \{(k, j) \mid j \in J, k \in K_j\}$ et vaut $x = (a_{k,j})_{(k,j) \in K}$.

Exemple 5 Soit $x = ab$. Le ζ-produit $\prod_\zeta x$ est le mot bi-infini $\dots abababab \dots$.

Exemple 6 Soit le mot infini $x = abbb\ldots$ composé d'un a et d'une suite infinie de b. Le mot $\prod_{-\omega} x$ est le mot $y = (a_j)_{j\in J}$ de longueur $J = \sum_{-\omega} \omega$ tel que pour tout $j \in J$, $a_j = b$ si j a un prédécesseur et $a_j = a$ sinon. Ce mot est de longueur limite à gauche : $\cdots\ abbb\ldots\ abbb\ldots\ abbb\ldots$

Soit $x = (a_j)_{j\in J}$ un mot indexé par un ordre linéaire. Pour tout intervalle non vide I de J, le mot $(a_j)_{j\in I}$ est un *facteur* de x et pour toute coupure (K, L) de $\hat{J}$, les mots $(a_j)_{j\in K}$ et $(a_j)_{j\in L}$ sont appelés respectivement *préfixe* et *suffixe* de x.

Plus généralement, une factorisation d'un mot x est une décomposition de x en une suite de mots dont le produit vaut x. Pour tout ordre linéaire J, une suite $(x_j)_{j\in J}$ de mots est une *J-factorisation* de x si $x = \prod_{j\in J} x_j$.

Etant donné un alphabet A, on note $A^\diamond$ l'ensemble de tous les mots sur l'alphabet A indexés par un ordre dénombrable total et dispersé. On note aussi A^* l'ensemble des mots finis sur A et $A^+ = A^* \setminus \{\epsilon\}$. Pour tout ordre linéaire J, A^J dénote l'ensemble des mots sur A de longueur J. Si E est un ensemble d'ordres linéaires, on note aussi A^E l'ensemble des mots sur A dont la longueur est un ordre appartenant à E. Ainsi, $A^{\mathcal{N}} = A^*$ et $A^{\mathcal{S}} = A^\diamond$. Par extension, le rang d'un mot est le rang de sa longueur. Par exemple, pour tout ordinal α, A^{W_α} est l'ensemble des mots de rang inférieur ou égal à α.

1.6 Théorème de Ramsey

Soit $(x_i)_{i\in\omega}$ une ω-factorisation d'un mot $x = \prod_{i\in\omega} x_i$. Une autre factorisation $(y_i)_{i\in\omega}$ de x est appelée une *factorisation extraite* ou *superfactorisation* si elle est obtenue à partir de la précédente en regroupant les facteurs i.e. s'il existe une suite d'entiers strictement croissante $(k_i)_{i\in\omega}$ telle que $y_0 = x_0 \ldots x_{k_0}$ et $y_i = x_{k_{i-1}+1} \ldots x_{k_i}$ pour tout $i \geq 1$:

$$x = \underbrace{(x_0 x_1 \ldots x_{k_0})}_{y_0}\underbrace{(x_{k_0+1} x_{k_0+2} \ldots x_{k_1})}_{y_1}\ldots$$

Soit $\varphi : A^\diamond \to E$ une application de $A^\diamond$ dans un ensemble fini E. Une ω-factorisation $(y_i)_{i\in\omega}$ d'un mot x est dite *ramseyenne pour* φ s'il existe un élément de e tel que quels que soient les entiers strictement positifs $i \leq j$, $\varphi(y_i y_{i+1} \ldots y_j) = e$.

Les éléments de E sont souvent appelés des couleurs et le théorème de Ramsey établit que tout mot infini admet une factorisation où tous les facteurs sont de la même couleur, excepté peut-être le premier. On donne ici un énoncé particulier de ce théorème avec sa preuve empruntés de [32].

Théorème 2 *(Ramsey, [35]) Soit $\varphi : A^+ \to E$ une application de A^+ dans un ensemble fini E. Tout mot de A^ω admet une ω-factorisation ramseyenne pour φ.*

Preuve. Soit $x = (a_i)_{i\in\omega}$ avec $a_i \in A$ pour tout $i \in \omega$. On définit par récurrence la suite $(n_i, U_i)_{i\in\omega}$ de paires composées d'un indice n_i et d'un ensemble infini d'entiers U_i de la façon suivante. Posons $U_0 = \mathbb{N}$ et $n_0 = 0$. Soit $i \in \omega$. Supposons que U_i soit défini et que $n_i = min(U_i)$. Pour tout élément $e \in E$, on définit l'ensemble

$$T_e = \{n \in U_i \mid \varphi(a_{n_i} a_{n_i+1} \dots a_{n-1}) = e\}.$$

Comme E est fini, il existe au moins un élément $e_i \in E$ tel que l'ensemble T_{e_i} soit infini. On pose alors $U_{i+1} = T_{e_i}$ et $n_{i+1} = min(U_{i+1})$. La suite $n_0, n_1, \dots$ obtenue est illustrée en figure 1.6. Par construction, quels que soient les entiers

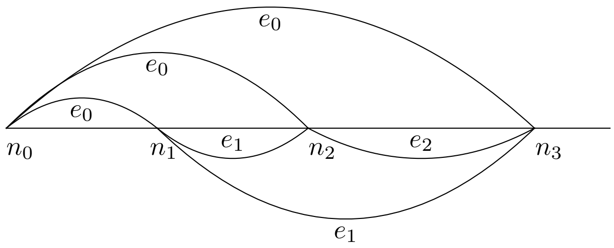


FIG. 1.5 – Suite d'indices.

$i \geq 0$ et $j > 0$, $\varphi(a_{n_i} a_{n_i+1} \dots a_{n_{i+j}-1}) = e_i$. Comme l'ensemble E est fini, il existe au moins un élément $e \in E$ tel que $e = e_i$ pour une infinité d'entiers i. Ces entiers définissent une sous-suite $n_{i_0}, n_{i_1}, \dots$ qui donne la factorisation recherchée $y_j = a_{n_{i_j}} a_{n_{i_j}+1} \dots a_{n_{i_{j+1}-1}}$. $\square$

Il existe de nombreuses versions de ce théorème dont la plupart sont largement détaillées dans [32]. Dans la suite, ce théorème sera largement utilisé et les énoncés les plus appropriés seront reprécisés. La version suivante est l'adaptation du théorème 2 aux mots sur les ordres linéaires dispersés.

Théorème 3 *Soit $\varphi : A^\diamond \to E$ une application de l'ensemble des mots sur A indexés par un ordre linéaire dispersé dans un ensemble fini E. Tout mot de $A^\diamond$ possédant une ω-factorisation admet une superfactorisation Ramseyenne pour φ.*

Preuve. Soit $x \in A^\diamond$ admettant une ω-factorisation $x = \prod_{i\in\omega} x_i$. On reprend la preuve du théorème 2 avec

$$T_{(i,e)} = \{n \in U_i \mid \varphi(x_{n_i} x_{n_i+1} \dots x_{n-1}) = e\}.$$

Avec les mêmes notations, la superfactorisation obtenue est $x = \prod_{j\in\omega} y_i$ avec $y_j = x_{n_{i_j}} x_{n_{i_j}+1} \dots x_{n_{i_{j+1}-1}}$. $\square$

1.7 Définitions algébriques

Nous rappelons dans ce paragraphe des définitions élémentaires d'algèbre universelle notamment celles d'opération, de type algébrique, de morphisme et de congruence.

Une *opération* d'*arité* finie n sur un ensemble E est une application de E^n dans E. Une opération binaire f sur E est *associative* si quels que soient $e_1, e_2, e_3 \in E$, $f(e_1, f(e_2, e_3)) = f(f(e_1, e_2), e_3)$. Dans la suite, nous considérerons des opérations d'arité 1 ayant pour argument une suite indexée par un ordre linéaire. Une telle opération f sur un ensemble E est une application qui associe un élément de E à toute suite d'éléments de E dont la longueur est un ordre linéaire.

Un *type algébrique* C est une classe d'ensembles munis de relations et d'opérations d'arités définies. Une *algèbre* de type C est un élément de C. Soient A et A' deux algèbres de même type C. Un *morphisme* $\varphi : A \to A'$ d'éléments de type C est une application qui préserve les relations et les opérations. Si A et A' sont munis de relations, respectivement notés $\mathcal{R}$ et $\mathcal{R}'$, quels que soient les éléments a_1 et a_2 de A, $a_1 \mathcal{R} a_2$ si et seulement si $\varphi(a_1)\mathcal{R}'\varphi(a_2)$. Si les ensembles A et A' sont équipés d'opérations f et f' d'arité finie n, φ est un morphisme si quels que soient les éléments a_1, a_2,..., a_n de A,

$$\varphi(f(a_1, a_2, \ldots, a_n)) = f'(\varphi(a_1), \varphi(a_2), \ldots, \varphi(a_n)).$$

Par exemple pour une opération binaire noté multiplicativement, $\varphi(e_1 * e_2) = \varphi(e_1) *' \varphi(e_2)$. Dans la cas où les ensembles A et A' sont équipés d'opérations f et f' d'arité 1 et prennant une suite indexée par des ordres linéaires en argument, un morphisme φ doit vérifier pour tout ordre linéaire J et toute suite $(e_j)_{j \in J}$ d'éléments de A,

$$\varphi(f((e_j)_{j \in J})) = f'((\varphi(e_j))_{j \in J}).$$

Dans la suite, nous repréciserons la définition de morphisme dans les structures algébriques que nous utiliserons. Les morphismes seront notés par les lettres grecques φ, ψ, ... Un *isomorphisme* est un morphisme bijectif. Soient A et A' deux éléments d'une algèbre de type C. S'il existe un isomorphisme de A dans A', A et A' sont *isomorphes*. La relation "être isomorphe à" est une relation d'équivalence. Soit $\varphi : A \to A'$ un morphisme d'algèbre de type C. Alors φ *reconnaît* un sous-ensemble X de A si $\varphi^{-1}(\varphi(X)) = X$. On dit que A' reconnaît $X \subseteq A$ s'il existe un morphisme $\varphi : A \to A'$ qui reconnaît X et X est dit *reconnaissable* s'il est reconnu par une algèbre finie. L'algèbre A' est une *sous-algèbre* de A s'il existe un morphisme injectif de A' dans A. A' est *quotient* de A s'il existe un morphisme surjectif de A sur A'. On dit que A *divise* A' et on note $A \prec A'$ si A est quotient d'une sous-algèbre de A'. La relation de division est transitive et deux algèbres se divisant mutuellement sont isomorphes.

Une *congruence* d'algèbre de type C est une relation d'équivalence qui est stable par les opérations. Soit A une algèbre de type donné. Une relation $\sim$ sur A est une congruence si quelle que soit l'opération f sur A d'arité finie n et quels que soient les éléments a_1,..., a_n,a'_1,..., a'_n de A,

$$\forall 1 \leq i \leq n, a_i \sim a'_i \Longrightarrow f(a_1, \ldots, a_n) \sim f(a'_n, \ldots, a'_n).$$

Dans le cas où A est équipé d'une opération f prenant en argument une suite indexée par des ordres linéaires, une congruence doit satisfaire pour toutes les suites $(e_j)_{j \in J}$ et $(e'_j)_{j \in J}$ d'éléments de A indexées par un ordre linéaire J,

$$\forall j \in J, e_j \sim e'_j \Longrightarrow f((e_j)_{j \in J}) \sim f((e'_j)_{j \in J}).$$

L'ensemble des classes d'équivalence noté $E/\sim$ est une algèbre de type C. L'application qui associe à tout élément de E sa classe d'équivalence est un morphisme surjectif de E sur $E/\sim$ donc $E/\sim$ est appelé quotient de E par $\sim$.

Chapitre 2

Automates et expressions rationnelles

La théorie des automates commença en 1956 lorsque Kleene établit l'équivalence entre expressions rationnelles et automates pour les ensembles de mots finis [21]. Depuis lors, ce célèbre théorème de Kleene a été généralisé aux mots infinis [11, 25], bi-infinis [19, 26], aux mots indexés par des ordinaux [13], et enfin aux mots sur les ordres linéaires dispersés [8]. Dans ce chapitre, nous redonnons, pour chaque classe d'ordres linéaires dispersés, les opérateurs rationnels et les automates définissant la classe de langages rationnels correspondante.

Büchi et Muller furent les premiers à définir des automates acceptant des mots de longueur ω. Büchi généralisa ensuite les automates pour les mots sur les ordinaux. Ces derniers possèdent non seulement les transitions classiques des automates sur les mots finis c'est-à-dire de la forme $p \xrightarrow{a} q$ où p et q sont des états et a une lettre de l'alphabet, mais également des transitions de la forme $P \to p$ où p est un état et P un ensemble d'états. Entre deux états successifs d'un chemin, il existe une transition étiquetée par la lettre correspondante. Au niveau des ordinaux limites, un chemin possède une transition $P \to p$ aux points où l'ensemble limite d'états est P. Un chemin possède toujours une fin, même lorsque son étiquette est de longueur limite. Ces automates acceptent les langages définis par les expressions rationnelles construites avec l'union finie, l'étoile, le ω-produit et l'itération sur les ordinaux dénombrables dénotée $\natural$. Ces automates sur les ordinaux dénombrables sont déterminisables ce qui montre que cette classe est fermée par complémentation et qu'elle forme une algèbre de Boole.

Plus récemment, cette théorie a été généralisée de façon naturelle aux mots sur les ordres linéaires dénombrables dispersés par Bruyère et Carton. De façon symétrique, ces automates possèdent des transitions limites à gauche de la forme $P \to p$ et des transitions limites à droite de la forme $p \to P$ où p est un état et P un ensemble d'état. La définition de chemin, basée sur la notion de coupures, généralise celle des chemins de longueurs finies et ordinales. Ils se construisent en insérant un état entre deux lettres consécutives de l'étiquette. Autrement dit, pour tout mot x indexé par un ordre linéaire I, un chemin d'étiquette x est une suite d'états indexés par l'ordre $\hat{I}$ des coupures de I. Lorsqu'un état p n'a pas de prédécesseur (respectivement pas de successeur) il doit exister une transition transition limite à gauche (respectivement à droite) menant de l'ensemble cofinal à gauche vers p (respectivement de p vers l'ensemble cofinal à droite). Aux opérateurs rationnels des mots sur les ordinaux dénombrables, il faut ajouter les opérateurs inverses : le $-\omega$-produit et l'itération ordinale dénombrable inverse $-\natural$ ainsi qu'une sorte d'itérateur sur tous les ordres linéaires dispersés noté $\diamond$. On retrouve un théorème de Kleene pour les ensembles de mots sur des ordres linéaires dispersés.

Lorsqu'on restreint l'ensemble des opérations rationnelles utilisées dans les expressions, on obtient toute une hiérarchie de classes d'ensembles rationnels. En imposant des conditions sur les transitions limites des automates, on retrouve un théorème de Kleene pour chacune des classes. Par exemple, la classe des langages rationnels de rangs finis est définie par les expressions rationnelles utilisant l'union finie, le produit fini, l'étoile, le ω-produit et le $-\omega$-produit. Ces langages sont exactement ceux acceptés par des automates sur les ordres linéaires dont les transitions limites de la forme $p \to P$ ou $P \to p$ vérifient $p \notin P$.

Ce chapitre rappelle les définitions d'ensembles rationnels de mots finis, infinis, bi-infinis, mots sur les ordinaux et mots sur les ordres linéaires dispersés. Pour chaque classe d'ordre, les expressions rationnelles et les automates sont redéfinis.

2.1 Mots finis

Nous définissons dans cette partie la notion de langage rationnel de mots finis sur un alphabet A. Ces ensembles sont définis de façon équivalente par des expressions rationnelles ou par des automates. Pour les exemples, on prendra l'alphabet $A = \{a, b\}$.

2.1.1 Expressions rationnelles

Tout ensemble rationnel peut être décrit par une expression rationnelle formée des lettres de l'alphabet et de symboles représentant les opérateurs. Pour deux ensembles X et $Y \subseteq A^*$, on définit les opérations suivantes :

$$\begin{array}{lrcl} \text{l'union} & X + Y & = & \{z \mid z \in X \text{ ou } z \in Y\} \\ \text{le produit} & X \cdot Y & = & \{xy \mid x \in X \text{ et } y \in Y\} \\ \text{l'itération} & X^* & = & \{\prod_{i=1}^{n} x_i \mid n \in \mathbb{N} \text{ et pour tout } 1 \leq i \leq n \text{ , } x_i \in X\} \end{array}$$

Définition 1 La classe $\mathcal{R}at(A^*)$ des langages *rationnels* de A^* est la plus petite famille contenant l'ensemble $\mathcal{P}(A)$ des parties de A et fermée par union finie, produit fini et itération finie.

Exemple 7 L'expression rationnelle A^*aA^* représente l'ensemble des mots finis sur l'alphabet A contenant la lettre a.

Exemple 8 L'ensemble $X \in \mathcal{R}at(A^*)$ des mots possédant un nombre pair d'occurrences de la lettre a est représenté par l'expression $(ab^*a + b)^*$ et son complémentaire $A^* \setminus X$ par l'expression $b^*a(ab^*a + b)^*$.

Notons que la classe $\mathcal{R}at(A^*)$ des langages rationnels est strictement incluse dans celle des langages de A^*. Par exemple, l'ensemble des mots contenant exactement autant de a que de b n'est pas rationnel.

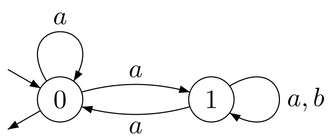
2.1.2 Automates

Nous définissons ici les automates sur les mots finis qui permettent, comme les expressions rationnelles, de décrire de façon finie les ensembles de mots.

Définition 2 Un *automate* est un quintuplet (Q, A, E, I, F) où
- Q est un ensemble fini d'états,
- A est un alphabet fini,
- $E \subseteq (Q \times A \times Q)$ est un ensemble de transitions,
- $I \subseteq Q$ est un ensemble d'états initiaux,
- $F \subseteq Q$ est un ensemble d'états finaux.

Un automate est représenté par un diagramme où les états sont représentés par des cercles et les transitions par des flèches. Les états initiaux sont marqués par une flèche entrante et les états finaux par une flèche sortante.

Exemple 9 L'automate $\mathcal{A} = (Q = \{0, 1\}, A = \{a, b\}$, $E = \{(0, a, 0), (0, a, 1),$ $(1, a, 1), (1, b, 1), (1, a, 0)\}, I = \{0\}, F = \{0\})$ est représenté par le diagramme suivant :



Un automate $\mathcal{A} = (Q, A, E, I, F)$ possède un *chemin* menant de l'état $p \in Q$ à l'état $q \in Q$ s'il existe un entier n et une suite d'états $(q_i)_{0 \leq i \leq n}$ tels que $q_0 = p$, $q_n = q$ et pour tout $0 \leq i \leq n - 1$, il existe $a_i \in A$ tel que $(q_i, a_i, q_{i+1}) \in E$. Le mot $u = a_0a_1 \dots a_{n-1}$ est l'*étiquette* de ce chemin et n sa *longueur*. L'ensemble $\{q_0, q_1, \dots, q_n\}$ est le *contenu* du chemin.

Un chemin est dit *acceptant* s'il mène d'un état initial à un état final. Un mot est *accepté* par un automate s'il est l'étiquette d'un chemin acceptant.

Exemple 10 L'automate de l'exemple 9 accepte exactement les mots commençant et finissant par la lettre a ainsi que le mot vide ϵ.

Les notions d'automates et d'expressions rationnelles sont équivalentes.

Théorème 4 *(Kleene, [21]) Un ensemble de mots finis est rationnel si et seulement s'il est accepté par un automate.*

Exemple 11 L'automate de la figure 2.1 accepte le langage rationnel $(ab^*a+b)^*$ des mots contenant un nombre pair de a.

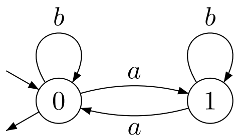


FIG. 2.1 – Automate acceptant le langage $(ab^*a + b)^*$

On note $p \xrightarrow[\mathcal{A}]{a} q$ ou $p \xrightarrow{a} q$ la transition (p, a, q) de l'automate $\mathcal{A}$ et $p \overset{u}{\underset{\mathcal{A}}{\Longrightarrow}} q$ ou $p \overset{u}{\Longrightarrow} q$ l'existence d'un chemin menant de p à q et d'étiquette u dans $\mathcal{A}$.

Un automate $\mathcal{A} = (Q, A, E, I, F)$ est dit *déterministe* si quels que soient $q \in Q$ et $a \in A$, il existe au plus un état $p \in Q$ tel que $(q, a, p) \in E$ et si l'ensemble I des états initiaux est réduit à un singleton. Il est *complet* si quels que soient $q \in Q$ et $a \in A$, il existe au moins un état $p \in Q$ tel que $(q, a, p) \in E$. Un automate est *co-déterministe* si quels que soient $p \in Q$ et $a \in A$, il existe au plus un état $q \in Q$ tel que $(q, a, p) \in E$. Il est *non-ambigu* si pour tous les états p, $q \in Q$ et tout mot $x \in A^*$, il existe au plus un chemin menant de p à q d'étiquette x.

Exemple 12 L'automate de l'exemple 9 n'est pas déterministe puisqu'en partant de l'état 0 et en lisant la lettre a les transitions vers les états 0 et 1 sont autorisées. Il n'est pas complet non plus car il n'existe aucune transition partant de l'état 0 et d'étiquette b.

Théorème 5 *Tout ensemble rationnel de mots finis est accepté par un automate déterministe et complet.*

Exemple 13 L'automate de la figure 2.2 est déterministe, complet et accepte le même langage que l'automate de l'exemple 9.

Soit X un sous-ensemble rationnel de A^* et soit $\mathcal{A}$ un automate complet et déterministe acceptant X. Il suffit d'échanger les états finaux pour obtenir un automate acceptant le complémentaire $A^* \setminus X$.

Exemple 14 L'automate de la figure 2.3 accepte le complémentaire $A^* \setminus (aA^*a + \epsilon) = bA^* + A^*b$ du langage accepté par l'automate de l'exemple 9.

Proposition 6 *La classe des ensembles rationnels de mots finis est fermée par union finie, intersection, produit fini, étoile et complémentation dans A^*.*

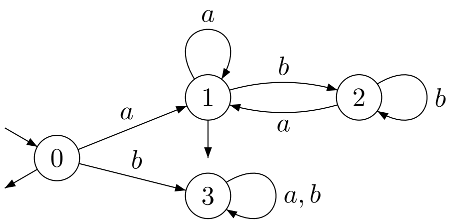


FIG. 2.2 – Un automate déterministe et complet acceptant $aA^*a + \epsilon$.

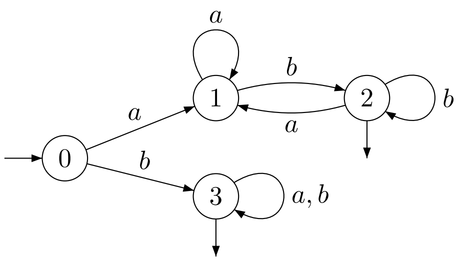


FIG. 2.3 – Un automate acceptant $A^* \setminus (aA^*a + \epsilon)$.

Les automates étudiés dans la suite pourront toujours être complétés en ajoutant un état "puits" où convergent toutes les transitions ajoutées. C'est le cas de l'état 3 de l'automate de la Figure 2.2. Lorsque les automates sont équivalents à des automates déterministes, la fermeture par complémentation de la classe de langages correspondante s'obtiendra donc par la méthode dite de déterminisation. D'autre part, notons qu'une forme canonique existe pour les automates sur les mots finis. Parmi tous les automates déterministes et complets acceptant un ensemble rationnel donné, il en existe un unique, dit minimal qui possède strictement moins d'états que les autres.

2.2 Mots de longueur ω

Les ensembles rationnels de mots de longueur ω ont été principalement étudiés par Büchi [12], Muller [25] et McNaughton [23]. Différents types d'automates ont été définis, largement détaillés dans [32]. Nous rappelons rapidement ceux de Büchi et de Muller par souci historique.

2.2.1 Expressions rationnelles

L'opérateur ω est ajouté à l'ensemble des opérateurs rationnels définis pour les mots finis. Le ω-produit d'un sous-ensemble X de A^* est défini par

$$X^\omega = \{ \prod_{i \in \omega} x_i \mid x_i \in X \setminus \{\epsilon\}\}$$

Définition 3 Un sous-ensemble de A^ω est *ω-rationnel* si et seulement s'il est une union finie d'ensembles de la forme XY^ω où X et Y sont des sous-ensembles rationnel de A^* et de A^+.

On note $\mathcal{R}at(A^\omega)$ la classe des ensembles ω-rationnels.

Exemple 15 L'ensemble $X \in \mathcal{R}at(A^\omega)$ des mots de longueur ω possédant un nombre fini d'occurrences de la lettre b est représenté par l'expression A^*a^ω et son complémentaire $A^\omega \setminus X$ par l'expression $(A^*b)^\omega$.

2.2.2 Automates

Nous redonnons dans ce paragraphe les deux types d'automates acceptant des mots de longueur ω introduits indépendamment par Büchi [12] et Muller [25]. Tous deux sont définis à partir d'automates sur les mots finis et se distinguent par leurs conditions d'acceptation d'un chemin. Soit $\mathcal{A}$ un automate sur les mots finis. Un *chemin de longueur* ω est une suite d'états $(q_i)_{i\in\omega}$ telle que pour tout $i \in \omega$, il existe une lettre $a_i \in A$ telle que $(q_i, a_i, q_{i+1}) \in E$. Le mot $u = a_0a_1\ldots$ est l'*étiquette* de ce chemin. On dit qu'il *passe infiniment souvent* par un état p si pour tout $i \in \omega$, il existe $j > i$ tel que $q_j = p$.

Automates de Büchi :

Un automate de Büchi est un automate $(Q, A, E \subseteq Q\times A\times Q, I \subseteq Q, F \subseteq Q)$ acceptant les étiquettes des chemins passant infiniment souvent par au moins un état de F.

Exemple 16 L'automate de Büchi de la figure 2.4 accepte exactement l'ensemble des mots possédant un nombre fini d'occurrences de la lettre b. Il est complet.

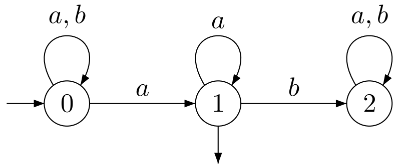


FIG. 2.4 – Automate de Büchi acceptant A^*a^ω

Automates de Muller :

Un automate de Muller est un automate (Q, A, E, I, F) avec $E \subseteq Q \times A \times Q$, $I \subseteq Q$ et où l'ensemble d'états finaux $F \subseteq \mathcal{P}(Q)$ est remplacé par un sous-ensemble des parties de Q. Un chemin est acceptant si l'ensemble des états par lequel il passe infiniment souvent appartient à F.

Exemple 17 L'automate de Muller de la figure 2.5 ayant pour ensemble final $F = \{\{1\}\}$ accepte les mots de A^ω possédant un nombre fini d'occurrences de la lettre b.

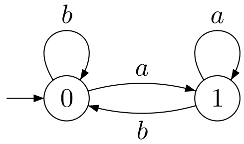


FIG. 2.5 – Automate de Muller acceptant A^*a^ω

McNaughton a montré que ces deux définitions sont équivalentes et que les notions d'automates et d'expressions rationnelles restent équivalentes pour les mots de A^ω.

Théorème 7 *(McNaughton, [23]) Un sous-ensemble de A^ω est ω-rationnel si et seulement s'il est accepté par un automate (de Büchi ou de Muller).*

Si les deux définitions d'automates sont équivalentes, ils ne possèdent pas les même propriétés. En particulier les automates de Muller sont déterminisables ce qui n'est pas le cas des automates de Büchi.

Proposition 8 *(McNaughton, [23]) Tout sous-ensemble rationnel de A^ω est accepté par un automate de Muller déterministe.*

La fermeture par complémentation s'obtient alors de la méthode de déterminisation.

Proposition 9 *La classe des ensembles ω-rationnels forme une algèbre de Boole.*

2.3 Mots bi-infinis

Dans ce paragraphe nous rappelons brièvement la notion de rationnalité pour les ensembles de mots de longueur ζ et renvoyons le lecteur à [32] pour plus de détails.

2.3.1 Expressions rationnelles

L'opérateur $-\omega$ est le symétrique de l'opérateur ω et correspond au produit infini à gauche. Le $-\omega$-produit d'un sous-ensemble X de A^* est défini par

$$X^{-\omega} = \{ \prod_{i \in -\omega} x_i \mid x_i \in X \setminus \{\epsilon\}\}$$

Définition 4 Un sous-ensemble de A^ζ est *ζ-rationnel* si et seulement s'il est une union finie d'ensembles de la forme $X^{-\omega}YZ^\omega$ où X, Y et Z sont respectivement des sous-ensembles rationnel de A^+, A^* et A^+.

On note $\mathcal{R}at(A^\zeta)$ la classe des ensembles ζ-rationnels.

Exemple 18 L'ensemble $X \in \mathcal{R}at(A^\zeta)$ des mots de longueur ζ sur l'alphabet $A = \{a, b\}$ possédant un nombre fini d'occurrences de la lettre b est représenté par l'expression $a^{-\omega}A^*a^\omega$ et son complémentaire $A^\omega \setminus X$ par l'expression $A^{-\omega}(A^*b)^\omega + (A^*b)^{-\omega}A^\omega$.

2.3.2 Automates

Soit $\mathcal{A} = (Q, A, E, I, F)$ un automate fini. Un *chemin de longueur* ζ et d'étiquette $u = \dots a_{-1}a_0a_1 \dots$ est une suite bi-infinie d'états $(q_i)_{i\in\zeta}$ telle que pour tout $i \in \zeta$, la transition (q_i, a_i, q_{i+1}) appartienne à E. Un chemin est *acceptant* si $q_i \in I$ pour une infinité de i négatifs et $q_i \in F$ pour une infinité de i positifs.

Exemple 19 L'automate de la Figure 2.6 accepte l'ensemble des mots de longueur ζ possédant un nombre fini d'occurrences de la lettre b.

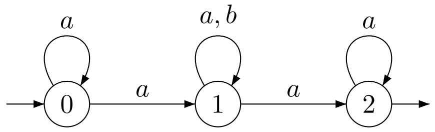


FIG. 2.6 – Automate acceptant le langage $a^{-\omega}A^*a^{-\omega}$

Théorème 10 *Un sous-ensemble de A^ζ est ζ-rationnel si et seulement s'il est accepté par un automate.*

Dans le cas des mots finis et des mots de longueur ω, la propriété de déterminisme est suffisante pour garantir que chaque mot est l'étiquette d'au plus un chemin acceptant. Ce n'est par le cas pour les automates sur les mots bi-infinis pour lesquels la notion de non-ambiguïté est appropriée. Un automate est ζ*-non-ambigu* si tout mot est étiquette d'au plus un chemin acceptant. Le théorème 8 a été reformulé pour les mots bi-infinis par Carton et Michel :

Proposition 11 *[15] Tout ensemble ζ-rationnel est accepté par un automate ζ-non-ambigu.*

2.4 Mots sur les ordinaux

Pour un exposé plus détaillé sur la théorie des langages de mots de longueur plus petite que ω_1, on renvoie à [14, 45, 4, 7].

2.4.1 Expressions rationnelles

Les expressions rationnelles pour les mots transfinis ont été introduites par Wojciechowski [45]. Elles étendent les expressions rationnelles pour les mots finis avec de nouveaux opérateurs. Dans [45], l'itération ordinale est dénotée par le symbole $\sharp$. L'ensemble $X^\sharp$ est défini comme les concaténations de suites de longueur ordinale de mots de X. Dans la suite comme dans [7], on s'intéresse uniquement aux mots indexés par des ordinaux dénombrables. C'est pourquoi on note différemment $X^\natural$ l'itération restreinte aux ordinaux dénombrables. L'ensemble de tous les mots indexés par un ordinal dénombrable est donc noté $A^\natural$. On rappelle ici ces expressions régulières dans un souci de clarté. Pour deux ensembles X et Y de mots sur les ordinaux dénombrables, on définit les opérations suivantes :

l'union	$X + Y$	$=$	$\{z \mid z \in X \text{ ou } z \in Y\}$
le produit	$X \cdot Y$	$=$	$\{xy \mid x \in X \text{ et } y \in Y\}$
l'étoile	X^*	$=$	$\{\prod_{i=1}^{n} x_i \mid n \in \mathbb{N} \text{ et } x_i \in X\}$
ω-produit	X^ω	$=$	$\{\prod_{i \in \omega} x_i \mid x_i \in X \setminus \{\epsilon\}\}$
l'itération ordinale	$X^\natural$	$=$	$\{\prod_{\gamma < \alpha} x_\gamma \mid \alpha < \omega_1 \text{ et } x_\gamma \in X\}$

Définition 5 La classe $\mathcal{R}at(A^\natural)$ des langages *rationnels* de $A^\natural$ est la plus petite famille contenant l'ensemble $\mathcal{P}(A)$ des parties de A et fermée par union finie, produit fini, itération finie, ω-produit et itération ordinale dénombrable.

Exemple 20 L'expression rationnelle $A^\natural a A^\natural$ représente l'ensemble des mots indexés par un ordinal dénombrable contenant la lettre a.

Exemple 21 L'ensemble $(A^\natural)^\omega$ contient l'ensemble des mots de longueur limite à droite et son complémentaire $A^\natural A$ l'ensemble des mots dont la longueur est un ordinal successeur.

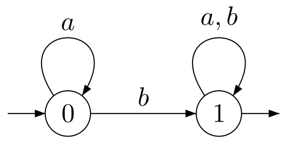
2.4.2 Automates

Dans le cas des automates sur les mots finis, les transitions de la forme $p \xrightarrow{a} q$ où p et q sont des états, sont insuffisantes pour construire des chemins de longueur supérieur à ω. Büchi [13] a été le premier à généraliser l'ensemble de transitions et à définir ainsi des automates pour les mots transfinis. Il définit des transitions de la forme $P \xrightarrow{a} q$ où P est un ensemble d'états et q un état. Nous utiliserons ici une définition des automates dont les transitions limites seront de la forme $P \to p$ i.e. non étiquetées. Les deux définitions sont équivalentes [2, 4] pour les mots sur des ordinaux dénombrables.

Définition 6 Un *automate* sur les ordinaux dénombrables est un quintuplet (Q, A, E, I, F) où Q est un ensemble fini d'états, A est un alphabet fini, $E \subseteq (Q \times A \times Q) \cup \mathcal{P}(Q) \times Q$ est l'ensemble des transitions et $I \subseteq Q$ et $F \subseteq Q$ sont respectivement des ensembles d'états initiaux et finaux.

Les transitions appartenant à $Q \times A \times Q$ sont appelées *transitions successeurs* et celle de $\mathcal{P}(Q) \times Q$ *transitions limites.*

Exemple 22 L'automate représenté en figure 2.7 possède quatre transitions successeurs $0 \xrightarrow{a} 0$, $0 \xrightarrow{b} 1$, $1 \xrightarrow{a} 1$, $1 \xrightarrow{b} 1$ et deux transitions limites $\{0\} \to 1$ et $\{1\} \to 1$.



$\{0\} \to 1$

$\{1\} \to 1$

FIG. 2.7 – Automate acceptant $(a^\omega + a^*b)A^\natural$

Les transitions limites sont définies à partir de la notion de cofinalité qui généralise celle de répétition infinie utilisée pour les automates acceptant les mots de longueur ω. Les états atteints en un point limite γ d'un chemin dépendent des états apparaissant cofinalement avec γ.

Définition 7 Soit α un ordinal et soit $c = (q_\beta)_{\beta<\alpha}$ une suite d'états de longueur α. L'ensemble *cofinal* de c en un ordinal $\gamma < \alpha$ est défini par

$$\mathrm{cof}_\gamma(c) = \{q \mid \forall\beta < \gamma\ ,\ \exists\beta',\ \beta < \beta' < \gamma \text{ et } q_{\beta'} = q\}.$$

Un état q appartient à $\mathrm{cof}_\gamma(c)$ si pour tout ordinal $\beta < \gamma$, il existe un ordinal $\beta < \beta' < \gamma$ tel que $q_{\beta'} = q$. Notons que si γ est un ordinal successeur, $\mathrm{cof}_\gamma(c) = \emptyset$ pour toute suite d'états c de longueur α. Cette définition, analogue à celle de valeur d'adhérence d'un ensemble, est illustrée à travers l'exemple suivant.

Exemple 23 La suite $(q^\omega p)^\omega q$ a pour ensemble cofinal $\{q\}$ aux positions ωk pour tout entier positif k et $\{q, p\}$ à la position ω^2.

$$q\ q\ q\ q\ q\ \cdots \underset{\omega}{p}\ q\ q\ q\ q\ q\ \cdots \underset{\omega 2}{p}\ q\ q\ q\ q\ \cdots \underset{\omega 3}{p}\ \cdots\ \underset{\omega^2}{q}$$

FIG. 2.8 – Suite $(q^\omega p)^\omega q$.

Un chemin d'ensemble cofinal P en un point limite γ doit utiliser une transition limite de la forme $P \to q$.

Définition 8 Soit $\mathcal{A}$ un automate et soit $x = (a_\gamma)_{\gamma<\alpha}$ un mot indexé par un ordinal α. Un *chemin* d'étiquette x dans $\mathcal{A}$ est une suite d'états $c = (q_\beta)_{\beta<\alpha+1}$ de longueur $\alpha + 1$ telle que

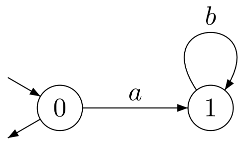
– pour tout $\beta < \alpha$, $q_\beta \xrightarrow{a_\beta} q_{\beta+1}$ est une transition successeur de $\mathcal{A}$,

– pour tout ordinal limite β, $\mathrm{cof}_\beta(c) \to q_\beta$ est une transition limite de $\mathcal{A}$.

Exemple 24 Reprenons l'automate de l'exemple 22. Le mot a^ω est l'étiquette du chemin $0^\omega 1$: la suite 0^ω ayant pour ensemble cofinal $\{0\}$ en ω et le chemin utilise la transition limite $\{0\} \to 1$. De même, le mot b^ω est l'étiquette du chemin $01^\omega 1$ qui termine par la transition limite $\{1\} \to 1$. Les transitions limites peuvent aussi s'imbriquer comme dans le cas du mot $(a^\omega)^\omega$ qui est l'étiquette du chemin $0^\omega(1^\omega)^\omega 1$ de longueur $\omega^2 + 1$.

Remarquons que la longueur d'un chemin est toujours un ordinal successeur. Un chemin possède donc un premier et un dernier état. Un chemin est *acceptant* s'il mène d'un état initial à un état final. Un mot est *accepté* par l'automate s'il est l'étiquette d'un chemin acceptant. Le théorème de Kleene se généralise aux mots sur les ordinaux.

Théorème 12 *(Wojciechowski, [45]) Un ensemble de mots indexés par des ordinaux dénombrables est rationnel si et seulement s'il est accepté par un automate.*



$\{1\} \to 1$

$\{1\} \to 0$

$\{0,1\} \to 0$

Fig. 2.9 – Automate acceptant $(a(b^\natural)^\omega)^\natural$

Un automate sur les ordinaux est *déterministe* s'il possède un unique état initial, si pour tout état p et toute étiquette a il existe au plus un état q tel que $p \xrightarrow{a} q$ soit une transition successeur et si pour tout ensemble d'états P, il existe au plus un état q tel que $P \to q$ soit une transition limite. L'automate de la figure 2.7 est déterministe alors que celui de la Figure 2.9 possédant les transitions $\{1\} \to 1$ et $\{1\} \to 0$ ne l'est pas.

Théorème 13 *(Büchi, [13]) Tout sous-ensemble rationnel de $A^\natural$ est accepté par un automate déterministe.*

Büchi avait déjà montré ce résultat mais de nouveaux algorithmes de déterminisation ont été donnés par Bedon et Carton [7, 4]. La méthode de déterminisation prouve ainsi que la classe des ensembles rationnels de mots indexés par des ordinaux dénombrables est fermée par complémentation.

Corollaire 14 *La classe des ensembles rationnels de $A^\natural$ forme une algèbre de Boole.*

2.4.3 Ordinaux de rang fini

Büchi avait surtout introduit les automates sur les ordinaux pour prouver des résultats de décidabilité de calcul séquentiel. Choueka [18] les a plus étudiés d'un point de vue de la théorie des automates dans le cas des mots indexés par des ordinaux de rang fini. Rappelons que pour tout entier n, un mot indexé par un ordinal est de rang au plus n si les imbrications de ω sont de profondeur inférieure ou égale à n. Les opérations rationnelles correspondantes sont donc l'union, le produit fini et le ω-produit. Les ensembles rationnels de rang fini n sont acceptés par des automates définis par Choueka structurés en $n+1$

niveaux. Soit Q un ensemble d'états. On note $[Q]^0 = Q$ et pour tout $n \in \mathbb{N}$, $[Q]^{n+1} = \mathcal{P}([Q]^n)$ l'ensemble des parties de $[Q]^n$. En notant $[Q]_0^n = \bigcup_{i=0}^{n} [Q]^n$, l'ensemble des transitions des automates de Choueka est un sous ensemble de $[Q]_0^n \times A \times Q$. L'utilisation d'une transition limite en un ordinal limite ne dépend plus de l'ensemble cofinal d'états comme dans les automates précédents mais de l'ensemble des états apparaissant infiniment souvent. On renvoie à [4] pour une description plus détaillée de ces automates qui nous ont paru moins pertinents. En particulier, ils ne semblent pas se généralisent aux ordinaux de rang supérieur. Par contre, on retrouve facilement les ensembles rationnels de rang fini en utilisant les automates précédents sur les ordinaux. Il sera montré bien plus tard qu'un ensemble rationnel de mots sur des ordinaux de rang fini est rationnel si et seulement s'il est accepté par un automate sur les ordinaux dont les transitions limites de la forme $P \to q$ vérifient $q \notin P$ [44, 45, 9].

2.5 Mots sur les ordres linéaires

Bruyère et Carton [8] ont introduit des expressions rationnelles et des automates pour les mots sur les ordres linéaires dénombrables. Il ont généralisé le théorème de Kleene aux ensembles de mots sur les ordres linéaires dénombrables et dispersés. Plus précisément, ils ont défini toute une hiérarchie d'ensembles rationnels [9]. Pour chaque sous-ensemble d'opérateurs rationnels, ils considèrent la classe d'ensembles rationnels correspondant et définissent les transitions des automates acceptant les mêmes langages.

2.5.1 Expressions rationnelles

Les opérations rationnelles utilisées pour les expressions de mots sur les ordres linéaires contiennent bien sûr celles définies pour les mots sur les ordinaux. Trois nouveaux opérateurs sont également nécessaires : le $-\omega$-produit, l'itération ordinale inverse notée $-\natural$ ainsi qu'un dernier opérateur binaire qui est l'équivalent d'une itération sur tous les ordres linéaires dispersés noté $\diamond$. Pour tous les ensembles X et Y de mots indexés par des ordres linéaires dénombrables, on définit les opérations

$$\begin{array}{ll}
X + Y = \{z \mid z \in X \text{ ou } z \in Y\} & \\
X \cdot Y = \{xy \mid x \in X \text{ et } y \in Y\} & X^* = \{\prod_{i=1}^{n} x_i \mid n \in \mathbb{N} \text{ et } x_i \in X\} \\
X^\omega = \{\prod_{i \in \omega} x_i \mid x_i \in X \setminus \{\epsilon\}\} & X^{-\omega} = \{\prod_{i \in -\omega} x_i \mid x_i \in X \setminus \{\epsilon\}\} \\
X^\natural = \{\prod_{i \in I} x_i \mid I \in \mathcal{O} \text{ et } x_i \in X\} & X^{-\natural} = \{\prod_{i \in -I} x_i \mid I \in \mathcal{O} \text{ et } x_i \in X\}
\end{array}$$

$$X \diamond Y = \{\prod_{i \in I \cup \hat{I}^*} z_i \mid I \in \mathcal{S} \setminus \{0\} \text{ et } z_i \in X \text{ si } i \in I \text{ et } z_i \in Y \text{ si } i \in \hat{I}^*\}$$

Rappelons la notation $\hat{I}^* = \hat{I} \setminus \{(\emptyset, I), (I, \emptyset)\}$. Un mot x appartient à $X \diamond Y$ si et seulement s'il existe un ordre linéaire dénombrable dispersé non vide I tel que

x soit le produit d'une suite de mots de longueur $I \cup \hat{I}^*$où chaque mot indexé par un élément de I appartient à X et chaque mot indexé par un élément de $\hat{I}^*$ appartient à Y. Notons que si le mot vide ϵ n'appartient pas à X, il n'appartient

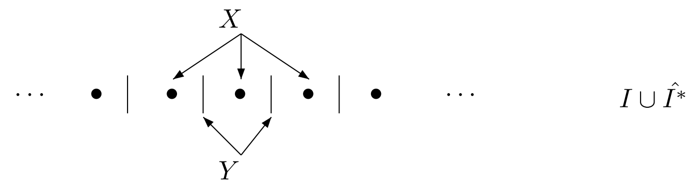


FIG. 2.10 – Opération $X \diamond Y$

pas non plus à l'ensemble $X \diamond Y$. On notera $X^\diamond$ l'expression $(X \diamond \epsilon) + \epsilon$.

$$X^\diamond = \{ \prod_{i \in I} x_i \mid I \in \mathcal{S} \text{ et } x_i \in X \}$$

Ainsi l'expression $A^\diamond$ représente l'ensemble des mots indexés par des ordres linéaires dénombrables et dispersés.

Définition 9 La classe $\mathcal{R}at(A^\diamond)$ des ensembles rationnels de mots sur les ordres linéaires dispersés est la plus petite famille contenant l'ensemble $\mathcal{P}(A)$ des parties de A et fermées par les opérateurs $+$, $\cdot$, $*$, ω, $-\omega$, $\natural$, $-\natural$ et l'opérateur binaire $\diamond$.

Chaque expression rationnelle denote un ensemble de mots défini de façon inductive par les opérateurs rationnels. Un ensemble de mots est rationnel s'il peut être défini par une expression rationnelle. Comme précédemment, les accolades et le point du produit de concaténation finie sont omis. Par exemple, l'expression $(\{a\} + \{a\}\{b\}^\omega)^\natural$ s'écrit $(a + ab^\omega)^\natural$. La classe $\mathcal{R}at(A^\diamond)$ contient par définition les classes $\mathcal{R}at(A^*)$ et $\mathcal{R}at(A^\natural)$.

Exemple 25 L'expression $A^\diamond a A^\diamond$ représente l'ensemble des mots possédant au moins une occurence de la lettre a.

Exemple 26 L'expression $(A^\diamond)^\omega A^\diamond$ représente l'ensemble des mots de $A^\diamond$ possédant un facteur de longueur limite à droite et symétriquement $A^\diamond(A^\diamond)^{-\omega}$ les mots possédant un facteur de longueur limite à gauche. Ainsi $(A^\diamond)^\omega A^\diamond + A^\diamond(A^\diamond)^{-\omega}$ contient les mots de longueur non finie, i.e. c'est le complément de A^* dans $A^\diamond$.

2.5.2 Automates sur les ordres linéaires

Les automates sur les ordres linéaires généralisent ceux définis pour les mots finis et les mots sur les ordinaux. Dans le paragraphe précédent, on a vu que les automates sur les ordinaux sont des automates sur les mots finis possédant des transitions limites de la forme $P \to q$ où P est un ensemble d'états et q un état. Dans le cas des ordres linéaires, on introduit de façon symétrique un ensemble de transitions limites de la forme $q \to P$ où P est un ensemble d'états et q un état.

Définition 10 Un *automate $\mathcal{A}$ sur les ordres linéaires* est un quintuplet $\mathcal{A} = (Q, A, E, I, F)$ où Q est un ensemble fini d'états, A un alphabet fini, $E \subseteq (Q \times A \times Q) \cup (\mathcal{P}(Q) \times Q) \cup (Q \times \mathcal{P}(Q))$ est l'ensemble des transitions et $I \subseteq Q$ et $F \subseteq Q$ sont respectivement les ensembles initiaux et finaux d'états.

Les transitions appartenant à $\mathcal{P}(Q) \times Q$ de la forme $P \to q$ sont appelées *transitions limites à gauche* et celles de $Q \times \mathcal{P}(Q)$ de la forme $q \to P$ *transitions limites à droite.*

Exemple 27 : L'automate $\mathcal{A} = (Q, A, E, I, F)$ représenté en figure 2.11 possède une transition limite à droite $0 \to \{1\}$ et une transition limite à gauche $\{2\} \to 3$.

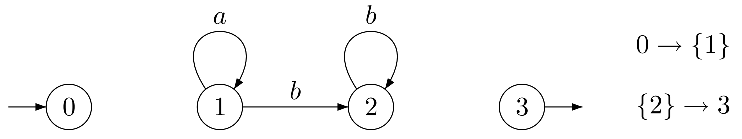


FIG. 2.11 – Automate acceptant le langage $a^{-\omega}b^{\omega}$

Intuitivement, un chemin de longueur limite à gauche (en lisant les étiquettes de la droite vers la gauche) peut être précédé d'une transition limite à droite. Ainsi, un chemin de cet automate menant de l'état 0 à l'état 1 est d'étiquette $a^{-\omega}$. Les transitions limites à gauche sont utilisées comme pour les automates sur les ordinaux et donc cet automate accepte le mot $a^{-\omega}b^{\omega}$.

Considérons un chemin de longueur finie $q_0 \xrightarrow{a_1} q_1 \xrightarrow{a_2} \dots \xrightarrow{a_n} q_n$ d'étiquette $x = a_1 \dots a_n$. Remarquons qu'un état est inséré entre deux lettres consécutives de x. En d'autres termes, à chaque factorisation $x = (a_1 \dots a_k) \cdot (a_{k+1} \dots a_n)$ de x est associé un état q_k. Cette définition de chemin est généralisée dans le cas des automates sur les ordres linéaires de la façon suivante : soit x un mot indexé par un ordre linéaire dispersé I. A chaque factorisation $x = y \cdot z$ de x est associée une coupure de I. Ainsi, un chemin d'étiquette x est un mot sur Q de longueur $\hat{I}$. Il faut distinguer ici les notions de cofinalité à droite et à gauche pour définir précisément les chemins.

Définition 11 Soit I un ordre linéaire dispersé et soit $\gamma = (q_c)_{c \in \hat{I}}$ une suite d'état de longueur $\hat{I}$. Les ensembles *cofinaux à gauche et à droite* de γ en une coupure $c \in \hat{I}$ notés respectivement $\mathrm{cof}_{c^-}(\gamma)$ et $\mathrm{cof}_{c^+}(\gamma)$ sont définis par

$$\mathrm{cof}_{c^-}(\gamma) = \{q \in Q \mid \forall c' < c,\ \exists c'', c' < c'' < c \text{ tel que } q = q_{c''}\}$$

$$\mathrm{cof}_{c^+}(\gamma) = \{q \in Q \mid \forall c', c < c',\ \exists c'', c < c'' < c' \text{ tel que } q = q_{c''}\}$$

Exemple 28 Reprenons l'automate de la figure 2.11. La suite des états $\gamma = 01^{-\omega}2^{\omega}3$ de longueur $\hat{\zeta}$ a pour ensembles cofinaux $\mathrm{cof}_{(\emptyset,\zeta)^+}(\gamma) = \{1\}$ et $\mathrm{cof}_{(\zeta,\emptyset)^-}(\gamma) = \{2\}$.

Comme dans les automates sur les ordinaux, les transitions limites utilisées dans un chemin doivent correspondre aux ensembles cofinaux.

Définition 12 Soient $\mathcal{A} = (Q, A, E, I, F)$ un automate, J un ordre linéaire dispersé et soit $x = (a_j)_{j \in J}$ un mot de longueur J sur A. Un *chemin* γ d'étiquette x dans $\mathcal{A}$ est un mot $\gamma = (q_c)_{c \in \hat{J}}$ sur Q de longueur $\hat{J}$ tel que pour toute coupure (K, L) de $\hat{J}$, on a les deux conditions suivantes :

s'il existe $l \in L$ tel que $(K \cup \{l\}, L \setminus \{l\}) \in \hat{J}$ alors $q_{(K,L)} \xrightarrow{a_l} q_{(K \cup \{l\}, L \setminus \{l\})} \in E$ sinon $q_{(K,L)} \to \mathrm{cof}_{(K,L)^+}(\gamma) \in E$;

s'il existe $k \in K$ tel que $(K \setminus \{k\}, L \cup \{k\}) \in \hat{J}$ alors $q_{(K \setminus \{k\}, L \cup \{k\})} \xrightarrow{a_k} q_{(K,L)} \in E$ sinon $\mathrm{cof}_{(K,L)^-}(\gamma) \to q_{(K,L)} \in E$.

Si une coupure possède un prédécesseur ou un successeur alors les transitions successeurs sont utilisées, sinon le chemin se construit sur les transitions limites à droite ou à gauche.

Exemple 29 Le chemin $01^{-\omega}2^{\omega}3$ de l'automate de la figure 2.11 d'étiquette $a^{-\omega}b^{\omega}$ utilise la transition limite à droite $0 \to \{1\}$ à la coupure $(\emptyset, \zeta)$ et la transition limite à gauche $\{2\} \to 3$ au point $(\zeta, \emptyset)$. Il est représenté par le diagramme suivant où les coupures sont symbolisées par un trait vertical.

$$\begin{array}{cccccccccc}
 & \cdots & a\,| & a\,| & a\,| & b\,| & b\,| & b\,| & b & \cdots \\
0 & \cdots & 1 & 1 & 1 & 2 & 2 & 2 & \cdots & 3 \\
0 \to \{1\} & & & & & & & & & \{2\} \to 3
\end{array}$$

FIG. 2.12 – Chemin d'étiquette $a^{-\omega}b^{\omega}$.

Quel que soit l'ordre linéaire $J \in \mathcal{S}$, l'ordre $\hat{J}$ possède toujours un plus petit élément $(\emptyset, J)$ et un plus grand élément $(\emptyset, J)$. Un chemin possède donc un premier état et un dernier état. De même que pour les ordinaux, un chemin est *acceptant* lorsqu'il mène d'un état initial à un état final et un mot est *accepté* par l'automate s'il est l'étiquette d'un chemin acceptant. On notera $p \overset{x}{\Longrightarrow} q$ l'existence d'un chemin menant de l'état p à q et d'étiquette x. Le *contenu* d'un chemin $\gamma = (q_c)_{c \in \hat{J}}$ est l'ensemble $\{q_c \mid c \in \hat{J}\}$ des états apparaissant dans le chemin et on notera $p \underset{P}{\overset{x}{\Longrightarrow}} q$ l'existence d'un chemin menant de p à q, d'étiquette x et de contenu P.

Exemple 30 Le seul chemin acceptant de l'automate de la figure 2.13 est le chemin $\gamma = (01^{-\omega})^{\omega}2$ d'étiquette $x = (a^{-\omega}b)^{\omega}$. La longueur de x est $|x| = \sum_{\omega} -\omega$. A chaque coupure de la forme $((-\omega) * n, \sum_{\omega} -\omega)$ où n est un entier positif, ce chemin a pour ensemble cofinal à droite $\{1\}$ et utilise la transition $0 \to \{1\}$. Au niveau de la coupure maximale $(|x|, \emptyset)$, l'ensemble cofinal à gauche est $\{0, 1\}$ et la transition $\{0, 1\} \to 2$ termine le chemin.

Bruyère et Carton ont généralisé le théorème de Kleene pour les langages de $A^{\diamond}$.

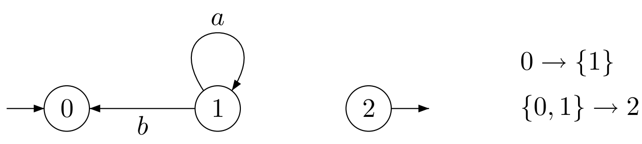


FIG. 2.13 – Automate acceptant le mot $(a^{-\omega}b)^{\omega}$.

	$\cdots$	a	a	a	b	$\cdots$	a	a	a	b		$\cdots$	
0	$\cdots$	1	1	1	0	$\cdots$	1	1	1	0		$\cdots$	2
$0 \to \{1\}$					$0 \to \{1\}$								$\{0,1\} \to 2$

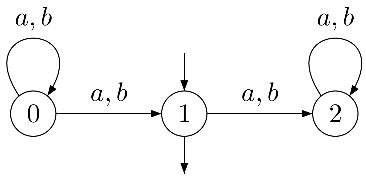
FIG. 2.14 – Chemin d'étiquette $(a^{-\omega}b)^{\omega}$.

Théorème 15 *[8] Un ensemble de mots indexés par des ordres linéaires dénombrables et dispersés est rationnel si et seulement s'il est accepté par un automate.*

La preuve de ce théorème donne la construction d'un automate correspondant à chaque expression rationnelle. La classe des langages rationnels de $A^{\diamond}$ est donc fermée par toutes les opérations rationnelles. Par contre, la fermeture par complémentation dans $A^{\diamond}$ restait un problème ouvert dans [8]. Si les automates sur les ordres linéaires généralisent bien ceux sur les ordinaux, ils ne sont par contre pas toujours équivalents à des automates déterministes. Le principal objet de cette thèse est de montrer que cette classe des langages rationnels sur des ordres linéaires dispersés est fermée par complémentation. L'équivalence entre automates et expressions rationnelles est illustrée à travers quelques exemples.

Exemple 31 Considérons un chemin de l'automate de la figure 2.15 partant de l'état initial 1. S'il emprunte la transition limite $1 \to \{0\}$, il revient à l'état 1 en lisant $A^{-\omega}$. Sinon il peut emprunter les transitions successeurs, passer par l'état 2 et revenir à l'état 1 avec la transition limite $\{2\} \to 1$ en lisant A^{ω}. En concaténant ces chemins un nombre fini de fois, l'automate accepte le langage $(A^{\omega} + A^{-\omega})^{*}$.

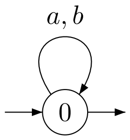
Exemple 32 Concernant, l'exemple de la Figure 2.17, remarquons l'égalité $A^{\diamond} \setminus A^{\diamond}A = (A^{\diamond})^{\omega}$. En effet, un mot de longueur limite à droite ne possède pas de dernière lettre et réciproquement, en utilisant le théorème 1, on peut montrer par induction sur le rang que tout mot sans dernière lettre admet une ω-factorisation.



$1 \to \{0\}$

$\{2\} \to 1$

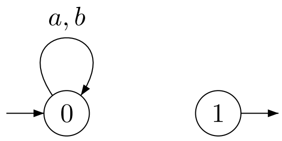
FIG. 2.15 – $(A^\omega + A^{-\omega})^*$.



$\{0\} \to 0$

$0 \to \{0\}$

FIG. 2.16 – Tous les mots $A^\diamond$.

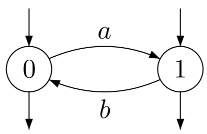


$\{0\} \to 0$

$0 \to \{0\}$

$\{0\} \to 1$

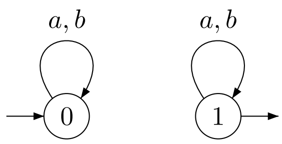
FIG. 2.17 – $A^\diamond \setminus A^\diamond A = (A^\diamond)^\omega$.



$\{0, 1\} \to 0, 1$

$0, 1 \to \{0, 1\}$

FIG. 2.18 – Mots sans facteur aa ou bb.



$0 \to \{1\}$

$\{0\} \to 1$

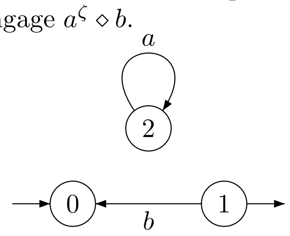
$1 \to \{1\}$

$\{1\} \to 1$

FIG. 2.19 – Mots non finis : $A^\diamond - A^* = A^\diamond(A^\diamond)^{-\omega} + (A^\diamond)^\omega A^\diamond$.

Exemple 33 Considérons un chemin de l'automate de la figure 2.19 partant de l'état initial 0. Pour atteindre l'état final 1, ce chemin doit nécessairement emprunter une des deux transitions limites $0 \to \{1\}$ ou $\{0\} \to 1$. Cet automate accepte donc le langage $A^*(A^\diamond)^{-\omega}A^\diamond + A^\omega A^\diamond$ correspondant à l'ensemble des mots non finis $A^\diamond(A^\diamond)^{-\omega}A^\diamond + A^\diamond(A^\diamond)^\omega A^\diamond$ s'écrivant aussi $A^\diamond(A^\diamond)^{-\omega} + (A^\diamond)^\omega A^\diamond$.

Exemple 34 Considérons l'automate de de la figure 2.20 avec les seules transitions limites $0 \to \{2\}$ et $\{2\} \to 1$. Avec ces transitions, l'automate accepte le langage $a^\zeta(ba^\zeta)^*$. Les deux autres transitions limites $0 \to \{0,1,2\}$ et $\{0,1,2\} \to 1$ permettent d'accepter tous les mots sur $A^\diamond$ composés d'occurrence de a^ζ séparées par un b, c'est-à-dire le langage $a^\zeta \diamond b$.



$$0 \to \{2\}$$
$$0 \to \{0,1,2\}$$
$$\{2\} \to 1$$
$$\{0,1,2\} \to 1$$

FIG. 2.20 – $a^\zeta \diamond b$.

2.5.3 Hiérarchie de langages rationnels

Dans [9], toute une hiérarchie de langages rationnels de mots sur les ordres linéaires a été établie. Elle met en relation les classes d'ordres linéaires, les opérateurs rationnels et des conditions sur les transitions des automates. On a déjà considéré la classe des ensembles rationnels sur les ordinaux et sur les ordinaux de rang finis. Dans chaque cas, les langages rationnels sont décrits par un sous-ensemble d'opérateurs rationnels et par des automates avec restriction sur les transitions (par exemple, que des transitions limites à gauche pour les ordinaux). Sans détailler toute la hiérarchie, on retiendra en particulier que la classe des ensembles rationnels de rangs finis sont décrits par les opérateurs rationnels $+$, $\cdot$, $*$, ω et $-\omega$ et sont acceptés par des automates sur les ordres linéaires dont les transitions limites de la forme $P \to q$ ou $q \to P$ avec $q \notin P$. On notera également que l'opérateur $\diamond$ n'a pas le même pouvoir d'expression lorsqu'il est utilisé en opérateur binaire ou en opérateur unaire : lorsqu'on se restreint au $\diamond$ unaire, la classe de langages obtenus est une sous-classe stricte des langages rationnels. Le lecteur est renvoyé à [9, 44] pour plus de détails.

Chapitre 3

Complémentation des langages de rang fini

Une des principales propriétés des ensembles rationnels est la clôture par complémentation. Cela signifie que pour tout automate $\mathcal{A}$, il existe un autre automate $\mathcal{B}$ acceptant exactement toutes les structures qui ne sont pas acceptées par $\mathcal{A}$. Cette propriété de clôture est vérifiée pour presque toutes les structures : mots finis et infinis, arbres finis et infinis et même pour les mots indexés par les ordinaux. La méthode de déterminisation n'étant pas facile à étendre aux mots infinis, Büchi [12] a utilisé une autre méthode basée sur une congruence sur les mots infinis et sur le théorème de Ramsey. McNaughton a étendu la méthode de déterminisation aux mots infinis [23] prouvant que tout automate de Büchi est équivalent à un automate de Muller déterministe. Cette méthode est réutilisée pour les mots infinis [2] et pour les mots sur les ordinaux de rang fini par Büchi [13] mais elle devient alors très complexe. Dans [9], la question de la complémentation des ensembles rationnels de mots indexés par des ordres linéaires dénombrables et dispersés était laissée ouverte. Dans ce chapitre, on prouve le résultat pour les mots sur les ordres linéaires dénombrables et dispersés de rang fini [16, 17]. Cette classe généralise celle des mots sur les ordinaux de rang fini acceptés par les automates de Choueka. Dans ce cas, les ensembles rationnels sont définis par des expressions rationnelles utilisant uniquement l'union finie, le produit de concaténation fini, le ω-produit et le $-\omega$-produit. Ils sont acceptés par des automates sur les ordres linéaires dont les transitions limites possèdent certaines restrictions. Nous prouvons tout d'abord que tout automate sur les ordres linéaires n'est pas nécessairement équivalent à un automate déterministe. La méthode de déterminisation ne peut donc pas être appliquée. Nous étendons ensuite la preuve de Büchi [12] en utilisant en plus une induction sur le rang pour prouver que les ensembles de mots indexés par des ordres linéaires dispersés de rang fini sont fermés par complémentation.

3.1 Déterminisme

La méthode classique pour obtenir un automate acceptant le complément d'un ensemble de mots finis accepté par un automate $\mathcal{A}$ se fait par déterminisation. Une fois obtenu l'automate déterministe et complet, il suffit d'échanger les états terminaux et non terminaux pour obtenir un automate acceptant le complément. Dans notre cas, cette méthode ne peut pas être appliquée.

Définition 13 *Un automate sur les ordres linéaires* $\mathcal{A} = (Q, A, E, I, F)$ *est* déterministe *s'il possède un unique état initial et si pour tout état* $q \in Q$ *et tout mot* $u \in A^\diamond$*, il existe au plus un chemin partant de* q *et d'étiquette* u*.*

Il s'avère que les automates sur les ordres linéaires ne sont pas toujours déterminisables même dans le cas des rangs finis. Considérons par exemple le langage $(a^{-\omega})^{-\omega}$ qui est accepté par l'automate de la Figure 3.1

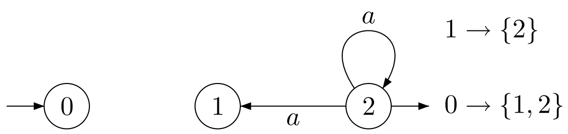


FIG. 3.1 – Automate acceptant le langage $(a^{-\omega})^{-\omega}$

Intuitivement, on sent que cet automate ne peut pas être déterminisé puisque lorsqu'on a lu le mot $a^{-\omega}$, la lecture d'un nouveau a ne change pas le mot accepté.

Proposition 16 *Il n'existe pas d'automate déterministe acceptant le langage* $(a^{-\omega})^{-\omega}$*.*

Preuve. Par l'absurde, supposons l'existence d'un automate déterministe $\mathcal{A}$ acceptant $(a^{-\omega})^{-\omega}$. On dispose d'un chemin γ d'étiquette $(a^{-\omega})^{-\omega}$ que l'on représente de la façon suivante :

$$\gamma : q \ldots \qquad \ldots \qquad \ldots\ q_{-1,-2} \xrightarrow{a} q_{-1,-1} \xrightarrow{a} q_{-1,0} \qquad \ldots\ q_{0,-2} \xrightarrow{a} q_{0,-1} \xrightarrow{a} q_{0,0}$$

Si on supprime la dernière transition $q_{0,-1} \xrightarrow{a} q_{0,0}$ de γ, on obtient un nouveau chemin également d'étiquette $(a^{-\omega})^{-\omega}$ et menant à l'état $q_{0,-1}$. Comme l'automate est déterministe, les deux chemins sont égaux et $q_{0,-1} = q_{0,0}$. De même, pour tout entier m négatif, on obtient $q_{0,m} = q_{0,0}$. Si on supprime toute la dernière partie de γ d'étiquette $a^{-\omega}$ le chemin obtenu mène à l'état $q_{-1,0}$ et est toujours étiqueté $(a^{-\omega})^{-\omega}$ donc $q_{-1,0} = q_{0,0}$. Par récurrence, on obtient ainsi que tous les états du chemin γ sont égaux à $q_{0,0}$ excepté peut-être le premier état q. L'automate $\mathcal{A}$ possède donc la transition limite $q_{0,0} \longrightarrow \{q_{0,0}\}$ ce qui signifie qu'il accepte aussi l'ensemble $a^{-\#}$. $\square$

Afin de contourner cette difficulté, on utilise des méthodes différentes pour montrer la clôture par complémentation. La première, basée sur des classes d'équivalence, permet de résoudre le cas des rangs finis.

3.2 Complément des ensembles de rang fini

Dans le cas des mots finis, Büchi a donné une preuve différente de la clôture des ensembles rationnels par complémentation. Celle-ci n'utilise pas la propriété de déterminisabilité. Elle est basée sur une relation d'équivalence définie pour tout automate $\mathcal{A} = (Q, A, E, I, F)$ sur les mots finis par :

$$u \sim v \text{ si et seulement si } \forall p \in Q, \forall q \in Q,\, p \overset{u}{\Longrightarrow} q \iff p \overset{v}{\Longrightarrow} q$$

Notons que si un mot u est l'étiquette d'un chemin acceptant de $\mathcal{A}$, il en est alors de même pour tout mot équivalent à u. Donc une classe d'équivalence est soit contenue dans le langage L accepté par $\mathcal{A}$ soit disjointe de L. Comme chaque classe d'équivalence est rationnelle, il prouve ainsi que le complément de L est rationnel en tant qu'union finie de classes d'équivalence. Nous généralisons cette preuve au cas des automates sur les ordres linéaires dispersés de rang fini.

Soit $\mathcal{A} = (Q, A, E, I, F)$ un automate sur les ordres linéaires acceptant le langage L. Rappelons qu'un chemin menant de l'état p à l'état q d'étiquette et de contenu P est noté $p \overset{u}{\underset{P}{\Longrightarrow}} q$. Comme les contenus des chemins sont nécessaires pour les transitions limites, on définit la relation d'équivalence $\sim$ par :

$$u \sim v \text{ si et seulement si } \forall p \in Q, \forall q \in Q, \forall P \subseteq Q,\, p \overset{u}{\underset{P}{\Longrightarrow}} q \iff p \overset{v}{\underset{P}{\Longrightarrow}} q.$$

Remarquons tout d'abord que cette relation d'équivalence possède un nombre fini de classes. La classe d'un mot u dépend de l'existence d'un chemin d'étiquette u menant de p à q et de contenu P pour chaque triplet (p, q, P). Comme il y a $n^2 2^n$ triplets de ce type, la relation $\sim$ possède au plus $2^{n^2 2^n}$ classes d'équivalence. On note $\mathcal{C}$ l'ensemble de toutes les classes d'équivalence de $\sim$. Pour toute classe $C \in \mathcal{C}$ et tout entier r, l'ensemble $C \cap A^{W_r}$ est appelé *classe d'équivalence de rang* r même si cet ensemble contient des mots de rang au plus r. Pour tout entier r, on note aussi $\mathcal{C}_r = \{C \cap A^{W_r} | C \in \mathcal{C}\}$ l'ensemble des classes d'équivalence de rang r. Le cardinal de $\mathcal{C}_r$ est inférieur ou égal à celui de $\mathcal{C}$. Comme dans le cas des mots finis, chaque classe est soit incluse dans L soit disjointe de L. On a donc les deux égalités suivantes :

$$L = \bigcup_{C \in \mathcal{C}, C \cap L \neq \emptyset} C \quad \text{et} \quad \bar{L} = A^{\diamond} \setminus L = \bigcup_{C \in \mathcal{C}, C \cap L = \emptyset} C$$

Les mêmes égalités se retrouvent pour les mots de rang inférieur ou égal à r.

$$L \cap A^{W_r} = \bigcup_{C \in \mathcal{C}_r, C \cap L \neq \emptyset} C \text{ et } A^{W_r} \setminus L = \bigcup_{C \in \mathcal{C}_r, C \cap L = \emptyset} C.$$

Pour tout entier r, la famille $\mathcal{C}_r$ contient un nombre fini de classes. Pour montrer que $A^{W_r} \setminus L$ est rationnel, il suffit de prouver que chaque classe $C \in \mathcal{C}_r$ est rationnelle. On montrera ce résultat par induction sur r en s'appuyant sur le théorème de Ramsey. Dans le cas où $r = 0$, le résultat est facile à obtenir. Il suffit de construire un automate sur les mots finis qui mémorise le contenu des chemins dans ses états. L'étape d'induction est basée sur l'idée suivante. Notons $\mathcal{C}_r = \{C_1, ..., C_m\}$ l'ensemble des classes de rang r. On définit des expressions rationnelles utilisant les C_i pour lettres. Une expression élémentaire est une expression de la forme C_i, C_i^{ω} ou $C_i^{-\omega}$ où C_i est une classe de $\mathcal{C}_r$. On note B l'ensemble des expressions élémentaires. Considérons l'ensemble B^* de toutes les expressions obtenues par concaténation finie des expressions élémentaires.

Par exemple supposons qu'il existe deux classes de rang r, $\mathcal{C}_r = \{C_1, C_2\}$. L'ensemble des expressions élémentaires est $B = \{C_1, C_1^{\omega}, C_1^{-\omega}, C_2, C_2^{\omega}, C_2^{-\omega}\}$ et un exemple typique d'élément de B^* serait $C_2^{\omega}C_1C_2^{-\omega}C_1C_2^{-\omega}$. Chaque élément de B^* est une expression rationnelle sur les C_i représentant un ensemble de mots de rang au plus $r+1$. Par un léger abus de langage, on dira qu'un mot appartient à une expression R de B^* lorsqu'il appartient à l'ensemble représenté par R. Les deux lemmes suivants seront nécessaires dans la preuve de la proposition 19. Le lemme 17 montre tout d'abord que tout mot de rang inférieur ou égal à $r + 1$ appartient au moins à une expression de B^*.

Lemme 17 $A^{W_{r+1}} = \bigcup_{R \in B^*} R.$

Preuve. Soit $x \in A^{W_{r+1}}$. Comme la longueur J de x appartient à la classe W_{r+1}, c'est une somme finie d'ordres linéaires appartenant à U_{r+1} : $J = \sum_{i=1}^{n} K_i$ où $\forall 1 \leq i \leq n$, $K_i \in U_{r+1}$. Soit $x = \prod_{i=1}^{n} x_i$ la factorisation associée c'est-à-dire que pour tout $1 \leq i \leq n$ on a $|x_i| = K_i$.

Il suffit de montrer que chaque x_i appartient à une expression R_i de B^* puisque x appartient alors à l'expression $R_1 \dots R_n$ qui est dans B^*.

On prouve que tout facteur x_i de x appartient à une expression de B^* :

- Cas 1 : Le rang de x_i est inférieur ou égal à r, $|x_i| \in W_r$. Dans ce cas x_i appartient à une classe d'équivalence C de rang r et $C \in B^*$.
- Cas 2 : x_i est un ω-produit de mots de rang r :
  L'ordre K_i peut être décomposé $K_i = \sum_{j \in \omega} K_{i,j}$ où $K_{i,j} \in W_r$ pour tout $j \in \omega$. Donc $x_i = \prod_{j \in \omega} x_{i,j}$ avec $x_{i,j} \in A^{W_r}$. Comme le nombre de classes d'équivalence de $\mathcal{C}_r$ est fini, il existe d'après le théorème 3, une factorisation extraite ramseyenne $x_i = \prod_{j \geq 0} x'_{i,j}$ où tous les $x'_{i,j}$ d'indice $j \geq 1$ appartiennent à une même classe d'équivalence C de $\mathcal{C}_r$. En notant C' la classe d'équivalence de $x_{i,0}$, on a $x_i \in C'C^{\omega}$ et $C'C^{\omega} \in B^*$.
- Cas 3 : x_i est un $-\omega$-produit de mots de rang r :
  De façon symétrique, on montre que x_i appartient à une expression de la forme $C^{-\omega}C'$ où C,$C' \in \mathcal{C}_r$.

$\square$

Dans le lemme 18, on montre que deux mots appartenant à la même expression R de B^* sont $\sim$-équivalents. Cela signifie que tout ensemble représenté par une expression R de B^* est inclus dans une seule $\sim$-classe.

Lemme 18 *Si deux mots x et y de rangs inférieurs ou égaux à $r+1$ appartiennent à une même expression R de B^*, alors ils satisfont $x \sim y$.*

Preuve. Soit R une expression de B^*. Soient x et y deux mots satisfaisant $x \in R$ et $y \in R$. On montre d'abord que $x \sim y$ lorsque R est une expression élémentaire de B.

- Cas 1 : $R = C$ pour une certaine classe C de $\mathcal{C}_r$.
  Comme x et y appartiennent à C, on obtient $x \sim y$.
- Cas 2 : $R = C^\omega$ pour une certaine classe C de $\mathcal{C}_r$.
  Soient $x = \prod_{i \in \omega} x_i$ et $y = \prod_{i \in \omega} y_i$ les factorisations satisfaisant pour tout $i \in \omega$, $x_i, y_i \in C$.
  Soit $\gamma : p \overset{x}{\underset{P}{\Longrightarrow}} q$ un chemin de $\mathcal{A}$ d'étiquette x. Le chemin γ peut être décomposé selon la factorisation de x en un chemin
  $$p = p_0 \overset{x_0}{\underset{P_0}{\Longrightarrow}} p_1 \overset{x_1}{\underset{P_1}{\Longrightarrow}} p_2 \ \ldots \ q$$
  avec $P_1 \cup P_2 \cup \ldots = P$ et finissant par une transition limite $P' \longrightarrow q$.
  Pour tout $i \in \omega$, on a $x_i \sim y_i$ donc $p_i \overset{y_i}{\underset{P_i}{\Longrightarrow}} p_{i+1}$. Finalement $p \overset{y}{\underset{P}{\Longrightarrow}} q$ est un chemin de $\mathcal{A}$ ce qui montre que $x \sim y$.
- Cas 3 : $R = C^{-\omega}$ pour une certaine classe C de $\mathcal{C}_r$. De façon symétrique au cas précédent, on obtient $x \sim y$.

Supposons à présent que R est un produit fini d'expressions élémentaires de B : $R = \prod_{i \geq o}^{n} R_i$.

Il existe deux factorisations $x = \prod_{i=1}^{n} x_i$ et $y = \prod_{i=1}^{n} y_i$ telles que pour tout $1 \leq i \leq n$, $x_i \in R_i$ et $y_i \in Ri$.

Supposons l'existence d'un chemin $p \overset{x}{\underset{P}{\Longrightarrow}} q$ dans $\mathcal{A}$. Décomposons ce chemin selon la factorisation précédente de x :

$$p = q_0 \overset{x_1}{\underset{P_1}{\Longrightarrow}} q_1 \overset{x_2}{\underset{P_2}{\Longrightarrow}} q_2 \ \ldots \ \overset{x_n}{\underset{P_n}{\Longrightarrow}} q_n = q$$

Comme les R_i sont des expressions élémentaires, on a $x_i \sim y_i$ pour tout $1 \leq i \leq n$. Donc $q_{i-1} \overset{y_i}{\underset{P_i}{\Longrightarrow}} q_i$ pour tout $1 \leq i \leq n$ et il s'ensuit que $p \overset{y}{\underset{P}{\Longrightarrow}} q$.

On peut conclure que $x \sim y$. $\square$

D'après les lemmes 17 et 18, on sait que chaque classe C de $\mathcal{C}_{r+1}$ satisfait

$$C = \bigcup_{R \in B^*, C \cap R \neq \emptyset} R$$

Cependant, ce n'est pas une expression rationnelle puisque il y a une infinité d'expressions R incluses dans C. Dans la proposition suivante, on montre que l'ensemble des expressions rationnelles incluses dans une quelconque classe C peut être décrit par une expression rationnelle sur les expressions élémentaires.

Proposition 19 *Chaque classe d'équivalence de $\mathcal{C}_r$ est rationnelle.*

Preuve. La preuve se fait par induction sur le rang r.

- $r = 0$:
  Considérons les classes de mots finis. Etant donnés $p \in Q$, $q \in Q$ et $P \subseteq Q$, on note $L_{p,q,P}$ l'ensemble des étiquettes des chemins de $\mathcal{A}$ de longueur finie menant de p à q et de contenu P. On peut facilement montrer que $L_{p,q,P}$ est rationnel en construisant un automate sur les mots finis mémorisant les contenus des chemins. Donc toute $\sim$-classe de mots finis est rationnelle en tant que combinaison booléenne des ensembles $L_{p,q,P}$.
- $r > 0$:
  Soit $\mathcal{C}_r = \{C_1, C_2, ..., C_m\}$ l'ensemble des classes d'équivalence des mots de rang inférieur ou égal à r. Par hypothèse d'induction, toute classe de rang r est rationnelle :

  $$\forall C \in \mathcal{C}_r, C \in Rat(A^{W_r})$$

  Soit $B = \{C_1, C_1^{\omega}, C_1^{-\omega}, ..., C_m, C_m^{\omega}, C_m^{-\omega}\}$. A partir de $\mathcal{A}$, on construit un automate $\mathcal{B}$ qui relie chaque expression R de B^* avec les mots de $A^{W_{r+1}}$ appartenant à R. Ensuite les classes de $\mathcal{C}_{r+1}$ seront définies comme des ensembles rationnels d'éléments de B^*.
  Soit $\mathcal{B} = (Q \times \mathcal{P}(Q), B, E^{'}, I^{'}, F^{'})$ l'automate défini par :

  $$E' = \{(p, S) \xrightarrow{R} (p', S \cup T) \mid \exists x \in R \text{ tel que } p \underset{T}{\overset{x}{\Longrightarrow}} p' \text{ in } \mathcal{A}\}$$

  $$\text{avec } I' = \{(p, \emptyset) \mid p \in Q\} \text{ et } F' = Q \times \mathcal{P}(Q)$$

  Remarquons tout d'abord que l'étiquette d'un chemin de $\mathcal{B}$ est un élément de B^*. Donc c'est une expression représentant un ensemble de mots de $A^{W_{r+1}}$. Soit $L_{(p,\emptyset),(q,P)}$ l'ensemble des étiquettes des chemins menant de l'état $(p, \emptyset)$ à l'état (q, P) dans $\mathcal{B}$ et soit $L^{r+1}_{p,q,P}$ l'ensemble des étiquettes des chemins de $\mathcal{A}$ menant de p à q, de contenu P et de rang inférieur ou égal à $r + 1$. On va montrer qu'un mot x de A^{W_r} appartient à $L^{r+1}_{p,q,P}$ si et seulement si x appartient à une expression R de $L_{(p,\emptyset),(q,P)}$. Donc l'ensemble $L^{r+1}_{p,q,P}$ est défini comme l'ensemble des mots appartenant à une étiquette de $L_{(p,\emptyset),(q,P)}$. Le résultat en découle puisque une classe C de $\mathcal{C}_{r+1}$ est définie comme combinaison booléenne des ensembles rationnels $L_{(p,\emptyset),(q,P)}$ de mots finis.
  On montre à présent que pour tous $p \in Q$, $q \in Q$, $P \subseteq Q$ et $R \in B^*$, $(p, \emptyset) \overset{R}{\Longrightarrow} (q, P)$ dans $\mathcal{B}$ si et seulement s'il existe un mot x dans R tel que $p \underset{P}{\overset{x}{\Longrightarrow}} q$ dans $\mathcal{A}$.

Soit x un mot de rang inférieur ou égal à $r+1$. D'après le lemme 17, il existe $R \in B^*$ telle que $x \in R$. Par construction de $\mathcal{B}$, l'existence d'un chemin $p \underset{P}{\overset{x}{\Longrightarrow}} q$ dans $\mathcal{A}$ implique que $(p, \emptyset) \overset{R}{\Longrightarrow} (q, P)$ dans $\mathcal{B}$.

Réciproquement, soit $R \in B^*$ et soit $x \in R$. Supposons que $(p, \emptyset) \overset{R}{\Longrightarrow} (q, P)$. Par construction de $\mathcal{B}$, il existe $y \in R$ tel que $p \underset{P}{\overset{y}{\Longrightarrow}} q$ dans $\mathcal{A}$. De plus, d'après le lemme 18, on a $x \sim y$ donc $p \underset{P}{\overset{x}{\Longrightarrow}} q$ dans $\mathcal{A}$.

□

On est à présent capable de montrer que l'ensemble des langages rationnels de rang fini est clos par complémentation.

Théorème 20 *Soit L un langage rationnel de mots indexés par des ordres linéaires et soit r un entier positif. Le complément $A^{W_r} \setminus L$ de L dans l'ensemble des mots de rang inférieur ou égal à r est rationnel.*

Preuve. Soit $\mathcal{A}$ un automate sur les ordres linéaires acceptant un langage L et soit r un rang fini. Soit $\mathcal{C}_r$ l'ensemble des classes d'équivalence de rang r selon $\mathcal{A}$. D'après la proposition 19, on sait que chaque classe $\mathcal{C}_r$ est rationnelle. De plus, si on considère la définition de $\sim$, on voit que si un mot u est l'étiquette d'un chemin acceptant de $\mathcal{A}$, il en est de même pour tout mot équivalent à u. Donc toute classe d'équivalence est soit incluse dans L soit disjointe de L. On en déduit une expression rationnelle de $A^{W_r} \setminus L$ comme union finie de classes de $\mathcal{C}_r$:

$$A^{W_r} \setminus L = \bigcup_{C \in \mathcal{C}_r, C \cap L = \emptyset} C.$$

□

L'exemple suivant illustre la construction des expressions rationnelles des classes d'équivalence :

Exemple 35 Soit $\mathcal{A} = (Q, A, E, I, F)$ l'automate de la figure 3.2 acceptant le langage $L = (a^\zeta)^\#$. On cherche une expression rationnelle des classes d'équivalence de rang 1.

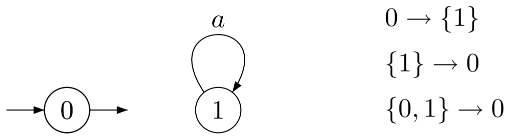


FIG. 3.2 – Automate sur les ordres linéaires acceptant le langage $(a^\zeta)^\#$.

Soit $\mathcal{C}_0$ l'ensemble des classes d'équivalence des mots finis. L'automate a deux classes d'équivalence : $\mathcal{C}_0 = \{a^+, \epsilon\}$. Définissons l'ensemble $B = \{a^+, a^\omega, a^{-\omega}, \epsilon\}$.

L'automate sur les mots finis correspondant $\mathcal{B}$ est donné à la figure 3.3. Comme les états terminaux ne sont pas nécessaires ici, ils n'ont pas été marqués sur la figure.

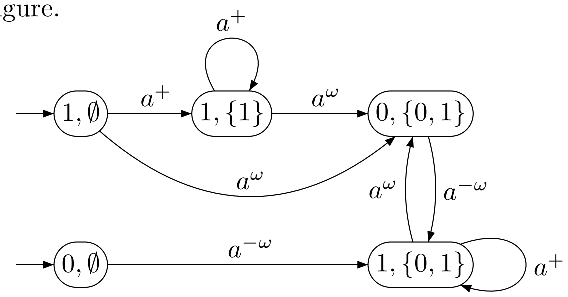


FIG. 3.3 – Automaton $\mathcal{B}$

On en déduit les $\sim$-classes de rang 1 suivantes : $\mathcal{C}_1 = \{C_1, C_2, C_3, C_4, C_5, C_6\}$.

$$\begin{array}{lll} C_0 = \epsilon, & C_1 = a^+, & C_2 = (a^\zeta)^* a^{-\omega}, \\ C_3 = a^\omega (a^\zeta)^*, & C_4 = (a^\zeta)^+, & C_5 = a^\omega (a^\zeta)^* a^{-\omega} \end{array}$$

et $C_6 = (a + a^\omega + a^{-\omega})^* \cdot (a^\omega \cdot a + a \cdot a^{-\omega}) \cdot (a + a^\omega + a^{-\omega})^*$

où la dernière classe contient les mots qui ne sont pas étiquette d'un chemin de $\mathcal{A}$. Considérons l'automate $\mathcal{B}$ ayant tous ses états terminaux. Comme c'est un automate sur les mots finis, C_7 est obtenu grâce à son complément.

Exceptées les classes C_0 et C_4 qui sont incluses dans L, toutes les autres classes de $\mathcal{C}_1$ sont disjointes de L. On obtient les expressions rationnelles suivantes de L et de son complément dans l'ensemble des mots de rang inférieur ou égal à 1 :

$$L \cap A^{W_1} = C_0 + C_4$$

$$A^{W_1} \setminus L = C_1 + C_2 + C_3 + C_5 + C_6.$$

Chapitre 4

Equivalence algébrique

Un semigroupe est un ensemble muni d'une opération interne binaire associative. Les semigroupes permettent de définir les ensembles rationnels de mots finis de façon équivalente aux automates et expressions rationnelles. Un ensemble de mots finis est rationnel si et seulement s'il est reconnaissable, c'est-à-dire s'il existe un semigroupe fini S et un morphisme de semigroupe de A^+ dans S le reconnaissant. De plus, il est possible d'associer à chaque ensemble rationnel un semigroupe minimal le reconnaissant : le semigroupe syntaxique. Cette approche algébrique est très utile dans l'étude des ensembles rationnels. Non seulement elle est utilisée pour montrer de nombreux résultats de décidabilité, mais surtout l'étude des propriétés algébriques des semigroupes syntaxiques a permis de classifier les ensembles rationnels. Les semigroupes n'autorisant pas de produit infinis, ils ne sont pas adaptés à l'étude des langages rationnels de mots infinis. Wilke [42], Perrin et Pin [31] ont alors introduit une structure algébrique dans laquelle le ω-produit est autorisé : les ω-semigroupes. Un ensemble de mots de longueur ω est rationnel si et seulement s'il est reconnu par un ω-semigroupe fini. A tout ensemble rationnel de mots infinis, on peut ainsi associer un objet canonique, l'ω-semigroupe syntaxique, et généraliser la classification des ensembles rationnels. Cette approche a également été généralisée aux mots sur les ordinaux par Bedon et Carton [3, 4, 7]. Un ω_1-semigroupe S est un ensemble muni d'une opération interne qui autorise les produits de toute suite d'éléments de S indexée par un ordinal dénombrable. En montrant l'équivalence entre ω_1-semigroupes finis et automates sur les ordinaux, ils donnent en particulier un algorithme de déterminisation des automates sur les ordinaux et une preuve de la fermeture par complémentation des ensembles rationnels sur les ordinaux dénombrables.

Dans ce chapitre, cette approche algébrique est généralisée aux ensembles de mots indexés par des ordres linéaires dispersés. On étend les semigroupes et ω_1-semigroupes aux $\diamond$-semigroupes dont le produit est défini pour toute suite indexée par un ordre linéaire dénombrable et dispersé. On montre que, lorsqu'ils sont finis, ces $\diamond$-semigroupes sont équivalents aux automates sur les ordres linéaires dénombrables et dispersés. Par analogie avec le cas des mots finis et des mots sur les ordinaux, on montre qu'un $\diamond$-semigroupe canonique minimal peut être associé à chaque ensemble rationnel : le $\diamond$-semigroupe syntaxique. On obtient alors qu'un ensemble sur les ordres linéaires dénombrables dispersés est rationnel si et seulement si son $\diamond$-semigroupe syntaxique est fini [36]. Cette caractérisation algébrique prouve la fermeture par complémentation mais elle donne également la possibilité d'étendre les résultats de classification de Schützenberger et d'Eilenberg à partir des propriétés du $\diamond$-semigroupe syntaxique. Le fait que les ordres soient dispersés et dénombrables sont des hypothèses nécessaires. Büchi a déjà montré que la classe des ensembles de mots transfinis de longueur supérieure à ω_1 (le plus petit ordinal non dénombrable) acceptés par un automate n'est pas fermée par complémentation. On peut montrer que l'ensemble des mots sur les ordres linéaires dispersés n'est pas rationnel en tant que sous-ensemble des mots sur les ordres linéaires alors que son complément l'est. Notre preuve de la fermeture par complémentation est effective. Etant donné un automate $\mathcal{A}$, elle construit un automate $\mathcal{B}$ acceptant exactement les mots qui ne sont pas acceptés par $\mathcal{A}$. Elle donne une autre preuve de la décidabilité de l'équivalence entre ces automates.

Ce chapitre est divisé en trois parties. Dans la première, les ensembles rationnels de mots finis, infinis et ordinaux dénombrables sont redéfinis comme les ensembles reconnaissables. La deuxième partie est consacrée à des définitions algébriques avancées sur les semigroupes qui seront utilisées pour les preuves, notamment les relations de Green. Enfin, la dernière partie généralise la notion de reconnaissabilité aux mots sur les ordres linéaires. On définit tout d'abord la structure algébrique des $\diamond$-semigroupes. On montre que lorsque un $\diamond$-semigroupe S est fini, son produit est défini par un produit fini et par deux fonctions compatibles à droite et à gauche avec S. Ceci permet de décrire le produit d'arité infinie de façon finie. On utilise ensuite une méthode classique pour prouver que tout langage rationnel est reconnu par un $\diamond$-semigroupe fini : les éléments du $\diamond$-semigroupe construit sont des matrices mémorisant les contenus des chemins. Vient alors le résultat le plus difficile de cette thèse : pour construire un automate acceptant un ensemble reconnaissable donné, on utilise une induction sur la $\mathcal{D}$-classe en plus de l'induction sur le rang. De cette preuve d'équivalence, on obtient la fermeture par complémentation. Enfin, on définit la congruence syntaxique et on montre l'effectivité du $\diamond$-semigroupe syntaxique associé à un ensemble rationnel.

4.1 Reconnaissabilité

Ce paragraphe fait la synthèse des structures algébriques reconnaissant les ensembles rationnels : semigroupes, ω-semigroupes, ω^n-semigroupes et ω_1-semigroupes donnent respectivement une caractérisation algébrique des ensembles rationnels de mots finis, infinis et ordinaux. Pour chaque classe, le semigroupe syntaxique correspondant est redéfini.

4.1.1 Semigroupes

Un *semigroupe* est un ensemble S muni d'une opération binaire associative. L'opération est notée multiplicativement $s \cdot t$ ou souvent st et appelée produit. Par exemple pour tout alphabet A, l'ensemble A^+ muni du produit de concaténation est un semigroupe. Un élément 1 d'un semigroupe S est *neutre* si pour tout élément s de S, $s \cdot 1 = 1 \cdot s = s$. Un *monoïde* est un couple $(M, 1)$ formé d'un semigroupe M et d'un neutre pour son opération noté 1. Par exemple l'ensemble A^* est un monoïde ayant pour neutre le mot vide. Le semigroupe S auquel un élément neutre a été ajouté est noté S^1. Un élément 0 de S est un *zéro* si pour tout élément s de S, $s \cdot 0 = 0 \cdot s = 0$.

Exemple 36 On notera U_1 le semigroupe $U_1 = \{0, 1\}$ possédant un zéro et un neutre. Son produit correspond à l'opérateur booléen : $00 = 01 = 10 = 0$ et $11 = 1$.

Etant donnés deux semigroupes S et T, on rappelle qu'une application $\varphi : S \to T$ est un *morphisme de semigroupe* si pour tout $s, t \in S$, $\varphi(st) = \varphi(s)\varphi(t)$. Pour les monoïdes, un morphisme $\varphi : S \to T$ doit aussi vérifier $\varphi(1_S) = 1_T$. Dans la suite, on identifiera deux semigroupes isomorphes.

Lorsqu'ils sont finis, les semigroupes permettent de définir les ensembles rationnels de façon équivalente aux automates et expressions rationnelles. Rappelons qu'un semigroupe S *reconnaît* un langage $X \subseteq A^+$ s'il existe un morphisme de semigroupe $\varphi : A^+ \to S$ et un ensemble $P \subseteq S$ tels que $X = \varphi^{-1}(P)$. Notons que si φ est une application de A dans un semigroupe S, il existe un unique morphisme $\overline{\varphi} : A^+ \to S$ qui prolonge φ (i.e. pour tout $a \in A$, $\overline{\varphi}(a) = \varphi(a)$).

Exemple 37 L'ensemble $X = A^*aA^*$ des mots contenant la lettre a est reconnu par le semigroupe U_1. En effet, en définissant le morphisme $\varphi : A^+ \to S$ prolongeant l'application $\varphi : A \to S$ définie par $\varphi(a) = 0$ et $\varphi(b) = 1$, on obtient $X = \varphi^{-1}(\{0\})$.

Les ensembles rationnels sont alors redéfinis comme ceux reconnus par un semigroupe **fini**.

Théorème 21 *Un ensemble de mots finis est rationnel si et seulement s'il est reconnaissable.*

De même que pour tout ensemble rationnel X de mots finis il existe un automate minimal acceptant X, on peut également définir un semigroupe minimal reconnaissant X.

Définition 14 Pour tout langage X de A^+, la relation $\sim_X$ est définie pour tous mots x, y de A^+ par :

$$x \sim_X y \text{ si et seulement si } \forall u\ ,\ v \in A^*\ ,\ uxv \in X \iff uyv \in X$$

Lorsque X est rationnel, la relation $\sim_X$ est une congruence de semigroupe appelée *congruence syntaxique* de X. Le semigroupe $A^+/\sim_X$ des classes d'équivalence de $\sim_X$ est appelé *semigroupe syntaxique* de X et noté $S(X)$. C'est le plus petit semigroupe reconnaissant X au sens de la division.

Théorème 22 *Un langage X de A^+ est reconnu par un semigroupe S si et seulement si $S(X)$ divise S.*

Les ensembles rationnels sont alors caractérisés par les propriétés de leurs semigroupes syntaxiques. Le théorème précédent montre en particulier qu'un ensemble de mots finis est rationnel si et seulement si son semigroupe syntaxique est fini.

Exemple 38 Le semigroupe syntaxique de l'ensemble $X = (ab)^+$ est $S(X) = \{s, e, t, f, 0\}$ dont les éléments représentent respectivement les $\sim_X$-classes $(ab)^*a$, $(ab)^+$, $b(ab)^*$, $(ba)^+$ et $A^+aaA^+ \cup A^+bbA^+$. Le produit est défini par $ee = e$, $ff = f$, $st = e$, $ts = f$, $es = s$, $te = t$, $ft = t$ et $sf = s$, les autres produits valant tous 0. Le semigroupe $S(X)$ est fini et reconnaît X : Soit $\varphi : A^+ \to S(X)$ le morphisme défini sur A par $\varphi(a) = s$ et $\varphi(b) = t$. Il vient $X = \varphi^{-1}(e)$.

4.1.2 ω-semigroupes

Les semigroupes constituent un très bon outil pour l'étude des ensembles de mots finis mais ils ne sont pas adaptés pour travailler sur les mots de longueur ω. Pécuchet [27, 28] fût le premier à étudier les ensembles reconnaissables de mots infinis par les semigroupes. Cependant, les deux principales structures algébriques adaptées, les ω-semigroupes et les algèbres de Wilke, ont été introduites par Perrin et Pin [31] et par Wilke [42, 43]. Nous présentons ici rapidement les ω-semigroupes et renvoyons à [32] pour une description plus détaillée de ces deux structures équivalentes.

Définition 15 Un *ω-semigroupe* S est une algèbre à deux composantes $S = (S_+, S_\omega)$ équipée des opérations suivantes :

- une opérations binaire définie sur S_+ et notée multiplicativement,
- une application $S_+ \times S_\omega \to S_\omega$ appelée *produit mixte*, qui à un couple (s, t) associe un élément de S_ω noté st,
- une application $\pi : (S_+)^\omega \to S_\omega$ appelée *ω-produit*.

Ces opérations doivent vérifier les propriétés suivantes :

- S_+ équipé de l'opération binaire est un semigroupe,
- quels que soient s, $t \in S_+$ et $u \in S_\omega$, $s(tu) = (st)u$,
- pour toute suite strictement croissante d'entiers $(k_i)_{i\in\omega}$ et pour toute suite $(s_i)_{i\in\omega}$ d'éléments de S_+,

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

- pour tout $s \in S_+$ et pour toute suite $(s_i)_{i\in\omega}$ d'éléments de S_+,

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

Ces propriétés généralisent l'associativité du produit de semigroupe et permettent de simplifier par la suite la notation $\pi(s_0, s_1, s_2, \dots)$ par $s_0 s_1 s_2 \dots$. Intuitivement, un ω-semigroupe est une sorte de semigroupe qui autorise les produits infinis d'éléments de S_+ mais où le produit à droite d'un élément de S_ω est interdit.

Exemple 39 Pour tout alphabet A, le couple (A^+, A^ω) est naturellement muni d'une structure d'ω-semigroupe par le produit de concaténation.

Etant donnés $S = (S_+, S_\omega)$ et $T = (T_+, T_\omega)$ deux ω-semigroupes, un *morphisme d'ω-semigroupe* est donc un couple $\varphi = (\varphi_+, \varphi_\omega)$ où $\varphi_+ : S_+ \to T_+$ est un morphisme de semigroupe et $\varphi_\omega : S_\omega \to T_\omega$ est une application qui préserve le ω-produit : pour toute suite $(s_i)_{i\in\omega}$ d'éléments de S_+,

$$\varphi_\omega(s_0 s_1 s_2 \ldots) = \varphi_+(s_0)\varphi_+(s_1)\varphi_+(s_2)\ldots$$

Elle préserve également le produit mixte : pour tout $s \in S_+$ et pour tout $t \in S_\omega$,

$$\varphi_+(s)\varphi_\omega(t) = \varphi_\omega(st).$$

La notion de reconnaissabilité s'étend alors aux ensembles de mots infinis.

Exemple 40 Soit $A = \{a, b\}$ et soit $S = (\{s, t\}, \{u, v\})$ l'ω-semigroupe défini par les opérations $ss = s$, $st = ts = tt = t$, $s^\omega = u$, $t^\omega = v$, $su = tu = u$, $sv = tv = v$. Soit $\varphi : (A^+, A^\omega) \to S$ le morphisme d'ω-semigroupe défini sur A par $\varphi(a) = s$ et $\varphi(b) = t$. L' ω-semigroupe S reconnaît chacun des ensembles $\varphi^{-1}(s) = a^+$, $\varphi^{-1}(t) = A^*bA^*$, $\varphi^{-1}(u) = A^*a^\omega$, $\varphi^{-1}(v) = (A^*b)^\omega$ ainsi que leurs unions.

Notons que les ω-semigroupes permettent également de reconnaître des ensembles de mots finis. On retrouve la triple équivalence entre expressions rationnelles, automates et ω-semigroupes **finis** pour les sous-ensembles de A^ω.

Théorème 23 *Un ensemble de mots de longueur ω est ω-rationnel si et seulement s'il est reconnaissable.*

La théorie des automates sur les mots de longueur ω ne donne pas l'existence d'un automate minimal acceptant un ensemble ω-rationnel donné. L'approche algébrique vient compléter cette théorie et permet de construire un objet canonique associé à chaque ensemble ω-rationnel. La définition de la congruence syntaxique d'un sous-ensemble reconnaissable de A^ω est due à Arnold [1].

Définition 16 Pour tout sous-ensemble X de A^ω, la relation $\sim_X$ sur (A^+, A^ω) est définie de la façon suivante :
pour tous mots x, y de A^+, $x \sim_X y$ si et seulement si $\forall u, v \in A^*$, $\forall w \in A^+$,

$$uxvw^\omega \in X \iff uyvw^\omega \in X \quad \text{et} \quad u(xv)^\omega \in X \iff u(yv)^\omega \in X$$

et pour tous mots x, y de A^ω, $x \sim_X y$ si et seulement si $\forall u \in A^*$,

$$ux \in X \iff uy \in X.$$

Lorsque X est reconnaissable, la relation $\sim_X$ est une congruence d'ω-semigroupe d'indice fini. L'ω-semigroupe syntaxique $S(X) = (A^+, A^\omega)/\sim_X$ est alors le plus petit semigroupe reconnaissant X au sens de la division. Notons que les ω-semigroupes ont été adaptés par Bedon [4] pour reconnaître des ensembles de mots indexés par des ordinaux de rang fini. Etant donné un entier n, un *ω^n-semigroupe* est partagé en $n+1$ semigroupes S_0, S_1, …, S_n de sorte que pour tout $i < n$, l'ω-produit d'éléments de S_i soit un élément de S_{i+1}. Les ensembles reconnaissables correspondant sont les ensembles rationnels de rang inférieur ou égal à n. A chaque ensemble reconnaissable est également associé un ω^n-semigroupe syntaxique. Nous renvoyons à [3] pour plus de détails.

4.1.3 ω_1-semigroupes

Les semigroupes ont été généralisés par Bedon et Carton [4, 7] pour reconnaître des ensembles de mots sur les ordinaux dénombrables. La structure algébrique introduite autorise les produits de toute suite de longueur ordinale dénombrable. Un ω_1-semigroupe est un ensemble S équipé d'un produit qui associe un élément de S à chaque suite d'éléments de S indexée par un ordinal dénombrable.

Définition 17 Un ω_1-*semigroupe* est un ensemble S muni d'une application $\pi : S^\natural \to S$ satisfaisant les propriétés suivantes :
- Pour tout élément $S \in S$, $\pi(s) = s$,
- Pour tout mot x sur S et pour toute factorisation $x = \prod_{\gamma<\alpha} x_\gamma$,

$$\pi(x) = \pi(\prod_{\gamma<\alpha} \pi(x_\gamma))$$

L'application π est appelée ω_1-produit de l'ω_1-semigroupe.

La première propriété, similaire à celle des ω-semigroupes, impose que le produit d'un élément de S reste lui-même. La seconde condition est une généralisation de la propriété d'associativité des semigroupes : lorsque qu'une suite x se factorise $x = \prod_{\gamma<\alpha} x_\gamma$, son produit $\pi(x)$ peut se calculer soit directement sur x, soit en appliquant d'abord π aux facteurs x_γ puis en appliquant π à la suite obtenue. Comme pour les semigroupes, on confond une suite d'éléments d'un ω_1-semigroupe et le produit de ces éléments.

L'ensemble $A^\natural$ des mots sur l'alphabet A indexés par des ordinaux dénombrables muni du produit de concaténation est un ω_1-semigroupe.

Exemple 41 L'ensemble $\{0,1\}$ muni du produit défini pour tout $x \in \{0,1\}^\natural$ par $\pi(x) = 0$ si x possède un 0 et $\pi(x) = 1$ sinon est un ω_1-semigroupe.

Exemple 42 L'ensemble $\{0,1\}$ muni du produit défini pour tout $x \in \{0,1\}^\natural$ par $\pi(x) = 0$ si x possède une longueur limite et $\pi(x) = 1$ sinon est un ω_1-semigroupe.

Un ω_1-semigroupe n'est pas une algèbre au sens habituel puisque son produit n'est pas d'arité finie. La description de son produit n'est pas finie puisqu'il faut donner le produit de toute suite de longueur ordinale dénombrable. Dans le cas d'un ω_1-semigroupe fini, on peut cependant définir le produit de façon finie à l'aide des fonctions compatibles :

Définition 18 Soit S un semigroupe. Une fonction $\tau : S^\natural \to S$ est *compatible à droite* avec S si et seulement si quels que soient les éléments s et t de S et quel que soit l'entier n, $s(ts)^\tau = (st)^\tau$ et $(s^n)^\tau = s^\tau$.

Le produit d'un ω_1-semigroupe (S, π) fini se définit alors par un produit fini sur S et une fonction compatible à droite avec S.

Théorème 24 *Soit (S, π) un ω_1-semigroupe fini. Le produit binaire défini par $st = \pi(st)$ induit de façon naturelle une structure de semigroupe sur S et la fonction τ définie par $s^\tau = \pi(s^\omega)$ est compatible à droite avec ce semigroupe.*

Réciproquement, soient S un semigroupe fini et τ une fonction compatible à droite avec S. Le semigoupe S induit de façon unique une structure d'ω_1-semigroupe (S, π) telle que $\pi(s^\omega) = s^\tau$.

Exemple 43 Le produit de l'ω_1-semigroupe $(\{0, 1\}, \pi)$ de l'exemple 41 est défini par le produit fini $00 = 01 = 10 = 0$, $11 = 1$ et par la fonction compatible définie par $0^\tau = 0$ et $1^\tau = 1$.

Exemple 44 Le produit de l'ω_1-semigroupe $(\{0, 1\}, \pi)$ de l'exemple 42 est défini par le produit fini $00 = 01 = 10 = 11 = 1$ et par la fonction compatible définie par $0^\tau = 0$ et $1^\tau = 0$.

Les notions de morphisme d'ω_1-semigroupe, de reconnaissabilité, division,..., proviennent directement de l'algèbre universelle même si le produit n'est pas d'arité finie. Soient (S, π) et (T, π') deux ω_1-semigroupes. Un *morphisme d'ω_1-semigroupe* $\varphi : S \to T$ est donc une application telle que pour tout mot $x \in S^\natural$ et pour toute factorisation $x = \prod_{\gamma<\alpha} x_\gamma$, $\varphi(x) = \pi'(\prod_{\gamma<\alpha} \varphi(x_\gamma))$.

Exemple 45 Reprenons l'ω_1-semigroupe $S = (\{0, 1\}, \pi)$ de l'exemple 43 et soit $\varphi : A^\natural \to S$ le morphisme d'ω_1-semigroupe étendant l'application définie sur A par $\varphi(a) = 0$ et $\varphi(b) = 1$. Les ensembles $\varphi^{-1}(0) = A^\natural a A^\natural$ des mots contenant au moins une occurrence de la lettre a et son complément $\varphi^{-1}(1) = b^\natural$ sont reconnaissables.

Exemple 46 Pour l'ω_1-semigroupe S de l'exemple 44, on définit le morphisme $\varphi : A^\natural \to S$ par $\varphi(a) = 0$ et $\varphi(b) = 0$. Il reconnaît les ensembles $\varphi^{-1}(0) = (A^\natural)^\omega$ des mots de longueur limite et son complément $\varphi^{-1}(1) = A^\natural A$.

Bedon et Carton ont montré que lorsqu'ils sont finis, les ω_1-semigroupes sont équivalents aux automates de Büchi sur les ordinaux dénombrables.

Théorème 25 *Un ensemble de mots indexés par des ordinaux dénombrables est reconnu par un ω_1-semigroupe fini si et seulement si il est rationnel.*

Deux preuves différentes de ce résultat sont présentées dans [4] et [7]. Etant donné un ω_1-semigroupe fini reconnaissant un ensemble $X \subseteq A^\natural$, les deux démonstrations construisent un automate déterministe reconnaissant X. Elles utilisent des propriétés algébriques que nous détaillerons plus tard, notamment le théorème de Ramsey qui sera également indispensable dans le cas des ordres linéaires. La congruence syntaxique est généralisée aux ω_1-semigroupes.

Définition 19 Soit (S, π) un ω_1-semigroupe et soit $X \subseteq S$. Pour tous les éléments x et y de S, on dit que $x \sim_X y$ si et seulement si pour tout entier n et tous les éléments $s_0 \dots s_n \in S^1$,

$$s_0(\dots((xs_1)^\omega s_2)^\omega \dots s_{n-1})^\omega s_n \in X \iff s_0(\dots((ys_1)^\omega s_2)^\omega \dots s_{n-1})^\omega s_n \in X$$

Lorsque l'ensemble X est reconnaissable, il suffit de vérifier l'égalité précédente pour un nombre fini d'entiers n. L'ω_1-semigroupe $S/\sim_X$ est alors effectif.

Théorème 26 *Soit X un sous-ensemble reconnaissable de $A^\natural$. La relation $\sim_X$ est une congruence d'ω_1-semigroupe d'indice fini. Le quotient $A^\natural/\sim_X$, appelé ω_1-semigroupe syntaxique, reconnaît X et divise tout ω_1-semigroupe reconnaissant X.*

4.2 Définitions algébriques

On rappelle ici des définitions algébriques avancées sur les semigroupes dont on aura besoin par la suite. Certaines propriétés classiques des semigroupes finis sont tout d'abord détaillées avant de présenter les relations de Green, largement utilisées dans l'étude algébrique des ensembles rationnels. Enfin, on présente une version algébrique du théorème de Ramsey pour les mots infinis.

4.2.1 Idempotents

Un élément e d'un semigroupe S est un *idempotent* si $ee = e$ et l'ensemble des idempotents de S est noté $E(S)$.

Si S est un semigroupe cyclique, engendré par un unique élément s, $S = \{s, s^2, s^3, \dots\}$ est l'ensemble des puissances de s. Si S est infini, il est isomorphe au semigroupe des entiers strictement positifs muni de l'addition. Si S est fini, il existe des entiers i et $p > 0$ tels que $s^{i+p} = s^i$. Les plus petits entiers i et p vérifiant ces propriétés sont appelés respectivement l'*indice* et la *période* de s. La structure du semigroupe S est représentée par le schéma de la figure 4.1.

Proposition 27 *Dans un semigroupe fini, tout élément admet une puissance idempotente.*

Preuve. Soient i et p l'indice et la période d'un élément s. Pour tout $k \geq i$, $s^k = s^{k+p}$. En particulier, si k est un multiple de p, $(s^k)^2 = s^{2k} = s^{k+qp} = s^k$. $\square$

La proposition 27 a deux conséquences importantes.

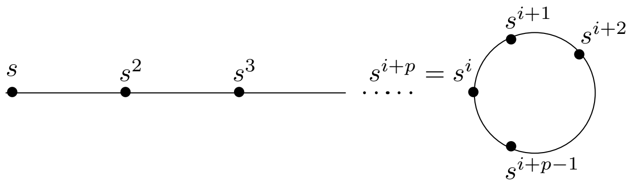


FIG. 4.1 – Semigroupe engendré par s.

Corollaire 28 *Tout semigroupe fini non vide possède un idempotent.*

Proposition 29 *Pour tout semigroupe fini S, il existe un entier π tel que pour tout élément $s \in S$, s^π est idempotent.*

Preuve. D'après la proposition 27, chaque élément $s \in S$ admet une puissance idempotente s^{n_s}. Il suffit de définir π comme le plus petit multiple commun des n_s pour $s \in S$. □
Le plus petit entier π vérifiant cette propriété est appelé l'*exposant* de S.

4.2.2 Relations de Green

Ce paragraphe est consacré aux relations de Green et à certains résultats classiques sur les semigroupes. On redonne certaines preuves pour familiariser le lecteur mais on renvoie à [33] pour plus de détails. Soit S un semigroupe. On définit sur S les quatre relations de préordre suivantes :

$$\begin{array}{llllll} s \le_{\mathcal{R}} t & \iff & \exists a \in S^1,\ s = ta & \iff & sS^1 \subseteq tS^1 \\ s \le_{\mathcal{L}} t & \iff & \exists a \in S^1,\ s = at & \iff & S^1 s \subseteq S^1 t \\ s \le_{\mathcal{J}} t & \iff & \exists a, b \in S^1,\ s = atb & \iff & S^1 s S^1 \subseteq S^1 t S^1 \\ s \le_{\mathcal{H}} t & \iff & s \le_{\mathcal{R}} t \text{ et } s \le_{\mathcal{L}} t. \end{array}$$

Les relations de Green sont les relations d'équivalence associées : pour tout $\mathcal{K} \in \{\mathcal{R}, \mathcal{L}, \mathcal{J}, \mathcal{H}\}$, $s\ \mathcal{K}\ t$ si et seulement si $s \le_{\mathcal{K}} t$ et $t \le_{\mathcal{K}} s$.

$$\begin{array}{lllll} s\ \mathcal{R}\ t & \iff & \exists a, b \in S^1,\ s = ta \text{ et } t = sb & \iff & sS^1 = tS^1 \\ s\ \mathcal{L}\ t & \iff & \exists a, b \in S^1,\ s = at \text{ et } t = bs & \iff & S^1 s = S^1 t \\ s\ \mathcal{J}\ t & \iff & \exists a, b, c, d \in S^1,\ s = atb \text{ et } t = csd & \iff & S^1 s S^1 = S^1 t S^1 \\ s\ \mathcal{H}\ t & \iff & s\ \mathcal{R}\ t \text{ et } s\ \mathcal{L}\ t \end{array}$$

Pour tout $\mathcal{K} \in \{\mathcal{R}, \mathcal{L}, \mathcal{J}, \mathcal{H}\}$, on note $s <_{\mathcal{K}} t$ si et seulement si $s \le_{\mathcal{K}} t$ et non $t \le_{\mathcal{K}} s$.

Proposition 30 *Les relations $\le_{\mathcal{R}}$ et $\mathcal{R}$ sont stables à gauche et les relations $\le_{\mathcal{L}}$ et $\mathcal{L}$ sont stables à droite.*

Preuve. Supposons $s \le_{\mathcal{R}} t$. On dispose de $a \in S^1$ tel que $s = ta$. Pour tout $u \in S^1$, $us = uta$ et donc $us \le_{\mathcal{R}} ut$. □
On note $s\mathcal{R}\mathcal{L}t$ si il existe $r \in S$, tel que $s\ \mathcal{R}\ r$ et $r\ \mathcal{L}\ t$.

Proposition 31 *Dans un semigroupe, les relations $\mathcal{R}$ et $\mathcal{L}$ commutent ($\mathcal{RL} = \mathcal{LR}$).*

Preuve. Soit S un semigroupe et soient $s, t, r \in S$. Supposons $s \mathrel{\mathcal{R}} r$ et $r \mathrel{\mathcal{L}} t$. Il existe $a, b, c, d \in S^1$ tels que $s = ra$, $r = sb$, $r = ct$ et $t = dr$. Considérons $u = dra = ds = ta$. Il vient $s \mathrel{\mathcal{L}} u$ car d'une part $u = ds$ et d'autre part $s = ra = cta = cu$. De plus, $u \mathrel{\mathcal{R}} t$ car $u = ta$ et $t = dr = dsb = ub$. D'où $s \mathrel{\mathcal{L}} u$ et $u \mathrel{\mathcal{R}} t$ donc $s \mathrel{\mathcal{LR}} t$. □

On pose $\mathcal{D}$=$\mathcal{RL}$=$\mathcal{LR}$. Ainsi, $\mathcal{D}$ est une relation d'équivalence et c'est la plus petite contenant $\mathcal{R}$ et $\mathcal{L}$.

Proposition 32 *Dans un semigroupe fini, $\mathcal{D}$=$\mathcal{J}$.*

Preuve. Soit S un semigroupe fini et soient $s, t \in S$. Supposons $s \mathrel{\mathcal{J}} t$. Il existe $a, b, c, d \in S^1$ tels que $t = asb$ et $s = ctd$. Il vient $s = ctd = casbd = (ca)s(bd)$. On peut itérer et obtenir pour tout entier n, $(ca)^n s(bd)^n$. D'après la proposition 29, on dispose d'un entier π tel que $(ca)^\pi = e$ et $(bd)^\pi = f$ soient idempotents. En multipliant à gauche par e, $es = eesf = esf = s$ d'où $es = s$. On déduit que $s \mathrel{\mathcal{L}} as$ puisque $s = (ca)^\pi s = ((ca)^{\pi-1})c)as$. De même, le raisonnement dual donne $s \mathrel{\mathcal{R}} sb$. Comme $\mathcal{R}$ est stable à gauche, on obtient $as \mathrel{\mathcal{R}} asb$ et $s \mathrel{\mathcal{L}} as$ d'où $s \mathrel{\mathcal{D}} asb = t$. Réciproquement, supposons $s \mathrel{\mathcal{D}} t$. On dispose de $r \in S$ tel que $s \mathrel{\mathcal{R}} r$ et $r \mathrel{\mathcal{L}} t$. Il existe $a, b, c, d \in S^1$ tel que $s = ra$, $r = sb$, $t = cr$ et $r = dt$. D'où $s = ra = dta$ et $t = cr = csb$ ce qui montre que $s \mathrel{\mathcal{J}} t$. □

Les classes d'équivalence des relations $\mathcal{R}$, $\mathcal{L}$, $\mathcal{J}$, $\mathcal{H}$ et $\mathcal{D}$ sont appelées $\mathcal{R}$-classes, $\mathcal{L}$-classes, ... Pour tout élément s d'un semigroupe S, on note R_s la $\mathcal{R}$-classe de S contenant s, L_s la $\mathcal{L}$-classe de s...

Proposition 33 *Soient R une $\mathcal{R}$-classe et L une $\mathcal{L}$-classe. Alors $R \cap L \neq \emptyset$ si et seulement si R et L sont contenues dans une même $\mathcal{D}$-classe.*

Une $\mathcal{D}$-classe est traditionnellement représentée par une "boîte d'oeufs" dans laquelle chaque colonne est une $\mathcal{L}$-classe, chaque ligne une $\mathcal{R}$-classe et chaque intersection de ligne et de colonne une $\mathcal{H}$-classe. La présence d'un idempotent

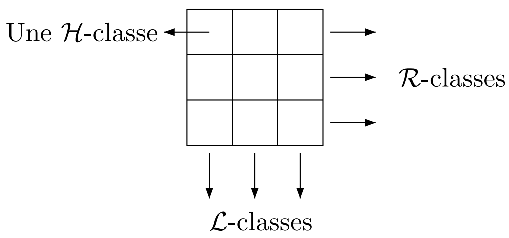


FIG. 4.2 – Représentation d'une $\mathcal{D}$-classe.

dans une $\mathcal{H}$-classe est signalée par une étoile.

Définition 20 Une $\mathcal{D}$-classe d'un semigroupe est dite *régulière* si elle contient un idempotent.

Proposition 34 *Si D est une $\mathcal{D}$-classe régulière d'un semigroupe S, alors chaque $\mathcal{R}$-classe et chaque $\mathcal{L}$-classe de D contient un idempotent.*

Proposition 35 *Soit D une $\mathcal{D}$-classe d'un semigroupe S et soient s et t des éléments de S. Si $s \mathcal{R} t$ et si $s = tp$ et $t = sq$, les applications $x \to xp$ et $x \to xq$ définissent des bijections inverses entre L_s et L_t et ces bijections préservent les $\mathcal{H}$-classes.*

Preuve. Soit $u \in L_s$. Il existe $a \in S^1$ tel que $u = as$. Donc $uqp = asqp = atp = as = u$. De façon symétrique, si $v \in L_t$, il existe $b \in S^1$ tel que $v = bt$ et donc $vpq = btpq = bsq = bt = v$. De plus, la proposition 30 montre que les applications $x \to xp$ et $x \to xq$ préservent les $\mathcal{H}$-classes. □

Proposition 36 *Soit S un semigroupe fini et soient $s, t \in S$.*
- *Si $s \mathcal{J} t$ et $s \leq_{\mathcal{R}} t$, alors $s \mathcal{R} t$.*
- *Si $s \mathcal{J} t$ et $s \leq_{\mathcal{L}} t$, alors $s \mathcal{L} t$.*

Preuve. Supposons $s \mathcal{J} t$. On dispose de $a, b \in S^1$ tels que $t = asb$. Si $s \leq_{\mathcal{R}} t$, il existe $c \in S^1$ tel que $s = tc$. Comme $t = atcb$, pour tout entier n, on a $t = a^n t (cb)^n$. D'après la proposition 29, on dispose d'un entier π tel que $a^\pi = e$ et $(cb)^\pi = f$ soient idempotents. Il vient $t = etf = (etf)f = tf = tc(bc)^{\pi-1}b = s(bc)^{\pi-1}b$. D'où $t \leq_{\mathcal{R}} s$ donc $s \mathcal{R} t$.

Définition 21 Soit S un semigroupe. Un couple $(s, e) \in S \times S$ est *lié à droite* si $e \in E(S)$ et si $se = s$. Deux couples (s_1, e_1) et (s_2, e_2) liés à droite sont *conjugués à droite* s'il existe a, $b \in S^1$ tels que $e_1 = ab$, $e_2 = ba$, $s_1 a = s_2$ (et $s_2 b = s_1$). Symétriquement, un couple $(e, s) \in E(S) \times S$ est *lié à gauche* si $es = s$ et deux couples liés à gauche (e_1, s_1) et (e_2, s_2) sont *conjugués à gauche* s'il existe a, $b \in S^1$ tels que $e_1 = ab$, $e_2 = ba$, $bs_1 = s_2$ (et $as_2 = s_1$).

Notons que si deux couples liés (s_1, e_1) et (s_2, e_2) sont conjugués à droite, alors $e_1 \mathcal{D} e_2$ et $s_1 \mathcal{R} s_2$ (respectivement à gauche, $s_1 \mathcal{L} s_2$).

Proposition 37 *[27] Les relations de conjugaison à droite et à gauche sont des relations d'équivalence.*

Preuve. On montre que la relation de conjugaison à droite est une relation d'équivalence. La réflexivité et la symétrie sont triviales par définition. Soient (s_1, e_1), (s_2, e_2) et (s_3, e_3) des couples liés à droite. Supposons que (s_1, e_1) et (s_2, e_2) sont conjugués à droite et que (s_2, e_2) et (s_3, e_3) aussi. Il existe des éléments $a, b, c, d \in S$ tels que

$$s_1 a = s_2, s_2 b = s_1, e_1 = ab, e_2 = ba, s_2 c = s_3, s_3 d = s_2, e_2 = cd, e_3 = dc.$$

Il vient $(ac)(db) = a(cd)b = a(ba)b = e_1$ et symétriquement, $(db)(ac) = e_3$. De plus, $s_1(ac) = s_2 c = s_3$ et $s_3(db) = s_1$ ce qui prouve la transitivité. □

Proposition 38 *[15] Soit S un semigroupe fini et soit D une $\mathcal{D}$-classe de S. Soient (s, e) et (t, f) deux couples liés à droite (respectivement liés à gauche) tels que $s,\ t,\ e,\ f \in D$. Les couples sont conjugués si et seulement si $s \mathrel{\mathcal{R}} t$ (respectivement $s \mathrel{\mathcal{L}} t$).*

Preuve. Nous avons déjà vu que lorsque deux couples sont conjugués à droite, alors $s_1 \mathrel{\mathcal{R}} s_2$. Réciproquement, supposons que $s_1 \mathrel{\mathcal{R}} s_2$. Comme $e_1 \in D$, D est régulière et la $\mathcal{R}$-classe de s_1 et s_2 contient un idempotent e. On montre que les couples (s_1, e_1) et (e, e) sont conjugués à droite. Comme $s_1 e_1 = s_1$, $s_1 \leq_{\mathcal{L}} e_1$ et d'après la proposition 36, $s_1 \mathrel{\mathcal{L}} e_1$. Donc $s_1 \in R_e \cap L_{e_1}$ et il existe $a, b, c, d \in S^1$ tels que $s_1 a = e$, $eb = s_1$, $cs_1 = e_1$ et $de_1 = s_1$. Soit $u = cs_1 a$. Il vient $s_1 u = s_1 c s_1 a = s_1 e_1 a = s_1 a = e$ et d'après la proposition 35, $us_1 = ueb = cs_1 aeb = ceeb = ceb = cs_1 ab = cs_1 = e_1$. De plus, $es_1 = eeb = eb = s_1$ ce qui montre que les couples (s_1, e_1) et (e, e) sont conjugués à droite. De même, on obtient que les couples (s_2, e_2) et (e, e) sont conjugués à droite, ce qui prouve que (s_1, e_1) et (s_2, e_2) sont conjugués.

□

4.2.3 Factorisations ramseyennes

On donne ici une version algébrique du théorème de Ramsey utilisant les semigroupes et sa preuve empruntée à [32].

Théorème 39 *Soit S un semigroupe fini et soit $\varphi : A^+ \to S$ un morphisme de semigroupe. Pour tout mot $x \in A^\omega$, il existe un couple lié à droite $(s, e) \in S \times E(S)$ et une factorisation $x = \prod_{i \in \omega} x_i$ telle que $\varphi(x_0) = s$ et pour tout $i > 0$, $\varphi(x_i) = e$.*

Preuve. En utilisant le théorème 2, on dispose d'éléments $r, t \in S$ tels que $x \in \varphi^{-1}(r)\varphi^{-1}(t)^\omega$. D'après le théorème 27, l'élément t admet une puissance idempotente notée e. En regroupant les facteurs successifs d'image t, on obtient une factorisation $x = \prod_{i \in \omega} x_i$ telle que $\varphi(x_0) = r$ et $\varphi(x_i) = e$ pour tout $i > 0$. Le couple lié est défini en posant $s = re$ puisque $se = ree = re = s$. Finalement, la factorisation recherchée est définie par $y_0 = x_0 x_1$ et pour tout $i > 1$, $y_i = x_{i+1}$.
□
De plus, notons que deux couples liés associés à des factorisations d'un même mot sont conjugués.

Proposition 40 *Soit S un semigroupe fini et soit $\varphi : A^+ \to S$ un morphisme de semigroupe. Si un mot $x \in A^\omega$ admet deux factorisations ramseyennes pour φ associées à deux couples liés, alors ces couples sont conjugués.*

Le lecteur est renvoyé à [32] page 81 pour une preuve détaillée.

4.3 Complémentation par l'approche algébrique

L'équivalence algébrique est ici généralisée aux mots sur les ordres linéaires dispersés. Après avoir défini les $\diamond$-semigroupes, on montre qu'un sous-ensemble de $A^\diamond$ est rationnel si et seulement son $\diamond$-semigroupe syntaxique est fini.

4.3.1 $\diamond$-semigroupes

Les semigroupes étant munis d'un produit fini, ils ne sont pas adaptés pour reconnaître des mots sur les ordres linéaires. On généralise le produit à toute suite indexée par un ordre linéaire dispersé.

Définition 22 Un $\diamond$-*semigroupe* est un ensemble S muni d'un produit $\pi : S^\diamond \longrightarrow S$ qui associe un élément de S à chaque mot indexé par un ordre total et dispersé et tel que

- pour tout élément s de S, $\pi(s) = s$;
- pour tout mot x sur S ayant pour longueur un ordre total et dispersé et pour toute factorisation $x = \prod_{j \in J} x_j$ où $J \in \mathcal{S}$,

$$\pi(x) = \pi(\prod_{j \in J} \pi(x_j)).$$

Cette dernière condition est une généralisation de l'associativité.

Par exemple, l'ensemble $A^\diamond$ muni de la concaténation est un $\diamond$-semigroupe.

Exemple 47 L'ensemble $S = \{0, 1\}$ muni du produit π défini pour tout $u \in S^\diamond$ par $\pi(u) = 0$ si u possède la lettre 0 et $\pi(u) = 1$ sinon est un $\diamond$-semigroupe.

Exemple 48 L'ensemble $S = \{0, 1\}$ muni du produit π défini pour tout $u \in S^\diamond$ par $\pi(u) = 1$ si $u \in 1^\#$ et $\pi(u) = 0$ sinon est un $\diamond$-semigroupe.

Pour tous les éléments s et t d'un $\diamond$-semigroupe (S, π), le produit fini $\pi(st)$ est simplement noté st.

Proposition 41 *Une application $\varphi : A \to S$ d'un alphabet A dans un $\diamond$-semigroupe S peut être étendue de façon unique en un morphime de $\diamond$-semigroupe $\varphi : A^\diamond \to S$.*

Preuve. Soit (S, π) un semigroupe et soit $\varphi : A \to S$ une application. Pour tout ordre linéaire $J \in \mathcal{S}$ et tout mot $x = (a_j)_{j \in J}$ de longueur J, on pose $\varphi(x) = \pi((\varphi(a_j))_{j \in J})$. $\square$

$\diamond$-semigroupes finis

Un $\diamond$-semigroupe (S, π) est dit fini si son support S est fini. Même quand S est fini, la fonction π n'est pas facile à décrire puisqu'il faut donner le produit de toute suite de longueur totale et dispersée. Il s'avère que la fonction π peut-être décrite par une structure de semigroupe sur S avec en plus deux fonctions (notées τ et $-\tau$) de S dans S. Ceci donne une description finie de la fonction π. Les fonctions τ et $-\tau$ sont l'équivalent des ω et $-\omega$-produits. Cette description finie est basée sur le lemme suivant qui découle directement du théorème de Ramsey [35] et qui sera largement utilisé dans les preuves.

Lemme 42 *Soit $\varphi : A^\diamond \longrightarrow S$ un morphisme de $A^\diamond$ dans un $\diamond$-semigroupe fini S. Pour tout mot de longueur limite à droite $x \in (A^\diamond)^\omega$ et toute factorisation $x = \prod_{i\in\omega} x_i$, il existe un couple lié à droite $(s, e) \in S \times E(S)$ et une factorisation extraite $x = \prod_{i\in\omega} y_i$ tels que $\varphi(y_0) = s$ et $\varphi(y_i) = e$ pour tout $i > 0$.*

Preuve. Soit $x \in (A^\diamond)^\omega$ et soit une factorisation $x = \prod_{i\in\omega} x_i$. En appliquant le théorème 3 à $\pi : S^+ \longrightarrow S$ et au mot $u = \prod_{i\in\omega} \varphi(x_i)$, on dispose d'un couple lié à droite $(s, e) \in S \times E(S)$ et d'une superfactorisation $u = \prod_{i\in\omega} u_i$ telles que $\pi(u_0) = s$ et $\pi(u_i) = e$ pour tout $i > 0$. Notons pour tout $i \geq 0$, $y_i = x_{i,1} \dots x_{i,m_i}$. Il vient $x = \prod_{i\in\omega} y_i$, $\varphi(y_0) = \varphi(x_{0,1} \dots x_{0,m_i}) = \varphi(x_{0,1}) \dots \varphi(x_{0,m_0}) = \pi(u_0) = s$ et pour tout $i > 0$, $\varphi(y_i) = \varphi(x_{i,1} \dots x_{i,m_i}) = \pi(u_i) = e$. $\square$

Une telle factorisation est appelée *factorisation ramseyenne*. De façon symétrique, tout mot de longueur limite à gauche admet une factorisation ramseyenne associée à un couple lié à gauche.

Proposition 43 *Soit (S, π) un $\diamond$-semigroupe fini et soit $\varphi : A^\diamond \to S$ un morphisme de $\diamond$-semigroupe. Si un mot $x \in A^\diamond$ admet deux factorisations ramseyennes pour φ associées à deux couples liés, alors ces couples sont conjugués.*

Preuve. Soient $x = \prod_{i\in\omega} x_i$ et $x = \prod_{i\in\omega} y_i$ deux factorisations ramseyennes pour φ associées respectivement aux couples liés (s, e) et (t, f). Quitte à prendre des superfactorisations, on peut supposer que $|x_0| < |y_0| < |x_0x_1| < |y_0y_1| < |x_0x_1x_2| < |y_0y_1y_2| \dots$. Ainsi pour tout $i > 0$, $x_i = z_iz'_i$ avec $y_0 = x_0z_1$ et $y_i = z'_iz_{i+1}$. Le mot $u = \varphi(x_0)\varphi(z_1)\varphi(z'_1)\varphi(z_2)\varphi(z'_2) \dots \in S^\omega$ admet les deux factorisations $u = \varphi(x_0) \prod_{i>0} (\varphi(z_i)\varphi(z'_i))$ et $u = (\varphi(x_0)\varphi(z_1)) \prod_{i>0} (\varphi(z'_i)\varphi(z_{i+1}))$ ramseyennes pour π associées respectivement aux couples liés (s, e) et (t, f). D'après le proposition 40, (s, e) et (t, f) sont conjuguées. $\square$

Nous introduisons à présent la définition de fonctions compatibles, qui permettra de décrire le produit d'un $\diamond$-semigroupe fini de façon finie.

Définition 23 Soit S un semigroupe. Une fonction $\tau : S \longrightarrow S$ (respectivement $-\tau : S \longrightarrow S$) est *compatible à droite avec* S (respectivement *à gauche*) si et seulement si quels que soient s et t appartenant à S et quel que soit l'entier n, τ vérifie les propriétés suivantes : $s(ts)^{\tau} = (st)^{\tau}$ et $(s^n)^{\tau} = s^{\tau}$ (respectivement $(st)^{-\tau}s = (ts)^{-\tau}$ et $(s^n)^{-\tau} = s^{-\tau}$).

Lemme 44 *Soient S un $\diamond$-semigroupe fini, τ une fonction compatible à droite avec S et soient $(s, e) \in S \times E(S)$ et $(t, f) \in S \times E(S)$ deux couples liés à droite. Si les deux couples sont conjugués, alors $se^{\tau} = tf^{\tau}$. Symétriquement, si τ est une fonction compatible à gauche avec S et si deux couples $(e, s) \in E(S) \times S$ et $(f, t) \in E(S) \times$ sont conjugués à gauche alors $e^{-\tau}s = f^{-\tau}t$.*

Preuve. Soient $(s, e) \in S \times E(S)$ et $(t, f) \in S \times E(S)$ deux couples conjugués à droite. On dispose de a, $b \in S^1$ tels que $e = ab$, $f = ba$, $sa = t$ et $tb = s$. Comme τ est compatible à droite, on a

$$se^{\tau} = s(ab)^{\tau} = (sa)(ba)^{\tau} = tf^{\tau}$$

Le cas des couples conjugués à gauche est montré de façon symétrique. $\square$

Le produit d'un $\diamond$-semigroupe fini S peut être décrit de façon finie par des fonctions compatibles à droite et à gauche avec S.

Théorème 45 *Soit (S, π) un $\diamond$-semigroupe fini. Le produit binaire défini pour tous les éléments s et t de S par $s \cdot t = \pi(st)$ induit de façon naturelle une structure de semigroupe et les fonctions τ et $-\tau$ définies respectivement par $s^{\tau} = \pi(s^{\omega})$ et $s^{-\tau} = \pi(s^{-\omega})$ sont compatibles respectivement à droite et à gauche avec S.*

Réciproquement, soit S un semigroupe fini et soient τ et $-\tau$ des fonctions compatibles respectivement à droite et à gauche avec S. Alors il existe une unique façon de munir S d'une structure de $\diamond$-semigroupe (S, π) de sorte que $s^{\tau} = \pi(s^{\omega})$ et $s^{-\tau} = \pi(s^{-\omega})$.

Preuve : Soit (S, π) un $\diamond$-semigroupe fini. Pour tout éléments s_1, s_2 et s_3 de S, on a $(s_1 \cdot s_2) \cdot s_3 = \pi(\pi(s_1s_2)s_3) = \pi(s_1\pi(s_2s_3)) = s_1 \cdot (s_2 \cdot s_3)$. De plus, pour tout entier n, on a $(s_1s_2)^{\tau} = \pi((s_1s_2)^{\omega}) = \pi(s_1(s_2s_1)^{\omega}) = \pi(s_1\pi((s_2s_1)^{\omega})) = s_1 \cdot (s_2s_1)^{\tau}$ donc τ est compatible à droite avec S. La preuve concernant $-\tau$ est similaire.

Réciproquement, soit S un semigroupe fini et soient τ et $-\tau$ des fonctions respectivement compatibles à droite et à gauche avec S. On montre qu'il existe un unique produit $\pi : S^{\diamond} \longrightarrow S$ tel que (S, π) soit un $\diamond$-semigroupe, $\pi(s^{\omega}) = s^{\tau}$ et $\pi(s^{-\omega}) = s^{-\tau}$ pour tout élément s de S.

Définition de π :

Le produit d'un mot $u = (s_i)_{i \in I}$ sur S de longueur $I \in \mathcal{S}$ est défini par induction sur $\alpha \in \mathcal{O}$ pour tout $I \in W_{\alpha}$.

Soit $I \in W_0$ et soit $u \in S^I$. Il existe un entier m et des éléments s_1, …, s_m de S tels que $u = s_1 \dots s_m$. On pose $\pi(u) = s_1 \cdot s_2 \dots s_m$.

Soit $I \in W_\alpha$ où $\alpha \geq 1$ et soit $u \in S^I$. On définit d'abord $\pi(u)$ lorsque $I \in U_\alpha = \bigcup_{\beta<\alpha} W_\beta \cup (\bigcup_{\beta<\alpha} W_\beta)^\omega \cup (\bigcup_{\beta<\alpha} W_\beta)^{-\omega}$. Dans le cas où $I \in \bigcup_{\beta<\alpha} W_\beta$, le produit $\pi(u)$ est défini par hypothèse d'induction. Supposons que $I \in (\bigcup_{\beta<\alpha} W_\beta)^\omega$. L'ordre I peut être décomposé en une somme $J = \sum_{j\in\omega} K_j$ où pour tout $j \in \omega$, $K_j \in \bigcup_{\beta<\alpha} W_\beta$. Soit $u = \prod_{j\in\omega} u_j$ la factorisation telle que pour tout $j \in \omega$, $|u_j| = K_j$. D'après le théorème de Ramsey, on dispose d' une superfactorisation $x = \prod_{j\in\omega} v_j$ et d'un couple lié à droite $(s, e) \in S \times E(S)$ tels que $\pi(v_0) = s$ et $\pi(v_j) = e$ pour tout $j > 0$. On pose $\pi(u) = se^\tau$. De façon symétrique, lorsque $I \in (\bigcup_{\beta<\alpha} W_\beta)^{-\omega}$, on pose $\pi(u) = e^{-\tau}s$.

Finalement, on suppose que $I \in W_\alpha$. On dispose d'un entier n et d'ordres $K_1, \ldots, K_n$ appartenant à U_α tels que $I = \prod_{j=1}^{n} K_j$. Soit $u = \prod_{j=1}^{n} u_j$ la factorisation associée, c'est-à-dire que pour tout $1 \leq j \leq n$, $|u_j| = K_j$. On pose alors $\pi(u) = \prod_{j=1}^{n} \pi(u_j)$.

On montre que le produit π est bien défini et est associatif sur $S^\diamond$. Par induction sur $\alpha \in \mathcal{O}$, on montre que pour tout mot u sur S de longueur $I \in W_\alpha$, $\pi(u)$ est bien défini de façon unique et que pour tout ordre $J \in W_\alpha$ et pour toute factorisation $u = \prod_{j\in J} u_j$ on a bien $\pi(u) = \pi(\prod_{j\in J} \pi(u_j))$.

Soit $u = \prod_{i\in I} s_i$.

Unicité de la définition et associativité de π sur S^{W_0} :

Supposons que $I \in W_0$. Le produit $\pi(u)$ est bien défini par le produit fini $\cdot$ du semigroupe S et est associatif sur S^{W_0} par associativité de $\cdot$.

Unicité de la définition de π sur S^{U_α} où $\alpha > 0$:

Soit $\alpha \in \mathcal{O}$. Supposons que $I \in U_\alpha$.

L'ordre I peut être décomposé en une somme $I = \sum_{j\in J} K_j$ où $J \in \mathcal{N} \cup \{\omega, -\omega\}$ et où pour tout $j \in J$, $K_j \in \bigcup_{\beta<\alpha} W_\beta$. On montre d'abord que le produit est bien défini pour une décomposition donnée de J puis on montrera que la valeur du produit ne dépend pas de la décomposition choisie.

Soit $u = \prod_{j\in J} u_i$ la factorisation associée, i.e. pour tout $j \in J$, $|u_j| = K_j$.

- $J \in \mathcal{N}$: la valeur de $\pi(u)$ est bien définie par hypothèse d'induction puisque $I \in \bigcup_{\beta<\alpha} W_\beta$.
- $J = \omega$: Supposons qu'il existe deux superfactorisations ramseyennes de u associées à deux couples liés à droite (s_1, e_1) et (s_2, e_2). D'après la proposition 43, ces deux couples sont conjugués et d'après le lemme 44,

$s_1 e_1^\tau = s_2 e_2^\tau$.

– $J = -\omega$: De façon symétrique, on obtient $e_1^{-\tau} s_1 = e_2^{-\tau} s_2$.

Il reste à montrer que la valeur de $\pi(u)$ ne dépend pas de la décomposition $J = \sum_{i \in I} K_i$ choisie. Supposons qu'il existe une autre décomposition $J = \sum_{l \in L} K'_l$ avec $L \in \mathcal{N} \cup \{\omega, -\omega\}$ et pour tout $l \in L$, $K'_l \in \bigcup_{\beta < \alpha} W_\beta$.

- Si $I \in \mathcal{N}$ ou si $L \in \mathcal{N}$ alors J est de rang strictement inférieur à α : $J \in \bigcup_{\beta<\alpha} W_\beta$ et le résultat est obtenu par hypothèse d'induction.
- Si $I = \omega$ et $L = -\omega$ (ou symétriquement $I = -\omega$ et $L = \omega$), le résultat est également obtenu par hypothèse d'induction puisque le rang de J est plus petit que α. En effet, soit $u = \prod_{i \in \omega} u_i = \prod_{l \in -\omega} v_l$ les factorisations associées, c'est-à-dire telle que pour tout $i \in \omega$, $|u_i| = K_i$ et pour tout $l \in -\omega$, $|v_l| = K'_l$. Il suffit de remarquer l'existence d'un indice $l_0 \in -\omega$ tel que $\sum_{l<l_0} K'_l \subseteq K_0$. L'ordre J est alors une somme finie d'ordres de rang strictement inférieur à α.
- Si $I = \omega$ et $L = \omega$ (ou symétriquement $I = -\omega$ et $L = -\omega$), chacune des factorisations admet une superfactorisation ramseyenne. En notant (s_1, e_1) et (s_2, e_2) les couples liés correspondants, on peut montrer comme précédemment que $s_1 e_1^\tau = s_2 e_2^\tau$.

Donc π est bien défini sur S^{U_α}. On montre qu'il est associatif sur ce domaine.

Associativité de π sur S^{U_α} :

Soit $I \in U_\alpha$ et $u \in S^I$. Si $I \in \bigcup_{\beta<\alpha} W_\beta$ l'associativité est obtenue par hypothèse d'induction. D'autre part, on peut supposer que $u \in (\bigcup_{\beta<\alpha} W_\beta)^\omega$ puisque le cas $u \in (\bigcup_{\beta<\alpha} W_\beta)^{-\omega}$ peut être traité de façon symétrique. Par définition, on dispose d'une ω-factorisation $u = \prod_{i\in\omega} u_i$ et d'un couple lié à droite (s, e) tels que $\pi(u_0) = s$ et pour tout $i > 0$, $\pi(u_i) = e$ et $\pi(u) = se^\tau = \pi(\prod_{i\in\omega} \pi(u_i))$. On note pour tout $i \in \omega$, $|u_i| = K_i$. Soit $J \in U_\alpha$ et $u = \prod_{j\in J} v_j$ une J-factorisation de u. On note pour tout $j \in J$, $|v_j| = K'_j$. On veut montrer que $\pi(u) = \pi(\prod_{j\in J} \pi(v_j))$. On montre d'abord ce résultat lorsque $J \in \mathcal{N} \cup \{\omega, -\omega\}$.

Soient $C_1 = \{(\bigcup_{0\le i\le i_0} K_i, \bigcup_{i_0<i} K_i) \mid i_0 \in \omega\}$ et $C_2 = \{(\bigcup_{\substack{j\in J,\\ j\le j_0}} K'_j, \bigcup_{\substack{j\in J,\\ j_0<j}} K'_j) \mid j_0 \in J\}$ les ensembles de coupures correspondants aux factorisations $u = \prod_{i\in\omega} u_i$ et $u = \prod_{j\in J} v_j$. Soit $C = C_1 \cup C_2$ muni de l'ordre sur $\hat{I}$ et soit $u = \prod_{h\in H} w_j$ la factorisation associée. Notons que $\prod_{i\in\omega} u_i$ et $\prod_{j\in J} v_j$ sont des superfactorisations

de $\prod_{h\in H} w_h$. Donc pour tout $i\in\omega$, on dispose d'un ordre H_i tel que $u_i = \prod_{h\in H_i} w_h$ et pour tout $j\in J$ on dispose d'un ordre H'_j tel que $v_j = \prod_{h\in H'_j} w_h$.

Quand $J\in\mathcal{N}\cup\{\omega,-\omega\}$, on montre que $\pi(\prod_{i\in\omega}\pi(u_i)) = \pi(\prod_{h\in H}\pi(w_h))$ et que $\pi(\prod_{j\in J}\pi(v_i)) = \pi(\prod_{h\in H}\pi(w_h))$, ce qui prouve que $\pi(u) = \pi(\prod_{j\in J}\pi(v_j))$.

- $J = \{0,\ldots,m\}\in\mathcal{N}$: Dans ce cas $H=\omega$. Comme $u = \prod_{h\in H} w_h$ avec $|w_h|\in\bigcup_{\beta<\alpha} W_\beta$, par définition on dispose d'une superfactorisation ramseyenne associée à un couple lié à droite (s',e'). Les couples (s,e) et (s',e') sont conjugués et de même que précédemment, on peut montrer que $\pi(u) = s'e'^{\tau}$. L'hypothèse d'associativité sur $\bigcup_{\beta<\alpha} W_\beta$ permet de conclure que $\pi(u) = \pi(\prod_{h\in H}\pi(w_h))$.
  D'autre part, le même argument donne $\pi(v_m) = \pi(\prod_{h\in H_m=\omega}\pi(w_{m,h}))$ puis par hypothèse d'induction on obtient $\pi(\prod_{j\in J}\pi(v_i)) = \pi(\prod_{h\in H}\pi(w_h))$.
- $J=\omega$: Dans ce cas $H=\omega$. Les deux ω-factorisations $u = \prod_{j\in\omega} v_j = \prod_{h\in\omega} w_h$ admettent des superfactorisations associées à des couples liés conjugués avec (s,e). Le résultat est montré de la même façon que précédemment pour l'unicité de la définition de π.
- $J=-\omega$: Dans ce cas $H=\zeta$. Comme $|u_0|\in\bigcup_{\beta<\alpha} W_\beta$, on peut utiliser l'hypothèse d'induction pour montrer que $\pi(u_0) = \pi(\prod_{h\in H_0=-\omega}\pi(w_h))$.
  De même que précédemment on obtient alors $\pi(u) = \pi(\prod_{h\in H}\pi(w_h))$. De même, en utilisant le cas $J\in\mathcal{N}$ et l'hypothèse d'induction, on obtient $\pi(\prod_{j\in J}\pi(v_i)) = \pi(\prod_{h\in H}\pi(w_h))$.

On a donc montré l'associativité sur U_α quand $J\in\mathcal{N}\cup\{\omega,-\omega\}$. En particulier on a le résultat quand $J\in W_0$ et on va le montrer par induction sur la classe de J. Soit $J\in U_\alpha$. On peut supposer que $u\in(\bigcup_{\beta<\alpha} W_\beta)^\omega$. On dispose d'une somme $J=\sum_{l\in\omega} K_l$ avec pour tout $l\in\omega$, $K_l\in\bigcup_{\beta<\alpha} W_\beta$.

$$\begin{array}{rcll}
\pi(\prod_{j\in J}\pi(v_j)) & = & \pi(\prod_{l\in\omega}\prod_{k\in K_l}\pi(v_{l,k})) & \\
 & = & \pi(\prod_{l\in\omega}\pi(\prod_{k\in K_l}\pi(v_{l,k}))) & \text{d'après le cas } J=\omega \\
 & = & \pi(\prod_{l\in\omega}\pi(\prod_{k\in K_l} v_{l,k})) & \text{par hypothèse d'associativité sur } \bigcup_{\beta<\alpha} W_\beta \\
 & = & \pi(\prod_{l\in\omega}(\prod_{k\in K_l} v_{l,k})) & \text{d'après le cas } J=\omega \\
 & = & \pi(u). &
\end{array}$$

Donc π est bien défini et associatif sur S^{U_α}.

Unicité de la définition de π sur S^{W_α} :

On suppose alors que $I \in W_\alpha$. Supposons qu'il existe deux entiers n et m et deux décompositions $I = \sum_{i=0}^{n} K_i = \sum_{j=0}^{m} K'_j$ avec pour tout $0 \leq i \leq n$, $K_i \in U_\alpha$ et pour tout $0 \leq j \leq m$, $K'_j \in U_\alpha$. De même que dans la preuve d'associativité précédente, on considère la plus grande sous-factorisation commune et on conclut l'égalité par associativité de π sur U_α.

Associativité de π sur S^{W_α} :

Soit $I \in W_\alpha$ et $u \in S^I$. On dispose d'un entier n et d'une factorisation $u = \prod_{j=1}^{n} u_j$ telle que pour tout $0 \leq i \leq n$, $|u_i| \in U_\alpha$. Par définition, $\pi(u) = \pi(\prod_{i=1}^{n} \pi(u_i))$. Par induction pour tout $0 \leq \beta \leq \alpha$, on montre que pour tout ordre $J \in W_\beta$ et toute factorisation $u = \prod_{j \in J} v_j$, on a $\pi(u) = \pi(\prod_{j \in J} \pi(v_j))$.

Dans le cas où $J \in \mathcal{N}$, la preuve est similaire au cas précédent. On peut considérer la plus grande sous-factorisation commune et conclure par unicité de la définition de π sur S^{W_α}.

Lorsque $J \in W_\alpha$, on dispose d'un entier n et d'une somme $J = \sum_{0 \leq l \leq n} K_l$ avec pour tout $0 \leq l \leq n$, $K_l \in U_\alpha$.

$$
\begin{array}{rcll}
\pi(\prod_{j \in J} \pi(v_j)) & = & \pi(\prod_{l=0}^{n} \prod_{k \in K_l} \pi(v_{l,k})) & \\
& = & \pi(\prod_{l=0}^{n} \pi(\prod_{k \in K_l} \pi(v_{l,k}))) & \text{d'après le cas } J \in \mathcal{N} \\
& = & \pi(\prod_{l=0}^{n} \pi(\prod_{k \in K_l} v_{l,k})) & \text{par hypothèse d'associativité sur } U_\alpha \\
& = & \pi(\prod_{l=0}^{n} (\prod_{k \in K_l} v_{l,k})) & \text{d'après le cas } J \in \mathcal{N} \\
& = & \pi(u). &
\end{array}
$$

Donc π est associatif sur S^{W_α} ce qui conclut la preuve. □

Exemple 49 Reprenons l'exemple 47. Ce $\diamond$-semigroupe $S = \{0, 1\}$ fini est défini par le produit fini $00 = 01 = 10 = 0$ et $11 = 1$ et par les fonctions compatibles τ et $-\tau$ définies par $0^\tau = 0^{-\tau} = 0$ et $1^\tau = 1^{-\tau} = 1$.

Exemple 50 Le $\diamond$-semigroupe $S = \{0, 1\}$ de l'exemple 48 est défini par le produit fini $00 = 01 = 10 = 0$ et $11 = 1$ et par les fonctions compatibles τ et $-\tau$ définies par $0^\tau = 0^{-\tau} = 1^{-\tau} = 0$ et $1^\tau = 1$.

Reconnaissabilité

On étend la notion de reconnaissabilité aux ensembles de mots indexés par des ordres linéaires dispersés.

Définition 24 Soient S et T deux $\diamond$-semigroupes. Le $\diamond$-semigroupe T *reconnaît* un sous-ensemble X de S si et seulement s'il existe un morphisme de $\diamond$-semigroupe $\varphi : S \longrightarrow T$ et un sous-ensemble $P \subseteq T$ tels que $X = \varphi^{-1}(P)$. Un ensemble $X \subseteq A^\diamond$ est *reconnaissable* si et seulement s'il existe un $\diamond$-semigroupe **fini** le reconnaissant.

Exemple 51 Soit $S = \{0, 1\}$ le $\diamond$-semigroupe fini de l'exemple 49. Définissons le morphisme de $\diamond$-semigroupe $\varphi : A^\diamond \longrightarrow S$ par $\varphi(a) = 1$ et $\varphi(b) = 0$. L'ensemble $a^\diamond$ est reconnaissable puisque $a^\diamond = \varphi^{-1}(\{1\})$.

Exemple 52 Soit $S = \{0, 1\}$ le $\diamond$-semigroupe fini de l'exemple 50 et soit $\varphi : A^\diamond \longrightarrow S$ par $\varphi(a) = 1$ pour tout $a \in A$. L'ensemble $A^\#$ est reconnaissable : $A^\# = \varphi^{-1}(\{1\})$.

Pour tout alphabet A, la classe des sous-ensembles de $A^\diamond$ reconnaissables par un $\diamond$-semigroupe fini est noté $Rec(A^\diamond)$. On sait qu'un ensemble de mots finis (respectivement de mots sur les ordinaux) est rationnel si et seulement si il est reconnu par un semigroupe fini (respectivement un ω_1-semigroupe fini). On généralise ce résultat au cas des mots sur les ordres linéaires dispersés.

Théorème 46 *Un ensemble de mots indexés par des ordres linéaires dispersés est rationnel si et seulement si il est reconnu par un $\diamond$-semigroupe fini.*

Les deux paragraphes suivants sont consacrés à la preuve de ce théorème.

4.3.2 Des automates vers les $\diamond$-semigroupes

Etant donné un automate acceptant un ensemble X de mots finis, on construit classiquement le semigroupe des matrices booléennes $Q \times Q$, où Q est l'ensemble des états. L'ensemble X est reconnu par le morphisme φ tel que pour tout mot x, l'entrée (q, q') de la matrice $\varphi(x)$ est vraie si et seulement s'il existe un chemin menant de q à q' et d'étiquette x. Cette construction a été généralisée pour les mots sur les ordinaux [4, 7] et s'adapte aussi à notre cas par symétrie.

Soit $\mathcal{A} = (Q, A, E, I, F)$ un automate sur les ordres linéaires acceptant $X \subseteq A^\diamond$. L'existence des transitions limites impose de mémoriser les contenus des chemins. Comme les automates ne sont pas déterminisables, il faut stocker tous les contenus possibles. Soit $T = \mathcal{P}(Q)$ l'ensemble des parties de Q et soit $K = \mathcal{P}(T)$ l'ensemble des parties de T. L'ensemble $K = \mathcal{P}(\mathcal{P}(Q))$ est équipé du produit fini et de l'union finie définis de la façon suivante pour tout $k, k' \in K$,

$$kk' = \{P \cup P' \mid P \in k,\ P' \in k'\} \ \text{ et } \ k + k' = k \cup k'.$$

Soit S l'ensemble des matrices $Q \times Q$ dont les entrées sont des éléments de K muni du produit fini défini pour tout $m, m' \in S$ par :

$$(m \cdot m')_{q,q'} = \bigcup_{p \in Q} m_{q,p} \cdot m'_{p,q'} = \{P \cup P' \mid \exists p \in Q,\, P \in m_{q,p},\, P' \in m'_{p,q'}\}$$

Le semigroupe S est fini et d'après le théorème 45, il suffit de définir des fonctions compatibles à droite et à gauche avec S pour induire une structure de $\diamond$-semigroupe. Soient les fonctions τ et $-\tau$ définies par :

$$m^{\tau}_{q,q'} = \{P \cup \{q'\} \mid \exists p \in Q,\, \exists P' \subseteq P,\, P \in m_{q,p},\, P' \in m^{\pi}_{p,p} \text{ et } P' \longrightarrow q' \in E\}$$

$$m^{-\tau}_{q,q'} = \{P \cup \{q\} \mid \exists p \in Q,\, \exists P' \subseteq P,\, P \in m_{p,q'},\, P' \in m^{\pi}_{p,p} \text{ et } q \longrightarrow P' \in E\}$$

où π est le plus petit entier tel que m^{π} soit une matrice idempotente. Les fonctions τ et $-\tau$ sont respectivement compatibles à droite et à gauche avec S. On définit le morphisme $\varphi : A^{\diamond} \longrightarrow S$ pour tout $a \in A$ par la matrice $\varphi(a)$ dont l'entrée (q, q') est égale à $\{\{q, q'\}\}$ si $q \xrightarrow{a} q' \in E$ et à l'ensemble vide sinon.

Lemme 47 *Pour tout $x \in A^{\diamond}$ et tous les états $q \in Q$ et $q' \in Q$,*

$$\varphi(x)_{q,q'} = \{P \mid q \underset{P}{\overset{x}{\Longrightarrow}} q'\}.$$

Preuve. Montrons tout d'abord que la propriété est stable par produit fini. Soient $x \in A^{\diamond}$ et $y \in A^{\diamond}$ vérifiant le lemme et soient $q, q' \in Q$.

$$\begin{aligned}(\varphi(x) \cdot \varphi(y))_{q,q'} &= \{P \cup P' \mid \exists p \in Q,\, P \in \varphi(x)_{q,p},\, P' \in \varphi(y)_{p,q'}\} \\ &= \{P \cup P' \mid \exists p \in Q,\, q \underset{P}{\overset{x}{\Longrightarrow}} p,\, p \underset{P'}{\overset{y}{\Longrightarrow}} q'\} \\ &= \{P \mid q \underset{P}{\overset{xy}{\Longrightarrow}} q'\}.\end{aligned}$$

On montre le lemme par induction sur le rang de $x \in A^{\diamond}$. Par ce qui précède et par définition de φ, le résultat est vrai pour tout $x \in A^*$.

Soit $n \in \mathbb{N}$ et $x \in A^{W_{n+1}}$. D'après ce qui précède, et comme tout ordre linéaire de W_{n+1} est un produit fini d'ordres linéaires de U_{n+1}, on peut supposer que $x \in A^{U_{n+1}}$. De plus par hypothèse d'induction et par symétrie, il suffit de considérer le cas où $x \in (A^{W_n})^{\omega}$. D'après le lemme 42, on dispose d'une factorisation $x = \prod_{i \in \omega} x_i$ et d'un couple lié à droite $(s, e) \in S \times E(S)$ tels que $\varphi(x_0) = s$ et pour tout $i > 0$, $\varphi(x_i) = e$. On montre le résultat pour $x' = \prod_{i>0} x_i$ d'image $\varphi(x') = e^{\tau}$. Supposons $q \underset{P}{\overset{x'}{\Longrightarrow}} q'$. Ce chemin est de la forme $q \underset{P_1}{\overset{x_1}{\Longrightarrow}} q_1 \underset{P_2}{\overset{x_2}{\Longrightarrow}} q_2 \ldots$ et termine par une transition limite $P' \to q'$. La suite $(q_i)_{i>1}$ admet au moins un élément p apparaissant une infinité de fois et quitte à refactoriser, on peut supposer que le chemin est de la forme $q \underset{P'_1}{\overset{x_1}{\Longrightarrow}} p \underset{P'}{\overset{x_2}{\Longrightarrow}} p \underset{P'}{\overset{x_3}{\Longrightarrow}} p \ldots$ avec $P'_1 \cup \{q'\} = P$. Par hypothèse d'induction, $P'_1 \in e_{q,p}$ et $P' \in e_{p,p}$ donc $P \in e^{\tau}_{q,q'}$. Réciproquement, supposons que $P \in e^{\tau}_{q,q'}$. Il existe $p \in Q$, $P_1 \in e_{q,p}$ et $P' \in e_{p,p}$ tels que $P = P_1 \cup \{q'\}$, $P' \subseteq P_1$, et $P' \to q' \in E$. Par hypothèse d'induction, $q \underset{P_1}{\overset{x_1}{\Longrightarrow}} p$ et pour tout $i > 1$, $p \underset{P'}{\overset{x_i}{\Longrightarrow}} p$ d'où l'existence d'un chemin

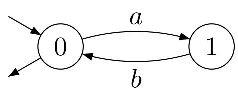
$q \underset{P}{\overset{x'}{\Longrightarrow}} q'$. $\square$

Un mot $x \in A^\diamond$ est accepté par l'automate si et seulement si il existe un état initial i et un état final f tels que $\varphi(x)_{i,f}$ soit non vide. Donc X est reconnu par S et $Rat(A^\diamond) \subseteq Rec(A^\diamond)$.

Proposition 48 *Tout langage rationnel de mots indexés par des ordres linéaires dispersés est reconnu par un $\diamond$-semigroupe fini.*

La construction est illustrée à travers quelques exemples.

Exemple 53 Soit $\mathcal{A}$ l'automate de la figure 4.3 acceptant le langage $X = (ab)^\diamond$. Soit S le $\diamond$-semigroupe des matrices $\{0,1\} \times \{0,1\}$. Notons



$$0 \to \{0,1\}$$

$$\{0,1\} \to 0$$

FIG. 4.3 – Automate acceptant $(ab)^\diamond$

$$s = \begin{bmatrix} \emptyset & \{\{0,1\}\} \\ \emptyset & \emptyset \end{bmatrix},\ t = \begin{bmatrix} \emptyset & \emptyset \\ \{\{0,1\}\} & \emptyset \end{bmatrix},\ e = \begin{bmatrix} \{\{0,1\}\} & \emptyset \\ \emptyset & \emptyset \end{bmatrix} \text{ et } f = \begin{bmatrix} \emptyset & \emptyset \\ \emptyset & \{\{0,1\}\} \end{bmatrix}$$

En notant 0 la matrice dont toutes les entrées sont l'ensemble vide, le produit du $\diamond$-semigroupe S est défini par :

$$\begin{array}{llll} st = e, & ts = f, & ee = e, & ff = f, \\ es = s, & ft = t, & sf = s, & te = t \end{array}$$

Les autres produits binaires valent tous 0. Pour les fonctions τ et $-\tau$, les valeurs non égales à 0 sont :

$$e^\tau = e, \quad e^{-\tau} = e, \quad f^\tau = t \quad f^{-\tau} = s$$

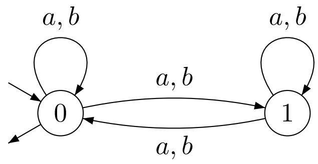
En définissant le morphisme par $\varphi(a) = s$ et $\varphi(b) = t$, il vient $X = \varphi^{-1}(e)$.

Exemple 54 Soit $\mathcal{A}$ l'automate de la figure 4.4 acceptant le langage $X = A^{-\#}(A^{\#}A^{-\#})^*$. Notons

$$a = \begin{bmatrix} \{\{0\}\} & \{\{0,1\}\} \\ \{\{0,1\}\} & \{\{1\}\} \end{bmatrix},\ e = \begin{bmatrix} \{\{0\},\{0,1\}\} & \{\{0,1\}\} \\ \{\{0,1\}\} & \{\{1\},\{0,1\}\} \end{bmatrix},$$

$$d = \begin{bmatrix} \emptyset & \{\{0,1\}\} \\ \emptyset & \{\{0,1\}\} \end{bmatrix},\ g = \begin{bmatrix} \{\{0\},\{0,1\}\} & \{\{0,1\}\} \\ \{\{0,1\}\} & \{\{0,1\}\} \end{bmatrix},$$

$$s = \begin{bmatrix} \{\{0,1\}\} & \{\{0,1\}\} \\ \{\{0,1\}\} & \{\{1\},\{0,1\}\} \end{bmatrix},\ h = \begin{bmatrix} \emptyset & \{\{0,1\}\} \\ \emptyset & \{\{0,1\}\} \end{bmatrix} \text{ et } f = \begin{bmatrix} \{\{0,1\}\} & \{\{0,1\}\} \\ \{\{0,1\}\} & \{\{0,1\}\} \end{bmatrix}$$



$0 \to \{0\}$

$1 \to \{0\} \qquad \{1\} \to 1$

FIG. 4.4 – Automate acceptant $X = A^{-\#}(A^{\#}A^{-\#})^{*}$

En notant 0 la matrice dont toutes les entrées sont l'ensemble vide, X est reconnu par le $\diamond$-semigroupe $S = \{a, e, g, d, s, h, f, 0\}$ dont le produit est défini par :

$e = aa = aer = ea,$
$d = ad = ed = dd = sd = a^{\tau} = e^{\tau} = d^{\tau} = s^{\tau},$
$g = ag = eg = ga = ge = gg = a^{-\tau} = e^{-\tau} = g^{-\tau},$
$s = sa = se = as = es = ds = ss = da = de,$
$h = gd = ah = eh = gh = hd = sh = fd = hh,$
$f = dg = sg = gs = ha = he = af = ef = fa = fe = hg = gf = fg = dh = df = hs$
$f = sf = fs = hf = fh = ff.$

Les autres produits étant tous égaux à 0. En définissant le morphisme $\varphi(a) = a$ et $\varphi(b) = a$, il vient $X = \varphi^{-1}(\{a, e, g, s, f\})$.

Exemple 55 On rajoute à l'automate de l'exemple 54 les transitions limites $0 \to \{0, 1\}$ et $\{0, 1\} \to 1$. Cet automate accepte alors l'ensemble $X = A^{-\#} \diamond A^{\#}$. On reprend le $\diamond$-semigroupe S auquel on rajoute les éléments

$$u = \begin{bmatrix} \emptyset & \{\{0,1\}\} \\ \emptyset & \emptyset \end{bmatrix} \text{ et } t = \begin{bmatrix} \{\{0,1\}\} & \{\{0,1\}\} \\ \emptyset & \emptyset \end{bmatrix}$$

Le produit du $\diamond$-semigroupe précédent est modifié par
$g^{\tau} = h$, $d^{\tau} = u$, $s^{\tau} = t$, $h^{\tau} = h$, $h^{-\tau} = u$, $f^{\tau} = h$ et $f^{-\tau} = t$ et complété par

$h = au = eu = gu = su = fu,$
$t = ua = ue = ug = us = uf = ta = te = tg = ts = tf = tt = t^{-\tau},$
$u = ud = uh = td = th = tu = t^{\tau},$
$f = at = et = gt = st = ft.$

Les autres produits étant tous égaux à 0. En définissant le même morphisme φ, il vient $X = \varphi^{-1}(\{a, e, g, s, f, t\})$.

4.3.3 Des $\diamond$-semigroupes vers les automates

Soit (S, π) un $\diamond$-semigroupe fini. D'après le théorème 45, le produit π est défini par un produit fini, une fonction τ compatible à droite avec S et une fonction $-\tau$ compatible à gauche avec S.

Soit X un langage de $A^{\diamond}$ reconnu par S. On dispose d'un morphisme de $\diamond$-semigroupe $\varphi : A^{\diamond} \longrightarrow S$ reconnaissant X. On va montrer que X est rationnel. Comme les ensembles rationnels sont fermés par union finie, on montrera que pour tout élément s de S, l'ensemble $\varphi^{-1}(s)$ est rationnel. Il suffit pour cela de prouver que pour tout $s \in S$, l'ensemble $\pi^{-1}(s)$ des mots de $S^\diamond$ ayant pour produit s, est rationnel.

Lemme 49 *Soit (S, π) un $\diamond$-semigroupe fini, $\varphi : A^\diamond \longrightarrow S$ un morphisme de $\diamond$-semigroupe et soit $s \in S$. Si $\pi^{-1}(s) \in Rat(S^\diamond)$ alors $\varphi^{-1}(s) \in Rat(A^\diamond)$.*

Preuve. Soit h la substitution rationnelle qui associe à chaque élément s de S l'ensemble des lettres de A d'image s par φ :

$$\begin{array}{rccc} h: & S & \rightarrow & \mathcal{P}(A) \\ & s & \mapsto & \varphi^{-1}(s) \cap A \end{array}$$

On montre que l'égalité suivante est vérifiée pour tout $s \in S$:

$$\varphi^{-1}(s) = h(\pi^{-1}(s) \cap \varphi(A)^\diamond)$$

Soit $x \in \varphi^{-1}(s)$. Notons $x = \prod_{j \in J} a_j$ où J est un ordre linéaire dispersé et où $a_j \in A$ pour tout $j \in J$. On a

$$x \in \prod_{j \in J}(\varphi^{-1}(\varphi(a_j)) \cap A) \subseteq \prod_{j \in J} h(\varphi(a_j)) = h(\prod_{j \in J} \varphi(a_j)) \subseteq h(\pi^{-1}(s) \cap \varphi(A)^\diamond).$$

Réciproquement, soit $x \in h(\pi^{-1}(s) \cap \varphi(A)^\diamond)$. Il existe $u \in \pi^{-1}(s) \cap \varphi(A)^\diamond$ tel que $x \in h(u)$. Notons $u = \prod_{j \in J} s_j$ où J est un ordre linéaire dispersé et où $s_j \in S$ pour tout $j \in J$. Comme $x \in \prod_{j \in J} h(s_j)$, on a

$$\varphi(x) = \pi(\prod_{j \in J} \varphi(h(s_j))) = \pi(\prod_{j \in J} \varphi(\varphi^{-1}(s_j) \cap A)) = \pi(\prod_{j \in J} s_j) = \pi(u) = s.$$

□

Illustrons cette réduction du problème.

Exemple 56 Soit $S = (\{t, e, f, 0\}, \pi)$ le $\diamond$-semigroupe dont le produit π est défini par le produit fini $ee = e$, $ff = ef = fe = f$, $tt = te = et = tf = ft = 0$ où l'élément 0 est un zéro, et par les fonctions compatibles suivantes : $e^\tau = e^{-\tau} = f$, $f^\tau = f^{-\tau} = t^\tau = t^{-\tau} = 0^\tau = 0^{-\tau} = 0$. On remarque que quel que soit $s \in S$, l'ensemble $\pi^{-1}(s)$ est rationnel : par exemple $\pi^{-1}(t) = t$, $\pi^{-1}(e) = e^+$ et $\pi^{-1}(f) = (e^* f e^* + e^* e^{-\omega} + e^\omega e^*)^+$.

Soit $\varphi : A^\diamond \rightarrow S$ le morphisme défini par $\varphi(a) = e$ et $\varphi(b) = t$ et soit $X = \varphi^{-1}(\{t, e, f\})$. On retrouve une expression rationnelle de X en utilisant la substitution h du lemme 49 :

$$\begin{array}{rcllllll} X & = & \varphi^{-1}(t) & \cup & \varphi^{-1}(e) & \cup & \varphi^{-1}(f) \\ & = & h(t) & \cup & h(e^+) & \cup & h((e^* e^{-\omega} + e^\omega e^*)^+) \\ & = & b & \cup & a^+ & \cup & (a^* a^{-\omega} + a^\omega a^*)^+ \end{array}$$

d'où $X = b + (a + a^{-\omega} + a^\omega)^+$.

Pour montrer qu'un langage reconnu par un $\diamond$-semigroupe fini S est rationnel, il suffit donc de construire, pour tout élément s de S, un automate sur les ordres linéaires acceptant le langage $\pi^{-1}(s)$. Dans le cas où S est un semigroupe -muni d'un produit fini- on construit naturellement l'automate sur les mots finis ayant S^1 pour ensemble d'états et $\{s \xrightarrow{t} s \cdot t \mid s \in S^1, t \in S\}$ pour transitions. Un mot $u = s_1 s_2 \ldots s_m$ est l'étiquette d'un chemin menant de 1 à s si et seulement si $s_1 \cdot s_2 \cdot \ldots s_m = s$.

Exemple 57 Soit $S = (e, f, 0)$ le semigroupe défini par $ee = e$, $ff = ef = fe = f$ et 0 est un zéro. L'automate fini $\mathcal{A} = (S^1, S, \{s \xrightarrow{t} s \cdot t\})$ est représenté en figure 4.5.

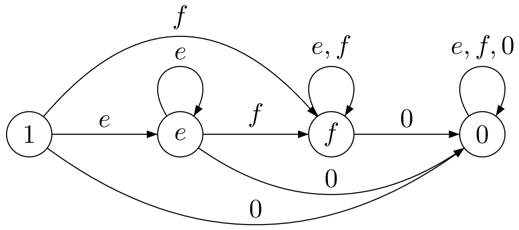


FIG. 4.5 – Automate $\mathcal{A} = (S^1, S, \{s \xrightarrow{t} s \cdot t\})$.

Par exemple, l'ensemble $\{s_1 s_2 \ldots s_m \mid s_1 \cdot s_2 \cdot \ldots s_m = f\}$ des mots finis dont le produit vaut f est l'ensemble des étiquettes des chemins menant de 1 à f c'est-à-dire le langage $e^* f (e + f)^*$.

Dans le cas des $\diamond$-semigroupes, on cherche à définir en plus des transitions limites. La principale difficulté provient du fait que l'ensemble cofinal d'un chemin ne suffit pas à en déterminer l'étiquette.

Exemple 58 Reprenons l'exemple 57 et étendons le semigroupe $S = (e, f, 0)$ en un $\diamond$-semigroupe en définissant les fonctions τ et $-\tau$ par $e^\tau = e^{-\tau} = f$, $f^\tau = f^{-\tau} = 0$ et bien sûr $0^\tau = 0^{-\tau} = 0$. Essayons de définir sur l'automate $\mathcal{A} = (S^1, S, \{s \xrightarrow{t} s \cdot t\})$ des transitions limites à gauche de sorte que si un mot u est l'étiquette d'un chemin de 1 à s alors $\pi(u) = s$. Intuitivement, on définirait la transition $\{e\} \to f$. En effet, tout chemin partant de 1, de longueur limite à droite et d'ensemble cofinal $\{e\}$ est d'étiquette e^ω dont le produit vaut $e^\tau = f$. Par contre, un chemin d'ensemble cofinal $\{f\}$ peut avoir pour étiquette f^ω aussi bien que fe^ω. Or $f^\tau = 0$ et $fe^\tau = f$. On ne peut donc pas définir la transition $\{f\} \to 0$.

Il faut ajouter une condition pour que toute étiquette u d'un chemin partant de 1 et terminant par la transition $P \to q$ vérifie $\pi(u) = q$. Lorsque deux chemins de longueur limite à gauche et d'étiquettes u et v admettent le même ensemble cofinal P, on va imposer que u et v admettent des factorisations ramseyennes associées à des couples conjugués. Cette condition garantira que $\pi(u) = \pi(v)$ et permettra de définir la transition $P \to q$.

Supposons qu'on cherche à définir des transitions limites sur un automate vérifiant que toute étiquette u d'un chemin menant de 1 à s est de produit $\pi(u) = s$ - c'est le cas de l'automate sur les mots finis-. Considérons deux chemins d'étiquettes u et v, de longueur limite à gauche et de même ensemble cofinal P. Quel que soit $p \in P$, il existe une infinité de préfixes w de u tels que $\pi(w) = p$ donc tous les éléments de P sont $\mathcal{R}$-équivalents. Or si u et v admettent des factorisations ramseyennes associées aux couples liés à droite (s, e) et (t, f), on remarque que $s \in P$ et $t \in P$ donc $s \mathcal{R} t$. Dans ce cas, la proposition 38 montre qu'il suffit d'imposer aux couples liés d'appartenir à la même $\mathcal{D}$-classe pour qu'ils soient conjugués.

On va construire, pour tout élément s de S, un automate sur les ordres linéaires acceptant le langage $\pi^{-1}(s)$ par induction sur la $\mathcal{D}$-classe de s. Pour chaque $\mathcal{D}$-classe D de S, on définit un automate tel qu'un chemin ne possède de transition limite que si son étiquette admet une factorisation ramseyenne associée à un couple lié de D. Pour cela, on ajoute aux états une composante booléenne qui vaut 1 lorsqu'on vient de lire un élément de D et 0 sinon. Une transition limite $P \to q$ ou $q \to P$ n'est définie que lorsque P contient un état dont la composante booléenne vaut 1. Cette condition garantit que le couple lié appartient à D et que l'automate calcule correctement le produit π des étiquettes de ses chemins.

Comme on doit également définir les transitions limites à gauche, on rajoute une composante à nos états pour lire les étiquettes des chemins de la droite vers la gauche. Une seule composante booléenne sera suffisante pour les deux cotés. Lorsqu'on lit de droite à gauche, les états de l'ensemble cofinal des chemins de longueur limite à gauche sont alors $\mathcal{L}$-équivalents et on utilisera le symétrique de la proposition 38.

Pour toute $\mathcal{D}$-classe D de S, on note

$$S_D = \{s \in S \mid \forall p \in D, s \geq_{\mathcal{J}} p\}$$

et on définit l'automate $\mathcal{A}_D = (Q_D, S_D, E_D)$ par :

$$\begin{aligned}
Q_D &= {S_D}^1 \times {S_D}^1 \times \mathbb{B} \text{ est l'ensemble des états où } \mathbb{B} = \{0,1\} \\
E_D &= \{(s, rt, b) \xrightarrow{r} (sr, t, b') \mid b \in \mathbb{B}, b' = (r \in D)\} \\
&\cup \ \{\{(s_i, t_i, b_i)\}_{1 \leq i \leq m} \longrightarrow (s, t, b) \mid b \in \mathbb{B}, \exists 1 \leq i \leq m \ , \ b_i = 1 \ , \\
&\qquad \exists 1 \leq k \leq m \ , \ \exists e \in E(D) \ , \ s_k e = s_k \ , \ e t_k = t_k \ , \ s = s_k e^{\tau} \text{ et } t_k = e^{\tau} t\} \\
&\cup \ \{(s, t, b) \longrightarrow \{(s_i, t_i, b_i)\}_{1 \leq i \leq m} \mid b \in \mathbb{B}, \exists 1 \leq i \leq m \ , \ b_i = 1 \ , \\
&\qquad \exists 1 \leq k \leq m \ , \ \exists e \in E(D) \ , \ s_k e = s_k \ , \ e t_k = t_k \ , \ s_k = s e^{-\tau} \text{ et } t = e^{-\tau} t_k\}.
\end{aligned}$$

Si on se restreint aux transitions successeurs, un mot $u \in S_D^*$ est l'étiquette d'un chemin menant de $(1, s, 0)$ à $\{(s, 1, b) \mid b \in \mathbb{B}\}$ si et seulement si $\pi(u) = s$. D'ailleurs lorsque la $\mathcal{D}$-classe n'est pas régulière, l'automate $\mathcal{A}_D$ n'accepte que des mots finis. Considérons un chemin de $\mathcal{A}_D$ de longueur limite à droite et d'ensemble cofinal $P = \{(s_i, t_i, b_i)\}_{1\leq i\leq m}$. Supposons que son étiquette admette une factorisation ramseyenne associée au couple lié (s_0, e_0). On a vu que s_0 était $\mathcal{R}$-équivalent à toutes les premières composantes $\{s_i\}_{1\leq i\leq m}$ de P. De plus, si P contient au moins un état dont la composante booléenne est positive, on sait que (s_0, e_0) appartient à D^2. Grâce à la proposition 38, on peut calculer $s = s_0 e_0^\tau$ à partir de n'importe quel couple lié (s_k, e) satisfaisant $s_k \mathcal{R} s$. Il reste à vérifier que l'étiquette est bien lue de droite à gauche avec $t_k = e^\tau t$. Montrons que cet automate calcule correctement le produit π de ses étiquettes dans les deux sens.

Lemme 50 *S'il existe un chemin de $\mathcal{A}_\mathcal{D}$ menant d'un état $q = (s, t, b) \in Q_D$ à un état $q' = (s', t', b') \in Q_D$ et d'étiquette $u \in S^\diamond$ alors $s' = s\pi(u)$ et $t = \pi(u)t'$.*

Preuve. On montre par induction sur $\alpha \in \mathcal{O}$ que le lemme est vérifié pour tout mot u de rang α.

Lorsque $\alpha = 0$, le résultat est immédiat par définition de $\mathcal{A}_D$.

Soit $\alpha > 0$. Supposons l'existence d'un chemin γ de $\mathcal{A}_\mathcal{D}$ menant d'un état $q = (s, t, b)$ à un état $q' = (s', t', b')$ et d'étiquette $u \in S^{W_\alpha}$. Notons $I = |u|$ la longueur de u.

Supposons dans un premier temps que $I \in U_\alpha$. L'ordre linéaire I peut s'écrire en une somme $I = \sum_{j\in J} K_j$ où pour tout $j \in J$, K_j est de rang strictement inférieur à α : $K_j \in \bigcup_{\beta<\alpha} W_\beta$ et où $J \in \mathcal{N} \cup \{\omega, -\omega\}$.

Dans le cas où $J \in \mathcal{N}$, on peut utiliser l'hypothèse d'induction puisque l'ordre I est alors de rang strictement inférieur à α.

Supposons que $J = \omega$ (le cas $J = -\omega$ est traité de façon symétrique). On dispose d'une ω-factorisation $u = \prod_{j\in\omega} u_j$ et d'un couple lié à droite (r, e) tel que $\pi(u_0) = r$ et $\pi(u_j) = e$ pour tout $j > 0$. Chaque facteur u_j est de rang strictement inférieur à α. Par définition du produit π, on a $\pi(u) = re^\tau$. On cherche à montrer que $s' = sre^\tau$ et $t = re^\tau t'$. Le chemin γ d'étiquette u se termine par une transition limite à gauche $P \longrightarrow q'$. L'ensemble cofinal $P = cof_{(I,\emptyset)^-}\gamma$ est noté $P = \{(s_1, t_1, b_1), \ldots, (s_m, t_m, b_m)\}$ et le chemin γ est représenté de la façon suivante :

$$\gamma : q = q_0 \overset{u_0}{\Longrightarrow} q_1 \overset{u_1}{\Longrightarrow} q_2 \overset{u_2}{\Longrightarrow} q_3 \ldots \; P \longrightarrow q'$$

Par hypothèse d'induction, pour tout $j > 0$, la première composante de l'état q_j que l'on note s_{q_j} satisfait $s_{q_j} = sr$, $s_{q_j} e = s_{q_j}$ et la seconde composante t_{q_j} vérifie $t = rt_{q_j}$ et $et_{q_j} = t_{q_j}$. Il existe donc un état de l'ensemble cofinal P vérifiant ces égalités. Soit $1 \leq k \leq m$ tel que l'état $p_k \in P$ vérifie $s_k = sr$, $s_k e = s_k$, $t = rt_k$ et $et_k = t_k$. Comme $P \longrightarrow q' \in E_D$, il existe au moins un état de P dont la composante booléenne est égale à 1. Par construction, cela signifie que $e \in D$. En effet, pour tout élément $i \in I$, il existe un indice supérieur $i' > i$ tel que la lettre de u indexée par i' soit dans D. Quitte à refactoriser, on peut supposer que pour tout $j > 0$, le facteur u_j contient une lettre $d_j \in D$. On a alors $\pi(u_j) = e \leq_J d_j$. D'autre part $e \geq_J sre$ et par hypothèse d'induction $sre = s_{q_j} \in S_D$ donc $e \in D$.

D'autre part, l'appartenance de la transition $P \longrightarrow q'$ à l'ensemble E_D des transitions donne l'existence d'un idempotent $f \in E(D)$ et de $1 \leq l \leq m$ tels que l'état $p_l \in P$ vérifie $s_l = s_l f$, $t_l = f t_l$, $s' = s_l f^\tau$ et $t_l = f^\tau t'$.

Montrons tout d'abord que $s' = s\pi(u)$. Comme $e \in D$ et $f \in D$ et que les couples (s_k, e) et (s_l, f) sont liés à droite, on sait que $s_k \in D$ et $s_l \in D$. Comme $p_k \in P$ et $p_l \in P$, on a $s_k \mathcal{R} s_l$ par construction. D'après le lemme 38, ces couples sont conjugués à droite. Enfin le lemme 44 donne $s_l f^\tau = s_k e^\tau$ i.e. $s' = sre^\tau = s\pi(u)$.

Symétriquement, (e, t_k) et (f, t_l) sont des couples liés à gauche de D et $t_k \mathcal{L} t_l$. D'après le lemme 38, ces couples sont conjugués à gauche donc on dispose de $a, b \in S$ tels que $e = ab$, $f = ba$, $t_l = bt_k$ et $t_k = at_l$. On obtient alors $t = rt_k = rat_l = raf^\tau t' = ra(ba)^\tau t' = re^\tau t' = \pi(u)t'$ ce qui conclut le cas où $|u| \in U_\alpha$.

Supposons maintenant que $I \in W_\alpha$. La longueur de u est une somme finie $I = \sum_{j=0}^{n} K_j$ où pour tout $0 \leq j \leq n$, $K_j \in U_\alpha$. Soit $u = \prod_{j=0}^{n} u_j$ la factorisation associée. Le chemin γ est noté $q = q_0 \overset{u_0}{\Longrightarrow} q_1 \overset{u_1}{\Longrightarrow} q_2 \ldots \overset{u_n}{\Longrightarrow} q_{n+1} = q'$. D'après le cas précédent, pour tout $0 \leq j \leq n$, $s_{q_j}\pi(u_j) = s_{q_{j+1}}$ et $t_{q_j} = \pi(u_j)t_{q_{j+1}}$. Donc $s\pi(u) = s'$ et $t = \pi(u)t'$. □

Pour tout $s \in S_D$, on note $\mathcal{A}_s$ l'automate $\mathcal{A}_D$ qui a $\{(1, s, 0)\}$ pour état initial et $\{(s, 1, b) \mid b \in \mathbb{B}\}$ pour ensemble d'états finaux. On note aussi L_s l'ensemble des mots acceptés par $\mathcal{A}_s$. Le lemme précédent montre que pour tout $s \in S_D$, $L_s \subseteq \pi^{-1}(s)$. Cette inclusion est stricte. A cause des restrictions sur les transitions limites, un mot $u \in S_D^\diamond$ dont le produit $\pi(u)$ appartient à D n'est pas toujours étiquette d'un chemin de $\mathcal{A}_D$. Afin de décrire l'ensemble L_s des mots acceptés, on donne quelques propriétés sur les chemins de $\mathcal{A}_D$.

Lemme 51 *Si un mot $u \in S^\diamond$ est l'étiquette d'un chemin $\gamma = ((s_c, t_c, b_c))_{c \in |\hat{u}|}$ de $\mathcal{A}_D$, alors quels que soient s et t appartenant à S_D, le chemin $((ss_c, t_c t, b_c))_{c \in |\hat{u}|}$ est également un chemin de $\mathcal{A}_D$ d'étiquette u.*

Preuve. La preuve se fait par induction sur $\alpha \in \mathcal{O}$ pour tout $u \in S^{W_\alpha}$.

Dans le cas où $\alpha = 0$, le résultat est immédiat par définition de $\mathcal{A}_D$.

Soit $\alpha > 0$, $I \in W_\alpha$ et soit $u \in S^I$. Supposons l'existence d'un chemin $\gamma = ((s_c, t_c, b_c))_{c \in \hat{I}}$ de $\mathcal{A}_\mathcal{D}$ d'étiquette u et soient s, $t \in S_D$.

On suppose dans un premier temps que $I \in U_\alpha$. L'ordre I peut être décomposé en une somme $I = \sum_{j \in J} K_j$ où pour tout $j \in J$, K_j est de rang strictement inférieur à α et où $I \in \mathcal{N} \cup \{\omega, -\omega\}$.

Lorsque $J \in \mathcal{N}$, le résultat est obtenu par hypothèse d'induction.

Supposons que $J = \omega$ (le cas $J = -\omega$ peut être traité de façon symétrique). Soit $u = \prod_{j \in \omega} u_j$ la factorisation telle que $|u_j| = K_j$ pour tout $j \in \omega$. On note $c_0 = (\emptyset, I)$ et pour tout $j > 0$ on note $c_j = (\bigcup\limits_{0\leq i<j} K_i, \bigcup\limits_{j\leq i} K_i)$ la coupure de $\hat{I}$ correspondant à la factorisation $u = (u_0 \ldots u_{j-1})(u_j \ldots)$. Par hypothèse d'induction, pour tout $j \in \omega$, $((ss_c, t_ct, b_c))_{c_j \leq c \leq c_{j+1}}$ est un chemin de $\mathcal{A}_D$ d'étiquette u_j. Il reste à montrer que le chemin $\gamma' = ((ss_c, t_ct, b_c))_{c\in\hat{I}\setminus\{(I,\emptyset)\}}$ est suivi d'une transition limite. Notons $P = cof_{(I,\emptyset)^-}\gamma$ et $P' = cof_{(I,\emptyset)^-}\gamma'$. Par construction, si $P = \{(s_i, t_i, b_i)\}_{1\leq i\leq m}$ alors $P' = \{(ss_i, t_it, b_i)\}_{1\leq i\leq m}$. Comme $P \longrightarrow (s_{(J,\emptyset)}, t_{(J,\emptyset)}, b_{(J,\emptyset)}) \in E_D$, il existe $1 \leq i \leq m$, $b_i = 1$. Il existe aussi $1 \leq k \leq m$ et $e \in E(D)$ tels que $s_ke = s_k$, $et_k = t_k$, $s_{(J,\emptyset)} = s_ke^\tau$ et $t_k = e^\tau t_{(J,\emptyset)}$. On a alors $ss_ke = ss_k$, $et_kt = t_kt$, $ss_{(J,\emptyset)} = ss_ke^\tau$ et $t_kt = e^\tau t_{(J,\emptyset)}t$ ce qui montre que $P' \longrightarrow (ss_{(J,\emptyset)}, t_{(J,\emptyset)}t, b_{(J,\emptyset)}) \in E_D$.

Supposons à présent que $I \in W_\alpha$. L'ordre I est une somme finie $I = \sum\limits_{j=0}^{n} K_j$ où pour tout $0 \leq j \leq n$, $K_j \in U_\alpha$. En gardant les même notations et en utilisant le cas précédent, $((ss_c, t_ct, b_c))_{c_j \leq c \leq c_{j+1}}$ est un chemin de $\mathcal{A}_D$ d'étiquette u_j pour tout $0 \leq j \leq n$. □

Notons également que la composante booléenne du premier état d'un chemin n'a pas d'importance, par définition de $\mathcal{A}_D$.

Lemme 52 *Si un mot $u \in S^\diamond$ de longueur I est l'étiquette d'un chemin $\gamma = ((s_c, t_c, b_c))_{c\in\hat{I}}$ de $\mathcal{A}_D$, alors pour tout booléen $b \in \mathbb{B}$, le chemin*

$$(s_{(\emptyset,I)}, t_{(\emptyset,I)}, b)((s_c, t_c, b_c))_{c\in\hat{I}\setminus\{(\emptyset,I)}$$

est également un chemin de $\mathcal{A}_D$ d'étiquette u.

Grâce aux lemmes 51 et 52 on peut concaténer les chemins de $\mathcal{A}_D$, un nombre fini de fois tout d'abord.

Lemme 53 $\forall s \in S_D$, $\forall t \in S_D$, $L_sL_t \subseteq L_{st}$.

Preuve. Soient $s, t \in S_J$. Supposons l'existence de chemins

$$(1, s, 0) \overset{u}{\Longrightarrow} (s, 1, b) \text{ et } (1, t, 0) \overset{v}{\Longrightarrow} (t, 1, b')$$

d'étiquettes $u \in L_s$ et $v \in L_t$. D'après les lemmes 51 et 52,

$$(1, st, 0) \overset{u}{\Longrightarrow} (s, t, b) \text{ et } (s, t, b) \overset{v}{\Longrightarrow} (st, 1, b').$$

sont également des chemins de $\mathcal{A}_D$ d'étiquettes u et v. D'où $(1, st, 0) \overset{uv}{\Longrightarrow} (st, 1, b')$ i.e. $uv \in L_{st}$. □

Supposons que D possède un idempotent e et considérons un mot $u \in L_e^\omega$. Le lemme précédent nous permet de construire un chemin de longueur limite à droite dont l'étiquette coïncide avec u. Si la suite des occurrences de lettres de u appartenant à D est cofinale, alors l'ensemble cofinal possède une composante booléenne positive et le chemin se termine par une transition limite. Il se peut également qu'il existe un indice à partir duquel u ne possède plus d'occurrence de lettres de D. Dans ce cas, toutes les composantes booléennes de l'ensemble cofinal sont nulles et la transition limite est interdite.

Pour tout $d \in D$, on note T_D l'ensemble des éléments de S strictement $\mathcal{J}$-supérieurs à d :

$$T_D = \{s \in S \mid \forall p \in D, s >_{\mathcal{J}} p\}.$$

Le mot u admet alors un suffixe sur l'alphabet T_D. Par construction, si un mot sur l'alphabet T_D est l'étiquette d'un chemin de $\mathcal{A}_D$, il est nécessairement de longueur finie. Notons E_d l'ensemble des mots finis sur l'alphabet T_D de produit d :

$$E_d = \{s_1 s_2 \dots s_m \in T_D^+ \mid s_1 \dots s_m = d\}$$

Lemme 54 *Soit $e \in E(D)$ tel que $e^\tau \in D$ (respectivement $e^{-\tau} \in D$). Alors ${L_e}^\omega \subseteq L_{e^\tau} \cup L_e {E_e}^\omega$ (respectivement ${L_e}^{-\omega} \subseteq L_{e^{-\tau}} \cup {E_e}^{-\omega} L_e$).*

Preuve. Soit $e \in E(D)$ tel que $e^\tau \in D$ et soit $u \in {L_e}^\omega$ un mot de longueur $I \in \mathcal{S}$. Soit $u = \prod_{j \in \omega} u_j$ la factorisation telle que pour tout $j \in \omega$, $u_j \in L_e$. Pour chaque $j \in \omega$, on dispose d'un chemin acceptant γ_j d'étiquette u_j :

$$\gamma_j : (1, e, 0) \xRightarrow{u_j} (e, 1, b_j).$$

D'après les lemmes 51 et 52, les chemins

$$\gamma'_0 : (1, e^\tau, 0) \xRightarrow{u_0} (e, e^\tau, b_0) \text{ et } \gamma'_j : (e, e^\tau, b_{j-1}) \xRightarrow{u_j} (e, e^\tau, b_j)$$

sont aussi d'étiquette u_j pour tout $j > 0$. On a donc construit un chemin γ' de longueur limite à droite dont l'étiquette coïncide avec u. Voyons s'il est suivi d'une transition limite. Soit $P = cof_{(I,\emptyset)^-}\gamma'$. L'ensemble $P = \{(s_i, t_i, b_i)\}_{1 \leq i \leq m}$ contient au moins un état de la forme (e, e^τ, b) pour un certain booléen b. Donc on dispose de $1 \leq k \leq m$ tel que $s_k = e$, $t_k = e^\tau$, $s_k e = e = s_k$ et $e t_k = e^\tau = t_k$. Il se dégage alors deux cas suivant que P contient au moins une composante booléenne égale à 1 ou non :

- Il existe $1 \leq l \leq m$ tel que $b_l = 1$. Alors $P \longrightarrow (e^\tau, 1, 0) \in E_D$ et $u \in L_{e^\tau}$.
- Pour tout $1 \leq l \leq m$, $b_l = 0$. Notons $u = \prod_{i \in I} a_i$ avec $a_i \in S$ pour tout $i \in I$. Par construction, il existe $i_0 \in I$ tel que pour tout $i > i_0$, $a_j \in T_D$. Donc $u \in L_e(L_e \cap T_D^\diamond)^\omega \subseteq L_e(L_e \cap T_D^+)^\omega \subseteq L_e {E_e}^\omega$.

L'inclusion ${L_e}^{-\omega} \subseteq L_{e^{-\tau}} \cup {E_e}^{-\omega} L_e$ est symétrique. □

Pour tout $d \in D$, le langage L_d accepté par $\mathcal{A}_d$ est strictement inclus dans l'ensemble $\pi^{-1}(d)$. Le lemme précédent montre que $\pi^{-1}(d)$ n'est pas inclu dans L_d. En effet, les mots sur l'alphabet T_D de rang non nul ne sont étiquettes d'aucun chemin de $\mathcal{A}_D$. On va résoudre ce problème en utilisant l'induction sur la $\mathcal{D}$-classe. Un mot de $\pi^{-1}(d)$ peut s'écrire comme le produit fini de mots de $\mathcal{D}$-classes strictement $\mathcal{J}$-supérieures à D. Pour tout $d \in D$, on définit

$$F_d = \bigcup_{\substack{s_1,\ldots,s_m >_{\mathcal{J}} d, \\ s_1 \ldots s_m = d}} \pi^{-1}(s_1) \ldots \pi^{-1}(s_m).$$

Dans le cas où il existe un idempotent e appartenant à une $\mathcal{D}$-classe strictement $\mathcal{J}$-supérieure à D tel que e^τ ou $e^{-\tau}$ tombe dans D, les mots de $\pi^{-1}(d)$ peuvent aussi provenir d'un ω-produit ou d'un $-\omega$-produit. Pour tout $d \in D$, on définit

$$G_d = \bigcup_{\substack{t,e >_{\mathcal{J}} d, \\ te^\tau = d}} \pi^{-1}(t)\pi^{-1}(e)^\omega \cup \bigcup_{\substack{t,e >_{\mathcal{J}} d, \\ e^{-\tau} t = d}} \pi^{-1}(e)^{-\omega}\pi^{-1}(t).$$

Pour tout $p \in D$, on substitue dans L_p toute lettre $d \in D$ par les ensembles F_d et G_d. On obtient ainsi une expression rationnelle de $\pi^{-1}(p)$.

Proposition 55 *Soit (S, π) un $\diamond$-semigroupe fini. Quel que soit $p \in S$, l'ensemble $\pi^{-1}(p)$ est rationnel.*

Preuve. Soit $p \in S$. On montre que $\pi^{-1}(p)$ est rationnel par induction sur la $\mathcal{D}$-classe D de p.

Supposons que D est $\mathcal{J}$-maximale i.e. $\forall s \in S$, si $p \leq_{\mathcal{J}} s$ alors $s \in D$. Dans ce cas, on montre que $\pi^{-1}(p) = L_p$. D'après le lemme 50, $L_p \subseteq \pi^{-1}(p)$. Réciproquement, on montre que pour tout $u \in \pi^{-1}(p)$, $u \in L_p$ par induction sur le rang de u. Par définition de l'automate $\mathcal{A}_D$, l'inclusion est vérifiée pour les mots finis. D'autre part lorsque D est $\mathcal{J}$-maximale, l'ensemble E_d est vide pour tout $d \in D$. Si D possède un idempotent e tel que $e^\tau \in D$ (respectivement $e^{-\tau} \in D$), le lemme 54 donne alors ${L_e}^\omega \subseteq L_{e^\tau}$ (respectivement ${L_e}^{-\omega} \subseteq L_{e^{-\tau}}$). Il s'ensuit que $L_p = \pi^{-1}(p)$ ce qui conclut le cas de base de l'induction sur la $\mathcal{D}$-classe.

Par hypothèse d'induction, pour tout $s \in T_D$, $\pi^{-1}(s)$ est rationnel. Notons que pour tout $d \in D$, F_d est aussi rationnel puisque l'ensemble E_d des mots finis sur T_D de produit d est accepté par l'automate fini $({T_D}^1, T_D, \{s \xrightarrow{r} sr\})$. L'ensemble G_d est également rationnel en tant qu'union finie d'ensembles rationnels. On définit la substitution rationnelle f par

$$\begin{array}{llll} f: & S_D & \longrightarrow & Rat({S_D}^\diamond) \\ & s & \longrightarrow & \begin{cases} \pi^{-1}(s) & \text{si } s \in T_D \\ \{s\} \cup F_s \cup G_s & \text{si } s \in D \end{cases} \end{array}$$

et on montre que $f(L_p) = \pi^{-1}(p)$.

D'après le lemme 50, $L_p \subseteq \pi^{-1}(p)$. Par définition, f préserve le produit : pour tout $u \in {S_D}^\diamond$, $\pi(f(u)) = \pi(u)$. Donc $f(L_p) \subseteq \pi^{-1}(p)$. Réciproquement, on montre que pour tout $u \in \pi^{-1}(p)$, $u \in f(L_p)$ par induction sur le rang de u. Soit u un mot fini satisfaisant $\pi(u) = p$. Par définition de f, $u \in f(u)$ et par définition de $\mathcal{A}_D$, $u \in L_p$ donc $u \in f(L_p)$. Ceci conclut le cas du rang 0.

Soit $\alpha > 0$ et soit $u \in S^{W_\alpha} \cap \pi^{-1}(p)$. Notons $I = |u|$ la longueur de u.

Supposons tout d'abord que $I \in U_\alpha$. L'ordre linéaire I peut être décomposé en une somme $I = \sum_{J \in J} K_J$ telle que pour tout $j \in J$, $K_j \in \bigcup_{\beta<\alpha} W_\beta$ et $J \in \mathcal{N} \cup \{\omega, -\omega\}$. Dans le cas où $J \in \mathcal{N}$, $I \in \bigcup_{\beta<\alpha} W_\beta$ et le résultat est obtenu par hypothèse d'induction. Supposons que $J = \omega$. On dispose alors d'un couple lié à droite$(s,e) \in S_D \times E(S_D)$ et d'une factorisation $u = \prod_{j\in\omega} u_j$ tels que $\pi(u_0) = s$ et pour tout $j > 0$, $\pi(u_j) = e$ avec $|u_j| \in \bigcup_{\beta<\alpha} W_\beta$ pour tout $j \in \omega$. Notons que $s \leq_{\mathcal{J}} e$ car (s,e) est un couple lié à droite et que $e^\tau \in S_D$ puisque $\pi(u) \in D$. On distingue les cas suivants :

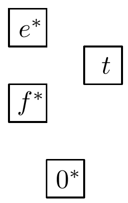
- $s \in T_D$ et $e \in T_D$:
  dans ce cas $u \in \pi^{-1}(s)\pi^{-1}(e)^\omega \subseteq G_p \subseteq f(p) \subseteq f(L_p)$.
- $s \in D$ et $e \in T_D$:
  par hypothèse d'induction sur le rang, $u \in f(L_s)\pi^{-1}(e)^\omega$. Notons $t = e^\tau$.
  - Si $t \in T_D$, alors $u \in f(L_s)\pi^{-1}(t)$.
  - Si $t \in D$, alors $u \in f(L_s)G_t$.

  Dans les deux cas $u \in f(L_s)f(t) \subseteq f(L_s)f(L_t) \subseteq f(L_p)$.
- $s \in D$ et $e \in D$:
  Par hypothèse d'induction sur le rang, $u \in f(L_s)f(L_e)^\omega = f(L_s L_e^\omega)$. D'après le lemme 54, $L_s L_e^\omega \subseteq L_s L_{e^\tau} \cup L_s E_e^\omega$.
  - Si $u \in f(L_s L_{e^\tau})$ alors $u \in f(L_p)$.
  - Sinon $u \in f(L_s)f(E_e)^\omega = f(L_s)F_e^\omega \subseteq f(L_s)f(e)^\omega = f(L_s)f(e^\omega) \subseteq f(L_s)f(L_{e^\tau}) \subseteq f(L_p)$.

Le cas $J = -\omega$ est symétrique ce qui conclut le cas où $I \in U_\alpha$.

Supposons que $I \in W_\alpha$. On dispose d'une factorisation finie $u = \prod_{j=1}^{n} u_j$ telle que pour tout $1 \leq j \leq n$, $|u_j| \in U_\alpha$ donc $u_j \in f(L_{s_j})$ en notant $s_j = \pi(u_j)$. Et $u \in \prod_{j=1}^{n} f(L_{s_j}) = f(L_{s_1} \dots L_{s_n}) \underset{Lemme\ 54}{\subseteq} f(L_p)$. $\square$

Exemple 59 Reprenons l'exemple 56. On représente la structure en $\mathcal{D}$-classes du $\diamond$-semigroupe $S = \{0, t, e, f\}$ de la façon suivante :



La $\mathcal{D}$-classe de t est $\mathcal{J}$-maximale. Comme elle n'est pas régulière, l'automate $\mathcal{A}_t$ n'a pas de transition limite et $L_t = t = \pi^{-1}(t)$. La $\mathcal{D}$-classe de e est aussi

$\mathcal{J}$-maximale mais régulière. Comme $e^\tau = e^{-\tau} = f >_{\mathcal{J}} e$, l'automate $\mathcal{A}_e$ est également un automate sur les mots finis et $L_e = e^+ = \pi^{-1}(e)$. De même pour f, on obtient $L_f = (e^* f e^*)^+$. En reprenant les même notations on a $E_f = F_f = \emptyset$ et $G_f = \pi^{-1}(e)^{-\omega} \cup \pi^{-1}(e)^{\omega} = e^{-\omega} + e^{-\omega}$. Il vient $f(t) = t$, $f(e) = e^+$ et $f(f) = f + e^{-\omega} + e^{-\omega}$ d'où $\pi^{-1}(f) = f(L_f) = (e^*(f + e^{-\omega} + e^{-\omega})e^*)^+$.

Exemple 60 Soit $S = \{0, a, b, s, e, f\}$ le $\diamond$-semigroupe fini ayant 0 pour zéro et dont le produit π est défini par le produit fini

$$aa = a \ , \ ee = sb = e \ , \ ff = bs = fa = af = f,$$

$$es = sf = sa = s \ , \ be = fb = ab = b,$$

les autres produits finis étant égaux à 0. On définit les fonctions τ et $-\tau$ par

$$a^\tau = a \ , \ a^{-\tau} = s \ , \ e^\tau = e \ , \ f^\tau = b,$$

les autres produits étant égaux à 0. La structure en $\mathcal{D}$-classe du $\diamond$-semigroupe S est représentée de la façon suivante :

a^*

s	e^*
f^*	b

0^*

L'automate $\mathcal{A}_a$ de la $\mathcal{D}$-classe $\mathcal{J}$-maximale de S est le suivant :

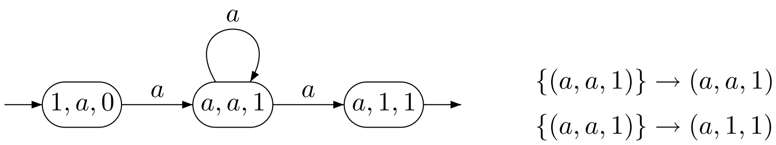
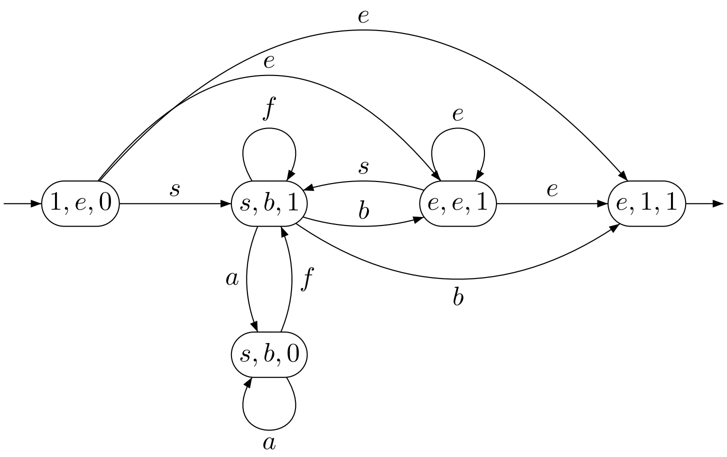


FIG. 4.6 – Automate $\mathcal{A}_a$ acceptant le langage $L_a = a^{\#}$.

On donne à présent l'automate $\mathcal{A}_e$ en figure 4.7.
Calculons l'ensemble $\pi^{-1}(e)$. Pour tout $d \in D = \{s, e, b, f\}$ les ensembles E_d et F_d sont vides. Le seul ensemble G_d non vide est $G_s = (\pi^{-1}(a))^{-\omega} = (a^{\#})^{-\omega}$. On en déduit $\pi^{-1}(e) = f(L_e) = (e + (s + (a^{\#})^{-\omega})(a^{\#} f)^* b)^{\#}$. En utilisant le lemme 49, on sait que tous les langages reconnus par S sont rationnels. En définissant le morphisme $\varphi : A^\diamond \to S$ par $\varphi(a) = a$ et $\varphi(b) = b$, on retrouve le langage $\varphi^{-1}(e) = ((a^{\#})^{-\omega} b)^{\#}$.



$\{(e,e,1)\} \to (e,1,1)$
$\{(e,e,1)\} \to (e,e,1)$
$\{(e,e,1),(s,b,1)\} \to (e,1,1)$
$\{(e,e,1),(s,b,1)\} \to (e,e,1)$
$\{(s,b,1)\} \to (e,1,1)$
$\{(s,b,1)\} \to (e,e,1)$

FIG. 4.7 – Automate $\mathcal{A}_e$ acceptant le langage $L_e = (e + s(a^*f)^*b)^{\#}$.

On déduit de la proposition 55 et du lemme 49, que tout langage reconnu par un $\diamond$-semigroupe fini est rationnel. La proposition 48 donne la réciproque ce qui termine la preuve du théorème 46.

On a obtenu une caractérisation algébrique des langages rationnels de $A^\diamond$ qui généralise la triple équivalence entre automates, expressions rationnelles et semigroupes finis. On a aussi prouvé que $Rat(A^\diamond)$ est fermé par complémentation.

Théorème 56 *L'ensemble des langages rationnels de mots indexés par des ordres linéaires dispersés est fermé par complémentation.*

Preuve. Soit $X \in Rat(A^\diamond)$. D'après le théorème **??**, on dispose d'un $\diamond$-semigroupe fini S reconnaissant X. Soient $\varphi : A^\diamond \to S$ et $P \subseteq S$ tels que $X = \varphi^{-1}(P)$. On a $A^\diamond \setminus X = \varphi^{-1}(S \setminus P)$. Donc $\bar{X}$ est rationnel. $\square$

On vient de montrer que pour tout langage rationnel de $A^\diamond$, il existe un $\diamond$-semigroupe fini le reconnaissant. On va montrer qu'il en existe un qui est minimal : le $\diamond$-semigroupe syntaxique.

4.4 $\diamond$-semigroupe syntaxique

Concernant les mots finis et les mots sur les ordinaux, on a vu que pour tout ensemble reconnaissable, il existe un semigroupe minimal le reconnaissant : le semigroupe syntaxique. On montre dans ce paragraphe que cette propriété se généralise aux mots indexés par les ordres linéaires. Soit X un sous-ensemble reconnaissable de $A^\diamond$. Parmi tous les $\diamond$-semigroupes reconnaissant X, il en existe un qui est minimal en nombre d'éléments au sens de la division. Il est appelé $\diamond$-*semigroupe syntaxique* de X et noté $S(X)$. Comme on ne peut pas minimaliser les automates sur les ordres linéaires, c'est le premier objet canonique associé à un ensemble rationnel de $A^\diamond$. Pour tout $\diamond$-semigroupe (S,π) et tout ensemble $P \subseteq S$, la relation d'équivalence $\sim_P$ est définie quels que soient s,t appartenant à S par $s \sim_P t$ ssi pour tout entier m,

$$\forall s_1, s_2, \ldots, s_m, t_1, t_2, \ldots, t_m \in S^1\ , \forall \theta_1, \theta_2, \ldots, \theta_{m-1} \in \{\omega, -\omega\} \cup \mathcal{N},$$

$$\pi(s_m(\ldots(s_2(s_1 s t_1)^{\theta_1} t_2)^{\theta_2} \ldots)^{\theta_{m-1}} t_m) \in P$$
$$\iff \pi(s_m(\ldots(s_2(s_1 t t_1)^{\theta_1} t_2)^{\theta_2} \ldots)^{\theta_{m-1}} t_m) \in P.$$

Vérifions que $\sim_P$ est bien une relation de congruence.

Proposition 57 *Soit (S,π) un $\diamond$-semigroupe fini et soit P un sous-ensemble de S. La relation $\sim_P$ est une congruence de $\diamond$-semigroupe.*

Preuve. Soit (S,π) un $\diamond$-semigroupe et soit $P \subseteq S$. D'après le théorème 45, le produit π est défini par des fonctions compatibles notées τ et $-\tau$. On veut montrer que $\sim_P$ est stable par produit i.e. que pour tout ordre linéaire dispersé J, si $s_j \sim t_j$ pour tout $j \in J$, alors $\pi(\prod_{j\in J} s_j) \sim \pi(\prod_{j\in J} t_j)$. Comme π est associatif, il suffit de montrer que $\sim_P$ est stable par produit fini et par les opérations τ et $-\tau$. Soient s, s', t et t' appartenant à S tels que $s \sim_P s'$ et $t \sim_P t'$.

On montre que $st \sim_P s't'$. Soit m un entier, s_1, s_2, …, s_m, t_1, t_2, …, $t_m \in S^1$ et θ_1, θ_2, …, $\theta_{m-1} \in \{\omega, -\omega\} \cup \mathcal{N}$.

Supposons que

$$\pi(s_m(\ldots(s_1 s t t_1)^{\theta_1} \ldots)^{\theta_{m-1}} t_m) \in P.$$

Comme $s \sim_P s'$, on a

$$\pi(s_m(\ldots(s_1 s' t t_1)^{\theta_1} \ldots)^{\theta_{m-1}} t_m) \in P.$$

De même, on a $t \sim_P t'$ donc

$$\pi(s_m(\ldots(s_1 s' t' t_1)^{\theta_1} \ldots)^{\theta_{m-1}} t_m) \in P.$$

La réciproque est similaire et finalement on obtient $st \sim_P s't'$.

On montre que $s^\tau \sim_P {s'}^\tau$.

Supposons que

$$\pi(s_m(\ldots(s_1 \pi(s^\omega) t_1)^{\theta_1} \ldots)^{\theta_{m-1}} t_m) \in P.$$

Posons $\theta_0 = \omega$ et $s_0 = t_0 = 1$. Par associativité de π, on a alors

$$\pi(s_m(\ldots(s_1(s_0 s t_0)^{\theta_0} t_1)^{\theta_1} \ldots)^{\theta_{m-1}} t_m) \in P.$$

Comme $s \sim_P s'$,

$$\pi(s_m(\ldots(s_1(s_0 s' t_0)^{\theta_0} t_1)^{\theta_1} \ldots)^{\theta_{m-1}} t_m) \in P$$

i.e. par associativité de π

$$\pi(s_m(\ldots(s_1 \pi(s'^\omega) t_1)^{\theta_1} \ldots)^{\theta_{m-1}} t_m) \in P.$$

La réciproque est similaire et donc $s^\tau \sim_P s'^\tau$.

De façon symétrique, $s^{-\tau} \sim_P s'^{-\tau}$ donc $\sim_P$ est une congruence. $\square$

Lorsque S est un $\diamond$-semigroupe fini, le quotient $S/\sim_P$ est un $\diamond$-semigroupe effectif.

Remarque 1 *Soit (S,π) un $\diamond$-semigroupe fini de cardinal k et soit $P \subseteq S$. Quels que soient s et t appartenant à S, $s \sim_P t$ si et seulement si pour tout m inférieur ou égal à k^2,*

$$\forall s_1, s_2, \ldots, s_m, t_1, t_2, \ldots, t_m \in S^1 \ , \forall \theta_1, \theta_2, \ldots, \theta_{m-1} \in \{\omega, -\omega\} \cup \mathcal{N},$$

$$\begin{aligned} &\pi(s_m(\ldots(s_2(s_1 s t_1)^{\theta_1} t_2)^{\theta_2} \ldots)^{\theta_{m-1}} t_m) \in P \\ \iff &\pi(s_m(\ldots(s_2(s_1 t t_1)^{\theta_1} t_2)^{\theta_2} \ldots)^{\theta_{m-1}} t_m) \in P. \end{aligned} \tag{4.1}$$

Preuve. Soient $s, t \in S$ et soit $P \subseteq S$. Si $s \sim_P t$ alors l'équation 4.1 est vérifiée de façon triviale pour tout $m \leq k^2$. Réciproquement, on montre par récurrence sur $m \geq k^2$ que si l'équation 4.1 est vérifiée pour tout $m \leq k^2$ alors elle l'est aussi pour tout entier m. Soit $m > k^2$, s_1, s_2, …, s_m, t_1, t_2, …, $t_m \in S^1$ et θ_1, θ_2, …, $\theta_{m-1} \in \{\omega, -\omega\} \cup \mathcal{N}$. Pour tout $1 \leq i \leq m$, on note

$$u_i = s_i(\ldots(s_2(s_1 s t_1)^{\theta_1} t_2)^{\theta_2} \ldots)^{\theta_{i-1}} t_i \quad \text{et} \quad v_i = s_i(\ldots(s_2(s_1 t t_1)^{\theta_1} t_2)^{\theta_2} \ldots)^{\theta_{i-1}} t_i$$

Comme $m > k^2$, on dispose de deux entiers $0 \leq i_1 < i_2 \leq m$ tels que $\pi(u_{i_1}) = \pi(u_{i_2})$ et $\pi(v_{i_1}) = \pi(v_{i_2})$. Donc

$$\pi(s_m(\ldots(s_{i_2+1} u_{i_1} t_{i_2+1})^{\theta_{i_2+1}} \ldots)^{\theta_{m-1}} t_m) = \pi(u_m)$$

$$\pi(s_m(\ldots(s_{i_2+1} v_{i_1} t_{i_2+1})^{\theta_{i_2+1}} \ldots)^{\theta_{m-1}} t_m) = \pi(v_m).$$

L'hypothèse de récurrence donne alors $\pi(u_m) \in P \iff \pi(v_m) \in P$. $\square$

On va montrer qu'un ensemble X est reconnaissable si et seulement si le $\diamond$-semigroupe quotient $A^\diamond/\sim_X$ est fini. Les preuves sont similaires aux cas des ordinaux : Si $X = \psi^{-1}(P)$ où P est un sous-ensemble d'un $\diamond$-semigroupe fini S, on montre que les quotients $A^\diamond/\sim_X$ et $S^\diamond/\sim_P$ ont même cardinal.

Lemme 58 *Soit (S,π) un $\diamond$-semigroupe fini et soit $\psi : A^\diamond \longrightarrow S$ un morphisme de $\diamond$-semigroupe surjectif. Soit P un sous-ensemble de S et soit $X = \psi^{-1}(P)$. Quels que soient les mots x,y appartenant à $A^\diamond$, $x \sim_X y$ si et seulement si $\psi(x) \sim_P \psi(y)$.*

Preuve. Soient $x, y \in A^\diamond$ tels que $x \sim_X y$. Soit $m \in N$, s_1, s_2, …, s_m, t_1, t_2, …, $t_m \in S^1$ et θ_1, θ_2, …, $\theta_{m-1} \in \{\omega, -\omega\} \cup \mathcal{N}$. Supposons que

$$\pi(s_m(\ldots(s_1(s_0 \psi(x) t_0)^{\theta_0} t_1)^{\theta_1} \ldots)^{\theta_{m-1}} t_m) \in P.$$

Soient u_1, …, u_m, v_1, …, $v_m \in A^\diamond$ tels que pour tout $1 \leq i \leq m$, $\psi(u_i) = s_i$ et $\psi(v_i) = t_i$. On a

$$\psi(u_m(\ldots(u_1(u_0 x v_0)^{\theta_0} v_1)^{\theta_1} \ldots)^{\theta_{m-1}} v_m) \in P$$

Comme $x \sim_X y$, il vient

$$\psi(u_m(\ldots(u_1(u_0yv_0)^{\theta_0}v_1)^{\theta_1}\ldots)^{\theta_{m-1}}v_m) \in P$$

donc

$$\pi(s_m(\ldots(s_1(s_0\psi(y)t_0)^{\theta_0}t_1)^{\theta_1}\ldots)^{\theta_{m-1}}t_m) \in P$$

c'est-à-dire $\psi(x) \sim_P \psi(y)$. La réciproque se montre de façon similaire, ce qui conclut l'équivalence. $\square$

Si X est un sous-ensemble reconnaissable de $A^\diamond$, alors le quotient $A^\diamond/\sim_X$ est fini et reconnaît X.

Proposition 59 *Un sous-ensemble X de $A^\diamond$ est reconnaissable si et seulement si la relation $\sim_X$ est une congruence de $\diamond$-semigroupe d'indice fini.*

Preuve. Soit $X \subseteq A^\diamond$.

Supposons que X est reconnaissable. On dispose d'un $\diamond$-semigroupe fini S et d'un morphisme de $\diamond$-semigroupe $\psi : A^\diamond \longrightarrow S$ reconnaissant X. Considérons le sous-$\diamond$-semigroupe $T = \psi(A^\diamond)$. Le morphisme $\psi : A^\diamond \longrightarrow T$ est alors surjectif et reconnaît toujours X. Soit $P \subseteq T$ tel que $X = \psi^{-1}(P)$. D'après les lemmes 57 et 58, $\sim_X$ est une congruence et $A^\diamond/\sim_X$ et $S/\sim_P$ ont même cardinal. Comme $\sim_P$ possède un nombre fini de classes d'équivalence, il en est de même pour $\sim_X$.

Réciproquement, supposons que le $\diamond$-semigroupe $A^\diamond/\sim_X$ est fini. Soit $\varphi : A^\diamond \longrightarrow A^\diamond/\sim_X$ le morphisme canonique qui associe à chaque mot de $A^\diamond$ sa $\sim_X$-classe. On montre que $\varphi^{-1}(\varphi(X)) = X$. L'inclusion de droite à gauche est triviale. Soit $y \in \varphi^{-1}(\varphi(X))$. On dispose de $x \in X$ tel que $y = \varphi^{-1}(\varphi(x))$. On a alors $\varphi(x) = \varphi(y)$ i.e. $x \sim_X y$ d'où $y \in X$. Donc $A^\diamond/\sim_X$ reconnaît X. $\square$

Pour tout sous-ensemble X de $A^\diamond$, le $\diamond$-semigroupe $A^\diamond/\sim_X$ est le plus petit $\diamond$-semigroupe reconnaissant X au sens de la division.

Proposition 60 *Soit X un sous-ensemble reconnaissable de $A^\diamond$. Un $\diamond$-semigroupe fini T reconnaît X si et seulement $S(X)$ divise T.*

Preuve. Soit X un sous-ensemble reconnaissable de $A^\diamond$ et soit T un $\diamond$-semigroupe fini.

Supposons que T reconnaît X. On dispose d'un morphisme de $\diamond$-semigroupe $\psi : A^\diamond \longrightarrow T$ reconnaissant X. Soit U le sous-$\diamond$-semigroupe $\psi(A^\diamond)$ de T. Alors le morphisme $\psi : A^\diamond \longrightarrow U$ est surjectif et reconnaît toujours X : on dispose de $P \subseteq U$ tel que $X = \psi^{-1}(P)$. On montre que $S(X)$ est un quotient de U. Soit $\varphi : A^\diamond \longrightarrow S(X)$ le morphisme canonique qui associe à tout mot de $A^\diamond$ sa $\sim_X$-classe. Soit $u \in U$. Comme ψ est surjectif, on dispose de $x \in A^\diamond$ tel que $\psi(x) = u$. Supposons qu'il existe $y \in A^\diamond$ tel que $\psi(y) = u$. Alors $\psi(x) = \psi(y)$ et d'après le lemme 58, $\varphi(x) = \varphi(y)$. Donc $\varphi \circ \psi^{-1}$ est un morphisme surjectif de U dans $S(X)$ ce qui montre que $S(X) < T$.

Réciproquement, supposons que $S(X) < T$. Il existe un sous-$\diamond$-semigroupe U de T et un morphisme surjectif $\psi : U \longrightarrow S(X)$. Comme ψ est surjectif, pour tout $a \in A$, l'ensemble $\psi^{-1} \circ \varphi(a)$ est non vide. Donc, il existe une application $\phi : A \longrightarrow U$ telle que pour tout $a \in A$, $\varphi(a) = \psi \circ \phi(a)$. De plus, ϕ peut être étendue de façon unique en un morphisme de $\diamond$-semigroupe $\phi : A^\diamond \longrightarrow U$ et $\varphi = \psi \circ \phi$. Soit $P = \psi^{-1} \circ \varphi(X)$. Il vient

$$\phi^{-1}(P) = \phi^{-1}(\psi^{-1} \circ \varphi(X)) = \varphi^{-1} \circ \varphi(X) = X$$

donc U reconnaît X d'où T reconnaît X . $\square$

En particulier, la relation $\sim_X$ est la congruence la plus grossière dont le quotient $S(X)$ reconnaît X. On va pouvoir caractériser un langage rationnel de $A^\diamond$ en donnant des propriétés sur son $\diamond$-semigroupe syntaxique.

Chapitre 5

Langages sans étoile

L'approche algébrique est un outil puissant pour l'étude des ensembles rationnels. Contrairement à la théorie des automates, elle permet d'associer un objet canonique à chaque ensemble rationnel : le semigroupe syntaxique. Les chapitres précédents ont montré qu'un ensemble est rationnel si et seulement si son semigroupe syntaxique est fini. Dans le cas des mots finis, l'étude des propriétés algébriques des semigroupes syntaxiques a permis d'établir toute une hiérarchie des ensembles rationnels. Schützenberger fût le premier à considérer les ensembles dont les semigroupes syntaxiques sont finis et apériodiques, c'est-à-dire ne contenant pas de groupes. Il a montré que ces ensembles étaient exactement les langages sans étoile [38], définis par des expressions rationnelles n'utilisant que les opérations booléennes et le produit fini. Cette classe est importante d'un point de vue logique puisqu'elle contient exactement les ensembles définissables par des formules logiques du premier ordre [24]. Ces résultats ont été étendus aux mots de longueur ω par Ladner [22], Thomas [41] et Perrin [29] et même aux mots ordinaux par Bedon [5]. Etant donné un entier n, les ensembles sans étoile de mots indexés par des ordres linéaires dispersés de rang inférieur ou égal à n, sont construits à partir des lettres de l'alphabet en utilisant le produit fini et les opérations booléennes où le complément est pris dans l'ensemble des mots non vides de rang au plus n. Dans ce chapitre, on montre que le $\diamond$-semigroupe syntaxique d'un ensemble X de rang fini n est apériodique et fini si et seulement si X est sans étoile dans l'ensemble des mots de rang au plus n. Pour montrer que le semigroupe syntaxique d'un ensemble sans étoile de rang fini est apériodique, le produit de Schützenberger est généralisé aux $\diamond$-semigroupes. La preuve de la réciproque, inspirée de [4], utilise une induction sur le rang et n'est donc pas adaptable aux ordres linéaires de rang quelconque.

5.1 Semigroupes apériodiques

Définition 25 Un semigroupe S est *apériodique* s'il existe un entier k tel que pour tout s appartenant à S, $s^{k+1} = s^k$.

Un $\diamond$-semigroupe (respectivement ω-semigroupe, ω_1-semigroupe) est apériodique si son semigroupe de support est apériodique.

5.2 Langages sans étoile

Un ensemble sans étoile est défini par une expression rationnelle où l'opérateur $*$ est interdit mais dans laquelle la complémentation peut être utilisée. Nous rappelons les définitions des langages sans étoile de mots finis, infinis et ordinaux de rang fini avant de les généraliser aux mots sur les ordres linéaires dispersés de rang fini.

5.2.1 Mots finis

Définition 26 La classe $SF(A^*)$ des ensembles *sans étoile* de mots finis est la plus petite famille contenant l'ensemble $\mathcal{P}(A)$ des parties de A et fermée par produit fini et par les opérations booléennes où la complémentation est prise par rapport à A^*.

Notons que l'ensemble vide est sans étoile et son complémentaire A^* aussi.

Exemple 61 L'ensemble $(ab)^+$ est sans étoile puisqu'il est défini par l'expression $(aA^* \cap A^*b) \setminus (A^+aaA^+ \cup A^+bbA^+)$.

Théorème 61 (Schützenberger) *[38] Un ensemble de mots finis est sans étoile si et seulement si son semigroupe syntaxique est apériodique et fini.*

Exemple 62 Le semigroupe syntaxique de l'ensemble $(ab)^+$, défini à l'exemple 38 par les $\sim_X$-classes $(ab)^*a$, $(ab)^+$, $b(ab)^*$, $(ba)^+$ et $A^*aaA^* \cup A^*bbA^*$ est apériodique : tout élément s vérifie $s^2 = s^3$.

5.2.2 Mots infinis

Pour les mots de longueur ω, le produit fini n'est autorisé que du coté gauche.

Définition 27 La classe $SF(A^\omega)$ des ensembles *sans étoile* de A^ω est la plus petite famille contenant $\mathcal{P}(A)$ et fermée par les opérations booléennes où la complémentation est prise par rapport à A^ω, et fermée par produit fini à gauche des ensembles sans étoile de A^*.

De même que A^* appartient à $SF(A^*)$, l'ensemble A^ω appartient à $SF(A^\omega)$ en tant que complémentaire de l'ensemble vide.

Exemple 63 L'ensemble $(ab)^\omega$ est sans étoile : $(ab)^\omega = aA^\omega \setminus (A^*aaA^\omega \cup A^*bbA^\omega)$.

Théorème 62 (Perrin) *[29] Un ensemble de mots de longueur ω est sans étoile si et seulement si son semigroupe syntaxique est apériodique et fini.*

5.2.3 Mots sur les ordinaux de rang fini

Pour tout entier n, notons $A^{<\omega^n}$ l'ensemble des mots non vide sur A indexés par un ordinal strictement inférieur à ω^n.

Définition 28 Soit n un entier naturel. La classe $SF(A^{<\omega^{n+1}})$ des ensembles *sans étoile* de mots indexés par des ordinaux inférieur à ω^{n+1} est la plus petite famille contenant l'ensemble $\mathcal{P}(A)$ des parties de A et fermée par les opérations booléennes où la complémentation est prise par rapport à $A^{<\omega^{n+1}}$, et fermée par produit fini.

Exemple 64 L'ensemble $X = ((ab)^+ + (ab)^\omega)^+$ appartient à $SF(A^{<\omega^2})$. En effet, en notant $L = A^{<\omega^2} \setminus A^{<\omega^2}A$ l'ensemble des mots de $A^{<\omega^2}$ de longueur limite, il vient

$$X = abA^{<\omega^2} \setminus (LbA^{<\omega^2} + A^{<\omega^2}a + A^{<\omega^2}aaA^{<\omega^2} + A^{<\omega^2}bbA^{<\omega^2}).$$

Théorème 63 (Bedon) *[6] Un ensemble de mots indexés par des ordinaux de rang fini est sans étoile si et seulement si son semigroupe syntaxique est apériodique et fini.*

La preuve de ce théorème utilise une approche logique et montre également que les langages sans étoile de mots transfinis sont définis par des formules logiques du premier ordre.

5.2.4 Mots sur les ordres linéaires

Définition 29 Soit A un alphabet fini et soit n un entier naturel. La classe $SF_n(A)$ (ou SF_n) des ensembles sans étoile de rang au plus n est le plus petit ensemble contenant $\mathcal{P}(A)$ des parties de A et fermé par produit fini et par les opérations booléennes où le complément est pris par rapport à l'ensemble A^{W_n}.

Bien sûr, l'ensemble A^{W_n} appartient à SF_n puisque c'est le complément de $\emptyset \in SF_n$.

Exemple 65 Pour tout entier naturel n, les ensembles G et D de mots de rang au plus n et de longueur limite à gauche et à droite sont sans étoile :

$$G = A^{W_n} \setminus AA^{W_n} \ , \ \ D = A^{W_n} \setminus A^{W_n}A.$$

L'ensemble A^* des mots finis appartient aussi à SF_n : c'est l'ensemble des mots ne contenant aucun facteur de longueur limite :

$$\begin{aligned} A^* &= A^{W_n} \setminus A^{W_n}(G \cup D)A^{W_n} \\ &= A^{W_n} \setminus (A^{W_n}(G \cup D)A^{W_n} \cup GA^{W_n} \cup A^{W_n}D). \end{aligned}$$

Lemme 64 *Quels que soient les entiers naturels n et $k < n$, les ensembles A^{W_k}, $(A^{W_k})^\omega$ et $(A^{W_k})^{-\omega}$ appartiennent à SF_n.*

Preuve. Soit $n \in \mathbb{N}$. L'exemple 65 montre déjà que $A^* \in SF_n$. On montre que pour tout $0 \leq k < n$, si $A^{W_k} \in SF_n$, alors $(A^{W_k})^\omega, (A^{W_k})^{-\omega}$ et $A^{W_{k+1}}$ appartiennent à SF_n. Soit $0 \leq k < n$. Supposons $A^{W_k} \in SF_n$. Son complément $\overline{A^{W_k}} = A^{W_n} \setminus A^{W_k}$ appartient donc aussi à SF_n. On montre que $(A^{W_k})^\omega \in SF_n$ en utilisant l'expression sans étoile suivante :

$$(A^{W_k})^\omega = \overline{A^{W_k}} \setminus (\overline{A^{W_k}}\ \overline{A^{W_k}} \cup \overline{A^{W_k}} A^{W_k})$$

L'inclusion de gauche à droite est directe. Réciproquement, soit x appartenant au membre droit de l'égalité. Comme $x \in A^{W_n} \setminus A^{W_k}$, il existe un entier $k < r \leq n$ tel que $|x| \in W_r \setminus W_{r-1}$. Notons $|x| = \sum_{i=1}^{m} K_i$ où $K_i \in U_r$ pour tout $1 \leq i \leq m$. Comme $x \in \overline{A^{W_k}} \setminus \overline{A^{W_k} A^{W_k}}$, il existe un unique entier $1 \leq i_0 \leq m$ tel que $K_{i_0} \in U_r \setminus W_k$. De plus, $i_0 = m$ puisque $x \in \overline{A^{W_k}} \setminus \overline{A^{W_k}} A^{W_k}$. On sait donc que x admet une factorisation $x = x'x''$ où $|x'| \in W_k$ et $|x''| \in U_r \setminus W_k$. Par définition de U_r, $|x''| = \sum_{j \in J} I_j$ où pour tout $j \in J$,$I_j \in W_{r-1}$ et où $J \in \mathcal{N} \cup \{\omega, -\omega\}$. Si $J \in \mathcal{N}$ alors $|x| \in W_{r-1}$ ce qui est une contradiction. Donc $J \in \{\omega, -\omega\}$. Par l'absurde, supposons qu'il existe $j_0 \in J$ tel que $I_{j_0} \in W_{r-1} \setminus W_k$. si $J = \omega$ alors les ordres linéaires $\sum_{j \leq j_0} I_j$ et $\sum_{j_0 < j} I_j$ appartiennent tous les deux à $W_{r-1} \setminus W_k$. Cela signifierait que $x \in \overline{A^{W_k} A^{W_k}}$ ce qui est une contradiction. L'argument symétrique pour $J = -\omega$ montre que $x'' \in (A^{W_k})^\omega$ ou que $x'' \in (A^{W_k})^{-\omega}$. Comme $x \in \overline{A^{W_k}} \setminus \overline{A^{W_k}} A^{W_k}$, on conclut que $x \in (A^{W_k})^\omega$ et que $(A^{W_k})^\omega \in SF_n$. De façon symétrique on obtient $(A^{W_k})^{-\omega} \in SF_n$.

Il reste à prouver que $A^{W_{k+1}} \in SF_n$. Pour cela, on montre l'égalité suivante :

$$A^{W_n} \setminus A^{W_{k+1}} = A^{W_n}(G \cup D)A^{W_n}$$

$$\text{où } G = A^{W_n} \setminus (A^{W_k} \cup (A^{W_k})^{-\omega})A^{W_n} \text{ et } D = A^{W_n} \setminus A^{W_n}(A^{W_k} \cup (A^{W_k})^\omega)$$

L'inclusion de gauche à droite découle du fait que tout mot appartenant à G ou à D est de rang strictement supérieur à $k+1$. Réciproquement, soit $x \in A^{W_n} \setminus A^{W_{k+1}}$. Alors x possède au moins un facteur y appartenant à $(A^{W_{k+1}} \setminus A^{W_k})^\omega$ ou à $(A^{W_{k+1}} \setminus A^{W_k})^{-\omega}$. Or $(A^{W_{k+1}} \setminus A^{W_k})^\omega \subset D$ et $(A^{W_{k+1}} \setminus A^{W_k})^{-\omega} \subset G$. D'où $x \in A^{W_n}(G \cup D)A^{W_n}$, ce qui conclut la preuve.

□

Pour tout $X \in SF_n$, on dispose d'une expression sans étoile de X où le complément est pris par rapport à l'ensemble A^{W_n}. Comme $A^{W_n} \in SF_{n+1}$, on peut remplacer dans celle-ci toute expression de la forme $A^{W_n} \setminus Y$ par $(A^{W_{n+1}} \setminus Y) \cap A^{W_n}$ et obtenir une expression sans étoile de X où le complément est pris par rapport à l'ensemble $A^{W_{n+1}}$.

Corollaire 65 *Quel que soit $n \in \mathbb{N}$, $SF_n \subseteq SF_{n+1}$.*

Soient $n \in \mathbb{N}$ et $X \in SF_n$. On définit f inductivement par

$$\begin{array}{rl} \forall a \in A, & f(a) = a \\ \forall X_1, X_2 \in SF_n, & f(X_1 \cdot X_2) = f(X_1) \cdot f(X_2), \\ & f(X_1 \cup X_2) = f(X_1) \cup f(X_2), \\ & f(X_1 \cap X_2) = f(X_1) \cap f(X_2), \\ & f(X_1 \setminus X_2) = f(X_1) \cap (A^{W_{n+1}} \setminus f(L_2)) \cap A^{W_n}. \end{array}$$

d'après le lemme 65, $A^{W_n} \in SF_{n+1}$ d'où $X = f(X) \in SF_{n+1}$. □

Les paragraphes suivants sont consacrés à la preuve du théorème suivant :

Théorème 66 *Un ensemble de mots indexés par des ordres linéaires dispersés de rang fini est sans étoile si et seulement si son* $\diamond$*-semigroupe syntaxique est fini et apériodique.*

5.3 Des $\diamond$-semigroupes apériodiques vers les langages sans étoile

Soit S un $\diamond$-semigroupe fini apériodique reconnaissant un ensemble X de rang fini r. Il existe un morphisme de $\diamond$-semigroupe $\varphi : A^\diamond \longrightarrow S$ et un ensemble $P \subseteq S$ tels que $X = \varphi^{-1}(P)$. Pour tout entier $n \in \mathbb{N}$ et tout élément $p \in S$, on note $X_{n,p}$ l'ensemble des mots de rang au plus n et d'image p par φ :

$$X_{n,p} = \varphi^{-1}(p) \cap A^{W_n}.$$

Comme $X = \bigcup_{p \in P} X_{r,p}$ et que les ensembles sans étoile sont clos par union fini, il suffit de montrer que pour tout $p \in P$, $X_{r,p} \in SF_r$. On va montrer par récurrence sur n que pour tout $n \in \mathbb{N}$ et pour tout $p \in S$, $X_{n,p} \in SF_n$. Le cas de base sera donné par le théorème de Schützenberger. Pour le pas de récurrence, on va d'abord montrer que pour tout idempotent $e \in E(S)$, si $X_{n,e} \in SF_n$ alors $X^\omega_{n,e} \in SF_{n+1}$. La preuve est une adaptation de celle du Théorème 5.4 de [32].

Lemme 67 *Quels que soient* $n \in \mathbb{N}$ *et* $e \in E(S)$, $(X_{n,e})^\omega = (A^{W'_n})^\omega \setminus Z$ *où*

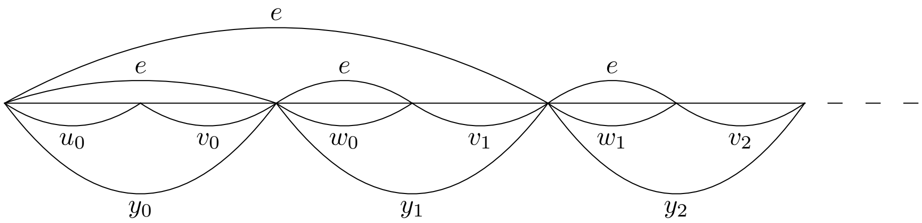
$$Z = \bigcup_{s \in S} X_{n,s}[(A^{W_n})^\omega \setminus (\bigcup_{\substack{t \in S, \\ st = e}} X_{n,t}) X_{n,e} (A^{W_n})^\omega].$$

Preuve. Soient $n \in \mathbb{N}$ et $e \in E(S)$. Soit $x \in X^\omega_{n,e}$. Par l'absurde, supposons que $x \in Z$. On dispose d'une factorisation $x = x'x''$ telle que pour toute factorisation $x'' = yy'y''$, $\varphi(x'y) \neq e$ ou $\varphi(y') \neq e$. Autrement dit la longueur du plus long préfixe de x appartenant à $X_{n,e}X_{n,e}$ est bornée. La contradiction avec $x \in X^\omega_{n,e}$ assure que $x \in (A^{W_n})^\omega \setminus Z$.

Réciproquement, soit $x \in (A^{W_n})^\omega \setminus Z$ et soit $x = u_0 x_0$ une factorisation quelconque. Notons $s = \varphi(u_0)$. Comme $x \notin Z$, il existe $t \in S$ tel que $st = e$ et $x_0 \in X_{n,t} X_{n,e} (A^{W_n})^\omega$. On dispose donc de v_0, $w_0 \in A^{W_n}$ tels que $\varphi(u_0 v_0) = e$, $\varphi(w_0) = e$ et $x_0 = v_0 w_0 x_1$ pour un certain suffixe $x_1 \in (A^{W_n})^\omega$ de x. On peut réappliquer ce raisonnement à la factorisation $x = u_1x_1$ et obtenir de façon récursive des suites $(u_i)_{i\geq 0}$, $(x_i)_{i\geq 0}$, $(v_i)_{i\geq 0}$ et $(w_i)_{i\geq 0}$ telles que pour tout $i \geq 0$,

$$x = u_ix_i \ , \quad u_{i+1} = u_iv_iw_i \ , \varphi(u_iv_i) = e \text{ et } \quad \varphi(w_i) = e$$

Soit $x = \prod_{i\in\omega} y_i$ la factorisation définie par $y_0 = u_0v_0$ et pour tout $i > 0$, $y_i = w_{i-1}v_i$.



On dispose d'une superfactorisation $x = \prod_{i\in\omega} y'_i$ et d'une paire liée à droite $(r, f) \in S \times E(S)$ tels que $\varphi(y'_0) = r$ et pour tout $i > 0$, $\varphi(y'_i) = f$. On montre que $(r, f) = (e, e)$. Comme il existe $m \in \mathbb{N}$ tel que $\varphi(y'_0) = \varphi(y_0y_1 \dots y_m) = \varphi(u_mv_m)$, on obtient $r = e$. De même, on dispose de $m_1 \leq m_2$ tels que $f = \varphi(y_{m_1} \dots y_{m_2}) = \varphi(w_{m_1-1})\varphi(v_{m_1}y_{m_1+1} \dots y_{m_2}) = eg$ où $g \in S$. D'où $f = eg = eeg = ef$. Comme (e, f) est un couple lié à droite, $e = ef$ ce qui conclut $e = f$. □

Etant donné $n \in \mathbb{N}$, on a montré que pour tout $p \in S$, si $X_{n,p} \in SF_n$, alors l'ensemble $Y_{n+1,p}$ défini par

$$Y_{n+1,p} = \varphi^{-1}(p) \cap A^{U'_{n+1}} = X_{n,p} \cup \bigcup_{\substack{(t,e)\in S\times E(S),\\ te^{\tau}=p}} X_{n,t}X^{\omega}_{n,e} \cup \bigcup_{\substack{(t,e)\in S\times E(S),\\ e^{-\tau}t=p}} X^{-\omega}_{n,e}X_{n,t}$$

appartient à SF_{n+1}. Le reste de la preuve est inspirée de [30] et utilise les deux lemmes ci-dessous.

Lemme 68 *Soit S un semigroupe apériodique. Quels que soient p, r et s appartenant à S, si $p = rps$ alors $p = rp = ps$.*

Preuve. Soit S un semigroupe apériodique et soient p, r, $s \in S$. On dispose d'un entier k tel que $r^{k+1} = r^k$. Il vient $p = rps = r^kps^k = r^{k+1}ps^k = rp$. De façon similaire, on obtient $p = ps$. □

Dans un semigroupe apériodique, si deux éléments p et s vérifient $s \leq_{\mathcal{R}} p$, $s \leq_{\mathcal{L}} p$ et $p \leq_{\mathcal{R}} s$ alors ils sont égaux.

Lemme 69 *Soit S un semigroupe apériodique. Quel que soit p appartenant à S, $p = (pS^1 \cap S^1p) \setminus \{r \in S \mid p \notin S^1rS^1\}$.*

Preuve. Soit S un semigroupe apériodique et soit $p \in S$. L'inclusion de gauche à droite est évidente. Réciproquement, soit s appartenant au membre droit de l'égalité. On dispose de a, b, c, $d \in S^1$ tels que $s = pa = bp$ et $p = csd$. Donc $s = csda$ et d'après le lemme 68, $s = cs$. On obtient $s = sd$ de façon similaire d'où $s = csd = p$. $\square$

Pour tout $p \in S$, on donne une expression de $X_{n+1,p}$ à partir des ensembles $Y_{n+1,s}$ pour $s \in S$.

Lemme 70 *Quels que soient $n \in N$ et $p \in S$,*

$$X_{n+1,p} = (RA^{W_{n+1}} \cap A^{W_{n+1}}L) \setminus A^{W_{n+1}}JA^{W_{n+1}}$$

où

$$R = \bigcup_{s\mathcal{R}p} Y_{n+1,s} \cup \bigcup_{\substack{rs\mathcal{R}p,\\ r>_{\mathcal{R}}p}} X_{n+1,r}Y_{n+1,s}$$

$$L = \bigcup_{s\mathcal{L}p} Y_{n+1,s} \cup \bigcup_{\substack{sr\mathcal{L}p,\\ r>_{\mathcal{L}}p}} Y_{n+1,s}X_{n+1,r}$$

$$J = \bigcup_{s<_{\mathcal{J}}p} Y_{n+1,s} \cup \Big(\bigcup_{\substack{rs\geq_{\mathcal{J}}p,\\ st\geq_{\mathcal{J}}p,\\ rst<_{\mathcal{J}}p}} Y_{n+1,r}X_{n+1,s}Y_{n+1,t}\Big) \cup \Big(\bigcup_{\substack{r\geq_{\mathcal{J}}p,\\ s\geq_{\mathcal{J}}p,\\ rs<_{\mathcal{J}}p}} Y_{n+1,r}Y_{n+1,s}\Big).$$

Preuve. Soient $n \in \mathbb{N}$ et $p \in S$. Soit $x \in X_{n+1,p}$. On commence par montrer que $x \in RA^{W_{n+1}}$. On dispose d'un entier m et d'une factorisation $x = \prod_{i=1}^{m} x_i$ telle que pour tout $1 \leq i \leq m$, $x_i \in Y_{n+1,p}$. Si $m = 1$, alors $x \in R$. Sinon comme $\varphi(x) = p$, il existe $1 \leq i_0 \leq m$ tel que pour tout $1 \leq i < i_0$, $\varphi(x_1 \dots x_i) >_{\mathcal{R}} p$ et pour tout $i_0 \leq i \leq m$, $\varphi(x_1 \dots x_i)\ \mathcal{R}\ p$.

- Si $i_0 = 1$ alors $\varphi(x_1)\ \mathcal{R}\ p$ et $x \in \bigcup_{s\mathcal{R}p} Y_{n+1,s}A^{W_{n+1}}$.
- Sinon $i_0 > 1$: Notons $r = \varphi(x_1 \dots x_{i_0-1})$ et $s = \varphi(x_{i_0})$. Comme $r >_{\mathcal{R}} p$, $rs\ \mathcal{R}\ p$ et $x_{i_o} \in Y_{n+1,s}$ il vient $x \in \bigcup_{\substack{rs\mathcal{R}p,\\ r>_{\mathcal{R}}p}} X_{n+1,r}Y_{n+1,s}A^{W_{n+1}}$.

D'où $x \in RA^{W_{n+1}}$. La preuve de $x \in A^{W_{n+1}}L$ peut se faire symétriquement. De plus, si $x \in A^{W_{n+1}}JA^{W_{n+1}}$ alors $\varphi(x) \neq p$ par définition de J. On obtient ainsi la première inclusion.

Réciproquement, soit $x \in (RA^{W_{n+1}} \cap A^{W_{n+1}}L) \setminus A^{W_{n+1}}JA^{W_{n+1}}$. Comme $x \in RA^{W_{n+1}}$, il vient $\varphi(x) \in pS$. De même, $x \in A^{W_{n+1}}L$ donc $\varphi(x) \in S^1p$. D'après le lemme 69, il suffit de montrer que $\varphi(x) \geq_{\mathcal{J}} p$ pour conclure que $\varphi(x) = p$. Par l'absurde, supposons que $\varphi(x) <_{\mathcal{J}} p$. Si $|x| \in U_n$, alors $x \in Y_{n+1,\varphi(x)} \subseteq J$ ce qui est une contradiction. Sinon x il existe un entier m et une factorisation $x = \prod_{i=1}^{m} x_i$ telle que pour tout $1 \leq i \leq m$, $|x_i| \in U_n$. Comme $\varphi(x) <_{\mathcal{J}} p$, il existe $1 \leq i_0 \leq j_0 \leq m$ tels que $\varphi(x_{i_0} \dots x_{j_0}) <_{\mathcal{J}} p$,$\varphi(x_{i_0+1} \dots x_{j_0}) \geq_{\mathcal{J}} p$ et $\varphi(x_{i_0} \dots x_{j_0-1}) \geq_{\mathcal{J}} p$.

– Cas 1 : $i_0 = j_0$. Alors $\varphi(x_{i_0}) <_{\mathcal{J}} p$ donc $x \in A^{W_{n+1}} Y_{n+1,\varphi(x_{i_0})} A^{W_{n+1}} \subseteq A^{W_{n+1}} J A^{W_{n+1}}$.
– Cas 2 : $i_0 + 1 = j_0$. Alors $x \in A^{W_{n+1}} Y_{n+1,\varphi(x_{i_0})} Y_{n+1,\varphi(x_{j_0})} A^{W_{n+1}}$ avec $\varphi(x_{i_0})\varphi(x_{j_0}) <_{\mathcal{J}} p$, $\varphi(x_{i_0}) \geq_{\mathcal{J}} p$ et $\varphi(x_{j_0}) \geq_{\mathcal{J}} p$.
– Cas 3 : $j_0 - i_0 \geq 2$. Soit $r = \varphi(x_{i_0})$, $s = \varphi(x_{i_0+1} \dots x_{j_0-1})$ et $t = \varphi(x_{j_0})$. On a $rs \geq_{\mathcal{J}} p$, $st \geq_{\mathcal{J}} p$ et $rst <_{\mathcal{J}} p$. De plus, $x_{i_0} \in Y_{n+1,r}$ et $x_{j_0} \in Y_{n+1,t}$ donc $x \in A^{W_{n+1}} J A^{W_{n+1}}$ ce qui est une contradiction.

□

Pour tout $p \in S$, on obtient une expression de $X_{n+1,p}$ contenant des ensembles de la forme $Y_{n+1,s}$ où $s \in S$ mais aussi de la forme $X_{n+1,s}$. Le lemme suivant montre que les ensembles $X_{n+1,s}$ apparaissant dans l'expression vérifient $s >_{\mathcal{J}} p$.

Lemme 71 *Quel que soit S $\diamond$-semigroupe fini apériodique et quels que soient p, r, s et t des éléments de S,*
i) *si $p \leq_{\mathcal{R}} r$ alors $p <_{\mathcal{J}} r$;*
ii) *si $p \leq_{\mathcal{L}} r$ alors $p <_{\mathcal{J}} r$;*
iii) *si $rs \geq_{\mathcal{J}} p$, $st \geq_{\mathcal{J}} p$ et $rst <_{J} p$ alors $s >_{\mathcal{J}} p$.*

Preuve. Soit S un $\diamond$-semigroupe fini apériodique et soient p, r, s, $t \in S$.
i) Supposons que $p <_{\mathcal{R}} r$. Alors $p \leq_{\mathcal{J}} r$ donc $p \leq_{\mathcal{J}} r$. De plus $p \mathrel{\mathcal{J}} r$ et $p \leq_{\mathcal{R}} r$ implique $p \mathrel{\mathcal{R}} r$. Donc $r >_{\mathcal{J}} p$.
ii) Similaire à **i)**.
iii) Supposons que $rs \geq_{\mathcal{J}} p$, $st \geq_{\mathcal{J}} p$ et $rst <_{\mathcal{J}} p$. Il existe $a, b, c, d \in S$ tels que $p = arsb = cstd$, donc $p \leq_{\mathcal{J}} s$. Par l'absurde, si $p \mathrel{\mathcal{J}} s$, il existe $u, v \in S$ tels que $s = upv$ donc en utilisant le lemme 68 on obtient $s = uarsbv = uars$, $p = cuarstd$ et $p \leq_{\mathcal{J}} rst$ ce qui est une contradiction. D'où $s >_{\mathcal{J}} p$. □

On obtient alors le résultat en utilisant une induction sur la $\mathcal{D}$-classe.

Proposition 72 *Quel que soit le $\diamond$-semigroupe fini apériodique S et quel que soit le morphisme de $\diamond$-semigroupe $\varphi : A^\diamond \longrightarrow S$, pour tout $n \in \mathbb{N}$ et pour tout $p \in S$, $X_{n,p} \in SF_n$.*

Preuve. Soit S un $\diamond$-semigroupe fini apériodique et soit $\varphi : A^\diamond \longrightarrow S$. On montre par récurrence sur $n \in \mathbb{N}$ que quel que soit $p \in S$, $X_{n,p} \in SF_n$. Supposons que $n = 0$ et soit $p \in S$. L'ensemble $\varphi^{-1}(p) \cap A^*$ est reconnu par le semigroupe fini apériodique support de S. Le théorème de Schützenberger donne alors $X_{0,p} \in SF_0$. Soit $0 \leq n$. Par hypothèse de récurrence, pour tout $p \in S$, $X_{n,p} \in SF_n$ donc $X_{n,p} \in SF_{n+1}$ d'après le corollaire 65. D'autre part, le lemme 64 donne $(A^{W_n})^\omega \in SF_{n+1}$ et $(A^{W_n})^{-\omega} \in SF_{n+1}$. En utilisant le lemme 67, pour tout $p \in S$, $Y_{n+1,p} \in SF_{n+1}$. On montre que pour tout $p \in S$, $X_{n+1,p} \in SF_{n+1}$ par induction sur la $\mathcal{D}$classe de p. Notons que le lemme 70 peut se récrire avec les même notations :

$$\begin{aligned} X_{n+1,p} = \quad & ((R \cup RA^{W_{n+1}}) \cap (L \cup A^{W_{n+1}} L)) \\ & \setminus (J \cup A^{W_{n+1}} J \cup JA^{W'_{n+1}} \cup A^{W_{n+1}} J A^{W_{n+1}}). \end{aligned}$$

Il suffit donc de montrer que R, L et J appartiennent à SF_{n+1} pour conclure.

- Supposons que pour tout $s \in S$, $p \geq_{\mathcal{J}} s$. D'après le lemme 71, $R = \bigcup_{s\mathcal{R}p} Y_{n+1,s}$, $L = \bigcup_{s\mathcal{L}p} Y_{n+1,s}$ et $J = \bigcup_{s<_{\mathcal{J}}p} Y_{n+1,s} \cup \bigcup_{\substack{r\geq_{\mathcal{J}}p,\\ s\geq_{\mathcal{J}}p,\\ rs<_{\mathcal{J}}p}} Y_{n+1,r}Y_{n+1,s}$ donc $X_{n+1,p} \in SF_{n+1}$.
- Soit $p \in S$. D'après l'hypothèse d'induction sur la $\mathcal{D}$-classe, pour tout $s \in S$ tel que $s >_{\mathcal{J}} p$, $X_{n+1,s} \in SF_{n+1}$. D'après les lemmes 70 et 71, on obtient $X_{n+1,p} \in SF_{n+1}$.

$\square$

On a montré que tout ensemble de rang fini n reconnu par un $\diamond$-semigroupe fini apériodique appartient à SF_n.

Proposition 73 *Tout ensemble de rang fini et reconnu par un $\diamond$-semigroupe fini apériodique est sans étoile.*

La construction est illustrée par l'exemple suivant.

Exemple 66 Soit $S = \{e, f, s, 0\}$ le $\diamond$-semigroupe dont le produit est défini par

$$ee = e \ , \ \ ff = sf = ef = f \ , \ \ ss = fs = fe = es = se = s \text{ et } \ e^\tau = f$$

tous les autres produits valant 0.

$$\boxed{e^*}$$

$$\boxed{s^*}\boxed{f^*}$$

$$\boxed{0^*}$$

Soit φ le morphisme défini pour tout $a \in A$ par $\varphi(a) = e$. Comme S ne contient que des idempotents, il est apériodique et on cherche une expression sans étoile de $X = \varphi^{-1}(f) = (A^\omega)^+$. Pour le rang 0, le seul ensemble non vide est $X_{0,e} = A^+$ d'où $Y_{1,e} = A^+$, $Y_{1,f} = A^\omega$ et $Y_{1,0} = A^{-\omega}$. Pour $X_{1,e}$ avec les notations du Lemme 70, $R = A^+$, $L = A^+$ et $J = A^{-\omega} \cup A^{-\omega}$ d'où $X_{1,e} = A^+ = (A^+A^{W_1} \cap A^{W_1}A^+) \setminus A^{W_1}(A^{-\omega} \cup A^{-\omega})A^{W_1}$. Pour $X_{1,f}$, il vient, $R = A^\omega$, $L = A^\omega$ et $J = A^{-\omega}$ d'où $X = X_{1,f} = (A^\omega A^{W_1} \cap A^{W_1}A^\omega) \setminus A^{W_1}A^{-\omega}A^{W_1}$.

5.4 Des langages sans étoile vers les $\diamond$-semigroupes apériodiques

Quel que soit l'entier naturel n, rappelons que SF_n est le plus petit ensemble contenant $\{a\}$ pour tout $a \in A$ et fermé par produit fini et par les opérations booléennes où le complément est pris par rapport à l'ensemble A^{W_n}. On prouve que l'ensemble des langages reconnus par un $\diamond$-semigroupe fini apériodique vérifie ces propriétés de fermeture. Le produit de Schützenberger de deux semigroupes S et T, noté $S \diamond T$ est l'ensemble $S \times \mathcal{P}(S^1 \times T^1) \times T$ muni du produit fini défini par :

$$(s_1,\ P_1,\ t_1) \cdot (s_2,\ P_2,\ t_2) = (s_1s_2,\ s_1P_2 \cup P_1t_2,\ t_1t_2)$$

où $s_1P_2 = \{(s_1s_2,\ t_2) \mid (s_2,\ t_2) \in P_2\}$ et $P_1t_2 = \{(s_1,\ t_1t_2) \mid (s_1,\ t_1) \in P_1\}$.
Un élément $(s,\ P,\ t)$ de $S \diamond T$ est représenté par la matrice $\begin{bmatrix} s & P \\ 0 & t \end{bmatrix}$ et le produit fini est le produit matriciel. Pour tous les $\diamond$-semigroupes finis S et T, on étend $S \diamond T$ en un $\diamond$-semigroupe en définissant les fonctions τ et $-\tau$ pour tout $(s, P, t) \in S \diamond T$ par

$$\begin{bmatrix} s & P \\ 0 & t \end{bmatrix}^\tau = \begin{bmatrix} s^{\tau_s} & \{(s^k s', t' t^{\tau_t}) \mid k \in \mathbb{N}, (s', t') \in P\} \cup \{(s^{\tau_s}, 1)\} \\ 0 & t^{\tau_t} \end{bmatrix}$$

$$\begin{bmatrix} s & P \\ 0 & t \end{bmatrix}^{-\tau} = \begin{bmatrix} s^{-\tau_s} & \{(s^{-\tau_s} s', t' t^k) \mid k \in \mathbb{N}, (s', t') \in P\} \cup \{(1, t^{-\tau_t})\} \\ 0 & t^{-\tau_t} \end{bmatrix}$$

où τ_s, $-\tau_s$, τ_t et $-\tau_t$ représentent les fonctions définissant S et T.

Lemme 74 *Soient S et T des $\diamond$-semigroupes finis. Les fonctions τ et $-\tau$ sont compatibles respectivement à droite et à gauche avec $S \diamond T$.*

Preuve. On montre que τ est compatible à droite avec $S \diamond T$. La preuve concernant $-\tau$ peut se faire de façon symétrique. Soit $r \in S \diamond T$ et soit $n \in \mathbb{N}$. On montre que $(r^n)^\tau = r^\tau$. Notons $r = (s, P, t)$.

$$(r^n)^\tau \quad = \quad \begin{bmatrix} s^n & \bigcup_{i=0}^{n-1} s^i P t^{n-1-i} \\ 0 & t^n \end{bmatrix}^\tau \quad = \quad \begin{bmatrix} s^{\tau_s} & E \cup \{(s^{\tau_s}, 1)\} \\ 0 & t^{\tau_t} \end{bmatrix}$$

avec

$$\begin{aligned} E = \quad & \{(s^{nk} s', t'(t^n)^{\tau_t}) \mid k \in N, (s', t') \in \bigcup_{i=0}^{n-1} s^i P t^{n-1-i}\} \\ = \quad & \bigcup_{i=0}^{n-1} \{(s^{nk+i} s', t' t^{n-1-i} t^{\tau_t}) \mid k \in N, (s', t') \in P\} \\ = \quad & \bigcup_{i=0}^{n-1} \{(s^{nk+i} s', t' t^{\tau_t}) \mid k \in N, (s', t') \in P\} \\ = \quad & \{(s^k s', t' t^{\tau_t}) \mid k \in N, (s', t') \in P\} \end{aligned}$$

On montre à présent que quels que soient $r_1, r_2 \in S \diamond T$, $r_1(r_2r_1)^\tau = (r_1r_2)^\tau$.
Soient $r_1, r_2 \in S \diamond T$ notés $r_1 = (s_1, P_1, t_1)$ et $r_2 = (s_2, P_2, t_2)$.

$$\begin{aligned} r_1(r_2r_1)^\tau \quad & = \begin{bmatrix} s_1 & P_1 \\ 0 & t_1 \end{bmatrix} \begin{bmatrix} (s_2s_1)^{\tau_s} & \{((s_2s_1)^k s', t'(t_2t_1)^{\tau_t}) \mid (s', t') \in s_2P_1 \cup P_2t_1\} \cup \{((s_2s_1)^{\tau_s}, 1)\} \\ 0 & (t_2t_1)^{\tau_t} \end{bmatrix} \\ & = \begin{bmatrix} (s_1s_2)^{\tau_s} & E \cup \{((s_1s_2)^{\tau_s}, 1)\} \\ 0 & (t_1t_2)^{\tau_t} \end{bmatrix} \end{aligned}$$

avec

$$\begin{array}{rlcl}
E = & s_1\{((s_2s_1)^k s', t'(t_2t_1)^{\tau_t}) \mid (s',t') \in s_2P_1 \cup P_2t_1\} & \cup & P_1(t_2t_1)^{\tau_t} \\
= & \{(s_1(s_2s_1)^k s_2 s', t'(t_2t_1)\tau_t) \mid (s',t') \in P_1\} & \cup & \{(s', t'(t_2t_1)^{\tau_t}) \mid (s',t') \in P_1\} \\
\cup & \{(s_1(s_2s_1)^k s', t't_1(t_2t_1)^{\tau_t}) \mid (s',t') \in P_2\} & & \\
= & \{((s_1s_2)^{k+1} s', t'(t_1t_2)^{\tau_t}) \mid k \in N, (s',t') \in P_1t_2\} & \cup & \{(s', t'(t_1t_2)^{\tau_t}) \mid (s',t') \in P_1t_2\} \\
\cup & \{((s_1s_2)^k s', t'(t_1t_2)^{\tau_t}) \mid (s',t') \in s_1P_2\} & & \\
= & \{((s_1s_2)^k s', t'(t_1t_2)^{\tau_t}) \mid k \in N, (s',t') \in P_1t_2 \cup P_1t_2\} & &
\end{array}$$

d'où $r_1(r_2r_1)^\tau = (r_1r_2)^\tau$ ce qui conclut que τ est compatible à droite avec $S \diamond T$.
$\square$

D'après la proposition 45, le $\diamond$-semigroupe $S \diamond T$ est donc bien défini. De plus le produit de Schützenberger de deux $\diamond$-semigroupes apériodiques est apériodique.

Proposition 75 *L'ensemble des $\diamond$-semigroupes finis apériodiques est fermé par produit de Schützenberger.*

Preuve. Soient S et T deux $\diamond$-semigroupes finis apériodiques. On dispose de deux entiers k_s et k_t tels que quel que soit $s \in S$, $s^{k_s} = s^{k_s+1}$ et quel que soit $t \in T$, $t^{k_t} = t^{k_t+1}$. On montre que pour tout $r \in S \diamond T$, $r^{k_s+k_t+1} = r^{k_s+k_t+2}$.

Soit $(s, P, t) \in S \diamond T$. Notons que quel que soit $k \in \mathbb{N}$,

$$\begin{bmatrix} s & P \\ 0 & t \end{bmatrix}^k = \begin{bmatrix} s^k & \bigcup_{i=0}^{k-1} s^i P t^{k-1-i} \\ 0 & t^k \end{bmatrix}$$

Puisque $s^{k_s+k_t+1} = s^{k_s+k_t+2}$ et $t^{k_s+k_t+1} = t^{k_s+k_t+2}$, il suffit de montrer l'égalité suivante :

$$\begin{array}{rclcl}
\bigcup_{i=0}^{k_s+k_t+1} s^i P t^{k_s+k_t+1-i} & = & \bigcup_{i=0}^{k_s} s^i P t^{k_s+k_t+1-i} & \cup & \bigcup_{i=k_s+1}^{k_s+k_t+1} s^i P t^{k_s+k_t+1-i} \\
& = & \bigcup_{i=0}^{k_s} s^i P t^{k_s+k_t-i} & \cup & \bigcup_{i=k_s+1}^{k_s+k_t+1} s^{i-1} P t^{k_s+k_t-(i-1)} \\
& = & \bigcup_{i=0}^{k_s} s^i P t^{k_s+k_t-i} & \cup & \bigcup_{i=k_s}^{k_s+k_t} s^i P t^{k_s+k_t-i} \\
& = & \bigcup_{i=0}^{k_s+k_t} s^i P t^{k_s+k_t-i} & &
\end{array}$$

$\square$

Soient $\varphi : A^\diamond \longrightarrow S$ et $\psi : A^\diamond \longrightarrow T$ deux morphismes de $\diamond$-semigroupe. L'application $\varphi \diamond \psi$ définie pour tout $a \in A$ par

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

peut être étendue de façon unique en un morphisme de $\diamond$-semigroupe vérifiant la propriété ci-dessous.

Lemme 76 *Pour tout entier n et tout mot u de rang au plus n,*

$$\varphi \diamond \psi(u) = \begin{bmatrix} \varphi(u) & \{(\varphi(v), \psi(w)) \mid u = vw\} \\ 0 & \psi(u) \end{bmatrix}$$

Preuve. Par récurrence sur $n \in \mathbb{N}$. Soit $u \in A^{W_0}$. Par définition de $\varphi \diamond \psi$, la propriété est satisfaite pour toute lettre $a \in A$. De plus, la propriété est stable par produit fini : quels que soient u et u' satisfaisant le lemme,

$$\begin{aligned}
\varphi \diamond \psi(uu') &= \varphi \diamond \psi(u)\ \varphi \diamond \psi(u') \\
&= \begin{bmatrix} \varphi(u) & \{(\varphi(v), \psi(w)) \mid u = vw\} \\ 0 & \psi(u) \end{bmatrix} \begin{bmatrix} \varphi(u') & \{(\varphi(v), \psi(w)) \mid u' = vw\} \\ 0 & \psi(u') \end{bmatrix} \\
&= \begin{bmatrix} \varphi(uu') & \{(\varphi(uv), \psi(w)) \mid u' = vw\} \cup \{(\varphi(v), \psi(wu')) \mid u = vw\} \\ 0 & \psi(uu') \end{bmatrix} \\
&= \begin{bmatrix} \varphi(uu') & \{(\varphi(v), \psi(w)) \mid uu' = vw\} \\ 0 & \psi(uu') \end{bmatrix}
\end{aligned}$$

Soient $n \in \mathbb{N}$ et $u \in A^{W_{n+1}}$. Par hypothèse de récurrence et par ce qui précède, il suffit de traiter le cas où $|u| \in (W_n)^\omega$ ou $|u| \in (W_n)^{-\omega}$. Comme ces deux cas sont symétriques, on suppose que $|u| \in (W_n)^\omega$. On dispose d'une factorisation $u = \prod_{i\in\omega} u_i$ et d'une paire liée à droite $(r, e) \in S\diamond T \times E(S\diamond T)$ tels que $\varphi\diamond\psi(u_0) = r$ et pour tout $i > 0$, $\varphi\diamond\psi(u_i) = e$. On montre le résultat pour $u' = \prod_{i>0} u_i$. Notons $e = (s, P, t)$. Comme $\varphi(u') = s^{\tau_s}$ et $\psi(u') = t^{\tau_t}$, il suffit de prouver que

$$\{(s^{\tau_s}, 1)\} \cup \{(s^k s', t' t^{\tau_t}) \mid k \in N, (s', t') \in P\} = \{(\varphi(v), \psi(w)) \mid vw = u'\}.$$

Soient v, $w \in A^{W_{n+1}}$ tels que $u' = vw$. Si $w = \epsilon$ alors $(\varphi(v), \psi(w)) = (s^{\tau_s}, 1)$. Sinon $w \neq \epsilon$ et il existe $k > 0$ et v', $w \in A^{W'_{n+1}}$ tels que $v = u_1 \ldots u_k v'$, $w = w' u_{k+2} \ldots$ et $v'w' = u_{k+1}$. Comme $\varphi \diamond \psi(u_{k+1}) = e$ et $|u_{k+1}| \in W_n$, l'hypothèse d'induction donne $(\varphi(v'), \psi(w')) \in P$ et donc $(\varphi(v), \psi(w)) = (\varphi(u_1 \ldots u_k)\varphi(v'),$ $\psi(w')\psi(u_{k+2} \ldots)) = (s^k s', t' t^{\tau_t})$ avec $(s', t') \in P$ ce qui conclut l'inclusion de la droite vers la gauche.

Réciproquement, soit $k \in \mathbb{N}$ et soit $(s', t') \in P$. D'après l'hypothèse d'induction, il existe $v', w' \in A^{W_n}$ tels que $(s', t') = (\varphi(v'), \psi(w'))$ et $u_{k+1} = v'w'$. Donc, $(s^k s', t' t^{\tau_t}) = (\varphi(u_1 \ldots u_k v'), \psi(w' u_{k+2} \ldots))$. De plus, $(s^{\tau_s}, 1) = (\varphi(u'), \psi(\epsilon))$ ce qui conclut la preuve. □

On déduit de ce lemme 76 que la classe des ensembles reconnus par des $\diamond$-semigroupes finis apériodiques est fermée par produit fini et opérations booléennes.

Théorème 77 *Un ensemble de rang fini est sans étoile si et seulement si il est reconnu par un $\diamond$-semigroupe fini et apériodique.*

Preuve. Soit $n \in \mathbb{N}$ et soient $X \subseteq A^{W_n}$ et $Y \subseteq A^{W_n}$. Supposons que X et Y sont reconnus par des $\diamond$-semigroupes finis apériodiques S et T. On dispose de φ :

$A^\diamond \longrightarrow S$, $\psi : A^\diamond \longrightarrow T$, $P_S \subseteq S$ et $P_T \subseteq T$ tels que $X = \varphi^{-1}(P_S)$ et $Y = \psi^{-1}(P_T)$. D'après le lemme 76, $X \cup Y = \varphi \diamond \psi^{-1}(\{(s,P,t) \in S \diamond T \mid s \in P_S \text{ ou } t \in P_T\})$, $X \cap Y = \varphi \diamond \psi^{-1}(\{(s,P,t) \in S \diamond T \mid s \in P_S \text{ et } t \in P_T\})$, $XY = \varphi \diamond \psi^{-1}(\{(s,P,t) \in S \diamond T \mid P_S \times P_T \cap P \neq \emptyset\})$ et $A^\diamond \setminus X = \varphi \diamond \psi^{-1}(\{(s,P,t) \in S \diamond T \mid s \notin P_S\})$. De plus, A^{W_n} est reconnu par le $\diamond$-semigroupe fini apériodique $S_n = \{e_0, e_1, \ldots, e_{n+1}\}$ ne contenant que des idempotents tel que pour tout $0 \leq i \leq n-1$, $e_i^\tau = e_i^{-\tau} = e_{i+1}$ et $e_{n+1}^\tau = e_{n+1}^{-\tau} = e_{n+1}$ et pour tout $0 \leq j \leq n$, $e_i e_j = e_j e_i = e_{max(i,j)}$. Donc $A^{W_n} \setminus X = A^{W_n} \cap A^\diamond \setminus X$ est reconnu par $S_n \diamond S$. D'après le lemme 75, la classe des ensembles reconnus par des $\diamond$-semigroupes finis apériodiques est fermée par produit fini, opérations booléennes et complémentation par rapport à A^{W_n}. Comme elle contient $\{a\}$ pour tout $a \in A$, elle contient SF_n par définition. La réciproque est donnée par la proposition 73. □

On revient alors à la preuve du théorème 66 : un ensemble de rang fini est sans étoile si et seulement si son $\diamond$-semigroupe syntaxique est fini et apériodique.

Preuve : Soit $n \in \mathbb{N}$ et soit $X \subseteq A^{W_n}$. Supposons que $S(X)$ est fini et apériodique. D'après la proposition 60, $S(X)$ reconnaît X donc X est sans étoile. Réciproquement, supposons que $X \in SF_n$ et montrons que $S(X)$ est apériodique. On dispose d'un $\diamond$-semigroupe fini apériodique S reconnaissant X. D'après la proposition 60, $S(X)$ divise S. On dispose d'un sous-$\diamond$-semigroupe U de S et d'un morphisme surjectif $\psi : U \longrightarrow S(X)$. Soit $s \in S(X)$. On dispose de $u \in U$ tel que $\psi(u) = s$. Comme S est apériodique, U l'est aussi. Il existe donc un entier m tel que $u^m = u^{m+1}$ et l'égalité suivante est vérifiée :

$$s^{m+1} = \psi(u)^{m+1} = \psi(u^{m+1}) = \psi(u^m) = \psi(u)^m = s^m.$$

□

Conclusion et perspectives

Dans le chapitre 4, nous avons montré qu'un ensemble de mots indexés par des ordres linéaires dénombrables et dispersés est rationnel si et seulement si son $\diamond$-semigroupe syntaxique est fini. Dans le chapitre 5, nous avons partiellement étendu le théorème de Schützenberger aux ensembles de rang fini. Une question naturelle est bien sûr de savoir si ce résultat s'étend à tous les ordres dispersés : un sous-ensemble de $A^\diamond$ serait sans étoile si et seulement si son $\diamond$-semigroupe syntaxique est fini et apériodique.

Dans le cas des mots finis, l'approche algébrique mène à une classification des ensembles reconnaissables en fonction des propriétés de leur semigroupe syntaxique. Les classes des ensembles rationnels et des ensembles sans étoile ne sont pas les seules étudiées. Simon a par exemple défini les ensembles testables par morceaux comme les ensembles dont le semigroupe syntaxique est fini et $\mathcal{J}$-trivial. Une structure naturelle pour formuler ce type de résultat est la théorie des variétés, développée par Eilenberg. Cette théorie consiste à étudier des classes d'algèbres fermées par sous-algèbre, quotient et produit direct fini. Le théorème d'Eilenberg montre une correspondance entre ces variétés et les familles de langages rationnels. Un premier pas serait de généraliser ce résultat aux ensembles de mots sur les ordres linéaires dispersés.

D'autre part, la preuve de fermeture par complémentation de la classe des ensembles rationnels de mots indexés par les ordres linéaires dispersés permet d'aborder les problèmes de logique sur ces structures. Dans le cas des mots finis, il est bien connu que les automates ont le même pouvoir d'expression que la logique monadique du second ordre. Un prolongement naturel de nos travaux serait d'étendre ce résultat aux automates sur les ordres linéaires. Un autre objectif est de généraliser l'équivalence entre langages sans étoile et logique du premier ordre.

Bibliographie

[1] André Arnold. A syntactic congruence for rational ω-languages. *Theoret. Comput. Sci.*, 39 :333–335, 1985.

[2] Nicolas Bedon. Finite automata and ordinals. *Theoret. Comput. Sci.*, 156 :119–144, 1996.

[3] Nicolas Bedon. Automata, semigroups and recognizability of words on ordinals. *Int. J. Alg. Comput.*, 8(1) :1–21, 1998.

[4] Nicolas Bedon. *Langages reconnaissables de mots indéxés par des ordinaux.* Thèse, Université de Marne-la-Vallée, 1998.

[5] Nicolas Bedon. Logic over words on denumerable ordinals. *Computer and System Science*, 63 :394–431, 2001.

[6] Nicolas Bedon. Star-free sets of words on ordinals. *Inform. Comput.*, 166 :93–111, 2001.

[7] Nicolas Bedon and Olivier Carton. An Eilenberg theorem for words on countable ordinals. In Cláudio L. Lucchesi and Arnaldo V. Moura, editors, *Latin'98 : Theoretical Informatics*, volume 1380 of *Lect. Notes in Comput. Sci.*, pages 53–64. Springer-Verlag, 1998.

[8] Véronique Bruyère and Olivier Carton. Automata on linear orderings. In Jiří Sgall, Aleš Pultr, and Petr Kolman, editors, *MFCS'2001*, volume 2136 of *Lect. Notes in Comput. Sci.*, pages 236–247, 2001.

[9] Véronique Bruyère and Olivier Carton. Hierarchy among automata on linear orderings. In Ricardo Baeza-Yates, Ugo Montanari, and Nicola Santoro, editors, *Foudation of Information technology in the era of network and mobile computing*, pages 107–118. Kluwer Academic Publishers, 2002. TCS'2002/IFIP'2002.

[10] Véronique Bruyère, Olivier Carton, and Géraud Sénizergues. Tree automata and automata on linear orderings. In Tero Harju and Juhani Karhumäki, editors, *WORDS'2003*, pages 222–231. Turku Center for Computer Science, 2003.

[11] J. Richard Büchi. Weak second-order arithmetic and finite automata. *Z. Math. Logik und grundl. Math.*, 6 :66–92, 1960.

[12] J. Richard Büchi. On a decision method in the restricted second-order arithmetic. In *Proc. Int. Congress Logic, Methodology and Philosophy of Science, Berkeley 1960*, pages 1–11. Stanford University Press, 1962.

[13] J. Richard Büchi. Transfinite automata recursions and weak second order theory of ordinals. In *Proc. Int. Congress Logic, Methodology, and Philosophy of Science, Jerusalem 1964*, pages 2–23. North Holland, 1964.

[14] J. Richard Büchi and Dirk Siefkes. *The monadic second order theory of all countable ordinals*, volume 328 of *Lecture notes in mathematics*. Springer, 1973.

[15] Olivier Carton and Max Michel. Unambiguous Büchi automata. *Theoret. Comput. Sci.*, 297 :37–81, 2003. IGM report 2001–06.

[16] Olivier Carton and Chloé Rispal. Complementation of rational sets on scattered linear orderings of finite rank. In *LATIN'2004*, volume 2976 of *Lect. Notes in Comput. Sci.*, pages 292–301, 2004.

[17] Olivier Carton and Chloé Rispal. Complementation of rational sets on scattered linear orderings of finite rank. *Theoret. Comput. Sci.*, 2004. To appear in special issue of LATIN'O4.

[18] Yaacov Choueka. Finite automata, definable sets, and regular expressions over ω^n-tapes. *J. Comput. System Sci.*, 17(1) :81–97, 1978.

[19] Danièle Girault-Beauquier. Bilimites de langages reconnaissables. *Theoret. Comput. Sci.*, 33(2–3) :335–342, 1984.

[20] Felix Hausdorff. *Set Theory*. Chelsea, New York, 1957.

[21] Stephen C. Kleene. Representation of events in nerve nets and finite automata. In C.E. Shannon, editor, *Automata studies*, pages 3–41. Princeton university Press, Princeton, 1956.

[22] Richard E. Ladner. Application of model theoretic games to discrete linear orders and finite automata. *Information and Control*, 33 :281–303, 1977.

[23] Robert McNaughton. Testing and generating infinite sequences by a finite automaton. *Inform. Control*, 9 :521–530, 1966.

[24] Robert McNaughton and Seymour Papert. *Counter free automata*. MIT Press, Cambridge, MA, 1971.

[25] David Muller. Infinite sequences and finite machines. In Proc. of Fourth Annual IEEE Symp., editor, *Switching Theory and Logical Design*, pages 3–16, 1963.

[26] Maurice Nivat and Dominique Perrin. Ensembles reconnaissables de mots bi-infinis. In *Proceedings of the Fourteenth Annual ACM Symposium on Theory of Computing*, pages 47–59, 1982.

[27] Jean-Pierre Pécuchet. Etude syntaxique des parties reconnaissables de mots infinis. *Lect. Notes in Comput. Sci.*, 226 :294–303, 1986.

[28] Jean-Pierre Pécuchet. Variétés de semigroupes et mots infinis. *Lect. Notes in Comput. Sci.*, 210 :180–191, 1986.

[29] Dominique Perrin. Recent results on automata and infinite words. In M. P. Chytil and V. Koubek, editors, *Mathematical foundations of computer science*, volume 176 of *Lect. Notes in Comput. Sci.*, pages 134–148, Berlin, 1984. Springer-Verlag.

[30] Dominique Perrin. Finite automata. In J. van Leeuwen, editor, *Handbook of Theoretical Computer Science*, volume B, chapter 1, pages 1–57. Elsevier, 1990.

[31] Dominique Perrin and Jean-Éric Pin. Semigroups and automata on infinite words. In J. Fountain and V. A. R. Gould, editors, *NATO Advanced Study Institute Semigroups, Formal Languages and Groups*, pages 49–72. Kluwer Academic Publishers, 1995.

[32] Dominique Perrin and Jean-Éric Pin. *Infinite Words : Automata, Semigroups, Logic and Games*, volume 141 of *Pure and Applied Mathematics*. Elsevier, 2004.

[33] Jean-Éric Pin. *Varieties of Formal Languages*. North Oxford, London and Plenum, New-York, 1986. (Traduction de *Variétés de langages formels*).

[34] Micheal Ozer Rabin. Decidability of second-order theories and automata on infinite trees. *Trans. Amer. Math. Soc.*, 141 :1–35, 1969.

[35] F. P. Ramsey. On a problem on formal logic. *Proc. London Math. Soc.*, 30(2) :264–286, 1929.

[36] Chloé Rispal and Olivier Carton. Complementation of rational sets on countable scattered linear orderings. In C. S. Calude, E. Calude, and M. J. Dinneen, editors, *DLT'2004*, volume 3340 of *Lect. Notes in Comput. Sci.*, pages 381–392, 2004.

[37] Joseph G. Rosenstein. *Linear Orderings*. Academic Press, 1982.

[38] Marcel-Paul Schützenberger. On finite monoids having only trivial subgroups. *Inform. Control*, 8 :190–194, 1965.

[39] Waclaw Sierpiński. *Leçons sur les nombres transfinis*. Gauthier-Villars, Paris, 1950.

[40] Waclaw Sierpiński. *Cardinal and ordinal numbers*. Polish Scientific Publishers, Varsovie, 1965.

[41] Wolfgang Thomas. Star-free regular sets of ω-sequences. *Inform. Control*, 42 :148–156, 1979.

[42] Thomas Wilke. An Eilenberg theorem for ∞-languages. In *ICALP '91*, volume 510 of *Lect. Notes in Comput. Sci.*, pages 588–599, Berlin, 1991. Springer-Verlag.

[43] Thomas Wilke. An algebraic theory for regular languages of finite and infinite words. *Int. J. Alg. Comput.*, 3(4) :447–489, 1993.

[44] Jerzy Wojciechowski. The ordinals less than ω^ω are definable by finite automata. In J. Demetrovics, G. Katona, and A. Salomaa, editors, *Proc. Colloquium on "Algebra, Combinatorics and Logic in Computer Science"*, volume 2, pages 871–887, Amsterdam, 1983. North Holland.

[45] Jerzy Wojciechowski. Finite automata on transfinite sequences and regular expressions. *Fundamenta informaticæ*, 8(3-4) :379–396, 1985.