Table des matières

— ◦ —

Chapitre 2

Automate universel

En 1971, J.H. Conway [19] a défini la *matrice des facteurs* d'un langage rationnel. Son but était de donner un moyen de décider si un langage rationnel $\mathcal{L}$ appartenait à la clôture rationnelle d'une famille finie de langages rationnels $(\mathcal{L}_i)_{i\in I}$. Depuis, cette construction a été plusieurs fois retrouvée. O. Carton [12] a montré que cette construction pouvait être faite dans n'importe quel monoïde. J. Sakarovitch [55] a mis en lumière une autre propriété de cette matrice des facteurs (facilement transformable en automate) : il existe un morphisme de tout automate qui reconnaît le langage dans cet « automate des facteurs ». Cette propriété incite à donner à cet objet le nom d'automate universel que nous employons ici. Elle avait été exploitée dans [4] pour le problème de l'obtention d'un automate minimal non-deterministe. On trouvera dans [55] une étude complète des propriétés de l'automate universel.

Toutefois, par souci de cohérence et afin de fixer certaines conventions, nous reprenons ces résultats dans la première partie de ce chapitre.

Le paragraphe 1 présente des définitions et propriétés qui figurent déjà, sauf mention explicite, ailleurs ([19, 4, 12, 55]). Elles sont évidemment nécessaires à la compréhension des paragraphes suivants. Les résultats principaux en sont que, pour tout langage reconnaissable, il existe un automate fini qui reconnaît le langage et dont les états sont caractérisés par le couple formé de leur passé et de leur futur qui représente une *factorisation* maximale du langage. Cet automate doit son qualificatif d'automate universel au fait que tout automate émondé qui reconnaît le langage s'y envoie par morphisme.

Le paragraphe 2 présente une construction de l'automate universel à partir de l'automate minimal, en évitant le calcul du monoïde syntaxique, ce qui rend la complexité de la construction dépendante uniquement de la taille du résultat.

La paragraphe suivant présente l'*écorché* de l'automate universel. Il s'agit d'une représentation qui permet de faire apparaître les propriétés de l'automate universel en codant à l'aide de transitions spontanées un certain nombre de transitions « redondantes » de l'automate universel. D'un point de vue pratique, ceci permet de diminuer le nombre de transitions à stocker, mais aussi de pouvoir dessiner des automates universels relativement gros.

Le dernier paragraphe montre comment établir un lien entre l'automate universel d'un langage et un automate donné qui reconnaît ce même langage, ce qui permettra ensuite de retrouver dans le premier des propriétés héritées du second. Ceci est en fait la clé des raisonnements que nous allons effectuer par la suite. L'automate universel peut en effet être construit, d'une part, à partir d'un automate particulier (qu'on ne sait *a priori* pas construire) duquel il hérite certaines propriétés et, d'autre part, à partir de l'automate minimal reconnaissant le même langage, ce qui permet une construction effective.

— ◦ —

1 Définitions et propriétés de l'automate universel

1.1 Automate universel dans un monoïde quelconque

Factorisations d'un ensemble. La définition de l'automate universel, basée sur la notion de facteurs, est possible pour n'importe quel sous-ensemble $\mathcal{L}$ de n'importe quel monoïde N.

Naturellement, le nombre d'états de l'automate universel d'un ensemble quelconque n'est pas nécessairement fini, ni même dénombrable.

Définition 2.1 *Une* ***sous-factorisation*** *d'un sous-ensemble $\mathcal{L}$ de N est un k-uplet de sous-ensembles non vides de N, $(R_1,R_2,\ldots,R_k)$ appartenant à $(\mathcal{P}(N) \setminus \{\emptyset\})^k$, tel que $R_1.R_2\ldots R_k$ est inclus dans $\mathcal{L}$. Les sous-factorisations sont partiellement ordonnées par inclusion. Une* ***factorisation*** *de $\mathcal{L}$ est une sous-factorisation maximale. Si une factorisation de $\mathcal{L}$ est un couple (L,R), on dit que L est un* ***facteur gauche*** *et R un* ***facteur droit*** *de $\mathcal{L}$.*

Remarque 2.1 Dans ce qui suit, on considérera les factorisations qui sont des couples, sauf mention explicite. Puisqu'elle est maximale, une factorisation (L,R) de $\mathcal{L}$ est entièrement déterminée par la donnée de L (ou de R) :

$$L = \max\{X \subseteq N \mid X.R \subseteq \mathcal{L}\}\,, \quad R = \max\{X \subseteq N \mid L.X \subseteq \mathcal{L}\}.$$

Exemple 1.7 Les factorisations du langage $\mathcal{L}_1 = A^*abA^*$ sont

$$(A^*,A^*abA^*), \quad (A^*aA^*,A^*bA^*) \quad \text{et } (A^*abA^*,A^*).$$

En effet, pour toute sous-factorisation (L,R) de $\mathcal{L}_1$, s'il existe un mot de L ne contenant pas de a, tout mot de R contient un facteur ab et s'il existe un mot de R ne contenant pas de b, tout mot de L contient un facteur ab. Il n'est pas difficile de voir que les trois factorisations données sont bien maximales.

Exemple 7.2 Soit $N_3 = (\mathbb{Z}/3\mathbb{Z},+)$ et $\mathcal{L}_3 = \{1,2\}$. La figure 1 représente la table du monoïde; les éléments de $\mathcal{L}_3$ sont encadrés, les rectangles indiquent les factorisations. Cette façon « géométrique » a été étudiée dans [20].

+	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

FIG. 1 – *La table du groupe $\mathbb{Z}/3\mathbb{Z}$ et les factorisations de $\mathcal{L}_3$.*

Les factorisations sont donc les suivantes :

$$(\{0\},\{1,2\}),\quad (\{1\},\{0,1\}),\quad (\{2\},\{0,2\}),$$
$$(\{1,2\},\{0\}),\quad (\{0,1\},\{1\}),\quad (\{0,2\},\{2\}).$$

EXEMPLE 8.1 Soit a^* le monoïde librement engendré par a et $\mathcal{L}_2 = a^+$ l'ensemble des mots de A^* qui contiennent au moins une lettre. Les factorisations de $\mathcal{L}_2$ sont (a^*,a^+) et (a^+,a^*).

EXEMPLE 9 Soit N_5 le monoïde des fonctions de $\mathbb{R}$ dans $\mathbb{R}$ munies de la composition et $\mathcal{L}_5$ le singleton contenant la fonction identité. Les factorisations de $\mathcal{L}_5$ sont les couples maximaux (L,R) tels que

– L est un ensemble non vide d'applications injectives telles que quel que soit α et β appartenant à L, quel que soit (x,y) dans $\mathbb{R}^2$, $x\alpha = y\beta$ entraîne $x = y$.
– R est l'ensemble des fonctions réalisant l'inverse de α sur $\mathsf{Im}\,\alpha$, pour tout α dans L.

Le couple suivant est un exemple de factorisation de l'identité :

$$(\{\mathsf{arctg},\mathsf{arctg}+\pi\},\{\alpha \mid X \subseteq \mathsf{Supp}(\alpha) \subseteq \mathbb{R}, \forall x \in X, x\alpha = \mathsf{tg}(x)\}).$$

Le nombre de factorisations n'est ici pas dénombrable. En effet, chaque bijection α de $\mathbb{R}$ sur lui-même induit une factorisation (maximale) de l'identité : $(\{\alpha\},\{\alpha^{-1}\})$. On obtient donc ainsi un sous-ensemble non dénombrable des factorisations de l'identité.

— ◦ —

Définition de l'automate universel. Nous allons construire l'automate dont les états sont les factorisations de l'ensemble à reconnaître, qui a un maximum de transitions, et qui respecte les propriétés suivantes :

– Un état (L,R) est initial (*resp.* final) si et seulement si 1_N appartient à L (*resp.* à R).
– Un élément x de N étiquette une transition entre deux états (L,R) et (L',R') si et seulement si $L.x.R'$ est inclus dans $\mathcal{L}$.

Définition 2.2 *Soit $\mathcal{L}$ un sous-ensemble de N et Q l'ensemble de ses factorisations. L'**automate universel de** $\mathcal{L}$ est $\mathcal{U}_{\mathcal{L}} = \langle Q,N,E,I,T \rangle$, avec:*

$$\begin{aligned} I &= \{(L,R) \in Q \mid 1_N \in L\}, \\ T &= \{(L,R) \in Q \mid 1_N \in R\}, \\ E &= \{((L,R),X,(L',R')) \in Q \times \mathcal{P}(N) \times Q \mid X = \{x \in N \mid L.x.R' \subseteq \mathcal{L}\}\}. \end{aligned}$$

Remarque 2.2 Dans la définition précédente, la condition $L.x.R' \subseteq \mathcal{L}$ est équivalente, à cause de la maximalité des factorisations à l'une des conditions $L.x \subseteq L'$ ou $x.R' \subseteq R$. De même, la condition $1_N \in L$ est équivalente à $R \subseteq \mathcal{L}$ et $1_N \in R$ est équivalent à $L \subseteq \mathcal{L}$.

Exemple 1.8 L'automate universel de $\mathcal{L}_1$ a trois états (les trois factorisations de $\mathcal{L}_1$ présentées exemple 1.7 page 46). Il est représenté figure 2.

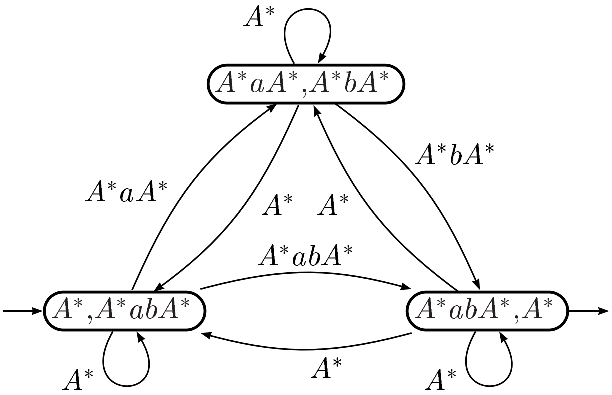


Fig. 2 – *L'automate universel de $\mathcal{L}_1$.*

Exemple 8.2 L'automate universel de $\mathcal{L}_2$ a deux états. Il est représenté figure 3.

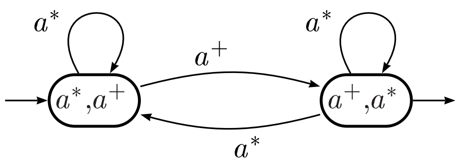


Fig. 3 – *L'automate universel de $\mathcal{L}_2$.*

Exemple 7.3 L'automate universel de $\{1,2\}$, dans $\mathbb{Z}/3\mathbb{Z}$, a six états (les six factorisations). Il est représenté figure 4. Sur la figure, chaque état est étiqueté par le facteur gauche de la factorisation à laquelle il correspond.

Parmi toutes les factorisations, il y en a deux qui vont jouer un rôle particulier.

Définition 2.3 *Soit $\mathcal{L}$ un sous-ensemble de N. Soit $L_i = \mathsf{max}\{X \subseteq N \mid X.\mathcal{L} \subseteq \mathcal{L}\}$ et $R_t = \mathsf{max}\{X \subseteq N \mid \mathcal{L}.X \subseteq \mathcal{L}\}$. Les couples $(L_i,\mathcal{L})$ et $(\mathcal{L},R_t)$ sont des factorisations*

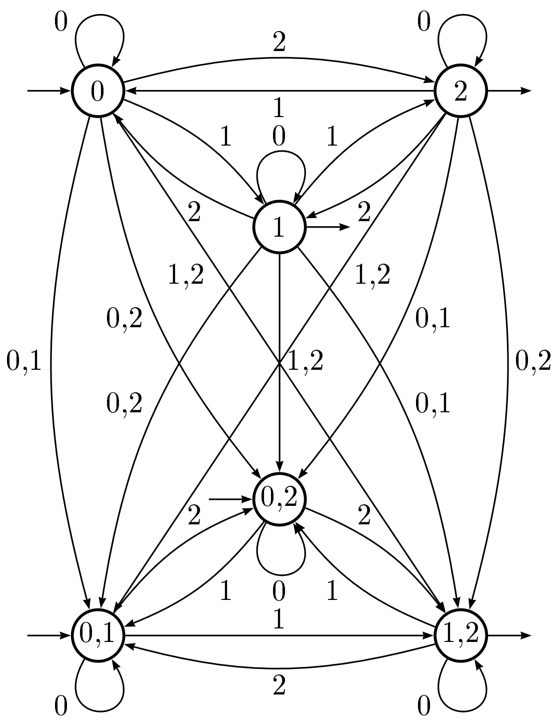


Fig. 4 – *L'automate universel de $\mathcal{L}_3$.*

et sont appelées respectivement ***factorisation initiale*** *et* ***factorisation terminale****. La factorisation initiale* (resp. *terminale) est un état initial* (resp. *final) de l'automate universel.*

— ∘ —

Propriétés. On vérifie à présent que l'automate construit a les propriétés voulues, c'est-à-dire qu'il reconnaît ce qu'on veut, bien sûr, mais aussi que les factorisations représentent effectivement le passé et le futur des états. C'est l'objet des deux propositions suivantes.

Proposition 2.1 *L'automate $\mathcal{U}_\mathcal{L}$ reconnaît l'ensemble $\mathcal{L}$.*

Démonstration. Soit $p_i = (L_i,\mathcal{L})$ et $p_t = (\mathcal{L},R_t)$ les factorisations initiale et terminale. $L_i.\mathcal{L}.R_t$ est inclus dans $\mathcal{L}.R_t$, donc dans $\mathcal{L}$. Le triplet $((L_i,\mathcal{L}),\mathcal{L},(\mathcal{L},R_t))$ est donc une transition de E, et tout mot de $\mathcal{L}$ est accepté par $\mathcal{U}_\mathcal{L}$.

Réciproquement, si u est reconnu par $\mathcal{U}_\mathcal{L}$, il existe des factorisations $(L_0,R_0),\ldots,(L_k,R_k)$ de $\mathcal{L}$ et $(u_1,\ldots,u_k)$ une factorisation de u en k éléments tels que $L_{i-1}.u_i.R_i$ est inclus dans $\mathcal{L}$ et 1_N appartient à L_0 et à R_k. Alors $L_0.u.R_k$ est inclus dans $\mathcal{L}$, donc u est un élément de $\mathcal{L}$. □

Proposition 2.2 *Pour chaque état $p = (L,R)$ de l'automate $\mathcal{U}_\mathcal{L}$, les facteurs L et R sont* *respectivement égaux au passé et au futur de p dans $\mathcal{U}_{\mathcal{L}}$:*

$$L = \mathsf{Past}_{\mathcal{U}_{\mathcal{L}}}(L,R)\ , \quad R = \mathsf{Fut}_{\mathcal{U}_{\mathcal{L}}}(L,R).$$

Démonstration. Soit $p_i = (L_i,\mathcal{L})$ et $p_t = (\mathcal{L},R_t)$ les factorisations initiale et terminale. Pour toute factorisation $p = (L,R)$, $L.R$ est inclus dans $\mathcal{L}$, donc $L_i.L.R$ aussi et (p_i,L,p) est une transition de $\mathcal{U}_{\mathcal{L}}$. Donc L est inclus dans le passé de p. De même, R est inclus dans le futur de p. Comme (L,R) est une factorisation (maximale) de $\mathcal{L}$ qui est l'ensemble reconnu par l'automate, la factorisation (L,R) est égale à la factorisation $(\mathsf{Past}_{\mathcal{U}_{\mathcal{L}}}(p),\mathsf{Fut}_{\mathcal{U}_{\mathcal{L}}}(p))$. □

Corollaire 2.3 *Tout automate universel est émondé.*

Démonstration. Chaque état correspond à une factorisation dont les facteurs représentent le passé et le futur de l'état. Ces facteurs ne sont pas l'ensemble vide, l'état est donc accessible et co-accessible. □

Nous allons maintenant voir ce qui fait de l'automate que nous avons défini un automate « universel » pour l'ensemble reconnu.

Proposition 2.4 *Soit $\mathcal{A}$ un automate émondé reconnaissant un sous-ensemble de $\mathcal{L}$. Alors, il existe un morphisme φ de $\mathcal{A}$ dans $\mathcal{U}_{\mathcal{L}}$. De plus, $\mathcal{U}_{\mathcal{L}}$ est le plus petit*[1] *automate reconnaissant $\mathcal{L}$ ayant cette propriété.*

Démonstration. Soit p un état de $\mathcal{A}$. Le passé et le futur de p forment une sous-factorisation de $\mathcal{L}$. On pose[2] :

$$R_p = \mathsf{max}\{X \subseteq N \mid \mathsf{Past}_{\mathcal{A}}(p).X \subseteq \mathcal{L}\} = \mathsf{Past}_{\mathcal{A}}(p)^{-1}\mathcal{L} \quad \text{et } L_p = \mathsf{max}\{X \subseteq N \mid X.R_p \subseteq \mathcal{L}\}.$$

Il s'agit bien d'une factorisation. On vérifie que si p est initial ou final, (L_p,R_p) aussi.
Si (p,u,q) est une transition de $\mathcal{A}$, comme $\mathsf{Past}_{\mathcal{A}}(p).u$ est inclus dans le passé de q, $\mathsf{Past}_{\mathcal{A}}(p).u.R_q$ est inclus dans $\mathcal{L}$, donc $u.R_q$ est inclus dans R_p et $L_p.u.R_q$ est inclus dans $\mathcal{L}$. L'élément u fait donc partie de l'étiquette de la transition entre (L_p,R_p) et (L_q,R_q).
Supposons qu'il existe un automate $\mathcal{A}'$ dans lequel tout automate reconnaissant $\mathcal{L}$ s'envoie par morphisme. S'il existe un morphisme non injectif de $\mathcal{U}_{\mathcal{L}}$ dans $\mathcal{A}'$, deux états p et q de $\mathcal{U}_{\mathcal{L}}$ ont une même image, dont le passé contient $L_p \cup L_q$ et le futur $R_p \cup R_q$. Comme (L_p,R_p) et (L_q,R_q) sont des factorisations, $(L_p \cup L_q).(R_p \cup R_q)$ n'appartient pas à $\mathcal{L}$ et $\mathcal{A}'$ reconnaît strictement plus que l'ensemble $\mathcal{L}$. Donc tout morphisme de $\mathcal{U}_{\mathcal{L}}$ dans $\mathcal{A}'$ est injectif. Le cardinal de l'ensemble des états et de l'ensemble des transitions de $\mathcal{A}'$ est donc supérieur ou égal à celui de $\mathcal{U}_{\mathcal{L}}$. □

L'automate universel peut, d'après ce qu'on en a dit, être infini. Nous allons voir sous quelle condition il est fini.

— ∘ —

1. Comme on n'a fait aucune supposition sur l'ensemble reconnu, l'automate est éventuellement infini. Cet automate est le « plus petit » dans le sens où il est inclus dans tout automate ayant cette propriété.

2. On voit que cette définition n'est pas symétrique. On peut ainsi définir deux morphismes canoniques (non nécessairement distincts) de $\mathcal{A}$ dans $\mathcal{U}_{\mathcal{L}}$. Ce ne sont d'ailleurs pas nécessairement les seuls morphismes de $\mathcal{A}$ dans $\mathcal{U}_{\mathcal{L}}$ (voir [55]).

1.2 Automate universel d'un ensemble reconnaissable

On suppose le sous-ensemble $\mathcal{L}$ de N reconnaissable. Ceci signifie qu'il existe un monoïde M fini, un morphisme φ de N dans M et P inclus dans M tel que $\mathcal{L} = P\varphi^{-1}$.

Proposition 2.5 *L'automate universel d'un ensemble a un nombre fini d'états si et seulement si cet ensemble est reconnaissable.*

Démonstration. Supposons l'ensemble $\mathcal{L}$ reconnaissable. Soit (L,R) une factorisation de $\mathcal{L}$. On a :

$$\begin{aligned} L.R &\subseteq \mathcal{L} \\ L\varphi.R\varphi &\subseteq \mathcal{L}\varphi \\ L\varphi\varphi^{-1}.R\varphi\varphi^{-1} &\subseteq \mathcal{L}\varphi\varphi^{-1} = \mathcal{L} \end{aligned}$$

Comme L est inclus dans $L\varphi\varphi^{-1}$, et R dans $R\varphi\varphi^{-1}$ et puisque (L,R) est une factorisation, $L = L\varphi\varphi^{-1}$ et $R = R\varphi\varphi^{-1}$. Les factorisations de $\mathcal{L}$ dans N sont donc les images inverses des factorisations de P dans M qui est fini. Elles sont donc en nombre fini.

Réciproquement, supposons le nombre de factorisations finies. A tout élément x du monoïde N, on associe la factorisation (L_x,R_x), telle que $R_x = \{y \in N \mid x.y \in \mathcal{L}\} = x^{-1}\mathcal{L}$. L'ensemble des factorisations munies de la loi induite par cette application est un monoïde fini qui reconnaît $\mathcal{L}$. En effet, x est un élément de $\mathcal{L}$ si et seulement si R_x contient 1_N. □

La proposition précédente assure que le nombre d'états de l'automate universel d'un ensemble reconnaissable est fini. Toutefois, pour que l'on puisse parler d'automate fini, il faut qu'on puisse donner une description finie, non seulement des états, mais aussi des transitions. La proposition suivante répond à cette contrainte.

Proposition 2.6 *Les étiquettes des transitions de l'automate universel d'un ensemble reconnaissable sont reconnaissables.*

Démonstration. L'argument est le même que dans la preuve précédente; si le triplet $((L_p,R_p),X,(L_q,R_q))$ est une transition,

$$L_p\varphi\varphi^{-1}.X\varphi\varphi^{-1}.R_q\varphi\varphi^{-1} = L_p.X\varphi\varphi^{-1}.R_q \subseteq \mathcal{L}.$$

Par maximalité de X, on obtient $X = X\varphi\varphi^{-1}$ qui est donc reconnaissable. □

Exemple 1.9 Soit M_1 le monoïde présenté exemple 1.4 page 22. Soit φ le morphisme de A^* dans M_1 défini par $a\varphi = x$ et $b\varphi = y$. Comme nous l'avons déjà vu page 31, le langage $\mathcal{L}_1$ est l'image inverse de 0_{M_1} par ce morphisme. Les factorisations de $\mathcal{L}_1$ dans A^* correspondent donc aux factorisations de 0_{M_1} dans M_1. Ce sont $(M_1,0)$, $(\{x,z,0\},\{y,z,0\})$ et $(0,M_1)$ qui correspondent bien aux factorisations données page 48. L'automate universel de 0 dans M_1 est présenté figure 5. Si on applique φ^{-1} à chaque étiquette de l'automate, on retrouve l'automate universel de $\mathcal{L}_1$ dessiné figure 2.

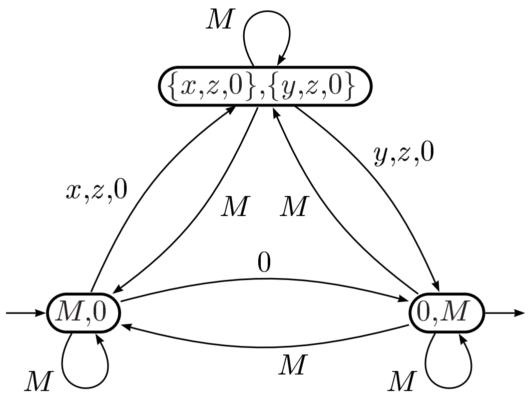


FIG. 5 – *L'automate universel de* 0 *dans* M_1.

EXEMPLE 7.4 Soit $A = \{a,b\}$ et $\mathcal{L}'_3 = \{u \in A^* \mid |u|_a \neq |u|_b \mod 3\}$. Le monoïde syntaxique de ce langage est $N_3 = \mathbb{Z}/3\mathbb{Z}$, avec φ_3 le morphisme de A^* sur N_3 défini par $a\varphi_3 = 1$ et $b\varphi_3 = 2$. L'image de $\mathcal{L}'_3$ par φ_3 est $\mathcal{L}_3 = \{1,2\}$. L'automate minimal, qui est dans ce cas le graphe de Cayley du monoïde syntaxique, est représenté figure 6.

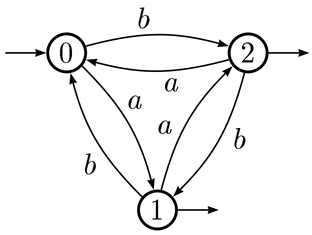


FIG. 6 – *L'automate minimal de* $\mathcal{L}'_3$.

On peut calculer les images inverses des éléments de N_3 par φ :

$$0\varphi^{-1} = \mathcal{L}_{3,0} = (ab + (a^2 + b)(ba)^*(b^2 + a))^*$$
$$1\varphi^{-1} = \mathcal{L}_{3,1} = (ab + (a^2 + b)(ba)^*(b^2 + a))^*(a + b^2)(ab)^*$$
$$2\varphi^{-1} = \mathcal{L}_{3,2} = (ab + (a^2 + b)(ba)^*(b^2 + a))^*(a^2 + b)(ba)^*$$

Pour obtenir les factorisations de $\mathcal{L}'_3$, il suffit donc de remplacer chaque élément de N_3 par son image inverse selon φ dans les factorisations données dans l'exemple 7.2 (page 46). De même, pour obtenir l'automate universel de $\mathcal{L}'_3$, il suffit de remplacer chaque étiquette de l'automate représenté figure 4 par son image inverse selon φ.

— ◦ —

1.3 Automate universel et générateurs du monoïde

Les transitions de l'automate universel d'un langage de A^* défini plus haut sont étiquetées par des langages. Lorsque l'on travaille avec un monoïde finiment engendré (en particulier dans le monoïde libre A^*), il est plus habituel de traiter des automates dont les étiquettes sont étiquetées par des lettres. On va voir qu'on peut définir un automate universel sur un alphabet.

DÉFINITION 2.4 *Soit $\mathcal{L}$ un sous-ensemble d'un monoïde N engendré par les éléments d'un ensemble A et Q l'ensemble des factorisations de $\mathcal{L}$ dans N. L'automate A-universel de $\mathcal{L}$ est $\mathcal{U}_{\mathcal{L}} = \langle\, Q,A,E,I,T \,\rangle$, avec :*

$$\begin{aligned} I &= \{(L,R) \in Q \mid 1_N \in L\}, \\ T &= \{(L,R) \in Q \mid 1_N \in R\}, \\ E &= \{((L,R),a,(L',R')) \in Q \times A \times Q \mid L.a.R' \subseteq \mathcal{L}\}. \end{aligned}$$

EXEMPLE 8.3 L'automate $\{a\}$-universel de $\mathcal{L}_2 = a^+$ est représenté figure 7.

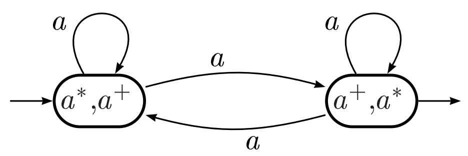


FIG. 7 – *L'automate $\{a\}$-universel de $\mathcal{L}_2$.*

EXEMPLE 7.5 L'automate $\{1\}$-universel de $\{1,2\}$, dans $\mathbb{Z}/3\mathbb{Z}$ est représenté figure 8.

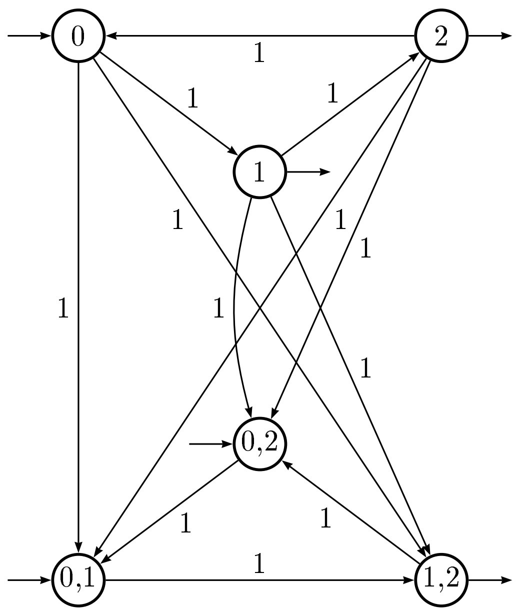


FIG. 8 – *L'automate $\{1\}$-universel de $\mathcal{L}_3$.*

Les états de l'automate A-universel sont les mêmes que ceux de l'automate universel. Les étiquettes des transitions sont exactement l'intersection des étiquettes de l'automate universel avec A. Nous allons voir que ceci ne diminue pas la puissance de l'objet.

Proposition 2.7 *L'automate A-universel de $\mathcal{L}$ reconnaît $\mathcal{L}$.*

Démonstration. Il est clair que l'automate reconnaît au plus $\mathcal{L}$ puisqu'il est une projection de l'automate universel . D'autre part, tout automate étiqueté par des éléments de A qui reconnaît cet ensemble s'envoie dans l'automate A-universel par morphisme. Celui-ci reconnaît donc l'ensemble $\mathcal{L}$. □

Remarque. Lorsqu'il n'y a pas d'ambiguïté, on appellera l'automate A-universel automate universel. C'est le cas, par exemple, lorsqu'on travaille avec le monoïde libre engendré par un alphabet fini A.

— ∘ —

1.4 Automate universel d'un langage rationnel de A^*

Un langage rationnel de A^* est, d'après le théorème de Kleene, reconnaissable. Il admet donc un automate universel fini et c'est la combinaison des deux paragraphes précédents qui forme le cadre dans lequel nous allons travailler. Nous utiliserons en effet l'automate A-universel des langages étudiés.

L'automate universel sera donc, dans ce cas, un automate fini dont les étiquettes sont des lettres.

Exemple 1.10 L'automate (A)-universel du langage $\mathcal{L}_1$ est obtenu à partir de l'automate de la figure 2 en remplaçant chaque étiquette par son intersection avec l'alphabet. Le résultat est présenté figure 9.

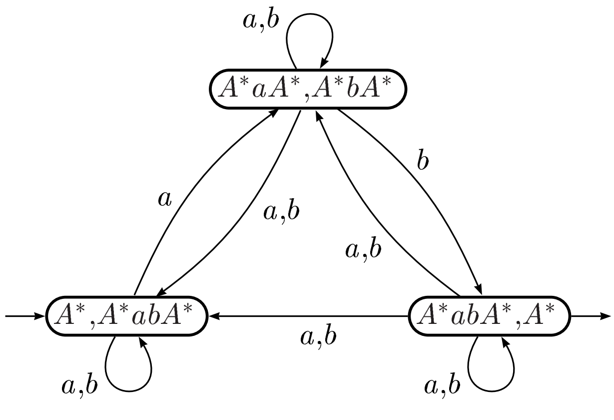


Fig. 9 – *L'automate (A)-universel de $\mathcal{L}_1$.*

EXEMPLE 7.6 L'automate (A)-universel du langage $\mathcal{L}'_3$ est obtenu en prenant l'intersection des langages qui étiquettent l'automate universel avec l'alphabet :$\mathcal{L}_{3,0} \cap A = \emptyset$,$\mathcal{L}_{3,1} \cap A = \{a\}$,$\mathcal{L}_{3,2} \cap A = \{b\}$. L'automate universel de $\mathcal{L}'_3$ présenté figure 10 est donc obtenu à partir de la figure 4 en remplaçant 1 par a, 2 par b et en supprimant les transitions étiquetées par 0.

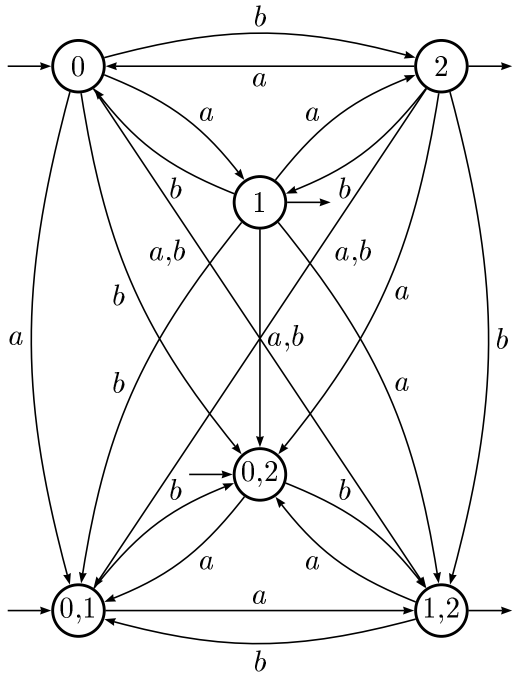


FIG. 10 – *L'automate universel de $\mathcal{L}'_3$.*

On peut remarquer que la définition de l'automate universel est symétrique, il n'y a donc pas d'automate « co-universel ». En d'autres termes, on a la proposition suivante :

PROPOSITION 2.8 *Soit $\mathcal{L}$ un langage rationnel de A^*. L'automate universel $\mathcal{U}_{\overline{\mathcal{L}}}$ de l'image miroir de $\mathcal{L}$ est l'automate transposé de l'automate universel $\mathcal{U}_{\mathcal{L}}$ de $\mathcal{L}$.*

Démonstration. Soit (L,R) et (L',R') deux factorisations de $\mathcal{L}$. Alors $(\overline{R},\overline{L})$ et $(\overline{R'},\overline{L'})$ sont des factorisations de $\overline{\mathcal{L}}$.

Si (L,R) est un état initial de $\mathcal{U}_{\mathcal{L}}$, le mot vide 1_{A^*} appartient à L donc à $\overline{L}$ et $(\overline{R},\overline{L})$ est un état final de $\mathcal{U}_{\overline{\mathcal{L}}}$. De même, si (L,R) est un état final de $\mathcal{U}_{\mathcal{L}}$, $(\overline{R},\overline{L})$ est un état initial de $\mathcal{U}_{\overline{\mathcal{L}}}$.

Si il existe $a \in A$ tel que $((L,R),a,(L',R'))$ est une transition de $\mathcal{U}_{\mathcal{L}}$, alors $L.a.R'$ est inclus dans $\mathcal{L}$, donc $\overline{R'}.a.\overline{L}$ est inclus dans $\overline{\mathcal{L}}$ et $((\overline{R'},\overline{L'}),a,(\overline{R},\overline{L}))$ est une transition de $\mathcal{U}_{\overline{\mathcal{L}}}$.

Donc l'automate transposé de l'automate universel $\mathcal{U}_{\mathcal{L}}$ est un sous-automate de $\mathcal{U}_{\overline{\mathcal{L}}}$. Réciproquement, par dualité, on obtient l'inclusion inverse, ces deux automates sont donc égaux. □

Pour n'importe quel automate $\mathcal{A}$ reconnaissant le langage $\mathcal{L}$, il existe un morphisme de $\mathcal{A}$ dans l'automate universel. En particulier, l'automate minimal d'un langage s'envoie dans l'automate universel. La propriété de minimalité signifie entre autre que tout morphisme non injectif donne un automate qui reconnaît un langage strictement plus grand. Comme l'automate universel ne reconnaît que $\mathcal{L}$, tout morphisme qui envoie l'automate minimal dans $\mathcal{U}_{\mathcal{L}}$ est injectif, il en est de même pour l'automate co-déterministe minimal. Ce fait peut être exprimé par la proposition suivante.

Proposition 2.9 *L'automate minimal et l'automate co-déterministe minimal d'un langage rationnel $\mathcal{L}$ sont des sous-automates de l'automate universel $\mathcal{U}_{\mathcal{L}}$.*

— ∘ —

2 Calcul effectif de l'automate universel

Nous donnons ici une nouvelle méthode de calcul de l'automate universel. Comme l'a fait remarquer O. Carton ([12]), on peut effectuer le calcul de l'automate universel d'un langage $\mathcal{L}$ dans le monoïde syntaxique $M_{\mathcal{L}}$ de $\mathcal{L}$. On calcule alors la table de ce monoïde et on repère les couples maximaux (L,R) de sous-ensembles de $M_{\mathcal{L}}$ tels que $L.R$ est inclus dans P, image canonique de $\mathcal{L}$ dans $M_{\mathcal{L}}$.

Toutefois, si on considère $M_{\mathcal{L}}$ comme monoïde de transition de l'automate minimal de $\mathcal{L}$ – ce qui est une façon courante de le calculer –, on peut remarquer deux choses. D'une part, si un élément d'image α appartient à un facteur gauche L, tout élément d'image β qui agit de la même façon (à droite) sur l'état initial i de l'automate minimal appartient au même facteur gauche :

$$\forall \alpha \in L, \forall \beta \in M_{\mathcal{L}}, \quad i\alpha = i\beta \Rightarrow \beta \in L.$$

D'autre part, si un élément d'image α appartient à un facteur droit R, tout élément d'image β qui agit de la même façon à gauche sur l'ensemble des état terminaux T de l'automate minimal appartient au même facteur droit :

$$\forall \alpha \in R, \forall \beta \in M_{\mathcal{L}}, \quad T\alpha^{-1} = T\beta^{-1} \Rightarrow \beta \in R.$$

On peut donc fixer deux relations d'équivalences sur $M_{\mathcal{L}}$:

$$\alpha \sim_i \beta \Leftrightarrow i\alpha = i\beta, \text{ et } \quad \alpha \sim_T \beta \Leftrightarrow T\alpha^{-1} = T\beta^{-1}.$$

Le produit de deux éléments α et β de $M_{\mathcal{L}}$ appartient à P si et seulement si l'intersection de $i\alpha$ et $T\beta^{-1}$ n'est pas vide. On peut donc calculer les factorisations dans $(M/\sim_i) \times (M/\sim_T)$. Remarquons au passage que $M/\sim_i$ est isomorphe à l'ensemble Q des états de l'automate minimal, puisque celui-ci est déterministe et que chaque état est accessible.

Comme, par maximalité, une factorisation est caractérisée par son facteur gauche, on peut décrire les factorisations par des ensembles d'états de l'automate minimal.

La nouvelle description que l'on obtient est donnée par la proposition suivante :

Théorème 2.1 *Soit $\mathcal{A}_{\mathcal{L}} = \langle Q,A,E,\{i\},T \rangle$ l'automate minimal d'un langage rationnel $\mathcal{L}$. Soit $P = \{u \cdot T \mid u \in A^*\}$ l'ensemble des états du co-déterminisé de $\mathcal{A}_{\mathcal{L}}$.*

Soit $P_\cap$ la clôture de P par intersection, privée de l'ensemble vide.
Alors, l'automate universel de $\mathcal{U}_\mathcal{L}$ est isomorphe à $\langle\, P_\cap, A, F, J, U\,\rangle$, avec :

$$\begin{aligned} J =& \{X \mid i \in X\} \\ U =& \{X \mid X \subseteq T\} \\ F =& \{(X,a,Y) \mid X \cdot a \subseteq Y \text{ et } \forall p \in X,\ p \cdot a \neq \emptyset\}. \end{aligned}$$

Lemme 2.10 *Soit $\mathcal{A} = \langle\, Q,A,E,i,T\,\rangle$ l'automate minimal d'un langage $\mathcal{L}$ et $P_\cap$ défini à partir de $\mathcal{A}$ comme dans le théorème ci-dessus. Il existe une bijection entre les éléments de $P_\cap$ et l'ensemble F des factorisations de $\mathcal{L}$:*

$$\begin{aligned} P_\cap &\longrightarrow F \\ X &\longmapsto \left(\bigcup_{p \in X} \mathsf{Past}_\mathcal{A}(p),\ \bigcap_{p \in X} \mathsf{Fut}_\mathcal{A}(p) \right) \end{aligned} \qquad \begin{aligned} F &\longrightarrow P_\cap \\ (L,R) &\longmapsto i \cdot L \end{aligned}$$

Démonstration. Examinons la première application. Soit X dans $P_\cap$ et son image (L,R). L'ensemble R n'est pas vide car X est un sous-ensemble d'un état du co-déterminisé de $\mathcal{A}$. L'ensemble L n'est pas vide car chaque élément de X est un état accessible de $\mathcal{A}$. Pour tout u dans L et tout v dans R, il existe un élément p de X tel que u est dans le passé de p et, par définition, v est dans le futur de p (comme de n'importe quel autre élément de X). Donc $u.v$ est un mot de $\mathcal{L}$ et (L,R) est une sous-factorisation.

Il faut maintenant montrer que (L,R) est une factorisation. On pose $L' = \{u \in A^* \mid u.R \subseteq \mathcal{L}\}$ et $R' = \{v \in A^* \mid L'.v \subseteq \mathcal{L}\}$; (L',R') est une factorisation qui domine (L,R). On pose $X' = i \cdot L'$. Si $X = X'$, alors $L = L'$ et

$$\begin{aligned} R' =& \{v \mid \forall u \in L, u.v \in \mathcal{L}\} \\ =& \{v \mid \forall u \in L, i \cdot (u.v) \in T\} \\ =& \{v \mid X \cdot v \in T\} \\ =& \bigcap_{p \in X} \mathsf{Fut}_\mathcal{A}(p) \\ =& R. \end{aligned}$$

Donc, si (L,R) n'est pas maximal, $X \subset X'$. Quel que soit v tel que $X \subseteq v \cdot T$, quel que soit u dans L', $u.v$ est un mot de $\mathcal{L}$, donc $X' \subseteq v \cdot T$. Donc X n'appartient pas à $P_\cap$, ce qui est contradictoire.

Cette application est injective. En effet, les passés des états de l'automate minimal sont disjoints, deux sous-ensembles de Q distincts donnent donc deux factorisations dont les facteurs gauches sont distincts.

Cette application est surjective. En effet, pour toute factorisation (L,R), on peut définir $X = i \cdot L$. Comme il s'agit d'une factorisation, $R = \{v \mid X \cdot R \subseteq T\}$; donc X apparaît aussi comme l'intersection des éléments de $\{Y \subseteq Q \mid \exists v \in R, Y = v \cdot T\}$; donc X appartient bien à $P_\cap$ et l'image de X est évidemment (L,R). $\square$

Démonstration du théorème 2.1. Soit $\mathcal{A} = \langle Q,A,E,i,T \rangle$ l'automate minimal de $\mathcal{L}$. Posons $\mathcal{B} = \langle P_\cap,A,F,J,U \rangle$ l'automate défini dans le théorème et $\mathcal{U}_\mathcal{L} = \langle S,A,G,K,V \rangle$ l'automate universel de $\mathcal{L}$. Le lemme précédent établit qu'il existe une bijection entre $P_\cap$ et S. Il suffit maintenant de montrer qu'il y a correspondance entre les états initiaux, les états terminaux et entre les transitions. Soit φ la bijection de $P_\cap$ sur S définie dans le lemme.

Si $X \in P_\cap$ est initial, i appartient à X et le mot vide appartient au facteur gauche de $X\varphi$; $X\varphi$ est donc initial.
Si une factorisation $p = (L,R)$ est un état initial de $\mathcal{U}_\mathcal{L}$, le mot 1_{A^*} est dans L, donc i appartient à $p\varphi^{-1}$ qui est donc un état initial de $\mathcal{B}$.

Si $X \in P_\cap$ est final, X est inclus dans T, donc le mot vide appartient au facteur droit de $X\varphi$; $X\varphi$ est donc final.
Si une factorisation $p = (L,R)$ est un état final de $\mathcal{U}_\mathcal{L}$, L est inclus dans $\mathcal{L}$, donc $p\varphi^{-1}$ est inclus dans T et est donc un état final de $\mathcal{B}$.

Si (X,a,Y) est une transition de $\mathcal{B}$, alors $X \cdot a \subseteq Y$ et, pour tout r dans X, $r \cdot a$ n'est pas vide. Donc, si L et L' sont les facteurs gauches respectifs de $X\varphi$ et $Y\varphi$, $L.a \subseteq L'$. Donc $X\varphi,a,Y\varphi$ est une transition de $\mathcal{U}_\mathcal{L}$.
Si (p,a,q) est une transition de $\mathcal{U}_\mathcal{L}$, avec L et L' facteurs gauches de p et q respectivement, on a $L.a \subseteq L'$; donc $p\varphi^{-1} \cdot a \subseteq q\varphi^{-1}$, et comme $L.a$ est un sous-facteur gauche du langage, pour tout r dans $p\varphi^{-1}$, $r \cdot a$ n'est pas vide. Donc $(p\varphi^{-1},a,q\varphi^{-1})$ est une transition de $\mathcal{B}$. □

Corollaire 2.11 *Soit $\mathcal{L}$ un langage rationnel. Soit $\mathcal{A}_\mathcal{L}$ son automate minimal et n le nombre d'états de $\mathcal{A}_\mathcal{L}$. Le nombre d'états de l'automate universel est au plus $2^n - 1$.*

Démonstration. Les états de l'automate universel sont, on l'a vu, indexés par des parties non vides d'états de $\mathcal{A}_\mathcal{L}$. Leur nombre est donc inférieur à $2^n - 1$. □

Remarque 2.3 Cette borne peut être effectivement atteinte pour tout n. Posons $\mathcal{A}_n = \langle \mathbb{Z}/n\mathbb{Z},\{a;b\},E,\{0\},[1;n-1] \rangle$, avec

$$E = \{(n,a,n+1) \mid n \in \mathbb{Z}/n\mathbb{Z}\} \cup \{(n,b,n+1) \mid n \in [1;n-1]\} \cup \{(0,b,0)\}$$

Cet automate est minimal. Les états de son co-déterminisé sont les sous-ensembles à $n-1$ éléments ainsi que $\mathbb{Z}/n\mathbb{Z}$ tout entier. A tout sous-ensemble non vide de $\mathbb{Z}/n\mathbb{Z}$ correspond une factorisation du langage.

Remarque 2.4 On peut effectuer la construction précédente sans calculer exactement l'ensemble des états de l'automate universel mais en considérant tous les sous-ensembles non vides de l'ensemble des états de l'automate minimal. On obtient alors un automate qui contient l'automate universel et qui compte exactement $2^n - 1$ états. Cet automate correspond à **l'automate des sous-ensembles d'ordre 0** présenté par R. Cohen et J. Brzozowski [18]. Nous allons par la suite généraliser cette construction pour obtenir, à partir de n'importe quel automate qui accepte le langage, un automate qui contient l'automate universel. Un autre automate, l'automate **fondamental**, présenté par O. Matz et A.Potthoff [45] et construit à partir de l'automate minimal, contient lui aussi l'automate universel.

Exemple 10.1 Soit $\mathcal{L}_4$ le langage dont l'automate minimal est présenté figure 11 a). Le monoïde de transition de l'automate minimal (engendré par les actions des lettres sur les états) compte vingt-quatre éléments. Il s'agit en effet de l'ensemble des applications d'un ensemble de trois éléments dans lui-même, hormi les trois transpositions. Utilisons le théorème 2.1 pour calculer l'automate universel. Les états du co-déterminisé de l'automate minimal sont :

$$\{p\},\{q\},\{r\},\{p,q\},\{p,r\},\{q,r\},\{p,q,r\}.$$

Cet ensemble est clos pour l'intersection (sauf le vide); ce sont donc les états de l'automate universel.

La table de transition de l'automate universel est la suivante :

	p	q	r	p,q	p,r	q,r	p,q,r
p	$-$	$a+c$	b	$a+c$	b	$a+b+c$	$a+b+c$
q	$b+c$	$-$	a	$b+c$	$a+b+c$	a	$a+b+c$
r	$a+c$	$-$	b	$a+c$	$a+b+c$	b	$a+b+c$
p,q	$-$	$-$	$-$	c	b	a	$a+b+c$
p,r	$-$	$-$	b	$a+c$	b	b	$a+b+c$
q,r	c	$-$	$-$	c	$a+b+c$	$-$	$a+b+c$
p,q,r	$-$	$-$	$-$	c	b	$-$	$a+b+c$

Les états initiaux sont ceux qui contiennent p et les états terminaux sont inclus dans $\{r\}$.

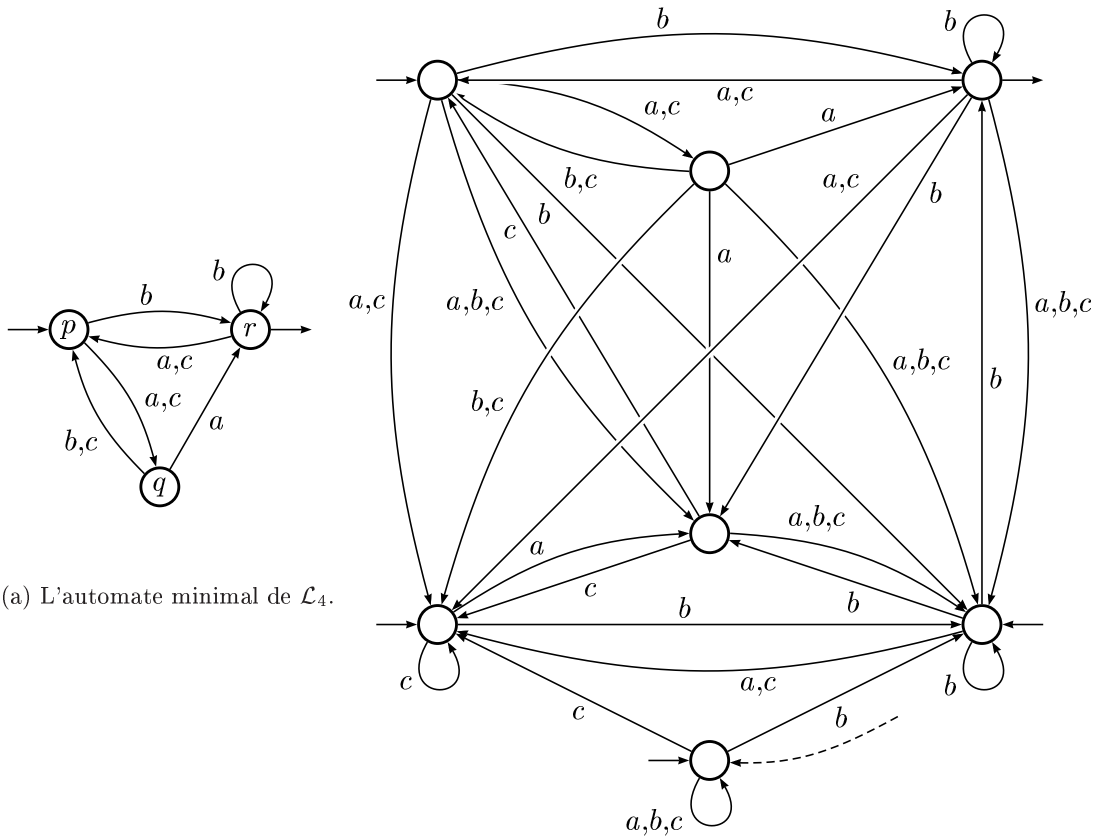
L'automate universel de ce langage compte sept états et est présenté figure 11 b). La flèche en pointillé arrivant sur l'état du bas indique qu'on peut arriver dans cet état en lisant n'importe quelle lettre à partir de n'importe quel état.

— ∘ —

3 Écorché de l'automate universel

Nous introduisons ici une nouvelle présentation de l'automate universel qui permettra de mettre en relief certaines propriétés des automates universels que nous verrons par la suite, mais aussi, ce qui n'est pas négligeable, de pouvoir dessiner des automates universels sans être asphyxié par la profusion des transitions.

Pour cela, il peut être intéressant de distinguer les transitions « significatives » et celles qui découlent d'une « perte d'information ». Ainsi, si (L,R) et (L',R') sont deux factorisations telles que $L \subseteq L'$, pour toute transition partant de (L',R') étiquetée par une lettre a et allant dans un état q, il existe une transition partant de (L,R) étiquetée par a et allant dans le même état q.



(a) L'automate minimal de $\mathcal{L}_4$.

(b) L'automate universel de $\mathcal{L}_4$.

FIG. 11 – *Deux automates canoniques de $\mathcal{L}_4$.*

On peut donc supprimer ces dernières et considérer qu'il existe une transition spontanée de (L,R) à (L',R'). Dans le même esprit, on peut ne retenir que les transitions spontanées minimales, c'est-à-dire celles dont la clôture permet de retrouver les autres. C'est le sens de la définition suivante.

DÉFINITION 2.5 *Soit $\mathcal{L}$ un langage de A^* et Q l'ensemble de ses factorisations. L'**écorché de l'automate universel** $\check{\mathcal{U}}_{\mathcal{L}} = \langle\, Q,A,E,I,T \,\rangle$[3] est défini par :*

$$\begin{aligned} I =& \{p\}, \text{ avec } p \text{ factorisation initiale,} \\ T =& \{q\}, \text{ avec } q \text{ factorisation finale,} \\ E =& \{((L,R),a,(L',R')) \in Q \times A \times Q \mid (L,R') \in \mathsf{max}\{(X,Y) \mid X.a.Y \subseteq \mathcal{L}\}\} \\ & \cup \{((L,R),1_{A^*},(L',R')) \in Q \times \{1_{A^*}\} \times Q \mid L' \in \mathsf{min}\{X \mid L \subset X\}. \end{aligned} \tag{1}$$

3. Nous utilisons cette métaphore anatomique pour souligner le fait que cette nouvelle présentation permet d'aller au-delà des apparences (touffues) de l'automate universel pour comprendre comment il « fonctionne ».

Remarque 2.5 Les facteurs gauches appartenant à l'ensemble $\mathsf{min}\{X \mid L \subset X\}$ correspondent aux facteurs droits de l'ensemble $\mathsf{max}\{Y \mid Y \subset R\}$. On peut définir un ordre sur les états de l'automate universel : $(L,R) < (L',R') \Leftrightarrow L \subset L'$. Les transitions spontanées de l'écorché de l'automate universel forment le graphe orienté de cet ordre.

Exemple 1.11 Les factorisations de $\mathcal{L}_1 = A^*abA^*$ sont :

$$p = (A^*,A^*abA^*), \qquad q = (A^*aA^*,A^*bA^*) \qquad \text{et} \qquad r = (A^*abA^*,A^*).$$

On voit que les facteurs gauches forment une chaîne décroissante; les transitions spontanées de l'écorché de l'automate universel sont donc $(q,1_{A^*},p)$ et $(r,1_{A^*},q)$. Les sous-factorisations maximales de $\mathcal{L}_1$ à trois facteurs avec a comme facteur central sont :

$$(A^*,a,A^*bA^*) \qquad \text{et} \qquad (A^*abA^*,a,A^*);$$

les transitions étiquetées par a sont donc (p,a,q) et (r,a,r). De même, les transitions étiquetées par b sont (p,b,p) et (q,b,r). L'écorché de l'automate universel de $\mathcal{L}_1$ est donc celui de la figure 12.

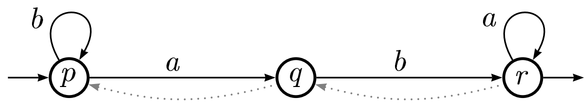


Fig. 12 – *Écorché de l'automate universel de* $\mathcal{L}_1$.

Exemple 8.4 Les factorisations du langage $\mathcal{L}_2 = a^+$ étant (a^*,a^+) et (a^+,a^*), le calcul de l'écorché de l'automate universel se fait simplement. Le résultat est présenté figure 13.

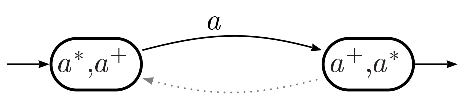
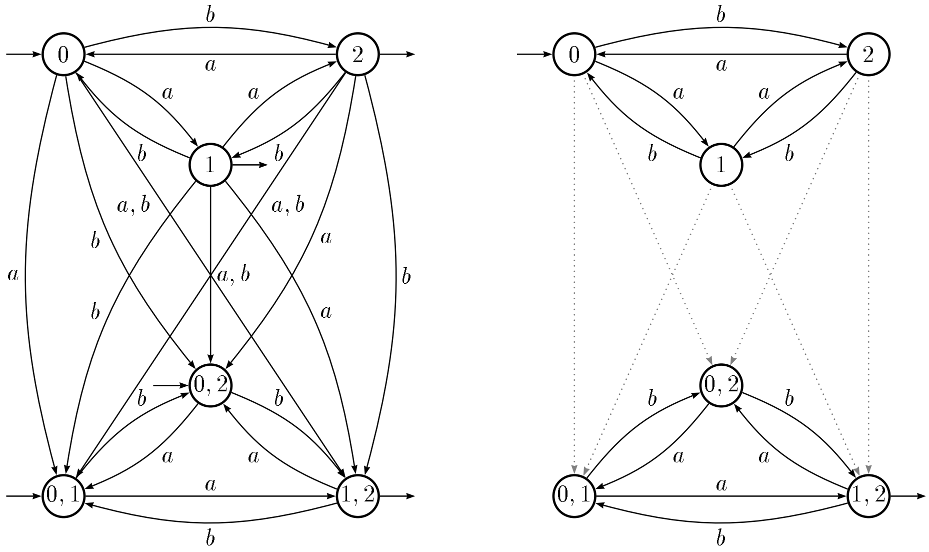


Fig. 13 – *Écorché de l'automate universel de* $\mathcal{L}_2$.

Exemple 7.7 La figure 14 montre d'une part l'automate $(\{a\})$-universel du langage $\mathcal{L}'_3$ et d'autre part l'écorché de ce même automate.

Exemple 10.2 Reprenons l'exemple du langage $\mathcal{L}_4$ présenté exemple 10.1 page 59. Les factorisations y étaient représentées par des sous-ensembles d'états de l'automate minimal. Les inclusions entre factorisations suivent les inclusions entre ensembles.

Si une lettre a apparaît dans deux cases distinctes de la table de transition de $\mathcal{U}_{\mathcal{L}_4}$ indexées par $(\mathcal{L}_{i_1},\mathcal{L}_{j_1})$ et $(\mathcal{L}_{i_2},\mathcal{L}_{j_2})$ et que $\mathcal{L}_{i_1} \subseteq \mathcal{L}_{i_2}$ (le passé de i_1 est inclus dans celui



(a) L'automate universel de $\mathcal{L}'_3$ ...

(b) ... et son écorché.

FIG. 14 – *Deux façons de représenter l'automate universel.*

de i_2) et $\mathcal{L}_{j_2} \subseteq \mathcal{L}_{j_1}$ (le futur de j_1 est inclus dans le futur de j_2), la lettre a ne figurera pas dans la case $(\mathcal{L}_{i_1},\mathcal{L}_{j_1})$ de la table de transition de $\check{\mathcal{U}}_{\mathcal{L}_4}$.

D'autre part, on indique une transition spontanée en $(\mathcal{L}_i,\mathcal{L}_j)$ si $\mathcal{L}_i \subset \mathcal{L}_j$ et qu'il n'existe pas k tel que $\mathcal{L}_i \subset \mathcal{L}_k \subset \mathcal{L}_j$.

La table de transition de $\check{\mathcal{U}}_{\mathcal{L}_4}$ est donc la suivante :

	p	q	r	p,q	p,r	q,r	p,q,r
p	–	$a+c$	–	1_{A^*}	1_{A^*}	–	–
q	b	–	a	1_{A^*}	–	1_{A^*}	–
r	a	–	–	–	1_{A^*}	1_{A^*}	–
p,q	–	–	–	–	–	a	1_{A^*}
p,r	–	–	b	a	–	–	1_{A^*}
q,r	c	–	–	–	a	–	1_{A^*}
p,q,r	–	–	–	c	b	–	a

Le seul état initial est celui qui correspond à p et le seul état final reste r. L'automate obtenu est représenté figure 15. Les transitions spontanées sont indiquées par des pointillés gris.

L'écorché de l'automate universel compte peu de transitions non spontanées. Plus précisément :

PROPOSITION 2.12 *Soit* $\check{\mathcal{U}}_{\mathcal{L}}$ *l'écorché de l'automate universel d'un langage rationnel de* A^*. *Alors, pour toute lettre* a *de* A, *pour tout état* p *de* $\check{\mathcal{U}}_{\mathcal{L}}$, *il y a au plus une transition étiquetée*

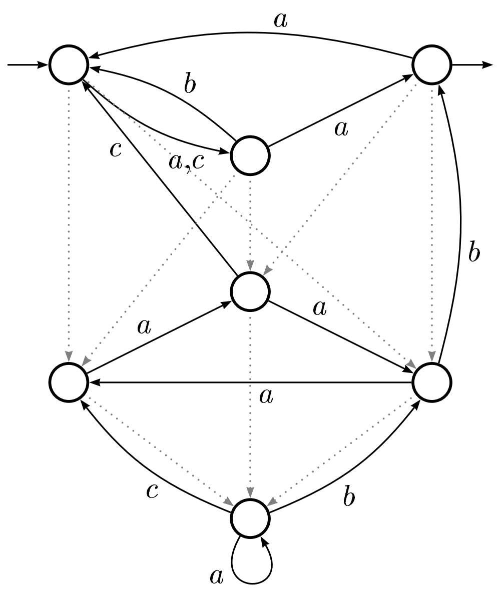


Fig. 15 – *Écorché de l'automate universel de $\mathcal{L}_4$.*

par a qui part de (resp. arrive dans) p.[4]

Démonstration. Soit $p = (L,R)$ un état de $\check{\mathcal{U}}_{\mathcal{L}}$ tel qu'il existe une transition sortant de p et étiquetée par a.

$$R' = \mathsf{max}\{Y \mid L.a.Y \subseteq \mathcal{L}\} \text{ et } L' = \mathsf{max}\{X \mid X.R' \subseteq \mathcal{L}\}.$$

Alors $q = (L',R')$ est bien une factorisation. Toute factorisation (L'',R'') telle que $L.a.R'' \subseteq \mathcal{L}$ vérifie $R'' \subseteq R'$, donc (p,a,q) est la seule transition étiquetée par a qui part de p. Donc, les transitions non spontanées de $\check{\mathcal{U}}_{\mathcal{L}}$ forment un automate déterministe, et, de même, co-déterministe. □

On va voir que retrouver l'automate universel d'un langage à partir de son écorché consiste à calculer la clôture des transitions par rapport aux transitions spontanées :

Proposition 2.13 *Soit $\mathcal{U}_{\mathcal{L}}$ l'automate universel d'un langage rationnel $\mathcal{L}$ et $\check{\mathcal{U}}_{\mathcal{L}}$ son écorché. Pour toute transition (p,a,q) de $\mathcal{U}_{\mathcal{L}}$, il existe deux états p' et q' et une transition (p',a,q') dans $\check{\mathcal{U}}_{\mathcal{L}}$ ainsi que des chaînes de transitions spontanées de p à p' et de q' à q. De même, pour tout état initial p de $\mathcal{U}_{\mathcal{L}}$ (resp. tout état final q de $\mathcal{U}_{\mathcal{L}}$), il existe une suite de transitions spontanées de p_i à p (resp de q à p_t) dans $\check{\mathcal{U}}_{\mathcal{L}}$, où p_i est la factorisation initiale et p_t la factorisation finale.*

Démonstration. Soit $p = (L,R)$ et $q = (L',R')$. Par définition de l'automate $\mathcal{U}_{\mathcal{L}}$, on a $L.a.R' \subseteq \mathcal{L}$. Soit (L_1,R_2) un couple maximal tel que $L \subseteq L_1$, $R' \subseteq L_2$ et $L_1.a.R_2$. Il existe une transition dans $\check{\mathcal{U}}_{\mathcal{L}}$ entre (L_1,R_1) et (L_2,R_2). D'autre part, comme $L \subseteq L_1$, il

4. S'il ne comporte pas de transition spontanée, un tel automate est dit réversible; nous reviendrons sur cette famille d'automate au chapitre suivant.

existe une chaîne de transitions spontanées entre p et (L_1,R_1) et, de même, une chaîne entre (L_2,R_2) et q.

Si $p = (L,R)$ est un état initial de $\mathcal{U}_{\mathcal{L}}$, et $p_i = (L_i,\mathcal{L})$ la factorisation initiale, L contient L_i, donc il existe une chaîne de transitions spontanées entre p_i et p. De même pour les états terminaux. □

— ∘ —

4 Développement d'un automate

4.1 Motivations et définitions

On peut généraliser la construction correspondant au théorème 2.1 pour n'importe quel automate qui accepte le langage :

PROPOSITION 2.14 *Soit $\mathcal{A} = \langle\, Q,A,E,I,T \,\rangle$ un automate qui accepte un langage rationnel $\mathcal{L}$. Soit $Q_d = \{Q \cdot u \mid u \in A^*\}$ les états du déterminisés de $\mathcal{A}$ et $Q_c = \{u \cdot Q \mid u \in A^*\}$ les états de son co-déterminisé.*
Soit S l'ensemble des couples maximaux (L,R) composés d'un sous-ensemble de Q_d et d'un sous-ensemble de Q_c tels que l'intersection de chaque élément de L avec chaque élément de R est non vide. Alors, l'automate universel de $\mathcal{U}_{\mathcal{L}}$ est isomorphe à $\langle\, S,A,F,J,U \,\rangle$, avec :

$$\begin{aligned} J =& \{(L,R) \mid I \in L\} \\ U =& \{(L,R) \mid \forall X \in L, X \cap T \neq \emptyset\} \\ F =& \{((L,R),a,(L',R')) \mid L \cdot a \subseteq L' \text{ et } \forall X \in L, X \cdot a \neq \emptyset\}. \end{aligned}$$

Démonstration. Soit $M_{\mathcal{A}}$ le monoïde de transition de $\mathcal{A}$ et φ le morphisme canonique de A^* dans $M_{\mathcal{A}}$. Les factorisations de $\mathcal{L}$ sont reconnues par $M_{\mathcal{A}}$, et, comme on l'a déjà dit, deux éléments α et β de $M_{\mathcal{A}}$ tels que $I\alpha = I\beta$ sont dans les images des mêmes facteurs gauches de $\mathcal{L}$. Ces images caractérisent donc les facteurs gauches; il s'agit des états du déterminisé de $\mathcal{A}$. De même, deux éléments α et β de $M_{\mathcal{A}}$ tels que $T\alpha^{-1} = T\beta^{-1}$ sont dans les images des mêmes facteurs droits. Les facteurs droits sont donc caractérisés par les états du co-déterminisé de $\mathcal{A}$. Pour qu'un couple de mots (u,v) forme une sous-factorisation du langage reconnu par $\mathcal{A}$, il faut que $I \cdot (u.v) \cap T \neq \emptyset$, c'est-à-dire que $I \cdot u \cap v \cdot T \neq \emptyset$. Il existe donc une bijection entre les factorisations du langage et l'ensemble S :

$$(L,R) \longmapsto (I \cdot L, R \cdot T)$$

Une factorisation représente un état initial de l'automate universel si et seulement si son facteur gauche contient le mot vide, c'est-à-dire si et seulement si l'élément de S correspondant contient l'élément I. De même, elle représente un état terminal si et seulement si son facteur gauche est inclus dans le langage, c'est-à-dire si et seulement si tout élément de l'élément de S correspondant a une intersection non vide avec T.

Il existe une transition étiqueté par a entre deux états de l'automate universel si et seulement si les facteurs gauches L et L' des factorisations correspondant à ces états respectent $L.a \subseteq L'$, c'est à dire si et seulement si, pour tout X dans $I \cdot L$, il existe Y dans $I \cdot L'$ tel que $X \cdot a = Y$. $\square$

On peut remarquer que, pour toute factorisation (L,R) du langage, pour tout mot u de L et tout mot v de A^*, si $I \cdot u \subseteq I \cdot v$, alors v appartient aussi à L. On peut donc représenter cette factorisation par l'ensemble $\mathsf{min}\{I \cdot u \mid u \in L\}$.

Chaque factorisation est ainsi caractérisée par une anti-chaîne de $\mathcal{P}(Q) \setminus \{\emptyset\}$.

Exemple 11 Supposons que l'ensemble des états d'un automate est $Q = \{p,q,r\}$. L'ensemble $\mathcal{P}(Q)$ est ordonné selon le schéma suivant :

$$\begin{tikzcd}
& \emptyset \arrow[dl, dotted] \arrow[d, dotted] \arrow[dr, dotted] & \\
\{p\} \arrow[d, dotted] \arrow[dr, dotted] & \{q\} \arrow[dl, dotted] \arrow[dr, dotted] & \{r\} \arrow[dl, dotted] \arrow[d, dotted] \\
\{p,q\} \arrow[dr, dotted] & \{p,r\} \arrow[d, dotted] & \{q,r\} \arrow[dl, dotted] \\
& \{p,q,r\} &
\end{tikzcd}$$

Les anti-chaînes de $\mathcal{P}(Q) \setminus \{\emptyset\}$ sont donc :

$$\begin{array}{ccc}
\{\{p\}\}, & \{\{q\}\}, & \{\{r\}\}, \\
\{\{p,q\}\}, & \{\{p,r\}\}, & \{\{q,r\}\}, \\
 & \{\{p,q,r\}\}, & \\
\{\{p\},\{q\}\}, & \{\{p\},\{r\}\}, & \{\{q\},\{r\}\}, \\
\{\{p\},\{q,r\}\}, & \{\{q\},\{p,r\}\}, & \{\{r\},\{p,q\}\}, \\
\{\{p,q\},\{p,r\}\}, & \{\{p,q\},\{q,r\}\}, & \{\{p,r\},\{q,r\}\}, \\
 & \{\{p\},\{q\},\{r\}\}, & \\
 & \{\{p,q\},\{p,r\},\{q,r\}\}, &
\end{array}$$

Évidemment, chaque anti-chaîne de $\mathcal{P}(Q)$ ne correspond pas nécessairement à un facteur gauche. Toutefois, on peut construire un automate dont l'ensemble des états est indexé par les anti-chaînes de $\mathcal{P}(Q)$. Celui-ci sera plus gros que l'automate universel (on verra qu'il le contient) et reconnaîtra le même langage que $\mathcal{A}$. Sa définition nous permettra de montrer qu'il hérite de certaines propriétés de $\mathcal{A}$.

— ∘ —

Construction du développé d'un automate

Définition 2.6 *Soit $\mathcal{A} = \langle\, Q{,}A{,}E{,}I{,}T\,\rangle$ un automate reconnaissant le langage rationnel $\mathcal{L}$. Soit S l'ensemble des anti-chaînes de $\{I \cdot u \mid u \in A^*\}$. Le* ***développé de l'automate*** *$\mathcal{A}$, noté $\mathcal{V}_{\mathcal{A}} = \langle\, S{,}A{,}F{,}J{,}U\,\rangle$, est défini par :*

$$\begin{aligned}
J =& \{P \in S \mid \exists Y \in P,\ Y \subseteq I\},\\
U =& \{P \in S \mid \forall X \in P,\ X \cap T \neq \emptyset\},\\
F =& \{(P{,}a{,}P') \in S \times A \times S \mid \forall X \in P,\ \exists Y \in P',\ Y \subseteq X \cdot a\}.
\end{aligned}$$

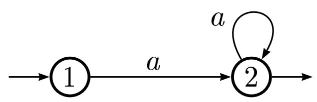
En fait, chaque anti-chaîne P représente l'ensemble des éléments minimaux d'un sous-ensemble $\overline{P}$ de $\mathcal{P}(Q)$ fermé supérieurement ($X \in \overline{P}$ et $X \subseteq Y \Rightarrow Y \in \overline{P}$). On peut alors réexprimer la définition du développé :

$$\begin{aligned}
J =& \{P \in S \mid \exists I \in \overline{P}\},\\
U =& \{P \in S \mid P \subseteq \overline{\{\{p\} \mid p \in T\}}\},\\
F =& \{(P{,}a{,}P') \in S \times A \times S \mid P \cdot a \subseteq \overline{P'}\}.
\end{aligned}$$

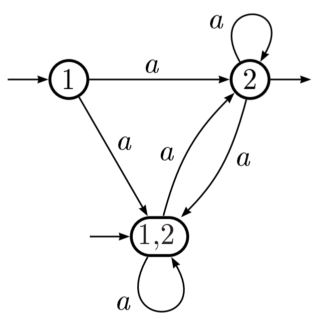
Remarque 2.6 On peut remarquer dès à présent que, d'une part, l'automate $\mathcal{V}_{\mathcal{A}}$ dépend fortement de l'automate $\mathcal{A}$ de départ et n'est donc en aucun cas un automate canonique du langage et que, d'autre part, la construction de l'automate $\mathcal{V}_{\mathcal{A}}$ est fortement orientée. On pourrait, en effet, construire un automate dual dont les états seraient les anti-chaînes des sous-ensembles obtenus à partir de T en faisant agir les lettres à gauche.

D'autre part, la structure de l'automate $\mathcal{V}_{\mathcal{A}}$ dépend en partie de la structure de $\mathcal{A}$. On peut par exemple montrer que $\mathcal{V}_{\mathcal{A}}$ contient le déterminisé de $\mathcal{A}$. (Il suffit de considérer le sous-automate de $\mathcal{V}_{\mathcal{A}}$ dont les états sont de la forme $\{I \cdot u\}$, pour u dans A^*.)

Exemple 8.5 La figure 16 présente l'automate minimal du langage $\mathcal{L}_2$ et son développé.



(a) L'automate minimal de $\mathcal{L}_2$.



(b) Le développé de l'automate minimal de $\mathcal{L}_2$.

Fig. 16 – *Deux automates reconnaissant $\mathcal{L}_2$.*

Exemple 1.12 Afin d'alléger les notations, on notera un sous-ensemble par un mot dont les lettres représentent les éléments du sous-ensemble. Par exemple, le mot pq représente l'ensemble $\{p{,}q\}$. L'automate $\mathcal{B}_1$ (figure 17) reconnaît le langage $\mathcal{L}_1 = A^*abA^*$. Calculons

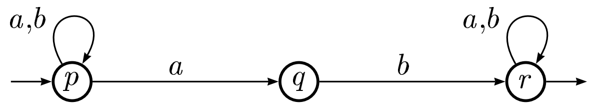


FIG. 17 – *L'automate $\mathcal{B}_1$ reconnaît $\mathcal{L}_1$.*

le développé de cet automate. L'ensemble $\{I \cdot u \mid u \in A^*\}$, qui est l'ensemble des états du déterminisé est

$R = \{p,pq,pr,pqr\}$. La table de transition du déterminisé est la suivante :

$$\begin{array}{c|c|c|c|c|} & p & pq & pr & pqr \\ \hline a & pq & pq & pqr & pqr \\ b & p & pr & pr & pr \end{array} \qquad (2)$$

Les anti-chaînes de R sont $S = \{\{p\},\{pq\},\{pr\},\{pqr\},\{pq,pr\}\}$. La définition du développé et la table de transition du déterminisé permettent de dire, par exemple : « Il y a une transition étiquetée par b de $\{pq\}$ vers n'importe quel état qui contient un ensemble plus petit que pr ». On a la matrice de transition suivante :

	p	pq	pr	pqr	pq,pr
p	a,b	a			a
pq	a,b	a	b		a,b
pr	a,b	a	a,b	a	a,b
pqr	a,b	a	a,b	a	a,b
pq,pr	a,b	a	b		a,b

De plus, l'état initial est $\{p\}$; les états finals sont $\{pqr\}$ et $\{pr\}$. (Par exemple, $\{pq,pr\}$ n'est pas final car l'intersection de pq avec T est vide.)

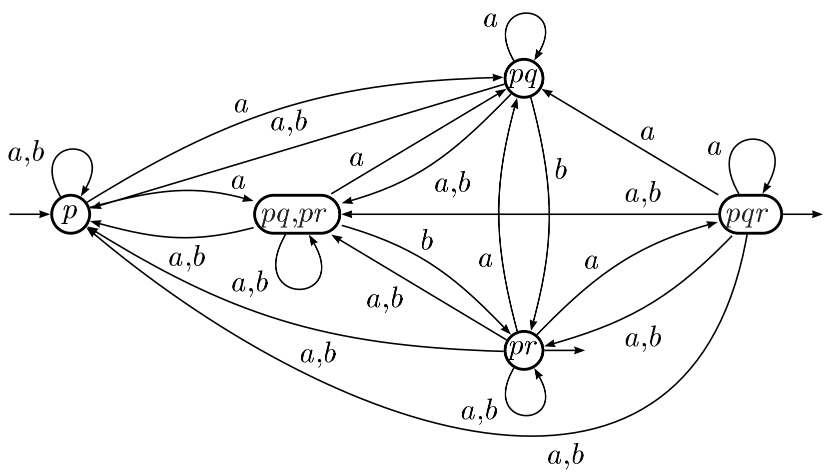


FIG. 18 – *Le développé de l'automate $\mathcal{B}_1$.*

On peut étendre la condition d'existence d'une transition aux chemins :

LEMME 2.15 *Soit P et P' deux états de $\mathcal{V}_{\mathcal{A}}$ et u un mot de A^*. Il existe un chemin étiqueté par u entre P et P' si et seulement si, pour tout X de P, il existe Y dans P' tel que $Y \subseteq X \cdot u$.*

Démonstration. La preuve est par récurrence sur la longueur de u.
Pour $|u| = 0$, la proposition est vraie (pour tout état p, $p \cdot 1_{A^*} = p$ est l'identité).
S'il existe un chemin entre P et P' étiqueté par $u = a.u'$, il existe un état P_1 tel que (P,a,P_1) est une transition et il existe un chemin étiqueté par u' entre P_1 et P'. Donc, pour tout X dans P, il existe Z dans P_1 et Y dans P' tels que $Y \subseteq Z \cdot u' \subseteq (X \cdot a) \cdot u' = X \cdot u$.

Réciproquement, si, pour tout X dans P, il existe Y dans P' tel que $Y \subseteq X \cdot u$, avec $u = a.u'$, on pose P_1 ensemble des états minimaux de $\{X \cdot a \mid X \in P\}$, P_1 est un état de $\mathcal{V}_{\mathcal{A}}$, (P,a,P_1) est une transition de $\mathcal{V}_{\mathcal{A}}$ et, quel que soit Z dans P_1, il existe X dans P et Y dans P' tel que $Z = X \cdot a$ et $Y \subseteq X \cdot (a.u') \subseteq Z \cdot u'$. Donc il existe un chemin entre P_1 et P' étiqueté par u'. □

— ◦ —

4.2 Propriétés de l'automate développé

Le lemme 2.15 nous permet d'exprimer le passé et le futur d'un état de $\mathcal{V}_{\mathcal{A}}$ de façon simple.

PROPOSITION 2.16 *Soit P un état de $\mathcal{V}_{\mathcal{A}}$ et u dans A^*.
u appartient à $\mathsf{Past}_{\mathcal{V}_{\mathcal{A}}}(P)$ si et seulement si il existe Y dans P inclus dans $I \cdot u$.
u appartient à $\mathsf{Fut}_{\mathcal{V}_{\mathcal{A}}}(P)$ si et seulement si, pour tout X dans P, $(X \cdot u) \cap T$ est non vide.
Ce qui peut s'écrire:*

$$\mathsf{Past}_{\mathcal{V}_{\mathcal{A}}}(P) = \bigcup_{Y \in P} \bigcap_{p \in Y} \mathsf{Past}_{\mathcal{A}}(p), \qquad \mathsf{Fut}_{\mathcal{V}_{\mathcal{A}}}(P) = \bigcap_{X \in P} \bigcup_{p \in X} \mathsf{Fut}_{\mathcal{A}}(p).$$

Démonstration. Si u appartient à $\mathsf{Past}_{\mathcal{V}_{\mathcal{A}}}(P)$, il existe un chemin étiqueté par u entre un état initial P_0 et P; il existe un élément X de P_0 inclus dans I et d'après le lemme 2.15, il existe Y dans P tel que $Y \subseteq X \cdot u \subseteq I \cdot u$.

Réciproquement, si u est tel qu'il existe Y dans P tel que $Y \subseteq I \cdot u$, il existe un chemin entre $\{I\}$, qui est initial, et P, donc u appartient à $\mathsf{Past}_{\mathcal{V}_{\mathcal{A}}}(P)$.

Si u appartient à $\mathsf{Fut}_{\mathcal{V}_{\mathcal{A}}}(P)$, il existe un chemin étiqueté par u entre P et un état final P_1. Donc, pour tout X dans P, il existe Y dans P_1 tel que $Y \subseteq X \cdot u$; comme $Y \cap T \neq 0_{\mathbb{B}}$, on obtinet $(X \cdot u) \cap T \neq 0_{\mathbb{B}}$,

Réciproquement, si u est tel que pour tout X dans P, $(X \cdot u) \cap T \neq 0_{\mathbb{B}}$, il existe clairement un chemin entre P et l'état correspondant aux éléments minimaux de $\{X \cdot u \mid X \in P\}$, qui est terminal; donc u appartient à $\mathsf{Fut}_{\mathcal{V}_{\mathcal{A}}}(P)$. □

Les propriétés suivantes de $\mathcal{V}_\mathcal{A}$ découlent directement de ces deux formules.

PROPOSITION 2.17 *Soit $\mathcal{A}$ un automate. L'automate $\mathcal{V}_\mathcal{A}$ reconnaît le même langage que l'automate $\mathcal{A}$.*

Démonstration. Considérons l'état $\{I\}$ qui est initial. D'après la proposition 2.16,

$$\mathsf{Fut}_{\mathcal{V}_\mathcal{A}}(\{I\}) = \bigcup_{p \in I} \mathsf{Fut}_\mathcal{A}(p) = \mathcal{L}(\mathcal{A}).$$

Donc l'automate $\mathcal{V}_\mathcal{A}$ accepte tous les mots du langage $\mathcal{L}(\mathcal{A})$.
Réciproquement, quel que soit l'état terminal P de $\mathcal{V}_\mathcal{A}$,

$$\begin{aligned}\mathsf{Past}_{\mathcal{V}_\mathcal{A}}(P) &= \bigcup_{X \in P} \bigcap_{p \in X} \mathsf{Past}_\mathcal{A}(p) \\ &\subseteq \bigcup_{X \in P} \bigcup_{p \in X \cap T} \mathsf{Past}_\mathcal{A}(p) \\ &\subseteq \bigcup_{p \in T} \mathsf{Past}_\mathcal{A}(p) = \mathcal{L}(\mathcal{A})\end{aligned}$$

Donc tous les mots acceptés par $\mathcal{V}_\mathcal{A}$ appartiennent à $\mathcal{L}(\mathcal{A})$. □

PROPOSITION 2.18 *Soit $\mathcal{A}$ un automate. Il existe un morphisme injectif de l'automate universel $\mathcal{U}_\mathcal{L}$ du langage $\mathcal{L}(\mathcal{A})$ dans l'automate $\mathcal{V}_\mathcal{A}$, défini par :*

$$\begin{aligned}\mathcal{U}_\mathcal{L} &\longrightarrow \mathcal{V}_\mathcal{A} \\ (L,R) &\longmapsto P_L = \mathsf{min}\{I \cdot u \mid u \in L\}.\end{aligned}$$

L'image inverse de ce morphisme est donné par les égalitées :

$$L = \mathsf{Past}_{\mathcal{V}_\mathcal{A}}(P_L)\ , \quad R = \mathsf{Fut}_{\mathcal{V}_\mathcal{A}}(P_L).$$

Démonstration. Soit (L,R) une factorisation de $\mathcal{L}$ et P_L l'ensemble des éléments minimaux de $\{I \cdot u \mid u \in L\}$.
D'après la proposition 2.16,

$$\begin{aligned}\mathsf{Past}_{\mathcal{V}_\mathcal{A}}(P_L) &= \bigcup_{X \in P_L} \bigcap_{p \in X} \mathsf{Past}_\mathcal{A}(p) \\ &= \bigcup_{u \in L} \bigcap_{p \in I \cdot u} \mathsf{Past}_\mathcal{A}(p) \\ &\supseteq \bigcup_{u \in L} u = L\end{aligned} \quad , \quad \begin{aligned}\mathsf{Fut}_{\mathcal{V}_\mathcal{A}}(P_L) &= \bigcap_{X \in P_L} \bigcup_{p \in X} \mathsf{Fut}_\mathcal{A}(p) \\ &= \bigcap_{u \in L} \bigcup_{p \in I \cdot u} \mathsf{Fut}_\mathcal{A}(p) \\ &\supseteq \bigcap_{u \in L} \{v \mid u.v \in \mathcal{L}\} = R\end{aligned}$$

Comme (L,R) est une factorisation (donc maximale) de $\mathcal{L}$, on obtient $L = \mathsf{Past}_{\mathcal{V}_\mathcal{A}}(P_L)$ et $R = \mathsf{Fut}_{\mathcal{V}_\mathcal{A}}(P_L)$.
Il existe donc une application injective des états de l'automate universel dans ceux de $\mathcal{V}_\mathcal{A}$.
Il reste à montrer que cette application est un morphisme.

Si (L,R) est initial ou final respectivement, l'état P_L l'est aussi.

Soit (L,R) et (L',R') deux factorisations telles que $L.a \subseteq L'$. Donc quel que soit u dans L, $u.a$ appartient à L'; donc quel que soit X dans P_L, il existe Y dans $P_{L'}$ tel que Y est inclus dans $X \cdot a$. Il y a donc une transition entre P_L et $P_{L'}$ étiquetée par a. □

Cette proposition peut se réexprimer de la façon suivante:

Corollaire 2.19 *Soit $\mathcal{A}$ un automate. L'automate $\mathcal{V}_{\mathcal{A}}$ contient l'automate universel du langage reconnu par $\mathcal{A}$.*

Non seulement ce résultat justifie la définition de $\mathcal{V}_{\mathcal{A}}$ comme « approximation » de l'automate universel, mais elle permet de voir les états de l'automate universel comme des parties de $\{I \cdot u \mid u \in \mathcal{L}\}$, ce qui donne un point d'ancrage supplémentaire pour travailler avec l'automate universel.

Remarque 2.7 La taille du développé d'un automate est difficile à évaluer. Le nombre de ses états est le nombre d'anti-chaînes de $\mathcal{P}(Q)$, où Q est un ensemble de n états. Avoir une expression close pour exprimer cette quantité en fonction de n est un problème ouvert connu sous le nom de problème de Dedekind. On trouvera en annexe un bref rappel de ce qui est connu sur ces nombres.

— ◦ —

4.3 Écorché du développé

De même qu'on a défini l'écorché de l'automate universel, on peut définir l'écorché du développé d'un automate. Comme la définition du développé est orientée, la définition de son écorché le sera aussi. Ainsi, les transitions non spontanées de l'écorché de l'automate universel formaient un sous-automate réversible (au plus une transition entrante et une transition sortante pour chaque lettre), alors que ces transitions formeront un sous-automate déterministe de l'écorché du développé.

Définition 2.7 *Soit $\mathcal{A} = \langle\, Q,A,E,I,T \,\rangle$ un automate reconnaissant le langage rationnel $\mathcal{L}$. Soit S l'ensemble des anti-chaînes de $\{I \cdot u \mid u \in A^*\}$. On ordonne partiellement les élément de S :*

$$P,P' \in S,\ P \preccurlyeq P' \Leftrightarrow \overline{P} \subseteq \overline{P'}.$$

*L'**écorché du développé** de l'automate $\mathcal{A}$, noté $\check{\mathcal{V}}_{\mathcal{A}} = \langle\, S,A,F,J,U \,\rangle$, est défini par :*

$$\begin{aligned}
J =& \{I\},\\
U =& \min_{\preccurlyeq}\{P \in S \mid \forall x \in P, x \cap T \neq \emptyset\},\\
F =& \{(P,a,P') \in S \times A \times S \mid P' = \min_{\subseteq}\{X \cdot a \mid X \in P\}\}\\
& \cup\{(P,1_{A^*},P') \in S \times A \times S \mid P' \in \min_{\preccurlyeq}\{R \mid P \prec R\}\}.
\end{aligned}$$

Remarque 2.8 Deux remarques sur cette définition :

i) Noter que selon l'ordre défini sur S, les singletons sont des éléments maximaux.

ii) T' n'appartient pas nécessairement à S; si c'est le cas, $U = \{T'\}$.

Exemple 1.13 En appliquant la définition et en se basant sur la table 2, on calcule l'écorché du développé de l'automate $\mathcal{B}_1$:

	p	pq	pr	pqr	pq,pr
p	b	a			
pq		a	b		1_{A^*}
pr			b	a	1_{A^*}
pqr		1_{A^*}	b,1_{A^*}	a	
pq,pr	1_{A^*}	a	b		

L'état initial est $\{p\}$; l'état final est $\{pr\}$.

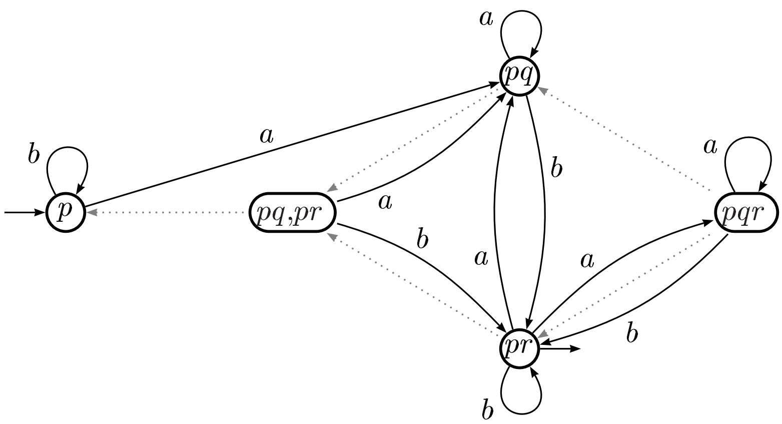


Fig. 19 – *L'écorché du développé de $\mathcal{B}_1$.*

Exemple 12.1 On considère l'automate $\mathcal{A}_{r_1}$ présenté figure 20.

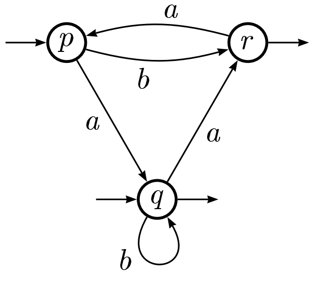


Fig. 20 – *L'automate $\mathcal{A}_{r_1}$*

La table de transition de l'écorché du développé de $\mathcal{A}_{r_1}$ (sans les transitions spontanées) est donnée par le tableau 3.

$$\begin{array}{c|c|c|c|c}
\text{États} & a & b & I & T \\
\hline
p & q & r & 0 & 0 \\
q & r & q & 0 & 0 \\
r & p & - & 0 & 0 \\
\hline
\{p,q\} & \{q,r\} & \{q,r\} & 0 & 0 \\
\{q,r\} & \{p,r\} & - & 0 & 1 \\
\{p,r\} & \{p,q\} & - & 0 & 0 \\
\hline
\{p,q,r\} & \{p,q,r\} & - & 0 & 0 \\
\hline
pq & qr & qr & 1 & 0 \\
qr & pr & q & 0 & 0 \\
pr & pq & r & 0 & 0 \\
\hline
\{p,qr\} & \{q,pr\} & \{q,r\} & 0 & 0 \\
\{q,pr\} & \{r,pq\} & \{q,r\} & 0 & 0 \\
\{r,pq\} & \{p,qr\} & - & 0 & 0 \\
\hline
\{pq,pr\} & \{pq,qr\} & r & 0 & 0 \\
\{pq,qr\} & \{pr,qr\} & q & 0 & 0 \\
\{pr,qr\} & \{pq,pr\} & \{q,r\} & 0 & 0 \\
\hline
\{pq,pr,qr\} & \{pq,pr,qr\} & \{q,r\} & 0 & 0 \\
\hline
\end{array} \tag{3}$$

L'automate obtenu est présenté figure 21. On peut remarquer que l'état $\{p,q,r\}$ n'est pas co-accessible. Il ne l'est donc pas non plus dans le développé. Cet état n'appartiendra donc pas à l'automate universel.

De même que pour l'automate universel, on peut retrouver le développé d'un automate à partir de son écorché en calculant la clôture des transitions par rapport aux transitions spontanées.

Proposition 2.20 *Soit $\mathcal{A}$ un automate, $\mathcal{V}_{\mathcal{A}}$ son développé et $\check{\mathcal{V}}_{\mathcal{A}}$ son écorché. Pour toute transition (p,a,q) de $\mathcal{V}_{\mathcal{A}}$, il existe un état q' et une transition (p,a,q') dans $\check{\mathcal{V}}_{\mathcal{A}}$ ainsi qu'une chaîne de transitions spontanées de q' à q.*
De même, pour tout état initial p de $\mathcal{V}_{\mathcal{A}}$ (resp. tout état final q de $\mathcal{V}_{\mathcal{A}}$), il existe une suite de transitions spontanées de l'état initial à p (resp. de p à un état final) dans $\check{\mathcal{V}}_{\mathcal{A}}$.

Démonstration. Soit (p,a,q) une transition de $\mathcal{V}_{\mathcal{A}}$. Par définition, pour tout X dans p, il existe Y dans q tel que $Y \subseteq X \cdot a$. Soit $q' = \mathsf{min}\{X \cdot a \mid X \in p\}$. Alors, (p,a,q') est une transition de $\check{\mathcal{V}}_{\mathcal{A}}$ et $q \preccurlyeq q'$, donc il existe une chaîne de transitions spontanées de q' à q.
Si p est un état initial de $\mathcal{V}_{\mathcal{A}}$, alors $I \in p$, donc $p \preccurlyeq \{I\}$ et il existe une chaîne de transitions spontanées de $\{I\}$ à p.
Si p est un état final de $\mathcal{V}_{\mathcal{A}}$, alors il existe un état final p' de $\check{\mathcal{V}}_{\mathcal{A}}$ tel que $p' \preccurlyeq p$ et il existe une chaîne de transitions spontanées de p à p'. □

— ◦ —

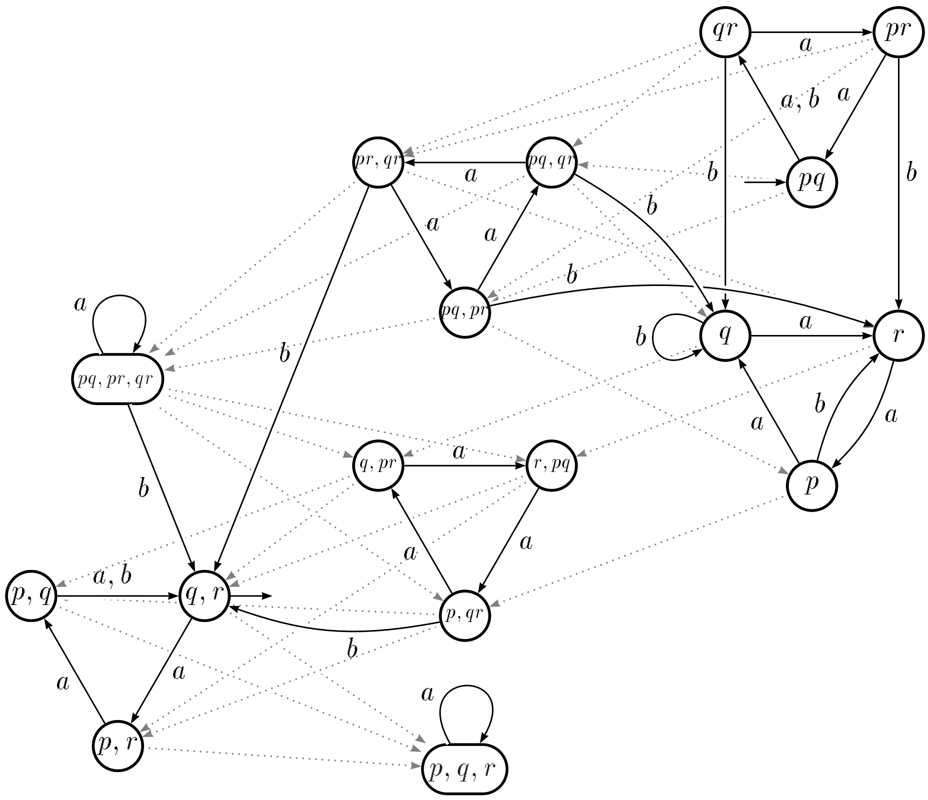


Fig. 21 – *Écorché du développé de* $\mathcal{A}_{r_1}$.

Annexe A

Le problème de Dedekind, posé en 1897, est généralement exprimé en termes de fonctions booléennes monotones, c'est-dire de fonctions réalisées uniquement avec les opérateurs logiques Et et Ou.

Le nombre de Dedekind $M(n)$ est le nombre de fonctions booléennes monotones à n variables distinctes.

On ne connaît pas de formule close pour exprimer $M(n)$. Les valeurs connues à ce jours [61] sont :

n	$M(n)$
0	2
1	3
2	6
3	20
4	168
5	7581
6	7828354
7	2414682040998
8	56130437228687557907788

Une fonction booléenne monotone est réalisée par une formule qu'on peut mettre sous forme disjonctive. Les constantes (Vrai et Faux) étant absorbantes ou élément neutre pour les opérateurs Et et Ou, les fonctions monotones qu'on peut réaliser par des formules non vides sont les fonctions monotones non constantes.

Chaque clause de la formule en forme normale disjonctive est formée d'un ensemble de littéraux; si les ensembles de littéraux de deux clauses sont inclus l'un dans l'autre, la clause ayant le plus grand de ces ensembles est inutile (si la plus petite est vraie, la formule aussi, si elle est fausse, la plus grande l'est aussi). Chaque formule est donc caractérisée par une anti-chaîne de sous-ensembles de l'ensemble des littéraux. On montre que deux anti-chaînes différentes correspondent à des fontions distinctes. Le nombre d'anti-chaînes (non vides) d'un ensemble à n éléments est donc $M(n) - 2$.

Table des exemples

Index

Bibliographie

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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