Khyra Gérardin

Automates. Séries rationnelles

Séminaire Delange-Pisot-Poitou. Théorie des nombres, tome 16, nº 2 (1974-1975), exp. nº G4, p. G1-G5

<http://www.numdam.org/item?id=SDPP_1974-1975__16_2_A7_0>

Séminaire DELANGE-PISOT-POITOU
(Groupe d'étude de Théorie des nombres)
16e année, 1974/75, n° G4, 5 p.
2 décembre 1974

AUTOMATES. SÉRIES RATIONNELLES

par Khyra GÉRARDIN

1. Définition d'un automate : Rappels.

Un automate $\mathfrak{A}$ est défini par la donnée de trois ensembles U, X, N, et de deux applications s, t :

U est l'ensemble des états.

X est l'ensemble des entrées.

N est l'ensemble des sorties.

t, une application de $U \times X \to U$, est dite application de transfert.

s, une application de $U \times X \to N$, est dite application de sortie.

Fonctionnement d'un automate. - Un automate fonctionne de la manière suivante :

Si R est une bande divisée en cases (R pouvant éventuellement être infinie), U un organisme pouvant changer d'état en se déplaçant vers la droite sur R, et lisant sur R une suite de symboles de X, lorsque U arrive sur R dans l'état a, devant la case i de R il lit la lettre x, et il écrit $s(a, x)$ dans cette case i. Il se déplace vers la droite en changeant d'état, et il se met dans l'état $t(a, x)$. Si la $(i+1)$-ième case de R n'est pas vide, le processus recommence avec l'état $t(a, x)$. Si la $(i+1)$-ième case est vide, l'automate s'arrête.

Automates dynamiques. - On notera X^* le monoïde libre engendré par X, N^* celui engendré par N.

On prolonge l'application $t : U \times X \to U$ à $U \times X^* \to U$ de la manière suivante : On notera T ce prolongement.

Si $x \in X$, $T(a, x) = t(a, x)$, $\forall a \in U$, $\forall x \in X$.

Sinon, on définit T par récurrence :

$$T(a, xp) = T(t(a, x), p), \quad a \in U, \quad x \in X, \quad p \in X^* .$$

$$T(a, e) = a, \quad \forall a \in U, \quad e \text{ étant le mot vide de } X^* .$$

De la même manière, on prolonge l'application $s : U \times X \to N$ à $U \times X \to N^*$: on notera S ce prolongement.

$S(a, x) = s(a, x)$, $\forall a \in U$, $x \in X$.

$S(a, xp) = s(a, x).S(t(a, x), p)$, $\forall a \in U$, $x \in X$, $p \in X^*$.

$S(a, e) = e'$, e étant le mot vide de X^*, e' celui de N^*.

Définition d'une application séquentielle. - Etant donnés deux monoïdes libres X^*, N^*, on dira qu'une application $f : X^* \longrightarrow N^*$ est séquentielle, si on a les deux conditions suivantes :

1° $f(p)$ a même longueur que p, $\forall\, p \in X^*$.

2° $f(pq) = f(p).\Gamma$.

Relations d'équivalence associée à une application séquentielle. - On désignera par $\mathcal{R}$ la relation d'équivalence sur X^*, définie par :

$$(p\ \mathcal{R}\ q) \Longleftrightarrow (\text{dernière lettre de } f(p) = \text{dernière lettre de } f(q)) .$$

Pour un automate U, X, N, on définira une relation d'équivalence associée à l'état a de la manière suivante : Sur N^*,

$$(p\ \mathcal{R}_a\ q) \Longleftrightarrow (\text{dernière lettre de } S(a,p) = \text{dernière lettre de } S(a,q))$$

où a est bien entendu un état de l'automate, et S le prolongement de la fonction de sortie de l'automate.

Une relation d'équivalence $\mathcal{R}$ sur un monoïde libre N^* est finie si $N^*/\mathcal{R}$ est fini. Une relation d'équivalence finie, compatible à droite avec la concaténation, est dite représentable.

PROPOSITION 1. - Toute relation représentable est associée à l'état a d'un automate possédant un nombre fini d'états, et réciproquement.

COROLLAIRE. - Si L_1 et L_2 sont deux relations représentables, $L_1 \cap L_2$ est représentable. Toute relation d'équivalence moins fine qu'une relation représentable est représentable.

2. Parties régulières d'un monoïde libre.

Définition. - On appelle partie régulière d'un monoïde libre X^*, les classes d'équivalence des relations représentables.

Conséquences. - L'intersection, la réunion finies de parties régulières sont régulières. Le complémentaire d'une partie régulière est régulier. Les parties finies sont régulières.

THÉORÈME 1. - Pour qu'une partie A d'un monoïde libre X^* soit régulière, il faut et il suffit qu'il existe un monoïde fini M, un morphisme φ de X^* dans M, telle que A soit saturé pour φ, c'est-à-dire

$$A = \varphi^{-1}\,\varphi(A) .$$

Supposons que l'automate $\mathfrak{A}$ soit sans sortie, que l'ensemble des entiers X soit fini, l'ensemble des états U est fini. Dans ce cas, une partie régulière A de X^* est associée à un état q_0 de l'automate de la manière suivante :

$$A = \{f \in X^* \ ;\ T(a_0\,,\,f) \in F\}\ ,\ \text{où}\ F \subset U .$$

A est, par définition, le langage de l'automate $\mathcal{A}$ par rapport à l'état initial a_0 et au sous-ensemble F des états finaux.

PROPOSITION 2. - Etant donné un automate $\mathcal{A}$ à états finis sans sortie, le langage associé à un état initial a_0 , où $a_0 \in U$, est une partie régulière.

PROPOSITION 3. - Le produit, la réunion de deux parties régulières sont parties régulières. Si A^* désigne le monoïde engendré par A , alors A^* est une partie régulière si A l'est.

THEORÈME DE KLEENNE. - L'ensemble des parties régulières d'un monoïde libre X^* est le plus petit sous-ensemble contenant X , stable par réunion, itération, produit fini.

3. Définition d'une grammaire.

Définition. - Une grammaire contexte libre ou "context free" est définie par les données suivantes :

1° Un ensemble V fini, pointé, par un élément initial a_0 dit axiome de la grammaire,

2° Un ensemble fini de règles P ou productions

$$a_i \longrightarrow p_i \ , \text{ où } \ a_i \in V \ , \quad p_i \in V^* \ .$$

On notera U le sous-ensemble de V dont les éléments figurent dans au moins une production. U est le vocabulaire non terminal de la grammaire, X est le complémentaire de U dans V . X est le vocabulaire terminal. On suppose que $a_0 \in U$. On notera une grammaire

$$G = (U \ , X \ , P \ , a_0) \ .$$

On définit une relation dans V^* de la manière suivante :

Si $p \ , q \in V^*$, alors $p \overset{*}{\Longrightarrow} q$, si

$$p = u a_i \, v \ , \quad q = u p_i \, v \quad \text{et} \quad a_i \longrightarrow p_i \in P \ , \quad u \ , v \in V^* \ .$$

$p \Longrightarrow q$ s'il existe une suite finie $p_1 \, p_2 \dots p_n$ telle que

$$p \overset{*}{\Longrightarrow} p_1 \ , \quad p_i \overset{*}{\Longrightarrow} p_{i+1} \ , \quad p_n \overset{*}{\Longrightarrow} q \ .$$

Par définition, le langage $L(G)$ de la grammaire $G = (U \ , X \ , P \ , a_0)$ est l'ensemble des mots f de X^* tel que $a_0 \Longrightarrow f$.

Définition. - Une grammaire est dite régulière si $G = (U \ , X \ , P \ , a_0)$, et si les productions ont l'une des deux formes suivantes :

$$a_i \longrightarrow p_j \ , \quad a_i \in U \ , \quad p_j \in X^*$$

ou

$$a_i \longrightarrow p_j \, a_k \ , \quad a_i \, a_k \in U \ , \quad p_j \in X^* \ .$$

PROPOSITION 4. - Les grammaires régulières définissent les parties régulières. Une partie régulière est appelée K-langage ou langage de Kleenne.

4. Liens. Séries rationnelles. Parties régulières. Automates.

Rappelons les notations :

X un ensemble fini, X^* le monoïde libre qu'il engendre, $\mathbb{Z}\langle\langle X\rangle\rangle$ l'algèbre large de X^* sur $\mathbb{Z}$, $\mathbb{Z}^{rat}\langle\langle X\rangle\rangle$ la sous-algèbre des séries rationnelles.

Si $a \in \mathbb{Z}\langle\langle X\rangle\rangle$, $a = \sum_{f\in X^*} (a , f)f$.

On note $\operatorname{Supp} a = \operatorname{Support} a = \{f \in X^* ; (a , f) \neq 0\}$.

a est une série positive, si $(a , f) \geq 0$, $\forall f \in X^*$.

La somme, le produit de deux séries positives sont séries positives.

Si $a , b \in \mathbb{Z}\langle\langle X\rangle\rangle$, et si a , b sont positives, alors, en notant $A = \operatorname{Supp} a$, $B = \operatorname{Supp} B$, il est facile de voir que

$A \cup B = \operatorname{Supp} (a + b)$,

$A \times B = \operatorname{Supp} (a \times b)$,

$A^* = \operatorname{Supp} a^*$, si a est quasi inversible,

$A \cap B =$ Support de $(a\odot b)$, où $a\odot b$ est le produit d'Hadamard de a et b.

PROPOSITION 5. - Les parties régulières d'un monoïde libre X^* sont les supports des séries rationnelles positives.

N.B. - De la même manière, on peut définir des séries algébriques, des grammaires "context free", des automates "push down" dans C-langages.

PROPOSITION 5'. - Les C-langages sont les supports des séries algébriques positives.

THÉORÈME. - L'intersection d'un K-langage et d'un C-langage sont un C-langage.

Ceci n'est autre que la traduction ensembliste du théorème de Jugen-Schützenberger [4] que nous rappelons ici.

THÉORÈME. - Le produit d'Hadamard d'une série rationnelle et d'une série algébrique est une série algébrique.

La partie concernant les séries algébriques n'a pas été exposée.

BIBLIOGRAPHIE

[1] FLIESS (Michel). - Sur certaines familles de séries formelles, Thèse Sc. math. Univ. Paris-7, 1972

[2] GÉRARDIN (Khyra). - Séries rationnelles, Séminaire Delange-Pisot-Poitou : Groupe d'étude de Théorie des nombres, 16e année, 1974/75, n° G1, 6 p.

[3] SCHÜTZENBERGER (M. P.). - On the definition of a famìlly of automata, Information and Control, t. 4, 1961, p. 245-270.

[4] SCHÜTZENBERGER (M. P.). - On a theorem of R. Jungen, Proc. of the American mathematical Society, t. 13, 1962, p. 885-890.

(Texte reçu le 6 février 1975)

Khyra GÉRARDIN
12 rue Beccaria
75012 PARIS