Schützenberger et la rationalité non commutative

Christophe Reutenauer

1996

Dans ses travaux sur le théorème de Kleene et les grammaires de Chomsky sur les langages formels et les automates, Schützenberger fit l'observation fondamentale qu'on se trouvait là en présence des concepts de rationalité et d'algébricité non commutatives.

Le théorème de Kleene affirme qu'un langage formel est reconnaissable par un automate fini si et seulement s'il est rgulier, c'est–dire si on peut l'obtenir à partir des langages finis par les opérations union, concaténation et l'opération étoile ("Kleene star operation", en fait le sous-monoïde engendré). Ce résultat fut généralisé par Schützenberger de la manière suivante: une série formelle en plusieurs variables non commutatives est reconnaissable (on dirait aujourd'hui: est une fonction représentative sur le monode libre) si et seulement si elle est rationnelle, c'est-à-dire si on l'obtient à partir des polynômes non commutatifs par les opérations algébriques (somme, produit) et l'inversion; autrement dit, l'algèbre des séries rationnelles est la clôture rationnelle de celle des polynômes non commutatifs. Plus généralement, l'inversion est remplacée par l'étoile, lorsqu'on ne dispose que d'un semi-anneau pour les coefficients, ce qui autorise de nombreuses applications et spécialisations, entre autres le théorème originel.

Il était là dans la non commutativité complète (pas d'identité polynomiale, ni de conditon à la Ore), et un apport immédiat de ces travaux à l'algèbre non commutative est l'observation fructueuse que pour construire des éléments rationnels, il faut inverser des matrices, plutôt que des éléments. Il étendit ainsi la théorie des automates en direction de l'arithmétique classique (une série rationnelle d'une variable n'est rien d'autre qu'une suite satisfaisant à une relation de récurrence linéaire) et des fonctions représentatives (il découle de ses travaux qu'un langage rationnel s'identifie à une fonction caractéristique représentative sur le monoïde libre).

De nombreux travaux de Schützenberger portent sur des raffinements du théorème de Kleene et des classifications des objets rationnels, langages ou séries. Dans une œuvre pléthorique, mentionnons quelques aspects:

1. L'élégant théorème d'apériodicité, qui est à l'origine de sa théorie des pseudo-variétés de monoïdes finis et de langages rationnels, en collaboration avec Eilenberg: l'absence de groupes dans le monoïde syntaxique du langage rationnel équivaut à l'expressibilité de celui-ci sans l'opération étoile, mais avec la complémentation.

2. L'étude du degré de croissance des séries rationnelles sur l'anneau des entiers l'a conduit à un classement très fin des monoïdes finiment engendrés de matrices sur cet anneau et il démontre l'égalité, dans le cas de la croissance polynomiale, du degré de croissance et de la longueur d'une suite à la Jordan-Hlder du monoïde; un corollaire en était la solution du problème de Burnside pour les matrices.

3. L'étude de la hauteur d'étoile des langages rationnels, essentiellement celui des fractions rationnelles non commutatives.

4. En collaboration avec Eilenberg, il refait toute la théorie des parties rationnelles des monodes abéliens.

Sortons un peu de la rationalité. Schützenberger a montré que les langages de Chomsky (context-free languages) et leurs généralisations en séries formelles sont en fait des éléments algébriques dans le cadre approprié. Il étendit un théorème de Jungen dans ce cadre: algébricité du produit de Hadamard d'une série formelle algébrique par une série formelle rationnelle. Ainsi la hiérarchie de Chomsky des langages formels (réguliers / hors contexte / contexte-sensitif / récursivement énumérables) s'apparentait à la classification des nombres en rationnels, algébriques et transcendants. Avec Chomsky, Schützenberger invente les automates à pile (une extension des automates finis): ceux-ci disposent d'une mémoire en forme de pile, i.e. à chaque instant, la machine n'a accès qu'au dernier symbole de la pile. Ils montrent que ces automates sont équivalentes aux grammaires hors-contexte, et qu'en un certain sens, les langages algébriques sont des spécialisations du langage de Dyck (celui-ci consiste en l'ensemble des expressions bien parenthésées), ce qui indique la pertinence de l'algébricité à propos de la compilation des langages de programmation.

La rationalité des langages reconnus par automates finis s'étend aux transducteurs, c'est-à-dire les automates finis avec sortie: le graphe des fonctions de mots calculées par ces machines est une partie rationnelle du monoïde produit. Dans cette direction, Schützenberger prouve que cette propriété est équivalente à la finitude du rang d'une matrice de Hankel canoniquement associée à la fonction. Il poursuit la classification des langages selon leur monoïde syntaxique dans la direction des transducteurs et des fonctions rationnelles de mots: croissance du cardinal de l'image, propriétés de continuité de la fonction vis-à-vis de diverses distances ayant un sens combinatoire (comme la distance préfixe dans les groupes hyperboliques), préservation des propriétés syntaxiques des langages sous ces fonctions. Une classe particulière, les fonctions sous-séquentielles, constitue le concept mathématique définitif de ce qu'on peut calculer sans mémoire de gauche à droite (comme la divison euclidienne).

Christophe Reutenauer