Marcel-Paul Schützenberger (1920-1996)

Marcel-Paul Schützenberger est mort le lundi 29 juillet 1996. C'était, pour reprendre l'expression d'André Lichnerowicz [1], un gentilhomme de la science, qui laisse en héritage une quantité de découvertes fondamentales et d'idées nouvelles. Il a travaillé durant sa vie dans un nombre incroyable de domaines différents et beaucoup, y compris parmi ses anciens élèves, ne connaissent qu'une partie de la fabuleuse diversité et de la profondeur des résultats de toutes sortes qu'il a pu obtenir. Je ne suis moi-même pas sûr de bien les connaître tous, comme par exemple sa participation à la découverte du gène de la trisomie au début des années 50.

Je vais essayer de retracer ici quelques éléments de son travail en informatique théorique. Une source d'autres informations est le volume *Mots* (Hermes, 1980). Il s'agit d'un recueil de textes en hommage à Marco Schützenberger préparé par ses amis et ses anciens élèves. On y trouve en particulier une liste de ses publications scientifiques (compilée par Imre Simon) complète jusqu'à la date de publication de l'ouvrage.

Le point de départ de son travail en informatique théorique est la théorie des codes à longueur variable. Il a présenté en 1955 au séminaire d'algèbre de Paul Dubreil un exposé intitulé 'Une théorie algébrique du codage' (*Séminaire Dubreil-Pisot*, année 1955-56, exposé n^o 15). Le texte contient déja un grand nombre des idées de son travail futur sur les automates. On y trouve en particulier la définition du semigroupe syntaxique d'une partie reconnaissable (le *'semigroupe fondamental'*). La preuve de l'égalité des ensembles rationnels et reconnaissables (ou, au moins, la partie essentielle qui est la cloture des ensembles reconnaissables par étoile) apparait là de façon presque simultanée avec la publication de l'article de Kleene (sur cette période héroïque, on peut consulter mon article 'Les débuts de la théorie des automates', *TSI*,**14**, 1995, 409-433). Il a certainement été le premier à voir que la propriété de décodage unique est de nature algébrique (en fait une propriété de base d'un monoïde libre, analogue à celle qui fait l'objet de la thérie de Nielsen-Schreier dans le groupe libre). Il a été fasciné dès le début par l'idée qu'il y avait quelque chose de plus à comprendre dans cette propriété et qu'on avait tort de se satisfaire du fait que l'algorithme de Sardinas et Paterson (sardinapoil pour les intimes, publié en 1953) permet de vérifier qu'un ensemble la possède. Il a aussi découvert dans ce domaine un jeu très riche de propriétés et de

[1] Hommage à M. P. Schützenberger, La Recherche, **280**, octobre 1996

points de vue qui s'articulent de façon étonnante. J'ai découvert moi-même, lorseque j'ai commençé à travailler sur les codes au début des années 70, les différents aspects algébriques (avec l'utilisation de semigroupes finis), combinatoires (avec la combinatoire des mots) ou probabilistes (liés à la théorie de l'information de Shannon). Les très nombreux résultats que Schützenberger a obtenus sur les codes ont fait l'objet de publications indépendantes et certains d'entre eux n'apparaissaient que dans des notes, comme celles du séminaire de Royan de 1965. L'ensemble constituait un corps assez difficile d'accès comportant, au dire d'Aldo De Luca et Antonio Restivo 'una tradizione orale'. La publication en 1984 de notre livre avec Jean Berstel ('Theory of Codes', Academic Press) a permis de donner une présentation plus accessible de l'ensemble de la théorie. L'ouvrage a, en fait, été préparé à l'initiative et avec la constante sollicitude de Marco qui, suivant son goût en la matière, a préféré que son nom ne figure pas parmi ceux des auteurs. Pour citer l'un des très beaux résultats de Schützenberger sur les codes, je choisirai celui-ci : *Un code maximal fini a un délai de déchiffrage qui est nul ou bien infini* (*J. Comb. Th.*, **1**, 1966, 437-442). Ce magnifique théorème avait été conjecturé par E. N. Gilbert et E. F. Moore dans leur article de 1959 (variable length binary encodings, *Bell System Tech. J.* **38**, 933-967). Il est, comme beaucoup de résultats combinatoires de nature extrêmale, susceptible de plusieurs preuves très différentes. L'une d'elles a d'ailleurs été découverte très récemment par Véronique Bruyère. Il peut aussi être interprété de diverses façons, en particulier comme un résultat sur l'optimalité des codes préfixes, qui sont ceux dont le délai de déchiffrage est nul. La théorie des codes à longueur variable reste aussi marquée par la conjecture de Schützenberger selon laquelle un code maximal fini est commutativement préfixe. Cette conjecture aparait déja dans le texte de Royan cité plus haut et a résisté jusqu'à présent à tous les efforts faits pour la démontrer, y compris ceux qui ont permis d'établir des théorèmes voisins, comme celui de Christophe Reutenauer. Les idées de Schützenberger sur les codes l'ont aussi conduit plus tard à développer une théorie des fonctions rationnelles sur les mots. Il s'agit de fonctions calculables par un automate fini, dont le décodage d'un code rationnel fournit un exemple typique. Ainsi, un des beaux outils qu'il a introduits sont les transducteurs non-ambigus, qui sont la généralisation naturelle des tranducteurs de décodage.

Marco Schützenberger a aussi commençé très tôt son célèbre travail sur les grammaires context-free. Le fameux théorème de Chomsky-Schützenberger, affirmant que tout langage context-free est un codage d'un langage de Dyck date de 1963. Son idée fondamentale est que les langages context-free sont une version non-commutative des séries algébriques, de la même façon que les langages rationnels sont la version non-commutative des séries rationnelles. Ainsi, pour donner un exemple, on écrira la grammaire des mots bien parenthésés

$$P \rightarrow (P)P + \epsilon$$

sous la forme de l'équation algébrique

$$\pi = a\pi b + 1$$

dont la solution, en variables commutatives, est la fonction algébrique

$$\pi = a \cfrac{1}{1 - a \cfrac{1}{1 - a \cfrac{1}{\ddots} b} b} b$$

Son travail sur les grammaires context-free est donc étroitement relié à celui sur les séries algébriques et rationnelles. C'est dans cet esprit qu'il a publié en 1962 un article dans lequel il montre que le produit de Hadamard d'une série algébrique et d'une série rationnelle non-commutatives est encore une série algébrique (On a theorem of R. Jungen, *Proc. Amer. Math. Soc.*, *13*, 885-890). Il s'agit en fait d'une version pour les séries de la propriété suivant laquelle l'intersection d'un langage context-free et d'un langage rationnel est encore context-free, qui aparait de la sorte comme une généralisation d'un résultat déja connu, et dû à R. Jungen, pour les séries en une variable.

Le magnifique théorème sur les ensembles apériodiques et sans-étoile est publié en 1965 (On finite monoids having only trivial subgroups, *Inf. and Control*, **8**, 190-194). Il caractérise les propriétés définissables par des opérations booléennes et le produit (et donc sans étoile) comme celles qui ne donnent pas lieu à une périodicité. Il s'agit en fait du premier résultat de la théorie des variétés qu'il allait développer avec Samuel Eilenberg et qui parut pour la première fois dans le second volume du livre d'Eilenberg (*Automata, Languages and Machines*, Academic Press, 1976).

La plupart du reste de l'oeuvre mathématique de M. P. Schützenberger est soit en algèbre, soit en combinatoire.

En algèbre, il a profondément influencé la théorie des semigroupes. Le nom de groupe de Schützenberger d'une $\mathcal{D}$-classe, ainsi que celui de représentation de Schützenberger, ont été choisis par A. H. Clifford et aparaissent dans le premier livre sur la théorie algébique des semigroupes (A. H. Clifford and G. B. Preston, *The algebraic Theory of Semigroups*, Amer. Math. Soc., 1961). Les techniques qu'il avait développé sur les semigroupes furent souvent son 'arme secrète' pour son travail sur les automates. Par exemple, la preuve du théorème mentionné plus haut sur les ensembles sans étoile utilise la théorie des semigroupes finis. Je ne connais d'ailleurs pas de preuve de ce résultat qui n'utilise pas les semigroupes (ce qui explique peut-être que le théorème ne soit pas aussi bien connu qu'on pourrait s'y attendre). Une autre partie du travail de Marco porte sur les algèbres de Lie. Il y a en fait, de nouveau, un lien avec les codes et les automates. De fait, l'un de ses plus beaux résultats relie les décompositions des algèbres de Lie et les factorisations des monoïdes libres (On a factorization of free monoids, *Proc. Amer. Math. Soc.*, **16**, 1965, 21-24).

Ses contributions en combinatoire sont probablement aussi importantes que celles d'informatique théorique, et Marco fut, aux dires de Herbert Wilf, 'l'un des combinatorialistes les plus créatifs et les plus importants de ce siècle' [2]. C'est probablement son travail sur les tableaux de Young qui est le plus célèbre. Une partie de ses premiers résultats est reproduite dans le volume 3 du livre de Donald Knuth (*Sorting and Searching*, Addison Wesley, 1975). Il n'était en fait pas si loin de son port d'attache, puisqu'il devait introduire dans ce domaine, dans son travail avec Alain Lascoux, un semigroupe (le *monoïde plaxique*) qui rend compte de nombre des propriétés des tableaux. On lui doit aussi l'idée des *preuves bijectives* consistant à rechercher une interprétation combinatoire naturelle des identités combinatoires en exhibant une bijection entre deux ensembles correspondant aux deux membres de l'identité. Les ensembles sont souvent des langages formels, comme dans le cas des énumérations de familles de graphes données par des langages context-free développées à sa suite par Robert Cori. .

Marco Schützenberger fut sans aucun doute le fondateur de la combinatoire des mots. Il a travaillé dans ce domaines à de très nombreuses reprises, en particulier avec André Lentin ou avec Roger Lyndon sur les équations en mots. Nombre des résultats sont présentés dans le volume collectif publié sous le pseudonyme de Lothaire en 1983 (*Combinatorics on Words*, Cambridge University Press). Marco avait l'habitude de me gratifier de titre de "protonotaire" de Lothaire. J'ai accepté longtemps ce titre mystérieux sans y penser mais j'ai eu récemment la curiosité de consulter le Robert. Voici le

[2] in *Electronic Journal of Combinatorics*, http://ejc.math.gatech.edu:8080/Journal

résultat:

> **PROTONOTAIRE** n. m. – 1390; lat. ecclés. *protonotarius*, de proto-, et lat. *notarius*. → Notaire
> ♢ **1.** Prélat de la cour romaine, du rang le plus élevé parmi ceux qui n'ont pas le caractère épiscopal. *Les protonotaires apostoliques participants* (de numero participiendum), *officiers de la cour pontificale, sont constitués en collège et sont chargés d'enregistrer les actes pontificaux dans les circonstances solennelles, à la congrégation des rites, de signer les bulles...–Protonotaires à l'instar* (ad instar participantium). *Protonotaires titulaires, honoraires* ou *protonotaires noirs* (dont l'habit ne comporte pas les ornements amarante des autres protonotaires).
> ♢ **2.** (1680). Hist. Premier notaire d'un empereur romain.
> (1869). Dignitaire laïc du moyen-âge (chef de la chancellerie, etc.)
> ♢ **3.** (Canada, 1795). Fonctionnaire chargé de l'enregistrement des actes dans un bureau régional.
> **DÉR. Protonotariat**

Une grande partie du travail de Marco en mathématiques est, comme nous l'avons vu, relié avec les automates, les mots ou les codes. Il serait certainement très difficile de dire si Marco Schützenberger a appliqué les mathématiques à l'informatique, ou si c'est plutôt l'inverse. Il m'a dit il y a longtemps qu'il considérait que son rôle était d'utiliser ses intuitions sur le traitement des données pour contribuer à enrichir les mathématiques d'objets et de propriétés nouvelles. Plaisantait-il? En fait, il avait dans toutes les matières des idées extrêmement originales et affirmées qui alimentaient une véritable passion pour la conversation, voire même les joutes oratoires. L'une des victimes fréquentes de ses sarcasmes était l'intelligence artificielle et tout ce qui s'en rapproche dans la croyance dans un modèle simpliste de la pensée et de la nature humaine, au point, disait-il, d'en arriver à oublier l'existence de différences entre les sexes.

Voilà pour les mathématiques. Ses contributions resteront vivantes très longtemps, même si le souvenir de son étonnante personnalité ne pourra survivre à ses nombreux amis et élèves. L'influence qu'il a eu sur chacun de nous dépasse de beaucoup notre éducation mathématique. Nous avons perdu un ami très attentionné, dont l'amour pour la discussion et les paradoxes nourrissait une nature merveilleusement généreuse.

Dominique Perrin
Paris, october 1996