An introduction to the scientific contribution of Marcel Paul Schützenberger

Robert Cori, Labri, Université de Bordeaux,
Dominique Perrin, IGM, Université Paris Est

August 22, 2019

1 Introduction

In march 2016 many former students of Marcel Paul Schützenberger gathered in Bordeaux, twenty years after his death, for a conference devoted to the scientific heritage of this prominent personality. This conference attracted many researchers from France, Belgium, Italy, United States, Israel, Chile, Canada who proposed at least 20 very different topics for which the contributions of M. P. Schützenberger (we will often refer to him as MPS) strongly influenced their scientific life. Also the participation of very young researchers showed how far the questions he proposed and and the ideas he developed are still relevant.

The list of publications of MPS contains about 300 papers appeared from 1943 to 2000, these were brought together and analyzed in an set of 12 volumes accessible on the site (where the publications are sorted by date).

`http://www-igm.univ-mlv.fr/ berstel/Mps/index.html.`

His first papers concern questions on the structure of lattices in algebra followed by an important number of papers applying statistical methods for Biological and Medical investigations. A remarkable property of Block designs was then his first contribution in Combinatorics. We will focus here on his work in Theoretical Computer Science and Combinatorics.

Each of the volumes of his collected papers begins with an introduction containing an abstract of the contents. These introductions were helpful for the writing of our text. We tried to make it accessible for those who are not specialists of subjects considered by giving many examples,and we hope that this will encourage many readers to read these publications.

One of the main conviction of MPS around 1960 was that the advent of the use of computers and the programming activity were just beginning and will widespread in the scientific community. It was then necessary to build mathematical tools as a model for these applications. Those mathematical models considered at that time were nor relevant for this task. They were suitable for application to Physics and Mechanics and not able to describe algorithms and programs. We can now realize how much premonitory was this observation as written by André Lichnerowicz: *Schützenberger avait le génie de la 'préscience' –du génie tout court : il a su dégager très tôt et comme par divination les fondements algébriques et combinatoires qui doivent sous-tendre tout le champ de l'informatique* (Schützenberger had the genius of 'premonition' –simply genius : he succeeded very early and by intuition to settle the algebraic and combinatorial foundations which were to underlay the whole field of computer science).

2 Words

Who could imagine, before the development of formal methods in syntax, that the mathematical structure whose basic object is a sequence of symbols and the main operation the concatenation of such sequences could rise difficult questions and lead to deep theorems used afterwards in many different fields?

The use of words in mathematics existed of course before the work of MPS, in particular in combinatorial group theory or with the remarkable contribution of Axel Thue. But he should be credited for identifying this structure as a field with rich and difficult problems, perhaps comparable to the beginnings of arithmetic.

The algorithmic questions raised by the problem of searching the occurrences of a given factor (that is a given subsequence of consecutive letters) in a word have also proved to form an ubiquitous problem in bio-informatics. The numerous algorithms presently used to solve this type of questions rely on constructions that he or his students have introduced.

A striking example of the influence that the work of MPS in combinatorics on words had later is *Dejean's conjecture*. In her thesis, prepared under the guidance of MPS, Françoise Dejean formulated a fascinating conjecture which was solved 40 years later after a sequence of efforts that we describe below as a testimony of the collective effort that was realized.

In 1906 and 1912, the Norvegian mathematician Axel Thue proved the existence of an infinite word on two letters not containing cubes (that is, factors repeated three times) and of an infinite word on three letters not containing squares (that is, factors repeated twice, see [20]). This type of question, considered by Thue as a purely intellectual challenge, happens to be frequently studied today in bioinformatics where squared factors, known as 'tandem repeats' play an important role for the analysis of molecular chains.

More generally, one may try to avoid fractional powers and not only integer ones. For this, one says that a word w has exponent p/q if $w = x^k y$ where $k \geq 1$ is an integer, y is a prefix of x and the length of w is p/q times the length of x. For example, the french word 'entente' has exponent 7/3 (with $x =$ 'ent', $y =$ 'e' and $k = 2$) and the english word 'entanglement' is a 4/3 power (with $x =$ 'entanglem', $y =$ 'ent' and $k = 1$). Dejean's conjecture is about the largest possible fractional number $e = \mathrm{RT}(k)$ such that every infinite word written on k letters has a factor of exponent e (the notation RT stands for 'repetition threshold').

The result of Thue can be reformulated as $\mathrm{RT}(2) = 2$. In fact every long enough binary word contains a square and there are infinite binary words (as the Thue-Morse word) without any factor of exponent > 2. In 1972, Françoise Dejean [13] proves that $\mathrm{RT}(3) = 7/4$. In other terms, on an alphabet with three letters, every long enough word (and Franoise Dejean verified this for every ternary word of length 39) contains a factor of exponent 7/4 or more, and there exists an infinite ternary word without any factor of exponent $> 7/4$. In the same work, she conjectures that $RT(4) = 7/5$ (result proved by Pansiot in 1984) and that $RT(k) = k/(k-1)$ for $k \geq 5$.

It is this last statement which is called the conjecture of Dejean, which was proved in 2009 and becomes now the theorem of Dejean. In summary, the repetition threshold takes the following values:

$$\begin{aligned} RT(2) &= 2, \\ RT(3) &= 7/4, \\ RT(4) &= 7/5, \\ RT(k) &= k/(k-1) \text{ for } k \geq 5. \end{aligned}$$

An infinite word which reaches this repetition threshold is called a word of Dejean. The first result, following that of Françoise Dejean, is due to Jean-Jacques Pansiot who showed that $RT(4) = 7/5$. Moulin Ollagnier proved the conjecture for $5 \leq k \leq 11$; Mohammad-Noori et Currie proved it for $12 \leq k \leq 14$. A real breakthrough occurred when Carpi proved the conjecture for $k \geq 33$. Currie et Rampersad later refined the construction of Carpi to extend it to $k \geq 27$. Even more recently, they used the technique of Moulin Ollagnier to solve the remaining cases $15 \leq k \leq 26$. By another technique, and almost simultaneously a proof of the remaining cases was given by Michal Rao[1].

3 Lyndon words and the theory of factorisations

One of the important notions concerning words is that of *Lyndon words*. Their definition is very simple: they are the words x such that any nontrivial cyclic shift of its letters gives a word which is larger in the alphabetic order. Thus, *abracedabra* is not a Lyndon word because placing its last letter in the first position gives the word *aabracadabr* which is both less easy to pronounce and smaller in the alphabetic order.

The notion was introduced by Roger Lyndon [21] and it follows from his results [7] that every word factorizes uniquely as a nonincreasing sequence of Lyndon words. Thus, the word *abracadabra* factorizes in *abracad abr a* where the three words *abracad*, *abr* and *a* are Lyndon words with $abracad' > abr < a$ in the lexicographic order.

This result was the starting point of the theory developed by MPS in 1965 [25] of *factorizations of free monoids*. Such a factorization is given by a totally ordered set X of words over a fixed alphabet A such that every word w on the alphabet A can be uniquely written

$$w = x_1 x_2 \cdots x_n$$

with $x_1 \geq x_2 \geq \ldots \geq x_n$. Thus the set of Lyndon words ordered by the alphabetic order provide an example of such a factorization. Many other

[1] The above paragraph is reproduced from `https://fr.wikipedia.org/wiki/Utilisateur:ManiacParisien/Brouillons/Math-12`. See also the blog of Jeff Shallitt `http://recursed.blogspot.fr/2009/05/dejeans-conjecture-solved.html`

examples had been studied independently , in particular in connection with the theory of Lie algebras (see the section Codes).

The main result of [25] states in particular that if X is a factorization of the free monoid on the alphabet A, then every word on A has a unique circular shift in the set X.

Factorizations of free monoids were later studied by Viennot [29] (see also [22] or [1]). An interesting application of Lyndon words is the Burrows-Wheeler transformation (see `https://en.wikipedia.org/wiki/Burrows%E2%80%93Wheeler_transform`).

4 Combinatorial algorithms

The *permutations* of a finite set have fascinated M. P. Schützenberger since these are central in Combinatorics and Computer Science. Indeed in the 60's it was often claimed that sorting some data occupied more than one half of the time of computers, so that it was important to determine the time complexity of the different sorting algorithms. One way to compute the average complexity of a sorting algorithm is to consider all permutations on n elements and to determine the sum of the number of elementary operations (namely comparisons and exchanges) performed by the algorithm sorting for each one of them. then to divided by the number of permutations (that is $n!$). In this analysis one uses some classical parameters on permutations like the number of left to right minima, the number of inversions, the descents, the lengths of increasing subsequences.

The main paper of MPS in this domain is the monograph he wrote in french with D. Foata on *Eulerian Polynomials* [16]. In this interesting text many of the classical results on these polynomials which were obtained by analytical computations are proved by combinatorial arguments. They concern mainly the enumeration of permutations with respect to some parameters. The methods developed use bijective proofs, so that this monograph may be considered as one of the first introduction for these new elegant proofs for which MPS and D. Foata were pioneers. The enumerative formulas of permutations were first obtained by celebrated mathematicians like Euler, Bernoulli, MacMahon or Stirling. The use of bijections allowed to avoid long analytic calculus.

Let us describe in a few words this method. It consists, in order to obtain an enumerative formula for the number of objects in a subset E, to find another subset F which one suspects to have the same number of elements as E, and for which a formula for the number of elements is known. The second step consists in building a bijection between the two sets. The building of the bijection is not always easy but it often gives a new insight on the structure of the objects in E. This technique may also be applied to obtain an efficient algorithm for random selection of an object in E by doing the selection in F when it is easier and then using the bijection to obtain the object in E.

Let us consider a central example of the bijections built by Foata and MPS. We take the one allowing to show that the number of permutations on n elements with k ascents is equal to those with k exceedances. These numbers are called the Eulerian numbers. A permutation α is a sequence

$$a_1, a_2, \ldots, a_n$$

where each number in $E_n = \{1, 2, \ldots, n\}$ appears exactly once.

An ascent in α is an i such that $a_i < a_{i+1}$ and an exceedance is an i such that $a_i > i$.

For example the permutation

$$\gamma = 7, 2, 5, 8, 6, 3, 9, 4, 1$$

has 3 ascents since $a_i < a_{i+1}$ for $i = 2, 3, 6$ and 5 exceedances since $a_i > i$ for $i = 1, 3, 4, 5, 7$

One may also consider a permutation as a mapping from E_n in into itself by denoting the image of i by $\alpha(i) = a_i$. In that case it could be represented by its cycles, that is the the circular lists $(i_1, i_2, \cdots, i_p)$ such that for $1 \leq j < p$, $\alpha(i_j) = i_{j+1}$ and $\alpha(i_p) = i_1$. So that we can write for our example

$$\gamma = (1, 7, 9)\ (2)\ (3, 5, 6)\ (4, 8)$$

The central bijection consists first in determining the left to right minima of α. A left to right minimum is an a_i such that $a_i < a_p$ for any $p < i$. The left to right minima of γ are $7, 2, 1$. Let $a_{i_1}, a_{i_2}, \ldots, a_{i_k}$ be the left to right minima of the permutation α. The subsequences beginning with a left to right minimum and ending just before the next one determine the cycles of a new permutation β whose number of exceedances is equal to the number of ascents in α. More precisely β is given by:

$$\beta(a_i) = \begin{cases} a_{i+1} & \text{if } i < n \text{ and } a_{i+1} \text{ is not a left to right minimum} \\ a_j = \min_{p \leq i} a_p & \text{otherwise} \end{cases}$$

The permutation δ constructed from γ is given by its cycles : (7), $(2, 5, 8, 6, 3, 9, 4)$, (1) and may be written as the sequence:

$$1, 5, 9, 2, 8, 3, 7, 6, 4$$

which has 3 exceedances.

The contributions of MPS in the domain of algorithms on permutations and their combinatorial structure is far from being limited to the Eulerian polynomials. Indeed in 1961 he is inspired by a publication of C. Schensted [23] which gives an algorithm to build the longest increasing subsequence of a sequence of integers. It uses a data structure which is now known as a Young Tableau. He examined the properties of this data structure and obtains remarkable results on Young tableau's. These are often quoted for instance by D. Knuth in The Art of Computer Programming (Vol. 3) [19]. This subject will be evoked in more details in the last section of our text.

The quest of bijective proofs for complicated enumerative formulas is a very vivid area in combinatorics since the work of MPS. Many researchers in combinatorics obtained very elegant bijective proofs for the enumeration of different families of combinatorial objects. This shed new light on the structures of the objects in the sets for which such bijections were found.

5 Automata

The idea of *mutiplicity* in formal models like automata or grammars is due to to Schützenberger [24]. The initial idea is to replace the Boolean notion of acceptance of an object (yes or no according to whether the object is recognized or not) by introducing the number of ways that the object is recognized.

This idea leads one naturally to that of automata or grammars with multiplicities. The notion of formal language is thus replaced by that of multiset of words, a function assigning to every word a multiplicity which can be an integer or a real number representing a probability.

The beauty in all this is that the notion of rational language will be replaced by that of rational series in noncommuting variables. The usual rational series in one variables will correspond to the case of an alphabet with just one letter. Boolean operations on languages will be replaced by algebraic operations (the sum replacing the union) and the star X^* by the operation of quasi inverse

$$(1 - X)^{-1} = 1 + X + X^2 + \dots$$

defined provided the multiplicity of the empty word in X is zero.

In the same way, the notion of context-free grammar will be replaced by systems of algebraic equations. Thus the grammar defining well parenthized words will write

$$X = (X)X + 1$$

instead of $X \to (X)|\varepsilon$ (see on this topic the section Formal Languages below).

On the other hand, the notion of finite automata is basically equivalent to that of linear representation of the free monoid, which leads to a generalization of Kleene's Theorem often called Kleene-Schützenberger Theorem.

This notion of noncommutative rationality leads also to the notion of inverse of matrices with formal series as coefficients. This is the basis of the principle of noncommutative algebra according to which, to build skew fields, one has to invert matrices and not only elements (as, in the commutative case, to build rational numbers starting from the integers). This principle was heavily used by Paul M. Cohn [11] (for the construction of the free field) and by his successors.

This point of view has been considerably developed since then. The theory of sets with multiplicity in a semiring was formulated in all its generality by Eilenberg in his book [14]. For an introduction to noncommutative rational series, one may refer to [2] and for algebraic ones to [15].

6 Semigroups

In the years following the last world war, the theory of semigroups developed, as part of the study of abstract algebraic structures. Although not always considered as a very important subject, especially in France where the Bourbaki school was leading the rules of good taste, it attracted a lot of research in the USA and independently in Russia (see the historical account of this curious situation in [18]). In the USA, it was pushed forward by the interest of John A. Clifford, a renown figure of algebra. In Russia the contribution of Suschkewitch came as early as 1928. In Europe, the main contribution was the thesis of J.A. Green in 1951, a student of Philip Hall in Cambridge.

Perhaps choosing a field not in the mainstream of research in France, Schützenberger, with a special inclination to singularity, chose to center his first contributions in algebra to lattices (1953) and semigroups (1959).

One of the first notable results in semigroup theory is the introduction by J.A. Green of several equivalence relations in a semigroup linked to its ideal structure. One of Green relations, denoted $\mathcal{H}$, has the property that if an $\mathcal{H}$-class contains an idempotent e (an element such that $e \cdot e = e$) then it is a group. The first result of Schützenberger was to generalize this result, showing that one may always associate a group to an $\mathcal{H}$-class, even if it does not contain an idempotent. At the same, he showed that one may associate to every $\mathcal{D}$-class (another of Green's relations) a representation of the semigroup by matrices with its nonzero elements in the group defined above.

This result was put in good place in the book of Clifford and Preston [10] with a section called the 'Schützenberger group' and another the 'Schützenberger representations'. The use of automata in computer science and the recent developments in modeling physical systems by discrete dynamical systems showed that the semigroup structure is central in the mathematics allowing to model such phenomena. Thus, the paper of Schützenberger of 1957 is quoted by Bond and Levine [3] as a reference to combinatorial models of physical systems with discrete parameters.

Approximately at the same time as Schützenberger was discovering these groups associated to a semigroup, the work of Kleene, followed by Rabin and Scott would create the theory of automata. Its cornerstone is Kleene's Theorem asserting that a language can be recognized by a finite automaton if and only if it can be obtained from the subsets of the alphabet by a finite number of unions, products and star (the star of X is the set $X \cup X^2 \cup X^3 \cup \ldots$).

The introduction of a semigroup associated to an automaton (its *transition semigroup*) is to be credited to Schützenberger. It allows to formulate properties of the automaton and the language that it recognizes by properties of this semigroup.

One of the most important results, next to Kleeene's Theorem, is the following: a language can be expressed by the operations of union, product and complement starting from finite languages (the *star-free* languages) if and only if the transition semigroup has only trival subgroups (the *aperiodic* languages). This result, proved in [26] is considered by Eilenberg as 'next to Klenne's Theorem, probably the most important result dealing with recognizable sets'.

It was the starting point of a great number of contributions where algebra and logic come into play. The link with logic is the starting point of a exciting travel. Finite automata allow one to define properties of words expressable in a logical language called the monadic second order logic of integers. In this logical language, quanitifiers may bear on second order variables provided they correspond to subsets (and not binary relations for example). One interprets the variables as integers with the relation $<$ and the relations $a(n)$ to express that the letter at position n is a. The equivalence between the two notions was established by Büchi [5] and it was shown by Robert McNaughton that star-free languages are those definable in the corresponding first-order logic. Thus one may define the language of words containing at least one b by the formula $\exists n b(n)$.

The theorem of MPS shows that one can decide if a rational language is definable at first order. For example, the language formed of words of even length is not aperiodic, thus not star-free, and thus not definable at first order. Waouh!

7 Codes

One of the favorite subjects for MPS was variable length codes. This problem appeared with the work of Shannon on information theory and concerns in its more elementary form the question of ambiguity or its contrary unambiguity. In fact, a code is a set X of words such that one cannot find a product $x_1 x_2 \cdots x_n$ of n words in X equal to another product $y_1 y_2 \cdots y_m$, unless $n = m$ and $x_1 = y_1, \ldots, x_n = y_n$. We are thus considering the unambiguity of the expression X^*. For example, $\{a, ba\}$ is a code but $\{a, ab, ba\}$ is not since $a(ba) = (ab)a$. As a very simple particular case, prefix codes are such that no word is a beginning of another one.

MPS formulated a number of conjectures concerning codes and in particular concerning maximal codes, which are such that one cannot add another word unless the result is not anymore a code. One may verify easily that a code formed of words on k letters is maximal by computing its Kraft sum $\sum_{n \geq 1} u_n k^{-n}$ (where u_n is the number of words of length n in the code) and checking that its value is 1. The main conjecture of MPS on maximal codes is still unsolved. It asserts that any finite maximal code may, by a rearrangement of the letters of its words, be transformed into a prefix code. Thus $X = \{aa, ab, aab, aba, bb\}$ is a maximal code (indeed, its Kraft sum is

$3 \times 1/4 + 2 \times 1/8 = 1$). It can be rearranged into a prefix code as below:

$$aa, ab, aab, aba, bb$$
$$aa, ab, baa, bab, bb.$$

(on this conjecture and what is known today, see [1]).

The notion of a code (and thus more generally that of unambiguity) is still full of mysteries. One may easily verify that a finite set of words is a code but there is no direct method producing them systematically. For example, one can prove that there is for each $n \geq 1$ a finite number of maximal codes having n elements on a given alphabet. But there is no known formula giving this number (for example using an induction on n). This is strikingly different in the case of prefix codes since the number of maximal prefix codes with n elements on a binary alphabet is the Catalan number

$$\frac{1}{n+1}\binom{2n+1}{n}$$

One of the essential contributions of MPS to this field is to have shown that the properties of codes were related to the field of noncommutative algebra. In particular, he showed that the property of synchronization, playing a role in the family of so-called *comma-free* codes, or more generally circular codes, were related to free Lie algebras. Circular codes are those for which the decoding can be realized in a unique way for a word written on a circle. Thus $\{ab, ba\}$ is a code but it is not a circular code since the word ab written on a circle can be decoded in two ways as (ab) or $a)(b$. And what is the free Lie algebra $L(A)$ on an alphabet A? It is obtained starting from the letters and forming expressions using ordinary sum but replacing the product by the *Lie bracket*

$$[xy] = xy - yx$$

Each of the terms obtained in this way has a degree. The degree of $[xy]$ is the sum of the degrees of x and y (the letters having degree 1). One obtains in this way a noncommutative algebra (one has even always $[xy] + [yx] = 0$) and non associative.

One of the results obtained by MPS [27], solving a conjecture of Golomb and Gordon [17], says that there exists a circular code having for every $n \leq N$, u_n words of length n if and only if the free Lie algebra $L(B)$ on an alphabet B with $\sum_{n \leq N} u_n$ elements can be embedded in $L(A)$ by an isomorphism sending u_n elements of B to monomials of degree n in $L(A)$. Thus there exists a circular code formed of one letter and two words of length 3 (like $\{b, aab, abb\}$) because with $B = \{u, v, w\}$, we embed $L(B)$ into $L(A)$ by the morphism $u \to a$, $v \to [a[ab]]$, $w \to [[ab]b]$.

This point of view has been substantially developed later by Viennot [29] and by Reutenauer [22].

8 Formal languages

In this part of his work, it is certainly the article he wrote with Noam Chomsky, entitled *The algebraic theory of context-free languages* [8] which is the most well known and the most frequently quoted. It has been translated in german and in french [9], also probably in other languages. This paper describes in a detailed manner the theory of context-free language from an algebraic point of view, this is somewhat different from the generative point of view which was mainly considered before. The authors suggest to work on the set of *formal power series* in non commuting variables.

The ambiguity of the grammar is then taken into account by the positive integer coefficient of a term of the series, this coefficient corresponds to the number of different ways it is generated by the corresponding grammar.

Indeed a context-free grammar was most often considered as a way to build words (or sentences in a natural language). In this paper the set of production rules are considered as algebraic equations, the variables of the grammar being unknowns. The solution of this system of equations is obtained by iterating the determination of an approximate result consisting of a finite set of words in the alphabet of terminal letters. This classical procedure in algebra is known as the determination of the *unique fixed point*for a system of equations.

This article contains many examples of closure properties of the family of context free languages are proved. It contains also the definition of *Dyck languages* and of local rational languages. The famous theorem stating that any context free language is the homomorphic image of the intersection of a Dyck language and a rational one is also a main result in this paper. Some summaries of many results obtained by MPS in previous articles are also presented.

The idea of considering a context-free language as a formal power series in non commuting variables proposed by Schützenberger introduced also a new method in Enumerative Combinatorics. He remarked that, when allowing the variables to commute in the series corresponding to the words of a non ambiguous context free language, one obtains a generating series enumerating these words with respect to the parameters corresponding to the number of occurrences of each letter of the alphabet.

Additionnally, when equating all the variables, the series obtained is the Taylor series of an algebraic function. One of the first results one can obtain using this method is to determine the number of Dyck words of a given length by solving a second degree equation (as we have already seen above in the section Automata).

Moreover the method proposed by MPS in order to enumerate some combinatorial objects consists by building a bijection form the set of these objects to the set of words generated by a non ambiguous context-free grammar, then to write the equation satisfied by the series corresponding to the grammar. One has to obtain the Taylor power series of the algebraic series satisfying the equation. This method turned out to be successful for the enumeration of many families of *planar maps* for which W. T. Tutte obtained nice enumerative formulas. See [28] for the generating functions and [12] for context-free grammars allowing to retrieve them .

One feeling of MPS was that planarity of maps and algebraicity of their generating power series were to facets of the same combinatorial properties. Some recents results have shown that this feeling was not correctly stated since some families of maps on the torus or on other surfaces with genus greater than 1 have also algebraic generating series. To make MPS intuition correct, one has to replace the constraint of planarity of the maps by that of having a genus less than a constant value (see [6]).

This method is in fact very efficient since many families of objects were enumerated in that way as it is illustrated in the paper of Mireille Bousquet-Mélou [4] and the book by Philippe Flajolet and Robert Sedgwick [15].

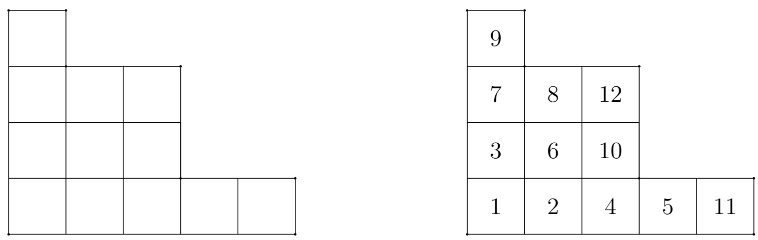
9 Algebraic combinatorics

Since 1977 the main interest of MPS was in the field of combinatorial problems related to the linear representation of finite groups. Alain Lascoux an algebraist geometer who worked among mathematicians in Grothendieck 's group, joined him in this direction. They both participated to an important movement including many other mathematicians like Adriano Garsia, Gian Carlo Rota, Richard Stanley,who created this new field of research mixing combinatorics and deep algebra now called Algebraic Combinatorics.

One of the first combinatorial objects on which MPS focused his work are the *Young Tableau's*, which we mentioned above as they are used in Schensted's algorithm for finding the longest increasing subsequence of a sequence of integers

It is worth noticing that this construction was also obtained by G. de B. Robinson (20 years before) in algebra for the linear representations of finite groups.

A *tableau* is a geometric figure associated to a partition of the integer n, this a weakly decreasing sequence $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)$, that is satisfying : $\lambda_i \leq \lambda_{i-1}$ for $1 < i \leq k$ and $\sum_{i=1}^{k} \lambda_i = n$. To obtain the tableau one draws k rows one above the other, each row i containing λ_i cells as in the figure below representing the partition $(5, 3, 3, 1)$ of 12:



A standard Young Tableau of shape λ is then a tableau of shape λ in which each cell contains one of the integers in $E_n = \{1, 2, \cdots, \}$, where the following conditions are satisfied. Any integer of E_n appears, those integers in the same row strictly increase from left to right and those in the same column strictly increase from down to up. An example of such a Young Tableau is given above.

Schensted's algorithm builds a Young Tableau from a permutation $\alpha = a_1, a_2, \cdots a_n$. It allows to determine the length its longest increasing subsequence and that of the longest decreasing one. This algorithm begins with a unique cell containing the value a_1. It uses an iteration by inserting one after the other the elements $a_2, \cdots, a_n$ of the permutation α in the Young Tableau, maintaining the increasing conditions on rows and columns. The lengths of the maximal increasing and decreasing subsequences of α are respectively the length of the first row and the number of rows of the Young Tableau associated to it.

Schützenberger proposed another algorithm which he called *"Jeu de taquin"* which led him to a new theory of these Tableaux by introducing a monoid defined by generators and some relations which may be interpreted as commutations This monoid was called the plactic monoid.

A few years later Schützenberger comes back to the domain of Young Tableaux by introducing the promotion operation on permutations which he illustrated by the Jeu de Taquin, now considered as a fruitful tool in algebraic combinatorics.

A remarkable property of Young Tableaux, which gave them a central role in algebra, is their relationship with a family of symmetric polynomials called the *Schur functions*. In order to define them one has to introduce non standard Young Tableaux by allowing the integers in the cells to appear more than once and to have a weak increasing sequence of values in each row and a strict increasing one in each column. A non standard Young Tableau is given below.

5				
4	5	7		
2	3	4		
1	1	3	3	5

D. E. Littlewood remarked that Schur's functions defined often as complicated determinants with difficult computations on them may be recovered by associating a monomial to a non standard Young Tableau T, this monomial is equal to $x_1^{i_1}, x_2^{i_2}, , \ldots , x_n^{i_n}$, such that i_j is the number of times the integer j appears in T. For instance $x_1^2 x_2 x_3^3 x_4^2 x_5^3 x_7$ is the monomial corresponding to the figure above.

The Schur function associated to a partition λ is then the sum of all monomials associated to all Young Tableau's with shape λ. This fact was used by MPS in order to give an enlightening proof of Littlewood Richardson formula which allows to obtain the coefficients of the product of two Schur functions with shapes λ and ν. This proof uses the plactic monoid which was noticed by I. J. MacDonald in the foreword of his celebrated book on symmetric functions writing :

"In the past years the combinatorial structure based on the jeu de taquin has become much better understood. Schützenberger, the main architect of this theory has recently published a complete exposition".

This is a remark in a paper published in 1979. Since then, and until the death of MPS in 1996, Lascoux and himself contributed to many remarkable results, giving a fundamental structure to this theory based on the "jeu de taquin". They proposed a great number of deep results on polynomials appearing in Geometrical Algebra giving a link between Combinatorics and the work of Grothendieck.

References

[1] Jean Berstel, Dominique Perrin, and Christophe Reutenauer. *Codes and automata*, volume 129 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2010.

[2] Jean Berstel and Christophe Reutenauer. *Noncommutative rational series with applications*, volume 137 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2011.

[3] Benjamin Bond and Lionel Levine. Abelian networks II: halting on all inputs. *Selecta Math. (N.S.)*, 22(1):319–340, 2016.

[4] Mireille Bousquet-Mélou. Rational and algebraic series in combinatorial enumeration. In *International Congress of Mathematicians. Vol. III*, pages 789–826. Eur. Math. Soc., Zürich, 2006. https://arxiv.org/abs/0805.0588.

[5] Julius Richard Büchi. On a decision method in restricted second order arithmetic. In *Logic, Methodology and Philosophy of Science (Proc. 1960 Internat. Congr .)*, pages 1–11. Stanford Univ. Press, Stanford, Calif., 1962.

[6] Guillaume Chapuy, Michel Marcus, and Gilles Schaeffer. A bijection for rooted maps on orientable surfaces. *SIAM J. Discrete Math.*, 23(3):1587–1611, 2009.

[7] Kuo-Tsai Chen, Ralph H. Fox, and Roger C. Lyndon. Free differential calculus. IV. The quotient groups of the lower central series. *Ann. of Math. (2)*, 68:81–95, 1958.

[8] Noam Chomsky and Marcel-Paul Schützenberger. The algebraic theory of context-free languages. In P. Braffort and D. Hirschberg, editors, *Computer Programming and Formal Systems*, pages 118–161. North-Holland, Amsterdam, 1963. Traduction: "Algebraische Theorie kontextfreier Sprachen", Kibern. Sb., Nov. Ser. 3, 195-242 (1966).

[9] Noam Chomsky and Marcel-Paul Schützenberger. Théorie algébrique des langages "context-free". In Maurice Gross, editor, *Les modèles en linguistique*, volume 9 of *Langages*, pages 77–118. Didier/Larousse, mars 1968. Traduit par G. Fauconnier.

[10] Alfred H. Clifford and Gordon B. Preston. *The Algebraic Theory of Semigroups. Vol. I.* Mathematical Surveys, No. 7. American Mathematical Society, Providence, R.I., 1961.

[11] Paul M. Cohn. *Free ideal rings and localization in general rings*, volume 3 of *New Mathematical Monographs*. Cambridge University Press, Cambridge, 2006.

[12] Robert Cori. *Un code pour les graphes planaires et ses applications.* Astérisque, Vol. 27. Société Mathématique de France, Paris, 1975.

[13] Françoise Dejean. Sur un théorème de Thue. *J. Combinatorial Theory Ser. A*, 13:90–99, 1972.

[14] Samuel Eilenberg. *Automata, Languages, and Machines. Vol. A.* Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York, 1974. Pure and Applied Mathematics, Vol. 58.

[15] Philippe Flajolet and Robert Sedgewick. *Analytic combinatorics.* Cambridge University Press, Cambridge, 2009.

[16] Dominique Foata and Marcel-Paul Schützenberger. *Théorie géométrique des polynômes eulériens.* Lecture Notes in Mathematics, Vol. 138. Springer-Verlag, Berlin, 1970.

[17] Solomon W. Golomb and Basil Gordon. Codes with bounded synchronization delay. *Information and Control*, 8:355–372, 1965.

[18] Christopher Hollings. *Mathematics Across the Iron Curtain*. Amer. Math. Soc., 2014.

[19] Donald E. Knuth. *The Art of Computer Programming. Vol. 3*. Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1973. Sorting and searching, Addison-Wesley Series in Computer Science and Information Processing.

[20] M. Lothaire. *Combinatorics on Words*. Cambridge University Press, second edition, 1997. (First edition 1983).

[21] R. C. Lyndon. On Burnside's problem. *Trans. Amer. Math. Soc.*, 77:202–215, 1954.

[22] Christophe Reutenauer. *Free Lie algebras*, volume 7 of *London Mathematical Society Monographs. New Series*. The Clarendon Press, Oxford University Press, New York, 1993. Oxford Science Publications.

[23] Craige Schensted. Longest increasing and decreasing subsequences. *Canad. J. Math.*, 13:179–191, 1961.

[24] Marcel-Paul Schützenberger. On the definition of a family of automata. *Information and Control*, 4:245–270, 1961.

[25] Marcel-Paul Schützenberger. On a factorisation of free monoids. *Proc. Amer. Math. Soc.*, 16:21–24, 1965.

[26] Marcel-Paul Schützenberger. On finite monoids having only trivial subgroups. *Information and Control*, 8:190–194, 1965.

[27] Marcel-Paul Schützenberger. Sur une question concernant certains sous-monoïdes libres. *C. R. Acad. Sci. Paris*, 261:2419–2420, 1965.

[28] William Tutte. A census of planar maps. *Canad. J. Math*, 31:249–271, 1963.

[29] Gérard Viennot. *Algèbres de Lie libres et monoïdes libres*, volume 691 of *Lecture Notes in Mathematics*. Springer, Berlin, 1978. Bases des algèbres de Lie libres et factorisations des monoïdes libres.