Additive Cellular Automata and Algebraic Series

Bruce Litow
University of Wisconsin
Computing Services Division
Milwaukee WI 53201 (USA)
Philippe Dumas
INRIA Rocquencourt
Algorithms Project
78153 Le Chesnay (France)

1 Introduction

A cellular automaton is a grid of elementary automata, each communicating only with a finite number of its neighbours. In the simplest models each elementary device, a cell or site, can take only two values and is updated at intervals according to a rule which expresses the actual value from its preceding value and that of its neighbours. Here the values are elements of a field $\mathbb{K}$, finite or infinite.

As O. Martin, A. Odlyzko and S. Wolfram [7] emphasized, the behaviour of a cellular automaton with a finite number of cells on a finite field is ultimately periodic. It is natural to consider also automata with cells in a line, which we call one dimensional automata. So a cell is indexed by an integer $n \in \mathbb{Z}$. At each time all the cells but a finite number are in state 0. Throughout the paper we make use of generating series and, from this point of view, a configuration is a Laurent polynomial

$$C(z) = \sum_{n \in \mathbb{Z}} c_n z^n \in \mathbb{K}[z, 1/z].$$

Here c_n is the state of the cell number n. We assume the rule $C(z) \to C^*(z)$ is additive so that it can be written

$$C^*(z) = Q(z)C(z),$$

where $Q(z) \in \mathbb{K}[z, 1/z]$ is a certain Laurent polynomial too. In plain words, the output of a cell depends linearly on the outputs of some of its neighbours at the previous stage. The most classical example is Wolfram's rule 90 [9] , which corresponds to $Q(z) = z + 1/z$ with $\mathbb{K} = \mathbb{F}_2$ and gives the Sierpinski triangle. If $Q(z)$ is an element of $\mathbb{K}[z]$ or $\mathbb{K}[1/z]$, the automaton is unilateral. We refer to the case, $Q(z) \in \mathbb{K}[1/z]$, as right-hand sided and the case, $Q(z) \in \mathbb{K}[z]$ as left-hand sided.

To study the behaviour of a cell in time we bring in a supplementary indeterminate x and the evolution of the automaton is encapsulated in the generating series

$$M(x, z) = \sum_{n,t} c_{n,t} z^n x^t.$$

Immediately we find, by the additivity of the rule,

$$M(x, z) = \sum_{t=0}^{+\infty} C_0(z) Q(z)^t x^t = \frac{C_0(z)}{1 - xQ(z)},$$

if $C_0(z)$ is the initial configuration. The vertical generating series, corresponding to the cell number n is the coefficient of z^n in $M(x, z)$, is given by

$$G_n(x) = [z^n]\, M(x, z).$$

We show here that the vertical generating functions can be computed directly both in the case of finite fields and in the case of the complex numbers. A number of explicit examples based on classical combinatorial sequences (Catalan, Motzkin) are also worked out. Notably we show how the paper folding sequence can be generated by an extended linear cellular automaton.

2 Automata on a number field

Theorem 1 *If the ground field is a number field, $\mathbb{K}$, (for example the complex numbers field $\mathbb{C}$) the vertical generating series terms, $G_n(x)$, are algebraic on $\mathbb{K}(x)$. Moreover they are rational if the automaton is unilateral.*

Proof : We call $-m$ the valuation of $Q(z)$, which means that $Q(z) = z^{-m} P(z)$ and $P(z)$ is a polynomial with $P(0) \neq 0$. The degree of $P(z)$ is d.

A. Bilateral case. First, we assume $d > m > 0$. The series

$$M(x, z) = \sum_{t=0}^{+\infty} C_0(z) Q(z)^t x^t$$

determines an analytic function in the product of an annulus by a disk, defined by $0 < r \leq |z| \leq R$ and $|x| \leq \epsilon$, for some r, R and ϵ. (The restriction on z makes $Q(z)$ bounded and the condition on x gives us the uniform convergence.) Then this function possesses a Laurent expansion

$$M(x,z) = \sum_{n=-\infty}^{+\infty} G_n(x) z^n$$

and the coefficient $G_n(x)$ is given by the Cauchy formula

$$G_n(x) = \frac{1}{2i\pi} \int_\gamma \frac{C_0(z)}{1 - xQ(z)} \frac{dz}{z^{n+1}}, \tag{1}$$

where γ is a circle centered at 0 with radius $\rho \in]r, R[$.

According to the implicit function theorem, the roots of the denominator inside the circle are analytic functions of x and taking the residues we obtain

$$G_n(x) = \sum_{i=1}^{m} \frac{C_0(\alpha_i)\alpha_i^{m-n-1}}{m\alpha_i^{m-1} - xP'(\alpha_i)} + r_n(x). \tag{2}$$

The term $r_n(x)$ comes from the residue at 0 and is a rational function of x. Formula (2) shows that G_n is an algebraic function on the field of rational functions $\mathbb{K}(x)$.

B. Unilateral case. If $m \geq d \geq 0$ (the automaton is right-hand sided), the sum is a symetrical function of the roots α_1, ..., α_m, then it can be expressed as a rational function of the coefficients of $z^m - xP(z)$, i.e., as a rational function of x. If $m \leq 0$ (left-hand side case) the sum disappears and the result is just $r_n(x)$. It should be noticed that, in the latter case, the Taylor formula also gives the result. □

It is worth noting that the sequence of the states of a given cell obeys a linear recurrence with polynomial coefficients, according to the theorem of Louis Comtet [3], and this permits us to look at the evolution of this particular cell without computing the values of the other cells.

3 Automata on a field of finite characteristic

The reader may have noticed a strong resemblance between the preceding result and an observation made by Furstenberg [4] about the diagonals of

bivariate complex power series. The diagonal of a formal series in m indeterminates

$$f(z_1, z_2, \ldots, z_m) = \sum_{n_1,n_2,\ldots,n_m} a_{n_1,n_2,\ldots,n_m} z_1^{n_1} z_2^{n_2} \cdots z_m^{n_m} \in \mathbb{K}((z_1, z_2, \ldots, z_m))$$

is the formal series in one indeterminate

$$\mathcal{D}f(w) = \sum_n a_{n,n,\ldots,n} w^n \in \mathbb{K}((w)).$$

Harry Furstenberg proved the next result for the case of finite characteristic. (The $\mathbb{C}$ case has been known for a long time.)

Proposition 1 *Let $f \in \mathbb{K}((z))$ a rational formal series in m indeterminates with coefficients in a field $\mathbb{K}$.*

If $m = 2$ and $\mathbb{K}$ is the complex number field, the diagonal $\mathcal{D}f(w)$ is algebraic on $\mathbb{K}(w)$.

If $m \geq 2$ and $\mathbb{K}$ has a finite characteristic, the diagonal $\mathcal{D}f(w)$ is algebraic on $\mathbb{K}(w)$.

Conversely if the ground field, $\mathbb{K}$, is the complex number field or a finite field, an algebraic series is the diagonal of a rational series in 2 indeterminates.

When the ground field is of finite characteristic, we can also treat r-dimensional automata. The cells are now indexed by r-uples from $\mathbb{Z}^r$. The letter z represents the family of indeterminates $(z_i)_{1 \leq i \leq r}$ and the exponents are multi-indices.

Theorem 2 *For a ground field, $\mathbb{K}$, of finite characteristic, and an r-dimensional automaton, the vertical generating series are algebraic on $\mathbb{K}(x)$.*

Proof : We use directly Furstenberg's theorem, because the coefficient of z^ν in $M(x, z)$ is the diagonal of $z^{-\nu} M(x z_1 \cdots z_r, z)$. □

We emphasize the fact that the algebraic character of the vertical generating series

$$G_\nu(x) = \sum_t c_t x^t$$

($\nu \in \mathbb{Z}^r$) translates into a linear recurrence

$$c_t = a_{0,1} c_{t-1} + a_{0,2} c_{t-2} + \cdots + a_{1,0} c_{t/p} + a_{1,1} c_{(t-1)/p} + \cdots,$$

when the field has characteristic p.

The most classical example in this area is the Thue-Morse sequence (μ_n). If an integer n has binary expansion $\epsilon_l \epsilon_{l-1} \cdots \epsilon_0$, then μ_n is the residue modulo 2 of the sum $\epsilon_l + \epsilon_{l-1} + \cdots + \epsilon_0$. The sequence obeys the recurrence

$$\begin{cases} \mu_0 &= 0 \\ \mu_{2k} &= \mu_k \\ \mu_{2k+1} &= 1 + \mu_k, \end{cases}$$

and its generating function

$$\mu(x) = \sum_{n \geq 0} \mu_n x^n \in \mathbb{F}_2[[x]]$$

satisfies the algebraic equation,

$$\mu(x) = (1+x)\mu(x)^2 + \frac{x}{1+x^2}.$$

Multiplying by $1 + x^2 = (1+x)^2$, we obtain the homogeneous recurrence

$$\mu_n + \mu_{n-2} = \mu_{n/2} + \mu_{(n-1)/2} + \mu_{(n-2)/2} + \mu_{(n-3)/2}$$

or, if one prefers,

$$\begin{cases} \mu_{2k} + \mu_{2k-2} &= \mu_k + \mu_{k-1} \\ \mu_{2k+1} + \mu_{2k-1} &= \mu_k + \mu_{k-1}. \end{cases}$$

This type of recurrence is a hybrid between the standard linear recurrences associated to the rational fractions and the divide-and-conquer recurrences, that are classical in computer science. If further the field is finite, we obtain exactly the automatic sequences introduced by Cobham, as results from the theorem of Christol, Kamae, Mendès France and Rauzy [1, 2].

Frequently, the generating series with a finite ground field appears as a reduction modulo p of a series with integral coefficient.

4 Examples

Generalization of the Sierpinski triangle If a and b are two integers and $g_{a,b}(x)$ is the series

$$\sum_{n \geq 0} \binom{a+bn}{n} x^n \in \mathbb{Q}[[x]],$$

we have [8]

$$g_{a,b}(x) = [z^0]\frac{(1+z)^a}{1-x(1+z)^b/z}.$$

So we can find $g_{a,b}(x)$ using a cellular automaton initialized by

$$C_0(z) = (1+z)^a$$

with the rule

$$C_t(z) = \frac{(1+z)^b}{z} C_{t-1}(z).$$

For $a = 0$ and $b = 2$, we obtain Wolfram's rule 90 and Sierpinski's triangle by reducing modulo 2.

Catalan numbers It is well known that the n-th Catalan number is the number of binary trees with size n [5, p. 111]. A generating series for these numbers is

$$\sum_{n\geq 0} \frac{1}{n+1}\binom{2n}{n} x^{2n+1} \in \mathbb{Q}[[x]]$$

and it satisfies the equation $y = x + xy^2$. According to Furstenberg's proof [4, p. 276] we can obtain it as the coefficient of z^0 in the rational series

$$z^2\frac{1-2xz}{z-x-xz^2}$$

and as the sequence of the states of the cell number 0 of the cellular automaton initialized by

$$C_0(z) = z$$

and the rule

$$C_t(z) = \left(\frac{1}{z} + z\right) C_{t-1}(z).$$

The diagram below illustrates the rule. (We will use the same convention for the other examples.) The coefficient of the box tagged by a circle $\circ$ is obtained by adding the coefficients in the boxes tagged by a cross $\times$.

$\times$	$\times$	$\times$	$\times$
	$\circ$		

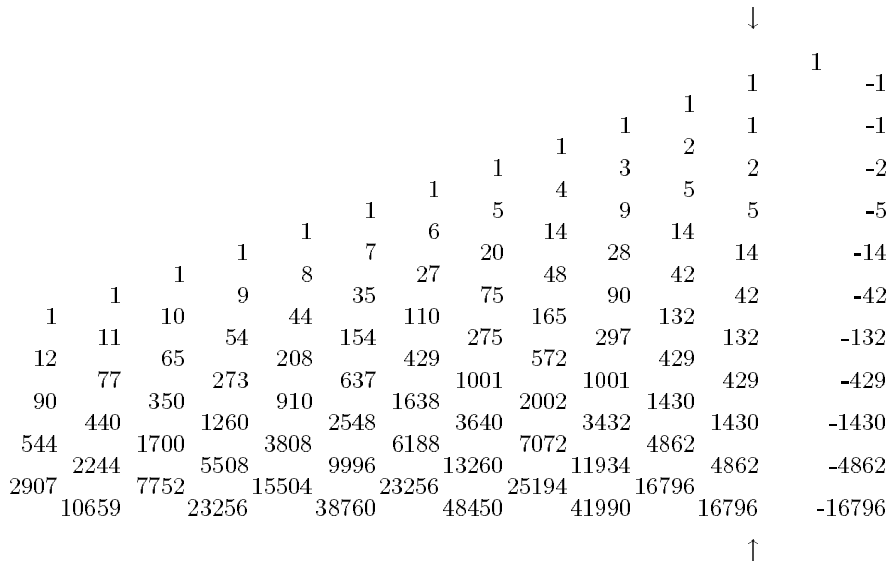


Figure 1: A cellular automaton for the Catalan numbers.

We see the evolution of the automaton for the cells, which have a number between -11 and 2, in the figure 1. (The arrows indicate the column number 0.)

More generally we can construct the generating series for the trees submitted to a condition on the degrees of the nodes, since this produces the equation

$$y = x + x(y^{d_1} + \cdots + y^{d_m}),$$

if we consider rooted, oriented trees each of whose nonleaf nodes has either d_1 or d_2 ...or d_m successors.

Paper folding sequence This example shows an extension of our model in which each automaton of the network can store a finite sequence of values instead of a single value and the rule is still additive. This could be realized by adding a limited memory.

The paper-folding sequence [6] is defined by the following procedure: one folds a sheet of paper an infinite number of times always in the same direction; next one unfolds it and one codes the sequence of the folds by

0 or 1 according to their up or down position . The sequence obeys the recurrence

$$\begin{cases} u_{4k} &= 1 \\ u_{4k+2} &= 0 \\ u_{2k+1} &= u_k, \end{cases}$$

and its generating series $u(x) \in \mathbb{F}_2[[x]]$ is a solution of the equation

$$x(1+x)^4y^2 + (1+x)^4y + 1 = 0.$$

The first few terms are

$$1 + x + x^3 + x^4 + x^7 + x^8 + x^9 + x^{12} + x^{15} + x^{16} + x^{17} + x^{19} + x^{20} + \cdots.$$

The shape of the equation is not directly compatible with a representation by an additive cellular automaton since the coefficient of x^0 is not a monomial. We modify the series by subtracting its first term and $Y = 1 + y$ satisfies the equation

$$x(1+x)^4Y^2 + (1+x)^4Y + x^5 + x^4 + x = 0;$$

the latter has the right shape. The series Y is the coefficient of z^0 in the rational series

$$\frac{z^2(x^4+1)}{z + x(z^2+1) + x^4(z+1) + x^5(z^2+1)}.$$

We initialize the automaton by the configurations

$$\begin{aligned} C_0(z) &= z, \\ C_1(z) &= 1 + z^2, \\ C_2(z) &= 1/z + z^3, \\ C_3(z) &= 1/z^2 + 1 + z^2 + z^4, \\ C_4(z) &= 1/z^4 + 1/z^2 + z^4 + z^6 \end{aligned}$$

and, for $t \geq 5$, we define the state at time t, C_t by the recurrence relation

$$C_t(z) = \left(\frac{1}{z} + z\right) C_{t-1}(z) + \left(\frac{1}{z} + 1\right) C_{t-4}(z) + \left(\frac{1}{z} + z\right) C_{t-5}(z).$$

Figure 2 shows the evolution of this automaton and the diagram illustrates the rule. One recognizes in the column indicated by arrows the paper-folding sequence, except for its first term, which is suppressed.

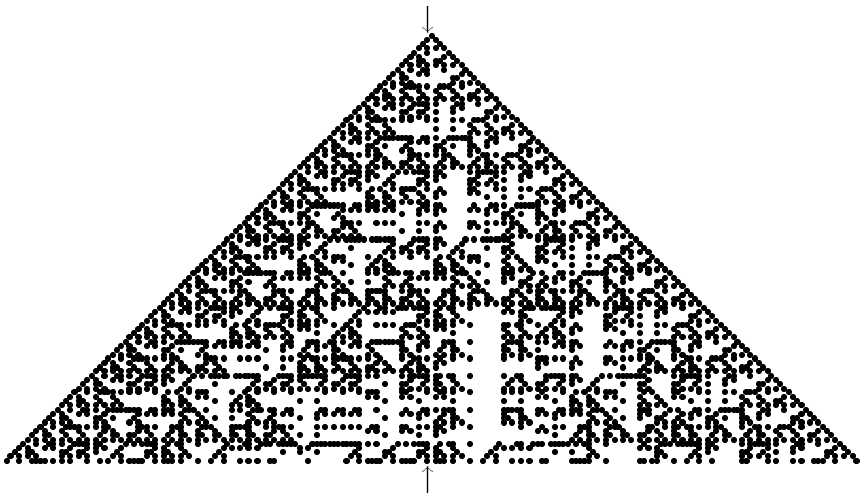


Figure 2: The paper folding cellular automaton.

×		×
×	×	
×		×
	○	

Motzkin numbers Here a path is a sequence $u_0 u_1 \dots u_l$ of points from $Z \times N$, which satisfies $u_{k+1} - u_k \in \{(1,-1),(1,0),(1,1)\}$ for $k = 0 \dots l-1$. The number l is the length of the path. The Motzkin number M_n is the number of paths of length n from some point with ordinate 0 to points with ordinate 0 [5, p. 309].

The generating series of the Motzkin numbers is

$$M(x) = \sum_{n \geq 0} M_n x^n = \frac{1 - x - \sqrt{1 - 2x - 3x^2}}{2x^2}.$$

Its first few terms are

$$1 + x + 2x^2 + 4x^3 + 9x^4 + 21x^5 + 51x^6 + 127x^7 + 323x^8 + 835x^9 + 2188x^{10} + \cdots.$$

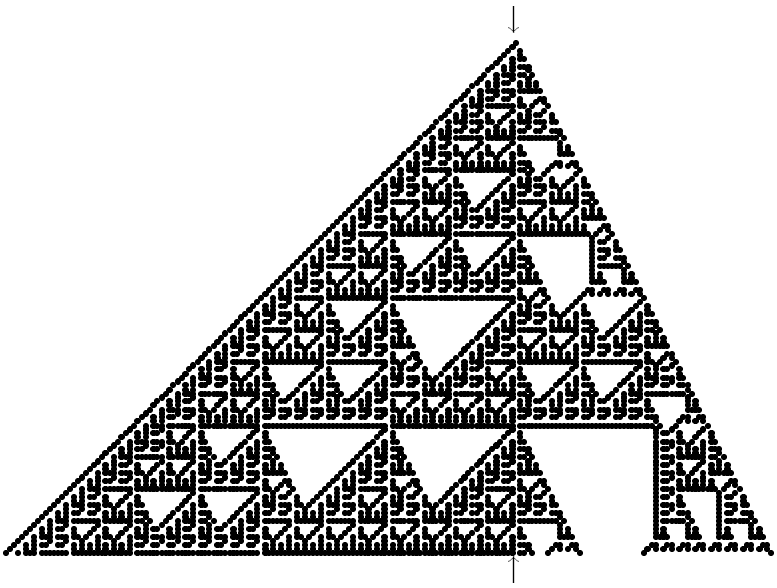


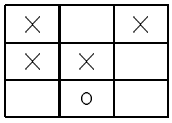
Figure 3: A cellular automaton for the Motzkin numbers reduced mod 2.

It is an algebraic series and satisfies $P(x, M(x)) = 0$ if

$$P(x, y) = 1 + (x - 1)y + x^2y^2.$$

As in the previous example, we suppress its first term to obtain an equation in which the constant term (relative to x) is a monomial.

The diagram illustrates the rule with the same convention as for the preceding example. Figure 3 shows the evolution of the automaton.



Acknowledgments We acknowledge the help of Philippe Flajolet who intoduced the authors to one another and gave them some useful hints like pointers to Pólya's work. We thank also Bruno Salvy who pointed out the Motzkin numbers as a nice example.

References

[1] Jean-Paul Allouche. Automates finis en théorie des nombres. *Expositiones Mathematicae*, 5:239–266, 1987.

[2] G. Christol, T. Kamae, M. Mendès France, and G. Rauzy. Suites algébriques, automates et substitutions. *Bulletin de la Société Mathématique de France*, 108:401–419, 1980.

[3] L. Comtet. Calcul pratique des coefficients de Taylor d'une fonction algébrique. *Enseignement Mathématique*, 10:267–270, 1964.

[4] Harry Furstenberg. Algebraic functions over finite fields. *Journal of Algebra*, 7:271–277, 1967.

[5] I. P. Goulden and D. M. Jackson. *Combinatorial Enumeration*. John Wiley, New York, 1983.

[6] A. van der Poorten M. Dekking, M. Mendès France. Folds! *Mathematical Intelligencer*, 4:130–138, 173–181, 190–195, 1982.

[7] Olivier Martin, Andrew M. Odlyzko, and Stephen Wolfram. Algebraic properties of cellular automata. *Commun. Math. Phys.*, 93:219–258, 1984.

[8] Georges Pólya. Sur les séries entières dont la somme est une fonction algébrique. *Enseignement Mathématique*, 1–2:38–47, 1921–1922. in Georges Pólya: Collected Papers Volume I: Singularities of Analytic Functions edited by R. P. Boas The MIT Press 1974.

[9] Stephen Wolfram. *Theory and Applications of Cellular Automata*, volume 1 of *Advanced series on complex systems*. World Scientific, 1986.