Pascal's Triangle, Automata, and Music

Jean-Paul Allouche
Université de Marseille

December 5, 1994

[summary by Philippe Dumas]

The reduction of Pascal's triangle modulo a prime number p, or a power of a prime number, has been intensively studied. It is known that the reduction produces a phenomenon called auto-similarity; natural homotheties with contraction ratio $1/p^k$ appear, and a limit set in the sense of the Hausdorff metric exists; moreover the limit set has fractal dimension $\log\binom{p+1}{2}/\log p$. The reduction modulo a composite number does not produce this phenomenon of auto-similarity; however it is still possible to make a limit set come out [7]. This talk shows a way to describe the complexity of such a double sequence. The basic tool is the concept of a double automatic sequence. First we introduce automatic sequences and complexity of sequences over a finite set; as an illustration it is shown that there are about $(m+n)^2$ distinct rectangular blocks with m rows and n columns in Pascal's triangle reduced modulo a prime number [1]. Next we give a statement about the automaticity of some linear cellular automaton, and specifically of Pascal's triangle reduced modulo an integer [2]. Finally an application to musical composition is mentioned [3, 5].

1. Automatic sequences

Formal language theory provides a way to define infinite words as fixed points of morphisms. As an example, take the alphabet $\{0, 1\}$ and the recurrence

$$w_0 = 0, \qquad w_{n+1} = w_n\overline{w_n},$$

where the bar means exchange 0 and 1; the first few terms of the sequence are the following words,

$$\begin{aligned} w_0 &= 0 \\ w_1 &= 01, \\ w_2 &= 0110, \\ w_3 &= 01101001, \\ w_4 &= 0110100110010110, \\ w_5 &= 01101001100101101001011001101001. \end{aligned}$$

Ultimately an infinite word appears. This word is the Thue-Morse word, which is a fixed point of the substitution [6]

$$\sigma(0) = 01, \qquad \sigma(1) = 10.$$

Another example is defined as follows. Two letters, a left brace $\{$ and a right brace $\}$ are the elements of the alphabet. The sequence of words A_n is defined by the rules

$$A_0 = \{\}, \qquad A_{n+1} = \{A_0 \dots A_n\}.$$

The limit sequence is a fixed point of

$$\lambda(\{) = \{\{\}, \qquad \lambda(\}) = \}\},$$

and provides the sequence of natural integers, as defined by Bourbaki.

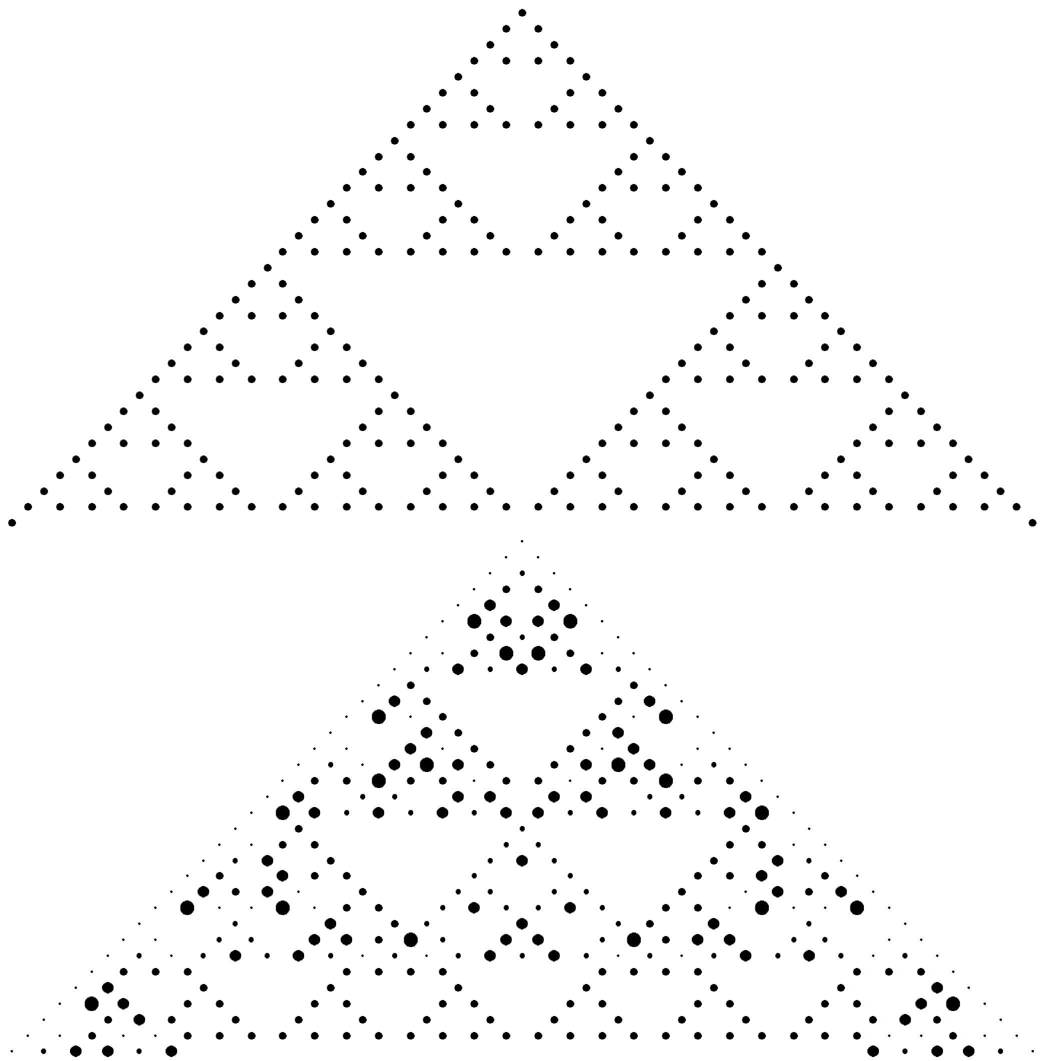

FIGURE 1. Pascal's triangle reduced modulo 2 (top) or modulo 6 (bottom); the size of a dot is proportional to the value of the residue it represents. In the first case the picture is auto-similar, but not in the second one. Nevertheless, in both cases, a limit set exists in the Hausdorff metric.

All these sequences are 2-automatic. (The 2 refers to the fact that the alphabet has two letters.) The definition may be adapted to double sequences. For instance the morphism

$$\beta(0) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \qquad \beta(1) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

applied to the starting point 1 gives Pascal's triangle reduced modulo 2,

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

2. Complexity

The complexity of a sequence over a finite alphabet is defined as another sequence $p(n)$, where $p(n)$ is the number of distinct factors with length n in the given sequence. Obviously the complexity satisfies

$$1 \leq p(n) \leq q^n,$$

if the alphabet is of size q. The complexity reflects how intricate the sequence is. For instance, if for any n the inequality $p(n) \leq n$ is satisfied, the sequence is ultimately periodic. In the case of the Thue-Morse sequence, the sequence of differences $p(n+1) - p(n)$ is 2-automatic. Cobham showed that every automatic sequence has complexity $O(n)$ [4].

For double sequences, the shape of block to use in the definition of complexity is rather arbitrary. A natural choice is to consider rectangular blocks. Then, the complexity is a double sequence $P(m,n)$ where $P(m,n)$ is the number of distinct rectangular blocks with m rows and n columns occurring in the given double sequence.

For Pascal's triangle reduced modulo 2, it is readily noticed that the complexity satisfies

$$P_2(m,n) = P_2(1, m+n-1).$$

Moreover the relation

$$(1+x)^{2t} = (1+x^2)^t$$

shows that row t determines rows $2t$ and $2t+1$; as a consequence the formula

$$P_2(1,n) = n^2 - n + 2$$

is satisfied. More generally, Pascal's triangle reduced modulo a prime p has complexity order n^2. The proof relies on the fact that differences of order 2 of $P_p(1,n)$ form a p-automatic sequence. If the modulus is not a prime but is square-free, for example if the modulus equals 6, the Chinese remainder theorem shows that

$$P_6(1,n) \leq P_2(1,n) P_3(1,n).$$

Actually, the quantities are equal, since the residues modulo 2 and modulo 3 may be considered as independent. More generally, the complexity $P_q(1,n)$ of Pascal's triangle reduced modulo a square free number q is shown to be of order $n^{2\omega(q)}$, where $\omega(q)$ is the number of prime factors of q. The case of prime powers may be tackled by a formula due to Kummer, namely $\left((1+x)^{p^{a-1}}\right)^p = (1+x^p)^{p^{a-1}}$ mod p^a. This result gives a mathematical meaning to the feeling that Pascal's triangle reduced modulo m is more and more complex as the number of prime factors of m increases.

3. Automaticity of linear cellular automata

Pascal's triangle is an example of a linear cellular automaton. There is an initial state, here $g(x) = 1$, and a rule $r(x) = 1 + x$. At time t, the state of the automaton is $g(x)r(x)^t$. To recover a more classical definition from this one, it suffices to consider that coefficients of the state at time t are the contents of cells arranged along the infinite line of integers $\mathbb{Z}$. Moreover the set of states of a cell is finite if the ring of coefficients is finite; here the ring is the ring of integers modulo m. In other words, the double sequence of binomial coefficients reduced modulo m shows the evolution of a classical linear cellular automaton.

It is shown that, when reduced modulo a prime power p^ℓ, a linear cellular automaton provides a p-automatic double sequence. The proof needs an additional concept: a polynomial $r(x)$ is said to have the m-Fermat property if it satisfies

$$r(x^m) = r(x)^m.$$

The Kummer formula above gives an example with $m = p$ a prime number. When the rule $r(x)$ has the m-Fermat property, then the associated double sequence is m-automatic.

As a consequence Pascal's triangle reduced modulo m is m-automatic if m is a prime power. Moreover the converse is true, and its proof relies on Cobham's theorem which asserts that a sequence both p-automatic and q-automatic, p and q being prime and distinct, is ultimately periodic. Here the sequence used is the sequence of central binomials $\binom{2n}{n}$. This result gives a precise formulation of the fact that Pascal's triangle reduced modulo a composite number is more complex than when reduced modulo a prime power.

4. Musical composition

Some composers have used finite automata to produce musical motifs. For instance Tom Johnson has used the morphism defined on a two-letter alphabet $\{+, -\}$ by

$$\mu(+) = + - +, \qquad \mu(-) = - - +.$$

A $+$ codes a melodic ascent, and a $-$ codes a melodic descent. In the same vein, he has used Pascal's triangle reduced modulo 7. The interest of such a composition is that automatic sequences are at the frontier between periodicity and chaos. But as Tom Johnson himself says, this can only be a tool and certainly not a way of composing music in a purely automatic fashion.

Bibliography

[1] Allouche (J.-P.). – Sur la complexité des suites infinies. *Bulletin of the Belgian Mathematical Society*, vol. 1, 1994, pp. 133–143.

[2] Allouche (J.-P.), von Haeseler (F.), Peitgen (H.-O.), and Skordev (G.). – Linear cellular automata, finite automata and Pascal's triangle. – To appear in Discrete Applied Mathematics, 1995.

[3] Allouche (Jean-Paul) and Johnson (Tom). – Finite automata and morphisms in assisted musical composition. – Preprint, 1995.

[4] Cobham (Alan). – Uniform tag sequences. *Mathematical Systems Theory*, vol. 6, n° 2, 1972, pp. 164–192.

[5] Johnson (Tom). – *Formulas for String Quartet*. – Editions 75, 75, rue de la Roquette 75011 Paris, 1994.

[6] Lothaire (M.). – *Combinatorics on Words*. – Addison-Wesley, 1983, *Encyclopedia of Mathematics and its Applications*, vol. 17.

[7] Willson (S.). – Cellular automata can generate fractals. *Discrete Applied Mathematics*, vol. 8, 1984, pp. 91–99.