The Informational Energy for Cellular Automata

Ioan Hălălae

"EFTIMIE MURGU" University, Faculty of Engineering, Traian Vuia Square 1-4, 320085 Resita, Romania
Phone: +40 255 210227, Fax: +40 255 210230, e-mail: i.halalae@uem.ro

Abstract: Cellular automata (CA) represent lately a widely used model for modelling the dynamic phenomena. Generally, they are used in a rather empirical manner. This paper proposes the adaptation of the informational energy concept for CA and its use as a predictor for the evolution of CA.

Keywords: cellular automata, randomness, prediction, informational energy, experimental mathematics

1 Introduction

In a larger context [1] I have studied the possibility of using CA as noise generators. The starting point was the project 'Looking for the lost noise' [2].

Our proposal is to use, as a generator of pseudo-random sequences, an uniform CA with 256 cells that can have two states, over which a global loop is closed. We add several new things to the classical use of a CA, in order to improve the 'random' evolution.

The random evolution is manifesting itself into a chaotic process, very sensitive to the initial state of a simple device working according to a strange rule.

The first problem is that CA are 'implicitly' considered as having a random evolution. Randomness is not rigorously defined for CA. More generally, the notion of evolution of CA either is not clearly defined (that is, formally, or, at least, functionally).

In order to make the monitoring, we have devised a new methodology. We have functionally devised the notions of evolution, randomness and cyclic character in CA. As regards randomness, we have adapted G. Chaitin's definition for binary strings.

2 Theoretical Concepts

2.1 Cellular Automata

An ‘elementary’ cellular automaton consists of a sequence of sites carrying values 0 or 1 arranged on a line. The configurations of the system are thus sequences of 0 and 1. In general, the sites of a cellular automaton may be arranged on any regular lattice, and each site may take on any discrete set of values.

They evolve in a series of time steps. At each step, the value of each site is updated according to a specific rule. The rule depends on the value of a site, and the values of, say, its two nearest neighbours. The value at each site evolves deterministically with time according to a set of definite rules involving the values of its nearest r neighbours. Different rules (lookup tables) generate different types of dynamics of CA, when the rules are iterated in time ([3], [4], [5], [6]).

2.2 Randomness and Prediction

G. Chaitin established in a large set of papers ([7]–[11]) a complete theory about randomness, connected with the capacity of programs for computing finite binary sequences. In short, he defines a string as being random if his shorter representation is the string itself. In other words, a string of bits is random as long as it does not repeat itself. When it starts repeating itself, it can be represented by an algorithm, and becomes predictable.

We adopt here the same functional point of view for CA: when it starts repeating itself it becomes predictable. Automaton’s evolution is random *until it starts cycling.*

We decide to stop the automaton in the moment it starts cycling, and *to consider its evolution as complete.*

2.3 Experimental Mathematics

Experimental Mathematic is a paradigm appearing in association "with the exploratory use of a computer" [12], especially "when one attempts to analyze experimentally algorithms" [13].

In our case, we have used the computer for simulating CA. The space of the solutions is of the 2^{256} dimension.

2.4 The Main Challenge

The main challenge is the one S. Wolfram formulated in [14]: ‘But what should be done is to find a characterization of those properties whose behaviour can be found by efficient methods, and those for which computational irreducibility makes explicit simulation the only possible approach, and precludes a simple description.’

2.5 The Project

The project presented here aims to explore the possibility of using the CA as random number generators. Simulation was made with linear CA, with 256 cells. We did not use genetic algorithms for obtaining better chromosomes from the older ones.

3 The Project Presentation

The starting point was the project 'Looking for the lost noise' by Gh. Stefan [2]. Let us see an abstract of this paper, which can be considered the starting point of our paper.

The proposal is to use a CA as pseudo-noise generator, to which a global loop was added (fig. 1).

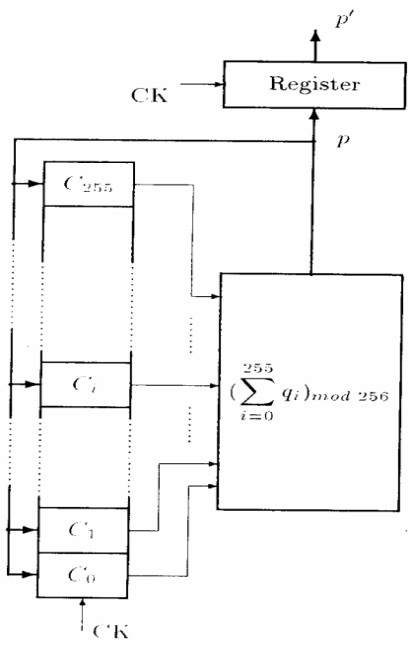

Figure 1

One of the simplest programmable CA, with a global loop closed through summation function

Our hope is that the 'feed forward' of the attached loop improves the random behaviour of the system.

The transition function over a loop being very simple, the random behaviour depends only on the initial state of the CA. In the design process the main step is to find a 'noisy' initial state.

Conclusions: the only way to present the machine behaviour is to make it work as a simulator. The formal method of using is the experimental mathematics.

S. Wolfram [15] was the first to use CA as noise pseudo-generators. He used the linear and uniform CA, with two states cells.

4 The Development of the Project

4.1 Chromosomes

We use a definition of randomness inspired by G. Chaitin: an automaton is random as long as *its evolution cannot be predicted*. From the moment when its evolution can be predicted, it is no longer random. An automaton is random *until it starts cycling*. So the problem is rephrased as follows: no matter the value, we stop an automaton when it becomes cyclic.

Practically, one way or another, we must:

- have an evidence of the states of the automaton, in the order of their appearance, and
- test a new state of the automaton (the current status, the last one resulted from the calculus process) if it appears from the first time or not.

If we find it in the 'history' of the automaton, it represents the end of the first cycle of the automaton and the generation of new states stops. If not, the new status is to be archived and the iterations continue.

We keep the whole automaton 'history' in a matrix, whose successive rows memorize the automaton's states in the order of their apparition. We will name this matrix 'the evolution matrix'.

In order to synthesize the results, we have to monitor the chromosomes in two different ways:

- the first one, we monitor for each chromosome the evolution as an internal mechanism of the cellular automaton
- the second one, we archive the results of each chromosome, individually, together with its initial configuration and the performance

And, of course, not to repeat the tests for the same chromosome.

As regards chromosomes, there are two problems:

- the generation of the initial states
- the concrete work with a CA.

We have systematically generated distinct chromosomes.

For actually working with a cellular automaton (monitoring the evolution of the automaton), we arrived at the following diagram:

1. An initial state is generated

2. The initial state is memorized in the evolution matrix, in line 1 of the matrix; a counter number_cycles_of_life is initialized with 1.
3. A calculating function is applied to a new state of the automaton; we name the new state of the automaton chromosome_current; the counter number_cycles_of_life is increased with 1.
4. The sequential chromosome_current is compared with the values from the evolution matrix, from position 1 to position number_cycles_of_life-1 the sequential chromosome_current is compared to the values in the evolution matrix, from position 1 to position number_cycles_of_life-1
5. If an equality does not appear, the chromosome_current is memorized in the evolution matrix; and we came back to step 3
6. If an equality appears (the value already exists; that means that the chromosome_current represents the closure of the first cycle), the experiment stops, but (number_cycles_of_life -1) represents the performance of the chromosome; we archive the chromosome

The archiving of the results is made in a database.

4.2 The Connection between the Automaton's Initial Structure and Performance

In all the discussions about the CA the local structure is mentioned. By local structure we understand the spatial distribution of values '0'/'1' (the CA configuration). Associated to the notion of local structure there appears that of density: the number of values '1' in the initial configuration of CA. Nevertheless, there was no systematic exploration of the combination of these two factors.

We generated 500.000 CA with similar densities, between 124 and 133. The results are those in table 1. By performance we understand the number of tacts until the automaton becomes cyclic. We have generated the data and counted the results by groups.

Performance	Number of cases
3-54	454. 828
55-106	0
107-158	10
159-210	0
211-262	42. 325
263-314	2. 825
315-366	0
367-418	0
419-470	0
471-518	12

Table 1

Statistics for the automata evolutions

The great variation of the performances of the chromosomes that have similar initial densities leads us to the necessity of a differentiating criterion between algorithms with the same initial density. A possibility of global characterizing of the automaton would implicitly represent a prediction instrument too.

5 Comportamental Factorization

We try to solve this problem: *how can we realize a differentiation between the cellular automata with the same density*.

We arrive at the specificity of cellular automata: the modification of the cell value by reaction to the neighbourhood. Thus, it appeared the idea of the classification of automata by factorization according to the initial value of the cells. In this way we have a possibility of differentiation of cellular automata with the same initial density.

Now, these results led us to the idea mentioned above: classification of automata by the factorization according to the value of the cells. Let us remember the calculation of the new value of a cell:

$C(i,n) = C(i-1,n-1) + C(i,n-1)* 2 + C(i+1,n-1)*2^2$

where we have noted C(j,k) the value of the j cell at a moment k. The value C(i,n) is the value of the cell from the calculus table, on a line corresponding to the density of the automaton.

Practically, we obtain the notations:

0 **0** 0 $\Rightarrow$ **0** noted v0

1 **0** 0 $\Rightarrow$ **1** noted v1

0 **1** 0 $\Rightarrow$ **2** noted v2

1 **1** 0 $\Rightarrow$ **3** noted v3

0 **0** 1 $\Rightarrow$ **4** noted v4

1 **0** 1 $\Rightarrow$ **5** noted v5

0 **1** 1 $\Rightarrow$ **6** noted v6

1 **1** 1 $\Rightarrow$ **7** noted v7

which we shall use below.

6 The Informational Energy in Cellular Automata

We studied the *problem of the structure according to the point of view of informational „disorder"*: we calculated the global informational energy of a cellular automaton, making an adaptation of the classical Onicescu formula.

First, let us remember the definition of the informational energy ([Onicescu '79]).

The global information of the system S with the states s_1, s_2, …, s_n having the following weigths p_1, p_2, …, p_n can be expressed by its informational energy calculated with the formula:

$$E_s = \sum_{j=1}^{n} p_j^2 \qquad (1)$$

(where $\sum_{j=1}^{n} p_j = 1$)

Let's go back, now, to the discussion from the preceding paragraph. In the classical formula (1) s_1, s_2, …, s_n, represent the states of the system. In the case of the cellular automaton, the state of the system is its interior configuration at a certain moment. We are in the situation to make an adaptation of the concept. The form in which we use the concept of informational energy is the following:

$$E = (v_0^2 + v_1^2 + v_2^2 + v_3^2 + v_4^2 + v_5^2 + v_6^2 + v_7^2)/ 256^2 \qquad (2)$$

where the notations v_0, v_1, …,v_7 represent the factorization of cells according to the local states.

6.1 Correlations between Performance and Energy

We arrive at the connection between the automaton's performance and the informational energy (calculated with formula 2).

We browsed again the database which contains the chromosomes. For each of them we calculated for each cell the classification v0 … v7. We counted the result in a file. Then, we calculated the corellation between the chromosome's performance and the initial structure: using the v0 … v7 counter, ve calculated the informational energy using formula 2. The results are in table 2.

Density	m/e Corelation
124	0.0261134
125	0.0754734
126	0.0474427
127	0.3759557
128	0.5101165

129	0.4001813
130	0.8669366
131	0.1601014
132	0.8857184
133	0.0557577

Table 2

The conection between performance (m) and informational energy (e).

Conclusions and directions for further research

Let us first see which are the totally new results brought about by this paper. On a conceptual level:

- A functional definition of evolution, randomness and cyclicity were given for CA. Until now, the notions of evolution and cyclicity were not clearly defined.
- There are no mathematical concepts for the calculations connected to the parallel phenomena. The informational energy introduced here is the first classic mathematical concept adapted to parrallel phenomena.
- Until now, discussions were moreover ‘intuitive’, based on a small number of experiments and on a graphical interpretation (visualization) of the results. We have refocused the discussions, from the domain of ‘intuition’, to that of figures.

On a practical level, we have, too, some immediate directions for further research:

- Regarding the global energy itself it remains to be studied the measure in which it can become a fine predictor for the characterisation of the evolution of automata, by doing again the calculations with partial regressions.
- The numerical results are dispersed. But this aspect is typical for phenomena with nonlinear dynamics.Remember that on some portions we have a remarkable correlation. This leads us tpo the idea of modeling the behaviour of cellular automata using neuronal nets.

References

[1] Halalae, I.: ”Studiul automatelor celulare ca generatoare de zgomot, folosind algoritmi genetici” (The study of CA as noise genrators, using gentical algorithms), PhD thesis, available by direct contact

[2] Stefan, Gh.: ”Looking for the lost noise”, CAS '98 Proceedings, Oct. 6 - 10, 1998, Sinaia, Romania. p. 579 - 582

[3] Wolfram S.: ”Cellular Automata as Simple Self-Organizing Systems”, Caltech preprint CALT-68-938 (1982)

[4] Wolfram S.: “Cellular Automata”, Los Alamos Science, 9 (Fall 1983) 2-21

[5] Davis L.: Handbook of genetic algorithms, Van Nostrand Reinhold, New York, 1991

[6] Mitchell M.: An Introduction to Genetic Algorithms, A Bradford Book, The MIT Press, 1999

[7] G. Chaitin, "On the length of programs for computing finite binary sequences", Journal of ACM 13 (1966), p. 547-569

[8] Chaitin G.: "On the length of programs for computing finite binary sequences: statistical considerations", Journal of ACM 16 (1969), p. 145-159

[9] Chaitin G.: "A theory of program size formally identical to information theory", Journal of the ACM 22 (1975), pp. 329-340

[10] Chaitin G.: Algorithmic Information Theory, 3rd printing, Cambridge University Press, 1990

[11] Chaitin G.: "Paradoxes of Randomness", in Complexity, Vol. 7, No. 5, May/June 2002, pp. 14-21

[12] Johnson D.: "A Theoretician's Guide to the Experimental Analysis of Algorithms", AT&T Labs Research, 1996. Available from www.research.att.com/¸dsj/papers/exper.ps

[13] Borwein J., Borwein P., Girgensohn R, S. Parnes S., "Experimental mathematics: A discussion". Mathematical Intelligencer 18, 4 (May 1996), 12--18

[14] Wolfram S.: "Twenty Problems in the Theory of Cellular Automata", in Physica Scripta, T9 (1985) 170-183

[15] Wolfram S.: "Random Sequence Generation by Cellular Automata", in Advances in Applied Mathematics, 7 (June 1986) 123-169

[16] Onicescu, O. "Estension of the Theory of Probabilities", in "Logic, Methodology and Philosophy of Science IV, Proceedings of the IV Congres, Bucharest 1971",

[17] Onicescu, O.: Principiile teoriei probabilităților, Editura Academiei, Bucureşti, 1969

[18] Sahleanu V.: Eseu de biologie informationala, Edit. Stiintifica, Bucuresti, 1973