ROMANIAN JOURNAL OF INFORMATION SCIENCE AND TECHNOLOGY
Volume **18**, Number 2, 2015, 166–181

On the Regularities and Randomness of the Dynamics of Simple and Composed CAs with Applications

Horia-Nicolai TEODORESCU[1,2]

[1] "Gheorghe Asachi" Technical University of Iasi, Romania, Iasi
[2]Romanian Academy, Romania, Iasi
E-mail: hteodor@etti.tuiasi.ro

Abstract. Several elementary dynamic properties of cyclic cellular automata are analyzed, with emphasis on the decorrelation of the cycles. Potential applications to the generation of pseudo-random sequences and to spread spectrum communications are briefly discussed.

Key-words: cellular automata, pseudo-random number generator, dynamics, spread spectrum, CDMA.

1. Introduction and preliminary results

The purpose of this study is to discuss the dynamics of a class of cellular automata (CAs) that exhibit complete periodicity (CP) and equal length cycles (ELCs), pointing toward applications including communications (spread spectrum modulation) and pseudo-random pattern generators (PRPGs). The characterization of the degree of independence of the cycles produced, which is one of the main points of discussion, is based on minimal distances, normalized minimal distances between the sequences, and on Pearson correlation of the cycles.

The main type of applications referred to are psudo-random pattern generators (PRPGs) and pseudo-random number generators (PRNGs) used in testing equipment and circuits, such as BIST (built-in self-test circuits) and communications, such as code division multiple access (CDMA). ELC-CAs were first used in [17], with the abbreviation ELC-CA and the corresponding name introduced in [9] and were presented in some recent papers as good candidates for PRNGs [14], [15], having the advantage that they produce several sequences of the same length that can be used as random

sequences. We show that this statement is arguable and that in many cases the cycles of ELC-CAs are less independent or decorrelated than one assumes in [14], [15].

Previous papers dealt with the randomness of the ELC-CA cycles, but not with the randomness of the family of cycles itself, that is, the independence (orthogonality) or decorrelation of the cycles in an ELC-CA set. The aim is to bring clarity to this problem and to its practical implications and to show several results in this line. Specifically, an aim is to investigate in depth the possibility of using equal length cycle CAs (ELC-CAs) for PRNGs, built-in self test circuits (BISTs), CDMA, and similar purposes by analizying the properties of the families of periodic attractors (cycles) generated. A special attention is paid to the properties of the keys generated. We also introduce 'staked' CAs in an attempt of obtaining, in a systematic if not optimal manner, CAs with improved decorrelation between the cycles.

The interest in CAs and various developments of the CAs, such as ELC-CAs and fuzzy CAs in recent years continues the trends already established, with various applications in cryptography [3], [4], [8], CDMA [18], and random number generators [14], communications [18], classification [9], and modeling [20], [23]. However, the fundamentals of the applicability of some of the newer sub-classes of CAs was not analyzed in depth. This article aims to contribute to the topic by analyzing several properties of the ELC-CAs.

The summary of the paper is as follows. Section 2 summarizes some concepts related to CAs and to ELC-CAs. Section 3 derives several elementary properties of the dynamics of CAs. Section 4 introduces notions related to the correlations of the generated cycles and exemplifies applications. In Section 5, hybrid CAs are proposed. The final section derives conclusions.

The following abbreviations and notations are used throughout the paper: CA cellular automaton, CAs cellular automata, CDMA code division multiple access, CPCA completely periodical CA, ELC equal length cycle, ELC-CA CA with all cycles equal, PRNG pseudo-random number generator, SsC spread spectrum communications.

Notations

$|\cdot|$ cardinal or modulus
$\circ$ function composition
$\vee$ logical OR
$\wedge$ logical AND
' denotes the next state, locally or globally, in decimal or binary representation.
$[0,0],[0,1],[1,0],[1,1]$ denote the four combinations of possible boundaries (edges), in configurations of the CA such as $[e_L = 0, \cdots, e_R = 0]$ a.s.o.
$*$ is whatever binary symbol
$_i$ index for the number of the current cell
b_i the i^th bit in the configuration (state) of the CA
$\vec{B}$ binary vector representing a state of the CA;
$\vec{B}_0$ - initial state
$C = (c_1, \cdots, c_m)$ a cycle of length m with decimal representation of the values of the CA 'output', c
C_{xy} correlation between x and y

C_{Pxy} Pearson correlation between x and y
$Cor(C_j, C_k)$ Pearson correlation between the cycles C_i and C_j
e_L value of the left boundary cell in the linear CA
e_R right boundary condition
L length (period) of a cycle
f global mapping of a CA
f^q the q^{th} iteration of f
$R(\cdot,\cdot,\cdot)$ is the local mapping named 'rule' (3-cell vicinity)
x_{-1} denotes x_{i-1}, x_0 denotes x_i, and x_1 denotes x_{i+1} when the value of i does not matter.

2. Basic Concepts and Definitions

Consider a linear, homogeneous, elementary CA with n cells and two extra 'boundary' cells establishing the boundary conditions (one of the four combinations of 0 and 1). Throughout this paper we are preoccupied only with finite CAs, that is, with CAs with finite number of cells. By convention, the boundary cells are not included in the count of 'automaton size', *i.e.*, the count is without these cells. However, when discussing mappings, the states of the boundary cells, even when fixed, are shown in the vector representing the variable of the map. Recall that a linear CA has linear topology, while a homogeneous CA has the same rule acting at each cell; the term elementary is used in the usual sense, meaning that the neighborhood of a cell is composed of three cells and that the rules of the cells correspond to one of the 256 rules in Wolfram's classification [25]. For further basic concepts on CAs, see also [2], [6], [10], [19].

Seen as an entity, the CA is a mapping from the set of binary vectors $(e_L, b_0, \cdots, b_n, e_R)$ to itself, where $b_i, i = 0, \cdots, n-1$ is the binary values of the cell i, and e_L, e_R are the binary left and right edge conditions, specified when defining the CA, $f(e_L, b_0, \cdots, b_{n-1}, e_R) = (e'_L, b'_0, \cdots, b'_{n-1}, e'_R)$.

When the boundary conditions are constant, $e_L = e'_L$ and $e_R = e'_R$; in case of periodic boundary conditions with period T, the n^{th} values of the boundaries, $e_L^{(n)}$, equal the first ones, $e_L^{(n)} = e_L, e_R^{(n)} = e_R$. The space $\{\vec{B}\}$ of all binary vectors $\vec{B} = (e_L, b_0, \cdots, b_{n-1}, e_R)$, with e_L, e_R specified, has the cardinal $|\{\vec{B}\}| = 2^n$. By converting the binary number $b_0, \cdots, b_{n-1}$ into its decimal one, p, the CA is a mapping with two parameters, e_L and e_R, from $\mathbf{N}$ to $\mathbf{N}$, $p' = f(e_L, e_R, p)$.

What is specific to the linear, homogeneous, elementary CAs is that f can be described locally by $b'_i = R(b_{i-1}, b_i, b_{i+1})$, where b_{i-1}, b_i, b_{i+1} is the neighborhood of b_i and $R(\cdot,\cdot,\cdot)$ is the local mapping named 'rule' and described solely in terms of logic or arithmetic operations; see also [**?**], [10], and [19], who use essentially the same terminology, but name f 'global mapping' and use the number of the rule in Wolfram's classification as an index for f.

Consider two CAs on the same input (definition) space, with the respective mappings denoted by f_1, f_2. Because the mappings are defined on the same space, both must have the same number of cells. The composed mappings $f_1 \circ f_2$, denoting

$f_1(f_2(\cdot))$, and $f_2 \circ f_1$ represent a composed CA with 'serried' connections between the two CAs.

A CA is said to be an equal length cycle (ELC) CA, briefly ELC-CA (also abbreviated ELCA [14]), when it produces for any initial condition (configuration), $\vec{B}_0$, a cycle and all its cycles have the same length. ELC-CAs are a special case of completely periodical CAs (CPCAs) introduced in the next Section. Because there are 2^n values in the definition domain (whose elements are named 'states of the CA'), the number of equal length cycles, when they occur, and the length of these cycles both must be powers of 2. Therefore, an n-cell ELC-CA may have 2^m cycles of length 2^{n-m}, for some value $0 \leqslant m$, see [9]. When $m = 0$, the single cycle is a maximal cycle.

3. Some Dynamic Properties of CAs

While the fundamentals of CAs have been extensively analyzed, see for example [1], [2], [7], [5], [6], [10], [13], the basics of ELC-CAs remain largely unexplored. In this Section, the class of completely periodical CAs is introduced, which includes the class of ELC-CAs, and a few of their dynamic properties are analyzed.

In the dynamics of CAs, two classes of CAs are of specific interest; these classes, defined below, are concerned in this paper. Throughout this section we refer only to finite, homogeneous, linear topology (one-dimensional) CAs, even when not explicitly mentioned.

Definition. A CA is named *completely periodical CA (CPCA)* when any point in its definition domain (any state) belongs to a cycle of period larger than 1.

This means that all its points (states) belong to nontrivial orbits, that is, to orbits with period larger than 1. Also, the reunion of all its orbits covers the entire domain of definition. Hence,

Property 1. The set of all orbits of a CPCA is a partition of the definition domain (set of all configurations) of the CA.

A similar concept to CPCA is that of multiple attractor cellular automata (MACA) [12]. However, CPCA is a subclass of MACA, because MACA is defined as 'transition graph of an MACA consists of a number of cyclic and non-cyclic states' [12], thus allowing transitory regimes.

Definition. (based on [9]) A CA is named *equal length cycle CA (ELC-CA)* when it is a completely periodical CA and all its cycles are equal.

The following properties are elementary, yet they are helpful throughout this study, especially in algorithms for systematic search for CPCAs and ELC-CAs.

The evolution of finite, homogeneous, linear CAs with a $2p + 1$ vicinity, $p \geq 1$, and with rules depending on left-hand variables only, that is $f(x_{i-p}, \cdots, x_{i+p}) = f(x_{i-p}, x_{i-p-1}, \cdots, x_i, *, *, \cdots, *)$, does not depend on the right-hand boundary, while the evolution of finite CAs with rules depending on right-hand variables only, $f(*, *, \cdots, *, x_i, x_{i-1}, \cdots, x_{i-p-1}, x_{i-p})$, does not depend on the left-hand boundary.

For example, CAs with rule 60 (3-cell vicinity), $f(x_{i-1}, x_i, x_{i+1}) = (\neg x_{i-1} \wedge x_i) \vee (x_{i-1} \wedge \neg x_i)$, shortly denoted by $\bar{x}_{i-1}x_i + x_{i-1}\bar{x}_i$ do not depend on the right-hand boundary; therefore, it produces the same results for boundaries $[0, 0]$ and $[0, 1]$, respectively for [1,0] and [1,1]. Also, CAs with rule 102, $\bar{x}_i x_{i+1} + x_i \bar{x}_{i+1}$, are independent of the left boundary. These two rules were also used for example in [17].

Property 2. A CPCA has no fixed point and no point on a transitory regime.

Assuming that there is a fixed point, that point (state) cannot occur in any cycle with period (length) larger than 1. Therefore, the CA cannot be a CPCA. Assume that there is a state $\vec{B}_r$ occurring in a transitory regime toward a cycle. Then, $\vec{B}_r$ cannot belong to another transitory regime or to any cycle; therefore, the CA is not a CPCA, which proves the property. □

In applications systematically searching for CPCAs or ELC-CAs, when a transitory regime is found for a CA, the respective CA should be rejected. (A CA with at least one transitory point cannot have periods (larger than 1) covering that point.)

Property 3. Any finite, homogeneous CAs with 3-cell vicinity rules depending on the left-hand variables only and satisfying the condition $f(0, 0, *) = 0$ or the condition $f(0, 1, *) = 0$ is not a CPCA for $e_L = 0$, whatever the right-side boundary condition is.

If the rule satisfies $f(0, 0, *) = 0$, then for any (finite, homogeneous) CA using that rule, $(0, \cdots, 0)$ is a fixed point; therefore, according to Property 2 the CA is not a CPCA. Also, if $f(0, 1, *) = 0$, $(0, \cdots, 0, 1)$ is a fixed point, thus the CA is not a CPCA. □

Example. The rule 60 satisfies both above conditions and thus has (at least) two fixed points for $e_L = 0$, thus it cannot produce CPCAs with $e_L = 0$.

Property 3 is easily extendable to rules with a vicinity larger than 3, for example with vicinity 5 and satisfying the condition $f(0, 0, \cdots, 0, 0, *, *) = 0$, which has fixed points $(0, \cdots, 0)$.

A finite, homogeneous CA with rules depending on left-hand variables only and satisfying one of the conditions in Property 3 cannot have a maximal period. In addition, any orbit including a point $(0_1, 0_2, \cdots, 0_p, 1, \cdots)$ where the index of 0 denotes its position in an $n > p$ cell CA, cannot have a length larger than 2^{n-p} (because only the last $n - p$ positions contribute to creating different results). Similarly for finite CA with rules depending on right-hand variables only.

Examples of application of the above properties are CAs with rules 60 and 102. Among others, these properties are used to reject cases of boundary conditions not discussed in this study.

4. Cycle Correlation

4.1. Definitions and comments

Although [18] proposes CAs for use in code division multiple access (CDMA) and [15] argues that the use of ELC-CAs increases randomness and entropy, we find that

such assertion may not apply for all cyclic CAs and ELC-CAs, because the cycles may be highly correlated. We show that low correlation between cycles may be achieved in some cases, specifically for some cases of the lowest number of cycles of given length. From the point of view of the dynamics, equal period cycles means that there is some minimal value $L = 2^m$ such that, for any initial condition $\vec{B}_0$, the L^{th} iteration of f, f^L, satisfies $f^L(\vec{B}_0) = \vec{B}_0$. This also means that for ELC-CAs there are 2^{n-m} basins of attraction and that they include the same number of points, with the basins of attraction constituting a partition of the domain of definition. We are interested how mixed are these basins of attraction; when they are strongly mixed, they may look more random and be more appropriate as a random pattern generator for testing, among others, the power of predictors or of classifiers.

For applications such as CDMA, a series of simple tests, such as runs of the partitions performed by the equal length cycles and the minimal distance between and correlation of symbolic sequences of the cycles are needed to ensure good quality of the cycles seen as code keys. We conjecture that the randomness of the partition of the definition (configurations) space of the ELC-CA, partition performed by the cycles is indicative on the chance that the cycles are decorrelated. Equivalently, the runs test of the partition may provide an indication of the randomness of the family of the cycles.

Definition. *The Euclidean distance between two cycles of equal length,* $C_{1L} = (c_{1,1}, \cdots, c_{1,h}, \cdots, c_{1,L})$ and $C_{2L} = (c_{2,1}, \cdots, c_{2,h}, \cdots, c_{2,L})$ is the minimal (Euclidean) distance between the cycles, when they are shifted one with respect to the other,

$$d^2(C_{1L}, C_{2L}) = min_k \sum_{h=1}^{L} (c_{1,h} - c_{2,h+k})^2$$

where the initial element of the cycles, $c_{1,1}, c_{2,1}$, are chosen arbitrarily, and by convention $c_{2,h+k} = c_{2,(h+k)} mod L$ (infinitely repeatable cycles).

In the definition, $c_{i,j}$ denote decimal numbers standing for the binary configurations of the CA. A normalized distance, $d^2_N(C_{1L}, C_{2L}) = \frac{1}{L} \times d^2(C_{1L}, C_{2L})$ is introduced for convenience. Then,

Property 4. For any ELC-CA with $L = 2^{n-m} < 2^n$, $d^2(C_{1L}, C_{2L}) \geqslant L$, $d^2_N(C_{1L}, C_{2L}) \geqslant 1$.

First notice that two different cycles cannot have any value in common, else they would not differ in any of their elements and order in the cycle; therefore, $(c_{1,h} - c_{2,h+k})^2 \geqslant 1 \, \forall k$. Hence, the sum of differences must be larger than the cycle length, $d^2(C_{1L}, C_{2L}) \geqslant L$. □

The use of the minimal distances between sequences is largely equivalent with the use of the maximal correlation value, when the sequences are repeated periodically and their cross-correlation is computed. The larger the distances between the cycles, the more the attractors (cycles) will be uncorrelated and thus useful in applications such as random pattern generators. Recall that the distance can be written as $d^2_k(C_{1L}, C_{2L}) =$

$\sum_{h=1}^{L}(c_{1,h}-c_{2,h+k})^2 = |\vec{c}_1|^2+|\vec{c}_2|^2-2\vec{c}_1\cdot\vec{c}_2$, where $vecc_1 = (c_{1,1}, c_{1,2}, \cdots, , c_{1,L})$ and similarly for $\vec{c}_2$, and $\vec{c}_1\cdot\vec{c}_2$ is the scalar product.

Several other characteristics of similarity of two cycles are easy to introduce. For example, one defines the maximal and average distance between EL cycles; moreover, one extends the distance to two ELC-CAs with the same EL cycle length as the minimal distance between their cycles (the two CAs do not need to have the same number of cycles or the same number of cells). Finally, the internal distance of the ELC-CA is defined as the minimal distance between all couples of cycles. To be suitable for PRNGs and other applications requiring randomness, an ELC-CA should have this distance as large as possible.

Definition. *The divergence based on distances (d-divergence) of a family of cycles (finite, periodical attractors)* $\{C_i\}_{i=1\cdots M}$ with the same period L is

$$D_d(\{C_i\}) = \frac{1}{LM(M-1)}\sum_{(j,k)\in M\times M} d(C_j, C_k).$$

In the definition, the factor $1/L$ accounts for the normalization with the length of the attractor (see Proposition 4), while the number of non-null distances is $M\times M-M$, hence the factor $1/M(M-1)$. Because of the equality $d(C_i,C_j)=d(C_j,C_i)$, each distance value occurs twice, for (i,j) and for $(j,i), i\neq j$. The divergence is null when all cycles are identical. The divergence is an indicator of the quality of a class of cycles for applications where numerous (highly) independent keys are needed, such as CDMA. When, in such applications, a minimal distance between the cycles, d_min, can be specified, another quality indicator of a family of cycles is

$$\mu(\{C_i\}) = M_{d_{min}}/|\{C_i\}|,$$

where $M_{d_{min}}$ is the maximal family of cycles that have between them a distance at least d_{min}.

Other measures of the similarity of the cycles are the maximal correlation value, for various shifts between them, $C_{xy} = \frac{max_k(\sum_{j=1}^{L} c_j c'_{j+k})}{\sqrt{\sum_{j=1}^{L} {c_j}^2}\sqrt{\sum_{j=1}^{L}(c'_j)^2)}}$ and the maximal Pearson correlation C_{Pxy} (obtained from the above formula by removing first the averages from the cycle values). The relationships between Euclidean distance, scalar product and Pearson correlation justify the use of the later one in:

Definition. *The divergence based on correlations (c-divergence) of a family of cycles* $\{C_i\}$ with the same period L is

$$D_c(\{C_i\}) = \frac{1}{M(M-1)}\sum_{(j,k)\in M\times M}(1-Cor(C_j,C_k)).$$

For example, for the 16 cycles produced by the ELC-CA with rule 60, $n=8$, boundary [1,*], the c-divergence is less than 0.03, showing that any other cycle can be easily derived with good precision from a specified one. ELC-CAs based on the rules 60 and 102 were first determined in [16].

The concepts of basin of attraction and phase diagram (log plot), along with other means [22] are useful in characterizing the dynamics of the iterations. An example of ELC-CA and of its basins of attraction on the natural set included in the real line is shown in Fig. 1. For the same ELC-CA, basins of attraction are shown in the phase plane $(B, f(B))$ after converting the binary vector in the corresponding decimal value. The ELC-CA uses rule 60, computed as $x_i = (x_{i-1} \wedge \neg x_i) \vee (\neg x_{i-1} \wedge x_i)$. For this rule, the effective neighborhood includes only two elements, at left. In the examples below, a small ($n = 4$) CA is used to maintain small EL cycles. With the boundary condition [1,1], the CA produces just two cycles ($m = 1$) of length $L = 2^{4-1} = 8$. The two cycles produce a partition on the set of definition. The randomness of the partition is a good indicator of the randomness of the cycles. The partition in Fig. 1 shows regularities in the grouping of the ordered elements in the two parts of the partition.

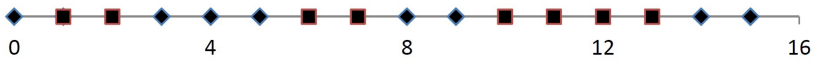


Fig. 1. Partition of the definition space by the two ELCs, for the ELC-CA $n = 4$, rule 60, boundary conditions [1,1].

Notice that, because of their intricacy, the planar mutual representations of the two cycles (with any time-shift) might provide could be useful as test patterns in determining the capabilities of classifiers and pattern recognition methods and algorithms.

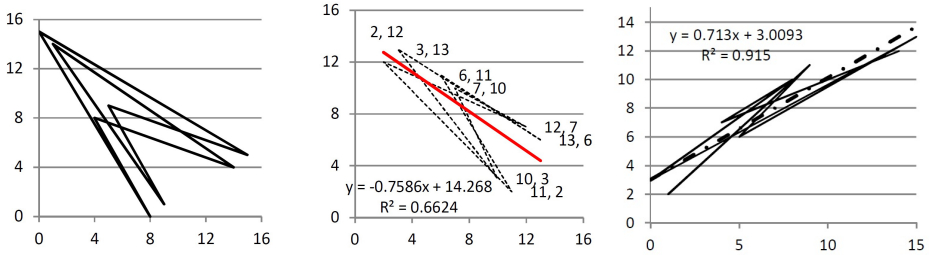


Fig. 2. Rule 60, boundary [1,1], and $n = 4$. a) Cycle (1,9,5,15,0,8,4,14) in the phase plane. b) The phase plot for the second cycle (6,13,3,10,7,12,2,11), and its best linear fit. c) The distribution of points of the phase plot for the two cycles when the second is shifted with 0,1, $\cdots$ 7 steps.

The diagrams in Fig. 2 are obtained according to the definition of the phase diagram for discrete processes, that is, the points have the coordinates $(c(t+1), c(t))$, where $c(t)$ is the decimal value corresponding to the configuration of the ELC-CA at time moment t.

4.2. Application: ELC-CAs with Rules 60 and 102

Consider the ELC-CAs with rule 60, as in the previous example. With the appropriate shift for achieving minimal distance, the coupling between the two cycles is visible in their joint graph, see Fig. 3. Also, the two cycles have close statistical properties, with the averages 7 and 8, and the standard deviations 5.20 and 3.87 respectively. Therefore, knowing one of the cycles may let one guess quite well the other cycle. This makes the two cycles of little use as independent pseudo-random series.

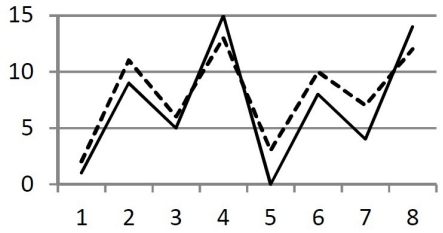


Fig. 3. Example of shifting of the cycles in Fig. 2 that minimizes the distance between them.

The strong linear correlation between a cycle and its shifted version shows that the cycle is easily predictable by a one-step-ahead linear predictor, thus, not random. The randomness of the partition, in case of two cycles, would require a uniform distribution of the points in the partition. Thinking that points in the first part are labeled a and points in the second by b. Then, as the parts are equal, each point has the same probability to be marked a or b. There are 7 runs in the above partition, $N_a = 8 = N_b$, $\mu = (2N_aN_b)/(N_a + N_b) + 1 = N_a + 1 = 9$. As there are about 20% fewer runs than expected, this indicates a trend, which is concordant with the linear trend in the phase diagram. Therefore, the distribution is not random.

With the appropriate shift for achieving minimal distance, the coupling between the two cycles is visible in their joint graph, see Fig. 3. Also, the two cycles have close statistical properties, with the averages 7 and 8, and the standard deviations 5.20 and 3.87 respectively. Therefore, knowing one of the cycles may let us guess quite well the other cycle. This makes the two cycles of little use as independent pseudo-random series.

For the two cycles discussed, $(1, 9, \cdots, 14)$ and $(2, \cdots, 12)$, with the best alignment of the cycles $(1 \leftrightarrow 2, 9 \leftrightarrow 11, \cdots 14 \leftrightarrow 12$, the correlation values are $C_{xy} = 0.971$ and $C_{Pxy} = 0.956$. The value $1 - C_{Pxy}$, which should be as close as possible to 1, is a good indicator of the quality of the cycles, that is, of their decorrelation.

4.3. Binary and symbolic sequences corresponding to one cycle

In communication applications, such as code division multiple access (CDMA), binary sequences are transformed in sequences of -1 and 1, with -1 corresponding to

0 in the binary sequence. If one wishes to use ELC-CAs for generating modulating sequences for code division multiple access, one is concerned with the correlation of the -1,1 sequences obtained from the whole cycles of an ELC-CA. For example, the first cycle of the ELC-CA with $n = 4$, rule 60 is (1-9-5-15-0-8-4-14), which converted into binary produces the binary sequence:

$$C_{1B} = (0,0,0,1,1,0,0,1,0,1,0,1,1,1,1,1,0,0,0,0,1,0,0,0,0,1,0,0,1,1,1,0)$$

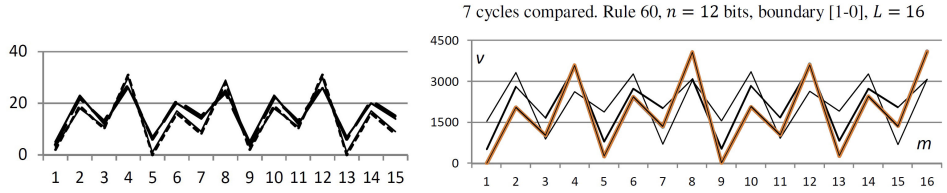
The binary sequence converts to the -1,1 type sequence as

$$C_{1D} = (-1,-1,-1,1,1,-1,-1,1,-1,1,-1,1,1,1,1,1,-1,-1,-1,-1,\cdots$$
$$\cdots 1,-1,-1,-1,-1,1,-1,-1,1,1,1,-1),$$

with the graphical appearance shown in Fig. 4. The sequences produced by the 4-cell CA have 32 elements as the cycle length is 8 and each digit is represented by 4 bits.

Fig. 4. The representation of the first complete cycle of ELC-CA $n = 4, R60$, as binary sequence, starting with the smallest value in the cycle, 1, and using the representation $0 \to -1, 1 \to 1$.

Converting in the same way the second cycle, and considering the sequences as vector representations, the scalar product of the vectors of the two cycles is indicative on their orthogonality. The scalar product results, however, $C_{1D}C_{2D} = 8$, showing the non-orthogonality. Similarly, the scalar product of the vectors $C_{1B}C_{2B} = 10$, showing a matching of the ones at 10 places in the two sequences. Also, the number of 1s, namely 14, and 0s $(32 - 14 = 18)$ in the sequence are significantly different (unbalanced sequence).



(a) **(b)**

Fig. 5. (a) The four cycles of length 8 obtained with different initial conditions, CA with $n = 5, L = 8$, R60; (b) Seven (various initial state) of the cycles of the same CA compared (R60, $[1, *], n = 12, L = 16$. Some cycles are so close that they are seen as superposed.

Some ELC-CAs produce less correlated cycles. For example, the ELC-CA with $n = 4$, rule 102, produces the cycles 14, 3, 4, 13, 6, 11, 12, 5, and 2, 7, 8, 9, 10, 15, 0,

1, which, when the second is rearranged to start with 8, have a Pearson correlation of only 0.0125, while the maximal correlation, 0.537, is obtained when the second cycle starts with 15. The two cycles and the corresponding phase diagrams are shown in Fig. 6. The minimal distance between the two cycles is 184, also showing their significant dissimilarity, with an average distance between values in the cycles larger than $\sqrt{184/8} \approx 5$, almost one third of the range of values. The phase diagrams are also very different, see Fig. 6.

Correlations between the orbits of ELC-CA for $n = 5$, rule 60, $L = 8$ are shown in Fig. 7, left. The four cycles (A,B,C,D) are highly correlated, with all series with an inter-correlation coefficient larger than 0.95 (after suitable shifting).

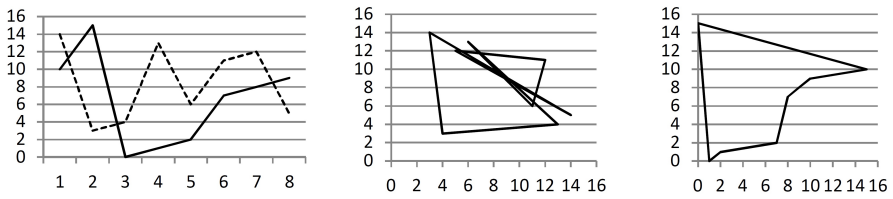


Fig. 6. Cycles for the ELC-CA $n = 4$, rule 102.

The low degree of correlation between the two cycles is reflected in their different evolution: the cycle 0, 1, 2, 7, 8, 9, 10, 15 is monotonic, a rarely seen property for what we experimented, while 3, 4, 13, 6, 11, 12, 5, 14 is not.

On the other hand, there is no guarantee that increasing the number of cells would help produce better decorrelated series. For example, using $n = 5$ and rule 60, with proper adjustment of the series 'shifts', one obtains the graph in Fig. 7, showing strong dependence between the cycles, which have the same symbolic representations (*udududududududud*, where *d* means *down* and *u* means *up*).

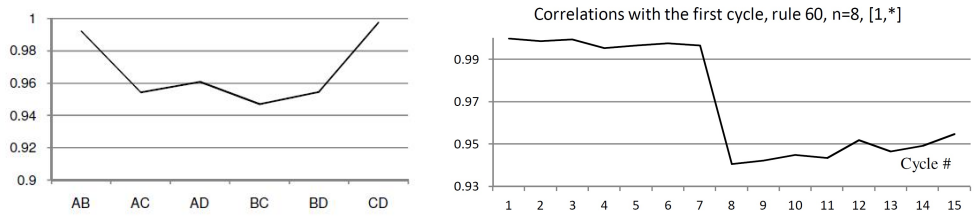
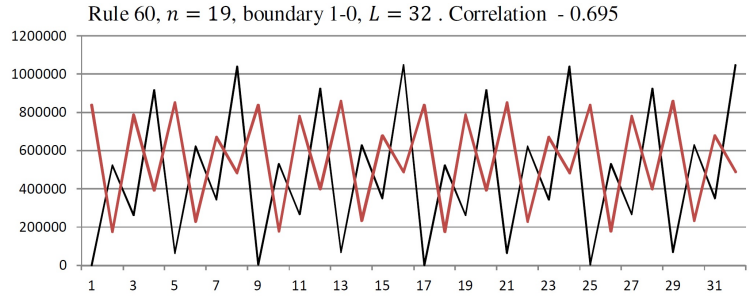


Fig. 7. Left: The correlation values for the four cycles of the CA with R60, $n = 5, L = 8$. Right: Correlations of the other cycles with the first cycle, for the CA with R60, $n = 8, [1, *], L = 16$.

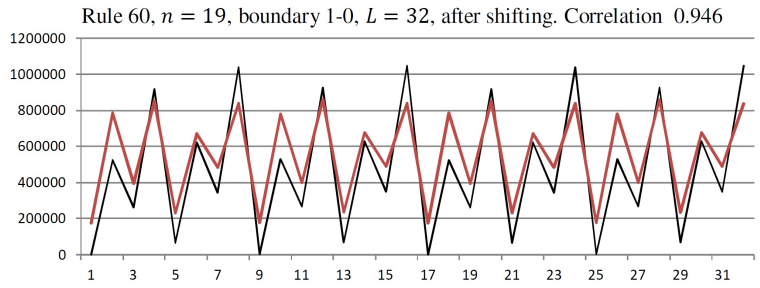
The use of the normalized symbolic distance $d_s = min_k \frac{1}{L} \sum_{j=1}^{L} (s_j - s'_{j+k})$, where s_j and s'_j are the respective symbols of the corresponding symbolic sequences and the difference is zero when the symbols are identical and 1 when they differ, is a good indicator of the randomness in the overall set of cycles, as the correlations of the cycles makes the whole set of patterns less random.

The difficulties of selecting good cycles produced by ELC-CAs are not specific to small ELC-CAs (small n values). For example, an ELC-CA with $n = 12$ produces cycles of length 16 in a large number ($2^{12-4} = 2^8 = 256$ cycles), but many of them are virtually identical. For example, in Fig. 8, the cycles starting with 0, 1, 2, 3, 4 cannot be seen as distinguished, and the symbolic series of two other cycles with large starting values randomly selected (412, 682) are identical with those for small starting values.

High cycle correlations occur even for very large number of cells in this type of CA. For example, using a CA with rule 60, $n = 19$, boundary [1,0], the cycle length is $L = 32$; for two cycles selected at random (see Fig. 9, left) the correlation is - 0.695, but after suitable shifting the correlation becomes 0.946 (Fig. 9, right). Thus, using a large number of cells does not significantly improves the cycle decorrelation. In fact, the large correlation values are due to the rules used in the exemplified CAs, as will be proved elsewhere [24].



(a)



(b)

Fig. 8. ELC-CA with $n = 19$, boundary [1,0], $L = 32$. (a) - before shifting; (b) - after shifting.

When there are more than two cycles, *e.g.*, for $4 \leq n < 8$, the number of low correlation orbits is 2 out of 2,4,8,16; similarly, when $8 \leq n < 16$, the number of low correlation cycles is only 4. We conjecture that cycles with low inter-correlation are obtained only when the smallest number of cycles is generated, that is, when n is equal to the smallest value that increases the length of the cycle.

For various other ways of assessing the properties of the dynamics of systems in engineering applications, see [22]. More details on the properties of the cycles of several types of ELC-CAs are given in [24]. For further interpretations of classical tests for randomness, see [21].

5. Hybrid CAs based on ELC-CAs

The reason for introducing hybrid (composed) CAs that may be candidates for having ELCs is that mixing EL cycles from different CAs might produce highly decorrelated ELCs. Various concatenation solutions for CAs have been proposed, *e.g.*, [8], [11], but not for ELC-CAs. One might hope that just concatenating two cycles from the same CA, whereas the cycles are obtained by using different initial conditions. However, we have shown that cycles of the same ELC-CA with different initial conditions are usually strongly correlated and thus are unsuitable for the purpose. This justifies the search for 'hybrid' CAs with ELC.

Consider two CAs with the same number of cells coupled in a feedback loop, such that the output of the first is input at the next step to second and vice-versa. The boundary conditions of the two CAs are not enforced to be the same. Such a construction is described by the composed mapping $f_1 \circ f_2$, where $p' = f_1(e_L, e_R; p)$, $q' = f_2(e'_L, e'_R; q)$, $p' = f_1(e_L, e_R; q')$, where e_L, e_R are seen as parameters. Such a system always has a periodical orbit; yet, its cycle is not necessarily a combination of two EL cycles of the two mappings in the combination. For example, if f_1 is an ELC-CA with $n = 4$ and rule 60 with edges [1,1], and f_2 is the ELC-CA with the same number of cells, rule 102 and edges [0,1], the period has the form 2, 10, 7, 12, 4, 14. This CA is not an ELC-CA because there is a short transitory regime, 15, 8, 0, 9, 3, 10 for the cycle above (recall that ELC-CAs cannot have transitory regimes, see Property 2), see the graph in Fig. 9.

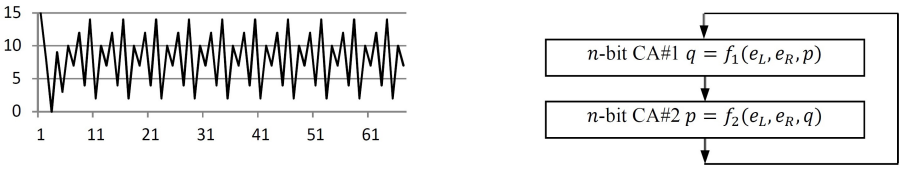


Fig. 9. Left: Sequence generated by a hybrid, staked CA using two ELC-CA with $n = 4$, one CA with rule 60 edge [1,1] and the other with rule 102 edge [0,1]. Right: Sketch of the configuration.

When the rule applied at time moment 1 and at all odd moments t is rule 60, boundary [1,1], while for t even the rule is 102, boundary [0,1], the change in rule is performed according to

$$if\ t\, mod\, 2 = 0,\ x'_i = (x_i \wedge \neg x_{i+1}) \vee (\neg x_i \wedge x_{i+1})$$

$$else\ x'_i = (x_{i-1} \wedge \neg x_i) \vee (\neg x_{i-1} \wedge x_i)$$

One can combine two ELC-CAs with different numbers of cells, for example the smaller CA seeing only the last bits from the larger one, while the larger one sees the output of the smaller as the less significant bits, but preserves the previous values of its most significant bits.

One can further envisage switched CAs, involving two ELC-CAs with the same periods (length of cycles) and number of bits. Essentially, the switched CA system has a counter and a comparator added to two ELC-CAs; for the duration of a period, one ELC-CA operates and produces the output; then, for the next period, the output to the other ELC-CA, which gets as input, at the first moment of being active, the last state produced by the first ELC-CA.

6. Discussion and Conclusions

The advent of IoT rises several technical challenges, including the expansion of the wireless communication to serve connected 'things'. Various communications techniques, such as spread spectrum communications, use digital keys to allow the communication only between the key holders, while preventing interference with other channels. These techniques require simple devices able to generate a very large number of keys, yet preserving desirable properties of the keys, such as orthogonality or low correlation.

The use of ELC-CAs in PRNGs and related applications, including communications (CDMA) is justified only when minimum distances between cycles of ELC-CAs are high (much larger than the distance value given by Property 4, that is, the normalized distance should be $\gg 1$). Equivalently, the correlation of the cycles should be low. Moreover, the number of runs in the partition should be low and the linear predictor in the phase space should have a low confidence (the R coefficient value).

Because ELC-CAS with R60 (same for others, *e.g.*, for R102) does not depend on the right edge value; thus, the right boundary can be fixed, or periodical, or even randomly variable, or can lack. Therefore, in software implementations there is no need to use memory and code (program memory) for that boundary. Maybe even more important for the cost reduction of the chips (hardware) potentially using ELC-CAs, is the fact that for there is no need for connection to ground (0 logic) or to the power line (1 logic), for the edge implementation; therefore, there will be fewer traces on the ELC-CA chips, and less power used by the last cell. Possibly, the ELC-CA will behave as the boundary is stochastic, due to the missing connection will pick up any environmental electromagnetic noise; but that has no effect on the operation, because the rules used do not take into account the right (or left) side neighbor.

The quality of the cycles produced by an ELC-CA may vary widely, both as the coupling (correlation) between the cycles of the ELC-CA and as properties of the individual EL cycles. The use of 'staked' CAs as proposed in the Section 5 may help produce larger and less correlated cycles, but in many cases the generated cycles are shorter than the ones of the original CAs that are staked.

Finally, from the applicability point of view, the high correlations between cycles of some CAs, as discussed in this article, shows a strong limit in the modeling capabilities

of such ELC-CAs: their spectrum of produced models is quite limited.

Concluding, special attention should be paid to the degree of independence and decorrelation of the cycles produces by CP-CAs in view of applications.

Acknowledgments. The author acknowledges the collaboration of Mr. A. Mitra on [16] and for comments on a preliminary version of this paper.

Conflict of interest. This research was not influenced by the grant support; the author declares no known conflict of interest.

References

[1] AKIN H., *The Entropy of Linear Cellular Automata with Respect to Any Bernoulli Measure*, Complex Systems, vol. **18**, pp. 237–244.

[2] ALLOUCHE J.-P., SKORDEV G., *Remarks on permutive cellular automata*, J. Comp. & System Sci., Vol. **67**, no. 1, Aug. 2003, pp. 174–182.

[3] ANGHELESCU P., IONITA S., SOFRON E., *Encryption Technique with Programmable Cellular Automata (ETPCA)*, J. Cellular Automata, Vol. **5**, no. 1–2, pp. 79–105, 2010.

[4] ANGHELESCU P., SOFRON E., RINCU C.I., IANA V.G., *Programmable Cellular Automata Based Encryption Algorithm*, CAS 2008 - Int.Semiconductor Conf., Proc., Sinaia, Romania, Oct. 13–15, 2008, pp. 351–354.

[5] BETEL H., FLOCCHINI P., *On the relationship between Boolean and fuzzy cellular automata*, E Notes Theor. Comp. Sci., Vol. **252**, 2009, pp. 5–21.

[6] DELORME M., MAZOYER J., OLLINGER N., THEYSSIER G., *Bulking II: Classifications of cellular automata*, Theor. Comp. Sci. **412** (2011) pp. 3881–3905.

[7] FATES N., *Experimental study of Elementary Cellular Automata Dynamics using the Density Parameter*, Discrete Mathematics and Theor. Comp. Sci. AB(DMCS), 2003, pp. 155–166.

[8] FUESTER-SABATER A., CABALLERO-GIL P., *Concatenated automata in cryptanalysis of stream ciphers*, Proc. ACRI'06 Proc. 7th Int. Conf. Cellular Automata for Research and Industry, pp. 611–616, Springer-Verlag Berlin, Heidelberg 2006.

[9] GHOSH S., BACHHAR T., MAITI N., MITRA I., CHAUDHURI P., *Theory and Application of Equal Length Cycle Cellular Automata (ELCCA) for Enzyme Classification*, in Bandini S. et al. (Eds.), *Cellular Automata*, Vol. **6350**, Lecture Notes Comp. Sci., pp. 46–57, Springer-Verlag Berlin, Heidelberg, 2010.

[10] KARI J., *Theory of cellular automata: A survey*, Theoretical Computer Science, vol. **334** (2005), no. 1–3, 15 Apr 2005, pp. 3–33.

[11] KUTRIB M., *Pushdown cellular automata*, Theor.Comp.Sci., Vol. **215**, no. 1–2, 28 Feb. 1999, pp. 239–261.

[12] MAJI P., SHAW C., GANGULY N., SIKDAR B. K., CHAUDHURI P. P., *Theory and Application of Cellular Automata For Pattern Classification*, Fundamenta Informaticae **58** (2003), pp. 321–354, IOS Press.

[13] MARTIN O., ODLYZKO A., WOLFRAM S., *Algebraic properties of Cellular automata*, Commun. Math. Phys., vol. **93**, pp. 219–258, 1984.

[14] MITRA A., KUNDU A., DAS C., *Cost effective PRNG using ELCA: A BIST application*, 1st Int. Conf. Autom., Control, Energy & Systems ACES, Feb 1–2, 2014, Hoogly, India, pp. 1–6.

[15] MITRA A., *Investigating Scopes for Automata Based Designs Targeting Image Security in Health System*, Proc. 5th IEEE Int. Conf. E-Health and Bioeng. EHB-2015, Iasi, Romania, Nov. 19–21, 2015.

[16] MITRA A., TEODORESCU H.N., unpublished.

[17] NANDI S., KAR B.K., PAL CHAUDHURI P., IEEE Trans. Computers, vol. **43**, no. 12, Dec. 2004, pp. 1346–1357.

[18] SARKAR S.K., MANNA G.C., SINGH S.S., DATTA T., NASKAR P.K., *CDMA Technology based Reliable Wireless Mobile Communication System on a Single Chip using Cellular Automata Concept*, J. Engng., Comp. and Architecture, Vol. **1**, no. 1, 2008.

[19] SAHOO S., CHOUDHURY P., *Issues on drawing the state transition diagram for arbitrary cellular automata*, `http://arxiv.org/ftp/arxiv/papers/0811/0811.1513.pdf`

[20] SOUZA-E-SILVA H., SAVINO W., FEIJÓO R.A., VASCONCELOS A.T.R., *A cellular automata-based mathematical model for thymocyte development*, PLoS ONE, vol. **4** (12), 2009: e8233. doi:10.1371/journal.pone.0008233.

[21] SYS M., RIHA Z., MATY V., MARTON K., SUCIU A., *On the Interpretation of Results from the NIST Statistical Test Suite*, Romanian J. Information Science and Technology, Vol. **18**, no. 1, 2015, pp. 18–32.

[22] TEODORESCU H.N., *Characterization of nonlinear dynamic systems for engineering purposes - a partial review*, Int.J. General Systems, Vol. **41**, No. 8, pp. 805–825, 2012.

[23] TEODORESCU H.N., *Type-D Fuzzy CAs for Medical and Social Sciences*, Proc. 5th IEEE Int. Conf. E-Health and Bioengineering EHB-2015, Iasi, Romania, Nov. 19–21, 2015.

[24] TEODORESCU H.N., *Analysis of the Behavior of Simple and Composed Equal Length Cycle Cellular Automata with Applications* (submitted).

[25] WOLFRAM S. (Ed.), *Theory and Application of Cellular Automata*, Reading, MA: Addison-Wesley, 1986.