NEW BOUNDARY CONDITIONS FOR 1-D CELLULAR AUTOMATA

K. Salman[1]

Abstract**—** *Cellular Automata (CA) have been suggested for Pseudo Random Sequence (PRS) generation as a new but characteristically different alternative to the classical Feedback Shift Registers (FSR). One-dimensional (1-D) as well (2-D) two-dimensional CAs has been used with remarkable success in generating PRS that have good statistical features. This research presents a new but notwithstanding remarkably simple design for PRS generation. The outputs of linear FSRs (LFSRs) act as continuous inputs or perturbations to the two extreme boundary cells of a 1-D CA. The results of sufficiently long evolutions time steps show superior randomness features as compared to the periodic boundary conditions classical approach. The output strings of such CAs under certain chaotic local transition rules have passed the stringent Diehard statistical battery of tests. The design is a viable candidate for parallel PRS generation, has strong correlation immunity, good order of asymptotic complexity and is inherently amenable for VLSI implementation.*

Index Terms — Boundary Conditions, Cellular Automata, Diehard Tests, Linear Feedback Shift Registers, Random Number Generators

INTRODUCTION

Both LFSRs and CAs have been used extensively in a wide area of applications, particularly random number generation for Mont Carlo simulation, communications, cryptography and network security [1]-[8]. LFSRs, although, are simple in structure and design were proven to have comparatively weak statistical features when utilized in the production of pseudo random numbers [2]. The weakness can be attributed to the linearity of the exclusive-or function used in the feedback network. Additionally, non-linear feedback shift registers are difficult to construct for long seed lengths and have their problems as well [1]. On the other hand, a uniform 1-D CA, where one rule is implemented throughout the spatiotemporal evolution of the CA, have shown unique and useful characteristics, and have been suggested by [3] and [4] and others for use in random number generation. A notable impediment however, is the input to the boundaries of the CA, when it is confined to a limited span. Hence, a long string of output will render the CA overly unsuitable. However, a shorter span implies a constant span length K and therefore a fixed and limited number of cells. Thus, inputs are needed to feed the two extremities of the CA. One approach attempted to solve this problem is to make the CA evolve in a continuous loop (referred to as autonomous or periodic), in which case the peripheral cells (i.e. the last and the first extreme cells) are made adjacent to each other. An alternative technique is to feed the peripheral cells with fixed inputs. Figure 1 depicts some common fixed boundaries from $GF(2)$ used in the literature.

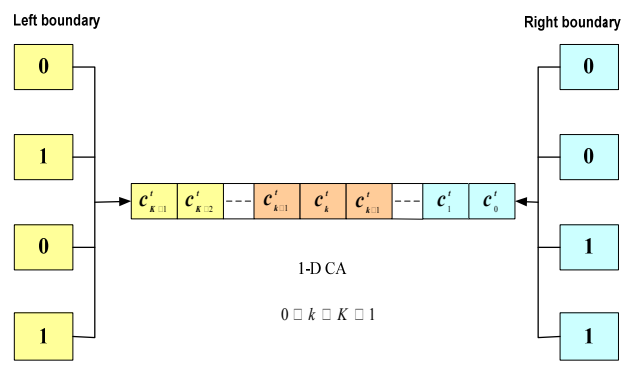


Figure 1. Common 1-D CA fixed boundaries.

All these methods running under chaotic rule 30 on uniform 1-D CAs have produced much shorter periods than a comparative LFSR and drastically failed the Diehard battery of tests. This paper reports the findings of a new method whereby a pair of uncorrelated LFSRs are used to generate the two boundary conditions. With this design the output string of the concatenated cells of the whole sufficiently large CA span K evolving for a time evolution steps $T \geq 2^K$, has passed the Diehard battery of tests and produced attractive parallelism, correlation properties, and strong cryptographic complexity.

[1]Middle Tennessee State University, Murfreesboro, TN 37132, (USA), ksalman@mtsu.edu

PRELIMINARIES

For the purpose of this paper we will restrict our attention towards one dimensional cellular automaton. The cells are arranged on a linear finite lattice, with a symmetrical neighborhood of three cells and radius $r = 1$. Each cell takes its value from the set $G = \{0,1,\ldots,p\}$ and we let $p = 1$ rendering the CA binary and therefore referred to as Elementary Cellular Automata (ECA). All cells are updated synchronously and they are, therefore restricted to local neighborhood interaction with no predetermined global communication. The CA will evolve according to one uniform neighborhood transition function, which is a local function (rule) $f : G^{2r+1}$ a G where the CA evolves after certain number of time steps *T*. Out of a total of $G^{G^{2r+1}}$ rules we use rule 30 as suggested by [3] and adopt the numbering scheme attributed to [3] and [4]. It follows that a 1-D CA is a linear register of K, K Î ¥ memory cells. Each cell is represented by c_k^t, where k Î K, t Î T and $T \in$ □ describes the content of memory location k at time evolution step t. Since $G = 2$ the cell takes one of two states from $GF(2)$. This implies the applicability of Boolean algebra to the design over $GF(2)$. A minimum Boolean representation of Rule30 in terms of the relative neighborhood cells is $c_k^{t+1} = c_{k-1}^t$ Å $(c_k^t + c_{k+1}^t)$, where 2 £ k £ $K - 2$, as depicted in figure 2. Furthermore, since the CA is actually a finite state machine then the present states of the neighborhood $(c_{k-1}^t, c_k^t, c_{k+1}^t)$ of cell c_k^t, at time step t and the next state c_k^{t+1} at time step $t+1$, can be analyzed by the state transition table and the state diagram depicted in figure. It can be seen from above that in order to evolve from the present time step t to the next time step $t+1$, each cell at lattice location k would require the present state of itself c_k^t as well as the present state of the other two cells in its neighborhood c_{k-1}^t and c_{k+1}^t. Therefore, if the CA is allowed to expand freely, left wise and right wise the total number of cells at each time step $t+1$, say *M* would need *M+2* cells at time step t. Hence, if the CA is unbounded then an unpractical span size of the CA would be required in order to accumulate a reasonable size of random numbers. Therefore, it is imperative that the CA has to be bounded. The open literature is rich with research on fixing the size of the CA and provides data for the extreme cells of the bounded CA. Figure 1 gives a brief account of some common fixed boundary conditions while figure 5 categorizes the boundary conditions to include the new boundary condition proposed in this paper using LFSR as a new source for boundary conditions. The autonomous category refers to the common so-called periodic boundary conditions and makes the extreme cells of the CA adjacent, as illustrated in figure 6. The resultant cellular automaton becomes circular at a single time step but renders the whole CA as cylindrical. The expression for the extreme left and right cells at time step $t+1$ are, respectively $c_0^{t+1} = c_1^t$ Å $(c_{K-1}^t + c_0^t)$ and $c_{K-1}^{t+1} = c_{K-2}^t$ Å $(c_{K-1}^t + c_0^t)$. The miscellaneous category includes either some *ad hoc* permutation of the fixed boundaries, some fixed sequence of inputs or even tapping the CA at variable points along the span during the time evolution. The published results of these different types of boundary conditions produced poor results when used as a source of generating random numbers. In this paper we are proposing a new source for the boundaries. We have used the well established LFSR as the source of inputs to the extreme cells of the fixed 1-D CA, as shown in figure 5. A LFSR of span N memory cells can be described by the following simple recurrence equation, $L_0^{t+1} = a_0 L_0^t$ Å $a_1 L_1^t$ Å ××× $a_i L_i^t$ Å ××× Å $a_{N-1} L_{N-1}^t$ Where a_i Î $GF(2)$. The choice of a_i are exactly the coefficients of a primitive polynomial of degree N. The extreme cells of the new design at time step *t+1* can now be described by $c_{K-1}^{t+1} = c_{K-2}^t$ Å $(L_0^t + c_{K-1}^t)$ and $c_0^{t+1} = R_0^t$ Å $(c_0^t + c_1^t)$

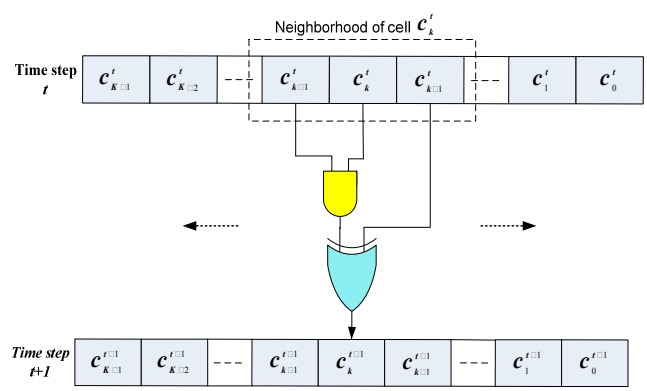


Figure 2. Rule 30 operating on two time steps.

Present State c_k^t	Neighborhood $c_{k-1}^t - c_{k+1}^t$			
	0-0	0-1	1-1	1-0
0	0	1	0	1
1	1	1	0	0

Next State

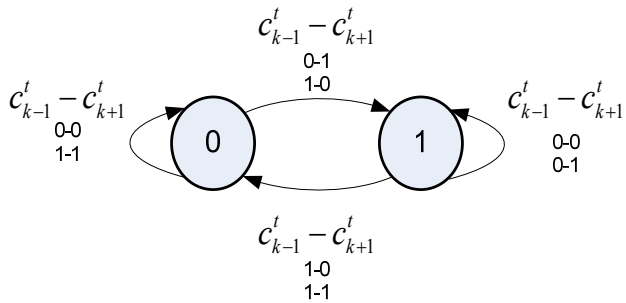


Figure 3. State machine analysis of Rule 30.

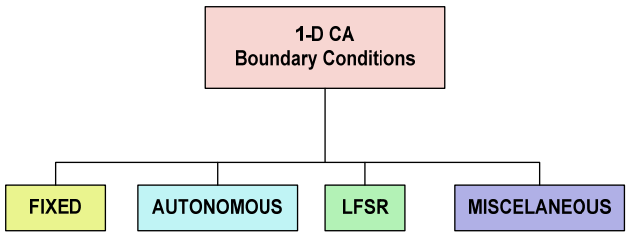


Figure 4. Categorization of a fixed span 1-D CA boundary condition sources.

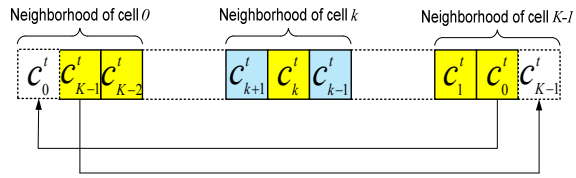


Figure 5. Autonomous (or periodic) boundary conditions at time step *t*.

RESULTS

In order to test the statistical properties of the new proposed design, we developed a research platform that encompasses a suite of programs emulating all known types of boundary conditions for wide range of spans for both the CA and the LFSRs, as described in figure 6.

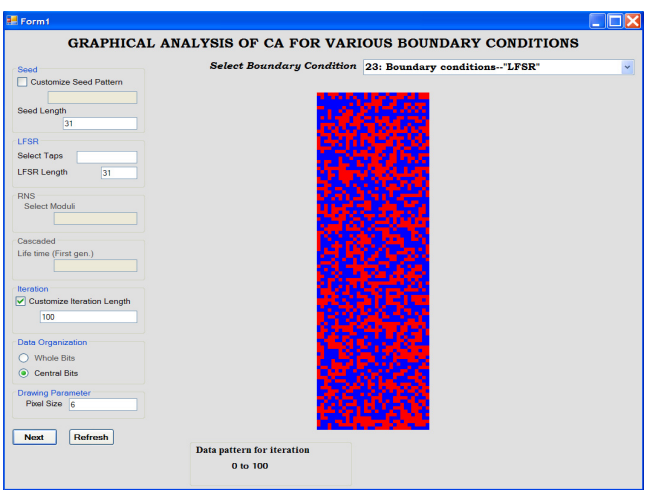


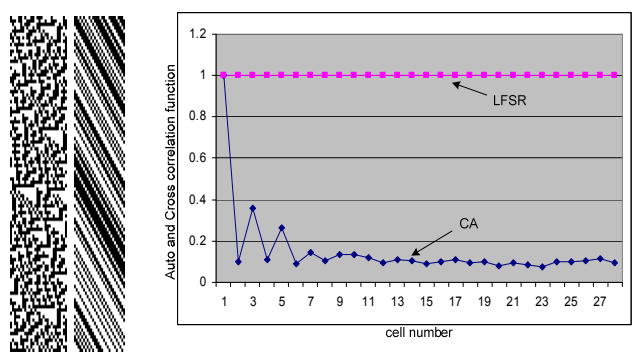
Figure 6. Spatiotemporal output of 1-D CA span 31-bit with two LFSRs as boundary inputs source of the same span. One hundred time steps are shown.

Included in this platform are selections that compute outputs for different types of random number generators for future work and are not presently described in this paper. The snapshot shown in the figure is a spatiotemporal image for a run of an LFSR bounded 1-D CA running under rule 30 with the span length of $K = R = L = 31$ bit and 100 time steps evolution, starting with a random seed. We will include snapshots of results obtained for representative runs on the Diehard battery of tests [9], which has been adopted in this paper due to its well established stringent requirements on the statistical randomness of the output string. Due to the restrictions imposed by this test the CA span K has to be at least 27-bit long evolving for a minimum of 2^K time steps. Table 1 shows the results of running the diehard tests on the autonomous boundary conditions for spans 32, 33, up to 512. The CA has not been able to pass all the tests even for a span of 256-bits. The results of running the diehard tests on the fixed boundary conditions have totally failed and therefore not worth reporting here. When running the CA using two LFSRs of span 3-bit each as the boundary conditions for various and increasing spans, the Diehard results were only marginally improved although the spatiotemporal image showed better randomness. It is clear that such boundary conditions will give better results than fixed boundary conditions but do not show vast improvement over the periodic boundary condition. However, when the LFSRs spans increased to 15-bits for

both registers some improvement became noticeable. Excellent results were obtained when the span of the LFSRs were increased to match the span of the CA. The CA has passed all tests with extremely superior P-values, as shown in table 2.

CONCLUSIONS

The string of contiguous stream data collected from the evolution of the 1-D CA for the center cell of various boundary conditions were tested by the 15 diehard battery of tests. The various fixed boundary conditions failed the diehard tests almost completely and were considered unworthy reporting. The autonomous boundary conditions have shown far better statistical properties than the fixed boundary conditions. However, it still falls far below the minimum requirements of the diehard tests for reliable considerations in producing dependable random numbers even for long spans of the CA (512-bit). When the boundaries were fed from LFSRs results did not improve significantly until the span of the LFSRs were comparable to that of the CA. The results steadily improved up to the upper bound when the two spans were comparable. It can be concluded that the new approach can produce random numbers even at modest size of the CA (i.e. 27-bit). More in depth study of the results show that the new approach produced superior *p*-values than the best of the autonomous results. Further assertion of the diehard results are also apparent from visual inspection of the spatiotemporal output as can be seen from figure 6. It is easy to expect that the fixed boundary conditions cause a CA running under Rule 30, which is considered a chaotic rule and belongs to group III, [4] and [5], to evolve into Group I or II (i.e. point attractors or limit cycles with extremely small periods. Therefore, such boundary conditions preclude these CAs from achieving strong random number generators. The autonomous boundary conditions, on the other hand gave better results which is indicative of better distribution during CA evolution. However, the periods of this type were far lower than the maximum length obtainable from LFSRs. The proposed design has an added favorable feature when considering the initial seeds. It is clear that all the possible 2^K *K-tuples* can be used as seeds including the all 0's and all 1's that usually yield quiescent states when used with the periodic boundary conditions. This is not possible with any other known boundary conditions including the autonomous type. All the tests were performed using a single one as the initial seed. This is admittedly not the case in a practical situation. Some patterns were observed during the initial evolution of the CA but did not persist. Although these initial patterns did not negatively impact the diehard tests it was found that avoiding the use of trinomials for the LFSRs and replacing them with primitive polynomials of better distribution of the coefficients managed to remove these patterns. One salient feature of the design is the almost total destruction of the cross-correlation between different cells as shown in figure 7(a). The strong correlation is an inherent feature of LFSRs that can be observed as maximum and constant between any two cells of the LFSR and as linear patterns on the diagonal ridge between the outputs of the LFSR cells, figure 7(b). An immediate consequence is the ability to use the CA as a parallel source of pseudo random numbers that can be considered a strong candidate for parallel data compaction (signature analysis) in VLSI testing [8]. This is justified since the structure as depicted in figure 8 presents a simple memory-based and inherently parallel design that is amenable to large scale integration. Inspection of rule 30 reveals that the function is surjective. Since reversibility implies bijection, it follows that the proposed system is not clear cut reversible. Hence analytical techniques may not be available to adequately and inversely describe the spatiotemporal data evolution in at most polynomial time. For a LFSR of span N, there are (2^N-1) N-tuple words as seeds. The two LFSRs are uncorrelated and running independently and synchronously, hence the effective input computational complexity from these registers to the CA would be $(2^N-1)^2$. The 1-D CA of span K can be initialized with a total of 2^K K-tuple words as initial seeds. There are a total of 2^{2^3} rules, which is the rule space of a1-D CA. Thus the computational asymptotic complexity of the system is $O((2^N-1)^2 \cdot 2^{2^3} \cdot 2^K) \approx O(2^{3K})$ for $K \ll N$, as compared to $2N$ for the LFSR and $O(2^K)$ for a 1-D CA with autonomous boundary conditions.



(a) (b)

Figure 7. Spatiotemporal images of CA and LFSR with span length of 28-bit (a) and correlation properties (b).

Table 2. Diehard tests results for CA of variable spans and with 2LFSRs for boundaries of the same as the CA spans.

Diehard Tests	Span Length K				
	27	28	32	64	128
T_1	0.242	0.43	0.3046	0.2398	0.2697
T_2	0.0744	0.4376	0.1128	0.5284	0.2123
T_3	0.8442	0.6365	0.3417	0.3317	0.5543
T_4	0.4688	0.47	0.4323	0.2713	0.0628
T_5	0.52235	0.4697	0.5166	0.4421	0.5454
T_6	0.4755	0.32	0.5486	0.5584	0.4654
T_7	0.6092	0.485	0.3151	0.4642	0.6849
T_8	0.5581	0.5083	0.4601	0.5009	0.6135
T_9	0.2253	0.6181	0.6947	0.5722	0.5413
T_10	0.8818	0.2469	0.9452	0.728	0.0897
T_11	0.7111	0.3403	0.1944	0.7524	0.5147
T_12	0.456	0.423	0.9646	0.9847	0.1522
T_13	0.3985	0.1387	0.2413	0.1063	0.3202
T_14	0.3026	0.6276	0.1753	0.3521	0.4801
T_15	0.343	0.5539	0.7578	0.4428	0.4845
Summary	**15 pass 0 fail**	**15 pass 0 fail**	**15 pass 0 fail**	**15 pass 0 fail**	**15 pass 0 fail**

p-values

Diehard Tests	Span Length K					
	32	33	64	128	256	512
T_1	0.4913	0.6089	0.4871	0.5683	0.5976	0.7166
T_2	1	1	1	1	1	1
T_3	0.759	0.7895	0.5035	0.4525	0.643	1
T_4	1	1	1	1	1	1
T_5	1	1	0.4973	0.5195	0.5777	0.7068
T_6	1	1	0.804	0.7769	0.4856	0.7756
T_7	1	1	0.999	1	1	1
T_8	1	1	1	1	1	1
T_9	1	1	0.6235	1	0.431	0.3777
T_10	1	1	0.4376	0.6587	0.489	0.5068
T_11	1	1	0.5549	0.5016	0.457	0.4066
T_12	1	1	1	0.019	1	1
T_13	0.3985	0.4106	0.337	1	0.2194	0.375
T_14	1	1	1	0.3576	1	1
T_15	1	0.8809	1	1	0.8697	0.8524
Summary	**3 pass 12 fail**	**4 pass 11 fail**	**8 pass 7 fail**	**8 pass 7 fail**	**9 pass 6 fail**	**8 pass 7 fail**

p-values

Table 1. Diehard results for 1-D CA of variable spans with periodic boundaries.

REFERENCES

1 SIEGENTHALER, T. : "Correlation Immunity of Nonlinear Combining Functions for Cryptographic Applications", IEEE Transactions on Information Theory, Vol. IT-30, No. 5, September 1984, pp. 776-780.

2 GUSTAVSON, F. G.: "Analysis of the Berlekamp-Massey Linear Feedback Shift-Register Synthesis Algorithm." *IBM J. Res. Dev.* **20**, Number 3, pp. 204-212, 1976.

3 WOLFRAM, S.: "A New Kind of Science". Champaign, IL: Wolfram Media, 2002.

4 WOLFRAM, S.: "Random Sequence Generation by Cellular Automata", *Advances in Applied Mathematics*. Volume 7, Issue 2, June 1986, Pages 123-169.

5 SEREDYNSKI, FRANCISZEK, BOUVRY PASCAL, and ZOMAYA, ALBERT Y.: "Cellular automata computations and secret key cryptography", Parallel Computing, Vol. 30, 2004, pp. 753-766.

6 LLACHIINSKI, Andrew: "Cellular Automata: A Discrete Universe", World Scientific, 2001, pp. 94.

7 HORTENSIUS, P.D., McLEOD, and CARD, H.C.:"Parallel Random Number Generation for VLSI Systems Using Cellular Automata", IEEE Transactions on Computers, Vol. 38, Issue 10, October 1989, pp. 1466-1473.

8 Marsaglia, "The Marsaglia Random Number CDROM including the Diehard Battery of Tests of Randomness", Florida State University, http://i.cs.hku.hk/~diehard/