77

LA-UR- 01-1381 c. 1

Approved for public release;
distribution is unlimited.

Title: ON GARDENS OF EDEN OF ELEMENTARY CELLULAR AUTOMATA

Author(s):

Christopher L. Barrett, D-2
William Y. C. Chen, T-DO
Christian M. Reidys, D-2

Submitted to: Discrete Models (Combinatorics, Computation & Geometry) Conference
Paris, France
July 2-5, 2001

Los Alamos
NATIONAL LABORATORY

On Gardens of Eden of Elementary Cellular Automata

Chris Barrett	D-2, MS M997 Los Alamos National Laboratory Los Alamos, New Mexico 87545, USA email: barrett@lanl.gov
William Y. C. Chen	Center for Combinatorics The Key Laboratory of Pure Mathematiccs and Combinatorics of Ministry of Education, China Nankai University Tianjin 300071, P. R. China Email: chencentral@hotmail.com and Theoretical Division, MS K710 Los Alamos National Laboratory Los Alamos, New Mexico 87545, USA
Christian Reidys	D-2, MS M997 Los Alamos National Laboratory Los Alamos, New Mexico 87545, USA email: reidys@lanl.gov

Please send correspondence to:

William Y. C. Chen
Center for Combinatorics
Nankai University
Tianjin 300071
P. R. China
Email: chencentral@hotmail.com

Abstract

Using de Bruijn graphs, we give a characterization of elementary cellular automata on the linear lattice that do not have any Gardens of Eden. It turns out that one can easily recoginze a CA that does not have any Gardens of Eden by looking at its de Bruijn graph. We also present a sufficient condition for the set of words accepted by a CA not to constitute a finite-complement language.

1 Introduction

Cellular automata (CAs) are discrete dynamical systems with simple construction but sophisticated behaviour. The classical model of CAs can be viewed as a dynamical system on a lattice (here we only consider the one-dimensional lattice) where the update of sites is done synchronously based on local rules.

The preimages of a CA are of fundamental importance to the study of surjectivity, reversibility, and the existence of Gardens-of-Eden [1, 2, 15, 6, 7, 8, 9, 10, 14, 17, 18]. The *de Bruijn graphs* have been employed to represent the rules of a CA [11]. E. Jen used the structure of de Bruijn graphs to enumerate preimages of a cellular automaton for the special case of elementary rules [5] (local rules for the nearest neighbor.) In general, she showed that the number of preimages of an arbitrary sequence can be computed by two uncoupled systems of linear recurrence relations with constants coefficients [4]. The idea of using de Bruijn graphs to encode a CA was also mentioned by S. Wolfram [16].

In this note, we obtain a very simple characterization of an elementary CA to have Gardens of Eden based its de Bruijn graph. We also present a sufficient condition to assert that the set of words accepted by a CA does not constitute a finite-complement language.

2 Gardens of Eden and de Bruijn Graphs

In this paper, we will be only concerned with elementary cellular automata on $F_2 = \{0, 1\}$. A rule is a map f from $F_2^3 \to F_2$, or a Boolean function in three variables. Given a sequence $X = \cdots x_{-1}x_0x_2 \cdots$ on F_2, the image of X under the mapping of the elementary CA with rule f is the sequence $X' = \cdots x'_{-1}x'_0x'_1 \cdots$, where $x'_i = f(x_{i-1}, x_i, x_{i+1})$. Given a cellular automaton T, and two sequences $S = s_1 \cdots s_n$ of length n and $X = x_0 \cdots x_{n+1}$ on F_2, if T maps X into S, then we call X a *preimage* of S.

Recall that a k-dimensional de Bruijn graph on a alphabet A is defined as a directed graph whose vertices are sequences on A of length k, an arc is of the form

$$x_1x_2 \cdots x_k \to yx_1x_2 \cdots x_{n-1}.$$

However, we will the reverse de Bruijn graph as in the CA literature. For an elementary CA, we associate it with a directed graph, called its **de Bruijn graph**, which is a directed graph with four vertices $\{00, 01, 10, 11\}$ the arc from wx to xy labeled by the value $f(w, x, y)$. Such a difinition easily extends to a general one-dimensional CA. For a path in the de Bruijn graph of a CA, its label sequence is defined as the list of labels occuring in the path as one travels along the path.

Our first result is the following observation:

Proposition 2.1 *There is a one-to-one correspondence between the preimages of S and the paths in the de Bruijn graph with label sequence S.*

For elementary rules, Erica Jen obtained formulas of counting the number of preimages of a CA based on an encoding of the local rules into a de Bruijn graph [5]. We are not getting into the direction of enumeration. To demonstrate the perspective of the above Proposition, we give an example of showing a result of Jen on the Gardens of Eden of a CA. A sequence S which does not have a preimage is called a *Garden of Eden.* A natural question to ask about Gardens of Eden is whether a CA does not have any Gardens of Eden. This is equivalent to the question whether any sequence S on F_2 is a label sequence of a path in the de Bruijn graph? If so, the CA does not have any Gardens of Eden.

Jen pointed out that for Rule 30, which is defined by

$$\{000, 101, 110, 111\} \to 0, \ \{001, 010, 011, 100\} \to 1,$$

the corresponding CA does not have any Gardens of Eden.

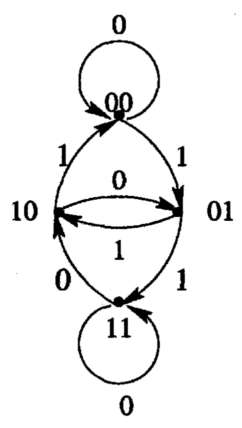

Figure 1: The de Bruijn graph for Rule 30

Here is a really simple reason. The de Bruijn graph has the following property: for any vertex, there are always an incoming arc with label 1 and an incoming arc with label

0. Given any sequence $S = s_1 s_2 \cdots s_n$, we may recover a path on the de Bruijn graph with label sequence S. Starting with any vertex, say u, finding the vertex v such that (v, u) has label s_n. Keeping moving backwards, we get the desired path.

Suppose that the de Bruijn graph has a similar property: for every vertex, there are always an outgoing arc with label 1 and an outgoing arc with label 0. Then we clearly see that such a CA does not have any Gardens of Eden.

The following theorem gives a characterization of elementary CAs that do not have Gardens of Eden.

Theorem 2.2 *An elementary CA, except for two special Rules* 51 *and* 204, *does not contain any Gardens of Eden, if and only if the de Bruijn graph has the property: Either every vertex has two outgoing arcs with different labels, or every vertex has two incoming arcs with different labels.*

Proof. The sufficiency is obvious. Given a de Bruijn graph as in Figure 2, let us assume that the corresponding CA does not have any Gardens of Eden.

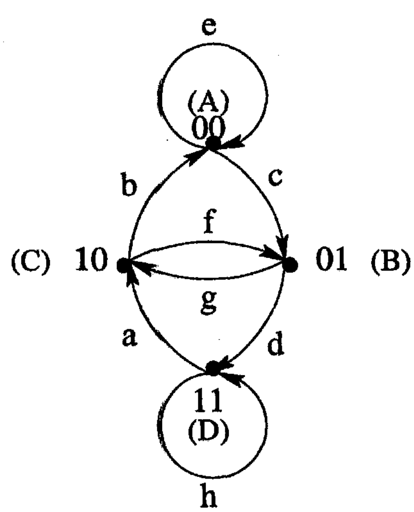

Figure 2: The de Bruijn graph for a general elementary rule.

Step One. For notational convenience, we use A, B, C, D to refer to the vertices 00, 01, 10, 11, and use GOE to refer to a Garden of Eden. We claim that if a CA does not have any Gardens of Eden, then the labels in the de Bruijn graph contains four 1's and four 0's. If not, we may assume that there are at most three 0's in the graph, because of the symmetry property between 0 and 1.

First, if there is no cycle with label 0, then 0000 has to be a GOE. Next, we assume that there is only one cycle with label 0 in the graph and $e = 0$.

1. If $c = 1$, let $S_1 = 00001$, then the end vertex must be B for there is only one cycle with label 0.

   If $g = d$, then the graph can generate $S_1\bar{g}$, where $\bar{g} = 1 - g$. So it is a GOE. Otherwise, we consider $S_1 g$. The end vertex must be C. If $b = f$, then $S_1 g\bar{b}$ would be a GOE. Otherwise, we consider $S_1 d0$. It must be a GOE, since b and f are different, d and g are different, implying that a and h must be 1.

   Hence, for all the cases with $c = 1$, there must exist a GOE.

   Here are more cases to consider.

2. If $e = c = 0$ and $f = 0$, then 0100 is a GOE.
3. If $e = c = 0$ and $g = 0$, then 010110 is a GOE.
4. If $e = c = 0$ and $d = 0$, then the sequence 010 is a GOE.
5. If $e = c = 0$ and $b = 0$, then 0101 is a GOE.
6. If $e = c = 0$ and $a = 0$, then 00100 is a GOE.

If there is only one cycle in the graph with label 0 and it is not e or h, then we have:

1. If $b = c = g = 0$, then 01010110 is a GOE.
2. If $c = f = g = 0$, then 01010 is a GOE.
3. If $b = f = g = 0$, then 0101 is a GOE.

For the cases that there are two cycle with label 0, we have:

1. If $e = 0$ and $f = g = 0$, then 011010110 is a GOE.
2. If $e = h = 0$ and $c = 0$, then 00001000010 is a GOE.
3. If $e = h = 0$ and $b = 0$, then 00001000010 is a GOE.
4. If $e = h = 0$ and $f = 0$, then 000010000 is a GOE.

From the above analysis, we reach the conclusion that there are four 1's and four 0's in the labels.

Step Two. Suppose that the above de Bruijn graph does not satisfy the conditions in the Theorem. We call a vertex X a *special vertex* if both of its outgoing arcs have the same label. For example, if $e = c = 1$, then we call A a special vertex with label 1. Because of the symmetry between 1 and 0, we may consider the following cases:

1. Case 1. A is a special vertex with label 0, D is a special vertex with label 1, and B, C are not special vertices.
   (a) For $f = g = 1$ and $b = d = 0$, 0101 is a GOE.
   (b) For $b = g = 1$ and $d = f = 0$, 101000 is a GOE.
   (c) For $f = g = 0$ and $b = d = 1$, 1010 is a GOE.
   (d) For $b = g = 0$ and $f = d = 1$, 10110 is a GOE.
2. Case 2. C is a special vertex with 0, D is a special vertex with 1, and A, B are not special vertices.
   (a) For $e = d = 1$ and $c = g = 0$, 01010 is a GOE.
   (b) For $c = d = 1$ and $e = g = 0$, 0101 is a GOE.
   (c) For $e = g = 1$ and $c = d = 0$, 0000 is a GOE.
   (d) For $e = d = 0$ and $c = g = 1$, 00000111 is a GOE.
3. Case 3. C is a special vertex with label 0, B is a special vertex with 1 and A, D are not special vertices.
   (a) For $a = c = 1$ and $e = h = 0$, 1111 is a GOE.
   (b) For $h = c = 1$ and $e = a = 0$, 1111101 is a GOE.
   (c) For $a = c = 0$ and $e = h = 1$, 0000 is a GOE.
   (d) For $a = e = 1$ and $c = h = 0$, 00000011 is a GOE.
4. Case 4. C is a special vertex with label 0, A is a special vertex with 1 and B, D are not special vertices.
   (a) For $a = g = 0$ and $d = h = 1$, 101 is a GOE.
   (b) For $a = g = 1$ and $d = h = 0$, 0000011 is a GOE.
   (c) For $a = d = 1$ and $g = h = 0$, it is just the graph with different incoming arcs for every vertex.
   (d) For $a = d = 0$ and $g = h = 1$, it is also the graph with different incoming arcs from every vertex.
5. For the cases with all the four vertices are special, one can see that the de Bruijn graphs are either the cases for which every vertex has different incoming arcs, or the special cases.

This completes the proof. ∎

The two special Rules 51 and 204 are also of particular interest. There is also a very good combinatorial reason for why they do not have any Gardens of Eden. Since Rule 51 and Rule 204 are symmetric with regard to the exchange of 1 and 0, we only need consider Rule 51. It is defined by the following local rule:

$$\{010, 011, 110, 111\} \to 0, \quad \{000, 001, 100, 101\} \to 1.$$

It is easy to see that for a sequence $S = 11\cdots1$ consisting of any positive number of 1's, there exists a path from the vertex 10 to 01 whose label sequence is S. On the other hand, there exists a path from vertex 01 to 10 such that the label sequence is $00\cdots0$ (of length at least 1). Therefore, any sequence on $\{0,1\}$ can be generated on the de Bruijn graph, implying that there is no Garden of Eden.

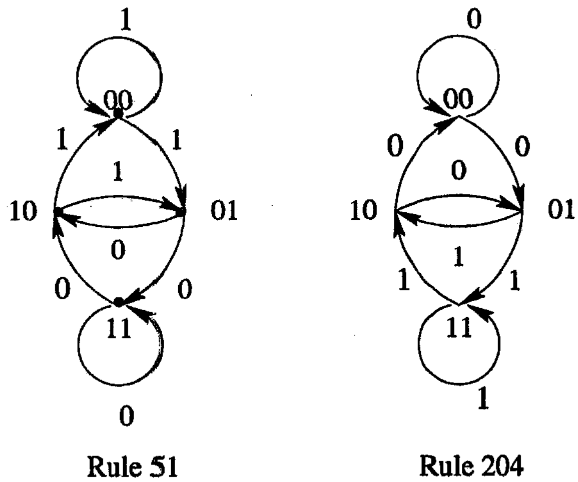

Figure 3: The de Bruijn graphs for Rule 51 and Rule 204

In conclusion of this section, we would like to propose the problem of classifying the general one dimensional CAs without Gardens of Eden, based on their de Bruijn graphs in the vein of Theorem 2.2.

3 Finite-complement Languages

Suppose a, b are two words on alphabet Σ. If there exist $u, v \in \Sigma^*$ such that $b = uav$, then we will say that a is a *subword* of b, denoted by $a \leq b$. If $a \leq b$ and $a \neq b$, we say a is a *proper subword* of b, denoted by $a < b$.

We recall the definition of a *finite-complement language.*

Definition 3.1 *Let L be a language and L^c be the complement of L. If there exist a subset $S \subseteq L^c$ which has the following properties:*

1. $|S| < \infty$.

2. *For every* $x \in L^c$, *there exists* $s \in S$ *such that* s *is the subword of* x.

Then we say that L *is a finite-complement language.*

Let $B_L = \{x \in L^c \,|\, \forall y < x, y \in L\}$, then we have

Lemma 3.2 *L is finite-complement if and only if* $|B_L| < +\infty$.

Proof. If $\epsilon \notin L$, then $B_L = \{\epsilon\}$ and the assertion becomes obvious. So we may assume that $\epsilon \in L$.

"$\Rightarrow$": We only need to show that B_L satisfies the second condition in Definition 3.1. Otherwise, there exists $x \in L^c$ such that for all $s \in B_L$, s is not the subword of x, hence there is one with minimal length. We still denote it by x. Since $\epsilon \in L$, $x \neq \epsilon$, it has a proper subword y. Since x is minimum, there exists $s \in B_L$ such that $s \leq y$, which implies $s \leq x$. A contradiction! Hence B_L satisfies the second condition.

"$\Leftarrow$": Suppose $S \subseteq L^c$ satisfies the conditions in Definition 3.1. For every $x \in B_L$, there exists $s \in S$ such that $s \leq x$. However, from the definition of B_L, s cannot be a proper subword of x, which means that $s = x$. It follows that $B_L \subseteq S$, and $|B_L| \leq |S| \leq +\infty$. ∎

Given a CA, if a sequence S has a preimage, we say that S is an *accepted word* of the CA. Here we give a sufficient condition to ensure that the set of words accepted by a CA does not constitute a finite-complement language. Note that this result is valid for any one dimensional CA on the linear lattice.

Theorem 3.3 *If there exists a cycle C in the de Bruijn graph such that for any label sequence S of the cycle and for sufficiently large n, S^n can be generated only by paths containing the cycle C. Let S_0 be a label sequence of C. (Note that C may have more than one label sequence.) If there exist sequences T_1 and T_2 on F_2 which satisfy the following conditions:*

1. *T_1S_0 is an accepted word and is a label sequence of a path with the beginning vertex on the cycle C.*

2. *S_0T_2 is an accepted word and is a label sequence of a path with the end vertex on the cycle C.*

3. *The beginning vertex and the end vertex in the above paths are different.*

Then the accepted words of the CA do not constitute a finite-complement language.

Proof. From the hypothesis and Condition 3, $T_1 S_0^n T_2$, $n \geq 1$ do not have preimages for sufficiently large n. Suppose S is a finite subset of sequences which do not have preimages and for any sequence x which does not have a preimage, there exists $s \in S$ such that $s \leq x$. Then there exists M such that the length of $s \in S$ is less than M. From Conditions 1 and 2, it follows that any subsequence of $T_1 S_0^n T_2$ of length less than S_0^n is accepted by the CA, which implies they cannot be $s \in S$. Choosing n sufficiently large, we deduce that $T_1 S_0^n T_2$ cannot have $s \in S$ as its subwords. A contradiction! This completes the proof. ∎

We give an example that has been studied by Jen [5].

Example 3.4 *For Rule* 22, *which is defined by*

$$\{000, 011, 101, 110, 111\} \to 0, \ \{001, 010, 100\} \to 1$$

the corresponding de Bruijn graph is shown in Figure 4.

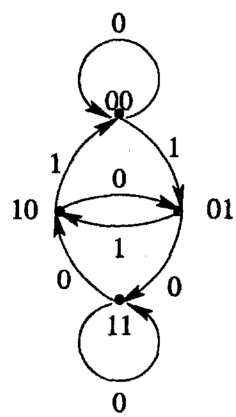

Figure 4: The de Bruijn graph for Rule 22

The words accepted by this CA do not constitute a finite-complement language.

Proof. We consider the cycle from 10 to 00 to 01 and then return to 10 as the cycle C in the above Theroem, which satisfies the condition as in the Theorem. For the basic sequence $S_1 = 111$, let $T_1 = 10$ and $T_2 = 0101$, it is easy to see that $T_1\{S_1\}^k$ can be accepted by the CA and the beginning vertex may be 00 or 01. The sequence $\{S_1\}^k T_2$ is also accpted by the CA and the ending vertex must be the point 10. Thus from Theorem 3.3, the set of words accepted by the CA according to Rule 22 does not form a finite-complement language. ∎

Acknowlegment. We would like to thank Q. H. Hou for valuable comments and suggestions.

References

1. S. Amoroso and G. Cooper, The Garden-of-Eden theorem for finite configurations, Proc. Amer. Math. Soc., **26** (1970) 158-164.
2. S. Amoroso, G. Cooper and Y. Patt, Some clarifications of teh concept of a Garden-of-Eden configuration, J. Computer and System Sci., **10** (1975) 77-82.
3. J. E. Hopcroft and J. E. Ullman, Formal Languages and Their Relation to Automata, Addison-Wesley Publishing Company, 1969.
4. E. Jen, Scaling of preimages in cellular automata, Complex Systems, **1** (1987) 1045-1062.
5. E. Jen, Enumeration of preimage in cellular automata, Complex Systems, **3** (1989) 421-456.
6. J. Milnor, Notes on surjective cellular automaton-maps, Institute for Advanced Study, preprint, June, 1984.
7. E. F. Moore, Machine models of self-reproduction, In: Proc. Sympos. Appl. Math., Vol. 14, Amer. Math. Soc., Providence, RI, 1962, pp. 17-33.
8. J. Myhill, The converse of Moore's Garden-of-Eden theorem, Proc. Amer. Math. Soc., **14** (1963) 685-686.
9. S. Skyum, Confusion in the Garden of Eden, Proc. Amer. Math. Soc., **50** (1975) 332-336.
10. K. Sutner, Linear cellular automata and the Garden-of-Eden, Math. Intelligencer, **11** (1989) (No. 2) 49-53.
11. K. Sutner, De Bruijn graphs and linear cellular automata, Complex Systems, **5** (1991) 19-30.
12. B. Voorhees, Predecessors of cellular automata states, I, Additive automata, Phys. D, **68** (1993) 283-292.
13. B. Voorhees, Predecessors of cellular automata states, II, Pre-images of finite sequences, Phys. D, **73** 136-151.
14. B. Voorhees, Computation of predecessor states for composite elementary cellular automata, Complex systems, **10** (1996) 207-217.
15. B. Voorhees and S. Bradshaw, Predecessors of cellular automata states, III, Garden of Eden classification of cellular automata, Phys. D, **73** (1994), 152-167.
16. S. Wolfram, Computation theory of cellular automata, Comm. Math. Phys., **96** (1984) 15-57.
17. S. Wolfram, Theory and Applications of Cellular Automata, World Scientific Press, Singapore, 1986.
18. T. Yaku, Inverse and injectivity of parallel relations induced by cellular automata, Proc. Amer. Math. Soc., **58** (1976) 216-220.