The Enumeration of Preimages and Gardens-of-Eden in Sequential Cellular Automata

Yan Deqin
Hou Qinghu*
Research Center for Combinatorics,
Nankai University,
Tianjin 300071, China

The enumeration of preimages in sequential cellular automata is approached. Two methods are given to enumerate the preimages of a rule. Preimage enumeration is simplified by using transform matrix expressions. The concept of factor graphs is presented in the study of De Bruijn graphs. A criterion for a rule having gardens-of-Eden is given.

1. Introduction

Cellular automata (CA) are discrete dynamic systems, of simple construction but varied behavior. Studying preimages of CA is important for learning their properties [2, 3]. Gardens-of-Eden (GOE) are sequences with no preimages and are related to the preimages of a rule in [1, 2, 4]. Recently in [5], Chris Barrett, William Y. C. Chen, and Christian Reidys made an approach on the GOE in sequential cellular automata (SCA), and proposed a question: What is a sufficient and necessary condition of a rule having no gardens-of-Eden (i.e., to be non-GOE)? For the enumeration of preimages in parallel CA, E. Jen gave a recurrence relation formula in [1]. But in SCA there is no formula to enumerate the preimages.

This paper presents a preimages enumeration formula for a rule in SCA, and gives two methods for enumerating the preimages of a rule. The enumeration is greatly simplified by using a transform matrix expression, especially when r and k are bigger than 1 and 2. The concept of factor graphs is presented in the study of sequential De Bruijn graphs. A criterion for a rule being nonGOE is given. That is, a certain kind of special structure of rule matrix is the sufficient and necessary condition for a rule to be nonGOE.

*Electronic mail address: hqh@public1.tpt.tj.cn.

2. Preliminaries

The general form of SCA is

$$x_i^{t+1} = f(x_{i-r}^{t+1}, \ldots, x_{i-1}^{t+1}, x_i^t, x_{i+1}^t, \ldots, x_{i+r}^t)$$

where x_i^t denotes the value of site i at time t. The values are restricted to the finite field $F_k = \{0, 1, \ldots, k-1\}$, f represents the "rule" defining the automaton and r is a nonnegative integer specifying the radius of the rule.

Just as in parallel CA, the rule number for primary CA ($r = 1$, $k = 2$) is defined as the decimal value of the term's binary expression: $a_7 a_6 \ldots a_0$, where $a_i = f(xyz)(i = 0, \ldots, 7)$. $xyz \in \{000, 001, \ldots, 111\}$. i is the decimal value of xyz. For example, according to a rule with $f(111) = f(110) = f(101) = f(100) = 0$, $f(011) = f(010) = f(001) = f(000) = 1$, the rule number is 15, and is called *rule 15*.

Definition 1. A $k^{2r} \times k^{2r}$ matrix, with rows and columns labeled by the sequences of length $2r$ on F_k, is called the *rule matrix* of rule f if $P_{i,j} = x_0 = f(m_0, m_1, \ldots, m_{2r-1}, x)$, for $i = m_0 m_1 \ldots m_{2r-1}$, $j = m_1 m_2 \ldots m_{r-1} x_0 m_{r+1} \ldots m_{2r-1} x$; and $P_{i,j} = 0$, otherwise. Also, let the vertices be the sequences of length $2r$ and draw an arc from i to j with the label x_0 for all i, j as above. Then the digraph is called a *sequential De Bruijn graph* of the rule.

When $r = 1, k = 2$, P is a 4×4 matrix. The row and column labels are 00, 01, 10, and 11. The element of row $(y1, y2)$ and column $(y0, y)$ is $y0 = f(y1, y2, y)$. If $y0 = 0$, the element is denoted by α, if $y0 = 1$, it is denoted by β, other elements are 0. For example, the rule matrix of rule 15 is

$$\begin{pmatrix} 0 & 0 & \beta & \beta \\ 0 & 0 & \beta & \beta \\ \alpha & \alpha & 0 & 0 \\ \alpha & \alpha & 0 & 0 \end{pmatrix}.$$

In fact, this rule matrix is the depiction of a De Bruijn graph. The De Bruijn graph in CA and in SCA of rule 121 are shown by D1 and D2 in Figure 1 respectively.

3. Preimage enumeration in sequential cellular automata

For an arbitrary integer $0 \le m \le k^{2r} - 1$, with

$$m = \sum_{i=0}^{2r-1} m_i k^{2r-i-1},$$

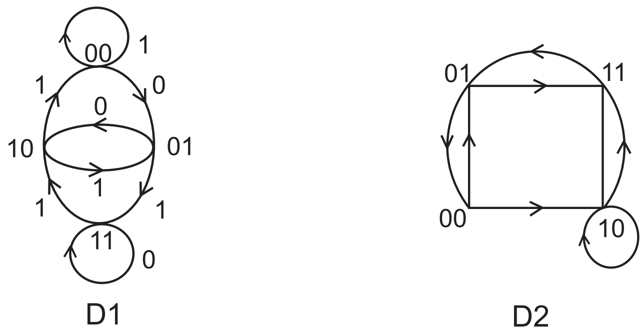

Figure 1. The De Bruijn graph of rule 121.

denote by $M = m_0 m_1 \ldots m_{2r-1}$ the symbol corresponding to its k-ary representation. The symbols M thus range over all possible blocks of length $2r$.

Given a sequence $S = s_n s_{n-1} \ldots s_1$, the number of its preimages is denoted by $N(S)$. Let L_M^j be the preimage of the sequence $s_j \ldots s_1$ beginning with M, from [1] we have

$$N(S) = \sum_{m=0}^{k^{2r}-1} L_M^n,$$

where

$$L_M^j = \sum_i L_{M_i}^{j-1} I_j(x_i), \tag{1}$$

and

$$I_j(x) = \begin{cases} 1, & x = x_j, \\ 0, & \text{otherwise.} \end{cases}$$

Definition 2. Let L^j be the column vector consisting of L_M^j, then equation (1) can be written as

$$L^j = P_j L^{j-1}. \tag{2}$$

We call P_j the *transform matrix*.

In fact, the rule matrix has the same structure as the transform matrix.

Theorem 1. On F_k, suppose $S = aa \ldots a$ is the sequence produced by a rule with radius r, then

$$L^n = A^n E,$$

where E is the k^{2r}-dimension column vector and A is the transform matrix of the rule on S.

Proof. From the structure of the transform matrix, we know $P_1 = P_2 = \cdots = P_j$, and by equation (2) we get $L^n = A^n E$. ■

Theorem 2. On F_k, for any n element sequence S, we separate S into blocks such that in one block all of the elements are the same. Let v_i be the value of the elements in block i and u_i be the length of the block. We have

$$L^n = A_{v_t}^{u_t} \dots A_{v_1}^{u_1} E, \tag{3}$$

where $v_i \in F_k$, t, u_i are integers, E is the k^{2r} element column unit vector (all elements are 1), and $A_{v_i}^{u_i}$ is the transform matrix on the elements of v_i $(i = 1, 2, \dots, t)$.

Proof. For a sequence S, the separation method means that if $s_i = s_{i-1} = \cdots = s_1$, $s_{i+j+1} \neq s_{i+j} = \cdots = s_{i+1} \neq s_i \dots$, there are $v_1 = s_1$, $u_1 = i$; $v_2 = s_{i+1}$, $u_2 = j \dots$. From Theorem 1 we have

$$\begin{aligned} L^{u_1} &= A_{v_1}^{u_1} E, \\ L^{u_2} &= A_{v_2}^{u_2} \cdot L^{u_1}, \dots, \\ L^{u_n} &= A_{v_n}^{u_n} \cdot L^{u_{n-1}}. \end{aligned}$$

Then

$$L^n = A_{v_t}^{u_t} \dots A_{v_1}^{u_1} E. \quad ■$$

There are two ways to enumerate the preimages in SCA. One is to apply the method used in [1]; the other is to use Theorem 2, which we call the *transform matrix method*. In the following we take $r = 1, k = 2$ and rule 12 as an example for enumerating the preimages by both methods.

Let S be an arbitrary sequence. From right to left, divide S into blocks of consecutive 0s and 1s, and let

$$\begin{aligned} a_i &= \text{number of consecutive 0s in the } i\text{th block}, \\ b_i &= \text{number of consecutive 1s in the } i\text{th block}. \end{aligned}$$

For example, with $S = 00101001100100$, set $a_1 = a_2 = a_3 = 2$, $a_4 = 1$, $a_5 = 2$ and $b_1 = 1$, $b_2 = 2$, $b_3 = b_4 = 1$.

Suppose there is a sequence S with the blocks of $b_t a_t \dots b_1 a_1$. On the block of 0s the transform matrix is A, and on the block of 1s the transform matrix is B.

By using the first method, we have

$$\begin{aligned} L_{00}^j &= (L_{00}^{j-1} + L_{01}^{j-1}) I_j(0), \qquad & L_{01}^j &= (L_{10}^{j-1} + L_{11}^{j-1}) I_j(1), \\ L_{10}^j &= (L_{00}^{j-1} + L_{01}^{j-1}) I_j(0), \qquad & L_{11}^j &= (L_{00}^{j-1} + L_{01}^{j-1}) I_j(0), \end{aligned}$$

and

$$\begin{aligned}
&U_{00}^{j} = U_{00}^{j-1} + U_{01}^{j-1},\\
&U_{01}^{j} = 0,\\
&U_{10}^{j} = U_{00}^{j-1} + U_{01}^{j-1},\\
&U_{11}^{j} = U_{00}^{j-1} + U_{01}^{j-1},\\
&V_{00}^{j} = 0,\\
&V_{01}^{j} = V_{10}^{j-1} + V_{11}^{j-1},\\
&V_{10}^{j} = V_{11}^{j} = 0.
\end{aligned}$$

Then

$$\begin{aligned}
&U_{00}^{1} = U_{10}^{1} = U_{11}^{1} = 2, U_{01}^{1} = 0,\\
&U_{00}^{a_1} = U_{10}^{a_1} = U_{11}^{a_1} = 2,\\
&U_{01}^{a_1} = 0,\\
&V_{00}^{b_1} = V_{10}^{b_1} = V_{11}^{b_1} = 0.
\end{aligned}$$

If $b_1 > 1$, there is $V_{01}^{b_1} = 0$. If $b_1 = 1$, there is $V_{01}^{b_1} = U_{10}^{a_1} + U_{11}^{a_1} = 4$.

For the first a_2 0s there is $U_{00}^{a_2} = \cdots = U_{00}^{1}$. After a_2 0s there is a 1, according to the rule, $U_{00}^{1} = V_{01}^{b_1} = 4$, namely $U_{00}^{a_2} = 4$. So we have

$$U_{10}^{a_2} = U_{11}^{a_2} = 0, \qquad U_{01}^{a_2} = 0.$$

From the recurrence relation of V_{01}^{j}, we know $b_i = 1$, $(i = 1, 2, \ldots)$. According to the rule, there is

$$\begin{aligned}
&V_{01}^{b_i} = U_{10}^{a_i} + U_{11}^{a_i}\\
&V_{00}^{b_i} = V_{10}^{b_i} = V_{11}^{b_i} = 0
\end{aligned}$$

so

$$U_{10}^{a_i} = U_{11}^{a_i} = V_{01}^{b_i-1}.$$

Then

$$\begin{aligned}
V_{01}^{b_i} &= 2U_{11}^{a_i} = 2V_{01}^{b_i-1} = 2^{i-1}V_{01}^{b_1} = 2^{i+1},\\
N(S) &= \sum_{i,j\in\{0,1\}} V_{ij}^{b_t} = V_{01}^{b_t} = 2^{t+1}.
\end{aligned}$$

By using the transform matrix method we have

$$A = \begin{pmatrix} 0 & 0 & 0 & 0\\ 0 & 0 & 1 & 1\\ 0 & 0 & 0 & 0\\ 0 & 0 & 1 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 1 & 0 & 0\\ 0 & 0 & 0 & 0\\ 1 & 1 & 0 & 0\\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

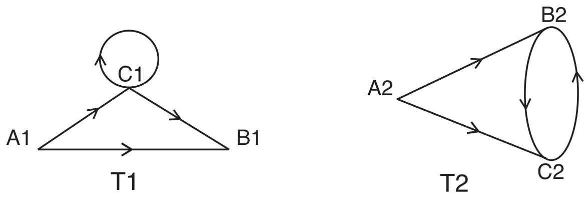

Figure 2. The factor graphs of the De Bruijn graph of a nonGOE rule.

It is easy to see that $B^i = B$ and $A^j = 0$ ($i = 1, 2, \ldots; j \geq 2$). We have $L^n = C^t E$, where

$$C = AB = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 2 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

and hence

$$L^n = \begin{pmatrix} 0 \\ 2^{t+1} \\ 0 \\ 0 \end{pmatrix}.$$

So we have

$$N(S) = 2^{t+1}.$$

4. The no gardens-of-Eden matrix expression

Given a rule, if there is $S = s_n \ldots s_1$ such that $N(S) = 0$, we say the rule has GOE, otherwise it is nonGOE.

We denote $0 \ldots 0$ and $1 \ldots 1$ as 0^* and 1^*. For any sequence S, if we find a path in the De Bruijn graph of a rule, and there is a bijective map between S and the consecutive labels of the path, it is denoted as $\{0, 1\}^*$. Since the preimages of the sequence S are the combinations of the vertices between which the labels hold consecutive maps to the elements of S, the rule is nonGOE.

Definition 3. We call graphs T1 and T2 in Figure 2 *the factor graphs* of a nonGOE rule.

Proposition 1. In graphs T1 and T2, one can find the path labeled 0^* or 1^*.

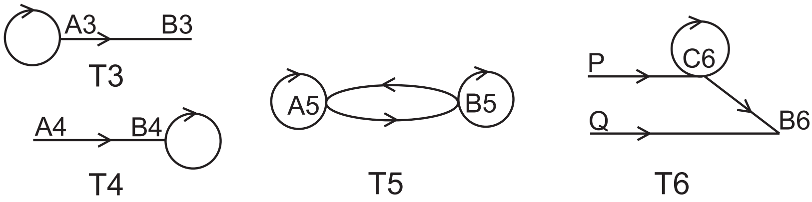

Figure 3. The graphs delivered from T1.

Proof. In graph T1, from A1 to B1 or A1 *via* C1 to B1 is the path. In graph T2, from A2 to B2 or A2 to C2 *via* several cycles made of B2 and C2 to one of B2 and C2 is the path. ■

Graphs T1 and T2 can be changed to graphs T3 and T4, T6 is another form of T1. Combinations of the factor graphs can produce the $\{0,1\}^*$ path shown in Figure 3.

Theorem 3. To obtain the graph containing the $\{0,1\}^*$ path by using graphs T1, T2, T3, and T4, the possible combinations are T1 + T2, T3 + T3, and T4 + T4.

Proof. Label the arcs or ring with 0 in graph T1, and 1 in graph T2. Let A1 connect B2 and B1 connect A2, then we can find the $\{0,1\}^*$ path in the new graph. Similarly, we can find the $\{0,1\}^*$ path in the new graph of T3 + T3 and T4 + T4.

Now we prove it is impossible to get the $\{0,1\}^*$ path from the other combinations.

T1 + T3: The only combination is vertex B3 connects A3. But, according to the rules in SCA, the label of the cycle at vertex A3 must be the same as the label in T1. So it is impossible.

T1 + T4: The only combinations are vertex B1 connects A4 and A1 connects B4. If so, there must be three arcs to get out from vertex B4, and it is impossible in SCA. For the same reason, T2 + T3, T2 + T4, and T3 + T4 are not qualified. ■

In SCA, according to the definition of the De Bruijn graph, there are four cases of combinations of T1 (or T6)+T2.

(a) B1 connects A2, A1 connects B2.

(b) C1 connects A2, A1 connects C2.

(c) B6 connects A2, Q connects B2, P connects C2.

(d) B6 connects A2, P connects B2, Q connects C2.

At first, we label all arcs and rings in graphs T1 and T6 with 0s, that is, let graphs T1 and T6 produce the 0^* path. Then graph T2 must produce the 1^* path, which implies that all arcs in factor graph T2 are labeled 1.

For cases (a) and (b), the vertex B1,C1 must be 00,01 or 01,00. The vertex A1,C2 can be any of 10,11. The rule matrices must be of the form

$$P_1 = \begin{pmatrix} \alpha & \alpha & . & . \\ 0 & 0 & . & . \\ \alpha & \alpha & . & . \\ 0 & 0 & . & . \end{pmatrix}, \qquad P_2 = \begin{pmatrix} 0 & 0 & . & . \\ \alpha & \alpha & . & . \\ \alpha & \alpha & . & . \\ 0 & 0 & . & . \end{pmatrix},$$

$$P_3 = \begin{pmatrix} \alpha & \alpha & . & . \\ 0 & 0 & . & . \\ 0 & 0 & . & . \\ \alpha & \alpha & . & . \end{pmatrix}, \qquad P_4 = \begin{pmatrix} 0 & 0 & . & . \\ \alpha & \alpha & . & . \\ 0 & 0 & . & . \\ \alpha & \alpha & . & . \end{pmatrix}.$$

In the matrix "." can be 0 or β. (This explanation is validated in the following.)

For the same reason, in cases (c) and (d) the vertex B1,C1 must be 00,01 or 01,00. The corresponding matrices are

$$P_5 = \begin{pmatrix} \alpha & \alpha & . & . \\ 0 & 0 & . & . \\ 0 & \alpha & . & . \\ \alpha & 0 & . & . \end{pmatrix}, \qquad P_6 = \begin{pmatrix} 0 & 0 & . & . \\ \alpha & \alpha & . & . \\ \alpha & 0 & . & . \\ 0 & \alpha & . & . \end{pmatrix},$$

$$P_7 = \begin{pmatrix} \alpha & \alpha & . & . \\ 0 & 0 & . & . \\ \alpha & 0 & . & . \\ 0 & \alpha & . & . \end{pmatrix}, \qquad P_8 = \begin{pmatrix} 0 & 0 & . & . \\ \alpha & \alpha & . & . \\ 0 & \alpha & . & . \\ \alpha & 0 & . & . \end{pmatrix}.$$

On the other hand, we let factor graph T2 produce the 0^* path, and T1 and T6 produce the 1^* path. The matrices with the structure of the combined graph using cases (a), (b), (c), and (d) are

$$P_9 = \begin{pmatrix} 0 & \alpha & . & . \\ \alpha & 0 & . & . \\ \alpha & \alpha & . & . \\ 0 & 0 & . & . \end{pmatrix}, \qquad P_{10} = \begin{pmatrix} 0 & \alpha & . & . \\ \alpha & 0 & . & . \\ 0 & 0 & . & . \\ \alpha & \alpha & . & . \end{pmatrix}.$$

Now we consider graph T5. T5 is the self-combination of T3 or T4, two T5s constitute a De Bruijn graph. When vertex A5 is 00 or 01, vertex B5 must be 10 or 11. Then the four vertices of the De Bruijn graph are 00,10 (or 00,11) and 01,10 (or 01,11). The corresponding matrices are

$$P_{11} = \begin{pmatrix} \alpha & . & . & \beta \\ . & . & . & . \\ . & . & . & . \\ \alpha & . & . & \beta \end{pmatrix}, \qquad P_{12} = \begin{pmatrix} . & . & . & . \\ . & \alpha & \beta & . \\ . & \alpha & \beta & . \\ . & . & . & . \end{pmatrix},$$

$$P_{13} = \begin{pmatrix} \alpha & . & \beta & . \\ . & . & . & . \\ . & . & \beta & . \\ \alpha & . & . & . \end{pmatrix}, \qquad P_{14} = \begin{pmatrix} . & . & . & . \\ . & \alpha & . & \beta \\ . & . & . & . \\ . & \alpha & . & \beta \end{pmatrix}.$$

It can be seen that a rule matrix determines the property of being nonGOE. However, any $\{0,1\}^*$ path must include the factor graphs; and, in order to get the $\{0,1\}^*$ path, we have discussed all possible combinations. So we get the conclusion in Theorem 4.

Theorem 4. A rule is nonGOE if and only if the structure of its rule matrix is the same as one of $P_1, \ldots, P_{14}$.

Example 1. The rule matrix of rule 150 is

$$\begin{pmatrix} \alpha & 0 & 0 & \beta \\ 0 & \alpha & \beta & 0 \\ 0 & \alpha & \beta & 0 \\ \alpha & 0 & 0 & \beta \end{pmatrix}.$$

Its structure is the same as P_{11} and P_{12}. So rule 150 is nonGOE. In fact, there are two factor graphs of T5 in its De Bruijn graph.

Acknowledgments

We are grateful to our advisor, professor William Y. C. Chen, for his help.

References

[1] E. Jen, "Enumeration of Preimages in Cellular Automata," *Complex Systems*, **3** (1989) 421.

[2] E. Jen, "Scaling of Preimages in Cellular Automata," *Complex Systems*, **1** (1987) 1045.

[3] S. Wolfram, *Theory and Applications of Cellular Automata* (World Scientific Press, Singapore, 1986).

[4] S. Amoroso and G. Cooper, "The Garden-of-Eden Theorem for Finite Configurations," *Proceedings of the American Mathematical Society*, **26** (1970) 158.

[5] Chris Barrett, William Y. C. Chen, and Christian Reidys, "A Combinatorial Study of Parallel and Sequential Cellular Automata" (preprint).