Complex Systems 4 (1990) 281–297

The Structure of the Elementary Cellular Automata Rule Space

Wentian Li
Santa Fe Institute, 1120 Canyon Road, Santa Fe, NM 87501, USA

Norman Packard
Center for Complex Systems Research, Physics Department, Beckman Institute, University of Illinois, Urbana, IL 61801, USA

Abstract. The structure of the elementary cellular automata rule space is investigated. The probabilities for a rule to be connected to other rules in the same class (intra-class), as well as rules in different classes (inter-class), are determined. The intra-class connection probabilities vary from around 0.3 to 0.5, an indication of the strong tendency for rules with the similar behavior to be next to each other. Rules are also grouped according to the mean-field descriptions. The mean-field clusters are classified into three classes (nonlinear, linear, and inversely linear) according to the "hot bits" in the rule table. It is shown that such classification provides another easy way to describe the rule space.

1. Introduction

Cellular automata (CA) as fully discrete dynamical systems with spatial degrees of freedom have become new models for the study of nonlinear complex systems [1]. Compared with other spatially extended dynamical systems, such as partial differential equations, cellular automata distinguish themselves by the possibility to have all kinds of functional forms including discontinuous ones. Even though the rules of cellular automata can be written as boolean functions (e.g., table 1 in [2]), the expression can be utterly irregular. Having modular operation as an ingredient in order to keep the value in a finite field, boolean functions are also intrinsically nonlinear, which makes the analysis difficult. As a result, much understanding of the cellular automata dynamics has been acquired by computer experiments.

Generally, a cellular automaton rule is represented by a lookup table: a list of state values to which block configurations are mapped. This list can be arranged in certain way as a sequence, which uniquely specifies the cellular automaton rule. We call this sequence *cellular automaton rule table.*

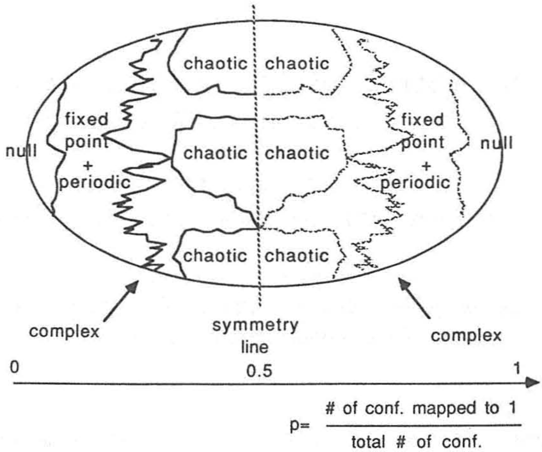

Figure 1: Schematic illustration of the typical structure of a cellular automaton rule space.

Sometimes, an analogy between the rule table and a DNA sequence is made, and we can also call it a *gene*. The characteristics of the rule table is the *genotype* and the dynamics shown by the rule is the *phenotype*.

If two CA rules are the same except that one maps a particular block configuration to state a_α while another maps the same block configuration to state $a_\alpha \pm 1$, we say the two rules are *next to each other*, with the Hamming distance between the two rule tables equal to 1. With this concept of distance, we can consider all the CA rules reside in a space, called the *rule space*. Each point in the rule space is a rule table, and all the points are arranged in such a way that nearby points have Hamming distance equal to 1. The operation of moving from one CA rule to its nearest rules can be called *flipping one bit in the rule table* (in particular for the two-state CAs), or, using the biological analogy, *mutation*.

For *elementary CA rules*, i.e., rules with the form $x_i^{t+1} = f(x_{i-1}^t, x_i^t, x_{i+1}^t)$ — the subscript i being the spatial position and the superscript t being the time — with $x_i \in (0,1)$, the rule table is a binary sequence with length $8 = 2^3$. We can write a rule table in the form: $(t_7t_6t_5t_4t_3t_2t_1t_0)$, which means that block configuration (000) maps to t_0, block configuration (001) maps to t_1, ... and (111) maps to t_7. Because there are two choices for each t_i, the total number of rule tables is $2^8 = 256$, which is also the total number of points in the elementary CA rule space.

We are interested in how these 256 rules are organized in their rule space, or *structure of the elementary CA rule space*, and whether two rules sitting next to each other are likely to have the same dynamical behavior, or more quantitatively, what the probability is for two nearby rules to have the similar behavior.

It has been observed [3] that different regions of the rule space have rules with different behaviors, such as regular, complex, and random. An activity parameter λ may be defined [4], in the case of binary states, as the density of 1's in the rule table. For example, the λ value for the rule table (10001001) is 3/8. Varying the λ value provides a way to move from one subset of the rule space to another.

Figure 1 shows schematically the typical structure of a CA rule space as parameterized by the λ. As λ gradually increases from 0 to 1, the dominate behavior of the rules changes from homogeneous fixed point to inhomogeneous fixed point, periodic, complex spatial-temporal dynamics, and chaotic dynamics, and when λ is larger than 0.5, the same process in reverse order. The reverse process is easy to understand because by switching 0 and 1, the rules at the tip with $\lambda \approx 1$ ("north pole") are equivalent to the rules at the tip with $\lambda \approx 0$ ("south pole"). Due to this equivalence between rules, one only has to examine the region from $\lambda = 0$ to $\lambda = 0.5$, and the remaining part of the rule space is just the mirror image of the former.

The abrupt change in global behavior when moving along a path in a CA rule space is highly reminiscent of bifurcation phenomena in smooth dynamical systems [5]. Actually, the "bifurcation-like" phenomena in CA rule space is more complicated and subtle than those in lower dimensional nonlinear systems with one or few parameters. For one thing, there is no unique "route to chaos," simply because there are so many paths one can pick to move from the "south pole" (simple behavior) to the "equator" (random behavior, or chaos). CA rule space provides us with a new challenger to understand bifurcation phenomena in multi-parameter, spatially extended dynamical systems.

A comprehensive discussion of the structure of the CA rule spaces will be in [3]. Here, we want to illustrate some of the main ideas by examining the simplest CA rule space — the elementary CA rule space. The paper is organized as follows. Section 2 introduces the concept of hypercubic space and folded hypercubic space. Section 3 outlines the classification scheme of CA, which is the basis for a description of the rule space structure. Section 4 discusses the transition within a class and between classes. Section 5 contains an attempt to characterize the rule space by the mean-field parameters.

2. Hypercubic spaces and "folded" hypercubic spaces

The set of all possible finite sequences with length n, $(t_1 t_2 \cdots t_n)$, consist a subset of the n-dimensional grid. When every t_i's can choose from the same state variables, $t_i \in \{a_\alpha\}$ $(\alpha = 0, 1, \cdots, m-1)$, this set has equal size m on each of the dimension axis, and we call it *the size-m n-dimensional hypercube*. In particular, if the state value is binary (m=2), it is the size-1 hypercube, or simply, *n-dimensional hypercube*.

Because of the application of finite sequences in many natural systems, the name hypercube appears in literature quite often, e.g., the study of binary DNA sequences [6]. Even two-dimensional patterns like those in the Ising spin model, the concept of hypercube is also applicable [7] because a two-dimensional pattern can be rearranged to a one-dimensional string. Since CA rule tables are sequences with finite length, CA rule spaces are also hypercubes (the size-m hypercube if the number of state is m).

An elementary CA rule f can be equivalent to another rule f_1 under the left-to-right transformation, if

$$f_1(x_{i-1}, x_i, x_{i+1}) = f(x_{i+1}, x_i, x_{i-1}) \tag{2.1}$$

is true for all 3-site blocks. f is equivalent to rule f_2 under the 0-to-1 transformation (represented by the overhead bar), if

$$f_2(x_{i-1}, x_i, x_{i+1}) = \overline{f(\overline{x_{i-1}}, \overline{x_i}, \overline{x_{i+1}})} \tag{2.2}$$

is always true, and it is equivalent to rule f_3 under the joint operation of both if

$$f_3(x_{i-1}, x_i, x_{i+1}) = \overline{f(\overline{x_{i+1}}, \overline{x_i}, \overline{x_{i-1}})} \tag{2.3}$$

holds. Suppose the rule table of f is $(t_7t_6t_5t_4t_3t_2t_1t_0)$, the three equivalent rules are $f_1 = (t_7t_3t_5t_1t_6t_2t_4t_0)$, $f_2 = (\bar{t_0}\bar{t_1}\bar{t_2}\bar{t_3}\bar{t_4}\bar{t_5}\bar{t_6}\bar{t_7})$, and $f_3 = (\bar{t_0}\bar{t_4}\bar{t_2}\bar{t_6}\bar{t_1}\bar{t_5}\bar{t_3}\bar{t_7})$. The spatial-temporal patterns for f_1, f_2 and f_3 are exactly the same by a mirror reflection transformation, or a white-to-black transformation, or a combination of both.

The rule space with only independent rules is smaller than the original rule space. The resulting rule space is called *folded rule space*, which is a *folded hypercubic space*. Notice that the two transformations or foldings are of different kinds. The 0-to-1 transformation folds the rule space with respect to the $\lambda = 0.5$ "plane." The left-to-right transformation operates on rules on the same "plane." For the 256 elementary CA rules, 88 of them remain independent.[1]

Figure 2 shows an actual folding process by grouping elementary "null rules" (the name will be explained in the next section) that are equivalent to each other. The rule numbers written in the figure are the decimal representation of the binary rule table as used in [8]. Rules in the same shaded rectangular in Figure 2(a) are equivalent by left-to-right transformations. The 0-to-1 transformation then folds these rules along the $\lambda = 0.5$ line and gives the resulting picture in Figure 2 (b).

Table 1 lists all the equivalence relations with the rules inside the parentheses are equivalent to the representative rule outside. Notice that the number of rules which are equivalent to each other can only be 1, 2, or 4 (see appendix for the explanation). To represent each cluster of equivalent rules, one can either (A) use the rule with the smallest decimal representation; or

[1]The number 88 will be explained in the appendix.

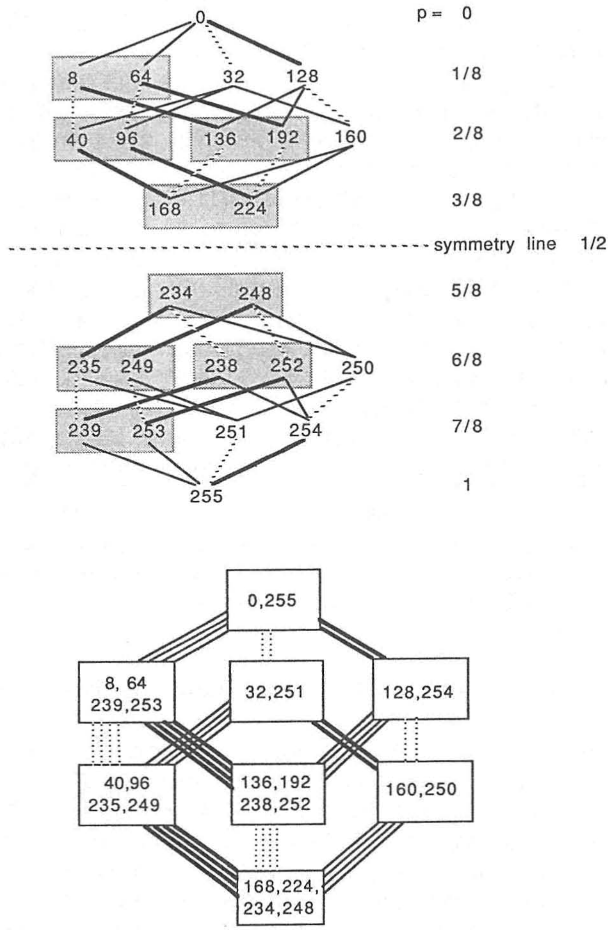

Figure 2: An example of the folding for the rule space. Thick lines are mutations by flipping bit t_0 (or t_7), thin lines for mutations at t_1 (or t_3, t_4, t_7), and dotted line for bit t_2 (or t_5).
(a) All the 24 null rules and their connections. Rules in the same shaded area are equivalent by the left-to-right transformations.
(b) Clusters and their connections in the folded hypercubic space.

(B) use the rule with the smallest decimal representation among these with the smaller λ value ("south sphere" in the hypercube). The previous studies use the first convention, e.g., table 2 of [2], but we adopt convention (B) in this paper. Rule-62, 94, 110, 122, 126 in convention (A) become Rule-131, 133, 137, 161, 129 in convention (B).

0(255), 1(127), 2(16,191,247), 3(17,63,119), 4(223), 5(95),
6(20,159,215), 7(21,31,87), 8(64,239,253), 9(65,111,125),
10(80,175,245), 11(47,81,117), 12(68,207,221), 13(69,79,93),
14(84,143,213), 15(85), 18(183), 19(55), 22(151), 23, 24(66,189,231),
25(61,67,103), 26(82,167,181), 27(39,53,83), 28(70,157,199),
29(71), 30(86,135,149), 32(251), 33(123), 34(48,187,243),
35(49,59,115), 36(219), 37(91), 38(52,155,211), 40(96,235,249),
41(97,107,121), 42(112,171,241), 43(113), 44(100,203,217),
45(75,89,101), 46(116,139,209), 50(179), 51, 54(147), 56(98,185,227),
57(99), 58(114,163,177), 60(102,153,195), 72(237), 73(109),
74(88,173,229), 76(205), 77, 78(92,141,197), 90(165), 104(233),
105, 106(120,169,225), 108(201), 128(254), 129(126), 130(144,190,246),
131(62,145,118), 132(222), 133(94), 134(148,158,214), 136(192,238,252),
137(110,124,193), 138(174,208,244), 140(196,206,220), 142(212),
146(182), 150, 152(188,194,230), 154(166,180,210), 156(198),
160(250), 161(122), 162(176,186,242), 164(218), 168(224,234,248),
170(240), 172(202,216,228), 178, 184(226), 200(236), 204, 232

Table 1: Clusters of all independent elementary rules. The rules outside the parenthesis are representative rules with the smaller λ value and the smaller decimal rule number. Those inside the parenthesis are rules equivalent to the representative rule.

3. Classification of the cellular automata rule dynamics

In this section, we summarize the classification schemes. A classification should be specified in order to determine the structure of the rule space. A full discussion of the issue of classification will be presented in [3]. It should be emphasized that by "classification of the CA dynamics," we mean the characterization of the rules by the dynamics from *typical* initial configurations ("random initial configurations"). It is not the classification according to the dynamics from all initial configurations [9], nor is it the mathematical characterization of the rule tables [10, 11]. In short, it is the *typical* phenotype rather than the genotype that is used for our classification.

Since there is nothing fundamentally different between the classification scheme discussed here as well as in [3] with the originally proposed four classes [12], the purpose for this section is to clarify a few points that often cause confusion, as well as to provide a quick reference.

The simplest phenotype classification scheme is to separate CA rules into two categories: those with periodic dynamics and those with nonperiodic dynamics. Since the dynamics for the CA with finite lattice length is always periodic (due to Poincare's recurrence time), the criterion becomes whether the dynamics has "short" or "long" periodicity. The "long" can be understood as the cycle length being exponentially divergent with the lattice length, and "short" as the cycle length being independent of the lattice width. It can be seen that the uncertainty in explaining the meaning of the "long" periodicity leaves room for ambiguity.

This classification scheme is too crude to incorporate the variety of the spatial configurations. After the spatial configuration is considered, we have the following five classes: (A) null rules: homogeneous fixed-point rules; (B) fixed-point rules: inhomogeneous fixed-point rules; (C) periodic rules; (D) locally chaotic rules: chaotic dynamics confined by the domain walls; and (E) global chaotic rules: rules with random-looking spatial-temporal patterns, or with exponentially divergent cycle lengths as lattice length is increased, or a non-negative spatial response to the perturbations.

The relationship between these five classes (A, B, C, D, E) and the four classes defined by Wolfram in [12] (I, II, III, IV) is the following: Wolfram's Class-I rules are the same as null rules (A); Class-II rules are fixed-point rules and periodic rules (B and C); Class-III rules are global chaotic rules (E); Class-IV rules, which typically have long transients, are difficult to include in any of these categories. If we choose the spatial response to the perturbation as the criterion for nonchaotic or chaotic, the Class-IV rules then belong to chaotic rules. On the other hand, if the periodicity for the limiting configuration is chosen as the criterion (suppose the time goes to infinity while the lattice length is finite), the Class-IV rules are more appropriately classified as periodic rules.

To classify CA rules with behaviors between simple and random is not easy. Nevertheless, the number of such rules is much smaller than those in other "obvious" groups. For elementary CA rules, only Rule-54 and Rule-137 (or Rule-110) have the typical Class-IV behaviors. In this paper they are classified as global chaotic rules.

As for the locally chaotic rules, only three are identified (Rule-26, Rule-73, and Rule-154). Similar to Class-IV rules, they can also belong to either the periodic class by one classification scheme or the global chaotic class by another. Actually, in counting the statistics for the transition probability in the next section, the locally chaotic rules are grouped with global chaotic rules. In section 5, they seem to fit well into the periodic rule class in the mean-field description. In any case, the ambiguity in defining these "boundary rules" (Class-IV rules, the locally chaotic rules, etc.) will not affect the general conclusion about the structure of the rule space.

Finally, table 2 lists all the 88 independent elementary CA rules in five classes (A, B, C, D, E). One thing to remember is that even when the locally chaotic rules (D) are combined with the chaotic rules (E), the resulting four classes are still not the same as Wolfram's four classes in [12].

4. Structure of the CA rule space: Intra-class and inter-class connections

With all the elementary CA rules being classified in the last section, each point in the rule space can be "colored" according to its class, and this "colored" rule space presents a structure. By this definition, once each rule

Class	Rule number
null	0, 8, 32, 40, 128, 136, 160, 168
fixed point	2, 4, 10, 12, 13, 24, 34, 36, 42, 44, 46, 56, 57, 58, 72, 76, 77, 78, 104, 130, 132, 138, 140, 152, 162, 164, 170, 172, 184, 200, 204, 232,
periodic	1, 3, 5, 6, 7, 9, 11, 14, 15, 19, 23, 25, 27, 28, 29, 33, 35, 37, 38, 41, 43, 50, 51, 74, 108, 131, 133, 134, 142, 156, 178
locally Chaotic	26, 73, 154
chaotic	18, 22, 30, 45, 54, 60, 90, 105, 106, 129, 137, 146, 150, 161

Table 2: Elementary cellular automata rules as classified to five classes.

is classified, the rule structure is determined. The problem is how to visualize the eight-dimensional hypercube and how to describe the main features of the structure. In this section, attempts are made to characterize the rule space by quantitatively determining the connections within the class and between the classes. In the next section, we will cluster rules by their mean-field parameters, especially by using the t_0 and t_7 bit in the rule table.

Figure 3 shows the five classes of rules and the connection within the class, or *intra-class connections*. Since there are only three locally chaotic rules, they are incorporated into the group with chaotic rules. The rules in the picture are arranged by their λ value: small λ rules are on the top and λ is increased while moving down. There is no rule with λ larger than 1/2 in the figure because these rules are folded upon the ones on the top with $\lambda < 1/2$.

We have the following observations:

Null rules consist of a cube.

Fixed-point rules are almost divided into two parts, with only Rule-172 and Rule-44 bridging them.

Most of the periodic rules are connected to one piece, except Rule-74 and Rule-108, which are disjointed from others.

Most of the chaotic rules are in two clusters, with the exception of Rule-45 and Rule-105.

Although it is difficult to draw some general conclusions, we postulate that some features mentioned above will preserve in larger CA rule spaces: for example, null rules dwindle into a hypercube with smaller dimensions; and chaotic rules scatter into more clusters than rules in other classes.

A more complete description of the rule space should also include the connection between different classes, or *inter-class connection*. Figure 4 shows such a diagram. It is clear that there are less connections between classes with completely different behaviors. There are four connections between null rules and the chaotic rules, but peculiarly, all by flipping the bit t_7 (or t_0) in the rule table. We will discuss this point in the next section. Notice that even in lower-dimensional dynamical systems like the logistic maps, there are also some similar rare transitions from short periodic cycles to chaotic dynamics by tuning the control parameter (e.g., the "crisis" [13]).

Table 3 lists the number of connections within the classes (N_{ii}) and between the classes (N_{ij}), for both the original rule space and the folded rule space. We define the average intra-class transition probabilities as

$$w_{ii} = \frac{2N_{ii}}{2N_{ii} + \sum_{k,k\neq i} N_{ik}} \tag{4.1}$$

which are the probabilities for each rule in class i to be connected to another rule in the same class i; and inter-class transition probabilities:

$$w_{ij} = \frac{N_{ij}}{2N_{ii} + \sum_{k,k\neq i} N_{ik}} \tag{4.2}$$

which are the probabilities for rule in the class i to be connected to rules in the class j. The reason to use $2N_{ii}$ rather than N_{ii} is because, from the point of view of each rule in a class rather than the class itself, the link between two rules in the same class is used twice and so should be N_{ii}.

By the above definition, the average intra-class transition probabilities $\{w_{ii}\}$ in the folded rule space are

0.429	(null rules)	
0.535	(fixed-point rules)	
0.508	(periodic rules)	
0.295	(chaotic and locally chaotic rules)	(4.3)

and in the original rule space are

0.417	(null rules)	
0.526	(fixed-point rules)	
0.511	(periodic rules)	
0.337	(chaotic and locally chaotic rules)	(4.4)

As mentioned before, the reason to combine the locally chaotic rules with the chaotic rules is to remove the boundary classes in order to have a better statistics.

Roughly speaking, the intra-class transition probabilities are between 0.3 to 0.5. It shows a strong tendency for rules to stick to the rules from the same class. Although elementary CA rule space maybe too simple to be general, we speculate that the intra-class probabilities for larger CA rule spaces are within the same range.

	null	fixed pt.	periodic	locally c.	chaotic	total
null 8 (24)	12 (40)	24 (88)	4 (12)	0 (0)	4 (12)	
fixed pt. 32 (97)		57 (204)	51 (192)	9 (32)	15 (56)	
periodic 31 (89)			50 (182)	5 (18)	37 (126)	
locally c. 3 (10)				1 (4)	6 (22)	
chaotic 14 (36)					13 (36)	
total 88 (256)						288 (1024)

Table 3: Numbers of connections between classes for the folded rule space (first line in every array) and for the original rule space (the number inside the parentheses in every second line). The number of rules for each class is listed in the first column. "locally c." refers to locally chaotic rules.

5. Grouping rules by mean-field clusters

For CA, as well as other dynamical systems, some details about the rule are not important for the overall behavior of the dynamics. By ignoring these details, many dynamical rules can be considered to belong to the same group, which will be called *mean-field cluster* in this section.

For example, Rule-10, 12, 24, 34, 36 map block (000) to 0, map one of the three blocks containing one 1 to 1, map one of the three blocks containing two 1's to 1, and map block (111) to 0. Following a notation in [11], we use $[n_0n_1n_2n_3]$ to refer to the mean-field cluster containing rules mapping n_i of the 3-site blocks with i 1's to site value 1. Clearly, $n_0, n_3 \in (0,1)$, and $n_1, n_2 \in (0,1,2,3)$. By this notation, the above rules belong to the cluster [0110].

The mean-field clusters consist a *mean-field cluster space*. It is a "hyper-rectangular-block," because the width of the cluster space is 2 along n_0 and n_3 axes, but 4 along n_1 and n_2 axes. Notice that the rules in the same mean-field cluster will not be next to each other in the original rule space. For example, the Hamming distance between Rule-2 and Rule-4 (they are in the same cluster [0100]) is 2. On the other hand, two rules in two nearby clusters can be next to each other in the original rule space. For example, the Hamming distance between Rule-40 and Rule-42 (they belong to clusters [0020] and [0120] respectively) is 1.

As in the case of rule space, the cluster space can also be folded by the equivalence between clusters. It is easy to show that $(n_0n_1n_2n_3)$ and $(1-n_3, 3-n_2, 3-n_1, 1-n_0)$ represent the same dynamics under the 0-to-1 transformation. The 64 clusters are thus reduced to 36 independent ones.

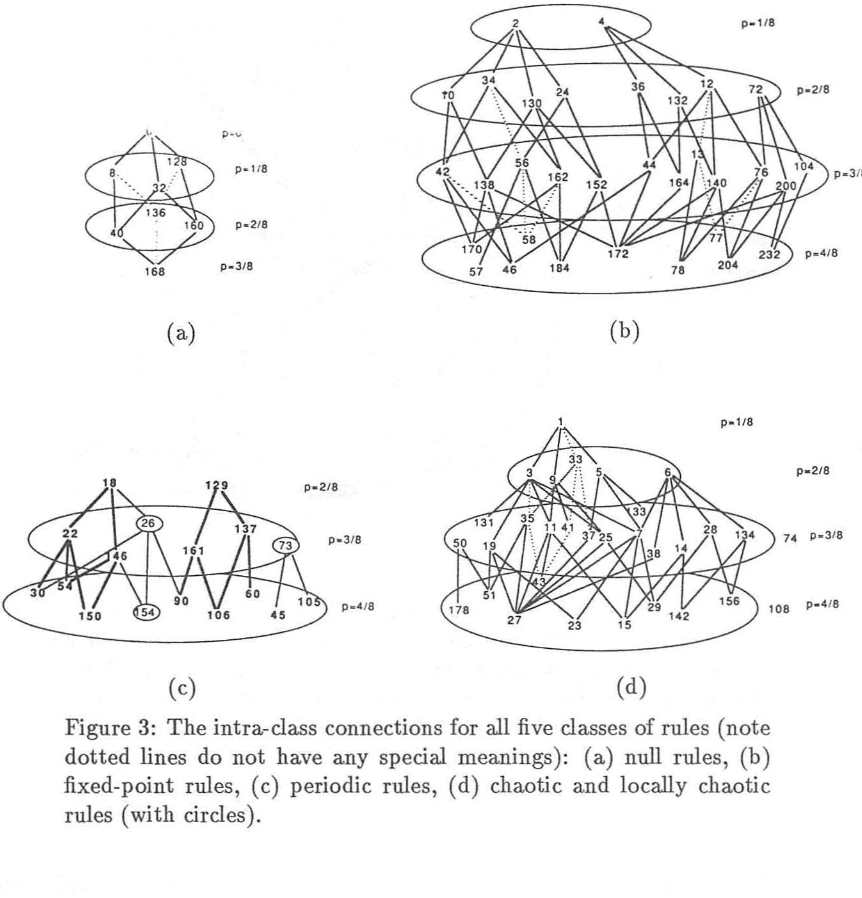

Figure 3: The intra-class connections for all five classes of rules (note dotted lines do not have any special meanings): (a) null rules, (b) fixed-point rules, (c) periodic rules, (d) chaotic and locally chaotic rules (with circles).

In order to view the space of mean-field cluster, three parts are sliced out: (a) $n_0 = n_3 = 0$; (b) $n_0 = 0$, $n_3 = 1$; (c) $n_1 = 1$, $n_3 = 0$. (Those with $n_0 = n_3 = 1$ can be transformed to case (a)). In each slice of the cluster space, two out of the four n_i's are fixed and the number of free parameters is 2. So these subspaces are very easy to visualize. Figure 4(a–c) show these three subspaces. Each cluster is "colored" by the dynamics of the rules. Since there are cases that rules in the same cluster have different dynamics (in another word, the mean-field theory fails to predict the behavior correctly by ignoring important details), more than one "color" is used.

In some sense, which site values block (000) and (111) map to is more crucial. So the bit t_0 and t_7 in the rule table are more important than other bits (they can be called *hot bits*). Once the hot bits are fixed, the features of the clusters are much easier to summarize.

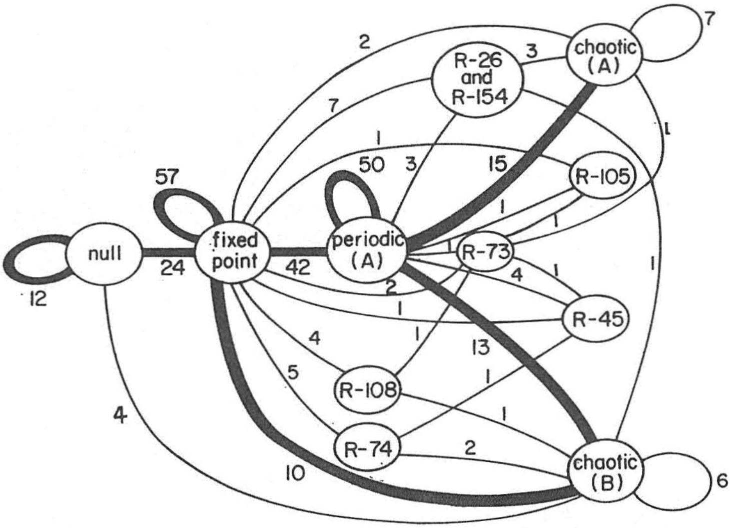

Figure 4: The inter-class connections. The number on each line is the number of connections between two clusters. A thick line indicates the number of connections being more than 10. The classes are aligned from left to right according to the degree of randomness. Cluster "Periodic(A)" refers to the biggest cluster for periodic rules, "Chaotic(A)" refers to Rule-18,22,30,54,146,150. "Chaotic(B)" refers to Rule-60,90,106,129,137,162.

The clusters with $n_0 = n_3 = 0$ can be called *nonlinear clusters*, which are shown in Figure 4(a), since the x_i^{t+1} versus $X_i^t = x_{i-1}^t + x_i^t + x_{i+1}^t$ function is reminiscent of the nonlinear logistic map. Nonlinear clusters are organized in such a way that by going down a path in the graph, one can have the transition from null to fixed point to periodic and then to chaotic rules. There are few cases when more random rules make transition to less random ones (e.g., from [0310] to [0320]), the overall tendency is nevertheless similar to other nonlinear systems with bifurcation phenomena.

The clusters with $n_0 = 0$ and $n_3 = 1$ (in Figure 4(b)) are called *linear clusters* for the similar obvious reason. Except for two cases (Rule-146 and Rule-150), most of the rules are null or fixed-point rules as well as a few periodic rules. It is consistent with the general consensus that linear systems would not generate complicated dynamics. Actually, $000 \rightarrow 0$ and $111 \rightarrow 1$ are necessary conditions for having fixed-point dynamics.

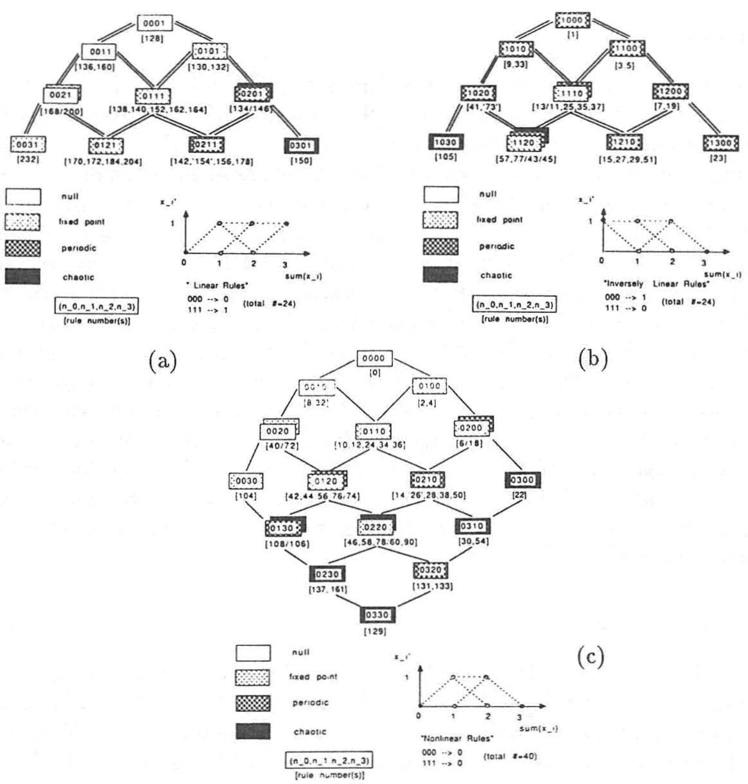

Figure 5: The CA rules grouped by the mean-field parameter $[n_0 n_1 n_2 n_3]$. (a) $n_0 = n_3 = 0$: nonlinear clusters. (b) $n_0 = 0, n_3 = 1$: linear clusters. (c) $n_0 = 1$, $n_3 = 0$: inversely linear clusters.

The clusters with $n_0 = 1$ and $n_3 = 0$ (in Figure 4(c)) are called *inversely linear clusters*, which are dominated by the periodic rules. These rules cannot have fixed-point dynamics simply because both $000 \rightarrow 1$ and $111 \rightarrow 0$ violate the invariant condition required by a fixed-point dynamics. It is probably somewhat related to the existence of oscillation in systems with negative feedback which might be responsible for the existence of many physiological rhythms [14].

To summarize this section, by fixing the "hot bits" t_0 and t_7, the structure of the cluster space becomes clear: the linear clusters tend to have rules with fixed-point dynamics, the inversely linear clusters tend to have rules with periodic dynamics, and the nonlinear clusters have the typical bifurcation transition with the increase of the parameters in the mean-field description.

Conclusion

We have studied the structure of the elementary CA rule space using two different representations: the original eight-dimensional rule space and the four-dimensional mean-field cluster space. In both cases, the separation of simple and random rules and the bifurcation-like phenomena when one moves around the rule space are observed. It is not clear, however, of how some other representations can change the structure of the rule space, for example, the one by "rotating" the eight axes, i.e., recombining the eight bits of the rule table in some symmetric form.[2] It is hoped that some axis transformation will reshuffle entries in the rules table in such a way that the separation between rules with different behaviors is "clean". The intra-class connection probability derived in this paper measures the stability of certain dynamical behavior under perturbations in the rule table. The range of the intra-class transition probabilities for elementary cellular automata (0.3 – 0.5) is amazingly close to the probability for an enzyme to preserve its function under single amino acid substitution (0.3 – 0.6) according to Ninio [15]. This similarity may reveal some deeper principle on how the phenotype is maintained when genotype is changed.

Acknowledgments

W.L. would like to thank Howard Gutowitz for discussion on the mean-field description of CA and Salman Burkie for discussion on the problem of counting the number of independent rules. This work is supported by National Science Foundation Grant PHY-86-58062 and Office of Naval Research Grant N00014-88-K-0293.

Appendix: The number of independent rules

In this appendix, we will show that 256 elementary rules are reduced to 88 independent rules under left-to-right, 0-to-1 transformation, and the joint operation of the two. For brevity, T_1 represents the 0-to-1 transformation and T_2 represents the left-to-right transformation. $T_3 = T_1 * T_2$ is the joint operation of the two.

Figure 6 shows all possibilities when the two transformations are applied to a CA rule R:

1. $T_1(R) = T_2(R) = R$, consequently $T_3(R) = R$.
2. $T_1(R) = T_2(R) \neq R$, so $T_3(R) = T_1(R) = T_2(R) \neq R$.
3. $T_1(R) = R \neq T_2(R)$, so $T_3(R) = T_2(R)$.

[2] A special case has been studied in [10].

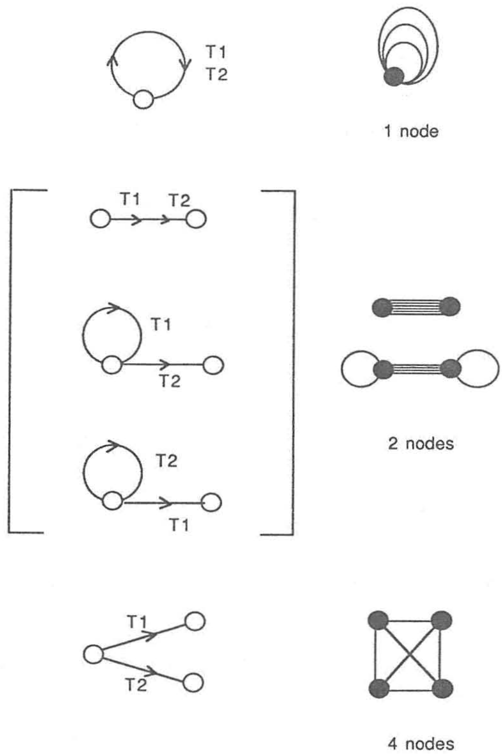

Figure 6: All possible outcomes of the 0-to-1 transformation (T_1) and the left-to-right transformation (T_2) applied on a CA rule. The graphs on the right side also include transformation $T_3 = T_1T_2$. There can only be 1, 2, or 4 rules in one equivalent group.

4. $T_2(R) = R \neq T_1(R)$, so $T_3(R) = T_1(R)$.

5. $T_1(R) \neq T_2(R) \neq R$, so $T_3(R) \neq T_1(R) \neq T_2(R) \neq R$. It can be proved by the method of contradiction, using the fact that $T_i(T_i(R)) = R$.

Case 1 will give clusters with 1 rule; Cases 2–4 lead to clusters with 2 rules; Case 5 gives clusters with 4 rules. Suppose the numbers of rules in the above five types are n_1, n_2, n_3, n_4, and n_5, then

$$n_1 + n_2 + n_3 + n_4 + n_5 = 256 \tag{5.1}$$

The total number of the clusters, or the number of independent rules, is $n_1 + (n_2 + n_3 + n_4)/2 + n_5/4$.

Now we determine the n_i's:

1. The rules satisfy $\bar{t_0} = t_0$, $\bar{t_7} = t_7$ and $t_1 = t_4 = \bar{t_6}$. There are $2^3 = 8$ possibilities, so $n_1 = 8$.

2. The rules satisfy $T_1(R) = T_2(R)$, or $\bar{t_7} = t_0, \bar{t_6} = t_4, \bar{t_5} = t_2$ and $\bar{t_3} = t_1$, is $2^4 = 16$. Subtracting the case where $T_1(R) = T_2(R) = R$, we get $n_2 = 16 - 4 = 8$.

3. The rules satisfy $T_1(R) = R$, which gives $2^4 = 16$ possibilities. Subtracting the case for $T_2(R) = T_1(R) = R$, we have $n_3 = 16 - 4 = 8$.

4. There are 2^6 rules with $T_2(R) = R$ ($t_0 = t_0$, $t_1 = t_4$, $t_2 = t_2$, $t_3 = t_6$, $t_5 = t_5$ and $t_7 = t_7$). Eight of them satisfy $T_1(R) = R$ and are subtracted, so $n_4 = 64 - 8 = 56$.

5. $n_5 = 256 - 8 - 8 - 8 - 56 = 176$.

Finally, the number of independent rules is $8 + (8 + 8 + 56)/2 + 176/4 = 8 + 36 + 44 = 88$.

References

[1] *Proceedings of an Inter-disciplinary Workshop on Cellular Automata*, J.D. Farmer, T. Toffoli, and S. Wolfram, eds., *Physica D*, **10** (1-2), (1984).

[2] Appendix of *Theory and Applications of Cellular Automata*, S. Wolfram, ed. (World Scientific, 1986).

[3] W. Li, N. Packard, and C. Langton "Transition phenomena in cellular automata rule space," *Physica D*, **44** (1990) to appear; W. Li, *Problems in Complex Systems*, Ph.D thesis (Columbia University, 1989); C. Langton, *Computation at the Edge of Chaos*, Ph.D Thesis (Michigan University, 1990).

[4] C. Langton, "Studying artificial life with cellular automata," in *Evolution, Games and Learning: Models for Adaptation in Machines and Nature*, J.D. Farmer, A. Lapedes, N. Packard, and B. Wendroff, eds. (North-Holland, 1986); *Physica D*, **22**(1-3) (1986) 120–149.

[5] R. May, "Simple mathematical models with very complicated dynamics," *Nature*, **261** (1976) 459–467.

[6] P. Schuster, "Structure and dynamics of replication-mutation systems," *Physica Scripta*, **35** (1987) 402–416.

[7] I.A. Campbell, J.M. Flesselles, R. Jullien, and R. Botet, "Random walks on a hypercube and spin glass relaxation," *Journal of Physics C*, **20** (1987) L47–L51.

[8] S. Wolfram, "Statistical mechanics of cellular automata," *Review of Modern Physics*, **55** (1983) 601–644.

[9] Karel Culik II and Sheng Yu, "Undecidability of CA classification schemes," *Complex Systems*, **2**(2) (1988) 177–190.

[10] Y. Aizawa and I. Nishikawa, "Toward the classification of the patterns generated by 1-dimensional cell automata," in *Dynamical Systems and Nonlinear Oscillations*, Giko Ikegami, ed. (World Scientific, 1986).

[11] Howard Gutowitz, "Classification of cellular automata according to their statistical properties," preprint, Center for Nonlinear Studies, Los Alamos National Lab, 1989.

[12] S. Wolfram, "Universality and complexity in cellular automata," *Physica D*, **10** (1984) 1–35.

[13] C. Grebogi, E. Ott, and J. Yorke, "Crises, sudden changes in chaotic attractors, and transient chaos," *Physica D*, **7** (1983) 181.

[14] Leon Glass and Michael C. Mackey, *From Clocks to Chaos: The Rhythms of Life* (Princeton University Press, 1988).

[15] J. Ninio, "Approches moléculaires de l'évolution," in *Collection de Biologie Evolutive*, **Vol 5** (Masson, 1979).