

Transition Phenomena in Cellular Automata Rule Space

Wentian Li^a, Norman H. Packard^{a,b,*}, Chris Langton^b

^aSanta Fe Institute, 1120 Canyon Road, Santa Fe, NM 87501;

^bComplex Systems Group, Theoretical Division,
Los Alamos National Laboratory, Los Alamos, NM 87545

published in Physica D, 45:77-94 (1990). figures are not included

Abstract. We define several qualitative classes of cellular automata (CA) behavior, based on various statistical measures, and describe how the space of all cellular automata is organized. As a cellular automaton is changed by varying entries in its rule table, abrupt changes in qualitative behavior may occur. These abrupt changes have the character of bifurcations in smooth dynamical systems, or of phase transitions in statistical mechanical systems. The most complex CA rules exhibit many of the characteristics of second-order transitions.

1. Introduction

Cellular automata were originally invented by von Neumann [18] for a particular task, to prove the existence of a self-reproducing universal computer. They have since caught the imagination of many, partly because of their rich and diverse phenomenology. One cellular automaton rule that has particularly rich phenomenology is Conway's game of life. Many other rules which display interesting phenomena are cataloged in [11].

In the space of all cellular automata, not all rules are interesting, of course. We will consider the question of how interesting a cellular automaton rule is on the basis of typical space-time configurations generated by the rule, *i.e.*, where a typical space-time configuration is generated by the rule acting on a random initial condition. In this case, there are two reasons for rules to be "uninteresting", and both reasons are based on the simplicity of a typical space-time configuration. One kind of simplicity is that the space-time configuration has a simple repetitive structure (possibly shifting in time). The other kind of simplicity is randomness, which is the lack of any discernible pattern.

Bifurcations occur in continuous dynamical systems as a parameter is varied to change from a particular dynamical system to a nearby one. Cellular

*Permanent address: Physics Department and the Center for Complex Systems Research, Beckman Institute, University of Illinois, 405 N. Mathews St., Urbana, IL 61801. Email n@complex.ccsr.uiuc.edu

automata are discrete dynamical systems; there are no continuous parameters to vary to move from one cellular automaton rule to an arbitrarily nearby rule. The space of rules is totally discrete; any change in a rule entails a “quantum jump” away from the rule.

There may nevertheless be changes in qualitative behavior as a rule is changed. In spite of the discrete nature of cellular automata, there is one unambiguous notion of proximity in the space of rules given by Hamming distance. This work is partly an exploration into the question of whether observables behave continuously in the limit that the size of the rule table becomes large and small changes (unit Hamming distance) are made between rules.

The term *bifurcation* has come to be used in a general sense within the context of continuous dynamical systems to denote any change in a qualitative behavior as a parameter in the system is varied, such as a change from periodic to chaotic behavior. Bifurcations have been studied in lattice systems with continuous variables at each lattice site [4], and in cellular automata obtained by discretizing continuous variables [2]. We will use the term bifurcation even more generally in the present work to include the qualitative change in the asymptotic behavior of a cellular automaton rule when one entry in its rule table is changed. We will call such a change a *unit transition* in the space of rules.

In the study of both phase transitions in statistical physics and bifurcations in dynamical systems, control parameters are needed. By varying these control parameters, the systems change from ordered states (or regular dynamics) to disordered states (or random dynamics.) Following [5], we will use a control parameter λ , which is defined as the percentage of all the entries in a rule table which map to non-zero states.

Suppose we are looking at 1-dimensional k -state, r -radius (or $(2r + 1)$ -neighbor) cellular automata. Each rule is specified by a rule table with k^{2r+1} entries, determining which state each possible neighborhood configuration is mapped to at the next time-step. If m out of the total k^{2r+1} neighborhood configurations map to a non-zero state, then λ is defined as:

$$\lambda = \frac{m}{k^{2r+1}} \quad (1.1)$$

The λ parameter can be compared with — though it is not equivalent to — the temperature in statistical physics, or the degree of nonlinearity in dynamical systems.

The rule spaces for cellular automata with 2 states per cell have a symmetry with respect to the point $\lambda = 0.5$. This is because rules with $\lambda = x$ are equivalent to rules with $\lambda = 1 - x$ for $0 \leq x \leq 0.5$; the roles of states 0 and 1 are simply reversed.

Most of the calculations included in this paper are for 2-state cellular automata, and we will only examine rules for which $\lambda \leq 0.5$. For cellular automata with $k > 2$ states, one can define a $k - 1$ dimensional parameter $\vec{\rho}$, which exhibits symmetry around the point at which each of the k states occurs equally, with frequency $1/k$, in the rule-table [7].

The simplest 1-D cellular automaton rule space is that for 2-state, 3-neighbor cellular automata (called “elementary rules” in [19]). As an illustration of the way in which the dynamics of these rules changes as λ changes, we plot the percentage of rules belonging to the four different classes as a function of λ (fig.1). (The overall classification will be discussed in the next section.) As λ is increased from 0 to 0.5, more and more rules exhibit chaotic, rather than periodic, dynamics. Since the elementary rule space is small (only $2^8 = 256$ possible rules in all), many features exhibited by this rule-space are non-generic, and may not be applicable to larger rule spaces. For example, the existence of a rather large percentage of periodic rules at $\lambda = 0.5$ is not observed in larger CA rule-spaces. This paper is devoted to the study of “large” cellular automata rule spaces. For more details about the elementary rule space, see [10].

The main point of this paper is to summarize experiments suggesting that, for sufficiently large CA rule-spaces, one observes a *phase-transition* between *ordered* and *disordered* dynamics as λ is varied from 0 to $1 - 1/k$. Furthermore, the various qualitative classes are ordered with respect to this transition regime, with the simplest behaviors (fixed and random) being located *away* from the transition regime while the most complex behaviors are found *within* the transition regime.

The paper is organized as follows: Section 2 discusses the classification of cellular automata rules; Section 3 discusses transition-like phenomena observed using difference patterns; Sections 4, 5, and 6 discuss the transition events in terms of entropy, mutual information, and fluctuations, respectively; Section 7 discusses the limit of large neighborhoods; and Section 8 discusses the inferred structure of CA rule-space. Appendix A gives the definitions of the quantities that are used in characterizing the transitions; and Appendix B presents a mean-field theory estimation of the spreading rates of difference patterns.

2. Classification

Classification of cellular automata is a tricky business. Wolfram developed a classification scheme consisting of four qualitative classes – homogeneous fixed point (class I), periodic (class II), chaotic (class III), and complex (class IV) [20]. Here, we list six classes of behavior typically observed in the dynamics of 1-D cellular automata which refine the Wolfram classes. One can provide quantitative descriptions which distinguish these classes in terms of various statistical measures, such as difference pattern growth rates, entropy, and mutual information (see Appendix A for definitions). Other approaches to the classification of CA rules are detailed in the papers by Chaté and Manneville [1], Gutowitz [3], and McIntosh [12] in this volume.

However, *classification alone is not enough!* What is needed is a deeper understanding of the structure of cellular automata rule space that provides an explanation for the existence of the observed classes and their relationship to one another. Choosing an appropriate parameterization of the space

of cellular automata rules, such as λ , allows direct observation of the way in which different statistical measures are related as a function of the parameter(s), and these relationships in turn provide an explanation for the existence and ordering of the various qualitatively distinguishable classes of cellular automata dynamics.

In the numerical experiments reported here, all measures are computed on cellular automata starting from random initial conditions, letting transients die away, and gathering statistics on a large enough space-time volume to achieve convergence. The question of statistical convergence is nontrivial, especially for the class 6 rules to be defined below.

Boundaries of classes are not necessarily sharp, and there is no known single statistic to distinguish between all classes. To decide empirically which class a rule belongs to on the basis of the statistical measures used here, one must inevitably choose threshold values for the measures, thus the classes are not sharp.

One can readily distinguish six classes on the basis of differences in the various measures as applied to the asymptotic behavior of 1-dimensional cellular automata:

1. *Spatially homogeneous fixed points.* The difference pattern spreading rate and all entropy and mutual information measures are identically zero.
2. *Spatially inhomogeneous fixed points,* or a uniform global shift of a fixed pattern. For this class, the difference pattern spreading rate is zero. Spatial entropy is finite, since the fixed pattern is random to a certain degree (as a result of the randomness of the initial condition). Entropy in other directions in space-time will be less than the spatial entropy, and goes to zero in the direction of periodicity, which is the direction of the net shift of the fixed pattern (the time direction, if the fixed pattern is unshifted). Mutual information will go to zero for all space-time directions as the separation between space-time patches becomes large.
3. *Periodic behavior,* or shifted periodic behavior. Typically this means regions with periodic behavior with unmoving walls between them. The behavior of statistical quantities will typically be the same as for the previous class, except that entropy and mutual information will go to nonzero values in the space-time direction of periodicity.
4. *Locally chaotic behavior.* These rules produce chaos between walls, where the walls can either be fixed or move with an overall shift. In two dimensions, this class is characterized by fixed or oscillating boundaries separating chaotic domains. In the one-dimensional case, since there is only a finite number of states at each site, the space-time pattern between the walls must be periodic, but the hallmark of these rules is that the period increases exponentially with the distance between the walls. For the one-dimensional case, entropy is positive in all directions,

mutual information is zero in all directions, and the difference pattern spreading rate is zero (because the difference pattern is stopped by the walls).

5. *Chaotic behavior.* The difference pattern spreading rate is high, entropy is high in all directions, and mutual information is zero in all directions.
6. *Complex behavior.* This class is characterized by long transients and complex space time patterns, including both oscillating and propagating structures. This class is also characterized by a lack of statistical convergence, for it is not clear that the assumptions needed for computation of statistics hold for this class of rules. When computation of statistics is attempted, entropy is moderate, the spreading rate is roughly zero, and the mutual information is large.

This classification scheme refines Wolfram's four classes. Roughly, class 1 above corresponds to Wolfram's class I; classes 2, 3, and 4 constitute Wolfram's class II; class 5 is equivalent to Wolfram's class III; and class 6 is Wolfram's class IV.

3. The transition using difference patterns

One statistical quantity that distinguishes chaotic behavior from ordered behavior is obtained from the average asymptotic motion of the difference pattern (see the Appendix A.1 for a definition). This motion generally describes how two configurations that are different on part of the lattice and the same on another part, become either increasingly different or increasingly the same under the action of the cellular automaton rule.

We compute the difference pattern spreading rate, γ , along a path in the space of rules, where each successive rule on the path has a higher value of λ (more 1's in the rule table) than the previous one. We may choose an arbitrary threshold for γ , above which we will say the rule is chaotic. The first rule on the path having γ above the threshold is the "transition point" for that path.

As discussed below, not all paths undergo a transition at the same value of λ . Fig.2 shows the transition to chaos averaged over many paths, where these paths are aligned by their transition points. Most paths exhibit a sharp jump in γ at the transition point. However, a few paths exhibit intermediate values of γ ; these are class 6 rules (complex rules), which exhibit the most irregular statistics.

For small λ , perturbations hardly ever spread, for large λ they always spread at roughly the same amount. Thus, away from the transition region, any particular value of λ is associated with a very narrow range of spreading rates. However, at the transition, one sees a *wide* range of possible spreading rates. Thus, rules that are characterized by a scaling relation in the size of their response to perturbations must lie in the transition region. Not all rules in the transition region, however, need have such uniform scaling.

4. Transition to chaos using entropy

The most commonly used quantity for measuring randomness is entropy (see Appendix A.2 for the definition). For spatially homogeneous fixed point rules, the entropy calculated from the spatial-temporal pattern is zero. For periodic rules, the entropy value is non-zero but low. For random rules, all possible configurations can occur and the entropy reaches its maximum.

Fig.3a shows the single site entropy as a function of λ for 50 different paths through the space of possible 2-D, 8-state, 5-neighbor cellular automata. The square lattice has 64 cells on a side, the first 500 transients are discarded, and the next 500 patterns are used to accumulate single site probabilities. There are some restrictions on the rules being chosen: the all-zero neighborhood maps to zero, and the rules are symmetric with respect to planar rotations [6]. As expected, the entropy generally increases with increasing λ , and one usually observes a sharp jump in the entropy at the transition from regular to chaotic dynamics.

The entropy reaches its maximum value $S_{max} = \log_2(8) = 3.0$ at $\lambda = 7/8$, when the rule tables are filled randomly and uniformly with respect to all 8 states. (For 2-state cellular automata, $S_{max} = \log_2(2) = 1.0$ at $\lambda = 1/2$, since only two symbols are used to fill the rule table randomly.)

Fig.3a is redrawn in fig. 3b by aligning paths according to the particular λ_c at which a transition from regular dynamics to chaotic dynamics occurs on each path. In this case, the transition point was chosen to be the λ value at which no cycle of length less than 500 was detected.

For most paths, the entropy jumps discontinuously, an indication of the sudden transition from regular dynamics (either fixed point or periodic) to chaotic dynamics. In statistical physics, such a discontinuous change in entropy accompanies a first-order phase transition. By contrast, a smooth change in entropy — such as is observed for a small percentage of the paths — is associated with a second-order phase transition.

An interpretation of this is that some paths through CA rule space pass through a region of “critical” rules, which are interposed between periodic rules and chaotic rules. When such a region is encountered, the entropy appears to change smoothly, i.e., the transition seems second-order. These “critical” rules are obviously not dense, if they were, *every* path would exhibit a smooth change in entropy. Fig.3b suggests that a first-order transition is more likely than a second-order transition, due to the low probability of passing through a “critical” region by chance.

The functional form of the entropy versus λ curve in fig.3a can be explained by a simple “mean-field theory” [21]. For cellular automata without any restrictions, a crude estimate of the probability of non-zero sites is simply equal to λ , with each of the $k - 1$ non-zero states appearing with equal probability. The estimate of entropy is obtained from

$$S_{m.f.} = -(1 - \lambda) \log(1 - \lambda) - \lambda \log\left(\frac{\lambda}{k - 1}\right). \quad (4.1)$$

As mentioned above, for the cellular automata whose entropy is plotted in

fig.3a, there is the restriction that the all-zero neighborhood maps to zero. A modified mean-field theory can be used to estimate $S_{m.f.}$ in this case (for more details see [21]). Notice that in the mean-field theory approximation entropy changes continuously, suggesting a second-order transition rather than the first-order transition which is observed more often in numerical simulations.

5. Transition to chaos using mutual information

Mutual information is a quantity for measuring correlation. If two random variables are statistically independent, the mutual information between them is zero. On the other hand, if the two are strongly correlated (e.g., one is a copy of another), the mutual information between them is large (see Appendix A for the definition of mutual information).

The spatial mutual information for homogeneous fixed point rules is zero, and low for both periodic (including inhomogeneous fixed point rules) and random rules, because these rules do not create spatial structures. On the other hand, most complex rules give rise to highly correlated structures and consequently large spatial mutual information.

As an example, we compute the mutual information for 1-D, 2-state, 11-neighbor cellular automata. In order to suppress fluctuations, we calculate the spatial mutual information between two blocks of sites (rather than between two individual sites). Also, many copies of the spatial configurations at different time steps (instead of just one configuration at a fixed time) are used in accumulating the probability distribution.

Fig.4a shows the spatial mutual information for rules along 120 paths moving toward larger values of λ . Each step corresponds to adding 7 non-zero transitions to the rule table, an increase in λ of $7/2^{11} = 0.0034$ per step. The mutual information is computed between two blocks of 3-sites each, separated by a distance of 10 sites. Lattice size is 601, 257 transient configurations are discarded, and the next 51 configurations are used to accumulate the probabilities.

From fig.4a, we can see that the spatial mutual information for different paths jumps at different values of λ , i.e., they have different values for λ_c . In order to characterize the behavior near the transition point, we may align the different paths by the relative λ value, $\Delta\lambda \equiv \lambda - \lambda_c$.

There are two ways to determine λ_c . The first method defines λ_c as the value of λ at the first jump of mutual information over some threshold value. The resulting plot is shown in fig.4b.

The second method defines λ_c as the value of λ at which the maximum value of mutual information is observed. The second method is equivalent to the first if the threshold value is raised to the maximum value of mutual information for that path. Fig.4c illustrates the second method.

As the second method results in a curve which is symmetric with respect to the critical point, suggesting a classic λ -transition, we will examine the second method in more detail. The first thing we want to know is whether the mutual information at the critical point diverges as a power law or as

an exponential. The data of fig. 4c is re-plotted in fig.5 using semi-log scales. This plot gives a better fit than a log-log plot (which is not included here). This means that the divergence is exponential rather than power law: $M(|\Delta\lambda|) \sim e^{-\alpha|\Delta\lambda|}$, with $\alpha \approx 12$.

If the mutual information, which is basically a difference between entropies, can be considered as a quantity analogous to specific heat in statistical physics (the derivative of entropy with respect to temperature),¹ the behavior of the mutual information at the critical point can be used to determine the order of the phase transition. As has been seen in the previous section, for most paths the entropy jumps discontinuously at the transition point, so the transition is usually first-order. Notice that second-order phase transitions in statistical physics typically have power law divergence of specific heat, instead of exponential divergence. Unlike specific heat, the mutual information will not diverge to infinity if the number of states is finite.

Fig.6b shows the spatial mutual information in semi-log scale similar to that in fig.5, except that the distance between the two 3-site-blocks is only 4 sites instead of 10. (Fig. 6a shows the original M versus λ plot similar to fig. 4a.) The average maximum value of the mutual information is more or less unchanged. This is probably because the spatial structures generated by complex rules have very long correlation lengths, so a change of distance from 10 to 4 does not change the mutual information value very much. On the other hand, the mutual information is greatly reduced at longer distances for periodic and random rules, as can be seen in the plots. As a result, the exponential decay rate around the maximum mutual information is different: $M(|\Delta\lambda|) \sim e^{-\alpha|\Delta\lambda|}$, where $\alpha \approx 8$.

Fig.7b shows the spatial mutual information in semi-log form for 15-neighbor rules. (Fig. 7a is the original M versus λ plot.) The separation between the two 3-site-blocks is again 10 sites. The exponential decay from the maximum mutual information seems to be faster than for $r = 5$ rules. It is not clear how the decay rates — which are somewhat analogous to the critical exponent — are related to the rule radius, block-length, or the separation between blocks.

6. Fluctuations at the transition to chaos

The transition to chaos appears to be marked by singular behavior in statistical measures such as entropy and difference pattern spreading rate. At the transition, however, there are problems computing these measures because of a lack of statistical convergence.

Rules at the transition have been conjectured to be capable of universal computation [5, 6, 14, 20]. If this is the case, for generic initial conditions, then the computation of these statistical measures is problematic, because the limit set of the automaton is uncomputable, and therefore all statistical quantities are uncomputable. Rules at the transition exhibit very long transients, possibly infinite.

¹A similar comparison in the context of 2-dimensional Ising model is made in [16].

One can attempt to compute statistical quantities in any case, with the usual prescription of filling probability histograms over volumes of space-time. Lack of convergence is then apparent in a large spread of values for the measures when computed over an ensemble of initial conditions. The distribution of values for a measure should be close to a delta function if the empirical estimate is valid, *i.e.* if statistical convergence is attained. The width of the distribution of values is a measure of the lack of convergence.

The poor convergence of statistical properties for complex rules is illustrated in fig.8, which plots the distribution of the spreading rate of the difference pattern for elementary rule-110 (a complex rule) and for elementary rule-30 (a chaotic rule). The distribution is accumulated over an ensemble of different initial conditions. It is clear that rule-110 has a wider distribution of spreading rate.

7. Large rule table limit

When more neighbors are involved in updating each site, site values become increasingly sensitive to sites at larger distances. One might suppose that this increased interdependence among sites would make random dynamics more likely. In this section, we will show that this is indeed the case, and that the critical point λ_c which separates regular and chaotic rules approaches zero in the large neighborhood limit (the “thermodynamic limit”).

Since spatial mutual information is a measure of correlation among the parts of a configuration, as discussed in Section 5, searching the peak in the mutual information relation is an effective way to locate the transition region. In order to study the effect of neighborhood size on the location of the transition region, we calculate the spatial mutual information versus λ for different radius rules.

The results for rules with radius $r = 3, 5, 7, 9$, and 11 are shown in fig.9 (the plot is taken from [8]). The mutual information is computed between two blocks of 3-sites each, separated by a distance of 8-sites. The lattice size is 167, 118 transient configurations are discarded, and the next 89 configurations are used to accumulate the probability. Rules are picked randomly for each λ value, rather than by repeatedly adding non-zero values in the rule table.

Fig.9 shows clearly that the transition region — as indicated by the peak of mutual information — moves toward smaller values of λ as r is increased. In other words, a larger proportion of cellular automata rules generate random dynamics as the radius of coupling r is increased. If we determine λ_c as the location of the peak in the spatial mutual information, then λ_c as a function of the number of neighbors $n = 2r + 1$ is shown in fig.10.

How does λ_c change in the $n \rightarrow \infty$ limit (fully connected cellular automata)? A numerical simulation becomes more and more difficult since the size of the rule table expands as 2^{2r+1} , which quickly reaches the limit of storage space. Nevertheless, some simple considerations lead to the conclusion

that

$$\lim_{n \rightarrow \infty} \lambda_c = 0. \quad (7.1)$$

There are many other ways to define λ_c . Another definition of λ_c is by the onset of non-zero spreading rates for difference patterns. By mean-field calculations, included in Appendix B, we can easily derive λ_c as a function of the radius r (Eq.(B.8)):

$$\lambda_c = \frac{1}{2} - \frac{1}{2} \sqrt{1 - \frac{2}{r+1}}. \quad (7.2)$$

In the large r limit, this λ_c goes to zero:

$$\lambda_c \approx \frac{1}{2(r+1)} + O\left(\frac{1}{r^2}\right). \quad (7.3)$$

The case for $\lambda_c = 0$ only means that it becomes increasingly hard to find rules with non-random dynamics in the large neighborhood limit. This does *not* imply that there are no non-random rules at all. We will discuss this point in the last section.

λ_c as determined by Eq.(7.3) is also plotted in fig.10, as well as $\lambda_c = 1/(2r+1)$ for the onset point of non-zero entropy estimated by mean-field theory [21]. Numerical estimates of the transition point always exceed the mean-field estimates.

8. Structure of rule space

Piecing the above results together, a clear picture of the fundamental structure of cellular automata rule spaces emerges, although there are still some details that need to be worked out.

There are two primary regimes of rules — *periodic* and *chaotic* — separated by a *transition regime*. This transition regime is *not* simply a smooth surface separating the other two domains, but itself has a complicated structure. Most of this transition regime seems to be simply a boundary between periodic and chaotic rules, containing no rules within it. Crossing the transition regime at such a boundary gives rise to a discrete jump in statistical measures of the dynamics, as is seen in first-order transitions.

However, other parts of this transition regime seem to have some “thickness”, in the sense that they contain the so-called “critical” rules. Crossing the transition regime through these areas gives rise to smooth changes in statistical measures, suggesting a second-order transition.

This basic picture is illustrated schematically in fig.11. Notice that this simple picture of a phase-transition separating a domain of ordered dynamics from a domain of disordered dynamics (each of which might be further subdivided), provides a simple explanation for the existence of the four qualitative classes of cellular automata behaviors identified by Wolfram and for the six qualitative classes identified earlier in this paper.

More importantly, however, this picture also provides an explanation for the relationship that obtains between these various qualitative classes, which has been lacking previously. The relative locations of these various classes are indicated in fig.11. The complex rules (Wolfram’s class IV and our class 6) are found in the transition regime separating the periodic rules (Wolfram’s classes I and II, and our classes 1,2,3, and 4) from the chaotic rules (Wolfram’s class III and our class 5).

Fig.12 plots one of our statistical measures (mutual information) over the $\vec{\rho}$ parameter space (discussed earlier) for 2-D, 3-state CAs, in order to illustrate empirically the relative locations and sizes of the various regimes of CA behaviors. The chaotic rules, the largest class, occupy the vast central depression. The fixed point and periodic rules lie to the outside of the peaks in the mutual information surface, towards the vertices of the base-triangle. The “complex” rules lie in the vicinity of the peaks in the mutual information surface. In this plot, the λ parameter would constitute a vertical slice through the surface, running from one of the vertices of the base-triangle to the center of the triangle (the middle of the central depression), with $\lambda = 0$ at the vertex and $\lambda = 1 - 1/k$ at the center. The mutual information surface has been smoothed in this figure by averaging over the runs of neighboring ρ -points.

The relative location of the different qualitative classes leads immediately to a fundamental conjecture about complex behavior in general. Since we have uncovered an apparent association between the most complex rules and second-order transitions, and since these complex rules seem to be associated with a capacity for computation — even universal computation — we have *a fortiori* uncovered a fundamental connection between complex behavior, computation, and phase-transitions — especially second-order transitions. Indeed, one can identify analogs for characteristic properties of computing systems in the phenomenology of phase-transitions [7]. Examples include the existence of different “complexity classes” associated with increases in transient time as one approaches a phase-transition, and an analog of Turing’s famous Halting problem in the “critical slowing down” associated with second-order transitions.²

Another observation that can clearly be made is that the λ parameter alone is insufficient for locating specific dynamical regimes precisely. For many physical systems exhibiting phase transitions, more than a single parameter is required to accurately reveal the phase-transition structure. For instance, the transition point from a solid to a fluid is not captured precisely by temperature alone, one must also control the pressure. For any specific pressure, there is a unique melting temperature, but if pressure is not being controlled carefully in an experiment, one will observe a range of temperatures at which melting will be observed to occur.

This suggests that we will have to find at least one more parameter affecting the dynamics of cellular automata before we can fill out all of the details of the transition or bifurcation structure of cellular automata rule-spaces.

²This latter association was observed independently by Vichniac *et al.* in [17].

The situation is illustrated schematically in fig.13. In this figure there are two primary domains of behavior, an ordered domain — the periodic rules — and a disordered domain — the chaotic rules. These two domains are separated by a transition regime with a complicated structure. Part of this transition regime — the dark-shaded region in the figure — contains the “critical” rules. The λ parameter controls the location along the abscissa, while some unknown parameter controls the location along the ordinate. We can see that, without controlling the specific value of the “mystery” parameter, we would expect to see transitions — both first and second-order — over a wide range of λ values, with the critical rules restricted to lie within the boundaries delimited by λ_{min} and λ_{max} .

There are various proposals for the “mystery” parameter(s), including the possibility of defining generalized versions of the thermodynamic quantities of temperature, pressure, density, energy, and so forth [7]. Other proposals involve various quantities derivable from mean-field theory [21], or replacing the 1-dimensional λ parameter by a subset of the mean-field theory parameters [3]. So far, none of the proposals provide a completely satisfactory theoretical framework.

9. Conclusion

The space of cellular automata rules seems to be divided into regions of similar behavior. Division into regions makes sense only with a metric on the space of rules. We use Langton’s λ parameter, which is simply the percentage of ones in the cellular automaton rule table (in the case of binary state automata.)

The largest region by far consists of chaotic behavior. The next largest region consists of fixed point (or shifting fixed point) behavior, and the other classes — periodicity, local chaos, and complex behavior — all appear to occupy a small region of the rule space. This region appears to contain something very much like a *phase-transition* between ordered and disordered CA dynamics.

Since the space of cellular automaton rules is finite and discrete, λ can take on only a finite number of values. One can, nevertheless, move in the space of rules along paths where λ increases monotonically in minimal increments. For a binary state automaton, this is accomplished by incrementing the number of ones in the cellular automaton rule table. This is a discrete analog of moving along a bifurcation arc in a space of smooth dynamical systems.

We find that as we increase the λ parameter, statistical quantities, such as entropy and difference pattern spreading rate, often undergo abrupt changes at the transition between ordered and disordered behavior. These discontinuous changes involve a jump, and have the character of a first-order phase transition. Occasionally, however, a smooth change in statistical measures is observed, indicating that something like a second-order phase-transition is possible. In these latter transitions, phenomena analogous to “critical slow-

ing down” and the droplet-formation responsible for “critical opalescence” are observed.

For “critical” rules — rules at or very near such a second-order transition — statistical quantities do not converge well. We propose a measure of the lack of convergence for such rules: the width of the distribution of values of a statistic. Using this measure, rules at the transition are seen to have poor statistical convergence. This provides evidence for the hypothesis that rules at the transition are capable of nontrivial (possibly even universal) computation, even with random initial conditions. At the transition, a mixture of stability with respect to some local configurations and instability with respect to others may well provide a cellular automaton rule with the necessary requirements to both store and transmit information, two requirements crucial to computation [5, 6, 14].

Although we have presented a general picture of cellular automaton rule space, the detailed structure is complicated. Since an elementary rule (1-D, 3-neighbor) can be considered as a 5-neighbor rule with no effects from the outer 2 sites, the elementary rule space is contained in the larger rule space of 5-neighbor cellular automata, which again is contained in even larger rule spaces. We know from fig.1 that there are periodic elementary rules at $\lambda = 0.5$, and these rules will persist at $\lambda = 0.5$ in larger rule spaces, even as $\lambda_c \rightarrow 0$. The implication is that as the rule space becomes larger, it is more difficult, but not impossible, to find periodic rules at $\lambda > \lambda_c$.

Although the λ parameter is not sufficient for locating the different dynamical regimes within cellular automata rule-spaces precisely, it is sufficient to reveal important structural details of these spaces. Furthermore, the structure revealed explains not only the existence of, but also the relationship between, previously proposed qualitative classes of cellular automata dynamics. Perhaps most importantly, the association of complex rules with second-order transitions in the space of rules suggests a fundamental connection between computation, complexity, and phase-transitions.

Acknowledgements

We would like to thank J. Crutchfield, D. Farmer, H. Gutowitz, H. Hartman, E. Jen, S. Kauffman, R. Shaw, and B. Wootters for discussions. The work is supported by NSF Grant PHY-87-14918 and DOE Grant DE-FG05-88ER25054 at SFI, by NSF Grant PHY-86-58062 and ONR Grant N00014-88-K-0293 at CCSR of Univ. of Illinois.

Appendices

A. Definition of statistical quantities

A.1 Spreading rate of difference patterns

The left-moving difference pattern is obtained by first taking a configuration $a = \{\dots a_{-1}, a_0, a_1, \dots\}$ and constructing another configuration $a' = \{\dots a'_{-1}, a'_0, a'_1, \dots\}$ with $a_i = a'_i$ for $i < 0$, $a_i \neq a'_i$ for $i = 0$, and a'_i chosen at random (independent of the a_i) for $i \geq 0$. The difference pattern at time t is then defined as $\delta^t = 0$ if $a^t = a'^t$ and $\delta^t = 1$ otherwise, where a and a' evolve in time according to the action of the cellular automaton rule. The location of the front of the difference pattern is $i_f^t = \min\{i | \delta_i^t = 1\}$. The left-moving difference rate is then defined as

$$\gamma_{left}(a) = \lim_{t, \tau \rightarrow \infty \text{ with } t/\tau \rightarrow 0} \frac{i_f^\tau - i_f^{\tau+t}}{t - \tau} \quad (\text{A.1})$$

providing the limit exists. The right-moving difference rate is defined similarly, with δ initially zero on the other half of the lattice, $\delta_i = 0$ for $i > 0$, and now letting $i_f^t = \max\{i | \delta_i^t = 1\}$

$$\gamma_{right}(a) = \lim_{t, \tau \rightarrow \infty \text{ with } t/\tau \rightarrow 0} \frac{i_f^{\tau+t} - i_f^\tau}{t - \tau} \quad (\text{A.2})$$

providing the limit exists. The total difference spreading rate is then defined to be

$$\gamma(a) = \gamma_{left}(a) + \gamma_{right}(a). \quad (\text{A.3})$$

Here, we will define a rule to be chaotic if $\gamma > 0$. $\gamma(a)$ is observed to be independent of a , if a is chosen at random, and we will often simply refer to the generic asymptotic spreading rate simply as γ . γ is measured empirically below by taking a and a' to differ initially at only one site, and measuring the rate that both the left-moving and right-moving fronts move at the same time. Transients are allowed to die out and an average is taken over many initial values of a and a' .

A.2 Entropy

Given a probability distribution $\{p_i\}$, the definition of entropy is well known:

$$S = - \sum_i p_i \log(p_i). \quad (\text{A.4})$$

There are many ways entropy can be applied to the spatial-temporal patterns of cellular automata. For example, the single site spatial-temporal entropy is calculated by counting the occurrence of all symbols $\{c_i\}$ ($i = 1, \dots, k$) from the spatial configurations at many time steps, and the probability distribution is $\{p_i\} = \{c_i / \sum_j c_j\}$.

Sometimes, we are interested in characterizing the randomness in a particular direction in the spatial-temporal pattern, then the probability distribution is accumulated along that direction, rather than the whole spatial-temporal pattern. For example, the spatial entropy, when we only analyze the configuration at a fixed time step; and temporal entropy, when only the time sequence on a particular site is studied. These entropies are called “directional entropies” [13].

A.3 Mutual information

Mutual information is a function of two probability distributions $\{p_i\}$ and $\{p_j\}$. In order to characterize the correlation between the two distributions, we also need the joint probabilities for both event i in the first variable and event j in the second variable to occur: $\{p_{ij}\}$. The mutual information between the two distributions is defined as [15]:

$$M = \sum_i \sum_j p_{ij} \log \frac{p_{ij}}{p_i p_j} \quad (\text{A.5})$$

To apply mutual information to the spatial-temporal patterns of the cellular automata, the two probability distributions can be two probabilities having k values on two sites separated by distance d [9]; they can also be probabilities having k^l values on two l -block's separated by d . If the distance is a spatial distance, we are calculating the spatial mutual information. If the distance is a time delay on the same site, we are calculating the temporal mutual information. In general, mutual information between two sites (or two blocks) separated by any spatial-temporal distances can be determined.

B. Estimation of the spreading rate of difference pattern by mean-field theory

Considering 1-dimensional, 2-state, $(2r + 1)$ -neighbor cellular automata, the maximum possible spreading rate of the difference pattern to the left (and right) is r , the range of the coupling, and the total maximum spreading rate is

$$\gamma_{max} = 2r. \quad (\text{B.1})$$

Typically, this maximum spreading rate of difference patterns will not be reached because expansions of perturbation with smaller spreading rates can also occur with non-zero probability. Suppose we are looking at rules with λ parameter value, i.e., λ percent of the $(2r + 1)$ -block's map to symbol 1, and $1 - \lambda$ percent of them map to symbol 0. Randomly choose two $(2r + 1)$ -block's, the probability that they map to the same symbol is

$$P_{same} = \lambda^2 + (1 - \lambda)^2, \quad (\text{B.2})$$

and the probability that they map to different symbols is

$$1 - P_{same} = 2\lambda(1 - \lambda). \quad (\text{B.3})$$

The probability that the left spreading rate is r is equal to the probability that block $(a_{-r}a_{-(r-1)} \cdots a_{r-1}1)$ and $(a_{-r}a_{-(r-1)} \cdots a_{r-1}0)$ map to different symbols, which is same with the probability that two randomly picked $(2r+1)$ -block's will map to different symbols, or $1 - P_{same}$. The contribution from the maximum spreading to the average left spreading rate is

$$(1 - P_{same})r. \quad (\text{B.4})$$

Similarly, the probability that the left spreading rate is $r-1$ is equal to the probability that block $(a_{-r}a_{-(r-1)} \cdots a_{r-1}1)$ and $(a_{-r}a_{-(r-1)} \cdots a_{r-1}0)$ map to the same symbol, which is P_{same} , multiplied by the probability that blocks $(a_{-r}a_{-(r-1)} \cdots 1a_r)$ and $(a_{-r}a_{-(r-1)} \cdots 0a_r)$ map to different symbols, $1 - P_{same}$. After counting all possibilities, including the contribution of the large negative spreading rates, the average left spreading rate of the difference pattern is

$$\gamma_{left} = \sum_{i=0}^{\infty} (r-i) P_{same}^i (1 - P_{same}) = r - \frac{P_{same}}{1 - P_{same}}. \quad (\text{B.5})$$

The average spreading rate taking into account of both left and right expansions is (inserting the P_{same} expression in terms of λ):

$$\gamma = \frac{(2r+2)\lambda - (2r+2)\lambda^2 - 1}{\lambda(1-\lambda)}. \quad (\text{B.6})$$

There are two interesting results which can be derived from the above formula. First, at $\lambda = 0.5$,

$$\gamma_{random} = 2(r-1) < 2r = \gamma_{max}, \quad (\text{B.7})$$

or, even in the most random cases, the maximum spreading rate will not be reached. Second, the critical λ_c can be defined as the onset of the non-zero value of γ . Setting $\gamma(\lambda_c) = 0$, we have the λ_c as a function of r :

$$\lambda_c = \frac{1}{2} - \frac{1}{2} \sqrt{1 - \frac{2}{r+1}}. \quad (\text{B.8})$$

References

- [1] H. Chaté and P. Manneville, Criticality in cellular automata, *Physica D*, 45 (1990) 122–135, these Proceedings.
- [2] J.P. Crutchfield and N.H. Packard, Bifurcations in discretized spatially-extended systems, Abstract, Cellular Automata’86 workshop (1986).
- [3] H. Gutowitz, A Hierarchical Classification of Cellular Automata, *Physica D*, 45 (1990) 136–156, these Proceedings. (this volume).
- [4] K. Kaneko, Pattern dynamics in spatio-temporal chaos: pattern selection, diffusion of defect and pattern competition intermittency, *Physica D* 34 (1989) 1–41.
- [5] C. Langton, Studying artificial life with cellular automata, *Physica D* 22 (1986) 120–140.
- [6] C. Langton, Computation at the edge of chaos: phase transitions and emergent computation, *Physica D* 42 (1990) 12–37.
- [7] C. Langton, Computation at the Edge of Chaos, Ph.D Thesis, University of Michigan (1990).
- [8] W. Li, Problems in Complex Systems, Ph.D Thesis, Columbia University (1989); University Microfilm International, Ann Arbor, MI.
- [9] W. Li, Mutual information functions versus correlation functions, *J. Stat. Phys.* 60 (1990) 823–837.
- [10] W. Li and N. Packard, Structure of the elementary cellular automata rule space, *Complex Systems* 4(3) (1990), in press.
- [11] N. Margolus and T. Toffoli, *Cellular Automata Machines* (MIT Press, Cambridge, MA 1987).
- [12] H. McIntosh, Wolfram’s class IV automata and a good Life, *Physica D* 45 (1990) 105–121, these Proceedings.
- [13] J. Milnor, Directional entropies of cellular automaton-maps, in: *Disordered Systems and Biological Organizations*, eds. E. Bienenstock et al. (Springer, Berlin, 1986).
- [14] N. Packard, Adaptation toward the edge of chaos, Center for Complex Systems Research Technical Report, University of Illinois, CCSR-88-5 (1988).
- [15] C.E. Shannon, The mathematical theory of communication, *Bell Syst. Tech. J.* 27 (1948) 379–423.
- [16] R. Shaw, Information density near a phase transition, Abstract, Cellular Automata’86 workshop (1986).
- [17] G. Vichniac and P. Tamayo and H. Hartman, Annealed and Quenched Inhomogeneous Cellular Automata (INCA), *J. Stat. Phys.* 45 (1986), 875–883.

- [18] J. von Neumann, *Theory of Self-reproducing Automata*, edited by A. Burks (Univ. of Illinois Press, Champaign, IL, 1966).
- [19] S. Wolfram, *Statistical Mechanics of Cellular Automata*, *Rev. Mod. Phys.* 55 (1983) 601–644.
- [20] S. Wolfram, *Universality and complexity in cellular automata*, *Physica D* 10 (1984) 1–35.
- [21] W.K. Wootters and C. Langton, *Is there a sharp transition for deterministic cellular automata?* *Physica D* 45 (1990) 95–104, these Proceedings.

Figure 1: Percentage of elementary rules which are null, time-invariant, periodic, and chaotic as a function of λ . There are two curves for each class of behavior for the original rule space and the folded rule space respectively (see [10] for the definition of folded rule space).

Figure 2: Difference pattern spreading rate averaged over 200 $r=7$ rules, with each path into chaos aligned to the point on the path where $\gamma > 2$.

Figure 3: Single site entropy S for 2-dimensional, 8-state, 5-neighbor cellular automata. The single site probabilities $\{p_i\}$ ($i = 1, 2, \dots, 8$) are accumulated from patterns at 500 consecutive times. The first 500 transient patterns are discarded. The square lattice has size 64×64 .
 (a) S as a function of λ for 20 different paths in the rule space.
 (b) S as a function of $\lambda - \lambda_c$.

Figure 4: Spatial mutual information M as a function of λ for 1-dimensional, 2-state, 11-neighbor cellular automata rules for 120 paths in the rule space. The lattice size is 601, 257 transient configurations are discarded, and the next 51 configurations are used to accumulate the probabilities. The mutual information is between two 3-block's separated by spatial distance 10. Each data point is M for a fixed rule averaging over 2 different initial conditions.

(a) M versus λ .
 (b) M versus $\Delta\lambda = \lambda - \lambda_c$ where λ_c is defined by the first λ value at which a threshold value of M is exceeded.
 (c) M versus $\Delta\lambda = \lambda - \lambda_c$ where λ_c is defined by the λ value with the maximum M .

Figure 5: Same data points as those in fig.4(c) are drawn in semi-log scales. The data can roughly be fit by $M(|\Delta\lambda|) \sim e^{-\alpha|\Delta\lambda|}$, with $\alpha \approx 12$.

Figure 6: Spatial mutual information M between two 3-block's separated by distance 4 instead of 10. (a) M as a function of λ . (b) M as a function of $\Delta\lambda = \lambda - \lambda_c$, where λ_c is defined by the maximum value of mutual information, in semi-log scale. The data can roughly be fit by $M(|\Delta\lambda|) \sim e^{-\alpha|\Delta\lambda|}$, with $\alpha \approx 8$.

Figure 7: Spatial mutual information M between two 3-block's separated by distance 10 for $r = 7$, instead of $r = 5$, rule space. (a) M as a function of λ . (b) M as a function of $\Delta\lambda = \lambda - \lambda_c$, where λ_c is defined by the maximum value of mutual information, in semi-log scale.

Figure 8: The distribution of spreading rates for elementary rules 110 (on the left side of the plot) and 30 (on the right side).

Figure 9: Spatial mutual information for randomly sampled rules with $r = 3, 5, 7, 9$, and 11 respectively. For each r value and fixed λ value, 51 rules are sampled. The lattice size is 167, 118 transient configurations are discarded, and the next 89 configurations are used to accumulate the probability distribution. The mutual information is between two 3-block's separated by distance 8.

Figure 10: Critical region λ_c as determined by the maximum value of the spatial mutual information plotted as the function of the neighborhood size $(2r+1)$ for 1-dimensional, 2-state, cellular automata rule spaces. Also plotted are two functions for the onset of non-zero entropy $(1/(2r+1))$ and non-zero expansion rate for difference pattern $(1/2 - 1/2\sqrt{1 - 2/r+1})$ according to mean field theory.

Figure 11: A schematic picture of the structure of cellular automata rule-space, indicating the relative location of the various qualitative classes. The λ parameter is plotted from left to right, and there is no special meaning associated with the vertical axis. Complex rules are associated with the small, shaded patches along the boundary between periodic and chaotic rules.

Figure 12: The smoothed mutual information surface plotted over the ρ parameter space for 2-D, 3-state, 5-neighbor CAs. The chaotic rules constitute the vast central depression, fixed-point, and periodic rules lie between the vertices of the base-triangle and the peaks in the mutual information surface, and the complex rules are found in the vicinity of these peaks.

Figure 13: A schematic diagram of a 2-D phase-space for CAs and its projection onto the λ parameter. The dark-shaded region in the middle is supposed to constitute the “critical” region. Without knowing the value of the “mystery” parameter on the ordinate, one would see both first- and second-order transitions over a range of λ values, as we observe experimentally. If we find other appropriate parameters governing CA dynamics, we should be able to locate the boundaries between the various dynamical regimes precisely.