Physica D 42 (1990) 12–37
North-Holland

COMPUTATION AT THE EDGE OF CHAOS: PHASE TRANSITIONS AND EMERGENT COMPUTATION

Chris G. LANGTON
Complex Systems Group, Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87454, USA

In order for computation to emerge spontaneously and become an important factor in the dynamics of a system, the material substrate must support the primitive functions required for computation: the transmission, storage, and modification of information. Under what conditions might we expect physical systems to support such computational primitives?

This paper presents research on cellular automata which suggests that the optimal conditions for the support of information transmission, storage, and modification, are achieved in the vicinity of a phase transition. We observe surprising similarities between the behaviors of computations and systems near phase transitions, finding analogs of computational complexity classes and the halting problem within the phenomenology of phase transitions.

We conclude that there is a fundamental connection between computation and phase transitions, especially second-order or "critical" transitions, and discuss some of the implications for our understanding of nature if such a connection is borne out.

1. Introduction

Most of the papers in these Proceedings assume the existence of a physical system with the capacity to support computation, and inquire after the manner in which processes making use of this capacity might emerge spontaneously.

In this paper, we will focus on the conditions under which this *capacity to support computation* itself might emerge in physical systems, rather than on how this capacity might ultimately come to be utilized.

Therefore, the fundamental question addressed in this paper is the following:

Under what conditions will physical systems support the basic operations of information transmission, storage, and modification constituting the capacity to support computation?

This question is difficult to address directly. Instead, we will reformulate the question in the context of a class of formal abstractions of physical systems: cellular automata (CAs). Our question, thus, becomes:

Under what conditions will cellular automata support the basic operations of information transmission, storage, and modification?

This turns out to be a tractable problem, with a somewhat surprising answer; one which leads directly to a hypothesis about the conditions under which computations might emerge spontaneously in nature.

1.1. Overview

First, we introduce cellular automata and a simple scheme for parameterizing the space of all possible CA rules. We then apply this parameterization scheme to the space of possible one-dimensional CAs in a *qualitative* survey of the different dynamical regimes existing in CA rule space and their relationship to one another. Next, we present a *quantitative* picture of these structural relationships, using data from an extensive survey of two-dimensional CAs. Finally, we review the observed relationships among dynamical regimes, and discuss their implications for the more general question raised in the introduction.

1.2. Results

We find that by selecting an appropriate parameterization of the space of CAs, one observes a *phase transition* between highly *ordered* and highly *disordered* dynamics, analogous to the phase transition between the *solid* and *fluid* states of matter. Furthermore, we observe that CAs exhibiting the most complex behavior – both qualitatively and quantitatively – are found *generically* in the vicinity of this phase transition. Most importantly, we observe that CAs in the transition region have the greatest potential for the support of information storage, transmission, and modification, and therefore for the emergence of computation.

These observations suggest that there is a fundamental connection between phase transitions and computation, leading to the following hypothesis concerning the emergence of computation in physical systems:

Computation may emerge spontaneously and come to dominate the dynamics of physical systems when those systems are at or near a transition between their *solid* and *fluid* phases, especially in the vicinity of a second-order or "critical" transition.

This hypothesis, if borne out, has many implications for understanding the role of information in nature.

Perhaps the most exciting implication is the possibility that life had its origin in the vicinity of a phase transition, and that evolution reflects the process by which life has gained local control over a successively greater number of environmental parameters affecting its ability to maintain itself at a critical balance point between order and chaos.

1.3. Cellular automata

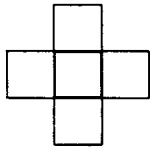
In this section, we review cellular automata, introduce a parameterization of the space of possible CA rules, and discuss computation in CAs.

Cellular automata are discrete space/time logical universes, obeying their own local physics [26, 3, 5, 27, 28].

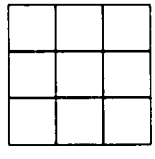
Space in CAs is partitioned into discrete volume elements called "cells" and time progresses in discrete steps. Each cell of space is in one of a finite number of states at any one time. The physics of this logical universe is a deterministic, local physics. "Local" means that the state of a cell at time $t+1$ is a function only of its own state and the states of its immediate neighbors at time t. "Deterministic" means that once a local physics and an initial state of a CA has been chosen, its future evolution is uniquely determined.

1.4. Formal definition of cellular automata

Formally, a cellular automaton is a D-dimensional lattice with a finite automaton residing at each lattice site. Each automaton takes as input the states of the automata within some *finite* local region of the lattice, defined by a neighborhood template $\mathcal{N}$, where the dimension of $\mathcal{N} \leq D$. The size of the neighborhood template, $|\mathcal{N}|$, is just the number of lattice points covered by $\mathcal{N}$. By convention, an automaton is considered to be a member of its own neighborhood. Two typical two-dimensional neighborhood templates are:



five cell neighborhood



nine cell neighborhood

Each finite automaton consists of a finite set of *cell states* Σ, a finite *input alphabet* α, and a *transition function* Δ, which is a mapping from the set of neighborhood states to the set of cell states. Letting $N = |\mathcal{N}|$:

$$\Delta\colon \Sigma^N \to \Sigma.$$

The *state* of a neighborhood is the cross product of the states of the automata covered by the neighborhood template. Thus, the input alphabet

α for each automaton consists of the set of possible neighborhood states: $\alpha = \Sigma^N$. Letting $K = |\Sigma|$ (the number of cell states) the size of α is equal to the number of possible neighborhood states

$$|\alpha| = |\Delta| = |\Sigma^N| = K^N.$$

To define a transition function Δ, one must associate a unique next state in Σ with each possible neighborhood state. Since there are $K = |\Sigma|$ choices of state to assign as the next state for each of the $|\Sigma^N|$ possible neighborhood states, there are $K^{(K^N)}$ possible transition functions Δ that can be defined. We use the notation $\mathcal{D}_N^K$ to refer to the set of all possible transition functions Δ which can be defined using N neighbors and K states.

1.5. Example

Consider a two-dimensional cellular automaton using 8 states per cell, a rectangular lattice, and the five-cell neighborhood template shown above. Here $K = 8$ and $N = 5$, so $|\Delta| = K^N = 8^5 = 32\,768$ and there are thus 32 768 possible neighborhood states. For each of these, there is a choice of 8 states as the next cell state under Δ, so there are $K^{(K^N)} = |\mathcal{D}_N^K| = 8^{(8^5)} \approx 10^{30\,000}$ possible transition functions using the 5-cell neighborhood template with 8 states per cell, an exceedingly large number.

2. Parameterizing the space of CA rules

$\mathcal{D}_N^K$, the set of possible transition functions Δ for a CA of K states and N neighbors, is fixed once we have chosen the number of states per cell and the neighborhood template. However, there is no intrinsic order within $\mathcal{D}_N^K$; it is a large, undifferentiated space of CA rules.

Imposing a *structure* on this undifferentiated space of CA rules allows us to define a natural ordering on the rules, and provides us with an index into the rule space. The ideal ordering scheme would partition the space of CA rules in such a manner that rules from the same partition would support similar dynamics. Such an ordering on $\mathcal{D}_N^K$ would allow us to observe the way in which the dynamical behaviors of CAs vary from partition to partition.

The location in this space of the partitions supporting the transmission, modification, and storage of information, relative to the location of partitions supporting *other* possible dynamical behaviors should provide us with insight into the conditions under which we should expect computation to emerge in CAs.

2.1. The λ parameter

We will consider only a subspace of $\mathcal{D}_N^K$, characterized by the parameter λ [18, 17].

The λ parameter is defined as follows. We pick an arbitrary state $s \in \Sigma$, and call it the *quiescent* state s_q. Let there be n transitions to this special quiescent state in a transition function Δ. Let the remaining $K^N - n$ transitions in Δ be filled by picking randomly and uniformly over the other $K - 1$ states in $\Sigma - s_q$. Then

$$\lambda = \frac{K^N - n}{K^N}. \tag{1}$$

If $n = K^N$, then *all* of the transitions in the rule table will be to the quiescent state s_q and $\lambda = 0.0$. If $n = 0$, then there will be *no* transitions to s_q and $\lambda = 1.0$. When all states are represented equally in the rule table, then $\lambda = 1.0 - 1/K$.

The parameter values $\lambda = 0.0$ and $\lambda = 1.0 - 1/K$ represent the most homogeneous and the most heterogeneous rule tables, respectively. The behavior in which we will be interested is captured between these two parameter values. Therefore, we experiment primarily with λ in this range.

2.2. Searching CA space with the λ parameter

In the following, we use the λ parameter as a means of sampling $\mathcal{D}_N^K$ in an ordered manner. We do this by stepping through the range $0.0 \leq \lambda \leq$

$1.0 - 1/K$ in discrete steps, randomly constructing Δ functions for each λ point. Then we run CAs under these randomly constructed Δ functions, collecting data on various measures of their dynamical behavior. Finally, we examine the behavior of these measures as a function of λ.

Δ functions are constructed in two ways using λ. In the "random-table method", λ is interpreted as a bias on the random selection of states from Σ as we sequentially fill in the transitions that make up a Δ function. To do this, we step through the table, flipping a λ-biased coin for each neighborhood state. If the coin comes up tails, with probability $1.0 - \lambda$, we assign the state s_q as the next cell state for that neighborhood state. If the coin comes up heads, with probability λ, we pick one of the $K - 1$ states in $\Sigma - s_q$ at uniform random as the next cell state.

In the "table-walk-through" method, we start with a Δ function consisting entirely of transitions to s_q, so that $\lambda = 0.0$ (but note restrictions below). New transition tables with higher λ values are generated by randomly replacing a few of the transitions to s_q in the current function with transitions to other states, selected randomly from $\Sigma - s_q$. Tables with *lower* λ values are generated by randomly replacing a few transitions that are *not* to s_q in the current table by transitions to s_q.

Thus, under the table-walk-through method, we progressively perturb "the same table", whereas under the random-table method, each new table is generated from scratch.

2.3. Further restrictions on CAs

In order to make our studies more tractable, we impose two further conditions on the rule space. First, a strong *quiescence condition*: all neighborhood states uniform in cell state s_i will map to state s_i. Second, an *isotropy condition*: all planar rotations of a neighborhood state will map to the same cell state. These restrictions mean that arrays uniform in any single state will remain so, and that the physics cannot tell which way is up, so to speak.

2.4. Discussion

λ is not necessarily the best parameter. One can improve on λ in a number of ways. For instance, Gutowitz [12,11] has defined a hierarchy of parameterization schemes in which λ is the simplest scheme, mean field theory constitutes the next simplest scheme, and so on.

However, λ suffices to reveal a great deal about the overall structural relationships between the various dynamical regimes in CA rule space, and it is very useful to get a feel for the "lay of the CA landscape" at this low-resolution level before increasing the resolution and surveying finer details. For one thing, λ helps restrict the area of search to a particularly promising "spot", which is useful because higher-order parameterizations map CA rule space onto *many* dimensions, whereas λ is a one-dimensional parameter.

λ discriminates well between dynamical regimes for "large" values of K and N, whereas λ discriminates poorly for small values of K and N. For example, for a 1D CA with $K = 2$, and $N = 3$, λ is only roughly correlated with dynamical behavior. This may explain why the relationships reported here were not observed in earlier work on classifying CA dynamics [29,28], as these investigations were carried out using CAs with minimal values of K and N.

For these reasons, we employ CAs for which $K \geq 4$ and $N \geq 5$, which results in transition tables of size $4^5 = 1024$ or larger.

2.5. Computation in CAs

Cellular automata can be viewed either as computers themselves or as *logical universes* within which computers may be embedded.

On the first view, an initial configuration constitutes the data that the physical computer is working on, and the transition function implements the algorithm that is to be applied to the data. This is the approach taken in most current applications of cellular automata, such as image processing.

On the second view, the initial configuration itself constitutes a computer, and the transition function is seen as the "physics" obeyed by the parts of this embedded computer. The algorithm being run and the data being manipulated are functions of the precise state of the initial configuration of the embedded computer. In the most general case, the initial configuration will constitute a universal computer.

We can always take the first point of view, but what we are interested in here is the question: when is it possible – even necessary – to adopt the second point of view to understand the dynamics of a CA?

That CAs are capable of supporting universal computation has been known since their invention by Ulam and von Neumann in the late 40's. Von Neumann's proof of the possibility of machine self-reproduction involves the demonstration of the existence of a universal computer/constructor in a 29-state CA [26]. Since then, Codd [5], Smith [24], Conway and co-workers [2], Fredkin and Toffoli [7] – to name but a few – have found much simpler CA rules supporting universal computation.

All of these proofs involve the embedding of a computer within the CA, or at least they show that all of the important parts of such a computer could be implemented and that those parts are sufficient to construct a computer. Some of these proofs involve the construction of Turing machines, others involve the construction of stored-program computers.

All of these constructs rely on three fundamental features of the dynamics supported by the underlying transition function physics. First, the physics must support the *storage* of information, which means that the dynamics must preserve local state information for arbitrarily long times. Second, the physics must support the *transmission* of information, which means that the dynamics must provide for the propagation of information in the form of *signals* over arbitrarily long distances. Third, stored and transmitted information must be able to interact with one another, resulting in a possible modification of one or the other.

These fundamental properties must be provided by *any* dynamical system if it is to support computation. Taken together, they require that any dynamical system supporting computation *must exhibit arbitrarily large correlation lengths in space and time*. These correlation lengths must be *potentially* infinite, but not *necessarily* so. Codd [5] refers to this situation as one in which the propagation of information must be *unbounded* in principle but *boundable* in practice.

2.6. *Wolfram's qualitative CA classes*

Wolfram [29] has proposed the following four qualitative classes of CA behavior:

Class I evolves to a homogeneous state.
Class II evolves to simple separated periodic structures.
Class III yields chaotic aperiodic patterns.
Class IV yields complex patterns of localized structures.

Wolfram finds the following analogs for his classes of cellular automaton behaviors in the field of dynamical systems.

Class I cellular automata evolve to *limit points*.
Class II cellular automata evolve to *limit cycles*.
Class III cellular automata evolve to *chaotic* behavior of the kind associated with *strange attractors*.
Class IV cellular automata "effectively have very long *transients*".

This association of class IV CAs with "very long transients" will figure "critically" in what follows.

Wolfram suggests that class IV CAs are capable of supporting computation, even universal computation, and that it is this capacity that makes their behavior so complex. This paper supports Wolfram's hypothesis, and offers an explanation for both the existence of these classes and their relationship to one another.

In their surveys of 1D and 2D CAs, Packard and Wolfram [23] hypothesized that class IV CAs constitute a set of measure 0. This means that class IV behaviors should be infinitely hard to find in the "thermodynamic limit" of an infinitely large CA rule space. However, it turns out that they are *not* hard to find in rule spaces that are far from the thermodynamic limit. By locating class IV behaviors in these non-limiting rule spaces and tracking the manner in which they become vanishingly rare as one goes to larger rule spaces, we can derive a general theory about where to locate rules likely to support computation in *any* CA rule space.

3. Qualitative overview of CA dynamics

In this section, we present a series of examples illustrating the changes observed in the dynamical behavior of one-dimensional CAs as we alter the λ parameter throughout its range using the table-walk-through method. For these CAs, $K = 4$, $N = 5$ (i.e. *two* cells on the left and *two* cells on the right are included in the neighborhood template). The arrays consist of 128 sites connected in a circle, resulting in periodic boundary conditions. Each array is started from a random initial configuration on the top line, and successive lines show successive time steps in the evolution.

For each value of λ, we show two evolutions. The arrays in fig. 1 are started from a uniform random initial configuration over all 128 sites, while those in fig. 2 are started from configurations whose sites are all 0, with the exception of a patch of 20 randomized sites in the middle.

Fig. 1 illustrates the kinds of structures that develop, as well as the typical transient times before these structures are achieved. Fig. 2 illustrates the relative spread or collapse of the area of dynamical activity with time. For those values of λ exhibiting long transients, we have reduced the scale of the arrays in order to display longer evolutions.

We start with $\lambda \approx 0.00$. Note that under the strong quiescence condition mentioned above we cannot have $\lambda = 0.00$ exactly. The primary features observed as we vary λ throughout its range are itemized below.

$\lambda \approx 0.00$ All dynamical activity dies out after a single time step, leaving the arrays uniform in state s_q. The area of dynamical activity has collapsed to zero.

$\lambda \approx 0.05$ The dynamics reaches the uniform s_q fixed point after approximately 2 time steps.

$\lambda \approx 0.10$ The homogeneous fixed point is reached after 3 or 4 time steps.

$\lambda \approx 0.15$ The homogeneous fixed point is reached after 4 or 5 time steps.

$\lambda \approx 0.20$ The dynamics reaches a periodic structure which will persist forever (fig. 1, $\lambda = 0.20$). Transients have increased to 7 to 10 time steps as well. Note that the evolution does not necessarily lead to periodic dynamics (fig. 2, $\lambda = 0.20$).

$\lambda \approx 0.25$ Structures of period 1 appear. Thus, there are now three different possible outcomes for the ultimate dynamics of the system, depending on the initial state. The dynamics may reach a *homogeneous* fixed point consisting entirely of state s_q, or it may reach a *heterogeneous* fixed point consisting mostly of cells in state s_q with a sprinkling of cells stuck in one of the other states, or it may settle down to periodic behavior. Notice that the transients have lengthened even more.

$\lambda \approx 0.30$ Transients have lengthened again.

$\lambda \approx 0.35$ Transient length has grown significantly, and a new kind of periodic structure with a longer period has appeared (fig. 1, $\lambda = 0.35$). Most of the previous structures are still possible, hence the spectrum of dynamical possibilities is broadening.

$\lambda \approx 0.40$ Transient length has increased to about 60 time steps, and a structure has appeared with a period of about 40 time steps. The area of dynamical activity is

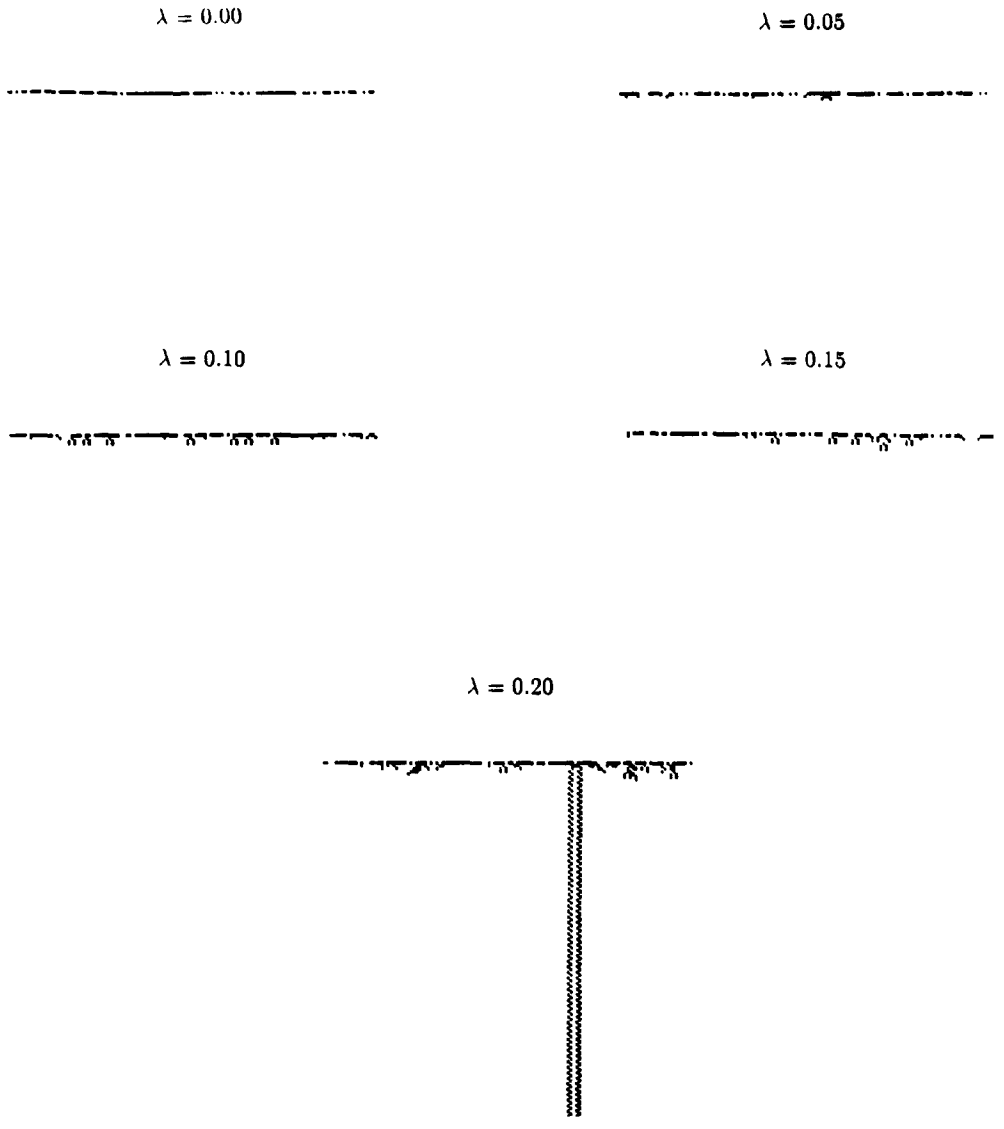


Fig. 1. Evolutions of one-dimensional, $K = 4$, $N = 5$ CAs from *fully* random initial configurations over $0.0 < \lambda \leq 0.75$. As λ is increased the structures become more complicated, and the transients grow in length until they become arbitrarily long at $\lambda \approx 0.50$. For $0.50 < \lambda \leq 0.75$, the transient lengths *decrease* with increasing λ, as indicated by the arrows to the right of the evolutions.

still collapsing down onto isolated periodic configurations.

$\lambda \approx 0.45$ Transient length has increased to almost 1000 time steps (fig. 1, $\lambda = 0.45$). Here, the structure on the right appears to be periodic, with a period of about 100 time steps. However, after viewing several cycles of its period, it is apparent that the whole structure is moving to the left, and so this pattern will not recur precisely in its same position until it has cycled at least once around the array. Furthermore, as it propagates to the left, this structure eventually annihi-

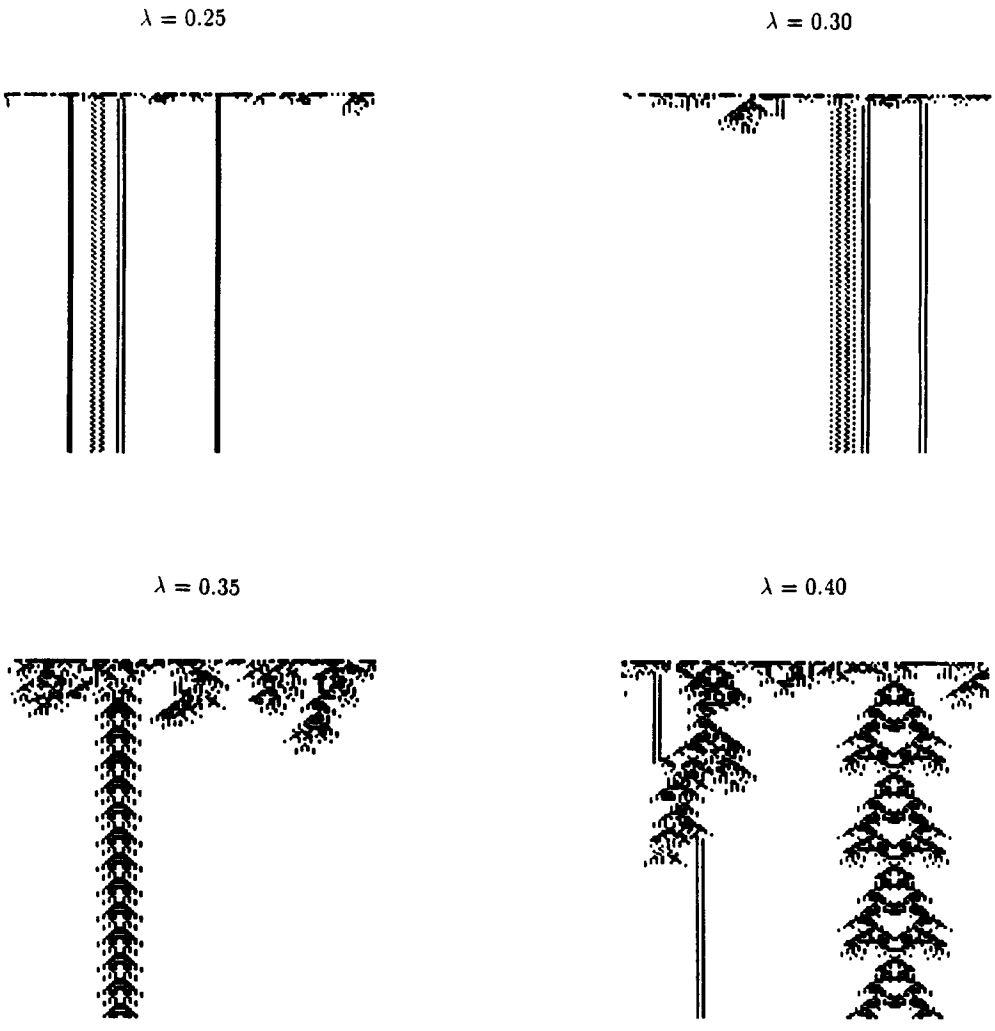


Fig. 1. Continued

lates a period-1 structure after about 800 time steps. Thus, the transient length before a periodic structure is reached has grown enormously. It turns out that even after one orbit around the array, the periodic structure does not return exactly to its previous position. It must orbit the array 3 times before it repeats itself exactly. As it has shifted over only 3 sites after its quasi-period of 116 time steps, the true period of this structure is 14 848 time steps. Here, the area of dynamical activity is at a balance point between collapse and expansion.

$\lambda \approx 0.50$ Typical transient length is on the order of 12 000 time steps. After the transient, the dynamical activity settles down to periodic behavior, possibly of period one as shown in the figure. Although, the dynamics eventually becomes simple, the transient time has increased dramatically. Note in fig. 2 that the general tendency now is that the area of

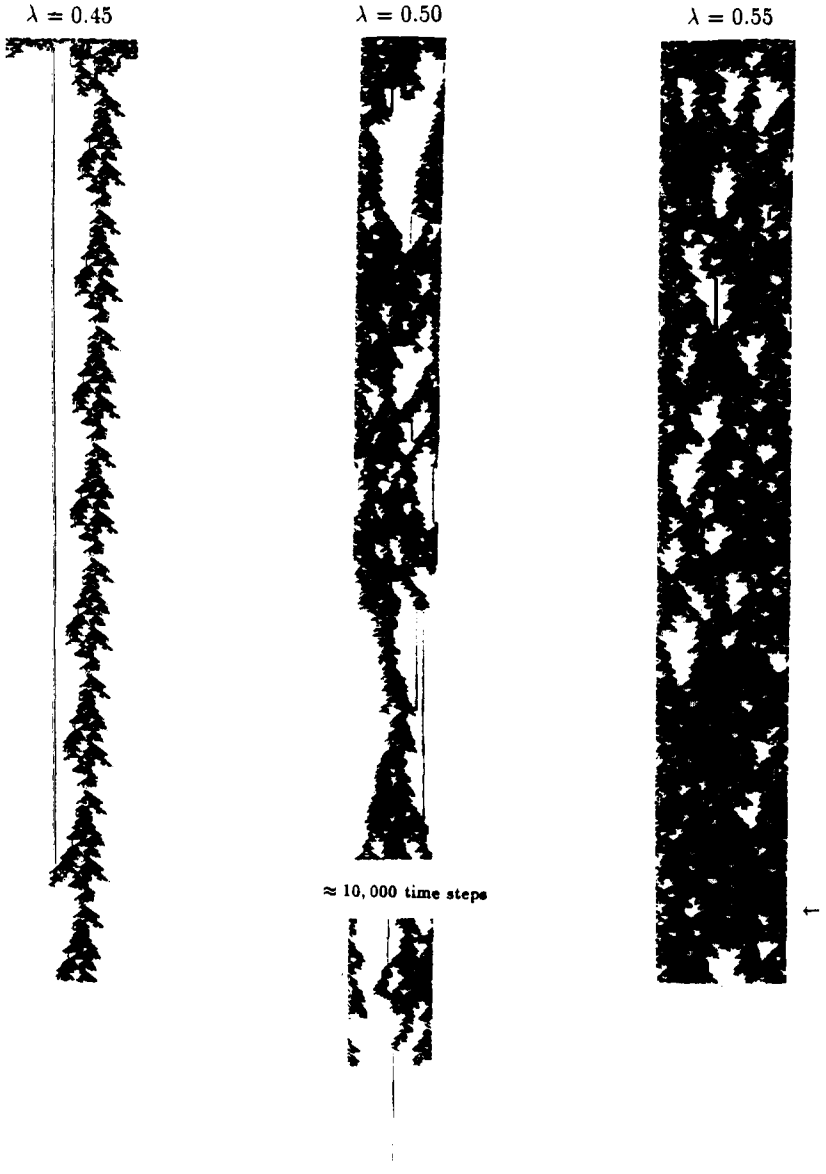


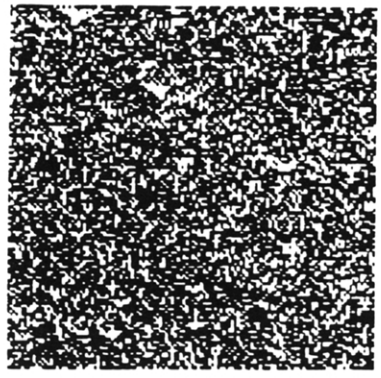
Fig. 1. Continued

dynamical activity *expands* rather than contracts with time. There are, however, large fluctuations in the area covered by dynamical activity, and it is these fluctuations which lead to the eventual collapse of the dynamics.

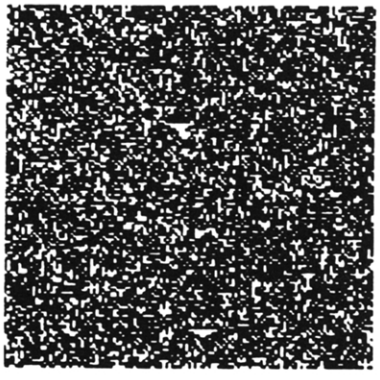
$\lambda \approx 0.55$ We have entered a new dynamical regime in which the transients have become so long that – for all practical purposes – they *are* the steady state behavior of the system over any period of time for which we can observe them. Whereas before, the dynamics *eventually* settled down to periodic behavior, we are now in a regime in which the dynamics typically settles down to ef-

$\lambda = 0.60$

$\lambda = 0.65$



$\lambda = 0.70$



$\lambda = 0.75$

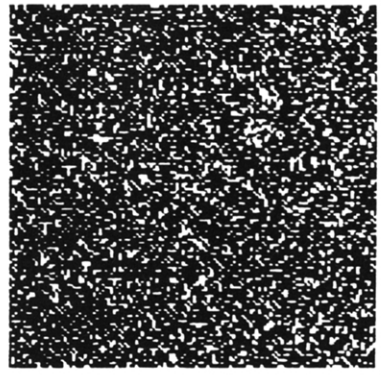


Fig. 1. Continued

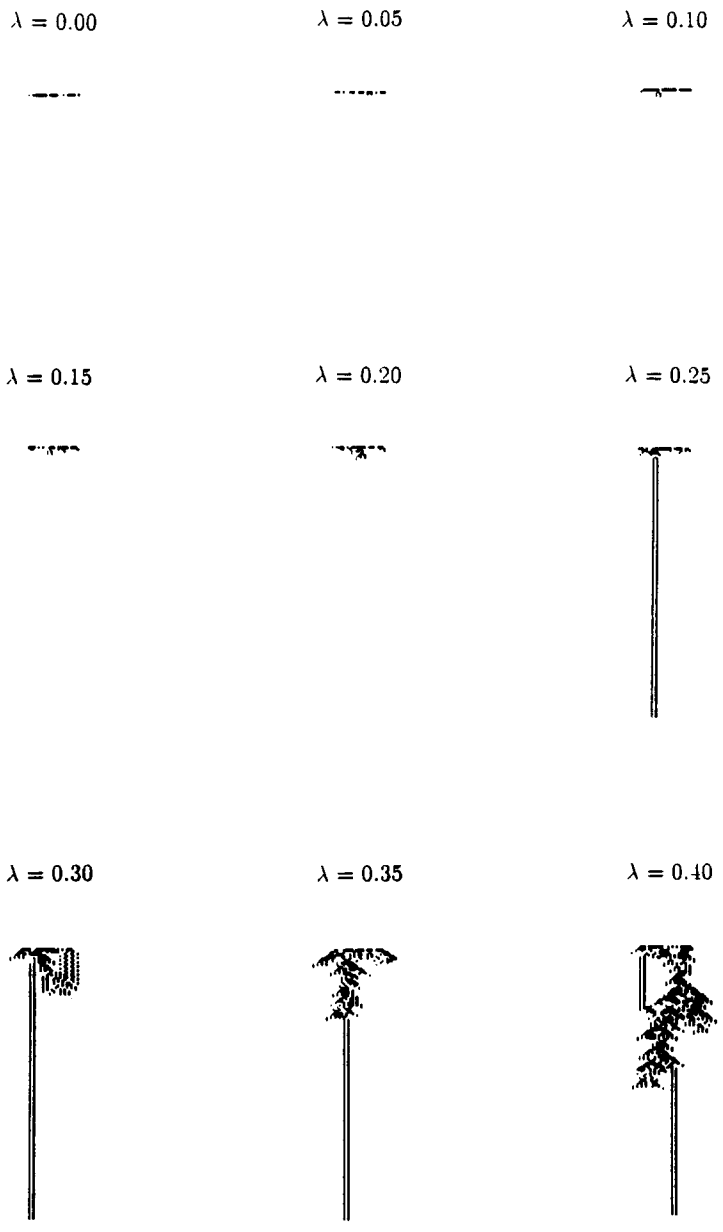


Fig. 2. Evolutions of one-dimensional, $K = 4$, $N = 5$ CAs from *partially* random initial configurations over $0.0 < \lambda \leq 0.75$. This series illustrates the change in the rate of spread of the dynamics from negative for $\lambda < 0.45$, to positive for $\lambda > 0.45$. For $\lambda \approx 0.45$, the dynamics is balanced between collapse and expansion, giving rise to particle-like solitary waves.

fectively *chaotic* behavior. Furthermore, the previous trend of transient length *increasing* with increasing λ is reversed. The arrow to the right of the evolutions of figs. 1, $\lambda = 0.55$–0.75 indicates the approximate time by which the site-occupation density has settled down to within 1% of its long-time average. Note that the area of dynamical activity expands more rapidly with time.

$\lambda \approx 0.60$ The dynamics are quite chaotic, and the transient length to "typical" chaotic be-

Fig. 2. Continued

havior has decreased significantly. The area of dynamical activity expands more rapidly with time.

$\lambda \approx 0.65$ Typical chaotic behavior is achieved in only 10 time steps or so. The area of dynamical activity is expanding at about one cell per time step in each direction, approximately half of the maximum possible rate for this neighborhood template.

$\lambda \approx 0.70$ Fully developed chaotic behavior is reached in only 2 time steps. The area

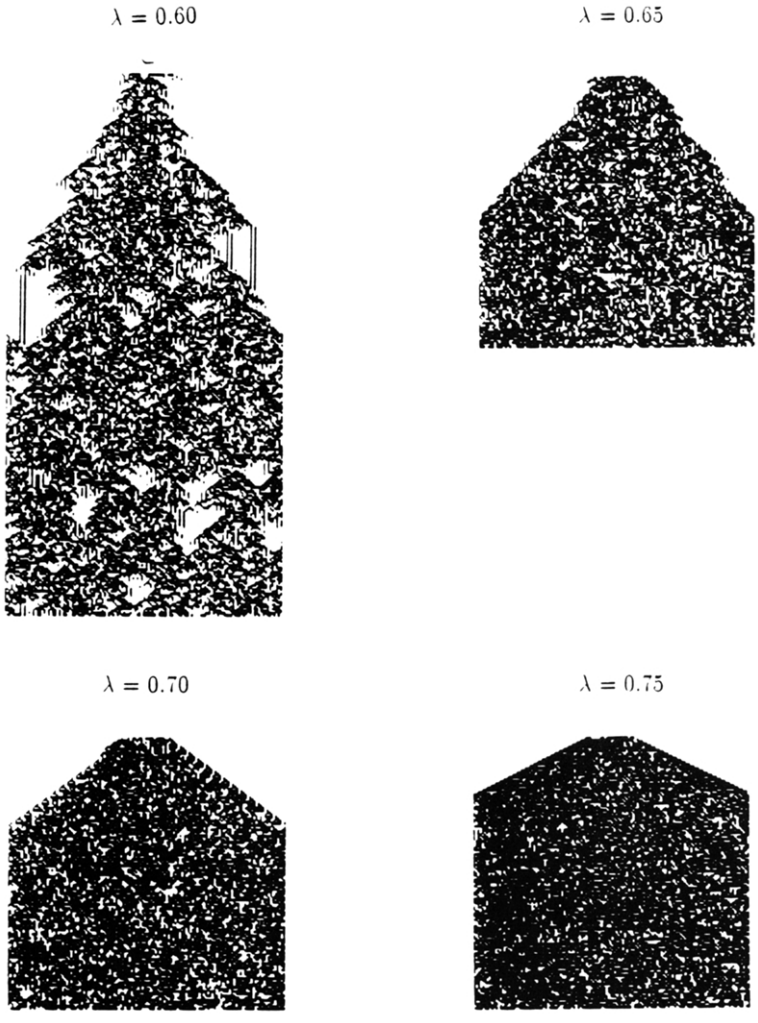


Fig. 2. Continued

of dynamical activity is expanding even more rapidly.

$\lambda \approx 0.75$ After only a single time step, the array is essentially random and remains so thereafter. The area of dynamical activity spreads at the maximum possible rate.

Therefore, by varying the λ parameter throughout $0.0 < \lambda \leq 0.75$ over the space of possible $K = 4$, $N = 5$, 1D cellular automata, we progress from CAs exhibiting the maximal possible order to CAs exhibiting the maximal possible disorder. At intermediate values of λ, we encounter a *phase transition* between periodic and chaotic dynamics, and while the behavior at either end of the λ spectrum seems "simple" and easily predictable, the behavior in the vicinity of this phase transition seems "complex" and unpredictable.

4. Comments on qualitative dynamics

There are several observations to be made about the ID examples of section 3.

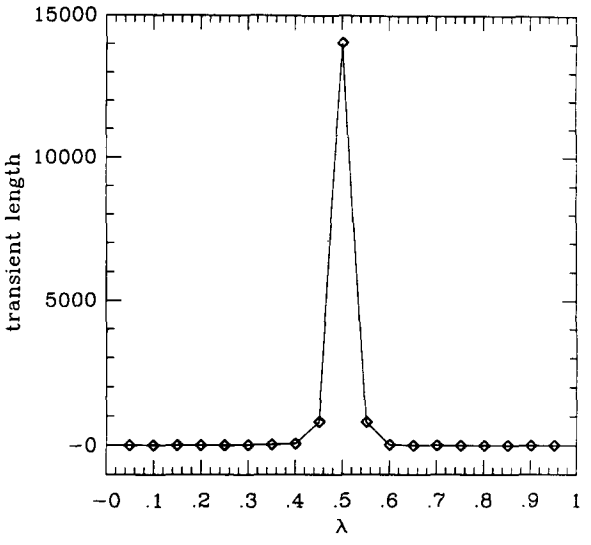


Fig. 3. Average transient length as a function of λ in an array of 128 cells.

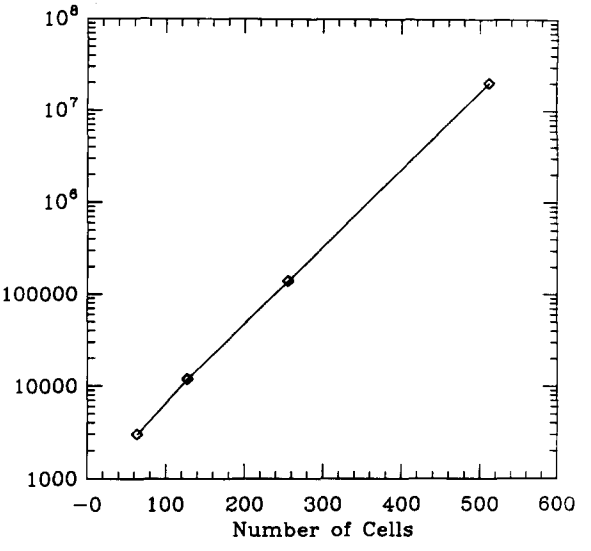


Fig. 4. Growth of average transients as a function of array size for λ = 0.50.

First, transients grow rapidly in the vicinity of the transition between ordered and disordered dynamics, a phenomenon known in the study of phase transitions as *critical slowing down*. The relationship between transient length and λ is plotted in fig. 3.

Second, the size of the array has an effect on the dynamics only for intermediate values of λ. For low values of λ, array size has *no* discernible effect on transient length. Not until $\lambda \approx 0.45$ do we begin to see a small difference in the transient length as the size of the array is increased. For $\lambda = 0.50$, however, array size has a significant effect on the transient length. The growth of transient length as a function of array size for $\lambda = 0.50$ is plotted in fig. 4. The essentially linear relationship on this log-normal plot suggests that transient length depends *exponentially* on array size at $\lambda = 0.50$. As we continue to raise λ beyond 0.50, although the dynamics is now settling down to effectively chaotic behavior instead of periodic behavior, the transient lengths are getting *shorter* with increasing λ, rather than longer. A number of statistical measures (see ref. [17]) reveal that the time it takes to reach "typical" behavior decreases as λ increases past the transition point. Furthermore, transient times exhibit decreasing dependence on array size as λ is increased past the transition point. By the time all states are represented uniformly in the transition table – at $\lambda = 0.75$ in this case – the transient lengths exhibit *no* dependence on array size – just as was the case for low values of λ.

Third, the overall evolutionary pattern in time appears more random as $\lambda \to 0.75$. This observation is borne out by various entropy and correlation measures (see section 5). $\lambda = 0.75$ represents the state of maximal dynamical disorder.

Fourth, the transition region supports both static and propagating structures (fig. 1, $\lambda = 0.45$.) These particle-like structures are essentially *solitary waves*, quasi-periodic patterns of state change, which – like the "gliders" in Conway's Game of Life [8] – propagate through the array, constantly moving with respect to the fixed background of the lattice. The λ value for the Game of Life ($\lambda_{\text{Life}} = 0.273$) lies within the transition region for $K = 2$, $N = 9$ 2D CAs. Fig. 5 traces the time evolution of an array of 512 sites, and shows that the rule governing the behavior of fig. 1, $\lambda = 0.45$ supports several different kinds of particles, which interact with each other and with static periodic

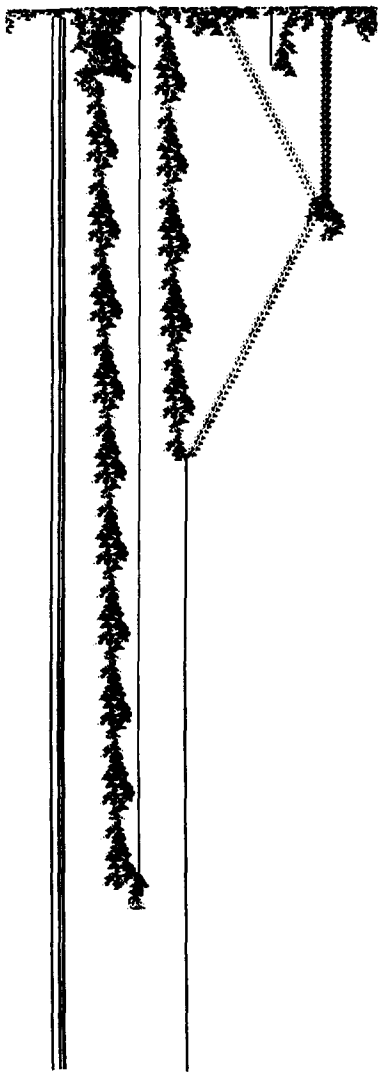


Fig. 5. Propagating structures and their interactions in an array of 512 cells with $\lambda = 0.45$.

structures in complicated ways. Note that the collision of a particle with a static periodic structure produces a particle traveling in the opposite direction. These propagating and static structures can form the basis for signals and storage, and interactions between them can modify either stored or transmitted information in the support of an overall computation. The proof that the Game of Life is computation-universal employs propagating "gliders" as signals and the period-2 "blinkers" as storage elements in the construction of a general purpose computer [2].

4.1. Complications

Finally, it must be pointed out that although the examples presented illustrate the general behavior of the dynamics as a function of λ, the story is not quite as simple as we have presented it here. The story is complicated by two factors, which will be detailed in the next section.

First, different traversals of λ space using the table-walk-through method make the transition to chaotic behavior at different λ values, although there is a well defined distribution around a mean value. Second, one does not always capture a second-order phase transition as neatly as in this example. Often, the dynamics jumps directly from fairly ordered to fairly disordered behavior, suggesting that both first- and second-order transitions are possible.

Despite these complications, the overall picture is clear: as we survey CA rule spaces using the λ parameter, we encounter a phase transition between periodic and chaotic behavior, and the most complex behavior is found in the vicinity of this transition, both qualitatively and quantitatively.

5. Quantitative overview of CA dynamics

In this section, we present a brief quantitative overview of the structural relations among the dynamical regimes in CA rule spaces as revealed by the λ parameter[#1].

The results of this section are based on experiments using 2D CAs with $K = 8$ and $N = 5$. Arrays are typically of size 64×64, and again, periodic boundary conditions are employed.

[#1] The results presented here summarize my Thesis research [17]. The reader is referred to that work for a more detailed presentation of the results in this section.

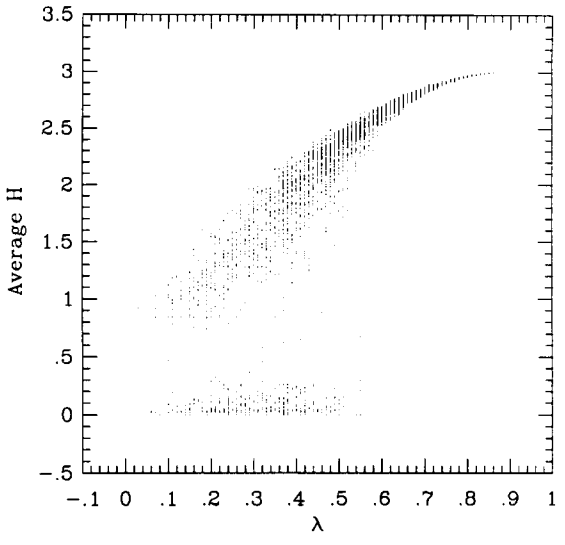


Fig. 6. Average single cell entropy $\overline{H}$ over λ space for approximately 10000 CA runs. Each point represents a different transition function.

5.1. *Measures of complexity*

The measures employed were chosen for their collective ability to reveal the presence of information in its various forms within CA dynamics.

5.1.1. *Shannon entropy*

We use Shannon's entropy H to measure basic information capacity. For a discrete process A of K states[#2]:

$$H(A) = -\sum_{i=1}^{K} p_i \log p_i. \qquad (2)$$

Fig. 6 shows the average entropy per cell, $\overline{H}$, as a function of λ for approximately 10000 CA runs. The random-table method was employed, so each point represents a distinct random transition table.

First, note the overall envelope of the data and the large variance at most λ points. Second, note the sparsely populated gap over $0.0 \lesssim \lambda \lesssim 0.6$ and between $0.0 \lesssim \overline{H} \lesssim 0.84$. This distribution appears to be bimodal, suggesting the presence of a phase transition. Third, note the rapid decrease in variability as λ is raised from ~0.6 to its maximum value of 0.875.

Two other features of this plot deserve special mention. First, the abrupt cutoff of low $\overline{H}$ values at $\lambda \approx 0.6$ corresponds to the *site–percolation* threshold $P_c \approx 0.59$ for this neighborhood template. Thus, we may suppose that, since λ is a dynamical analog of the site occupation probability P, the *dynamical* percolation threshold for a particular neighborhood template is bounded above by the *static* percolation threshold P_c. This is borne out by experiments with other neighborhood templates. For instance, the 9-neighbor template exhibits a sharp cutoff at $\lambda \approx 0.4$, which corresponds well with the site percolation threshold $P_c \approx 0.402$ for this lattice.

The second feature is the "ceiling" of the gap at $\overline{H} \approx 0.84$. This turns out to be the average entropy value for one of the most commonly occurring chaotic rules. In such rules the dynamics has collapsed onto only two states – s_q and one other – and the rule is such that a mostly quiescent neighborhood containing one non-quiescent state maps to that non-quiescent state. In 1D CAs, such rules give rise to the familiar triangular fractal pattern known as the Sierpiński gasket. There are many ways to achieve such rules, and they can be achieved at very low λ values. Most of the low-λ chaotic rules are of this type.

The entropy data of fig. 6 suggest an anomaly at intermediate parameter values, possibly a phase transition between two kinds of dynamics. Since there seems to be a discrete jump between low and high entropy values, the evidence points to a first-order transition, similar to that observed between the solid and fluid phases of matter. However, the fact that the gap is not completely empty suggests the possibility of second-order transitions as well.

The table-walk-through method of varying λ reveals more details of the structure of the entropy data. Fig. 7 shows four superimposed examples of the change in the average cell entropy as we vary the λ value of a table. Notice that in each of the four cases the entropy remains fairly close to zero

[#2] Throughout, log is taken to the base 2, thus the units are bits.

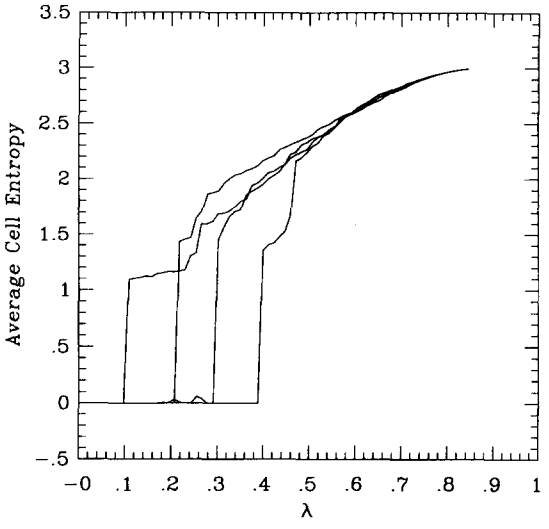


Fig. 7. Superposition of 4 transition events. Note the different λ values at which the transitions take place.

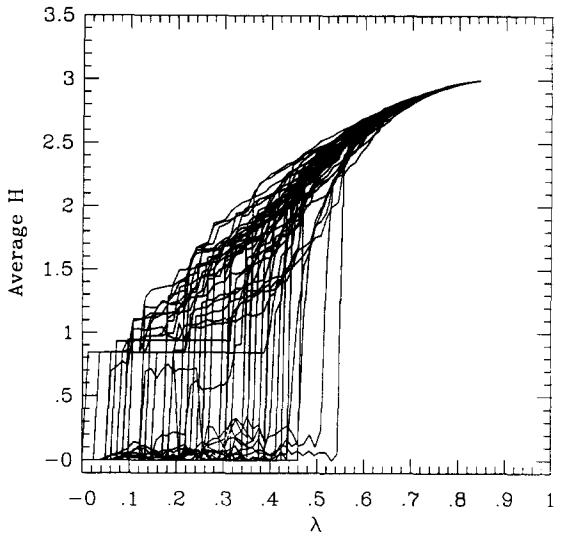


Fig. 8. Superposition of 50 transition events, showing the internal structure of fig. 6.

until – at some critical λ value – the entropy jumps to a higher value, and proceeds fairly smoothly towards its maximum possible value as λ is increased further. Such a discontinuity is a classic signature of a first-order phase transition. Most of our complexity measures exhibit similar discontinuities at the same λ value *within a particular table*.

Notice also that the λ value at which the transition occurs is different for each of the four examples. Obviously, the same thing – a jump – is happening as we vary λ in each of these examples, but it happens at different values of λ. When we superimpose 50 runs, as in fig. 8. we see the internal structure of the entropy data envelope plotted in fig. 6.

Since we have located the transition events, we may line up these plots by the events themselves, rather than by λ, in order to get a clearer picture of what is going on before, during, and after the transition. This is illustrated in fig. 9. The abcissa is now measured in terms of Δλ: the distance from the transition event. Fig. 10 shows the same data as fig. 8 but lined up by Δλ.

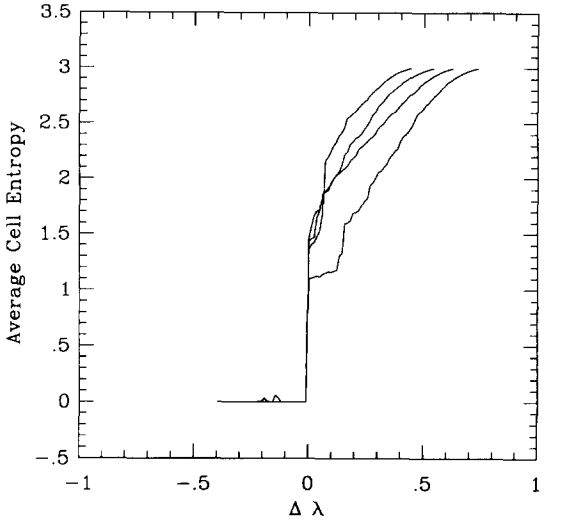


Fig. 9. Plots lined up by the transition event, rather than by λ. Δλ is the distance from the transition event.

5.1.2. *Mutual information*

In order for two distinct cells to cooperate in the support of a computation, they must be able to affect one another's behavior. Therefore, we should be able to find correlations between events taking place at the two cells.

The mutual information $I(A;B)$ between two cells A and B can be used to study correlations in systems when the values at the sites to be measured cannot be ordered, as is the case for the states of the cells in cellular automata [19].

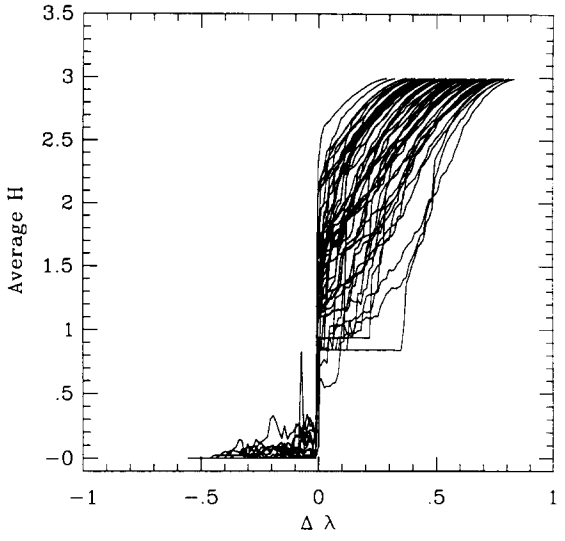


Fig. 10. Superposition of 50 transition events lined up by $\Delta\lambda$. Compare with fig. 8.

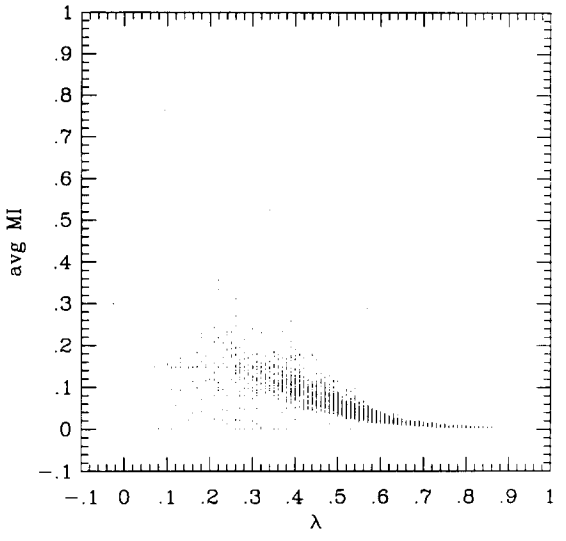


Fig. 11. Average mutual information between a cell and itself at the next time step.

The mutual information is a simple function of the individual cell entropies, $H(\mathrm{A})$ and $H(\mathrm{B})$, and the entropy of the two cells considered as a joint process, $H(\mathrm{A},\mathrm{B})$:

$$I(\mathrm{A};\mathrm{B}) = H(\mathrm{A}) + H(\mathrm{B}) - H(\mathrm{A},\mathrm{B}). \qquad (3)$$

This is a measure of the degree to which the state of cell A is correlated with the state of cell B, and vice versa.

Fig. 11 shows the average mutual information between a cell and itself at the next time step. Note the tight convergence to low values of the mutual information for high λ and the location of the highest values.

The increase of the mutual information in a particular region is evidence that the correlation length is growing in that region, further evidence for a phase transition.

Fig. 12 shows the behavior of the average mutual information as λ is varied, both against λ and $\Delta\lambda$. The average mutual information is essentially zero below the transition point, it jumps to a moderate value at the transition, and then decays slowly with increasing λ. The jump in the mutual information clearly indicates the onset of the chaotic regime, and the decaying tail indicates the approach to effectively random dynamics. The lack of correlation between even adjacent cells at high λ means that cells are *acting* as if they were independent of each other, even though they are causally connected. The resulting global dynamics is the same as if each cell picked its next state at uniform random from among the K states, with no consideration of the states of its neighbors. This kind of global dynamics is predictable in the same statistical sense that an ideal gas is globally predictable. In fact it is appropriate to view this dynamical regime as a hot gas of randomly flipping cells.

Fig. 13 shows the average mutual information curves for several different temporal and spatial separations. Note that the decay in both time and space is slowest in the middle region.

At intermediate λ values, the dynamics support the preservation of information locally, as indicated in the peak in correlations between distinct cells. If cells are cooperatively engaged in the support of a computation, they must exhibit some – but not *too* much – correlation in their behaviors. If the correlations are too strong, then the cells are overly dependent, with one mimicing the other – not a cooperative computational enterprise.

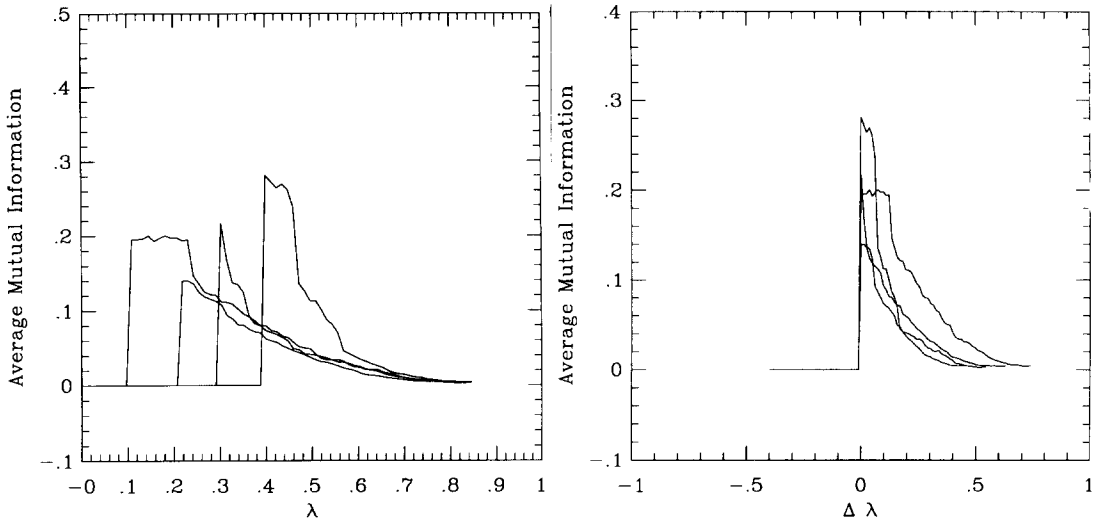


Fig. 12. Average mutual information versus λ and Δλ. The mutual information in this case is for a single time step at a single cell.

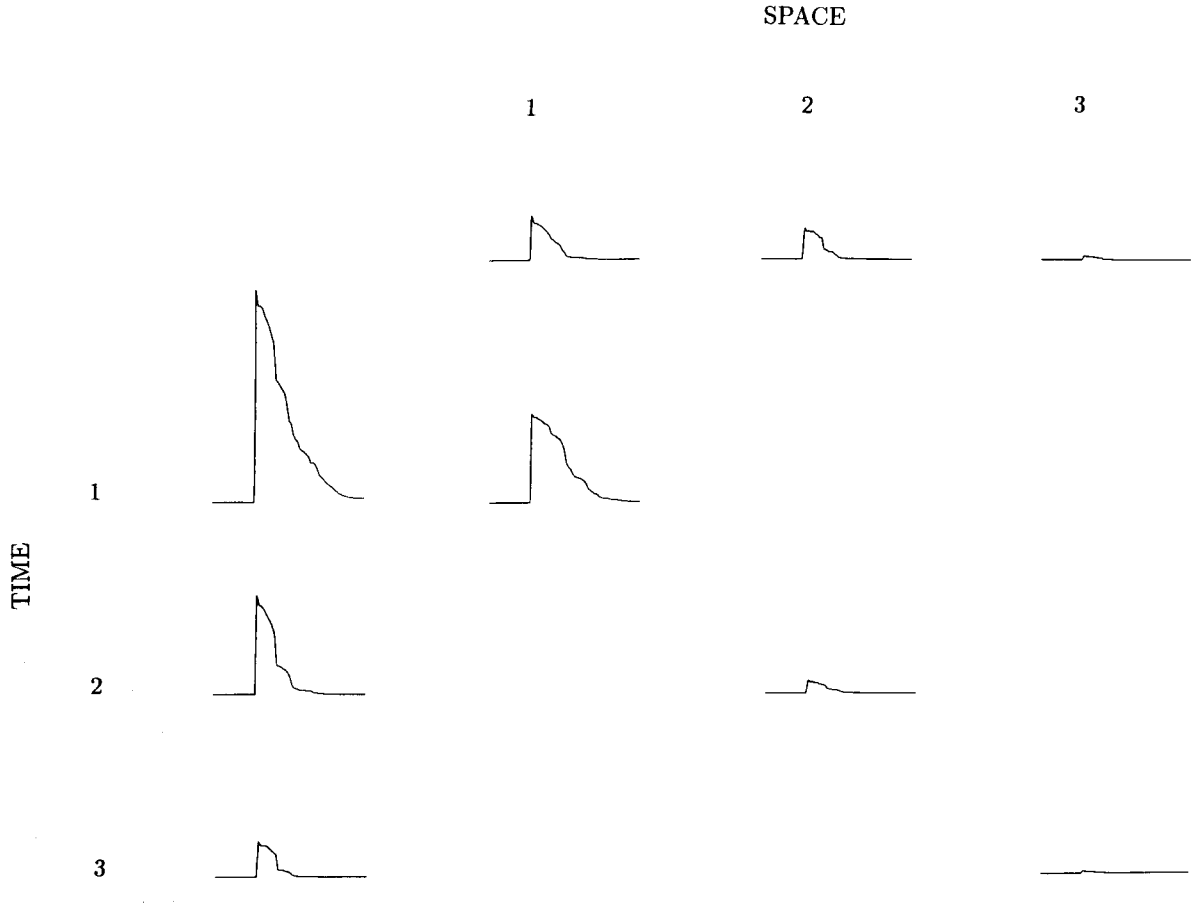


Fig. 13. Decay of average mutual information in space and time.

On the other hand, if the correlations are too small, then the cells are overly *independent*, and again, they cannot cooperate in a computational enterprise, as each cell does something totally unpredictable in response to the state of the other. Correlations in behavior imply a kind of common code, or protocol, by which changes of state in one cell can be recognized and understood by the other as a *meaningful signal*. With no correlations in behavior, there can be no common code with which to communicate information.

6. Mutual information and entropy

It is often useful to examine the way in which observed measures behave when plotted against one another, effectively removing the (possibly unnatural) ordering imposed by the control parameter.

Of the measures we have looked at, the most informative pair when plotted against each other are the mutual information and the average single cell entropy. The relationship between these two measures is plotted in fig. 14. Again, we see clear evidence of a phase transition.

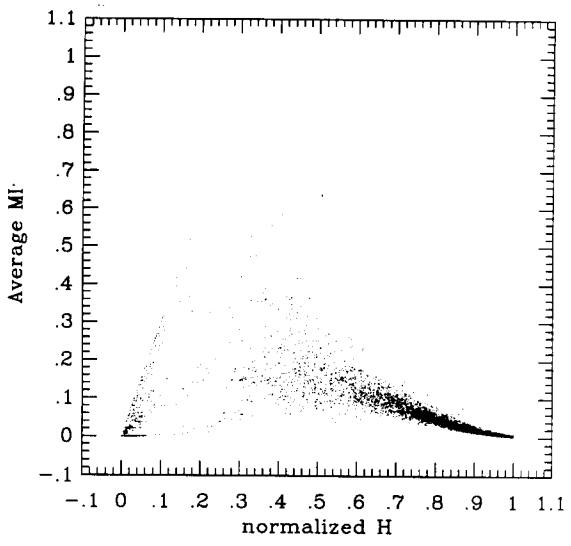


Fig. 14. Average mutual information versus average single cell entropy $\overline{H}$. The mutual information in this case is computed between a cell and itself at the next time step. The entropy is normalized to 1.0.

The envelope of the relationship is bounded below the transition by the linear bound that H places on the mutual information. All of the points on this line are for periodic CAs. This line intersects the curve bounding the envelope *above* the transition at an entropy value $H_c \approx 0.32$ on the normalized entropy scale.

This is a *very* informative plot. There is a clear, sharply defined maximum value of mutual information at a specific value of the entropy, and the mutual information falls off rapidly on either side. This seems to imply that there is an *optimal working entropy* at which CAs exhibit large spatial and temporal correlations. Why should this be the case?

Briefly, information storage involves *lowering* entropy while information transmission involves *raising entropy* [10]. In order to compute, a system must do both, and therefore must effect a trade-off between high and low operating entropy. It would seem from the work reported here that this trade-off is optimized in the vicinity of a phase transition.

A similar relationship has been observed by Crutchfield at Berkeley in his work on the transition to chaos in continuous dynamical systems [6]. This relationship is illustrated in fig. 15. Briefly, the ordinate of this plot – C – is a measure of the size of the minimal finite state machine required to recognize strings of 1's and 0's generated by a dynamical system (the logistic map, in this case) when these strings are characterized by the normalized per-symbol entropy listed on the abcissa. The observance of this same fundamental entropy/complexity relationship in these different classes of dynamical systems is very exciting.

These relationships support the view that, rather than increasing monotonically with randomness – as is the case for the usual measures of complexity, such as that of Chaitin and Kolmogorov [4, 16] – complexity increases with randomness only up to a point – *a phase transition* – after which complexity *decreases* with further in-

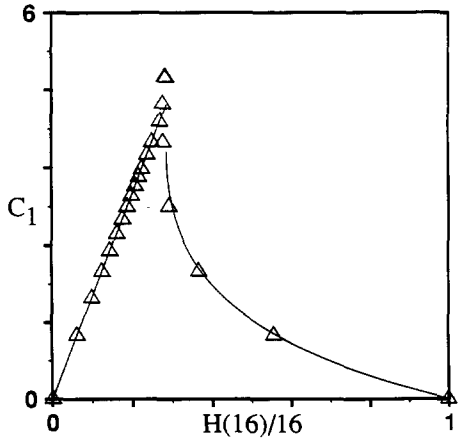


Fig. 15. Crutchfield's plot of machine complexity versus normalized per-symbol entropy for the logistic map. Compare with fig. 14.

creases in randomness, so that total disorder is just as "simple", in a sense, as total order. Complex behavior involves a mix of order and disorder.

7. Phase transitions and computation

What does all of this tell us about emergent computation? The answer is that *information* becomes an important factor in the dynamics of CAs in the vicinity of the phase transition between periodic and chaotic behavior. Only in the vicinity of this phase transition can information propagate over long distances without decaying appreciably. This allows for the long-range correlations in behavior, sensitivity to "size", extended transients, etc., which are necessary for the support of computation. By contrast, the ordered regime does not allow information to propagate at all, whereas the disordered regime propagates effects too well, causing information to decay rapidly into random noise.

If it is true that these phase-transition dynamics – especially "critical" or second-order dynamics – support the possibility of emergent computation, then we should be able to find analogs for various well-known features of computation in the phenomenology of phase transitions, and vice versa. In the following sections, we point out several possible analogs, and offer an interpretation which suggests that computation as we know it is really just a special case of a more universal physical phenomenon.

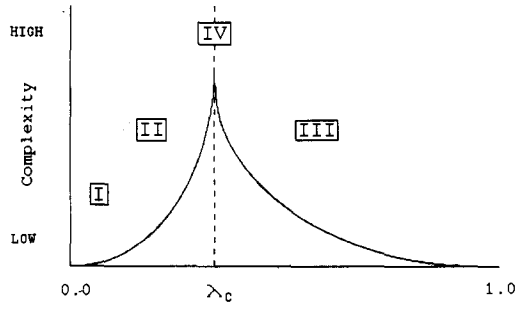


Fig. 16. Location of the Wolfram classes in λ space.

7.1. Locating the Wolfram classes

First, there is an obvious mapping of the Wolfram classes onto the spectrum of dynamical possibilities over the λ space: classes I and II constitute the *ordered* phase, while class III constitutes the *disordered* phase. Because of their long transients, propagating structures, large correlation lengths, and other statistical properties, the only logical choice for the location of class IV CAs is *at* the transition between these two phases of dynamical behavior. Fig. 16 shows how the Wolfram classes fit into the λ spectrum.

This also explains why one expects class IV CAs to constitute a set of measure 0. In the thermodynamic limit, the phase transition is located along a $(K-2)$-dimensional hyperplane in the rule space for K-state CAs (see ref. [17]). Hyperplanes embedded in higher-dimensional spaces constitute sets of measure 0. However, if we know where to look for a set of measure 0, we can find many instances. As we go to the thermodynamic limit, we can locate the phase transition more and more precisely, and hence we should be able to locate class IV CAs in arbitrarily large rule spaces even though they constitute a set of measure 0.

If Wolfram is correct in attributing the capacity for universal computation to class IV CAs, then when we locate class IV CAs at a phase transition, we are also locating universal computation at a phase transition.

7.2. Complexity classes

One obvious property of computations for which we would like to find an analog in phase-transition phenomena is the existence of the various complexity classes. Some computations may be performed using an amount of time or space which is only a linear – or even a constant – function of the "size" of the input, while other computations exhibit polynomial, or even exponential dependence [9]. Where can we find a natural analog of these complexity classes within the phenomenology of phase transitions?

The obvious answer is in the divergence of transient times as one approaches the phase transition. As illustrated in the qualitative dynamics of 1D CAs, for λ values far from the transition point, transients die out in time which is independent of the size of the array. As λ approaches the transition point, transients begin to show more and more dependence on array size. For values of λ very near a "critical" transition, this size dependence appears to be exponential or worse. This is true whether we approach the transition from the ordered regime or the disordered regime, which suggests that in addition to the familiar complexity-class hierarchy for halting computations, there should be a similar complexity-class hierarchy for *non*-halting computations.

7.3. The Halting problem

This last point brings up another property of computation which should be reflected in phase-transition dynamics.

Some computations halt, and some do not. For some computations, we can decide whether or not they will halt. However, Turing demonstrated that for certain classes of machines this "halting problem" is undecidable: there exist computations for which it is not possible to decide whether or not they will halt.

Thus, with respect to our ability to decide the ultimate outcome of computations, there are essentially three possibilities: we can determine that they will halt, we can determine that they will not halt, or we cannot determine whether or not they will halt.

As we have seen, there are three similar possibilities for the ultimate outcome of the evolutions of CAs. CAs below the transition point rapidly "freeze up" into short-period behavior from any possible initial configuration. On the other hand, CAs above the transition point will never freeze into periodic behavior, settling down rapidly instead to chaotic behavior. Thus, we can predict the ultimate dynamics of CAs away from the transition point with a high degree of certainty.

For CAs in the vicinity of the transition, however, both of these ultimate dynamical outcomes are possible, and because of the extended transients, it will be "effectively" undecidable whether a particular rule operating on a particular initial configuration will ultimately lead to a frozen state or not for this range of λ.

Thus, we can identify a natural analog of Turing's Halting problem in what we call the *Freezing problem*: for an arbitrary CA in the vicinity of the transition point, will the dynamics ultimately "freeze up" into short-period behavior or not? It is quite likely that the freezing problem is undecidable.

8. The natural domain of information

Let us now lay out in general outline an interpretation that will tie together all of these disparate phenomena into a coherent picture of the nature of computation. The reader should bear in mind that this interpretation, although strongly supported by evidence, is only a conjecture at this point; many details remain to be worked out.

8.1. Solids, fluids, and dynamics

We propose that the *solid* and *fluid* phases of matter, with which we are so familiar from everyday experience, are much more fundamental aspects of nature than we have supposed them to be. Rather than merely being possible states of matter, they constitute *two fundamental universality classes of dynamical behavior*.

We know solids and fluids primarily as states of matter because up until quite recently, everything that exhibited dynamical behavior was made up of some kind of material. Now, however, with the availability of computers, we are able to experiment with dynamics abstracted from any particular material substrate. The findings reported in this paper suggest that for dynamical systems in general – whether purely formal or manifestly material – there are primarily only two ultimate dynamical possibilities.

However, these two universality classes are separated by a *phase transition*. The dynamics of systems within this transition region – especially the "critical" systems – appear to support the basic mechanisms necessary for information transmission, storage, and modification, and therefore provide the capacity for emergent computation. Thus, a third possibility is that systems can be constructed in such a way that they manage to avoid either of the two primary dynamical outcomes by maintaining themselves on indefinitely extended transients.

It is a system's capacity for supporting a dynamics of information that allows complex behavior in the vicinity of a phase transition. This in turn allows for the possibility of the freezing problem. Since computers and computations are specific instances of material and formal systems respectively, they are also ultimately bound by these universality classes. Therefore, if this interpretation is correct, the halting problem can be seen as a specific instance of the more general freezing problem for dynamical systems. We can therefore view computations as special instances of the kinds of processes that occur in a physical system in the vicinity of a solid/liquid or a liquid/vapor transition.

8.2. Related work

Others have been working on the problem of finding structure in the rule spaces of cellular automata and other, similar spatially distributed dynamical systems.

In my initial investigations with the λ parameter [18], I suggested that Wolfram's class IV CAs constituted a transition between class II and class III, that is, between periodic and chaotic dynamics.

Kauffman [14, 13] has investigated a class of related dynamical systems known as *Boolean nets*, in which he finds a similar phase transition between ordered and disordered dynamics.

Vichniac, Tamayo and Hartman [25] discovered that the Wolfram classes could be recovered by varying the frequency of two simple rules in an inhomogeneous cellular automaton. They also suggested a relation between critical slowing down and the halting problem.

Packard and Li [20] have mapped out the space of "elementary" $K = 2$, $N = 3$, 1D CAs fairly completely, using a parameterization scheme similar to λ.

Packard [22] has also performed an interesting series of experiments in which he "adapts" CA rules by selecting for certain behaviors. He finds an initially random population of rules will drift towards the phase-transition region. His interpretation of this phenomenon is that it is easier to find rules which will *compute* the desired behavior – by making use of a general computational capacity – than it is to find rules that are "hard-wired" to produce *only* the desired behavior.

McIntosh [21] has applied the mean-field approach of Gutowitz [12, 11] and suggests that the Wolfram classes can be distinguished on the basis of simple features of the mean field theory curves.

Wootters [30] has applied mean-field theory to explain the results from the λ parameter, and has

been able to reproduce many of the features of fig. 6.

Together with Crutchfield's work mentioned earlier, these results collectively point to the existence of a phase transition in the spectrum of dynamical systems, and also suggest that the complex dynamics of systems in the vicinity of a phase transition rest on a fundamental capacity for processing information.

8.3. *Questions*

There are many questions that need to be addressed. For instance, can the "fluid" dynamical systems be further divided up into "gases" and "liquids"? There is some evidence for both solid/liquid and liquid/gas transitions in the space of CAs [17].

How might these issues be addressed by statistical mechanics, which has been very effective in treating phase transitions in general? Can analogs for temperature, pressure, volume, and energy be found? There is some evidence that equivalent measures can be defined [6, 17]. On the other hand, it is possible that statistical mechanics alone will not be able to fully treat phase-transition phenomena without being augmented by ideas from the theory of computation.

What are the implications for optimization techniques such as simulated annealing [15], which call for extended stays in the vicinity of the freezing point? It is interesting that this is the very point at which we would expect information processing to emerge spontaneously within the system being annealed – suggesting that the real reason for hovering in the vicinity of the freezing point is to allow the system to compute its own solution via an emergent computation.

How are the notions reported here related to Bak's self-organized criticality [1]? In many ways, it seems that Bak has discovered that dynamical systems can be made to *boil* when driven in the right way, which is a phenomenon we would expect at a phase transition. In fact, Bak has suggested that Conway's game of Life is a self-organized critical system, although he does not bring Life's computational capacity into the discussion.

Finally, what are the implications for understanding the origin and evolution of life? One of the most exciting implications of this point of view is that life had its origin in just these kinds of extended transient dynamics. Looking at a living cell, one finds phase-transition phenomena everywhere. The point of view advocated here would suggest that we ourselves are examples of the kind of "computation" that can emerge in the vicinity of a phase transition given enough time.

Now nature is not so beneficient as to maintain conditions at or near a phase transition forever. Therefore, in order to survive, the early extended transient systems that were the precursors of life as we now know it had to gain control over their own dynamical state. They had to learn to maintain themselves on these extended transients in the face of fluctuating environmental parameters, and to steer a delicate course between too much order and too much chaos, the Scylla and Charybdis of dynamical systems. Such transient systems must have "discovered" how to make use of their intrinsic information processing capability in order to sense and respond to their local environment. Evolution has been the process by which such systems have managed to gain local control over more and more of the environmental variables affecting their ability to maintain themselves on extended transients with essentially open futures.

9. Conclusion

Von Neumann observed that[#3]:

"There is thus this completely decisive property of complexity, that there exists a critical size below which the process of synthesis is degenerative, but above which the phenomenon of synthesis, if

[#3]John von Neumann, in his 1949 University of Illinois lectures on the Theory and Organization of Complicated Automata [26].

properly arranged, can become explosive, in other words, where syntheses of automata can proceed in such a manner that each automaton will produce other automata which are more complex and of higher potentialities than itself".

Although we are using a slightly different sense of "complexity" than von Neumann, the results of this paper support his observation. More importantly, however, we suggest that a similar observation can be made in the case of *too much* "complexity": *above* a certain level of "complexity", the process of synthesis is also degenerative.

In other words, we find that there exist an *upper* limit as well as a *lower* limit on the "complexity" of a system if the process of synthesis is to be non-degenerative, constructive, or open ended. We also find that these upper and lower bounds seem to be fairly close together and are located in the vicinity of a phase transition.

As the systems near the phase transition exhibit a range of behaviors which reflects the phenomenology of computations surprisingly well, we suggest that we can locate computation within the spectrum of dynamical behaviors at a phase transition here at the "edge of chaos".

Acknowledgements

Many people have contributed to the ideas presented here. I have benefitted greatly from conversations with Richard Bagley, Jim Crutchfield, Doyne Farmer, Howard Gutowitz, Hyman Hartman, Stuart Kauffman, Wentian Li, Norman Packard, Steen Rasmussen, Rob Shaw and Bill Wooters. Stephanie Forrest has been a long time intellectual companion and critic, and served as midwife in the delivery of this paper.

References

[1] P. Bak, C. Tang and K. Wiesenfeld, Self-organized criticality, Phys. Rev. A 38 (1988) 364–374.

[2] E. Berlekamp, J.H. Conway and R. Guy, Winning Ways for Your Mathematical Plays (Academic Press, New York, 1982).

[3] A.W. Burks, Essays on Cellular Automata (University of Illinois Press, Urbana, IL, 1970).

[4] G. Chaitin, J. Assoc. Comput. Mach. 13 (1966) 145.

[5] E.F. Codd, Cellular Automata (Academic Press, New York, 1968).

[6] J.P. Crutchfield and K. Young, Computation at the onset of chaos, in: Complexity, Entropy, and Physics of Information, ed. W. Zurek (Addison–Wesley, Reading, MA, 1990).

[7] E. Fredkin and T. Toffoli, Conservative logic, Int. J. Theor. Phys. 21 (1982) 219–253.

[8] M. Gardner, Mathematical games: The fantastic combinations of John Conway's new solitaire game 'Life', Sci. Am. 223(4) (October 1979) 120–123.

[9] M.R. Garey and D.S. Johnson, Computers and Intractability (Freeman, San Fransisco, 1979).

[10] L.L. Gatlin, Information Theory and the Living System (Columbia Univ. Press, New York, 1972).

[11] H.A. Gutowitz, A hierarchical classification of cellular automata, in: Proceedings of the 1989 Cellular Automata Workshop, ed. H.A. Gutowitz (North-Holland, Amsterdam, 1990), Physica D, to be published.

[12] H.A. Gutowitz, J.D. Victor and B.W. Knight, Local structure theory for cellular automata, Physica D 28 (1987) 18–48.

[13] S.A. Kauffman, Emergent properties in random complex automata, Physica D 10 (1984) 145–156.

[14] S.A. Kauffman, Metabolic stability and epigenesis in randomly constructed genetic nets, J. Theor. Biol. 22 (1969) 437–467.

[15] S. Kirkpatrick, C.D. Gelatt and M.P. Vecchi, Optimization by simulated annealing, Science 220 (1983) 671–680.

[16] A.N. Kolmogorov, Prob. Inf. Transm. 1 (1965) 1.

[17] C.G. Langton, Computation at the edge of chaos, Ph.D. Thesis, University of Michigan (1990).

[18] C.G. Langton, Studying artificial life with cellular automata, Physica D 22 (1986) 120–149.

[19] W. Li, Analyzing Complex Systems, Ph.D. Thesis, Columbia University (1989).

[20] W. Li and N.H. Packard, Structure of elementary cellular automata rule-space, Complex Systems, submitted for publication (1990).

[21] H.V. McIntosh, in: Proceedings of the 1989 Cellular Automata Workshop, ed. H.A. Gutowitz (North-Holland, Amsterdam), Physica D, to be published.

[22] N.H. Packard, Adaptation toward the edge of chaos, Technical Report, Center for Complex Systems Research, University of Illinois, CCSR-88-5 (1988).

[23] N.H. Packard and S. Wolfram, Two-dimensional cellular automata, J. Stat. Phys. 38 (1985) 901.

[24] A.R. Smith III, Simple computation-universal cellular spaces, J. Assoc. Comput. Mach. 18 (1971) 339–353.

[25] G.Y. Vichniac, P. Tamayo and H. Hartman, Annealed and quenched inhomogeneous cellular automata, J. Stat. Phys. 45 (1986) 875–883.

[26] J. von Neumann, Theory of self-reproducing automata, 1949 University of Illinois Lectures on the Theory and Organization of Complicated Automata, ed. A.W. Burks (University of Illinois Press, Urbana, IL, 1966).
[27] S. Wolfram, Statistical mechanics of cellular automata, Rev. Mod. Phys. 55 (1983) 601–644.
[28] S. Wolfram, ed., Theory and Applications of Cellular Automata (World Scientific, Singapore, 1986).
[29] S. Wolfram, Universality and complexity in cellular automata, Physica D 10 (1984) 1–35.
[30] W.T. Wootters and C.G. Langton, Is there a sharp phase transition for deterministic cellular automata?, in: Proceedings of the 1989 Cellular Automata Workshop, ed. H.A. Gutowitz, to appear in Physica D (1990).