in Physics Letters A **171** (1992) 52 — 60. UIUC-BI-CCSR-92–08

Unreconstructible at Any Radius

James P. Crutchfield

Center for Complex Systems Research
Beckman Institute
405 North Mathews
Urbana, Illinois 61801 USA[1]

Abstract

Modeling pattern data series with cellular automata fails for a wide range of deterministic nonlinear spatial processes. If the latter have finite spatially-local memory, reconstructed cellular automata with infinite radius may be required. In some cases, even this is not adequate: an irreducible stochasticity remains on the shortest time scales. The underlying problem is illustrated and quantitatively analyzed using an alternative model class called cellular transducers.

Contents

List of Figures iii
Section 1 Introduction 1
Section 2 Reconstructing the Spatio-Temporal Dynamic 1
Section 3 Spatio-Temporal Models 2
Section 4 Generating Instruments versus Reconstructibility 4
Section 5 Finitary Cellular Transducers 5
Section 6 Reconstructed CA and SCA 8
Section 7 Indeterminacy 10
Section 8 Concluding Remarks 11
References and Notes 12

List of Figures

Figure 1 Space-time diagram of elementary CA 18. The horizontal axis gives the spatial site index; the vertical, time increasing downward. $N = 200$ sites are shown for 150 iterations, after 100 transient steps. Black cells denote $q_t^i = 1$; white, $q_t^i = 0$. The initial pattern was arbitrary. 6

Figure 2 Space-time diagram of FCT 18\222. Same initial pattern and format as in Figure 1. 6

Figure 3 Space-time diagram of FCT 22\222. Same initial pattern and format as in Figure 1. 7

Figure 4 Space-time diagram of FCT 90\222. Same initial pattern and format as in Figure 1. 7

Figure 5 Simulation of the nearest neighbor SCA estimated from FCT 18\222 spatial data series. Same initial pattern and format as in Figure 2. In lexicographically increasing order of neighborhood pattern $p \in \mathcal{Q}^3$, the conditional transition probabilities $Pr(q_{t+1}^i = 0|p)$ were estimated to be $(1, 0, 1, 0.0055, 0, 1, 0.0058, 0.905)$. The LUT estimation used spatial data series over 10^6 iterations on an $N = 500$ site lattice starting from an arbitrary initial state. 8

Figure 6 Simulation of the next-nearest neighbor estimated SCA. The same initial condition and simulation parameters as in the preceding figure. 9

Figure 7 Indeterminacy $\Xi(r)$ versus template spatial radius r for five example systems: (i) A uniformly random stochastic cellular automaton: $Pr(q_{t+1}^i|p) = 2^{-2r-1}$ (diamonds along upper boundary); (ii) Elementary cellular automaton rule 18 (X's), cf. Figure 1; (iii) finitary cellular transducer 90\222 (squares), cf. Figure 4; (iv) finitary cellular transducer 22\222 (circles), cf. Figure 3; (v) finitary cellular transducer 18\222 (triangles), cf. Figure 2. $\Xi(r)$ estimated from patterns on an $N = 500$ site lattice over 10^6 iterations after 10^6 transient iterations from a random initial pattern. The same initial pattern was used in each case. . 11

Introduction

Cellular automata (CA) form one of the simplest model classes for spatial pattern generating processes. CA have been proposed as models of pattern formation in natural systems.[2–5] Verifying this has been largely a matter of comparing CA behavior, as revealed in (say) space-time diagrams, snapshots of spatial patterns, and various macroscopic statistics produced during computer simulation, with natural patterns. More recently, several authors suggested that effective CA equations of motion, consisting of a look up table that maps neighborhood templates to next site value, could be inferred from pattern data time series.[6–10] The learning paradigm employed, however, did not take into account the effect of measurement distortion common in obtaining experimental data. The latter, though, can have a fundamental effect on the success of CA estimation, in particular, and spatial modeling, generally.

Measurements are only indirect representations of a process's internal states. One practical consequence of this basic physical fact is that cellular transducers (CT) which explicitly account for the measurement process, rather than cellular automata (CA), should be used as the computational model class for reconstructing the spatio-temporal dynamic from pattern data series. The latter includes data generated by discrete-state systems and the spatio-temporal symbolic dynamics of continuum-state extended systems, such as map lattices,[11] oscillator chains, and partial differential equations. The main difficulty is that estimated CA look up tables (LUTs) misrepresent the dynamics even if the observed behavior was generated by a deterministic process with finite local memory. Examples of nearest-neighbor binary alphabet CT with two local states are given below that require an infinite CA LUT for their deterministic dynamics to be (i) effectively reconstructed, (ii) approximately reconstructed, and (iii) not reconstructed at all. In these cases any estimated CA is stochastic and, as such, fails to capture obvious spatio-temporal structure. This leads to an overestimation of the degree of intrinsic randomness underlying the spatial data series.

Reconstructing the Spatio-Temporal Dynamic

The following considers the problem of reconstructing the governing spatio-temporal dynamic from a given discretized data set: the discrete-state spatial model inference problem. Its main goal goes beyond data prediction to learn a model of the governing process. The primary task in this is to infer internal states that are hidden by measurement distortion.

This problem is closely allied to that of reconstructing (chaotic) attractors from a single continuous-time data series, as introduced in 12. The goal in attractor reconstruction is the discovery of a new state space using information only available from the data. If the resulting representation is of sufficiently low dimension, the attractor's geometry can be graphically studied and various statistical measures of information content and the degree of unpredictability can be estimated.[13,14] It has been shown that reconstruction from continuous-state noiseless time series yields an equivalent representation in the new state space given sufficient data, a linear measuring instrument with infinite precision, and prior knowledge of the underlying process's dimension.[15]

More recently, reconstruction from continuous-state time series in the presence of noise has been considered.[6,16,17] As soon as one admits the existence of extrinsic noise, a new question

presents itself — How to distinguish between it and unpredictability due to a deterministic (chaotic) mechanism? The first step to addressing this is to layout the prior knowledge that an observer brings to the analysis of a data set. This includes not only the data set size and measurement resolution, but also biases reflected in the selected model. Proper accounting of prior information allows one to balance model complexity against prediction error in order to estimate an effective noise level and effective deterministic nonlinear equations of motion. Contemporary model order selection procedures,[6,18] however, do not address the problem of selecting an incorrect representation. Should one use Fourier or wavelet functions, or possibly neural networks with radial basis functions? Unfortunately, model class selection can be the overwhelming determinant in the success of nonlinear modeling, as will be shown below.

In an attempt to circumvent these problems, which are typical of continuum-state modeling, reconstruction under the hypothesis of discrete-amplitude, discrete-time data series was developed for stationary nonlinear dynamical systems.[19] The following considers the spatio-temporal analog of this: modeling discrete-state, discrete-time spatial data series. Previous work reconstructing cellular automata from data, via the estimation of a LUT that maps local patterns to future site values, has met with mixed success. Difficulties in modeling with CA have been attributed to insufficient neighborhood probes, improper discretization of continuous observables, and infinite forecasting complexity.[6,8,20] Though these must be addressed, here I suggest that there is a more fundamental problem: the CA hypothesis leads to the misidentification of the underlying process since it does not directly account for measurement distortion.

The analysis of the problem proceeds as follows. The spatio-temporal model classes of interest and a measure of nondeterminism are introduced first. This ends with a mathematical statement of space-time reconstruction. Then several formal properties of cellular transducers are illustrated via examples and space-time diagrams of CT are compared to reconstructed CA. This demonstrates how CA estimation fails and it suggests a classification of failure scenarios. Finally, the conclusion outlines the solution to this problem and mentions a physical implication of this type of misidentified stochasticity.

Spatio-Temporal Models

The following adopts the view of an observer making a series of measurements. Each measurement returns a pattern consisting of discretized values over a set of spatial sites. In the simplest spatial case, the observer obtains a one-dimensional pattern $\mathrm{s}_t = s_t^0 s_t^1 \ldots s_t^{N-1}, s_t^i \in \mathcal{A}$, consisting of a sequence of symbols s_t^i in some measurement alphabet $\mathcal{A}$ observed at site $i = 0, 1, \ldots, N-1$, and time t. To emphasize simplicity, alphabets will be binary $\mathcal{A} = \{0, 1\}$. This choice, though, in no way conditions the following results: they are not essentially different for larger finite and countably infinite alphabets.

A pattern's entropy density $h_\mu(\mathrm{s}_t)$ measures its degree of unpredictability; while its statistical complexity $C_\mu(\mathrm{s}_t)$ indicates the amount of information that must be remembered to make optimal predictions. One can estimate these from a data set and they give an indication of the computational capability of the underlying process in terms of its rate of information production and its memory capacity, respectively.[19] An ideal random process, for example, has high entropy,

but zero statistical complexity. Several remarks later on require a passing familiarity with these two quantities.

The following also assumes the reader is familiar with deterministic CA.[3] At time t the global state $\mathbf{q}_t$ of a CA is a sequence of symbols in some local state alphabet: $\mathbf{q}_t = q_t^0 q_t^1 \ldots q_t^{N-1}, q_t^i \in \mathcal{Q}$, for an N site lattice. The global state's temporal evolution is specified by the CA's rule table $\phi : \mathcal{Q}^{2r+1} \to \mathcal{Q}$ that maps a neighborhood pattern $p = q^{-r} \ldots q^0 \ldots q^r \in \mathcal{Q}^{2r+1}$ of radius r to the value of the site at the next time. That is, the symbols in the local state at the next time are determined by the equations of motion $q_{t+1}^i = \phi\left(q_t^{i-r} \ldots q_t^i \ldots q_t^{i+r}\right)$. The LUT ϕ specifies the local space-time dynamics and is identified by an integer index, the CA rule number.[3] The temporal evolution of the entire pattern is governed by the induced global mapping, or transducer, $\mathbf{\Phi}_\phi$ from a global state to a global state[21]

$$\mathbf{q}_{t+1} = \mathbf{\Phi}_\phi(\mathbf{q}_t) \tag{1}$$

Now consider two additional model classes for discrete-state spatial systems. The first is a probabilistic generalization of CA.[22] A stochastic CA (SCA) is specified by its neighborhood conditional probability transition table

$$q_{t+1}^i = \begin{cases} 0 & \text{with } Pr\left(0 | q_t^{i-r} \ldots q_t^i \ldots q_t^{i+r}\right) \\ 1 & \text{otherwise} \end{cases} \tag{2}$$

where $Pr(q|p)$ is the probability of local state q conditioned on seeing neighborhood pattern p at the previous time.

The degree of an SCA's nondeterminism is measured by the indeterminacy[23]

$$\Xi_{\mathbf{q}_0}(r) = -\sum_{p \in \mathcal{Q}^{2r+1}} Pr(p) \sum_{q \in \mathcal{Q}} Pr(q|p) \log_2 Pr(q|p) \tag{3}$$

Since the neighborhood distribution $Pr(p)$ depends on the initial condition $\mathbf{q}_0$, then so does Ξ. $\Xi(r)$ measures the uncertainty, in bits, of a site's value at the next time step given the knowledge of the current neighborhood pattern. Note that it's bounds, $0 \leq \Xi(r) \leq 1$, do not depend on the radius. Furthermore, if $\Xi(r) = 0$, then the SCA reduces to a deterministic CA of radius r: there is no choice in the future site values. Finally, it is monotone decreasing with radius: $\Xi(r') = 0 \Rightarrow \Xi(r) = 0, \forall r > r'$ and $\Xi(r) \leq \Xi(r'), \forall r \geq r'$. Note that estimation of the indeterminacy $\Xi(r)$ requires the reconstruction of a radius r SCA since it uses the associated neighborhood conditional probability distribution.

The final and new spatial model class, the (deterministic) cellular transducer (CT), explicitly incorporates the measurement act, unlike CA and SCA. A CT's local state is a pair $\left(q_t^i, s_t^i\right)$ of symbols, one q_t^i from a set $\mathcal{Q}$ of internal states and the other s_t^i from a set $\mathcal{A}$ of observed measurements. The global internal state $\mathbf{q}_t$ evolves as in a CA: there is an internal state update rule ϕ that operates on an internal neighborhood pattern to produce the next internal state q_{t+1}^i. In contrast to CA, however, an observer does not have direct access to the internal states, but instead measures symbols that are a spatially-local function of the internal state neighborhood. That is, the observed global state $\mathbf{s}_t = s_t^0 s_t^1 \ldots s_t^{N-1}, s_t^i \in \mathcal{A}$, is determined by a observation function ψ

$$s_{t+1}^i = \psi\left(q_t^{i-r} \ldots q_t^i \ldots q_t^{i+r}\right) \tag{4}$$

The equations of motion for a CT can be compactly written, using global transducers, as

$$\begin{aligned} \mathbf{q}_{t+1} &= \Phi_\phi(\mathbf{q}_t) \\ \mathbf{s}_{t+1} &= \Phi_\psi(\mathbf{q}_t) \end{aligned} \tag{5}$$

In the following the initial condition $\mathbf{q}_0$ will be an arbitrary pattern and periodic boundary conditions, $q_t^0 = q_t^N$ and $s_t^0 = s_t^N$, will be used.

When the number of internal states is finite, $\|\mathcal{Q}\| < \infty$, one refers to a finitary cellular transducer (FCT). Let ϕ and ψ denote the CA rule number associated with the ϕ and ψ rule tables. Then the FCT considered here will be denoted $\phi\backslash\psi$ with the first number identifying the internal state rule table and the second, the observation function. Simply stated, then, such FCT are "cellular automata with measuring instruments".[24]

The definitions lend themselves to a formal statement of the discrete-state spatial model inference problem for CA. Given a pattern data series $\{\mathbf{s}_t : t = 0, 1, 2, \ldots, K\}$ produced by a deterministic process with a spatially-local dynamic, are the effective equations of motion equivalent to some CA or SCA? In other words, does $\mathbf{s}_{t+1} = \Phi_{\phi_r}(\mathbf{s}_t)$ for some radius r CA rule or SCA conditional distribution ϕ_r? In answering these questions, the additional and very practical requirement is added that ϕ_r be reconstructible from the given data $\{\mathbf{s}_t\}$.[25]

The following will show, by way of examples here and more formally elsewhere, that the answer to these questions in general is negative.

The consequences are direct. CA form too weak an inference class for modeling discretized spatio-temporal data. Except for a restricted class of spatial systems, randomness in the underlying process is overestimated. Furthermore, SCA, which add an element of stochasticity seemingly to compensate for CA's lack of a nondeterministic capacity, are similarly incapable of properly modeling a wide range of spatial systems.

More constructively, the following indicates that CT are the correct modeling class for discretized data series generated by deterministic spatial processes. This will not be established here, only made plausible. Rather the following will use FCTs to show how the CA modeling hypothesis fails. This will then establish that, despite its simplicity, there is an important problem in practice as well as in principle. One, in fact, that has been largely ignored to date.

Generating Instruments versus Reconstructibility

There are several limits on what an FCT's observation function can do to the internal state information. For example, the observed patterns cannot be more random than the internal dynamic allows: $h_\mu(\mathbf{s}_t) \leq h_\mu(\mathbf{q}_t)$. This refers to the global internal and external states. If one looks at small patches, such as a single site, then the observed entropy $H(\mathcal{A})$ can be larger, within specified bounds, than the internal single site entropy $H(\mathcal{Q})$: $H(\mathcal{A}) \leq 2rH(\mathcal{Q})$. The statistical complexity, however, may or may not be decreased by a deterministic observation function.

This brings us to the question of what characterizes good, or "generating", measuring instruments. Informally, a generating observation function ψ produces output symbols that allow for the reconstruction of the internal-state equations of motion. In statistical terms, a generating instrument produces patterns $\{\mathbf{s}_t\}$ that have the same space-time entropy and statistical

complexity as the internal states $\{\mathbf{q}_t\}$: $h_\mu(\mathbf{s}_t) = h_\mu(\mathbf{q}_t)$ and $C_\mu(\mathbf{s}_t) = C_\mu(\mathbf{q}_t)$. A generating space-time instrument, in this sense, is a goal for the design of spatial data acquisition systems. Using such an instrument the effective equations of motion can be reconstructed, even without direct measurements of the internal local states.

It is important to emphasize that the notion of generating instrument refers to *prior knowledge* of the equations of motion. If one is given only a time series of spatial patterns, then one can only hope to reconstruct a dynamic consistent with the given data. Whether the instrument in this case was generating or not, cannot be answered in principle. And, although a minimal consistent model can be estimated, the "real" underlying dynamic will never be known.

This leads to the notion of a reconstructible dynamic, as distinguished from a generating instrument. Fix a deterministic or stochastic model class $\mathcal{M}$. Given an arbitrarily long data series, the equations of motion are reconstructible with respect to $\mathcal{M}$ if, when minimized, they are finite and reproduce the observed data exactly, if $\mathcal{M}$ is deterministic, or statistically,[26] if $\mathcal{M}$ is stochastic.

Finitary Cellular Transducers

Now that the context of spatial modeling has been outlined, this section gives four examples of FCTs to illustrate their typical behavior. Typical spatial data series from these FCTs will then be analyzed by reconstructing CA and SCA models.

Elementary CA 18 is equivalent to FCT 18\204 since $\psi = 204$ is the nearest neighbor identity. It is shown in Figure 1 for reference. The most notable property of CA 18 is the strong spatial and temporal periodicity. Every other cell is 0, except in those regions where one finds $1(00)^*1$ patterns. The other apparent feature consists of triangles in which contiguous 0 sequences shrink in length with time. The spatial entropy density was estimated to be $h_\mu(\mathbf{q}_t) \approx 0.52$ bits per site and the statistical complexity was $C_\mu(\mathbf{q}_t) = 1.3$ bits per site.

The observed patterns $\mathbf{s}_t$ generated by FCT 18\222 are shown in Figure 2 in the same format as the preceding figure. For direct comparison, the same initial state was used in both figures. This is reflected in the space-time coincidence of the 0–triangles, for example, in the two diagrams. Other than this there is little superficial commonality to the space-time diagrams. The $\psi = 222$ observation function obscures the period 2 structure in the CA 18 state. And, in contrast to CA 18, Figure 2 shows that FCT 18\222 generates triangles with a base pattern of $\ldots 0(11)^n 0(11)^n 0 \ldots, n = 1, 2, 3, \ldots$, that shrinks with time. The spatial entropy density and statistical complexity of the internal states is given by CA 18, as noted above. The observed spatial entropy density and statistical complexity were estimated to be $h_\mu(\mathbf{s}_t) \approx 0.51$ and $C_\mu(\mathbf{s}_t) \approx 2.5$ bits per site, respectively. Thus, although the internal and observed patterns are equally unpredictable, more information must be remembered to predict the FCT's observed patterns.

The observed patterns $\mathbf{s}_t$ generated by FCT 22\222 are shown in Figure 3. 0-triangles are apparent, along with a high proportion of $q_t^i = 1$ sites. There are occasional long temporal correlations apparent in the vertical lines of 0s. The observed entropy and complexity were estimated to be $h_\mu(\mathbf{s}_t) \approx 0.71$ and $C_\mu(\mathbf{s}_t) \approx 0.55$ bits per site. This should be compared to the

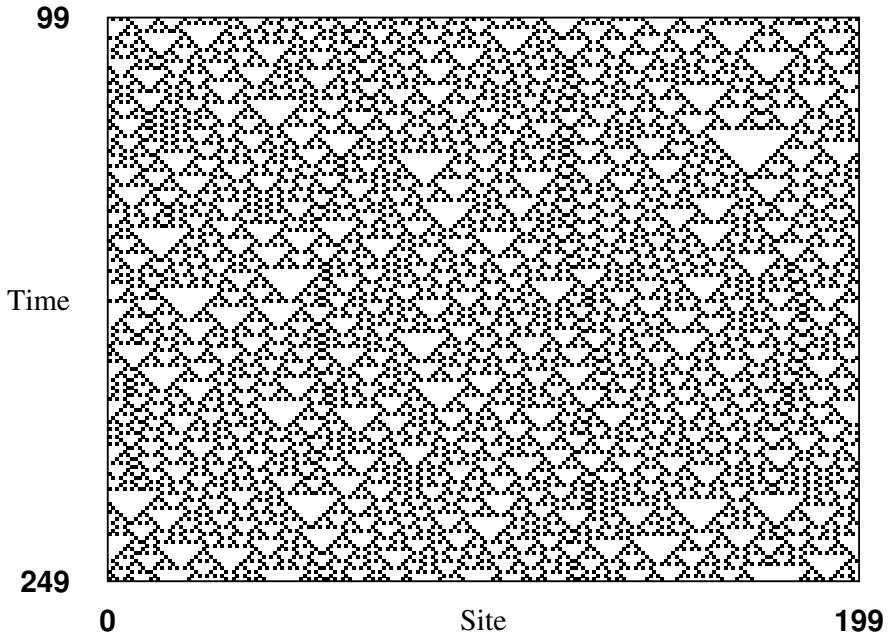

Figure 1 Space-time diagram of elementary CA 18. The horizontal axis gives the spatial site index; the vertical, time increasing downward. $N = 200$ sites are shown for 150 iterations, after 100 transient steps. Black cells denote $q_t^i = 1$; white, $q_t^i = 0$. The initial pattern was arbitrary.

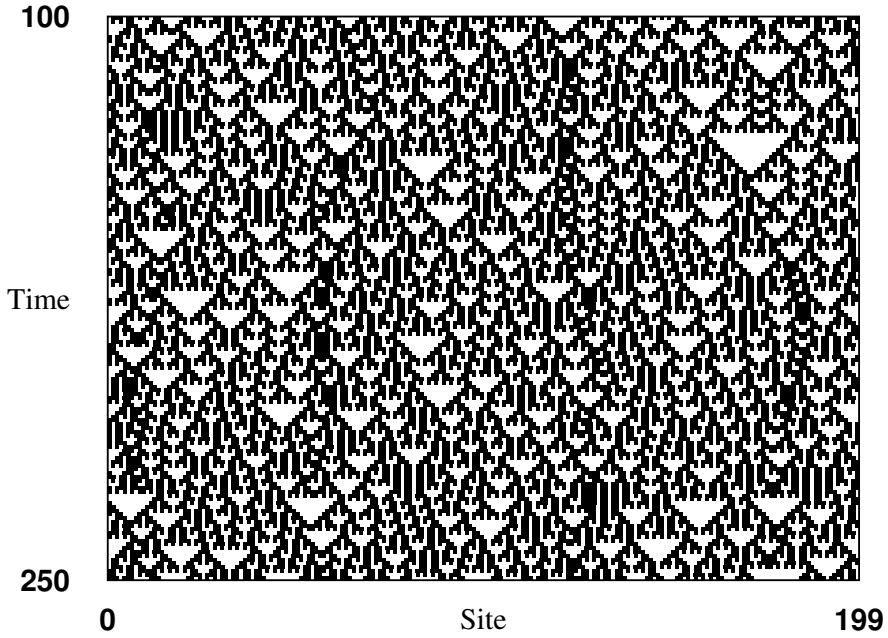

Figure 2 Space-time diagram of FCT 18\222. Same initial pattern and format as in Figure 1.

internal quantities which are equivalent to those for CA 22: $h_\mu(\mathbf{q}_t) \approx 0.78$ and $C_\mu(\mathbf{q}_t) \approx 0.92$ bits per site.

The observed patterns $\mathbf{s}_t$ generated by FCT 90\222 are shown in Figure 4. There are few 0-triangles, a very high proportion of $q_t^i = 1$ sites, and a number of isolated $q_t^i = 0$ sites.

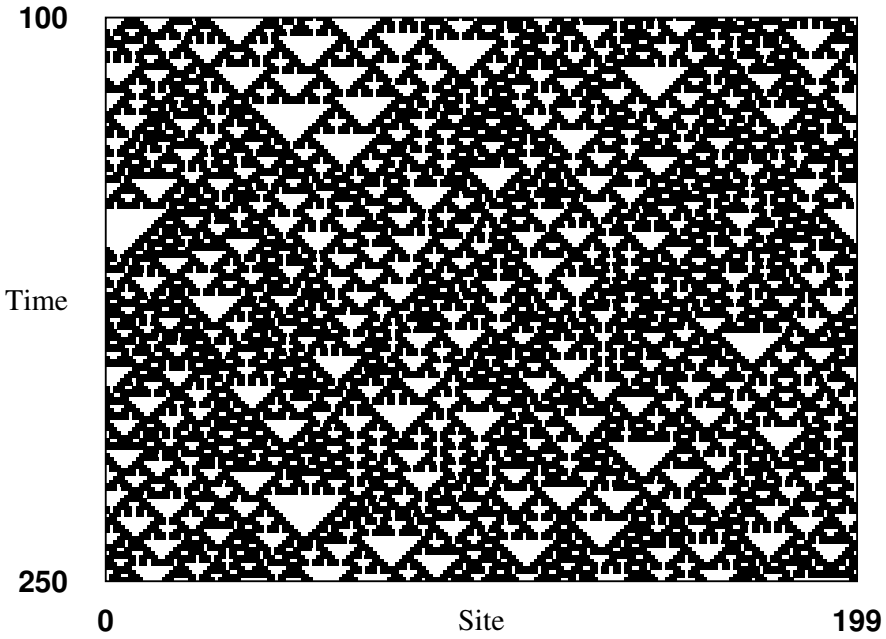

Figure 3 Space-time diagram of FCT 22\222. Same initial pattern and format as in Figure 1.

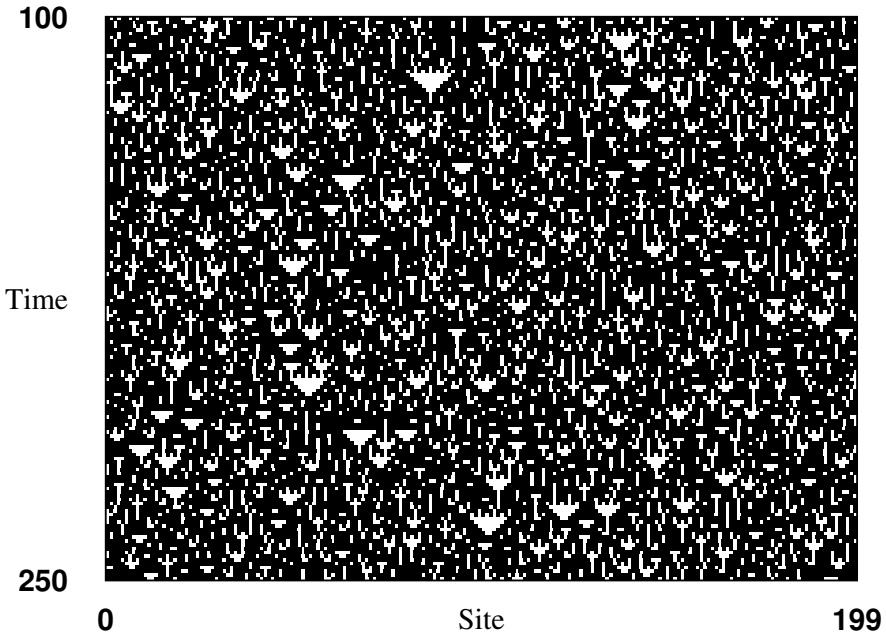

Figure 4 Space-time diagram of FCT 90\222. Same initial pattern and format as in Figure 1.

The observed entropy and complexity were $h_\mu(\mathrm{s}_t) \approx 0.76$ and $C_\mu(\mathrm{s}_t) \approx 0.13$ bits per site; the comparable internal quantities were $h_\mu(\mathrm{q}_t) \approx 1.00$ and $C_\mu(\mathrm{q}_t) \approx 0.00$ bits per site. These examples illustrate a property of deterministic instruments: the observed data's unpredictability cannot be larger than the internal process's; but the statistical complexity can be either increased or decreased.

Reconstructed CA and SCA

The hypothesis that the preceding spatial data series can be modeled within the class of CA or, more properly, SCA can now be evaluated. Recall the first figures (1 and 2) that compared CA 18 and FCT 18\222. There were notable deviations which were reflected in a large difference in complexity. Perhaps the observed patterns of FCT 18\222 could be generated by some different, possibly larger radius CA? After all, Figure 2 was obtained from figure 1 by a simple local function $\psi = 222$.

Estimating the nearest-neighbor conditional transition table gives an SCA with a finite indeterminacy of $\Xi(r) \approx 0.07, r \geq 1$. The estimated dynamic appears weakly stochastic. And so, one might expect it to capture a very large portion of FCT 18\222's behavior. For comparison, then, a simulation of the estimated SCA is shown in Figure 5. The same initial condition was used as that for FCT 18\222 in Figure 2. The figure indicates that the estimated SCA differs substantially from FCT 18\222. There are, for example, smaller scale structures, such as the 0-triangles, and a higher proportion of isolated 1s and 0s. These differences are reflected in a spatial entropy, $h_\mu(\mathbf{q}_t) \approx 1.0$, that is approximately 0.5 bits higher than found for FCT 18\222. The SCA's patterns are much less complex, $C_\mu(\mathbf{q}_t) \approx 0.0$, than those generated by FCT 18\222.

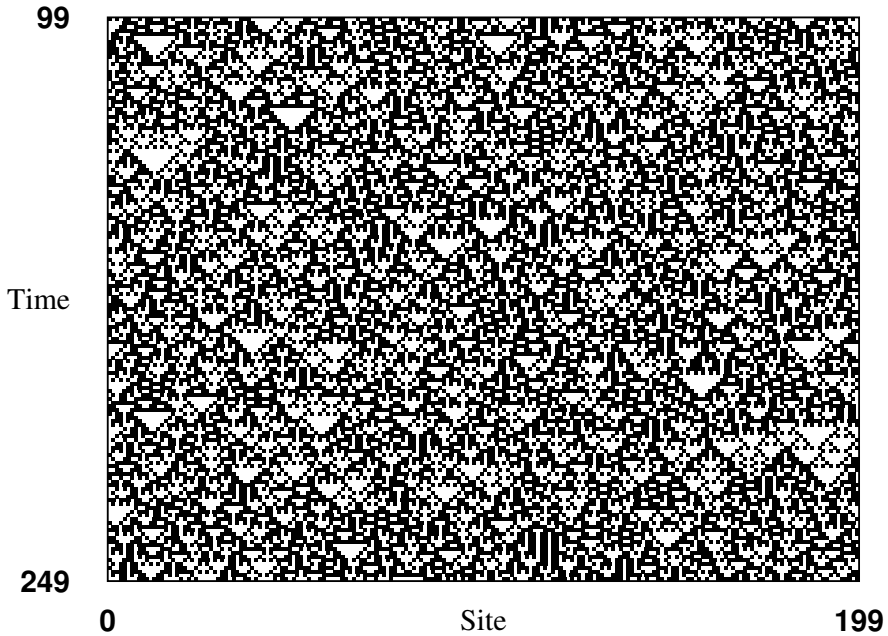

Figure 5 Simulation of the nearest neighbor SCA estimated from FCT 18\222 spatial data series. Same initial pattern and format as in Figure 2. In lexicographically increasing order of neighborhood pattern $p \in \mathcal{Q}^3$, the conditional transition probabilities $Pr(q^i_{t+1} = 0|p)$ were estimated to be $(1, 0, 1, 0.0055, 0, 1, 0.0058, 0.905)$. The LUT estimation used spatial data series over 10^6 iterations on an $N = 500$ site lattice starting from an arbitrary initial state.

But perhaps the situation will improve by estimating the next-to-nearest neighbor SCA? Figure 6 shows the result. It has $\Xi(r) = 0.06, r \geq 2$, $h_\mu(\mathbf{q}_t) \approx 0.76$, and $C_\mu(\mathbf{q}_t) \approx 0.90$. Certainly a smaller indeterminacy than found with the nearest-neighbor SCA and so even less stochastic.

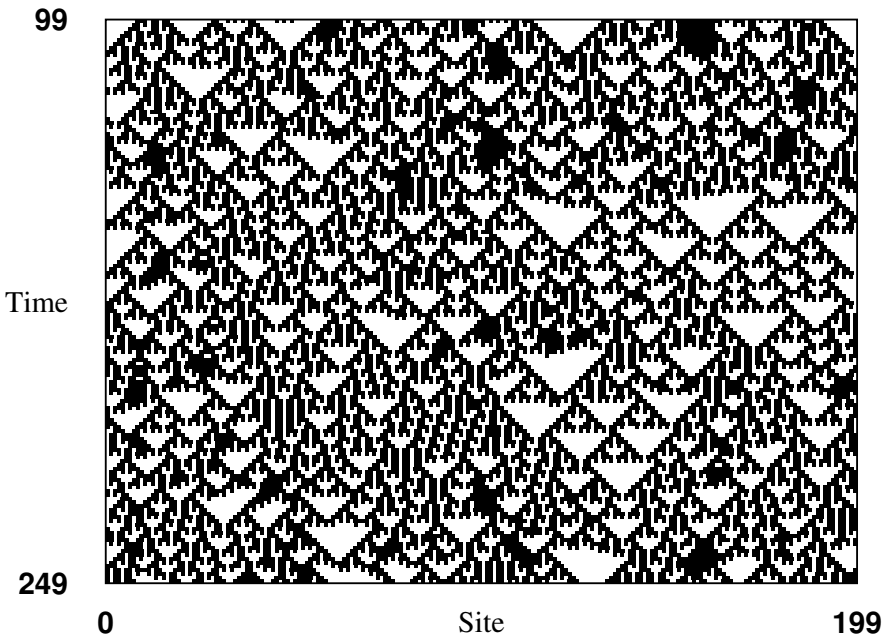

Figure 6 Simulation of the next-nearest neighbor estimated SCA. The same initial condition and simulation parameters as in the preceding figure.

Despite the small indeterminacy, the space-time structures still differ from FCT 18\222. There are now, in contrast to the near-neighbor SCA just shown, large 0-triangles, most of which do not have flat tops. In fact, the size distribution has shifted so that there is a preponderance of larger 0-triangles than in FCT 18\222. Additionally, there are larger space-time regions of contiguous 1s, which do not occur in FCT18\222. Compared to the near-neighbor SCA which exhibited more decorrelation than FCT 18\222, the next-to-nearest SCA errs in the other direction. There are highly structured patterns on longer space and time scales than in the nearest-neighbor SCA and in FCT 18\222. The structures are in radically different locations in the diagram, as expected for a chaotic system subject to even small amounts of extrinsic noise. The entropy density is now only 0.25 bits per site higher than FCT 18\222; the difference in complexities has lowered to 1.6 bits from 2.5 bits.

The indeterminacy decays rapidly with increasing rule table radius, e.g. $\Xi(10) \approx 0.003$, but significant structures observed in FCT 18\222 are still not captured by larger radius SCA. Estimated SCA apparently never reduce to a CA. The indeterminacy may very well vanish at infinite radius. But the result then is a model of infinite size being reconstructed from a process that is only nearest neighbor with four local states: two internal and two observed. The practical problem of estimating the low-indeterminacy SCA is exacerbated by the exponential growth in the SCA model's size with radius. At a minimum, exponentially more data is required to maintain the estimated transition probabilities' statistical accuracy.

Now consider the remaining two examples, FCT 22\222 and FCT 90\222. The indeterminacy is much higher and therefore their data series are even less well modeled by CA. FCT 22\222 has a slowly vanishing indeterminacy. For example, on an $N = 500$ site lattice that started from an arbitrary pattern and that was allowed to relax for 10^4 iterations, an indeterminacy of

$\Xi(10) \approx 0.30$ bits was estimated over 10^6 iterations. FCT 90\222 has the largest indeterminacy found: $\Xi(10) \approx 0.39$ bits. This was estimated using the simulation parameters just quoted.

The conclusion from space-time diagrams, indeterminacy, entropy and complexity, is that even large radius SCA, let alone CA, do not capture the structures generated by FCT. Large indeterminacy at large radius suggests, erroneously, that the mechanism underlying the FCT data series has, at a minimum, a large spatial radius dynamic coupled to a stochastic process. Thus, even though the FCTs considered have radius one, to an observer the patterns are generated by a relatively nonlocal SCA: important structure appears in neighborhoods of 21 sites rather than just 3 sites. It is also noteworthy that when the estimated dynamic is chaotic even the smallest indeterminacy leads to significant prediction errors between the given and the simulated data series.

Indeterminacy

These examples serve to illustrate an indeterminacy classification of reconstruction scenarios that can confront an observer. Four reconstruction types can be distinguished for CA; they are illustrated in Figure 7.

1. **Determinate**: indeterminacy vanishes at finite rule table radius.[27] Elementary CA 18 is an example of this (cf. X's in Figure 7).
2. **Effectively determinate**: indeterminacy decays exponentially. FCT 18\222 is an example of this (cf. triangles in Figure 7). As shown in the preceding space-time diagrams, even in this scenario very low indeterminacy does not imply accurate reproduction.
3. **Asymptotically determinate**: indeterminacy vanishes at infinite radius. FCT 22\222 appears to be an example of this (cf. circles in Figure 7), if one allows extrapolation of the downward trend to infinite radius. Note that a monotonic decrease in indeterminacy indicates the existence of increasingly longer spatial correlations.
4. **Unreconstructible**: indeterminacy is finite at infinite radius. FCT 90\222 appears to be an example of this case (cf. squares in Figure 7). The observed patterns are unreconstructible with respect to CA, if one assumes the plateau from $r = 7$ to $r = 10$ continues on to infinite radius.

With respect to SCA, asymptotic determinacy and unreconstructibility were indicated for FCT 22\222 and FCT 90\222, respectively, since it was found that the minimal stochastic LUT grows without bound. Over the radius range investigated estimated SCA continued to reveal new variations in the conditional transition LUT $Pr(q|p)$ and in the neighborhood distribution $Pr(p)$. And so correlations operate over corresponding spatial scales. The data series never appeared to break into statistically independent regions which would have indicated that a minimal SCA with finite radius was reconstructed. A simple example of the latter is the ideal random spatial process (cf. diamonds in Figure 7) which is unreconstructible with respect to CA, but is reconstructible at radius 0 with respect to SCA.

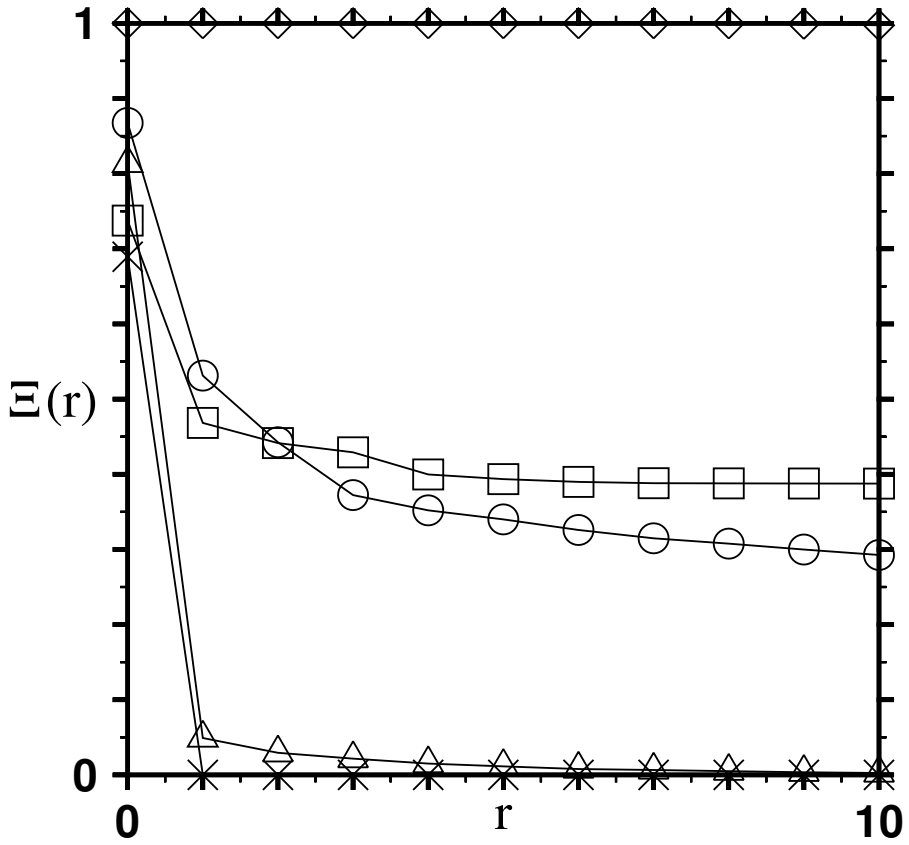

Figure 7 Indeterminacy $\Xi(r)$ versus template spatial radius r for five example systems: (i) A uniformly random stochastic cellular automaton: $Pr(q^i_{t+1}|p) = 2^{-2r-1}$ (diamonds along upper boundary); (ii) Elementary cellular automaton rule 18 (X's), cf. Figure 1; (iii) finitary cellular transducer 90\222 (squares), cf. Figure 4; (iv) finitary cellular transducer 22\222 (circles), cf. Figure 3; (v) finitary cellular transducer 18\222 (triangles), cf. Figure 2. $\Xi(r)$ estimated from patterns on an $N = 500$ site lattice over 10^6 iterations after 10^6 transient iterations from a random initial pattern. The same initial pattern was used in each case.

Concluding Remarks

These results illustrate the perils of working within an inadequate representation class for discrete-state spatial systems. Many nearest neighbor FCT with only a few states cannot be reconstructed as deterministic or stochastic CA. The overall argument here would be completed with the demonstration of a learning paradigm for estimating cellular transducers that, starting from the same data series, gives the expected and correct results for locally-finite discrete-state dynamics. In this case, the prediction error and indeterminacy would vanish, the given data could be exactly reproduced, and the observer would have inferred the hidden process. Although this must wait for the sequel, it can be noted that the alleged learning method for FCTs is a straightforward adaptation of ϵ-machine reconstruction to the context of space- and time-indexed measurements.[19]

The preceding discussion turned on a curious problem: locally-deterministic behavior viewed with a locally-deterministic instrument can appear random to all levels of approximation. What might the physical implications be? If local space-time states are obscured necessarily by the act of measurement, then microphysical reality would forever appear irreducibly uncertain. This would occur without invoking randomness; it would be a property of a purely deterministic world.

Information distortion during measurement could be due to some intrinsic nonlinearity of the measurement act on microphysical scales, analogous (say) to that seen in the unreconstructible FCT 90\222 or given by a measurement transducer more general than a LUT. Apparent randomness, even on the shortest time scales,[28] is consistent with underlying determinism. More to the point, a certain class of nonlinear interaction operating during measurement could underlie quantum indeterminacy. In a similar vein, the effective nonlocality seen in the slow vanishing of indeterminacy deserves further study.

Irreducible indeterminacy, as illustrated above, is consistent with internal deterministic dynamics, even though the latter may never be accessible, testable, or identifiable, using "reasonable" representations. Conversely, more sophisticated modeling techniques may be required for the discovery of internal structure than the estimation of local statistics. At the very least, the paradox considered here and these conjectures on its physical implications point to an important role that a measurement theory of nonlinear chaotic processes can play in basic physical theory. In concert with this, a systematic re-evaluation of how accepted model classes preclude the discovery of natural mechanisms appears necessary.

References and Notes

1. Impermanent address: Department of Physics, University of California, Berkeley, CA 94720. Internet address is chaos@gojira.berkeley.edu.
2. S. Wolfram. Cellular automata as models of complexity. *Nature*, 311:419, 1984.
3. S. Wolfram. *Theory and Applications of Cellular Automata*. World Scientific Publishers, Singapore, 1986.
4. T. Toffoli and N. Margolis. *Cellular Automata Machines: A New Environment for Modeling*. MIT Press, Cambridge, Massachusetts, 1987.
5. H. Gutowitz, editor. *Cellular Automata: Theory and Experiment*, New York, 1991. Bradford.
6. J. P. Crutchfield and B. S. McNamara. Equations of motion from a data series. *Complex Systems*, 1:417, 1987.
7. H. Chate and P. Manneville. Spatiotemporal intermittency in coupled map lattices. *Physica D*, 32:409, 1988.
8. T. F. Meyer, F. C. Richards, and N. H. Packard. A learning algorithm for the analysis of complex spatial data. *Phys. Rev. Lett.*, 63, 1989.
9. R. Livi and S. Ruffo. Probabilistic cellular automata models for a fluid experiment. In P. Coullet and P. Huerre, editors, *New Trends in Nonlinear Dynamics and Pattern-Forming Phenomena*, volume 237 of *NATO ASI Series B*, New York, 1990. Plenum Press.
10. F. Bagnoli, S. Ciliberto, R. Livi, and S. Ruffo. Phase transitions in convection experiments. In P. Manneville, N. Boccara, G. Y. Vishniac, and R. Bidaux, editors, *Cellular Automata and Modeling Complex Physical Systems*, volume 46 of *Springer Proceedings in Physics*, Berlin, 1990. Springer-Verlag.
11. J. P. Crutchfield and K. Kaneko. Phenomenology of spatio-temporal chaos. In Hao Bai-lin, editor, *Directions in Chaos*, page 272. World Scientific Publishers, Singapore, 1987.

12. N. H. Packard, J. P. Crutchfield, J. D. Farmer, and R. S. Shaw. Geometry from a time series. *Phys. Rev. Let.*, 45:712, 1980.
13. G. Mayer-Kress, editor. *Dimensions and Entropies in Chaotic Systems: Quantification of Complex Behavior*, New York, 1985. Springer-Verlag.
14. M. Casdagli and S. Eubank, editors. *Nonlinear Modeling and Forecasting*, volume XII of *Santa Fe Institute Studies in the Sciences of Complexity*, Reading, Massachusetts, 1992. Addison-Wesley.
15. F. Takens. Detecting strange attractors in fluid turbulence. In D. A. Rand and L. S. Young, editors, *Symposium on Dynamical Systems and Turbulence*, volume 898, page 366, Berlin, 1981. Springer-Verlag.
16. G. Sugihara and R. M. May. Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. *Nature*, 344:734, 1990.
17. M. Casdagli, S. Eubank, J. D. Farmer, and J. Gibson. State space reconstruction in the presence of noise. *Physica D*, 51:52, 1991.
18. J. Rissanen. *Stochastic Complexity in Statistical Inquiry*. World Scientific, Singapore, 1989.
19. J. P. Crutchfield and K. Young. Inferring statistical complexity. *Phys. Rev. Let.*, 63:105, 1989.
20. D. Zambella and P. Grassberger. Forecasting complexity in a class of simple models. *Complex Systems*, 2:269, 1988.
21. J. E. Hanson and J. P. Crutchfield. The attractor-basin portrait of a cellular automaton. *J. Stat. Phys.*, 66:1415, 1992.
22. L. S. Schulman and P. E. Seiden. Statistical mechanics of a dynamical system based on Conway's game of life. *J. Stat. Phys.*, 19:293, 1978.
23. The prediction error, defined as the per-site rate of error and estimated from $Pr(p)$ and $Pr(q|p)$, could be used.
24. Taking ϕ and ψ to be CA LUTs is a restriction of the class of FCTs imposed to simplify discussion here. In general, ϕ and ψ are finite state transducers: a proper superset of transducers based on LUTs, since LUTs cannot store past information.
25. The following considers only "pure-space" templates, i.e. those with sites at the previous time step. Note that under the assumption that perturbations propagate at some finite speed, there is a pure-space template representation of finite radius for a CA governed by a temporally deep neighborhood template.
26. Statistical agreement here includes reproducing the entropy and statistical complexity, along with various moments of neighborhood distribution $Pr(p)$.
27. Vanishing indeterminacy does not imply, interestingly enough, robust deterministic reconstruction. The exceptions, though, appear to be of zero probability.
28. This is not to be confused with the unpredictability appearing on moderate to long time scales that is due to deterministic chaos.

Acknowledgments: Thanks to Jim Hanson and Karl Young for helpful comments and to the Beckman Institute and its Center for Complex Systems Research for their generous hospitality. This work was partially funded under grant AFOSR 91-0293 and by a Visiting Research Professorship at the Beckman Institute, University of Illinois, Urbana-Champaign.