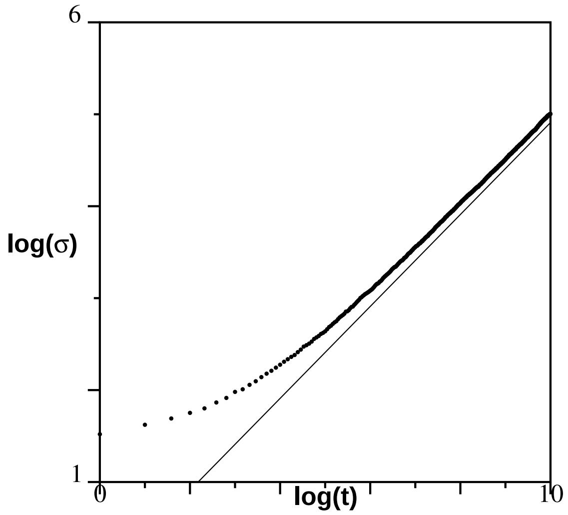

Fig. 19. Standard deviation of the dislocation histogram as a function of time. The measured slope of the asymptotic line is 0.501 ± 0.004 for $t > 100$. Logarithms are base 2. The straight line shown has slope $1/2$.

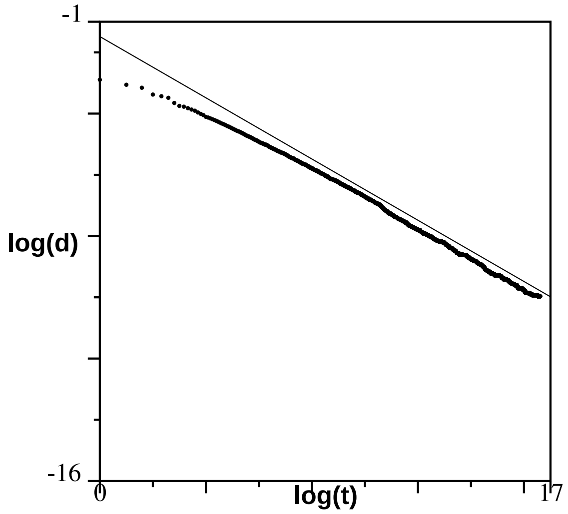

Fig. 20. Dislocation number density versus time on a periodic lattice of 10^5 sites. The initial pattern was random. The density is plotted at each of $100,000$ time steps. Again, logarithms are base 2. The straight line has slope $-1/2$.

The largest nonattracted fractions were found at lattice sizes N_i^* obeying the recursive rule $N_{i+1}^* = 2N_i^* + 2,\ N_1^* = 4$. Thus, although $f(N,t)$ has erratic, number-theoretic fluctuations, the fraction of nonattracted vicinity patterns decreases rapidly. We also found that for fully half of the lattice sizes the nonattracted patterns were a vanishing fraction. The second result is that, on a lattice of 1000 sites using 10,000 initial vicinity patterns, we found near-diffusive decay of $f\left(10^3, t\right) \propto t^{-\alpha}$ with $\alpha = 0.47 \pm 0.04$. Thus, when restricted to the vicinity, decay to the attractor also appears to be diffusive, with the fraction of nonattracted patterns decreasing monotonically in time and in lattice size. Similar results were found for odd lattices. Details will be reported elsewhere.

The figures and the estimated spreading, decay, and finite-size scaling rates give strong evidence that the dislocation trajectories are diffusive. Via the above results on domains, attractors, and vicinities, the diffusive behavior further supports the contention that the attractors Λ^0 and $\Lambda^{0,0}$ are the only statistically important attracting invariant sets of the system. As noted above, assuming that the dislocations are actually diffusive, the previous results on their attracting properties can be established in the statistical sense. That is, the fraction of initial configurations leading to diffusive motion of two or more dislocations that do not annihilate is vanishingly small. The Appendix gives an estimate for this fraction.

5.3 Computing with Dislocations

We have seen that space-time patterns generated by rule 18 are composed of domains Λ^0 separated by dislocations. The computational structure of a domain is extremely simple, since $\mathrm{M}(\Lambda^0)$ possesses only two recurrent states. In other words, the recurrent part of $\mathrm{M}(\Lambda^0)$ is capable of storing only 1 bit of information. This is clearly too simple to be harnessed for useful computation. If it exists, the useful computation must be performed by the dislocations at the domain boundaries.

We can make a quantitative estimate of the computational capability of rule 18 dislocations in the following way. After detecting dislocations with the transducer $\mathrm{T}_{\mathrm{domain}}(\Lambda^0)$, we recode the motion of a single trajectory as a symbol stream by assigning the symbol L to a move to the left and R to a move right. We then apply the machine reconstruction technique to this R-L data stream. The resulting machine, displayed in Fig. 21, shows an asymmetric structure: every R must be followed by an L. Noting that the topological entropy of this machine is $\log_2 \phi$, where $\phi = (1 + \sqrt{5})/2$ is the golden mean, we call it the golden mean machine M_{gm}. The asymmetry in M_{gm} is due to the asymmetry of the dislocation-recognizing transducer $\mathrm{T}_{\mathrm{domain}}(\Lambda^0)$ which identifies dislocations by a misplaced 1 in the spatial data set. As we noted above, other dislocation-recognizing algorithms can be used for which this constraint would not be present. Apart from this, however, we see that all trajectories consistent with this single constraint are allowed. Numerical estimates indicate a uniform branching probability from M_{gm}'s start state. Thus, one type of computation of which rule 18 is capable is the map from $\Lambda^{0,0}$ to $L(\mathrm{M}_{\mathrm{gm}})$. This is a very restricted class of computation, since it is a mapping between such simple regular languages.

In an early investigation of CA computational capability, Wolfram identified four classes of CA. He based his classification on visual examination of the space-time patterns they generate. Class 4, which he conjectured to be the most computationally capable, was defined by the presence of spatially isolated transients of many shapes evolving for arbitrarily long times. The

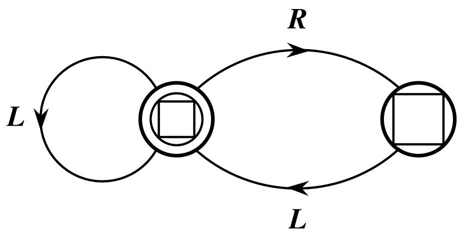

Fig. 21. The golden mean machine M_{gm} reconstructed from the space-time trajectory of a single dislocation. The symbol L denotes a dislocation move to the left; R a move right. Its entropy is $\log_2 \phi$, where $\phi = \left(1 + \sqrt{5}\right)/2$ is the golden mean.

conjecture was that, using these propagating "blobs," a model of universal serial computation could be constructed.[1]

A simple recoding of rule 18, however, produces space-time patterns satisfying the criteria of class 4. This recoding, elsewhere called the "blocking transformation",[1] is a special type of a decimation transducer. Recall that in the blocking transformation every other cell, for example, in space and time is excised. The new pattern represents the evolution of the recoded initial condition under a different CA rule, typically one with a larger neighborhood size. General blocking transformations replace a fixed length block of cells with a single symbol. Blocking transformations, though, are neither as versatile as transducers nor do they exploit a CA's qualitative dynamics.

Visually, blocked space-time patterns of arbitrary states evolving under rule 18 display spatially isolated patches of fluctuating cells moving against an invariant background of 0's. It is the spatial period 2 nature of Λ^0, however, that is responsible for this. When a domain is in phase with the blocking scheme's internal state, arbitrary spatial patterns will be observed. When it is out of phase, all the nonzero sites will be removed. The resulting blocked region is a homogeneous patch of 0's. Thus, the structures in the blocked pattern are caused by the wall dynamics. They are misinterpreted by the blocking transformation since the domains alternate between being in-phase and out-of-phase with the blocking. The evolution of the boundaries of an in-phase domain governs the evolution of the blocked patch between them. When the boundaries meet, they annihilate and the "spatially isolated transient" disappears.

Based on the appearance of propagating blobs under a simple recoding, one might be tempted to argue that rule 18 should be placed in class 4. However, its computational capability appears quite limited, being dominated by diffusive behavior of the dislocations. The simplicity of M_{gm} makes it doubtful that rule 18 can be harnessed to perform useful computations. The misleading "transients" in the blocked patterns are an artifact introduced by the blocking transformation. They occur because the transducer used in the blocking was unsuited to the underlying domain structure.

The ambiguities of Wolfram's classification suggest a reevaluation of how computation is identified in CA. The use of transducers in concert with machine reconstruction[26,41] gives a principled approach to analyzing a CA's intrinsic computational capability. The first step employs machine reconstruction to search for computational structures such as domains. When found, these are factored out using domain transducers similar to $\mathrm{T}_{\mathrm{decimate}}(\Lambda^0)$. The resulting encoded pattern then can be searched for higher-order structures, again using machine reconstruction. In this way, one builds a hierarchy of computational models, with structures in each level describing computational properties missed at the previous level. The hierarchy as a whole describes a CA's information processing architecture. The hierarchy's depth in this view is a simple measure of CA's global computational capacity.*

For intrinsic computation to be useful, there are the additional requirements that individual computations be robust and that a compiler exist for a class of nontrivial tasks. The first means that, at the very least, small perturbations do not change the computational class of the temporal evolution. This in turn requires an understanding of the global organization the CA induces in regular language space; that is, of the attractor-basin portrait. The second requirement, a compiler, determines how a computational task is mapped onto the CA's dynamics. Nontrivial tasks can be supported either by the CA's evolution toward attractors or by propagating wall structures. A dislocation compiler for rule 18, for example, is a mapping of a computational task specification T to an initial condition $\mathrm{s}_0 \in \Lambda^{0,0}$ such that the time series $\mathrm{s}_0, \mathrm{s}_1, \mathrm{s}_2, \ldots, \mathrm{s}_t, \ldots$ executes the task via the $R - L$ movement of wall and domain structures. The output is a string in $L(\mathrm{M}_{\mathrm{gm}})$. Since rule 18's computation appears quite limited, we must leave the question of useful intrinsic computation to a future analysis of more complex CA.

Although formulated in the present context of spatially-extended systems, the general questions raised concerning the design and existence of compilers for CA are germane to parallel programming and distributed processing. Should such compilers exist, then analyses along the lines developed here would indicate where useful computation can lie in high-dimensional state space. They also would give a global view of the geometry underlying the computation. The practical use of this appears to be in the analysis of errors and stability of a given CA implementation.

6 Concluding Remarks

Domains and dislocations are basic dynamical paradigms. In physics dislocations are observed as crystal defects and elementary particles, in chemistry they are seen in polymer chains and chemical waves, and in biology they are found in cellular membrane development and have been posited to be the information processing foundation of prebiotic evolution[42] (an additional brief list is given in [43]). These structures often result from lateral inhibition coupling between processing elements, to use the neurophysiological term. This same coupling is responsible for the ability of artificial neural networks to learn, although in these high-dimensional

*A related reconstruction hierarchy for the computational classification of continuum-state dynamical systems is presented elsewhere.[26,27]

systems their nonspatial architecture makes the domain and dislocation structures substantially more difficult to visualize. We have avoided this important difficulty by focusing on spatially-extended systems. The attempt in the preceding was to set up a constructive framework to study structures supporting complex information processing in high dimensional state spaces.

Let us recapitulate our results. First, we developed a structural analysis of CA based on elementary notions from computation theory. Second, this allowed us to describe the attractor-basin portrait and subbasin-portal structure of a specific CA. Third, along the way, we introduced specific tools for observing and manipulating CA structures: the various transducers or, more formally, endomorphisms of Sofic systems. Finally, as an application, these were used in a statistical analysis of dislocation trajectories within specific subspaces of the state space. The results, along with the attractor-subbasin-portal portrait, strongly suggest that there are only two statistically significant attracting invariant sets. A rigorous proof of this remains an open problem, though it appears to be within reach.

Within a larger arena, we have addressed one aspect of the general problem of recognizing structures in space-time. Consider, as another example, a soliton wave in a shallow-water channel. Typically, the recognition of a propagating structure is done intuitively via visual inspection. In some cases, such as with solitons, this identification eventually develops a formal expression. The propagating entity has a mathematically defined shape expressed in closed form. With this, one can search for those structures by convolving the space-time pattern with the soliton kernel. In a literal sense, the shapes form a basis for the space of recognizable patterns. The method of transducer machine recognition introduced above gives a different, but complementary, computational approach to space-time pattern recognition, as we have just described it, that is both general and reconstructive. It, along with machine reconstruction, suggests an automated method of identifying significant structures and tracking them in space-time. Although the present approach starts with discretized data, we note that rather than being a liability, this is also the domain of many important problems in time series and image analysis and processing. The authors are well aware of the need to extend the present techniques to continuum-state spatially-extended systems.

During the attractor-basin analysis, we referred to another general problem, concerning transients in high dimensions.[40,44,45] Rule 18 is one nontrivial, i.e. nonlinear, space-time system that apparently can be analyzed completely in this way. The key elements in this are the global linearization techniques of the intercalation operator,[36] the dynamical homogeneity of regular domains that we have introduced here, and the organization of subbasins and portals.[40] The analysis of domain-wall dominated CA hints at similar analyses for a wider class of spatially-extended systems, such as map lattices and soliton-bearing PDEs.

Implicit in our investigation of subspaces of initial configurations, represented by taking Ω_0 to be various regular process languages other than $\mathcal{A}^*$, is a set of techniques to study the geometry of high-dimensional spatially-extended systems. This was motivated, in fact, by techniques used in earlier work on continuous-state lattice dynamical systems[46] and their extremely long transients.[40,44] And these too seem applicable to a wide range of similar problems posed by high-dimensional dynamical systems.

In order to illustrate the basic ideas, we have used the example of rule 18. The question naturally arises of whether the analysis generalizes to other CA. The answer is yes. The additional rules can be seen by simply examining the space-time diagrams of other CA. Due to the reconstructive approach we have taken and the computation-theoretic formulation of the dynamical systems concepts, the interesting possibility arises that the analysis for these other CA can be automated. We have alluded to the way in which an automated machine reconstruction "expert system" could be used as an aid to analysis. The goal is for it to help with automated proofs of language invariance and attraction, with the derivation of basins and vicinities, and in the investigation of the computational structure of basins and separatrices.

The analysis we have given is not complete, though our and others' results strongly suggest the possibility that it is nearly so. Rather it outlines a range of problems and a constructive computational approach that integrates qualitative dynamics. We are left, therefore, with several future questions. We close by mentioning a few of the most pressing ones. Of what complexity are the basins of the two regular attracting domains and the null language 0^*? How about the separatrix? The near completeness of the analysis suggests that there should be a rigorous proof that the deterministic rule 18 is diffusive. This will require an exact analysis of the short time behavior and an analogous development of approximation of stochastic process languages.[39] The relationship between the intercalation linearization and the subbasin-portal picture is important to fully understand the global structure of the state space. Finally, how can we compute with dislocations? Does a dislocation compiler exist? If not, why is this structure not usable to perform computations?

7 Acknowledgments

The authors thank Peter Grassberger and Erica Jen for helpful discussions. Visits to the Center for Nonlinear Studies and the Theoretical Division at the Los Alamos National Laboratory are warmly acknowledged. They are also grateful to Professor Carson Jeffries for his continuing support. This work was funded in part by ONR contract N00014-90-J-1774, NASA-Ames University Interchange NCA2–488, and a Cray Research, Inc., University Research and Development grant. It was completed at the Santa Fe Institute where JPC was supported by a Robert Maxwell Foundation Visiting Professorship.

Appendix: Dislocation Decay Statistics

We estimate for an infinite lattice the probability that randomly diffusing dislocations have not annihilated by some time t. This also gives an estimate of the fraction, or basin measure, of those initial patterns which have decayed by t.

Consider the effect of flipping a single cell in each state of an ensemble consisting of a uniform distribution on Λ^0. We assume that the perturbation occurs in the region of compact support, i.e., there are nonzero cells on both sides of it. We count dislocations using the machine metric $\| \cdot \|_{\mathrm{M}(\Lambda^0)}$ defined above. With probability $1/2$, the cell flipped will be a wild-card, so the perturbation will not generate any dislocations. Otherwise, it will produce two dislocations a distance d apart, with $d-1$ equal to the number of adjacent 0's to the right of the perturbed cell. The uniform distribution of initial conditions implies that the wild-card cells are independently, identically distributed with $p\big(\sigma^i = 0\big) = 1/2$. This in turn implies that the probability that the dislocations are exactly d cells apart is given by

$$p(d) = \begin{cases} 2^{-(d+1)/2} & d \text{ odd} \\ 0 & d \text{ even} \end{cases}$$

Thus, the probability that two dislocations are d_t or fewer cells apart is

$$p(d \leq d_t) = \sum_{d=1}^{d_t} p(d) = 1 - 2^{-(d_t+1)/2}$$

for d_t odd.

Two dislocations will typically collide when their plumes overlap; that is, when the standard deviation $\sigma \approx d$. As noted in the main text, the standard deviation of the plume gives a time-dependence of $\sigma \approx 8\pi D t^{-1/2}$ with $D \approx 1/2$.

The probability that two dislocations have collided after t steps is the probability that their initial separation was σ or less, which is given by

$$p(d \leq \sigma) = 1 - 2^{-\left(8\pi D\sqrt{t}+1\right)/2}$$

For long times, the probability that the dislocations have not yet collided is therefore

$$p_{\mathrm{survive}}(t) \approx 2^{-2\pi\sqrt{t}}$$

This gives an estimate of the decay in the number of initial conditions that have not yet fallen onto the attractor, assuming diffusive motion. It clearly vanishes in the limit $t \to \infty$.

Bibliography

[1] S. Wolfram. *Theory and Applications of Cellular Automata*. World Scientific Publishers, Singapore, 1986.

[2] A. R. Smith. Simple computation-universal cellular spaces. *J. ACM*, 18:339, 1971.

[3] T. Toffoli. *Cellular Automata Mechanics*. PhD thesis, University of Michigan, Michigan, 1977. Tech. Rep. 208, Comp. Comm. Sci. Dept.

[4] S. Wolfram. Computation theory of cellular automata. *Comm. Math. Phys.*, 96:15, 1984.

[5] L. Hurd. Formal language characterizations of cellular automaton limit sets. *Complex Systems*, 1:69, 1987.

[6] E. R. Berlekamp, J. H. Conway, and R. K. Guy. *Winning Ways for your Mathematical Plays*, volume 2. Academic Press, New York, 1984.

[7] T. Toffoli and N. Margolis. *Cellular Automata Machines: A New Environment for Modeling*. MIT Press, Cambridge, Massachusetts, 1987.

[8] O. Martin, A. Odlyzko, and S. Wolfram. Algebraic properties of cellular automata. *Commun. Math. Phys.*, 93:219, 1984.

[9] K. Kaneko. Attractors, basin structures and information processing in cellular automata. In S. Wolfram, editor, *Theory and Applications of Cellular Automata*. World Scientific, Singapore, 1986.

[10] H. Ito. Intriguing properties of global structure in some classes of finite cellular automata. *Physica*, 31D:318, 1988.

[11] E. Jen. Cylindrical cellular automata. *Comm. Math. Phys.*, 118:569, 1990.

[12] S. Wolfram. Statistical mechanics of cellular automata. *Rev. Mod. Phys.*, 55:601, 1983.

[13] H. A. Gutowitz, J. D. Victor, and B. W. Knight. Local structure theory for cellular automata. *Physica*, 28D:18, 1987.

[14] W. Li, N. H. Packard, and C. G. Langton. Transition phenomena in cellular automata rule space. *Physica*, 45D:77, 1990.

[15] N. Boccara, J. Nasser, and M. Roger. Particle-like structures and their interactions in spatio-temporal patterns generated by one-dimensional deterministic cellular automaton rules. *Phys. Rev. A*, 44:866, 1991.

[16] J. Park, K. Steiglitz, and W. Thurston. Soliton-like behavior in automata. *Physica*, 19D:423, 1986.

[17] P. Grassberger. New mechanism for deterministic diffusion. *Phys. Rev. A*, 28:3666, 1983.

[18] C. G. Langton. Computation at the edge of chaos: Phase transitions and emergent computation. In S. Forrest, editor, *Emergent Computation*, page 12. North-Holland, Amsterdam, 1990.

[19]K. Lindgren. Correlations and random information in cellular automata. *Complex Systems*, 1:529, 1987.

[20]D. R. J. Chillingworth. *Differential Topology with a View to Applications*. Pitman, London, 1976.

[21]J. Guckenheimer and P. Holmes. *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*. Springer-Verlag, New York, 1983.

[22]J. P. Crutchfield and J. E. Hanson. Computational mechanics of cellular automata: Space and time machines. in preparation, 1991.

[23]P. Grassberger. Chaos and diffusion in deterministic cellular automata. *Physica D*, 10:52, 1984.

[24]D. Griffeath. *Additive and Cancellative Interacting Particle Systems*, volume 724 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1979.

[25]J. E. Hopcroft and J. D. Ullman. *Introduction to Automata Theory, Languages, and Computation*. Addison-Wesley, Reading, 1979.

[26]J. P. Crutchfield and K. Young. Computation at the onset of chaos. In W. Zurek, editor, *Entropy, Complexity, and the Physics of Information*, volume VIII of *SFI Studies in the Sciences of Complexity*, page 223, Reading, Massachusetts, 1990. Addison-Wesley.

[27]J. P. Crutchfield. Reconstructing language hierarchies. In H. A. Atmanspracher and H. Scheingraber, editors, *Information Dynamics*, page 45, New York, 1991. Plenum.

[28]R. E. Blahut. *Principles and Practice of Information Theory*. Addison-Wesley, Reading, Massachusetts, 1987.

[29]D. A. Lind. Applications of ergodic theory and sofic systems to cellular automata. *Physica*, 10D:36, 1984.

[30]M. G. Nordahl. Formal languages and finite cellular automata. *Complex Systems*, 3:63, 1989.

[31]L. P. Hurd. Appendix. table 10. regular language complexities. In S. Wolfram, editor, *Theory and Applications of Cellular Automata*. World Scientific, Singapore, 1986.

[32]N. H. Packard. Complexity in growing patterns in cellular automata. In J. Demongeot, E. Goles, and M. Tchuente, editors, *Dynamical Behavior of Automata: Theory and Applications*. Academic Press, 1984.

[33]S. Wolfram. Universality and complexity in cellular automata. *Physica*, 10D:1, 1984.

[34]F. Harary. *Graph Theory*. Addison-Wesley, Reading, Massachusetts, 1969.

[35]J. P. Crutchfield and J. E. Hanson. Computational mechanics of cellular automata: Space-time machines. in preparation, 1991.

[36]E. Jen. Aperiodicity in one-dimensional cellular automata. *Physica*, 44D:121, 1990.

[37]E. Jen. Exact solvability and quasiperiodicity of one-dimensional cellular automata. *Nonlinearity*, 4:251, 1990.

[38]K. Eloranta and E. Nummelin. The kink of cellular automaton rule 18 performs a random walk. Technical Report A296, Helsinki University of Technology, Institute of Mathematics, 1991.

[39]K. Lindgren and M. G. Nordahl. Complexity measures and cellular automata. *Complex Systems*, 2:409, 1988.

[40]J. P. Crutchfield. Subbasins, portals, and mazes: Transients in high dimensions. *J. Nucl. Phys. B*, 5A:287, 1988.

[41]J. P. Crutchfield and K. Young. Inferring statistical complexity. *Phys. Rev. Let.*, 63:105, 1989.

[42]A. G. Cairns-Smith. *Genetic takeover and the mineral origins of life*. Cambridge University Press, New York, 1982.

[43]J. P. Crutchfield. Spatio-temporal complexity in nonlinear image processing. *IEEE Trans. Circ. Sys.*, 37:770, 1988.

[44]J. P. Crutchfield and K. Kaneko. Are attractors relevant to turbulence? *Phys. Rev. Let.*, 60:2715, 1988.

[45]J. P. Crutchfield and K. Kaneko. Transients in high dimensions. in preparation, 1991.

[46]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.