Role of Irreversibility in Stabilizing Complex and Nonergodic Behavior in Locally Interacting Discrete Systems

Charles H. Bennett[(a)] and G. Grinstein
IBM T. J. Watson Research Laboratory, Yorktown Heights, New York 10598
(Received 18 March 1985)

Irreversibility stabilizes certain locally interacting discrete systems against the nucleation and growth of a most-stable phase, thereby enabling them to behave in a computationally complex and nonergodic manner over a set of positive measure in the parameter space of their local transition probabilities, unlike analogous reversible systems.

PACS numbers: 05.90.+m

The dynamics of a statistical-mechanical system in contact with a larger environment is often modeled as a random walk on the system's state space (e.g., kinetic Ising model). If the environment is at equilibrium, this random walk will be "microscopically reversible" (its matrix of transition probabilities being of the form DSD^{-1}, where D is diagonal and S symmetric), and the system's stationary distribution of states will be an equilibrium one (e.g., canonical ensemble) defined by a Hamiltonian simply related to the transition probabilities. On the other hand, if the environment is not at equilibrium, the system's transition matrix in general will be irreversible, and the resulting nonequilibrium stationary distribution may be very hard to characterize.

Though an irreversible system's distribution of states is not simply related to the transition probabilities, its distribution of histories is. More specifically, the stationary distribution of histories for any stochastic model, whether reversible or not, may be viewed as a canonical distribution under an effective Hamiltonian on the space of histories, in which each configuration interacts with its predecessor in time with an "interaction energy," equal simply to the logarithm of the corresponding transition probability.

In particular, we consider the case in which the underlying stochastic model is a probabilistic cellular automaton (CA), in other words, a d-dimensional lattice with finitely many states per site, in which each site, at each discrete time step, undergoes a transition depending probabilistically on the states of its neighbors. In this case[1,2] the stationary distribution of histories of the CA is equivalent to the equilibrium statistics of a corresponding generalized Ising model (GIM) in $d+1$ dimensions. This appears paradoxical, because CA's are known to be capable of complex, nonergodic behavior even when all local transition probabilities are positive,[3,4] whereas the behavior of GIM is generally simple and ergodic (a stochastic process is "ergodic" if its stationary distribution is unique). For example, a standard kinetic Ising model, at a generic point in its temperature-magnetic field parameter space, undergoes nucleation and growth of a unique most-stable phase, thereby relaxing to a stationary distribution independent of the initial conditions.

Here we note the resolution of the paradox, and illustrate an essential difference between reversible and irreversible systems by characterizing the phase diagram, equation of state, domain growth kinetics, and equivalent $(d+1)$-dimensional GIM of one of the simplest nonergodic irreversible CA, viz., Toom's north-east-center (NEC) voting model, The resolution of the paradox lies in the fact that when a d-dimensional CA is represented as a $(d+1)$-dimensional GIM, the parameters (coupling constants) of the latter system are not all independent, but are constrained in such a way as to cause the free energy of the $(d+1)$-dimensional system to be identically zero, no matter how the parameters (transition probabilities) of the underlying CA are varied. It is therefore possible for irreversible systems such as the NEC model to be nonergodic, and in particular to have two or more stable phases, over a finite region in their phase diagrams, whereas reversible systems can exhibit this behavior only over a subset of zero measure consisting of points in the phase diagram where two or more phases, by symmetry or accident, have exactly equal free energy.

North-east-center model, and reasons for its nonergodicity.—The NEC model is one of a class of voting rules for CA shown by Toom[3] to be nonergodic in the presence of small but arbitrary probabilistic perturbations. The model consists of a square lattice of spins, each of which may be up or down. The spins are updated synchronously, with a spin's future state decided by majority vote of the spins in an unsymmetric neighborhood, consisting of the spin itself and its northern and eastern neighbors. The rule just described is deterministic; we consider two-parameter noisy perturbations of the rule, in which a spin whose present neighborhood majority is up, instead of going up with certainty at the next time step, goes up with probability $1-p$ and down with probability p; and a spin whose present neighborhood majority is down goes down with probability $1-q$ and up with probability q. Alternatively, the noise may be characterized by its "amplitude" $p+q$, analogous to temperature, and its "bias" $(p-q)/(p+q)$, analogous to magnetic field.

Because probabilistic CA typified by the NEC model are not microscopically reversible, not all the formalism of equilibrium statistical mechanics can be applied to them. In particular, the stationary probability measure $\mu(X)$ on configuration space cannot in general be represented as the Boltzmann exponential of any locally additive potential. However, a stable phase of a probabilistic CA can still be defined in the thermodynamic limit as a probability measure μ on the space of configurations of the infinite lattice, that is (1) stationary under the transition rule, and (2) extremal in the sense of not being expressible as a linear combination of other stationary measures. This is analogous to the definition of a phase for Hamiltonian systems as a measure μ which is extremal among measures that can be obtained as the thermodynamic limit of the canonical ensemble under various boundary conditions.[5]

When p and q are small and positive, no local transition is entirely forbidden; hence the model is ergodic on any finite lattice (e.g., an N-by-N torus). However, Toom showed[3] that for sufficiently small p and q, the transition rate between the mostly up and mostly down states of the entire system tends to zero with increasing N, rendering the infinite system nonergodic, with two stable phases, like a conventional Ising model below its critical temperature.

When the noise is unbiased ($p=q$), the NEC system behaves like an Ising model in zero field: There is a critical noise amplitude at which the spontaneous magnetization vanishes continuously. Where the NEC system differs from equilibrium systems is in its response to biased noise, e.g., $0<q<p<\frac{1}{2}$. In an equilibrium system, such a symmetry-breaking perturbation (analogous to magnetic field) would cause the system to become ergodic, by rendering it susceptible to nucleation and growth of the more stable phase. The NEC system, on the other hand, remains nonergodic, with two stable phases, even in the presence of biased noise, if the noise amplitude is small enough. This difference in response to a symmetry-breaking perturbation can be understood by comparing the mechanisms by which the two systems suppress fluctuations of the minority phase.

In a zero-field Ising system below its critical point, as in any reversible system at a first-order phase transition, a flat interface between the two equally stable phases must have zero mean propagation velocity. Finite islands of the minority phase nevertheless shrink because of surface tension, an island of radius r shrinking with velocity $-dr/dt$ proportional to $1/r$. The NEC system's irreversibility, by contrast, allows a flat interface to drift even when the noise is unbiased ($p=q$), and this drift velocity depends on interface orientation in such a way as to cause islands of either phase to shrink at a rate $-dr/dt$ roughly independent of their radius. In both systems, a small symmetry-breaking perturbation s adds a constant term $\propto s$ to dr/dt for islands of the favored phase. This is sufficient to render the Ising system metastable, by favoring growth of islands larger than a certain critical radius $\propto 1/s$; but in the NEC system, dr/dt remains negative for all r, and so the system remains stable.

The NEC system's interface motions are easiest to understand in the case of zero noise. Here a 135° diagonal interface between up and down spins drifts southwestward with unit speed, regardless of which phase is on which side of the interface, because sites just southwest of the interface, at each instant of time, have neighborhood majorities dominated by their north and east neighbors on the other side. On the other hand, a vertical or horizontal interface in the same system does not drift.

These motions enable the noiseless NEC system to eliminate islands of either phase with a linear shrinkage velocity $-dr/dt$ independent of their size r. To see this, consider an island of, say, up spins of arbitrary size and shape. Let an isosceles right triangle, with the hypotenuse on the northeast, be circumscribed about the island, and let all other spins in this triangle also be flipped up, so that we are considering a somewhat larger island of up spins of a particular triangular shape. Because the NEC rule is monotonic, the addition of further up spins to an island of up spins cannot hasten its disappearance. Therefore, the lifetime of the circumscribed triangular island is an upper bound on the lifetime of the original arbitrary-shaped island. The fate of the triangular island in the noiseless NEC system is quite simple: Its southern and western borders remain fixed, while its northeastern border closes in with unit velocity, eliminating the island in time proportional to its original size r.

The same argument holds in the presence of noise, whether biased or unbiased, because, if the noise amplitude is small enough, each of the interface velocities will differ only slightly (linearly in p and q) from its value in the noiseless system. Under these conditions, an island of size r of either phase will disappear in time proportional to r by differential motion of its borders, the lifetime being longer for one phase than the other if the noise is biased.

Phase diagram of the NEC system.—Numerical studies were performed on the NEC model with use of CAM, a fast CA simulator.[6] Besides providing quantitative data, CAM's real-time display was most helpful in assessing qualitative features of the model.

The phase diagram shown in Fig. 1 was obtained by finding pairs of noise parameters (p,q) such that a large minority island of up to (25 000 sites out of total of $65\,536=256\times256$ sites) neither grew nor shrank on average, during runs of about 50 000 time steps. The two-phase region is bounded by a pair of first-order transitions (solid curves terminating at the critical point $p=q=0.90\pm0.003$), on which one phase becomes marginally stable, losing its ability to eliminate large islands of the other. Beyond the two-phase region is a narrow metastable zone (demarcated on the right by dashed lines), in which large islands of the favored phase grow but small islands shrink. A critical exponent of 3.0 ± 0.4 was found for the vertical width of the two-phase region as a function of noise amplitude $(p+q)$ below the critical point. Other runs on one-phase systems with unbiased noise $p=q$ yielded the value $\beta = 0.122 \pm 0.01$ for the exponent describing magnetization as a function of noise amplitude below the critical point.

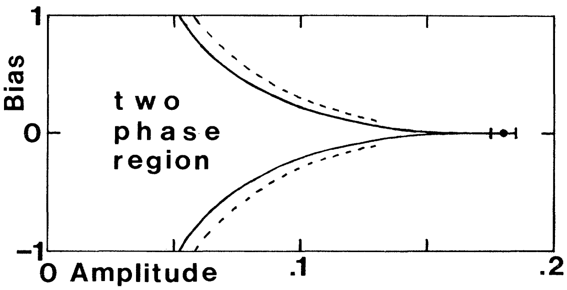

FIG. 1. Phase diagram of the NEC system, for noise parameters p and q, with amplitude $= p+q$ and bias $=(p-q)/(p+q)$.

Besides being irreversible, the NEC model differs from conventional kinetic Ising models in having synchronous updating. However, preliminary runs in which only a fraction ($\frac{1}{2}$ to $\frac{1}{16}$) of the spins are updated at each time step indicate that even a fully asynchronous NEC model would have a qualitatively similar phase diagram. We also explored an analytically solvable mean-field approximation[7] to the NEC rule, based on the recurrence relation

$$R(m) = -1 + [p(1-m)^3 + 3p(1-m)^2(1+m) + 3(1-q)(1-m)(1+m)^2 + (1-q)(1+m)^3]/4,$$

where $R(m)$ is the magnetization at time $t+1$ as a function of that at time t. Here, too, the phase diagram was similar, except that there was no metastable zone, and the critical exponents were $\frac{3}{2}$ (for the two-phase region width) and $\frac{1}{2}$ (for β).

The equivalent (d+1)-dimensional Hamiltonian, and its free energy.—We now review the construction of an equivalent $(d+1)$-dimensional Hamiltonian model for an arbitrary (in general irreversible) d-dimensional CA,[1,2] and show why the former can have multiple stable phases over a set of finite measure in the parameter space of the latter. The possible time histories $X(0), X(1), \ldots, X(t)$ of a d-dimensional CA can be viewed as configurations of a $(d+1)$-dimensional lattice with one boundary fixed at $X(0)$, the initial state of the CA. The probability $P(X(0), X(1), \ldots, X(t))$ of such a history may be expressed as the product

$$P(X(1)/X(0))P(X(2)/X(1)) \cdots P(X(t)/X(t-1)),$$

where $P(X(i+1)/X(i))$ denotes the conditional probability for the CA to be in state $X(i+1)$ at time $i+1$ given that it was in state $X(i)$ at time i.

By defining

$$H(X(i+1), X(i)) = -\ln[P(X(i+1)/X(i))],$$

we cast the history probability in the familiar form of a Boltzmann factor (taking $kT=1$):

$$P(X(0), X(1), \ldots, X(t)) = \exp\left(-\sum_{i=0}^{t-1} H(X(i+1), X(i))\right),$$

with H playing the role of an effective Hamiltonian coupling adjacent d-dimensional time slices. All properties of the CA can thus be expressed as canonical-ensemble averages of the $(d+1)$-dimensional system defined by Hamiltonian H. The $(d+1)$-dimensional model has the remarkable feature[8] that its free energy is identically zero regardless of the CA's initial condition or transition probabilities. This follows from the normalization of these probabilities [for each $X(i)$, the sum over $X(i+1)$ of $P(X(i+1)/X(i))$ must be 1], which in turn implies that the $(d+1)$-dimensional partition function is 1. In the thermodynamic limit any stable phase of the d-dimensional system is a stable phase of the $(d+1)$-dimensional system; therefore, if a d-dimensional CA has multiple stable phases, its corresponding $(d+1)$-dimensional Hamiltonian model will also, all with zero free energy.

The preceding argument holds whether $X(0)$, $X(1)$, etc., represent time steps of a synchronous model, or discrete-time snapshots of an asynchronous (master equation) model evolving in continuous time. However, in the synchronous case, because transitions at all sites in the same time slice are independent, the equivalent Hamiltonian is of a generalized Ising form, being a sum of local terms $H(y,\underline{x}) = -\ln P(y/\underline{x})$, where $P(y/\underline{x})$ is the conditional probability that a site will be in state y at time $t+1$, given that its neighborhood was in state $\underline{x}$ at time t. The normalization constraint, viz., that for each $\underline{x}$, the sum over y of $P(y/\underline{x})$ be 1, restricts the GIM to a lower-dimensional surface in the parameter space of its coupling constants, on which the free energy is zero.

In the case of metastable phases the $(d+1)$-dimensional free energy per space-time site is not 0 but $-\ln(1-\Gamma) \approx \Gamma$, where Γ is the nucleation rate for transitions out of the phase. This result, whose derivation will be given elsewhere, may be explained heuristically by saying that the expected number of futures per present is 1 in a stable phase and $1-\Gamma$ in a metastable phase, if we exclude futures not belonging to the phase. We can now summarize the unusual behavior of the $(d+1)$-dimensional free energy of a nonergodic irreversible system such as the Toom NEC model. Throughout the two-phase region of parameter space, the free energy of both phases is identically zero. Beyond this region, the free energy of the stable phase remains zero, while the other phase's free energy lifts off very smoothly from zero, as $\Gamma \approx \exp(-\text{const}/s^{d-1})$, where s (>0), is the distance in parameter space from the first-order phase boundary.

Irreversibility and complexity.—The relation of irreversibility to complexity has been extensively studied, especially in chemical reaction-diffusion systems.[9] The "dissipative structures" developed by such systems far from equilibrium exhibit macroscopic space-time ordering which persists over a set of positive measure in parameter space, but because the local interaction lacks the spatial asymmetry of the NEC rule, these systems remain ergodic in the thermodynamic limit. In other words, for a generic choice of parameters, one dissipative structure is stable, and the others are metastable. A closer chemical analog to the NEC system's generic nonergodicity can be seen in the "once for ever" selection exhibited by stirred (i.e., mean-field) nonlinear autocatalytic reaction systems.[10]

Probably the most comprehensive kind of complexity of which cellular automata or other discrete systems are capable is the capacity for universal computation. A computationally universal system[11] is one that can be programmed, through its initial conditions, to simulate any digital computation. Computational universality of course can only occur in a nonergodic system: In a computationally universal system, not only does the indefinite future depend on the initial condition, but it does so in an arbitrarily programmable way. For example, the computational universality of the well-known deterministic CA rule "life" implies that one can find an initial configuration for it that will evolve so as to turn a certain site on if and only if white has a winning strategy at chess. Universal automata can be programmed to mimic arbitrary kinds of nontrivial behavior observed in other systems, e.g., the scale invariance of the Ising model at its critical point.

The property of computational universality was originally demonstrated for systems (e.g., Turing machines, deterministic cellular automata) rather unlike those ordinarily studied in mechanics and statistical mechanics. Later, the property was demonstrated for certain noiseless classical mechanical systems[12] such as hard spheres with appropriate initial and boundary conditions. Very recently,[4] the computational universality has been shown to hold for certain noisy, locally interacting systems, i.e., irreversible CA in which all local transition probabilities are positive. Notable among these is the three-dimensional CA of Gacs and Reif,[13] which uses the NEC rule in two of its dimensions to correct errors in an arbitrary computation being performed along the third dimension.

Though all universal automata are equivalent in the computations they can perform, there may still be qualitative differences among them in the density of initial states that lead to nontrivial computations. These differences need to be better understood before one can make the tempting assertion that "self-organization" like that observed in nature is a spontaneous tendency of locally interacting irreversible systems, on a set of positive measure in the space of their transition probabilities and initial conditions. Characterizing the generic behavior of homogeneous locally interacting systems capable of universal computation is, we believe, the central problem in what might be called the new field of discrete computational statistical mechanics.

We wish to thank Peter Gacs for helpful discussions during all phases of this work, especially for bringing Toom's models to our attention and emphasizing the robust nature of their nonergodicity, compared to that of the Ising models with which we were familiar.

---

[a]Present address: Boston University, 111 Cummington Street, Boston, Mass. 02215.

[1]I. G. Enting, J. Phys. C **10**, 1379 (1977).

[2]E. Domany and W. Kinzel, Phys. Rev. Lett. **53**, 311 (1984).

[3]A. L. Toom, in *Multicomponent Random Systems,* edited by R. L. Dobrushin, in Advances in Probability, Vol. 6 (Dekker, New York, 1980), pp. 549–575.

[4]P. Gacs, University of Rochester Computer Science Department, Technical Report No. 132, 1983 (to be published).

[5]Ya. G. Sinai, *Theory of Phase Transitions: Exact Results* (Pergamon, New York, 1982), pp. 1–10.

[6]T. Toffoli, Physica (Amsterdam) **10D**, 195 (1984).

[7]E. Domany, Phys. Rev. Lett. **52**, 871 (1984).

[8]L. S. Schulman and P. E. Seiden, J. Stat. Phys. **19**, 293 (1978).

[9]G. Nicolis and I. Prigogine, *Self-Organization in Nonequilibrium Systems* (Wiley, New York, 1977); H. Haken, *Synergetics* (Springer-Verlag, New York, 1983), 3rd ed.

[10]M. Eigen and P. Schuster, *The Hypercycle* (Springer, New York, 1979).

[11]S. Wolfram, Rev. Mod. Phys. **55**, 601 (1983), and references therein.

[12]E. Fredkin and T. Toffoli, Int. J. Theor. Phys. **21**, 219 (1982); N. Margolus, Physica (Amsterdam) **10D**, 81–95 (1984).

[13]P. Gacs and J. Reif, in *Proceedings of the Seventeenth ACM Symposium on the Theory of Computing, Providence, Rhode Island, 1985* (Association for Computing Machinery, New York, 1985), pp. 388–395.