Complex Systems **10** (1996) 143–155

Information-theoretic Analysis of Phase Transitions

Dirk V. Arnold*
Department of Computer Science,
University of Dortmund, LS XI,
44221 Dortmund, Germany

Abstract. Despite using different formalizations and investigating very different kinds of systems, the same unimodal dependence between disorder and complexity has been found in several independently conducted studies. Maximally interesting behavior of complex systems was observed at "the edge of chaos," the onset of instability between the ordered and the chaotic regime. The particular shape of the complexity-disorder plot has led researchers to suggest that complex systems can have inherent nontrivial information processing capabilities in the vicinity of a phase transition. However, it has subsequently been pointed out that the observed kind of dependence is a consequence of the definition of disorder used in the studies and that a different definition would make the structure suggesting a phase transition vanish.

In this paper, the dependence between disorder and complexity for two-dimensional Ising spin systems is investigated. Measures not sharing the flaw pointed out above are used, and the hypothesis of maximally interesting behavior in the vicinity of a phase transition is confirmed for simple, spatially homogeneous systems with random noise. Moreover, evidence is presented that more complex systems in which frustration is present can show interesting behavior over a broad range of noise levels.

1. Introduction

Phase transitions are a phenomenon intensively studied in statistical mechanics. Large systems often show a sudden change in their behavior as a parameter (which is often called temperature) is gradually varied. While below the critical temperature the system shows a very simple kind of organization and regular structures, above this temperature seemingly random

*Current address: School of Computing Science, Simon Fraser University, Burnaby, BC, Canada V5A 1S6.

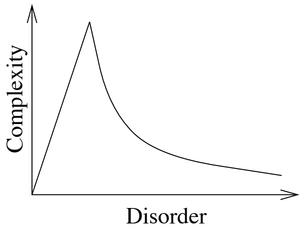

Figure 1: The dependence of complexity on disorder found in [4, 5, 11, 12]. The sharp peak at an intermediate disorder value was interpreted as evidence of a critical phase transition.

behavior can be observed. This fact can usually be attributed to two competing influences: symmetry-imposing laws intrinsic to the structure on one hand, and the disordering effect of thermal fluctuations on the other hand. As the amount of thermal fluctuations exceeds a certain threshold, the ordering powers have a negligible effect on the behavior of the system.

The notion of phase transitions has gained importance in complex systems theory as results reported in [4, 5, 11, 12, 22] sparked the idea that maximally interesting behavior of complex systems can be expected to emerge in the vicinity of phase transitions. By investigating the information processing capabilities of various kinds of systems, signs of globally coherent behavior and self-organization were found in a region for which Langton coined the term *edge of chaos*: the boundary between the ordered and chaotic regimes in the behavior of a system. These observations had considerable impact on the new field of artificial life.

Although Langton [11, 12] investigated the behavior of cellular automata (CA) while Crutchfield and Young's [4, 5] interest was on iterated function systems and widely differing complexity measures were employed, both studies arrived at the same, strikingly simple dependence sketched in Figure 1. The seemingly obvious conclusion was that neither very ordered systems with static structures that do not support information transmission, nor chaotic systems in which information cannot be persistently stored are capable of complex information processing tasks, and that only systems on the sharp borderline between the two extremes are computationally interesting.

However, in a closer investigation of disorder measures in [13], it is pointed out that the linear left boundary of both plots is purely due to the particular choice of a disorder measure, and that choosing a different measure would lead to the loss of the structure in those plots.

In this paper, the question of what results can be obtained when a physical system for which the existence of a second-order phase transition is well established is investigated. In particular, the behavior of information-theoretic measures applied to Ising spin systems at various temperatures is explored. The benefits of this approach are twofold: first, general insights into the behavior of information-theoretic measures in the vicinity of phase transitions are gained. It is shown that the complexity of a two-dimensional Ising ferromagnet indeed shows the expected unimodal dependence on disorder, with an accentuated peak near the critical temperature. Second, and maybe more important, the same methodology is employed to arrive at a better understanding of spin glasses for which the transition from the solid to the gaseous state is not abrupt but a gradual process. Evidence is presented that such nonhomogeneous systems can behave in a nontrivial manner over a wide range of temperatures.

Section 2 reviews measures of disorder and complexity and points out links to coding theory and dimension measures which make the definitions used in this paper seem the most natural candidates for an analysis of information processing capabilities of complex systems. In Section 3, Ising spin systems and an algorithm for their dynamical simulation on a massively parallel computer are described and the results from the information-theoretic analysis of Monte Carlo simulations are summarized. Section 4 concludes with a suggestion of how to interpret the results and some open problems.

2. Measures of disorder and complexity

The usual procedure for defining measures of disorder and complexity of processes or structures of any kind is to choose some type of model for the process or structure being investigated, compute the disorder and complexity of the model according to some model-specific definition, and then identify these with the disorder and complexity of the original process or structure. This procedure makes the definition of such measures inherently subjective, as different types of models are more or less well suited for representing the process or structure at hand. For example, the output of a pseudo random number generator is not disordered at all if the deterministic law according to which the sequence of pseudo random numbers is generated is known and used as a model for the process. However, using stochastic processes as a model, the output of a good generator will be classified as completely random.

The two classes of models most commonly used are stationary, ergodic stochastic processes and universal computers, leading to definitions of information-theoretic and automata-theoretic measures, respectively. The following is restricted to information-theoretic measures. However, it is well known that the information-theoretic disorder measure as defined below and the automata-theoretically defined Kolmogorov complexity are almost always equivalent. Furthermore, an informal argument can be given showing that their complexity counterparts, mutual information as defined below and Bennett's logical depth, are at least qualitatively similar. While the computation of automata-theoretic measures requires the cumbersome construction of automata which approximate the statistical properties of the object or structure being investigated, information-theoretic measures can be approximated without much effort. For a more extensive discussion of definitions of disorder and complexity measures see [8, 14].

A formal basis for the information-theoretic disorder measure is the entropy of a probability distribution, first defined in [18] as

$$H(S) = \sum_s p_S(s) \log \frac{1}{p_S(s)},$$

where S is a discrete random variable with distribution p_S and the summation extends over the whole range of possible values for S. Given a stationary, ergodic time series $(x_t)_{-\infty<t<\infty}$ of symbols from a finite alphabet Γ, let the random variable S_n range over the space Γ^n of words of length n occurring in the series and define the disorder of (x_t) as

$$H = \lim_{n\to\infty} \frac{H(S_n)}{n}. \tag{1}$$

As the size of the space Γ^n of all words of length n increases exponentially with n, this definition captures the same scale invariance properties as the information dimension

$$D = \lim_{\epsilon\to 0} \frac{H(\epsilon)}{\log(1/\epsilon)}$$

defined via box counting in continuous spaces. Instead of changing the scale by decreasing the box size ($\epsilon \to 0$), in the definition of disorder according to equation (1) we let $n \to \infty$ to achieve the effect of a change of scales by increasing the sequence length. Given stationarity and ergodicity the result is the same: both definitions compute the characteristic constant of a system under a change of scales.

Another justification for equation (1) stems from coding theory. As Hamming shows in [9], the disorder of a time series as defined above can be interpreted as its maximal compressibility by a general coding algorithm. In this way, disorder can be regarded as the average amount of information carried by each symbol of the series, or alternatively as a measure of randomness inherent in the series.

As a complexity measure, the mutual information

$$C = \lim_{m,n\to\infty} (H(S_m) + H(S_n) - H(S_{m+n})) \tag{2}$$

is widely used. It measures the amount of information a word of length m in the sequence contains about its continuation of length n for $m, n \to \infty$. Therefore, "shallow" sequences where consecutive symbols are uncorrelated have low complexity while sequences with inherent long-range correlations have high values of C.

The complexity of equation (2) can also be written as

$$C = \lim_{n\to\infty} (H(S_n) - nH) \tag{3}$$

so that for large enough n we have

$$H(S_n) \approx C + nH. \tag{4}$$

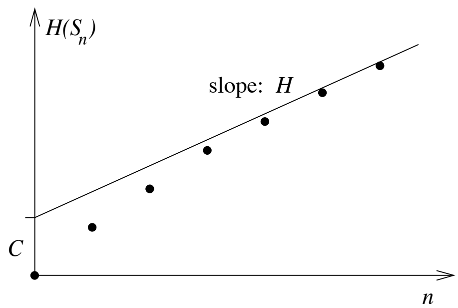

Figure 2: Typical relationship between the entropies $H(S_n)$ and the word length n. The slope of the limiting straight line is equal to the disorder H of the time sequence; the y-intercept is equal to the complexity C (plotted according to similar figures in [13, 19]).

Plotting the values of $H(S_n)$ over n typically leads to a diagram as shown in Figure 2. Equation (4) shows that H is the value of the slope of the limiting straight line that fits $H(S_n)$ as $n \to \infty$ and C is the value of the y-intercept of this straight line. As n is the logarithm of the size of the space in which the entropies are computed, Figure 2 effectively illustrates the concept of scale invariance in the definition of H.

Another interpretation which has been suggested for equation (4) draws from the distinction of the contents of computer memory into programs and data: The total information $H(S_n)$ inherent in a sequence of n symbols is composed of a part nH which equals the amount of information carried by n symbols (data) and a part C which is interpreted as the amount of information inherent in the code itself (program), and thus a measure for the difficulty of decoding the given sequence.

Neither [4, 5] nor [11, 12] use the disorder or complexity measures defined above. In [11, 12], scale invariance properties are not considered at all; the single-symbol entropy $H(S_1)$ is used as a disorder measure and the mutual information $H(S_1) + H(S_1') - H(S_1, S_1')$ where S_1 and S_1' denote *spatially* (and not temporally) related time series is used as a complexity measure. In [4, 5], $H(S_{16})/16$ is used as an approximation to the disorder measure defined above and probabilistic finite automata which reproduce statistical properties of the time series at hand are constructed for the definition of a complexity measure. For both disorder measures, all periodical time series are situated on the limiting straight line through the origin in Figure 1. As first pointed out in [13], using the disorder measure given by equation (1) instead, all periodical series have zero disorder, and the complexity-disorder relationship takes the form shown in Figure 3. In this figure, there is no indication of a phase transition.

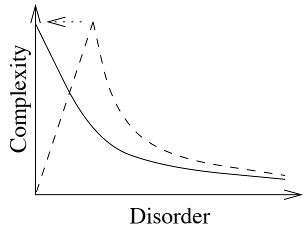

Figure 3: Change in the complexity-disorder relationship when using different measures. When using the disorder measure given by equation (1) instead of H_1 or $H_{16}/16$, the complexity-disorder relationship changes from the one depicted by the broken curve to the one depicted by the solid curve. The structure indicating a phase transition vanishes.

3. Analysis of Ising spin systems

A number of points make Ising spin systems an interesting candidate for the analysis of the complexity-disorder relationship of complex systems.

- Ising spin systems are naturally discrete; therefore, no discretization effects influence the simulation results.
- For the Ising ferromagnet, the existence of a second-order phase transition is a well-known fact. The behavior of the disorder and complexity measures in the vicinity of this transition can be studied and compared with their behavior when applied to spin glasses, the nature of which is not as well understood.
- Due to their regular structure, a relatively efficient implementation on massively parallel computers is possible. Unlike for more irregular systems such as random boolean networks or fully connected neural networks, experience can be gained from experiments over a wide temperature range and for a large number of systems.
- It can be expected that related systems such as Hopfield networks or simulated annealing strategies show qualitatively similar behavior, and that studying Ising spin systems furthers the understanding of their various relatives.

In this section, the Ising spin formalism is first outlined in general, then methods for the dynamical simulation of Ising spin systems are described, and afterwards results from the information-theoretic analysis of Ising ferromagnets and Ising spin glasses are presented.

3.1 Ising spin systems

Ising spin systems are large spatially distributed systems comprised of very simple, identical elements which can show very complex overall behavior due to an abundance of competing feedback links. They consist of a large number of binary variables $S_i \in \{\pm 1\}$ that are called spins. These form a regular n-dimensional grid and interact according to the energy function

$$E = -\frac{1}{2} \sum_{i,j} J_{ij} S_i S_j, \tag{5}$$

where the summation extends over all pairs of spins. The interaction coefficients J_{ij} are symmetric in the sense that $J_{ij} = J_{ji}$ and are usually monotonically decreasing in absolute value as the distance between spins increases. Vanishing interaction coefficients for spins more than one grid site apart are often assumed. The link between spins i and j is called ferromagnetic if $J_{ij} > 0$ and it is called antiferromagnetic if $J_{ij} < 0$. The statistical properties of Ising spin systems are described by the Gibbs distribution which states that the probability of configuration η is

$$p(\eta) = \frac{1}{Z} e^{-\beta E_\eta}$$

where Z is the partition function, $\beta = 1/T$ is the inverse temperature, and E_η is the energy of configuration η given by equation (5). It is easy to see that for low temperatures spins with a ferromagnetic connection tend to have the same value while spins with an antiferromagnetic connection tend to have opposite values.

Two important special cases which have been intensively studied in statistical mechanics as models for solids with magnetic properties are the Ising ferromagnet and spin glasses. While the two-dimensional Ising ferromagnet [10] for which all interaction coefficients between nearest neighbors are identically positive and longer-range interaction coefficients vanish is well understood and mathematically tractable [17], spin glasses are less amenable to mathematical analysis and are the subject of continuing research. For the state of the art in spin glass theory, see [6]. This reference also outlines the close relationship of spin glasses to neural network models and to simulated annealing strategies.

3.2 Computer experiments

Computer simulations are a useful tool for investigating the behavior of systems which are difficult or impossible to deal with using formal methods of mathematics. Monte Carlo methods have proven to be useful in a number of cases.

For a Monte Carlo simulation of Ising spin systems, a dynamical rule for updating configurations that captures the two competing influences of ordering spin interactions and disordering thermal fluctuations and that produces configurations with the probabilities of the Gibbs distribution has to be given. A rule known to lead to thermodynamic equilibrium is the Glauber rule

$$p(\eta \to \eta') = \frac{1}{1 + \exp(\beta \Delta E)} \tag{6}$$

which prescribes the probability of a state change from configuration η to configuration η' as a function of the energy change ΔE going along with the state change and the inverse temperature β.

This rule does not say anything about which spin flips to chose as candidates in a dynamical simulation. For exploring a large part of the configuration space, it would be most efficient to make big steps by updating as many spins as possible simultaneously. However, to be able to compute the energy difference ΔE associated with a configuration change, two neighboring spins must not flip at the same time step. Therefore, we follow the suggestion in [20] and employ probabilistic CA and a checker board updating pattern, changing the states of the spins on the even grid and on the odd grid alternately.

The experiments described below were carried out on two-dimensional grids of size 128×128 with periodical boundary conditions on a Connection Machine CM-2. Each spin interacted with its four nearest neighbors. After skipping the transient phase, 10,000 consecutive configurations were computed for each value of the temperature. Approximations to disorder and complexity were computed using the fast converging expressions

$$H = \lim_{n \to \infty} \left(H(S_{n+1}) - H(S_n) \right) \tag{7}$$

and

$$C = -\sum_{n=1}^{\infty} n \left(\left(H(S_{n+1}) - H(S_n) \right) - \left(H(S_n) - H(S_{n-1}) \right) \right). \tag{8}$$

Approximation errors in the computation of disorder have proven to be negligible. As for the complexity, it is of course not possible to extend the summation in equation (8) to infinity. Because all summands are positive, all that can be computed are lower bounds for the complexity. However, except for the immediate vicinity of the phase transition of the ferromagnet and the corresponding transition interval of the spin glass, the error resulting from summing up only the first few items is very small.

3.2.1 The Ising ferromagnet

The Ising ferromagnet is characterized by uniform interaction coefficients $J_{ij} = J > 0$ for neighboring spins and no direct interaction between spins more than one grid site apart. Two typical configurations of the system at different temperatures are shown in Figure 4. Its thermodynamic behavior is well understood, and the simulation results agree well with the theoretical predictions (see Figure 5).

Figure 6 shows that disorder increases monotonically with the temperature of the ferromagnet as expected, and that it resembles, not surprisingly, the curve of thermodynamic entropy when plotted over the temperature. The complexity as defined in equation (2) indeed has an accentuated peak near the critical temperature. Its height is underestimated due to the fact that the sum in equation (8) is only approximated.

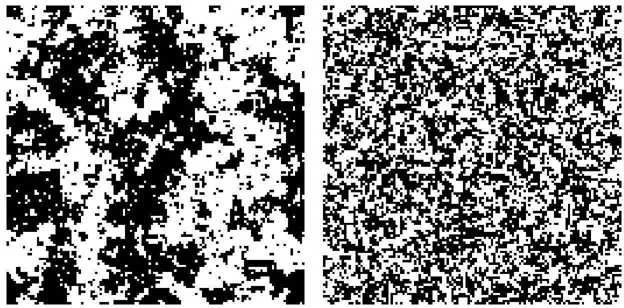

Figure 4: Typical configurations of a two-dimensional Ising ferromagnet near the critical temperature (left) and above the critical temperature (right). Spin variables of value 1 are shown in black, those with value -1 are depicted in white.

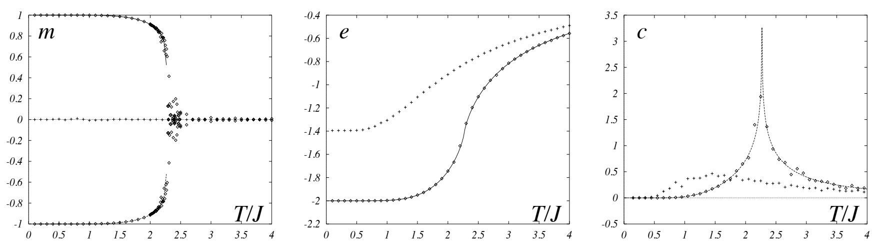

Figure 5: The dependence of some thermodynamic variables on temperature for the Ising ferromagnet (dots) and an Ising spin glass (crosses). The exact solution for the ferromagnet according to [17] is shown as a solid line. The variables shown are (from left to right) per spin magnetization m, per spin energy e, and per spin heat capacity c.

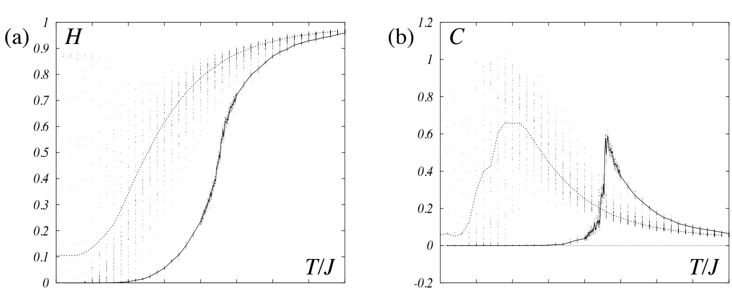

Figure 6: Dependence of disorder and complexity on temperature. (a) Disorder H over temperature T/J for the ferromagnet (lower curve) and a spin glass. (b) Complexity C over temperature for the ferromagnet (right peak) and a spin glass. The dots mark the measured values, the lines their arithmetic means.

3.2.2 Ising spin glasses

For the Ising spin glasses investigated here, the interaction coefficients J_{ij} between nearest neighbors are randomly distributed, taking values $\pm J$ with equal probability. Systems of this kind show frustration and a rugged energy landscape due to the presence of both positive and negative feedback links. Generally, their behavior is much less well understood than that of the considerably simpler ferromagnet, and even the existence of a phase transition in the strict sense of statistical mechanics is an open question (e.g., [6]).

Figure 5 shows the dependence of some thermodynamic variables of the simulated systems on temperature. For high temperatures, thermal fluctuations dominate the dynamical behavior and for $T \to \infty$ the ferromagnet and the spin glasses behave identically. However, for low temperatures the situation is quite different. While the energy landscape of the ferromagnet has only two very deep valleys, the rugged energy landscape of the spin glasses is comprised of a large number of valleys of widely varying sizes separated by energy barriers of widely varying heights. Therefore, even at low temperatures more than one energy valley is accessible to the system.

The comparably flat curve depicting the heat capacity does not seem to suggest the existence of an abrupt phase transition for Ising spin glasses. However, Figure 6 reveals a gradual phase transition accompanied by complexity values which even exceed those of the abrupt transition in the ferromagnet. High complexity values can be found at very low temperatures, but are most common in the temperature range of maximal heat capacity. The curves of both average disorder and average complexity for the spin glasses qualitatively resemble those for the ferromagnet. However, the average deviation of the measured values is substantially higher for the spin glasses. This is due to the fact that more than one valley of the energy landscape is accessible to the system already at comparatively low temperatures, and that the local structure of the energy landscape greatly influences the dynamics of the system and thereby the complexity of its trajectory. Given a sufficiently complex energy landscape, interesting behavior can occur over a wide temperature range.

4. Conclusions

The aim of this paper was to explore the relationship between complexity and disorder in large systems in the light of the hypothesis that complex systems can spontaneously develop the ability for nontrivial information processing in the vicinity of phase transitions. For this purpose, disorder and complexity measures were employed which are based on scale invariance properties of the investigated systems to study the behavior of dynamically simulated Ising spin systems. For the Ising ferromagnet, a system which exhibits a second-order phase transition, the hypothesis proved to be correct, although the structure of the complexity-disorder plot (Figure 7(a)) is qualitatively different from earlier results (e.g., [4, 11]) where the structure suggesting a phase transition is due to a peculiarity of the disorder measures employed.

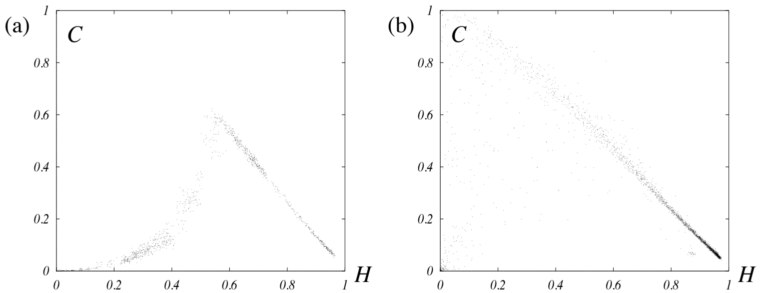

Figure 7: Dependence of complexity C on disorder H for (a) the ferromagnet and (b) a spin glass.

The conclusion from this seems to be that a sharp phase transition in the sense of statistical physics is no necessary precondition for interesting behavior in spatially nonhomogeneous systems with frustration, and that a gradual transition from ordered to disordered behavior may as well go along with complex behavior as an abrupt one. This coincides with the observation in [16] that "phase transitions which are not 'truly' critical are still interesting and maybe even more relevant for reality." If the frequency spectrum of the energy landscape is sufficiently broad, fluctuations of widely varying intensity will lead to interesting phenomena.

For the future, it will be interesting to examine the behavior of systems which are intermediate in complexity between the totally ordered ferromagnet and the structurally totally random spin glasses. Not only the temperature which controls the amount of random fluctuations but also the distribution of ferromagnetic and antiferromagnetic links can be subject to explicit manipulation, and the results reported here (as well as the classifications of CA by S. Wolfram [21] and C. Langton [11, 12], a recent study of random boolean networks by J. F. Lynch [15], and D. Amit's [1] work on the Hopfield model) hint at the existence of more interesting systems at the edge between order and chaos.

References

[1] D. H. Amit, H. Gutfreund, and H. Sompolinsky, "Storing Infinite Numbers of Patterns in a Spin-Glass Model of Neural Networks," *Physical Review Letters*, **55** (1985) 1530–1533.

[2] D. V. Arnold, "Informationstheoretische Analyse von Phasenübergängen" (Diploma thesis, University of Dortmund, Department of Computing Science, 1994).

[3] V. M. Alekseev and M. V. Yakobson, "Symbolic Dynamics and Hyperbolic Dynamic Systems," *Physical Reports*, **75** (1981) 287–325.

[4] J. P. Crutchfield and K. Young, "Inferring Statistical Complexity," *Physical Review Letters*, **63** (1989) 105–109.

[5] J. P. Crutchfield and K. Young, "Computation at the Onset of Chaos," in *Complexity, Entropy, and the Physics of Information*, edited by W. H. Zurek (Addison Wesley, Redwood City, CA, 1990).

[6] K. H. Fischer and J. A. Hertz, *Spin Glasses* (Cambridge University Press, Cambridge, MA, 1991).

[7] P. Grassberger, "Toward a Quantitative Theory of Self-Generated Complexity," *International Journal of Theoretical Physics*, **25** (1986) 907–938.

[8] P. Grassberger, "Problems in Quantifying Self-Generated Complexity," *Helvetica Physica Acta*, **62** (1989) 489–511.

[9] R. W. Hamming, *Coding and Information Theory* (Prentice-Hall, Englewood Cliffs, NJ, 1980).

[10] E. Ising, "Beitrag zur Theorie des Ferromagnetismus," *Zeitschrift für Physik*, **31** (1925) 253–258.

[11] C. G. Langton, "Computation at the Edge of Chaos," *Physica D*, **42** (1990) 12–37.

[12] C. G. Langton, "Life at the Edge of Chaos," in *Artificial Life II*, edited by C. G. Langton, C. Taylor, J. D. Farmer, and S. Rasmussen (Addison Wesley, Reading, MA, 1992).

[13] W. Li, "On the Relationship between Complexity and Entropy for Markov Chains and Regular Languages," *Complex Systems*, **5** (1991) 381–399.

[14] K. Lindgren and M. G. Nordahl, "Complexity Measures and Cellular Automata," *Complex Systems*, **2** (1988) 409–440.

[15] J. F. Lynch, "A Phase Transition in Random Boolean Networks," in *Artificial Life IV*, edited by R. A. Brooks and P. Maes (MIT Press, Cambridge, MA, 1994).

[16] K. Nagel and S. Rasmussen, "Traffic at the Edge of Chaos," in *Artificial Life IV*, edited by R. A. Brooks and P. Maes (MIT Press, Cambridge, MA, 1994).

[17] L. Onsager, "Crystal Statistics I: A Two-Dimensional Model with an Order-Disorder Transition," *Physical Review*, **65** (1944) 117–149.

[18] C. E. Shannon, "A Mathematical Theory of Communication," *The Bell System Technical Journal*, **27** (1948) 379–423 and 623–656.

[19] R. Shaw, *The Dripping Faucet as a Model Chaotic System* (Aerial Press, Santa Cruz, CA, 1984).

[20] G. Vichniac, "Simulating Physics with Cellular Automata," *Physica D*, **10** (1984) 96–115.

[21] S. Wolfram, "Universality and Complexity in Cellular Automata," *Physica D*, **10** (1984) 1–35.

[22] W. K. Wooters and C. G. Langton, "Is There a Sharp Phase Transition for Deterministic Cellular Automata?" *Physica D*, **40** (1990) 95–105.