Correlations and Random Information in Cellular Automata

Kristian Lindgren
*Physical Resource Theory Group, Chalmers University of Technology,
University of Göteborg, S-41296 Göteborg, Sweden*

Abstract. Infinite one-dimensional cellular automata are studied using information theory. The average information per cell is divided into contributions from different correlation lengths and random variations (measure entropy). It is shown that the measure entropy is non-increasing in time for deterministic rules, and constant for rules which are one-to-one mappings of their first or last argument (*almost reversible* rules). For probabilistic rules, there is no such general law, but for almost reversible rules where the states are randomly shifted, it is proven that the system evolves towards the maximally disordered state, independent of initial conditions.

It is discussed how some of the information-theoretical concepts are related to analogous concepts in algorithmic information theory, and an equality between algorithmic information and measure entropy is proved.

Numerical and analytical examples are given for specific rules.

1. Introduction

Cellular automata have been used to simulate a variety of physical systems [1], such as the microscopic motion in fluids [2], the macroscopic concentrations in chemical self-organizing systems [3], the growth of crystals [4], and abstract models for phase transitions [5,6]. Self-organizing systems or dissipative structures [7,8] have the property of evolving into spatially or temporally ordered states, but we do not know much about evolutionary criteria for such processes. The study of evolutionary rules and other mathematical properties of cellular automata may yield results which can be applied to more specific (e.g. chemical) self-organizing systems. Investigations of complexity measures and related concepts for cellular automata have revealed interesting properties of their spatial organization and temporal behavior [9–14].

In algorithmic information theory, concepts for measuring structure and complexity are defined [15] which have the advantage of being generally applicable, but they are usually not computable. It will be shown that these

concepts are identical to or related to computable information-theoretical measures [16] for infinite one-dimensional cellular automata at finite times. An equality between algorithmic information (Kolmogorov complexity or Chaitin complexity) and measure entropy has been stated [17], and a relation between these concepts has been proved for finite systems [18]. In section 3, a proof of this equality is given for a system which is the outcome of an infinite stationary stochastic process.

The second law of thermodynamics tells us that the entropy of a closed system cannot decrease. In a one-dimensional cellular automaton, the system can for most rules not be considered as closed in the thermodynamical sense. In a thermodynamical system, reversible microscopic rules, which may be influenced by noise, govern the time evolution of the system, which leads to a maximization of the entropy. However, the rules which are responsible for the time evolution in cellular automata are not generally reversible. In cellular automata correlations are often built up and the (measure) entropy decreases [9].

In a cellular automaton, the randomness, expressed by the measure entropy, may (partially) be irreversibly transformed to correlational information when complex structures evolve. The system is either closed or open with respect to random information (noise), giving deterministic or probabilistic rules. When the system is influenced by noise, correlations can be destroyed leading to an increase of the randomness. Only if the rules are probabilistic we can have a "second law" for cellular automata, and in section 4, it is shown that if an almost reversible rule is influenced by noise, the measure entropy increases until the system is completely randomized.

In sections 2 and 4, we apply concepts from information theory [13,16, 19,20] to cellular automata. The average information per cell is divided into contributions from different correlation lengths and random variations. Laws concerning the change of measure entropy in time for different classes of rules are proven. In section 5, this is applied to numerical and analytical examples for specific rules. Complex long-range behavior can be understood in terms of the concepts presented in section 2.

2. Information-theoretical concepts for lattice distributions

In this section, a summary of some concepts in [16] is given. Let $p(k)$ be a normalized probability distribution,

$$p(k) \geq 0$$

$$\sum_k p(k) = 1 \tag{2.1}$$

and let $p_0(k)$ be a reference distribution which is positive everywhere,

$$p_0(k) > 0$$

$$\sum_k p_0(k) = 1 \tag{2.2}$$

Then the contrast (Kullback-information, relative information) of p with respect to p_0 (log denotes base two logarithm),

$$K[p_0, p] = \sum_k p(k) \log \frac{p(k)}{p_0(k)} \tag{2.3}$$

is the information gain when an a priori distribution p_0 is replaced by an a posteriori distribution p.

The contrast has the property

$$K[p_0, p] \geq 0 \tag{2.4}$$

where equality holds only if p and p_0 are identical.

Consider a one-dimensional discrete system with an infinite number of lattice sites. Each site can be in either of two states: 0 or 1. (A higher number of states can be handled by binary coding.) The total information per site, one bit, can be decomposed into chemical, correlational, and textual contrast. The chemical contrast is the information due to an average density of zeroes and ones differing from 1/2. The correlational contrast is the amount of information present in all correlations within the system. The textual contrast is the random information or, if the system is produced by a language, the amount of information conveyed through the text [19,20]. The textual contrast is identical to the measure entropy [10]. (In the theory of dynamical systems, measure entropy (Kolmogorov-Sinai entropy) is the mean rate of creation of information in time [21], but here it is the spatial counterpart.)

Let $(i_1 \ldots i_m)$ be a certain sequence of zeroes and ones and $p_m(i_1 \ldots i_m)$ be the probability that a randomly chosen m-length sequence coincides with $(i_1 \ldots i_m)$. Suppose that the system is large scale homogenous, i.e. the result of a stationary stochastic process, so that $p_m(i_1 \ldots i_m)$ is well defined. Further, let $\tilde{p}(i_1 \ldots i_m)$ be the estimated (maximum entropy method) probability if correlations only up to $m-1$ are known,

$$\tilde{p}_m(i_1 \ldots i_m) = \frac{p_{m-1}(i_1 \ldots i_{m-1}) p_{m-1}(i_2 \ldots i_m)}{p_{m-2}(i_2 \ldots i_{m-1})} \tag{2.5}$$

The chemical contrast k_{ch} is the contrast of mean concentrations p_1 with respect to the reference distribution $p_1^{(0)} = \{1/2, 1/2\}$,

$$k_{\text{ch}} = K[p_1^{(0)}; p_1] \tag{2.6}$$

The correlational contrast k_{corr} can then be written as a sum of contributions from different correlation lengths,

$$k_{\mathrm{corr}} = \sum_{m=2}^{\infty} k_m \tag{2.7}$$

where k_m is the contrast of p_m with respect to $\tilde{p}_m$,

$$k_m = \mathrm{K}[\tilde{p}_m; p_m] \tag{2.8}$$

The remaining part, $1 - k_{\mathrm{ch}} - k_{\mathrm{corr}}$, is the textual contrast k_{tx} [19,20], which is an average over local textual contrasts [16], or the measure entropy s_μ,

$$s_\mu = \lim_{m \to \infty} \frac{1}{m} S_m = \lim_{m \to \infty} \Delta S_m \tag{2.9}$$

where

$$S_m = - \sum_{i_1 \ldots i_m} p_m(i_1 \ldots i_m) \log p_m(i_1 \ldots i_m) \tag{2.10}$$

$$\Delta S_m = S_m - S_{m-1} \tag{2.11}$$

The limes of ΔS_m should be used for numerical estimates of the measure entropy (2.9) since it converges faster than S_m/m. One can define a mean correlation length $\bar{m}$, where length is defined so that the distance between adjacent cells is 1,

$$\bar{m} = \sum_{m=1}^{\infty} m \, k_{m+1}/k_{\mathrm{corr}} \tag{2.12}$$

and if it is multiplied with the correlational contrast, one gets

$$\eta = \bar{m} \, k_{\mathrm{corr}} \tag{2.13}$$

which is the "effective measure complexity" defined by Grassberger [13]. Numerical calculations of ΔS_m and η have been performed for different cellular automaton rules [13,14].

3. Algorithmic information theory

The concepts of correlational and textual contrast are closely related to concepts in algorithmic information theory. The algorithmic information $H(\alpha_m)$ of a sequence α_m of m zeroes and ones is defined as the minimal program for a general-purpose computer that generates the sequence [15]. Consider an infinite sequence α_∞ which is the outcome of a stationary stochastic process and has $s_\mu > 0$. Let $h(\alpha_m)$ be the average algorithmic information per symbol of the sequence α_m, $h(\alpha_m) = H(\alpha_m)/m$. The correlational part, including the chemical contrast, of the total information is possible to describe, at least approximately, by a finite program, i.e. a finite amount of algorithmic information. However, since almost all sequences are "algorithmically random"

[22], the textual contrast must either be given explicitly, if the exact sequence is to be generated, or be available as the information in random numbers, if it is enough that the probabilities for all sequences are correctly generated. In both cases, the amount of information needed is infinite—in average, s_μ per lattice site. Thus, for almost all infinite sequences with $s_\mu > 0$ we can conclude that the average algorithmic information is equal to the measure entropy,

$$\lim_{m \to \infty} \frac{1}{m} H(\alpha_m) = h(\alpha_\infty) = s_\mu(\alpha_\infty) \tag{3.1}$$

We prove equation (3.1) by showing how to construct an algorithm with average length $h(\alpha_m)$ which generates the sequence α_m. The algorithm consists of one part which is a code for the sequence α_m, and another part which serves as a decoder. We divide the sequence α_m into n-length sequences ($n \ll m$), where each n-length sequence is denoted by a new symbol γ, $\gamma \in \{\gamma_1, \ldots, \gamma_N; N = 2^n\}$, so that a new sequence $\Gamma_{m,n}$ of γ-symbols is formed. Given the symbol β, the probability for the next symbol γ_i is the conditional probability $p_\beta(\gamma_i) = p(\beta\gamma_i)/(p(\beta))$, where $p(\beta\gamma_i)$ and $p(\beta)$ are the probabilities (defined as in section 2) for the sequences $\beta\gamma_i$ and β respectively. The code words for the symbols γ_i depend on the previous sequence β, and are chosen so that their lengths $l_\beta(\gamma_i)$ fulfill

$$-\log p_\beta(\gamma_i) \le l_\beta(\gamma_i) < -\log p_\beta(\gamma_i) + 1 \tag{3.2}$$

Since $\sum_i 2^{-l_\beta(\gamma_i)} \le 1$, we know from coding theory (e.g. [23]) that there exists an "instantaneous" code with the given lengths. In an instantaneous code, no word is prefix of another one so that the end of each word is given by the word itself. This is important since the coding depends on the previous symbol β. The average code word length l is

$$l = \sum_\beta p(\beta) \sum_{\gamma_i} p_\beta(\gamma_i) l_\beta(\gamma_i)$$

Taking the average of the inequality (3.2), we get for the logarithmic terms

$$-\sum_\beta p(\beta) \sum_{\gamma_i} p_\beta(\gamma_i) \log p_\beta(\gamma_i) = \sum_{\beta\gamma_i} p(\beta\gamma_i) \log \frac{p(\beta)}{p(\beta\gamma_i)} =$$

$$= S_2(\Gamma_{m,n}) - S_1(\Gamma_{m,n}) = \Delta S_2(\Gamma_{m,n})$$

where S_1 and S_2 are entropies defined as in equation (2.10). To make the algorithm self-delimiting, we add a prefix to each code word, which is one if the present symbol is the last one in the sequence $\Gamma_{m,n}$ and zero otherwise. Then, the average code length per symbol in α_m is $l_{m,n} = (l+1)/n$, and the inequality (3.2) can be written

$$\frac{1}{n}(\Delta S_2(\Gamma_{m,n}) + 1) \le l_{m,n} < \frac{1}{n}(\Delta S_2(\Gamma_{m,n}) + 2)$$

Writing the entropies as functions of the original sequence α_m, we get

$$\frac{1}{n}\Delta S_2(\Gamma_{m,n}) = \frac{1}{n}(S_{2n}(\alpha_m) - S_n(\alpha_m)) \rightarrow s_\mu(\alpha_\infty)$$

if $n, m \rightarrow \infty$ as $n \ll m$.

The length of the decoder grows exponentially with n and depends on the distributions $p_\beta(\gamma_i)$, but is independent of the length m of α_m. Thus, if we let m, n, $\log(m)/n \rightarrow \infty$, then $l_{m,n} \rightarrow h(\alpha_\infty)$ and $l_{m,n} \rightarrow s_\mu(\alpha_\infty)$, which gives equation (3.1).

Chaitin [15] has proposed a definition of structure at different levels using algorithmic information theory as follows. Divide the sequence α_m into sequences of length d, and let $H_d(\alpha_m)$ be the sum of the algorithmic information of these sequences, where mutual information between disjoint sequences is not used. Thus, the difference $\Delta_d H_d(\alpha_m) = H_d(\alpha_m) - H_{d-1}(\alpha_m)$ measures the structure at level d, which corresponds to the contribution to the correlational contrast from d-point distributions (2.8),

$$- \lim_{m\rightarrow\infty} \frac{1}{m}\Delta_d H_d(\alpha_m) \approx k_d = K[\tilde{p}_d; p_d] \qquad (3.3)$$

4. Information theory of cellular automata

Let us apply the information-theoretical concepts to the evolution of an infinite one-dimensional cellular automaton. The extension to higher dimensions is straightforward. The initial state is generated by a stochastic stationary process giving spatially independent probabilities for all possible sequences. Consider a rule R that depends on neighbors up to the finite distance $r > 0$. A sequence of $m + 2r$ sites with the state $\alpha_{m+2r} = (i_{1-r} \ldots i_{m+r})$ then determines the state in a sequence of m sites $\beta_m = R_m(i_{1-r} \ldots i_{m+r}) = (j_1 \ldots j_m)$. The probabilities for sequences at times t and $t + 1$ are then related as

$$p_m(\beta_m; t + 1) = \sum_{\alpha_{m+2r}} T_R(\alpha_{m+2r}, \beta_m) p_{m+2r}(\alpha_{m+2r}; t) \qquad (4.1)$$

$$T_R(\alpha_{m+2r}, \beta_m) = \delta(R_m(\alpha_{m+2r}), \beta_m) \qquad (4.2)$$

which gives the entropy at time $t + 1$,

$$S_m(t + 1) = - \sum_{\beta_m} p_m(\beta_m; t + 1) \log p_m(\beta_m; t + 1) \leq S_{m+2r}(t) \qquad (4.3)$$

The increase of measure entropy for one timestep, $\Delta_t s_\mu(t) = s_\mu(t+1) - s_\mu(t)$, is then

$$\Delta_t s_\mu(t) = \lim_{m \to \infty} \left(\frac{1}{m} S_m(t+1) - \frac{1}{m+2r} S_{m+2r}(t) \right) =$$

$$= \lim_{m \to \infty} \left(\frac{1}{m+2r} (S_m(t+1) - S_{m+2r}(t)) + \right.$$

$$\left. + \left(\frac{1}{m} - \frac{1}{m+2r} \right) S_m(t+1) \right) \le 0 \tag{4.4}$$

since, in the second line, the first term in the limes is negative or zero and the second term goes to zero as $m \to \infty$. Thus, the measure entropy cannot increase when a deterministic rule is applied, a fact that has been observed in numerous simulations [9,10].

Assume that the rule R is a one-to-one mapping of its last argument (a one-to-one mapping of the first argument is treated analogously); i.e., the rule is of the form

$$R(i_{-r} \ldots i_r) = i_r + f(i_{-r} \ldots i_{r-1}) \bmod 2 \tag{4.5}$$

where f is a mapping from $\{0,1\}^{2r}$ to $\{0,1\}$. These rules are surjective; i.e., for all sequences β_m there is a sequence α_{m+2r} such that $R_m(\alpha_{m+2r}) = \beta_m$, all m. (If f is a sum over (not necessarily all) elements in $\{i_{-r}, \ldots, i_{r-1}, 1\}$, R is additive.) As before, we have that $R_m(i_{1-r} \ldots i_{m+r}) = (j_1 \ldots j_m)$, but now we also have an inverse rule $\hat{R}$ that gives the state $\eta_m = (i_{1+r} \ldots i_{m+r})$ at time t, if the state $\gamma_{2r} = (i_{1-r} \ldots i_r)$ at time t and the state $\beta_m = (j_1 \ldots j_m)$ at time $t+1$ are known,

$$\eta_m = \hat{R}(\gamma_{2r}; \beta_m) \tag{4.6}$$

Thus, knowing the state at time $t+1$, it is possible to reconstruct the state at time t, using only a finite amount of information (at most, $2r$ bits) at time t. These rules are not reversible in the sense discussed in [9], so we call them *almost reversible*. The transfer matrix of (4.1) then has the properties

$$1 \le \sum_{\alpha_{m+2r}} T_R(\alpha_{m+2r}, \beta_m) \le 2^{2r}$$

$$\sum_{\beta_m} T_R(\alpha_{m+2r}, \beta_m) = 1 \tag{4.7}$$

The difference in entropy between time $t+1$ and t is

$$S_m(t+1) - S_{m+2r}(t) = -\sum_{\beta_m} p_m(\beta_m, t+1) \log p_m(\beta_m, t+1) +$$

$$+ \sum_{\alpha_{m+2r}} p_{m+2r}(\alpha_{m+2r}, t) \log p_{m+2r}(\alpha_{m+2r}, t) \ge -2r \tag{4.8}$$

where (4.7) is used. Thus,

$$\Delta_t s_\mu(t) \geq \lim_{m \to \infty} -\frac{2r}{m} = 0$$

But we have already shown that $\Delta_t s_\mu(t) \leq 0$, and hence, for almost reversible rules (4.5),

$$\Delta_t s_\mu(t) = 0 \tag{4.9}$$

These rules conserve randomness independent of initial state.

Assume that the rule R at every site has the probability q of making an error, i.e. the conjugate rule $\bar{R} = 1 - R$ is accidentally applied. The probabilistic rule formed in this way is denoted by (R,q). Then, the change in entropy for one time-step can be divided into one part coming from the deterministic rule R, as in equation (4.4), and another part due to the noise. The system is open to random information and, as will be shown below, the influence of noise gives a non-negative contribution to the change in measure entropy. When the states in randomly chosen lattice sites are shifted, the probabilities $p_m(\alpha_m)$ for sequences $\alpha_m = (i_1 \ldots i_m)$ are transformed according to

$$\hat{p}_m(\alpha_m) = \sum_{\beta_m} T_{\text{err}}(\alpha_m, \beta_m) p_m(\beta_m) \tag{4.10}$$

The transfer matrix is

$$T_{\text{err}}(\alpha_m, \beta_m) = q^{H(\alpha_m, \beta_m)}(1 - q)^{m - H(\alpha_m, \beta_m)} \tag{4.11}$$

where $H(\alpha_m, \beta_m)$ is the Hamming distance between α_m and β_m (the number of positions in which the sequences differ). Thus, the transfer matrix is symmetric and normalized,

$$\sum_{\beta_m} T_{\text{err}}(\alpha_m, \beta_m) = \sum_{\alpha_m} T_{\text{err}}(\alpha_m, \beta_m) = 1 \tag{4.12}$$

The entropy of $\hat{p}_m$, then, is

$$S[\hat{p}_m] = -\sum_{\alpha_m} \hat{p}_m(\alpha_m) \log \hat{p}_m(\alpha_m) \geq$$

$$\geq -\sum_{\beta_m} p_m(\beta_m) \log p_m(\beta_m) = S[p_m] \tag{4.13}$$

where the unequality comes from equation (4.12) and the convexity of S. If the entropy increase due to the noise ΔS_{noise} is written as

$$\Delta S_{\text{noise}} = S[\hat{p}_m] - S[p_m] =$$

$$= \sum_{\alpha_m} \hat{p}_m(\alpha_m) K[T(\alpha_m, \cdot); T(\alpha_m, \cdot) p_m(\cdot)/\hat{p}_m(\alpha_m)] \geq 0 \tag{4.14}$$

	Deterministic	Probabilistic (R, q)
Not *almost reversible*	$\Delta_t s_\mu(t) \leq 0$	—
Almost reversible	$\Delta_t s_\mu(t) = 0$	$\Delta_t s_\mu(t) \geq 0$

Table 1: The change of measure entropy in time for different classes of rules.

we immediately see that ΔS_{noise} vanishes only if $\hat{p}_m(\alpha_m) = p_m(\beta_m)$ for all α_m and β_m; that is, $\hat{p}_m = p_m = (1/2)^m$. For an almost reversible rule, this means that the system evolves to the maximally disordered state with spatial entropy $s_\mu = 1$ and chemical and correlational contrast $k_{\text{ch}} = k_{\text{corr}} = 0$.

The results concerning the change of s_μ in time are summarized in table 1.

To find $\Delta_t \eta(t) = \eta(t+1) - \eta(t)$ for almost reversible rules, we write equation (2.12) as

$$\eta(t) = \lim_{m \to \infty} S_m(t) - m s_\mu(t) \tag{4.15}$$

which together with equation (4.9) gives

$$\Delta_t \eta(t) = \lim_{m \to \infty} [S_m(t+1) - S_{m+2r}(t) + 2r\, s_\mu]$$

Equations (4.3) and (4.8) then give the following limits for $\Delta_t \eta(t)$ for almost reversible rules.

$$s_\mu - 1 \leq \frac{\Delta_t \eta(t)}{2r} \leq s_\mu \tag{4.16}$$

It is easy to construct an example where $\Delta_t \eta(t)$ is negative, and a case in which Grassberger's complexity η increases is shown in section 5.

5. Examples and discussion

In this section, we consider rules that are depending on nearest neighbors only. The rules are numbered according to Wolfram's notation [9].

The fact that $s_\mu(t)$ is time-independent for almost reversible rules implies that, starting with a completely random sequence with $s_\mu = 1$, all sequences will remain equally probable [9–13,24]. If an initial state without correlations but with different densities of zeroes and ones, $k_{\text{ch}} > 0$ and $k_{\text{corr}} = 0$, is chosen, the measure entropy, $s_\mu < 1$, will stay at the initial level. The dynamics may, however, allow for changes between chemical and different correlational contrasts.

For most rules, starting with $s_\mu = 1$, the measure entropy will decrease to a stationary level $\hat{s}_\mu < 1$ [10]. In these cases, correlations are built up, and random information is transformed to chemical and correlational contrast. An example of this is shown in figure 1, where rule 110 (see figure 2) leads
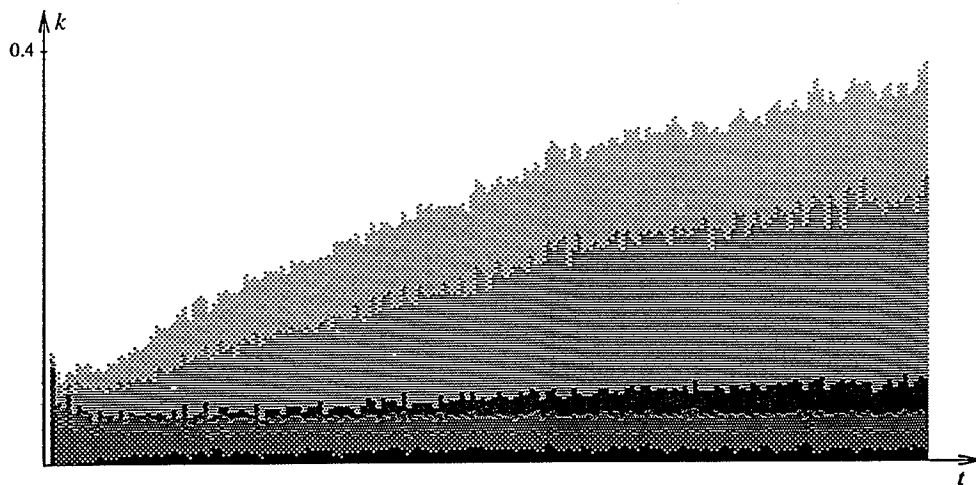
Figure 1: The evolution of chemical and correlational contrast for rule 110. At each time, the chemical contrast and contributions to the correlational contrast from sequences of length $m \leq 6$ are calculated and drawn on top of each other. The contrasts are represented by different grayness, from black for chemical contrast to light gray for correlational contrast from length 6. Initially, no correlations are present, but as the system evolves the correlational contrast increases, especially from lengths 5 and 6. The system approaches a pattern with periodicity 14 in space and 7 in time. In the periodic state, the only contribution to the correlational contrast comes from sequences of length $m \leq 6$. A system of 5000 lattice sites has been used.
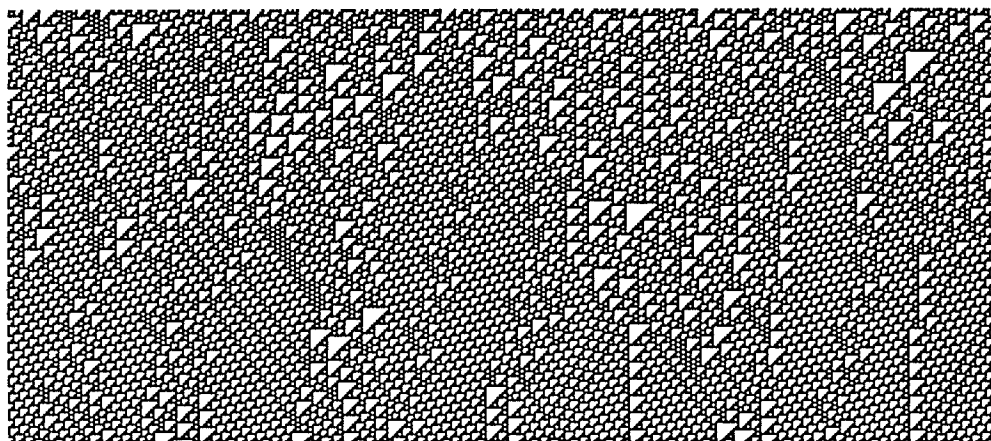

Figure 2: The space-time pattern of rule 110 clearly shows how correlations are built up. The periodic pattern mentioned in figure 1 is visible. (A system of 200 time steps and 400 lattice sites was used.)

to an increase in correlational contrast. In some cases, it can be shown that the rule at the stationary level simulates an almost reversible rule—e.g., rule 182 simulates rule 195.

We have shown that if noise may influence the evolution and the rule is almost reversible, the system reaches the maximally disordered state, independent of its initial state. This case is analogous to the time evolution in thermodynamical systems. One example is the simulation of particle motions in a fluid [2], if noise is added. Another example is found in a study of phase transitions in two-dimensional stochastic cellular automata, where an additive rule performing "turbulence" was observed [6]. For other rules, however, a balance is reached between entropy increase due to noise and entropy decrease due to irreversibility.

The evolution of the almost reversible rule 195 is studied numerically to show the dynamical behavior. We choose an initial state which is uncorrelated with a density $p = 0.1$ of zeroes. Since this is an additive rule, it is possible to find analytical expressions for probabilities of different sequences. For time steps $t^* = 2^n (n = 1, 2, \ldots)$, the chemical contrast is $k_{ch} \approx 0.320$, and one finds that the only contribution to correlational contrast comes from correlation lengths which are multiples of the time t^* [9], giving a correlational contrast $k_{corr} \approx 0.211$. Since the rule is almost reversible, the measure entropy is constant, $s_\mu \approx 0.469$. In figure 3, the evolution of chemical contrast and m-sequence correlational contrasts are shown for $m = 2, 3, \ldots, 6$. The distance between the level $1 - s_\mu$ and the shaded contrasts below it corresponds to correlational contrasts of order greater than 6. At the time steps t^*, the chemical contrast is relatively high and no correlations of lengths less than t^* are present. The mean correlation length $\bar{m}(t^*)$, equation (2.12), is proportional to t^* leading to an ever increasing η, equation (2.13). At these times, if correlations of higher order ($\geq t^*$) are neglected, the state can be regarded as an uncorrelated (initial) state, which explains the self-similarity in the figure. Obviously, the probabilities $p_m(\alpha_m)$ do not converge in the time evolution, which has been proven to be a general feature of additive rules if $p \neq 1/2$ [24]. Because of the self-similarity, we conjecture that, for almost all times as $t \to \infty$, there is no correlational information from finite lengths. Although $s_\mu < 1$, the system appears to be completely random.

In figure 4, it is illustrated how noise destroys the correlations, so that a maximally disordered state is reached. Each time the rule 195 is applied, the probability for making an error is $q = 0.01$, giving a probabilistic rule. The initial state has the same properties as the initial state of the system in figure 3, and a comparison between the figures reveals how sensitive the correlations are to random perturbations. The space-time patterns created by the deterministic and the probabilistic rule are shown in figure 5.

The two-dimensional space-time pattern of a one-dimensional deterministic rule always has measure entropy equal to zero, since it is sufficient to specify the states in lattice sites near the border of a space-time rectangle to achieve the state of the whole rectangle [10].

Figure 3: The evolution of chemical and correlational contrast, as in figure 1 (the distance that remains to level $1 - s_\mu$ corresponds to the information in correlations of higher order), for the additive rule 195. Although the measure entropy is constant, information may flow between different correlation lengths if the initial uncorrelated state has a density of zeroes p differing from $1/2$ (here $p = 0.1$). A system of 2000 lattice sites has been used.
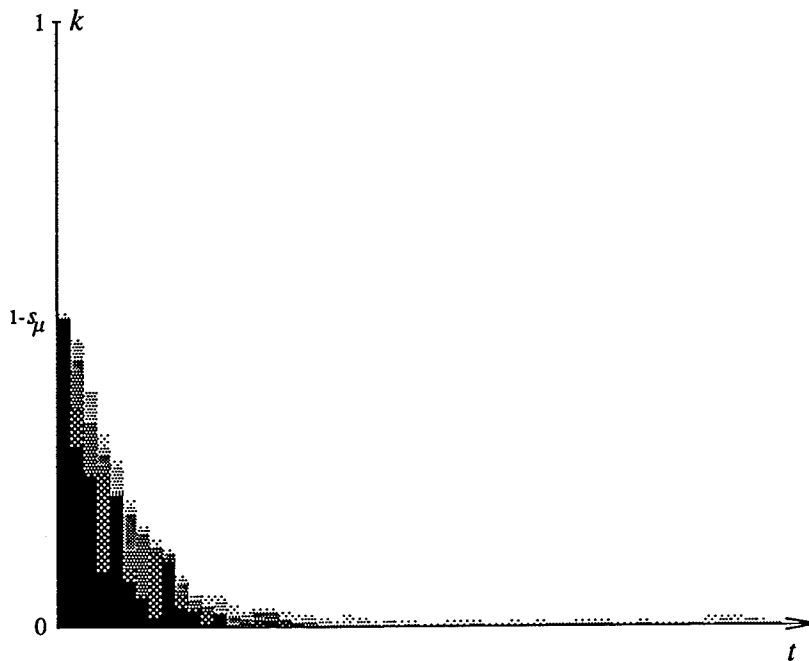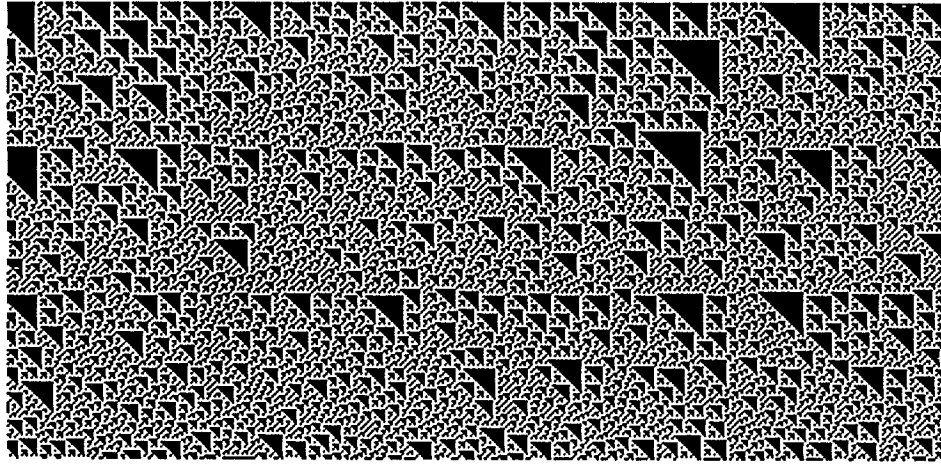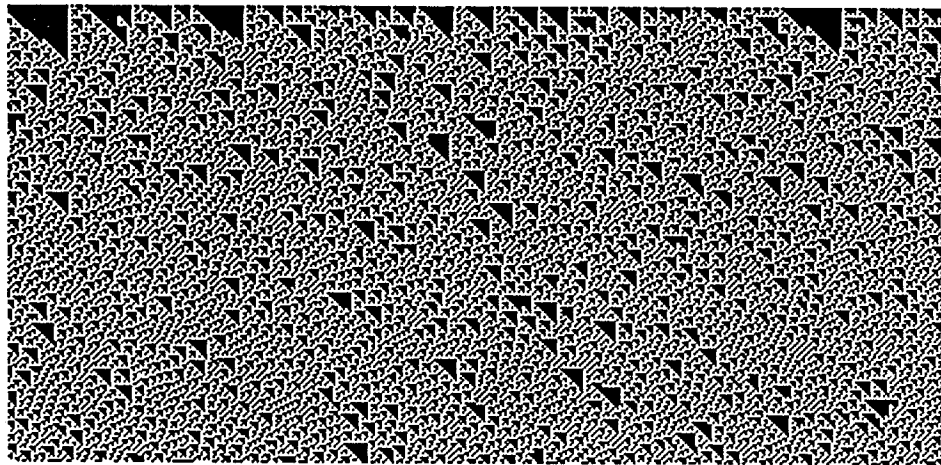

Figure 4: If the evolution of figure 3 is modified by noise, the correlations will be destroyed. Here, on average, the state in every hundredth lattice site is shifted.

(a)

(b)

Figure 5: The space-time patterns formed by the rules of figures 3 and 4. In 5a, the deterministic rule is applied, and in 5b, noise is added.

The necessity of probabilistic rules for an increase of algorithmic information has been discussed by Bennett [25]. This corresponds to equations (4.4) and (4.14) since the algorithmic information is equal to the measure entropy, equation (3.1).

Acknowledgments

References

[1] Stephen Wolfram, ed., *Theory and Applications of Cellular Automata*, (World Scientific, Singapore, 1986).

[2] J. B. Salem and Stephen Wolfram, "Thermodynamics and hydrodynamics of cellular automata" in *Theory and Applications of Cellular Automata*, (World Scientific, 1986).

[3] A. T. Winfree, E. M. Winfree, and H. Seifert, "Organizing centers in a cellular exitable medium", *Physica*, **17D** (1985) 109–115.

[4] Norman H. Packard, "Lattice models for solidification and aggregation" in *Theory and Applications of Cellular Automata*, (World Scientific, 1986).

[5] W. Kinzel, "Phase transitions of cellular automata", *Zeitschrift für Physik B*, **58** (1985) 229–244.

[6] K. Kaneko and Y. Akutsu, "Phase transitions in two-dimensional stochastic cellular automata", *Journal of Physics A*, **19** (1986) L69–L75.

[7] G. Nicolis and I. Prigogine, *Self-Organization in Non-equilibrium Systems*, (Wiley Interscience, New York, 1977).

[8] H. Haken, *Advanced Synergetics*, (Springer, Berlin, 1983).

[9] Stephen Wolfram, "Statistical mechanics of cellular automata", *Reviews of Modern Physics* **55** (1983) 601–644.

[10] Stephen Wolfram, "Universality and complexity in cellular automata", *Physica*, **10D** (1984) 1–35.

[11] Stephen Wolfram, "Random sequence generation by cellular automata", *Advances in Applied Mathematics*, **7** (1986) 123–169.

[12] Stephen Wolfram, "Computation theory of cellular automata", *Communications in Mathematical Physics*, **96** (1984) 15–57.

[13] P. Grassberger, "Towards a quantitative theory of self-generated complexity", *International Journal of Theoretical Physics*, **25** (1986) 907–938.

[14] P. Grassberger, "Long-range effects in an elementary cellular automaton", *Journal of Statistical Physics*, **45** (1986) 27–39.

[15] G. J. Chaitin, "Toward a mathematical definition of 'Life' " in *The Maximum Entropy Formalism* (edited by D. Levine and M. Tribus, MIT Press, Cambridge, MA, 1979).

[16] K. E. Eriksson and K. Lindgren, "Structural information in self-organizing systems", *Physica Scripta*, **35** (1987) 388–397.

[17] A. K. Zhvonkin and L. A. Levin, "The complexity of finite objects and the development of the concepts of information and randomness by means of the theory of algorithms", *Russian Mathematical Surveys*, **25** (1970) 83–124.

[18] S. K. Leung-Yan-Cheong and T. M. Cover, "Some equivalences between Shannon entropy and Kolmogorov complexity", *IEEE Transactions on Information Theory*, **24** (1978) 331–338.

[19] C. E. Shannon, "A mathematical theory of communication", *Bell System Technical Journal*, **27** (1948) 379–423.

[20] L. Brillouin, *Science and Information Theory* (2nd ed., Academic Press, New York, 1962) Chapter 3.

[21] J. P. Eckman and D. Ruelle, "Ergodic theory of chaos and strange attractors", *Reviews of Modern Physics*, **57** (1985) 617–656.

[22] P. Martin-Löf, "The definition of random sequences", *Information and Control*, **9** (1966) 602–619.

[23] R. Ash, *Information Theory* (Interscience, New York, 1965) Chapter 2.

[24] D. A. Lind, "Applications of ergodic theory and sofic systems to cellular automata", *Physica*, **10D** (1984) 36–44.

[25] C. H. Bennett, "The thermodynamics of computation—a review", *International Journal of Theoretical Physics*, **21** (1982) 905–940.