Thermodynamics and Complexity of Cellular Automata

Remo Badii[1] and Antonio Politi[2]
[1]*Paul Scherrer Institut, Villigen, Switzerland*
[2]*Istituto Nazionale di Ottica, I-50125 Firenze, Italy and INFN, Sezione di Firenze, Firenze, Italy*
(Received 9 September 1996)

The complexity exhibited by cellular automata is studied using both topological (graph-theoretical) and metric (thermodynamic) techniques. A novel topological classification, based on a hierarchy of languages, is introduced. In particular, it is shown that the elementary rule 22 is able to produce, upon iteration, a deep nesting of grammatical rules and that this asymptotically yields a phase transition when the thermodynamic formalism is applied to the limit spatial configuration. [S0031-9007(96)02188-6]

PACS numbers: 05.45.+b, 02.50.Le, 05.70.–a, 87.10.+e

Deterministic cellular automata (CAs) exhibit a remarkable ability to produce intricate, although not necessarily random patterns, even when acting upon a nearly uniform initial condition [1]. The apparent contrast between this behavior and the elementary form of the dynamical rules constitutes a challenge to the definition and characterization of complexity [2]. The structure of CAs makes them a fertile ground for testing conjectures and new methods of analysis both in a physical and in a computational context. On the one hand, in fact, they mimic more realistic models, such as partial differential equations and coupled-map lattices, while yielding quicker simulations; on the other hand, the discreteness of space, time, and field variables assimilates them to Turing machines [3], so that application of computational tools is straightforward.

Some of the schemes proposed to classify CAs can be viewed as complexity measurements. In particular, dynamical behavior lying "between" ordered (periodic) and chaotic evolution is sometimes regarded as "complex." In spite of all efforts made to formalize this conjecture [4], however, the very existence of complex cellular automata is still in doubt. Rigorous methods, in fact, are seldom usable in practice because of uncomputability problems [5]. Others, based on mean-field approximations [6], demonstrate the utility of a physical approach. Notwithstanding their different origin and motivation, these procedures present several interconnections to such an extent that, even without seeking exact results, much progress can be made by carefully combining them in the analysis.

In the present paper, we present a novel graph-theoretical technique for the characterization of generic one-dimensional symbol sequences which stresses the hierarchical nature of the underlying language. We then apply it to the limit set of the "elementary" CA 22, a one-dimensional, nearest-neighbor rule over two symbols which has gained the fame of being indeed complex, especially because of the difficulty of estimating its metric entropy [7]. For this reason, we also evaluate its "thermodynamic properties" through the entropy spectrum of the limit set using the nearest-neighbor method [8]. Our main findings are an unexpected grammatical structure of the spatial configuration even at finite times, given a random initial condition, and strong numerical evidence for the existence of a phase transition in the asymptotic spatial configuration, which appears to exhibit an infinite nesting of grammatical rules.

We study symbolic sequences of the type $S = s_1 s_2, \dots, s_n$, with $s_i \in \{0, 1, \dots, b-1\}$ and $n \gg 1$, such as those produced by a dynamical system endowed with a b-element phase-space partition [9] or by a one-dimensional CA. The latter consists of a dynamical rule which updates synchronously the variables $s_i(t)$, for all $i \in \mathbb{Z}$, where t is the discrete time [1]. Spatial configurations S at fixed time t, possibly with $t \gg 1$, will be studied through the properties of the associated language $\mathcal{L}$, the set of all finite subwords of S. In order for a statistical analysis and, in particular, for the thermodynamic formalism [10] to be applicable to the spatial patterns, these must be stationary (as it is always true with rule 22 whenever the initial condition is also stationary).

In Ref. [11], two indicators have been proposed to characterize the topological complexity of $\mathcal{L}$. The former, $C^{(1)}$, was identified with the topological entropy K_0, which represents the exponential growth rate of the number $N(n)$ of words of length n in $\mathcal{L}$ [12]; the latter, $C^{(2)}$, was defined as the topological entropy of the set $\mathcal{F}$ of all irreducible forbidden words (IFW) of S: These are words which do not belong to $\mathcal{L}$, although all their proper subwords do. Hence, denoting with $N_f(n)$ the number of IFWs of length n,

$$C^{(2)} = \lim_{n\to\infty} [\ln N_f(n)]/n \,. \quad (1)$$

Since $N_f(n)$ is, at most, equal to $bN(n-1)$, $C^{(2)} \leq C^{(1)}$ and the passage from $\mathcal{L}$ to $\mathcal{F}$ in the description of S represents, *de facto*, a compression of information. At this stage, the maximum degree of complexity is attained when no compression takes place, i.e., when the equality holds. On the contrary, $C^{(2)} = 0$ implies simplicity, since a compact description of the topology

 0031-9007/97/78(3)/444(4)$10.00

of $\mathcal{L}$ is achieved through $\mathcal{F}$. In particular, all systems with a finite $\mathcal{F}$ (called subshifts of finite type) belong to this class, random signals corresponding to the extreme situation $\mathcal{F} = \emptyset$. In general, $C^{(2)}$ can be seen to measure the difficulty of approximating the language $\mathcal{L}$ through subshifts of finite type with increasing memory (length of the IFWs).

Strictly positive $C^{(2)}$ can be found already in the class of regular languages (akin to Markov processes in a physical language). An example is given by the set $\mathcal{L}'$ of words that do not contain any expression of the type $101(1 + 00)^*101$, where the asterisk means arbitrary concatenations of the words 1 and 00. Indeed, none of these forbidden sequences contains any other one, because of the delimiter word 101, so that they are true IFWs belonging to a set $\mathcal{F}'$. After recovering $\mathcal{L}'$ from $\mathcal{F}'$, one finds $C^{(1)} \approx 0.6374$ and $C^{(2)} \approx 0.4812$, in agreement with the supposed inequality between the two exponents.

We now extend this scheme by supplying a recursive procedure which makes the approach truly hierarchical. The clue is to realize that the IFW language $\mathcal{F}$ can be analyzed in the same way as the original language $\mathcal{L}$. More precisely, we propose to determine the set [which we indicate with $\mathcal{F}(\mathcal{F})$] of all irreducible prohibitions found in $\mathcal{F}$ and to iterate the procedure on $\mathcal{F}(\mathcal{F})$ and on its possible descendants. Special care is required, however. In fact, $\mathcal{F}$ is neither factorial nor transitive; i.e., concatenations of words in $\mathcal{F}$ do not belong to $\mathcal{F}$ by definition, and, if u, $w \in \mathcal{F}$, then $uvw \notin \mathcal{F}$, $\forall v$. Languages $\mathcal{L}$ which originate from dynamical systems, instead, enjoy these two properties. The language $\mathcal{F}'$ above, for example, has a factorial, transitive component (given by all possible concatenations of 1 and 00) and "one-way" parts, corresponding to the prefix/suffix 101. In general, therefore, knowledge of all prohibitions in $\mathcal{F}$ [i.e., of $\mathcal{F}(\mathcal{F})$] does not permit unambiguous reconstruction of $\mathcal{F}$ itself. Notwithstanding this and the presumable lack of a general theory that covers all languages, the hierarchy $\mathcal{L} \to \mathcal{F} \to \mathcal{F}(\mathcal{F}) \to \ldots$ can be descended whenever the main source of diversity in each of its members consists of a finite number of factorial, transitive components. When this holds, the complexity of $\mathcal{F}$ originates from its own IFWs and a new index $C^{(3)}$ can be properly defined as the topological entropy of the language $\mathcal{F}(\mathcal{F})$.

This is, in particular, possible for regular languages, the words of which can be generated by following all paths on a finite graph G [13]. There is, in fact, a well-defined procedure to construct a graph G_f, the "dual" of G, which reproduces the language $\mathcal{F}$: As a consequence, $\mathcal{F}$ is also regular [2]. In general, G_f contains transient parts (arising from the lack of factoriality and transitivity of $\mathcal{F}$) as well as disjoint ergodic components (usually associated with different nodes of the original graph). The exponent $C^{(2)}$ is just the largest of the topological entropies of such components. Since each of them is a finite graph itself, the procedure can be iterated, thus obtaining a cascade of languages characterized by decreasing topological entropies $C^{(d)}$ ($d = 1, 2, \ldots$). We conjecture that, for each regular language, there exists a finite d_t such that $C^{(d_t+1)} = 0$; i.e., d_t is the number of nested hierarchical levels in the language $\mathcal{L}$ and may be seen as the "topological depth" of the language.

The elementary CA rule 22, defined by $s_i(k+1) = 1$ if $s_{i-1}(k) + s_i(k) + s_{i+1}(k) = 1$ and $s_i(k+1) = 0$ otherwise (k being the discrete time), illustrates the above ideas. Let us apply it on a random initial condition (which corresponds to a regular language with $\mathcal{F} = \emptyset$). It is known that any elementary cellular automaton yields, after k steps, a spatial configuration S_k in the class of the regular languages if the initial condition S_0 also belonged to it [1]. The number of IFWs, however, may increase in time. Indeed, the analysis of the language $\mathcal{L}_1$ obtained after one iteration of rule 22 reveals the existence of four hierarchical levels with topological exponents $C^{(1)} \approx 0.6508$, $C^{(2)} \approx 0.5297$, $C^{(3)} \approx 0.4991$, and $C^{(4)} \approx 0.1604$.

No rigorous results are instead known for the limit set $\Omega^{(\infty)}$ (the intersection of the sets $\Omega^{(k)}$ of all spatial configurations surviving after k steps, for $k = 0, \ldots, \infty$) of rule 22. Since, however, the size of the graph associated with S_k is a rapidly increasing function of k, one expects that both the topological depth $d_t = d_t(k)$ and the exponents $C^{(d)}(k)$ also increase with k. Rather than attempting a troublesome investigation of the convergence properties of the sequence $\{C^{(d)}(k)\}$ for increasing k, we have preferred to attack the more meaningful problem of constructing a hierarchical description of the limit set which, in the present context, is tantamount to the identification of all IFWs for increasing length. Unfortunately, one is immediately faced with a fundamental limitation: At variance with the case of dynamical systems with known generating partition [14], there is no algorithm which may determine all forbidden sequences in the limit set $\Omega^{(\infty)}$ of a generic automaton in a finite time. One could imagine to proceed as follows: Given a CA rule, all the preimages of a test sequence S of length n are computed for k backward iterates (their length being $n + 2rk$ in elementary automata with range r). Then, S is forbidden if an empty set is found for a finite k. This procedure, however, is not guaranteed to halt since there is no bound, in principle, to the smallest k which is necessary to reach in order to assess the legality of S. In practice, however, the detection of forbidden words is not too hard since knowledge of the previously identified ones may be used to speed up the computation. Usually, a few iterates are sufficient to identify the "short" prohibitions.

A more fundamental difficulty is represented by sequences that are only "asymptotically" forbidden. Let $P(S, n; k)$ denote the occurrence probability of a sequence S of length n in the kth image of a uniformly random initial configuration. Then,

$$P(S,n;k) = N_p(S,n;k)b^{-(n+2rk)}, \tag{2}$$

where $N_p(S,n;k)$ is the number of kth order preimages of S and the factor $b^{-(n+2rk)}$ represents the Lebesgue measure of each preimage. Then, S must be considered asymptotically forbidden if $P(S,n;k) \to 0$ for $k \to \infty$.

The existence of such sequences is easily verified for rule 22: $S' = 10101$, e.g., has only one preimage of order k, consisting of $5 + 2k$ alternating 0s and 1s. Accordingly, the probability of observing 10101 after k iterations of the rule is $P(10101, 5; k) = 2^{-5-2k}$, which vanishes exponentially for large k. Further asymptotically forbidden words are 010110001 and 100110011.

The preimages method enabled us to determine all IFWs up to length 25 in a spatial configuration S_k generated by rule 22 in $k = 10\,000$ iterations from a random initial condition. The results, illustrated in Fig. 1, yield $C^{(2)} \approx 0.545$. Before comparing this value with the topological entropy K_0, let us shift to the general framework offered by the thermodynamic formalism which provides a more complete characterization of symbolic signals. From the probability $P(s,n)$ of each subsequence of length n, one defines the local metric entropy $\kappa(S) = -\ln P(S,n)/n$ [8] and considers the number $dN_n(\kappa')$ of such sequences with $\kappa \in (\kappa', \kappa' + d\kappa')$. Setting $N_n(\kappa) \sim e^{ng(\kappa)}$ for $n \to \infty$, one may interpret κ as the thermodynamic energy E of a chain of n spins, and $g(\kappa)$, the "entropy spectrum," as the corresponding thermodynamic entropy $S(E)$ [2,10]. The function $g(\kappa)$ has been estimated for the asymptotic configuration S from the dimension spectrum $f(\alpha)$ [8] in the space of symbolic sequences S [the interval $x \in [0,1]$, with $x(S) = \sum_i s_i 2^{-i}$] using the relation $\kappa = \alpha \ln 2$ between κ and the local dimension α. We have analyzed $M = 10^6$ symbols of S using the fixed-mass method [15]. The local entropies were computed in 2×10^4 randomly chosen neighborhoods, each containing a mass $p = m/n_p$. Histograms of the entropies were constructed for $n_p \in [10^5, 9.8 \times 10^5]$ and $m \in \{20, 40, 60, 80\}$. The curve in Fig. 2 refers to $n_p \approx 8 \times 10^5$, $m = 40$, and represents the typical shape of the curves obtained for all pairs (m, n_p). The curves are most reliable in the region around the value $\kappa = K_1 \approx 0.51 \pm 0.01$ of the metric entropy, which has been estimated independently. More accurate results can be obtained with a finite-size scaling analysis. The topological entropy K_0, in particular, has been determined either as

$$K_0^{(a)}(n) = \ln \frac{N(n)}{N(n-1)}, \text{ or as } K_0^{(b)}(n) = \ln \mu_{\max}(n), \tag{3}$$

where $\mu_{\max}(n)$ is the largest eigenvalue of the transition matrix for the graph G that reproduces all prohibitions up to length n. Assuming an exponential convergence of the form $K_0^{(i)}(n) \sim K_0 + a_i e^{-\eta_i n}$, we have estimated the topological entropy to be $K_0 = 0.55 \pm 0.01$ and $\eta \approx 0.08$ in both cases (see Fig. 3). The agreement between the two approaches strengthens the validity of the numerical findings [16]. It is important to notice that $C^{(1)}$ is very close to $C^{(2)}$, which hints at the maximal degree of order-two complexity for the asymptotic sequence S.

Moreover, the approximately linear behavior of $g(\kappa)$ vs κ for $0.25 < \kappa < 0.5$ is suggestive of a (first-order) phase transition. This graph is, indeed, analogous to an entropy-energy diagram in which subsystems with local energies in a finite range coexist at the same temperature $T = (dS/dE)^{-1}$. Analogous curves are obtained for well-known dynamical systems [8]. The slow convergence of the thermodynamic functions (finite-size estimates show that η is indeed small for other generalized entropies K_q [8] as well) confirms this conjecture. Within the error bounds, we may presumably conclude that the presence of the phase transition is related to the validity of the exact equality $C^{(1)} = C^{(2)}$, which could not be directly assessed.

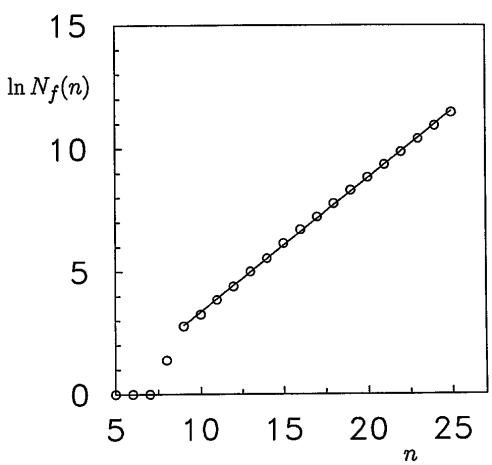

FIG. 1. Natural logarithm of the number $N_f(n)$ of IFWs of length n for the asymptotic spatial configuration of the elementary CA rule 22 vs n. The fitted slope is $C^{(2)} = 0.545$.

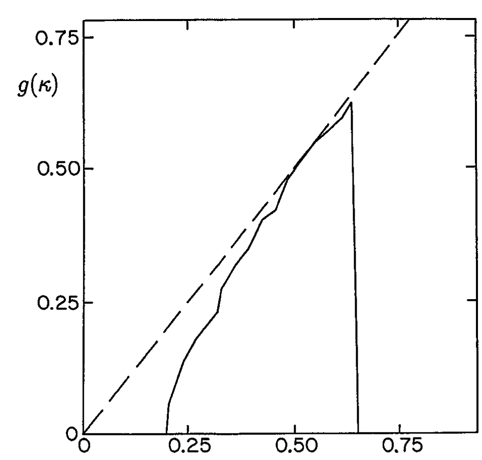

FIG. 2. Entropy spectrum $g(\kappa)$ vs κ for the elementary CA rule 22.

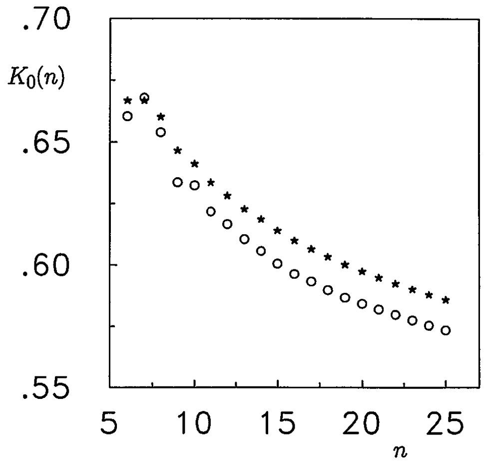

FIG. 3. Finite-size estimates $K_0(n)$ of the topological entropy K_0 for the asympototic spatial configuration of the elementary CA rule 22. Stars: logarithm of the largest eigenvalues of the directed graph constructed from all forbidden words of length $\ell \leq n$. Circles: logarithm of the ratio $N(n)/N(n-1)$ between the numbers of legal words of lengths n and $n-1$.

The reliability of the $g(\kappa)$ spectrum of Fig. 2 is further confirmed by the agreement of the boundaries $\kappa_{\min}$ and $\kappa_{\max}$ of its support with various approximate estimates. An upper bound to the minimum local entropy $\kappa_{\min}$ is given by the decay rate $\kappa^{(0)}$ of the probability $P(000\ldots)$ of a sequence composed of n zeros. This is the most frequently observed one (the spatio-temporal pattern, in fact, presents a large number of triangles filled with 0s) and, hence, it asymptotically dominates all others. The abundance of such triangles and the existence of a simple recognition algorithm for them permit achievement of reliable results even for large values of n. Careful simulations indicate that $\kappa^{(0)} = 0.267$ is indeed very close to $\kappa_{\min}$. An analytic lower bound to the topological entropy can be obtained by realizing that rule 22, when applied four times on an arbitrary concatenation $\tilde{S}$ of the two words $w_0 = 0000$ and $w_1 = 0001$, simulates the effect of rule 90 on the configuration S' obtained from $\tilde{S}$ by substituting each w_0 with 0 and each w_1 with 1 [3]. Since rule 90 is "linear," it prohibits no sequence, and the topological entropy of its limit set is $\ln 2$. Hence, all subsequences of $\tilde{S}$ in rule 22 contribute to the topological entropy with $(\ln 2)/4 \approx 0.173$. The value represents a lower bound to K_0 only if all such subsequences are really contained in the invariant set of rule 22, as it is confirmed by numerical simulations.

In this work, we have introduced a method which allows one to identify a hierarchy of nested levels of grammatical structure in a generic symbolic sequence. Its application to the language generated by a single iteration of rule 22 reveals four different levels. Investigation of the limit set of the same rule indicates that the spatial configuration is a good candidate for a second-order maximally complex language. The implications of this finding in the thermodynamic setting have been demonstrated with a direct estimate of the entropy spectrum which appears to exhibit a phase transition.

[1] S. Wolfram, *Theory and Applications of Cellular Automata* (World Scientific, Singapore, 1986).
[2] R. Badii and A. Politi, *Complexity: Hierarchical Structures and Scaling in Physics* (Cambridge University Press, Cambridge, England, 1997).
[3] S. Wolfram, Commun. Math. Phys. **96**, 15 (1984).
[4] C. G. Langton, Physica (Amsterdam) **22D**, 120 (1986).
[5] M. Hurley, Ergod. Th. Dynam. Syst. **10**, 131 (1990).
[6] H. A. Gutowitz, Physica (Amsterdam) **45D**, 136 (1990).
[7] G. Fahner and P. Grassberger, Complex Syst. **1**, 1093 (1987).
[8] P. Grassberger, R. Badii, and A. Politi, J. Stat. Phys. **51**, 135 (1988).
[9] V. M. Alekseev and M. V. Yakobson, Phys. Rep. **75**, 287 (1981).
[10] Ya. G. Sinai, Russ. Math. Surv. **27**, 21 (1972); D. Ruelle *Thermodynamic Formalism* (Addison-Wesley, Reading, MA, 1978).
[11] G. D'Alessandro and A. Politi, Phys. Rev. Lett. **64**, 1609 (1990).
[12] R. L. Adler, A. G. Konheim, and M. H. McAndrew, Trans. Am. Math. Soc. **114**, 309 (1965).
[13] J. E. Hopcroft and J. D. Ullman, *Introduction to Automata Theory, Languages, and Computation* (Addison-Wesley, Reading, MA, 1979).
[14] A. Politi, in *From Statistical Physics to Statistical Inference and Back*, edited by J.-P. Nadal and P. Grassberger (Kluwer, Dordrecht, 1994), p. 293.
[15] R. Badii and A. Politi, J. Stat. Phys. **40**, 725 (1985); R. Badii and G. Broggi, Phys. Lett. A **131**, 339 (1988).
[16] At variance with generic dynamical systems, where the graph provides better approximations [17], the more stringent upper bound to K_0 is here provided by the direct method.
[17] G. D'Alessandro, P. Grassberger, S. Isola, and A. Politi, J. Phys. A **23**, 5285 (1990).