Complex Systems 1 (1987) 1093–1098

Entropy Estimates for Dynamical Systems

Gerald Fahner
Peter Grassberger
Physics Department, Bergische Univerität, Gesamthochschule Wuppertal, Gauss-Strasse 20, Postfach 100127, 5600 Wuppertal 1, West Germany

Abstract. We apply a method proposed recently for estimating entropies of symbol sequences to two ensembles of binary sequences obtained from dynamical systems. The first is the ensemble of (0,1)-sequences in a generating partition of the Henon map; the second is the ensemble of spatial strings in cellular automaton 22 in the statistically stationary state. In both cases, the entropy estimates agree with previous estimates. In the latter case, we confirm a previous claim that the entropy of spatial strings in rule 22 converges to zero, although extremely slowly.

One of the most important characteristics of low dimensional chaotic dynamical systems is their entropy. In spatially extended systems, we have in addition spatial strings whose non-vanishing entropy is an important observable. In the first case, the entropy is most easily obtained indirectly as the sum over the positive Lyapunov exponents [1,2] (for attractors; for repellers, see reference [3]), provided the equations of motion are given analytically. If not, measuring the entropy can be non-trivial if block-entropies (defined below) converge slowly. The latter is, in particular, the case for some cellular automata, where either spatial strings at fixed time or temporal strings at fixed space points can have extremely slowly converging block entropies [4].

Given an ensemble of strings made up of symbols $s_i \in \{0,1\}$, with probabilities $p[s_1, s_2 \dots s_n]$ for finding blocks $(s_1 s_2 \dots s_n) \in \{0,1\}^n$ at arbitrarily chosen positions, block entropies H_n are defined as

$$H_n = -\sum_{s_i=0,1} p[s_1 \dots s_n] \log_2 p[s_1 \dots s_n]. \tag{1.1}$$

The entropy h is defined as

$$h = \lim_{n\to\infty} h_n, \quad h_n = H_{n+1} - H_n. \tag{1.2}$$

Since the number of terms in equation (1.1) increases exponentially with n, it is practically unfeasable to compute the H_n for large n ($n \gtrsim 20$, in our two examples and with our computational resources). If equation (1.2) does not converge sufficiently fast, computing the entropy via block entropies is thus not possible.

The purpose of the present letter is to apply a method which was recently proposed by one of us [5]. It is related on the one side to the Lempel-Ziv [6] entropy estimates, and on the other hand to dimension estimates via next-neighbor distances introduced by Badii and Politi [7] and others [8].

Consider a one-sided infinite string $S = (s_1 s_2 \ldots)$, its one-sided infinite substrings $S_i = (s_i s_{i+1} \ldots)$, $i = 1, 2, \ldots N$, and the substrings $S_{i,k} = (s_i \ldots s_{i+k-1})$ consisting of the first k digits of S_i. Effectively, the algorithm finds, for each i, the largest k such that $S_{i,k} = S_{j,k}$ for some other $j \in [1, N]$. In other words, it finds the nearest neighbor of S_i within the set of all S_j, $1 \le j \le N$, in a suitable metric.

Let us denote by $\langle k_{\max}^{(N)} \rangle$ the average maximal k,

$$\langle k_{\max}^{(N)} \rangle = \frac{1}{N} \sum_{i=1}^{N} k_{i,\max}, \quad k_{i,\max} = \max_j \{k | S_{j,k} = S_{i,k}\}. \tag{1.3}$$

Then, we have

$$h = \lim_{N \to \infty} \eta_N, \quad \eta_N = \frac{\log_2 N}{\langle k_{\max}^{(N)} \rangle} \tag{1.4}$$

and the somewhat more complicated but faster convergent expression

$$h = \lim_{N \to \infty} h(N), \quad h(N) = \frac{1}{\langle k_{\max}^{(N)} - k_{\max}^{(N/2)} \rangle}. \tag{1.5}$$

Actually, the algorithm finds the nearest neighbors by constructing a binary tree (see also [9]) in which each S_i corresponds uniquely to one path from the root to one of the leaves. The time required for searching for a nearest neighbor is reduced in this way from $O(N)$ (as in a naive search [10]) to $O(\ln N)$. The sequences analyzed in the present work had N up to 4×10^5. The CPU time required for one sequence of 4×10^5 binary digits was ~ 2.4 minutes on a μVAX, and the storage demand was 12 MB of (virtual) memory.

The Henon map is given by

$$x_{n+1} = 1 + b y_n - a x_n^2, \quad y_{n+1} = x_n. \tag{1.6}$$

We took the parameter values $a = 1.4$ and $b = 0.3$, for which the Lyapunov exponent is $\gamma = 0.4192 \pm 0.0001$.

For computing the entropy directly, one needs a generating partition of the plane. We use the binary partition proposed in [11]. Its division line is a polygon which passes through points of homoclinic tangencies. Applying the above algorithm to the (0,1)-sequence obtained in this way yields the results

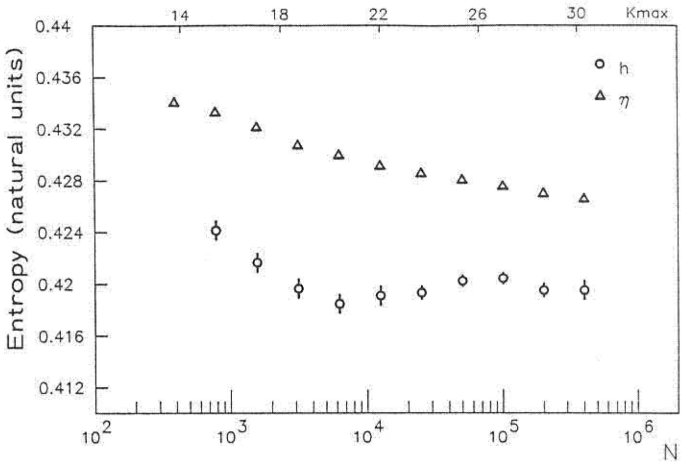

Figure 1: Entropy estimates (in natural units) for the Henon map with $a = 1.4$, $b = 0.3$. The data are from 40 runs with 4×10^5 iterations, 80 runs with 2×10^5 iterations, and 110 runs with 10^5 iterations each. The circles represent $h(N)$ (equation (1.5)); the triangles are η_N (equation (1.4)). The upper horizontal scale indicates $\langle k_{\max}^{(N)} \rangle$. Error estimates are not given when the error is smaller than the symbol.

shown in figure 1. We see that the $h(N)$ in equation (1.5) converge quickly to $h = 0.4194 \pm 0.0007$, in good agreement with the expectation $h = \gamma$. The values of η_N have much smaller error bars but converge less fast so that they are less useful.

This result may be contrasted with previous entropy estimates [11] using the same partition but using equation (1.2). The block probabilities needed there were simply obtained by counting the frequencies of all different sub-strings. Even with much higher statistics than in the present paper (ca. 10^8 vs. less than 4×10^7), it was not possible to go beyond a block length of 20. In our present approach, we have $\langle k_{\max}^{(N)} \rangle \approx 30$. This is the maximal correlation length to which our algorithm is sensitive with the present statistics. The increase by ten units causes a more reliable estimate of the entropy.

The second system attacked is the one-dimensional cellular automaton rule 22. The evolution of an initially random string is given by the local rule that each digit s_i has to take on the value '1' if and only if exactly one of s_{i-1}, s_i, and s_{i+1} in the previous generation was '1'; otherwise, s_i becomes '0'. All the s_i, $i = 1, 2, \ldots, N$, are updated simultaneously and periodic boundary conditions are used. The lattice sizes were $(1-4) \times 10^5$. The first 500 iterations were discarded as not yet stationary; after that, every 500$^{\text{th}}$ generation was analyzed. The results are shown in figure 2.

For our longest sequences of $N = 4 \times 10^5$ spins, the average $k_{\max}^{(N)}$ is about 24. This is to be compared to the maximal feasible block length of 17 units

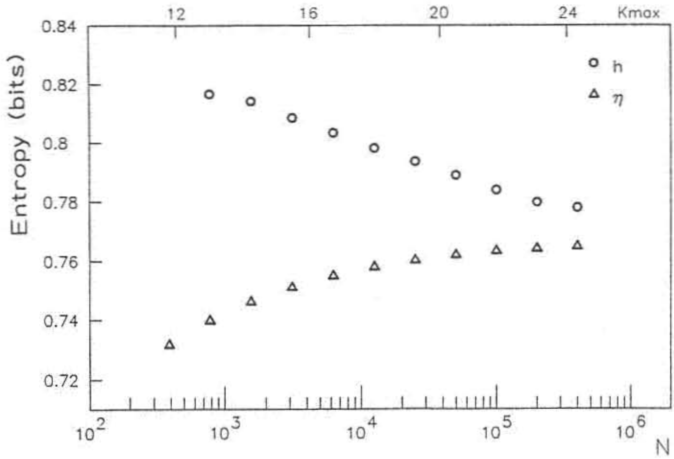

Figure 2: Entropy in bits for the one-dimensional cellular automaton rule 22. Here, we used 40 runs with 4×10^5 iterations and 80 runs with 10^5 iterations each. Error estimates are given as in figure 1.

in [4]. The entropy decreases very steadily with N, or rather, with $\langle k_{\max}^{(N)} \rangle$, as already observed in reference [4]. Our best estimate is $h \lesssim 0.778$ bits/spin, but there seems to be further decrease, indicating extremely long-ranging correlations. The best fit is obtained with the scaling law

$$h(N) \sim \langle k_{\max}^{(N)} \rangle^{-0.084}. \tag{1.7}$$

This agrees perfectly with the estimate of reference [4], provided we treat $\langle k_{\max}^{(N)} \rangle$ as an effective block length. It indicates that $h(N)$ might tend toward 0 for $N \to \infty$ and that the spatial strings in the stationary case actually are not random.

Recently, there has been much interest in generalized entropies and in treating fractal measures as multifractals [12]. In our case, generalized entropies $h^{(q)}$ are defined by

$$\lim_{N \to \infty} N^{1-q} \langle 2^{(q-1)h^{(q)} k_{\max}^{(N)}} \rangle = 1. \tag{1.8}$$

The crowding index (pointwise entropy, "singularity") α is defined as

$$\alpha(S_i) = \log_2 N / k_{i,\max} \tag{1.9}$$

and $f(\alpha)$ is defined as usual as the Hausdorff dimension of subsequences with crowding index α. It can be obtained either by Legendre transforming the function $(q-1)h^{(q)}$, or by using directly the numbers $m_{N,k}$ of subsequences with $k_{i,\max} = k$:

$$m_{N,k} = \frac{\text{const}}{\sqrt{k}}\, e^{kf(\alpha)} \tag{1.10}$$

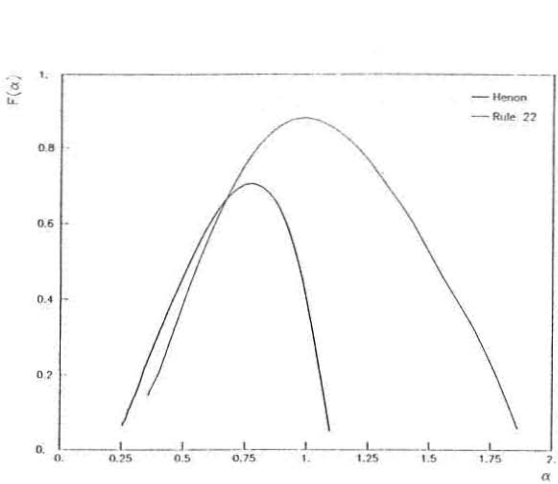

Figure 3: Entropy $f(\alpha)$-scaling functions for the Henon map (heavy line) and for rule 22 (light line), both obtained from equation (1.10). For the figure, we used strings of length 4×10^5; curves obtained from strings of length 10^5 are virtually identical. (The graph for the Henon map obtained from 10^5 iterations is slightly broader than that from 4×10^5, an effect of the smaller statistics.)

where const is fixed in practice by demanding $f(\alpha) = \alpha$ for $f'(\alpha) = 1$. Using the latter method, our results for $f(\alpha)$ for both systems are shown in figure 3. Their most striking feature is that $f(\alpha)$ is much wider for cellular automaton 22 than for the Henon map. We should, of course, point out that $f(\alpha)$ for the cellular automaton 22 is indeed ill defined if its entropy vanishes asymptotically, but due to the slow convergence, this does not effect the numerical estimate of $f(\alpha)$.

We conclude that the presented method yields entropy estimates in good agreement with previous and independent results. Compared to the standard method of counting frequencies of blocks of fixed length, its main advantage is that it needs much shorter strings to get estimates which take into account possible long-range correlations. In particular, we verified that long-range correlations are much more important in cellular automaton 22 than in the Henon map.

References

[1] Ya. B. Pesin, *Russ. Math. Surv.*, **32** (1977) 55.

[2] J. P. Eckmann and D. Ruelle, *Rev. Mod. Phys.*, **57** (1985) 617.

[3] R. Bowen and D. Ruelle, *Invent. Math.*, **29** (1975) 181; H. Kantz and P. Grassberger, *Physica* **17D** (1985) 75.

[4] P. Grassberger, *J. Stat. Phys.*, **45** (1986) 27.

[5] P. Grassberger, preprint WU B 87-11 (1987).

[6] A. Lempel and J. Ziv, *IEEE Trans. Inform. Theory*, **22** (1976) 75; **23** (1977) 337; **24** (1978) 530.

[7] R. Badii and A. Politi, *Phys. Rev. Lett.*, **52** (1984) 1661.

[8] A. Brandstæter et al., *Phys. Rev. Lett.*, **51** (1983) 1442; J. Guckenheimer, *Contemp. Math.*, **28** (1984) 357.

[9] M. Rodeh, V. R. Pratt, and S. Even, *J. ACM.*, **28** (1981) 16.

[10] F. Kaspar and H. G. Schuster, *Phys. Rev.*, **A36** (1987) 842.

[11] P. Grassberger and H. Kantz, *Phys. Lett.*, **113A** (1985) 235.

[12] R. Benzi, G. Paladin, G. Parisi, and A. Vulpiani, *J. Phys.*, **A17** (1984) 3521; T. C. Halsey et al., *Phys. Rev.*, **A33** (1986) 1141; J. P. Eckmann and I. Procaccia, *Phys. Rev.*, **A34** (1986) 659.