PHYSICAL REVIEW

LETTERS

VOLUME 64 | 2 APRIL 1990 | NUMBER 14

Hierachical Approach to Complexity with Applications to Dynamical Systems

G. D'Alessandro and A. Politi

Istituto Nazionale di Ottica, I-50125 Firenze, Italy

(Received 20 December 1989)

A hierarchical approach to complexity of infinite stationary strings of symbols is introduced by investigating the scaling behavior of suitable quantities. The topological entropy, which estimates the growth rate of the number of admissible words, corresponds to the first-order indicator $C^{(1)}$. At the second level, a novel indicator $C^{(2)}$ is introduced which measures the growth rate of the number of irreducible forbidden words. Finally, a detailed analysis of 2D maps reveals that $C^{(2)}$ can be expressed in terms of the Lyapunov exponents.

PACS numbers: 05.45.+b

Many efforts have been recently made to better understand dynamical behaviors intermediate between perfect deterministic predictability and complete randomness. The problem, in its simplest and nevertheless intriguing version, can be formulated as that of characterizing the complexity of an infinite sequence of symbols which, without losing generality, are assumed to be the "letters" 0 and 1.

Evolution of chaotic systems, for instance, can always be described in this way by using a generating partition to encode a generic orbit. The state of a 1D cellular automaton and DNA sequences belong, by definition, to the same class of problems as well.[1] However, even remaining confined to 1D sequences, the state of the art is still very preliminary: A satisfactory indicator of complexity has not yet been introduced (for a review on the subject, see Ref. 2). Grassberger noticed that a meaningful definition of complexity should allow one to recognize a purely random sequence as a simple pattern, because of the lack of rules behind its generation. As a consequence, approaches like those introduced by Kolmogorov[3] and Chaitin[3] or Lempel and Ziv[4] should be considered as insufficient, as they end up reproposing the concept of entropy, which associates a larger number to purely random strings. While agreeing with the previous request, we also point out that a plain sequence of 0's should anyhow be recognized as simpler than a random one. In fact, loosely speaking, many more inspections must be made to realize that a given sequence belongs to the class of random, rather than periodic, strings. Before entering a detailed discussion, let us note that we restrict our investigations to the problem of learning the rules of an *a priori* unknown language, leaving aside the problem of reproducing a sequence in terms of an *a priori* known set of rules. This because we think that prior to reproducing a sequence it is necessary to understand it. For the sake of simplicity, we also limit ourselves to describing topological properties, thus ruling out a number of definitions (like those given by Grassberger[5] and by Crutchfield and Young[6]) which deal with metric properties.

To overcome the above-mentioned difficulty, we propose a two-step procedure, and suggest the possibility of a future introduction of higher-order steps. Indeed, we conjecture that the learning process of a given infinite string of symbols is hierarchically organized, each step being naturally associated with a suitable indicator of complexity. At the first level, the "size" of the vocabulary is investigated, while at the second level the "size" of the underlying grammar is analyzed. Since the most general class of languages one can expect to find is presumably characterized by an infinite vocabulary, grammar, etc., we propose to define each complexity in terms of the scaling rate of the "sizes," rather than in terms of the "sizes" themselves.

More precisely, having defined the number $N_a(n)$ of admissible subsequences of length n, we introduce the first-order complexity $C^{(1)}$ as the asymptotic growth rate

$\ln N_a(n)/n$. It is nothing but the topological entropy, in agreement with the common belief that the entropy contributes to defining the concept of complexity. This indicator allows one to discriminate between periodic and random strings, while it fails to recognize random strings as simple objects. Such a task is accomplished at the next level, where we look for possible regularities of a given language. We think that the most objective way of defining the notion of a rule for finite-length words is that they should not contain irreducible forbidden words, where irreducible means that the word does not include any shorter forbidden sequence (FS). Analogously to the first level, we then define a second-order indicator $C^{(2)}$ as the growth rate of the number $N_f(n)$ of FS's,

$$C^{(2)}=\lim_{n\to\infty}\frac{\ln N_f(n)}{n}\,. \tag{1}$$

This *Ansatz* is suggested by the fact that the most general scaling behavior of $N_f(n)$ should be of the same type as the growth rate of admissible words, with the only obvious limitation $C^{(2)}<C^{(1)}$. In fact, we do not see any reason why $N_f(n)$ should exhibit a slower than exponential growth rate. As a consequence, it is at this stage that we discover that a purely random sequence is simple: No FS's at all are discovered and $C^{(2)}$ is 0. For generic chaotic dynamical systems, instead, we expect $C^{(2)}$ to be nonzero, as a finite Markov partition (i.e., a finite grammar) typically does not exist. The definition of $C^{(2)}$ is similar to the one suggested by Badii[7] who, however, restricts his analysis to periodic orbits organized in a suitable logical tree. Wolfram has given another analogous definition in terms of the number of nodes in the associated graph;[8] the main difference is that he assumes that all the rules are known *a priori*.

The introduction of a second level allows to compress the information required to characterize a given string, passing from the set of admissible sequences to the smaller set of irreducible FS's. The next step requires, in principle, to determine possible rules concatenating FS's of increasing length. However, a precise formulation of the problem is still lacking, and here we limit ourselves to raising a very preliminary question. Is a language, characterized by a set of randomly chosen irreducible FS's, to be considered as the most complex or as the simplest one, among all languages exhibiting the same $C^{(2)}$?

In the second part of this Letter, we apply the previous analysis to 2D maps, showing that

$$C^{(2)}=\lambda_+D/(1+D)\,, \tag{2}$$

where multifractal corrections have been neglected (usually yielding corrections of a few percent), and where λ_+ is the positive Lyapunov exponent, while D is the fractional part of the attractor's dimension. Before deriving Eq. (2), we consider the simple case of a 1D map, $y_{n+1}=F(y_n)$, of logistic type. Here it is not possible to have more than one irreducible FS of length n. By denoting the coordinate of the maximum with x_0 and its ith iterate with x_i, a trajectory is encoded by associating a 0 or a 1 to y_n, depending whether the point y_n belongs to $I_1\equiv(x_0,x_1)$ or $I_2\equiv(x_2,x_0)$, respectively. The existence of FS's follows from the fact that I_2 is, in general, not exactly mapped onto $I_1\cup I_2$. If x_0 does not belong to $F(I_2)=[x_3,x_1]$, there is one forbidden sequence of length 2 (namely, 00). To determine the FS's of length 3, it is sufficient to iterate $I_3\equiv F(I_2)$. If x_0 belongs to $F(I_2)$, there are no FS's of length 2, and we can write $F(I_2)$ as the union of two intervals ($[x_0,x_1]$, $[x_3,x_0]$) of which the first needs not be iterated, since its extrema are previously found iterates of the maximum. In this case we can find the FS's of length 3 by iterating $I_3\equiv[x_3,x_0]$. By applying the above sketched procedure to I_j, we can find a possible FS of length j and determine the next interval I_{j+1}. Finally, the existence of at most one FS of length j follows from the need to iterate only one interval at each step. Therefore, according to definition (1), 1D maps, whatever the dynamical behavior they exhibit, are never complex. This surprising conclusion follows from our definition of $C^{(2)}$ as an exponential growth rate. Anyway, this is not to be considered as a real drawback, since 1D maps represent too crude an approximation of physical systems, as they do not guarantee invertibility of the dynamical behavior. The more realistic 2D maps are instead associated with a nonzero complexity, which consistently vanishes for $D\to 0$, i.e., in the limit of a one-dimensional evolution.

Equation (2) will be derived with reference to a simple model. The extension to more general systems naturally follows. We investigate the map introduced by Tél,[9] which belongs to the class of Henon-type maps

$$x_{n+1}=y_n,\quad y_{n+1}=f(y_n)+bx_n\,, \tag{3}$$

with $f(y)=ay-\mathrm{sgn}(y)$, where $\mathrm{sgn}(y)$ denotes the sign of y. In the limit case $b=0$, the map reduces to the 1D Bernoulli shift. We have chosen the Tél map, since the dynamics along the stable and unstable manifolds are really coupled together, while the expansion and contraction rates are independent of the position in phase space. The attractor can be schematized as an infinite collection of oblique segments organized in an almost self-similar manner (see Fig. 1). Open trajectories are encoded according to the sign of y (0 for $y<0$ and 1 for $y>0$).

To discover the rules hidden in the language of this chaotic system, we construct progressive 1D approximations of the map, by using the infinite collection of segments which compose the unstable manifold of any periodic orbit of map (3) (for the sake of simplicity we have chosen the period-2 cycle P_1,P_2 shown in Fig. 1). If we iterate a number h of times the two segments passing through P_1,P_2, we find $N_h\simeq e^{\lambda_+h}$ segments, each one corresponding to a symbol sequence $\{S_i(h)\}$ which codes the last h steps of its history. Further iterates of such segments cluster around them in stripes of width

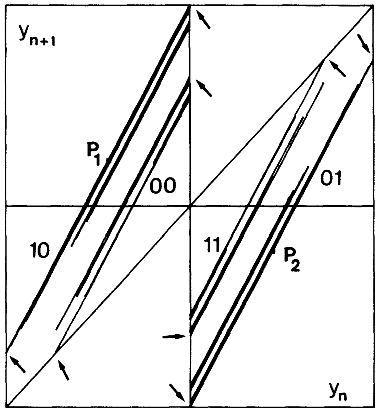

FIG. 1. Attractor of the Tél map (3), reconstructed for $a=1.4$ and $b=0.3$ from the unstable manifold of the period-2 cycle (P_1,P_2). The pairs of symbols indicate the four stripes composing the attractor at the order $h=2$. The arrows indicate the two extrema of each segment B_i which are iterated to determine the FS's.

$\delta_h = e^{\lambda_- h}$ (where λ_- is the negative Lyapunov exponent). If we assume that distances smaller than δ_h cannot be resolved, then we can approximate the attractor with segments B_i having for extrema those of the corresponding stripes (see Fig. 1). This static approximation can be straightforwardly transformed into a dynamical one, by observing that a point (x,y) of B_i is mapped onto B_j, where j is such that $\{S_j(h)\}$ is obtained by shifting $\{S_i(h)\}$ and adding one new symbol, determined according to the actual sign of y. In other words, the evolution in phase space can be interpreted as a one-dimensional rule, chosen among a finite series of possibilities, according to a code which is deterministically evaluated from the past history. Indeed, knowledge of the code of the initial segment allows one to determine y_{n+1} from y_n, while the sign of y_{n+1} plus the initial code yields the new code.

Before discussing the accuracy of this approximation, let us investigate the properties of the language generated by the above map. To detect the FS's, it is necessary first to determine the iterates of the extrema of each B_i, discovering whether they allow both a 0 and a 1, and then to repeat the same procedure for the iterates. However, analogously to the logistic map, where only the segment containing the nth iterate of the extremum x_1 of I_1 is to be considered, here only the images $B_i(n)$ containing the iterates of the initial extrema have to be tested. Indeed, all the other segments generated after n iterates coincide, within a distance δ_h, with some previously generated $B_i(n')$ $(n'<n)$. Therefore, we can at most expect as many FS's for each n as the number of the initial extrema. One can reasonably expect that the fraction of segments that do not break up remains finite for $n\to\infty$, so that the number of FS's is asymptotically of the same order as N_h. Moreover, as any FS is obtained by adding a single bit to an admissible sequence, it appears reasonable to conjecture that the probability that a FS is reducible remains strictly smaller than 1 in the limit of infinite length. Accordingly, the order of magnitude of the number of irreducible FS's is given by N_h.

We are now in the position to discuss the accuracy of our approximation. Being the extrema of each segment B_i are defined as the lowest and the highest among all points of the corresponding stripe, their forward iterates yield the largest possible segment lengths, thus preventing the recording of spurious FS's. The relevant difference from the true map arises whenever the iterate of a B_i is to be broken up into two parts (associated with different symbols), and the intercept with the x axis is to be computed. The finite width $\delta_h = e^{\lambda_- h}$ of the stripes induces an analogous uncertainty on the intercept itself which is amplified according to the positive Lyapunov exponent, when iterated. The indeterminacy is nothing but a memory effect due to the bits "older" than h time steps, and is crucial when it becomes of the same order as the distance from the x axis, preventing a conclusive decision on the existence of a FS. As the average distance from the x axis is ~ 1, the 1D map can be at most iterated k times, with k given by

$$k = -(\lambda_-/\lambda_+)h\,, \tag{4}$$

before grossly failing to determine the correct FS's. As the approximate map is built from all the segments of word length h, all FS's up to this length are automatically given, and the first k iterates allow one to reach a maximum length $n=h+k$. As discussed before, such a map presents an average number $e^{\lambda_+ h}$ of irreducible FS's, a number which can now be interpreted as the expected number of FS's of length n in the true map. By expressing h as a function of n, through Eq. (4), we finally find

$$N_f(n) = \exp\left(\frac{\lambda_+ n}{1-\lambda_-/\lambda_+}\right), \tag{5}$$

which leads to Eq. (2), recalling the Kaplan-Yorke relation for the fractional part of the dimension $D=\lambda_+/|\lambda_-|$. Equation (2) indicates that, keeping the positive Lyapunov exponent (i.e., the topological entropy) fixed, the language becomes more and more complex for increasing dimension, and it reaches the maximum for a conservative map where $C^{(2)}$ equals $\lambda_+/2$.

In order to check Eq. (2), we have performed accurate numerical simulations on the Tél map. To begin with, we have determined the first two segments of the unstable manifold of its period-2 cycle. Then, they have been iterated enough times to guarantee an accurate determination of the maximum and minimum among all points lying in the same stripe (the result of the procedure is sketched in Fig. 1 for $h=2$). As a final step, the extrema of the B_i's have been iterated, looking for possible crossings with the x axis. We then checked if each FS, obtained in the case of no crossing, was really

TABLE I. Comparison of the numerical ($C^{(2)}_{\text{num}}$) and theoretical [$C^{(2)}_{\text{th}}$, from Eq. (2)] estimates of complexity. Letters T and L indicate Tél and Lozi maps, respectively; a and b indicate the parameter values, while D is the fractional part of the dimension, and λ_+ is the positive Lyapunov exponent.

Map	a	b	λ_+	D	$C^{(2)}_{\text{th}}$	$C^{(2)}_{\text{num}}$
T	1.4	0.3	0.463	0.278	0.101	0.101 ± 0.005
T	1.8	0.1	0.617	0.211	0.108	0.110 ± 0.005
T	0.85	0.8	0.347	0.609	0.131	0.129 ± 0.007
L	1.6	0.4	0.414	0.311	0.098	0.105 ± 0.009
L	1.7	0.3	0.504	0.295	0.115	0.123 ± 0.01

irreducible. The results for three sets of parameter values and an approximation order $h=15$ are reported in Table I, and indicate good agreement with the theoretical estimates.

The same procedure has also been applied to the Lozi map [$f(y)=a-1-a|y|$], where multifractal fluctuations are present. The segments B_i have been determined by iterating the unstable manifold of the fixed point which belongs to the attractor. The results for two sets of parameter values are reported in Table I, again confirming formula (2) within numerical error and possible corrections due to the multifractal structure. As a consequence, all such results strongly suggest the generic charcter of Eq. (2). If cases with $C^{(2)}$ smaller than the value expressed by Eq. (2) certainly exist, we conjecture that they have zero Lebesgue measure in the parameter space.

In this Letter we have defined the first steps of a hierarchical approach to the problem of complexity, in terms of forbidden words of increasing length. We have applied such a definition to 2D maps, deriving a relation between the second-order complexity $C^{(2)}$ and Lyapunov exponents. By no means do we claim that this hierarchical procedure, even when it will be fully set up, completely exhausts the whole problem of characterizing complexity, which is a subject far from being understood. However, we have shed new light on a series of questions, not only introducing a new classification scheme, but also proposing a different way of looking at the properties of the many physical systems that can be described in terms of 1D sequences of symbols. One obvious generalization of this scheme goes, for instance, in the direction of accounting for the probability of each sequence, defining "metric complexities" as well.

Finally, we have also introduced a new approximation scheme for 2D maps which is particularly powerful in the derivation of long-length FS's and which, consequently, allows one to compute the topological entropy in the most precise way.

[1] *Measures of Complexity*, edited L. Peliti and A. Vulpiani (Springer-Verlag, Berlin, 1988).

[2] (a) P. Grassberger, in *Measures of Complexity* (Ref. 1); (b) D. Zambella and P. Grassberger, Wuppertal Report No. B 11, 1988 (to be published).

[3] (a) A. N. Kolmogorov, Probl. Inf. Transm. Engl. Transl. **1**, 1 (1965); (b) G. Chaitin, J. Assoc. Comput. Mach. **13**, 145 (1966).

[4] A. Lempel and J. Ziv, IEEE Trans. Inf. Theory **22**, 75 (1976).

[5] P. Grassberger, Int. J. Theor. Phys. **25**, 907 (1986).

[6] J. P. Crutchfield and K. Young, Phys. Rev. Lett. **63**, 105 (1989).

[7] R. Badii, in Proceedings of the Workshop on Quantitative Measures of Dynamical Complexity in Nonlinear Systems, Bryn Mawr, Pennsylvania, June 1989 (to be published).

[8] S. Wolfram, Commun. Math. Phys. **96**, 15 (1984).

[9] T. Tél, Phys. Lett. **97A**, 219 (1983).