

On an Approach to the Estimation of the Complexity of Curves (Using as an Example an Electroencephalogram of a Human Being)

B. S. Darkhovskii*, **A. Ya. Kaplan****, and **S. L. Shishkin****

**Institute for Systems Analysis, Russian Academy of Sciences, Moscow, Russia*

***Lomonosov State University, Moscow, Russia*

Received September 27, 2001

Abstract—A new approach to the estimation of the “complexity” of a continuous curve on a section is suggested. The basic idea of the approach consists in that the complexity of the curve must be estimated by a relative share of information, which is necessary for its recovery with a specified accuracy from a set of values in the finite number of points by means of a prescribed aggregate of methods. The suggested approach is illustrated by an estimate of the complexity of various fragments of an electroencephalogram (EEG) of a human being and can be used, in particular, for the macrostructural and microstructural analysis of the EEG.

1. BASIC IDEA

The notion of a “complexity” (or “simplicity”) of one object or another relates to the class of fundamental scientific notions. This problem was taken up by many prominent scientists, and there are numerous attempts to employ the “complex” approach for applied problems, including the analysis of an EEG of a human being.

Let us recall in short the basic idea involved with the notion of the complexity. It seems likely that the quantitative approach to the notion of a “complex physical system” first appeared in statistical physics when the notion of “entropy” arose. In the equilibrium statistical physics, by the entropy is meant a coefficient with the asymptotics of the logarithm of a number of configurations displaying certain properties when the number of the degrees of freedom of a system tends to infinity. The higher the entropy, the more complex the system.

In the Shannon theory of information, created in the 1940s, the notion of the entropy of probability distributions is introduced. It is established that the entropy is a certain measure of the “uncertainty degree” inherent in one or another probability distribution, and this measure is unique under natural conditions. It can be shown that the Shannon entropy is a coefficient with the asymptotics of the logarithm of the number of “typical” sequences of independent random quantities that have one and the same current distribution of probabilities as the length of the sequence increases (typical sequences are those sequences for which the deviation of the frequencies of events from their probabilities does not exceed a specified small number). It can be said again that the probability distribution is more “complex,” the higher is its entropy (from this viewpoint, the most complex distribution is a uniform one because it has the highest possible uncertainty).

The theory of dynamic systems, developed by Kolmogorov and Sinai (see, for example [1]), comes to the notion of the entropy of these systems. In essence, this is an extension of the Shannon entropy to dynamic systems. The entropy of a dynamic system is a coefficient with the asymptotics of the logarithm of the number of various types of trajectories of the dynamic system when the time tends to infinity. Here, again, the entropy of a dynamic system can serve as a measure of its complexity: the more complex the system, the greater the variety of its trajectories.

The notion of the entropy of a dynamic system (and such “nonlinear characteristics” related to the entropy as the correlation integral, the fractal dimension, the Lyapunov exponential curve, see, for example [2]) is used in the so-called nonlinear, or “chaotic,” dynamics. One of the problems of this science is the recovery of the description of a dynamic system from its implementations. Here again, the complexity of a system is identified with the complexity of its description, i.e., in the final analysis, with the entropy.

In the early 1980s, Kolmogorov [3] suggested a new algorithmic approach to the notion of the complexity of objects. The basic idea of this approach consists in that a “complex” object requires a large body of information for its description, and a “simple” one a small body of information. This idea is formalized in the language of the theory of algorithms, and the complexity, broadly speaking, is measured by the length of a program, which leads to the separation of the given object from a certain set.

It is this Kolmogorov idea that is the closest one to the idea that we intend to outline here. But before we proceed to do this, we will consider how “complex” approaches to the analysis of an EEG are used.

As is known, an EEG signal relates to the most complex physical signals. First of all, this is due to a high (and principal) unsteadiness of an EEG [4]. In view of the EEG unsteadiness, the methods based on the adjustment of various models to the fixed EEG are not fit in principle for the description of its new sections. The phenomenon of the unsteadiness led to the need for the preliminary segmentation of an EEG into (quasi)-stationary segments so as to select for each segment its own mathematical model after this segmentation. For purposes of the segmentation, attempts were taken to employ the notion of the complexity in order to separate in the EEG signal the fragments that are homogeneous in “complexity.”

The notions of the complexity used in the literature of the EEG analysis rely on the ideas presented above, primarily on the entropy (or close “nonlinear characteristics”) of dynamic systems and the Shannon entropy. We will note the main drawbacks of similar approaches. The basic assumption of the nonlinear dynamics (the theory of dynamic systems, the theory of determinate chaos) is the assumption of the fact that the dynamic system under study is stationary, i.e., does not change its properties with time. It is this assumption that forms the basis of the modern, rather developed, mathematical theory, which has numerous applications in physics and other fields. The disturbance of this condition renders the use of the above-mentioned approaches impossible, but it is the unsteadiness that is almost the main sign of an EEG signal. Here, we note that even for the stationary case, the problem of recovery of the description of a dynamic system from its observable trajectories is rather complex and is rather far from its solution in the practical sense (we recommend the reader to become familiar with an interesting and informative article [2], where, probably, the practical aspects of this problem were first dealt with at a rather strict mathematical level).

Thus, an attempt to use the entropy notions of complexity as a method of describing the dynamic system generating an EEG signal, in our opinion, is inconsistent *a priori*. On the other hand, the use of the classical Shannon entropy for the characterization of an EEG signal, in our opinion, does not also fully conform to those purposes which could be stated in dealing with such a signal. The thing is that the Shannon entropy specifies the complexity of the probability distribution or, as noted above, the abundance of “typical” trajectories of the sequence of independent random quantities. Even if we abstract ourselves from the unsteadiness problem, such a characteristic can only indirectly be used for measuring the complexity of the description of a signal itself, but it is the description complexity that is, in our opinion, the property by which we could specify various EEG segments.

We will pass on directly to the idea of our approach to the notion of the EEG complexity and, more generally, to the notion of the “complexity of a continuous curve.” Let us prescribe in the section $[0, T]$ a continuous function $x(t)$ (curve). We would wish to put in correspondence to this curve a certain number that specifies its complexity. We will proceed from the Kolmogorov idea of the fact that a complex object requires for its description a large amount of information, and a simple object a small amount of information. We will select a number $h > 0$ and consider the subdivision of the section $[0, T]$ by points $\{t_i\}$ into $n = [T/h]$ equal parts (here $[a]$ is the integral part of the number a). Let us assume that we measured our curve only at subdivision points t_i . At what accuracy can we recover the entire curve from this information? Let us prescribe a certain set \mathcal{F} of methods of the approximation of the curve according to the finite set of points (ideally, this is the set of all known methods of the recovery of the curve by its values in the finite number of points). We will consider an approximation error $\varepsilon(h) = \inf_{\mathcal{F}} \|x(t) - \hat{x}(t)\|$, where $\hat{x}(t)$ is an approximating curve plotted by one of the admissible approximation method over n given points, $\|\cdot\|$ is a certain integral norm in the functional space, while the infimum is taken relative to all the set of the specified approximation methods.

It is clear that if $x(t) \equiv \text{const}$, then for its recovery it is sufficient to use one point; in other words, the constant is the simplest curve, which quite agrees with the intuition. Therefore, it is expedient to eliminate the constants from the set of the curves of interest, for which purpose it is necessary to carry out a preliminary centering of $x(t)$. Next, to make it possible to compare with one another the approximation errors of various curves, we need to consider all of them at one scale, i.e., to carry out the preliminary normalization. For this purpose, after the centering, we should consider the normed curve $z(t) = x(t)/\|x(t)\|$ (it is obvious that $\|z(t)\| = 1$). After the centering and the normalization of the curve, the approximation error (the integral one) will not exceed 1, in which case a maximum error appears if the discreteness interval is $h = T$ (this means that we must use only one constant for the approximation, so that the best constant will be the zero one from consideration of symmetry; here, the error will prove to be equal to the norm of the initial curve, i.e., 1).

We will now plot a graph of the function $\varepsilon(h)$, increasing the quantity h (i.e., decreasing the number of the current n). It is clear that with a growth of h , the function $\varepsilon(h)$ must steadily increase (in fact, an increase in the discreteness step means that we select a more and more amount of information on the function and, hence, its approximation become worse and worse if the set of the approximation methods is fixed). In view of the presented remarks, the error $\varepsilon(h)$ will be a monotone positive function of its argument, which reaches its maximum equal to 1 at $h = T$. If now we prescribe a certain “acceptable” level of the approximation error $\Delta < 1$ (from the viewpoint of a specific user), then it is possible to define what admissible body of information that can be discarded without an appreciable loss of the quality of the curve recovery.

Now, the complexity (more exactly, $(\Delta, \mathcal{F}^{\|\cdot\|})$ -complexity) of the given specific curve $x(t)$ can be called the quantity $1 - h^*/T$, where $h^* = \min\{h \leq T : \varepsilon(h) = \Delta\}$. In other words, this is a relative share of information that is necessary for the recovery of the given curve with the specified accuracy by the prescribed set of methods.

Let us note that the suggested measure of the complexity is an individual characteristic of a particular curve rather than a certain set of curves generated by a certain mechanism, as this takes place with the use of the entropy of dynamic systems. Next, this measure is in no way related to possible mechanisms of the generation of the curve (for example, to the fact whether the curve is a piece of the implementation of any random process or a trajectory of any dynamic system). This circumstance appears to be important, in particular, as applied to an EEG in view of the above—mentioned considerations as to the specific features of this signal. Generally speaking,

the suggested measure of the complexity is “adjusted” to the perception of the user, i.e., directly accounts for the fact whether it is easy or not to work with discrete data for the curve.

Note 1. It could be possible to attempt to specify the complexity of a curve by its Shannon entropy; namely, it is possible to break up the axis of values of the curve into a finite number of intervals, count the rates of the fall of values of the curve into one or another interval and then estimate the entropy of the obtained frequency distribution by the Shannon formula. But it is easy to plot the curves (for example, periodic ones) in which the entropy (at the fixed number of the intervals of the subdivision of the axis of values) with practically not change with a decrease in the number of discrete counts (if there are still rather many counts of this type), so that the dependence of such a measure on the share of the remaining information will be very weak. At the same time, the approximation error must be sufficiently sensitive to a decrease in the accessible information, especially for complex curves.

Note 2. The suggested complexity measure obviously depends on the adopted functional norm, on the class \mathcal{F} , and on the level Δ . It may so happen that with a change of these parameters, the complexity relation for one and the same pair of curves will change to the opposite one. However, this, in our opinion, is not a deficiency because the functional norm and the level Δ are criteria of the approximation quality, which a particular user chose for himself, while the class \mathcal{F} describes the algorithmic capabilities of this user. Moreover, to analyze such complex curves as EEGs, what is important is not a particular numerical value of the complexity, but the dynamics of its changes in passing from one portion of the curve to another. In principle, it is this dynamics that can offer the possibility of detecting important changes in an EEG, which can signify some functional rearrangements. From this viewpoint, we should admit such a complexity index of the most acceptable type for the analysis of an EEG as that which detects its changes in passing from one segment of the EEG to another; in this case the numerical expression of the index is unessential for a specific segment, and of importance is only the degree of its sensitivity to a change in the character of the curve.

Note 3. The possibility of calculating the complexity of the curve on a segment of an arbitrary length enables the researcher to select that time scale for which he would wish to perform the analysis of the entire observed implementation. A possibility appears to investigate the dynamics of the complexity of an EEG at various time scales and, thus, to analyze various structures of this curve, which can specify various functional states of the brain.

2. ALGORITHM

In the practical implementation of the idea outlined above, we need to account for the fact that the implementation of a curve is always preset in a discrete time, i.e., the “curve” represents a finite set of counts (a finite-dimensional vector). In addition, from considerations of the simplicity of the implementation, not too large a class \mathcal{F} should be chosen for the first experiments. Further, as noted above, all analyzable curves need be brought to one scale.

The discreteness of the initial curve leads to the fact that the dependence of an approximation error on the number of discarded points will not be monotone if there is a small number of the remaining points. Therefore, it is desirable to have a fairly high frequency of the inquiry so that even with the removal of a large percent of data (actually, in our experiments, this removal reaches 70–90%), at which a satisfactory approximation is still possible, the discreteness of the recording of the curve should have a small effect on the monotonicity of the behavior of the error.

To reduce the curves to one scale, we used the centering and the subsequent normalization (in the metric of ℓ_1) of the finite-dimensional vector (curve), with the result that we obtained the vector from an appropriate unit sphere.

Thus, if the curve was preset by the vector $x = (x_1, x_2, \dots, x_n)$, then the first operation of the algorithm consisted in the transition to the vector $y = x - \bar{x}$, where $\bar{x} = n^{-1} \sum_{i=1}^n x_i$, and the second operation consisted in the transition to the vector $z = y/\|y\|$, $\|y\| = \sum_{i=1}^n |y_i|$.

We measured the error of the approximation of the vector z by the vector \hat{z} in the same ℓ_1 -metric, i.e., $\|z - \hat{z}\| = \sum_{i=1}^n |z_i - \hat{z}_i|$.

The discreteness of the recording of the initial continuous curve leads to a small modification of the formula for the complexity. First, with a sufficiently large number $n^* < n$ of the points left after the rejection, the ratio h/T can be replaced by the quantity $1/n^*$. Second, it makes sense to reason that if $n^* = n$ (i.e., one cannot discard any point from the initial recording in view of a substantial loss of the recovery quality), then the complexity of the curve is equal to 1. In view of these remarks, we used for the estimation of the complexity s of the curve, the formula $s = \frac{1-1/n^*}{1-1/n}$, where, we recall, n^* is the number of points left after the discarding, at which the curve still recovers with the acceptable accuracy, and n is the number of points in the initial recording of the curve.

3. EXPERIMENTAL RESULTS

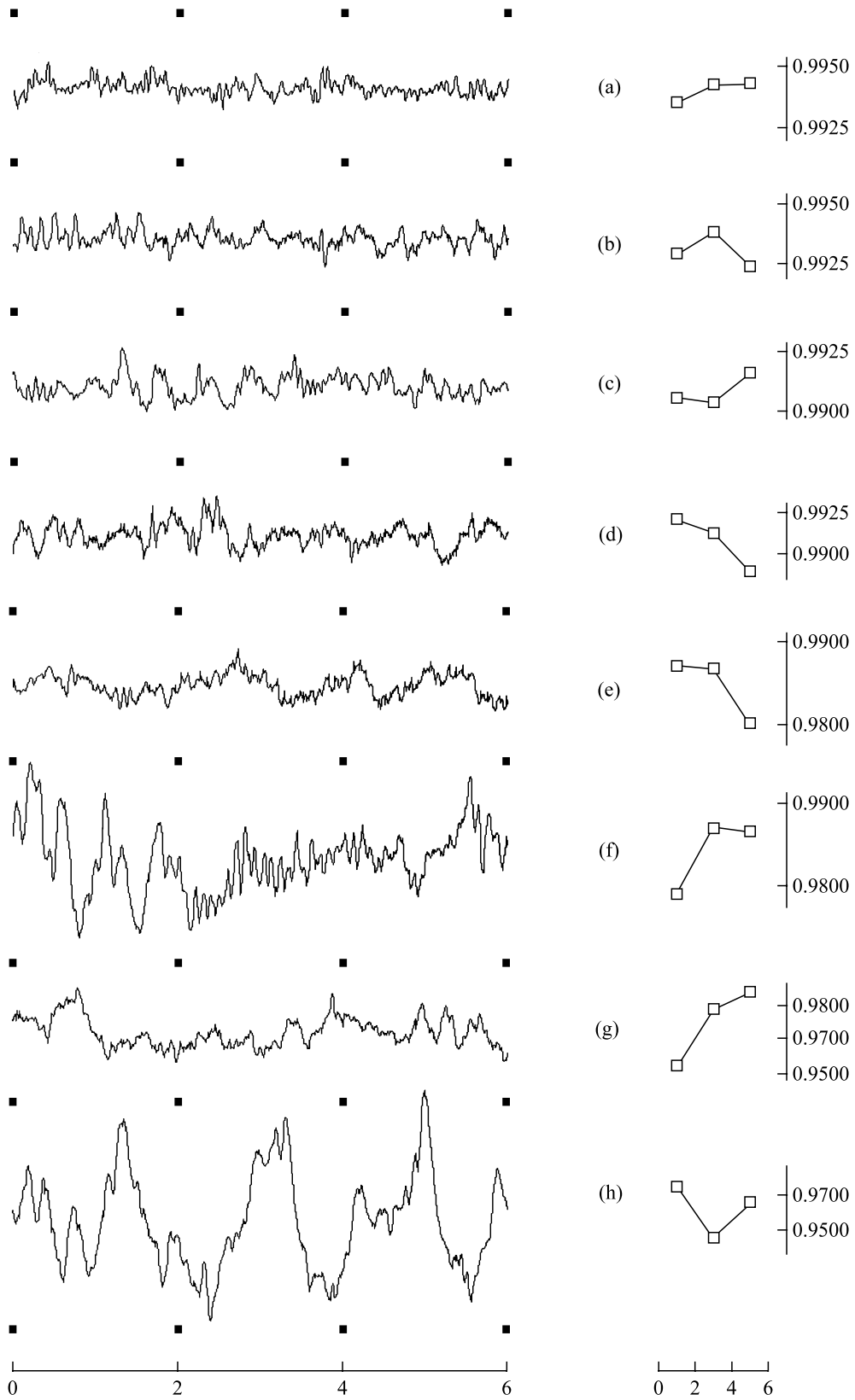
The class \mathcal{F} in the first version of our algorithm contained merely one method—the approximation of the curve by a step function. In other words, with the removal of some points from the initial set, we replaced values of the curve at the removed point by its value at the preceding point. In view of the discreteness, in the search for the best approximation, it is necessary to carry out the exhaustive search relative to the value at the initial point despite the fact that we use only one approximation method. This exhaustive search is carried out within that “window of thinning” which is used at the given instant (i.e., at the prescribed share of discarded points).

Later on, it is assumed to include into the class \mathcal{F} the piecewise—polynomial functions and the approximation by a section of the Fourier series with respect to various orthonormal systems of functions.

The testing was performed on an EEG recorded on a human being during his sleep. What is typical for this EEG is a wide range of patterns corresponding to various states of neurons of the cortex of large cerebral hemispheres [5]. The EEG was recorded in the right occipital lead at a sampling frequency of 100 Hz.

In the course of the survey, 25 nonartifactual fragments of a length from 4 to 16 s were selected (the median—9 s), each of which visually appeared comparatively homogeneous, but differed in pattern from the remaining ones. The complexity index was estimated in sequential “windows”—sections of a time length of $T = 2$ s. A value of the approximation error Δ was set at a level of 0.32.

The variance of the complexity estimates in “windows” of the analysis within the homogeneous fragments of the EEG proved to be markedly lower than the variance of the estimates averaged over each of these fragments: the median of the “intrainterval” standard deviation was equal to 0.0019, whereas the standard deviation (in all fragments) of values averaged over the fragment was equal to 0.0083. The complexity estimates rather strongly differed among 6 fragments, which had the most distinguished traits of a mild or a paradoxical sleep (the mean is 0.993, the standard deviation is 0.0004), and four fragments with the most clearly expressed, powerful delta-activity



Fragments of an EEG (a)–(h) in order of decreasing complexity estimate and the dynamics of this estimate. The EEG is on the left, the dynamic of the complexity estimate is on the right. The boundaries of windows in which the complexity was estimated, are marked by points for convenience of the comparison of the EEG and the dynamics of the complexity. The time scale is at the bottom (in seconds).

typical for the third and the fourth stage of the sleep (the mean is 0.973, the standard deviation is 0.009).

The relation between the EEG pattern and complexity estimates, which showed up in these results, found an additional verification when we arranged the EEG fragments in order of decreasing mean estimate of the complexity (see Figure where in an effort to save room, we showed only each third fragment and only the first three windows each of a length of 2 s). The mean estimate of the complexity (over all windows) for the fragments presented in the Figure, denoted by the letters (a)–(h), was equal to 0.9934, 0.9926, 0.9917, 0.9911, 0.9875, 0.9830, 0.9794, and 0.9635, respectively. It was found that the deepening of the sleep occurs approximately in this order, too, if we judge by the EEG signs (see on the left in Figure). The dynamics of single complexity estimates (see in the right part of the figure), as a rule, also followed the changes in the EEG.

The suggested complexity measure described here for the first time, even in the simplest version of the algorithm implementation, proved to be quite a working tool sensitive to the EEG pattern. As one might expect, the complexity estimates fell in parallel with the shift of the EEG to the side of the typically sleepy pattern.

Let us note once again that the new complexity measure represents an individual characteristic of a particular curve and is not related in any way to possible mechanisms of its outcome, which is especially important in the case of an EEG, the conventional model of the generation of which does not exist at present. The possibility to calculate the complexity of a curve on the segment of an arbitrary length enables the researcher to study the dynamics of this index of an EEG at various time scales and, hence, to analyze the time hierarchy of various functional states of the brain.

We express our gratitude to Professor J. Roschke (the University of the city of Mainz, Germany) for the submitted recordings of the carolic EEG.

We are grateful to Yu.S. Popkov for attention to the work.

REFERENCES

1. Sinai, Ya.G., *Vvedenie v ergodicheskuyu teoriyu* (Introduction to Ergodic Theory), Moscow: Fazis, 1996.
2. Demidov, E.E., Darevskaya, Yu.V., Morenkov, O.A., *et al.*, *Nelineinyi korrelyatsionnyi analiz. Obozrenie prikladnoi i promyshlennoi matematiki* (Nonlinear Correlation Analysis. The Review of Applied and Industrial Mathematics), Moscow: TVP, 1999, vol. 6, issue 1, pp. 4–57.
3. Kolmogorov, A.N., Combinatorial Bases of the Theory of Information and Probability Calculus, *Usp. Mat. Nauk*, 1983, vol. 38, issue 4, pp. 27–36.
4. Kaplan, A.Ya., Unsteadiness of an EEG: Methodological and Experimental Analysis, *Usp. Fiziol. Nauk*, 1998, vol. 29, no. 3, pp. 35–55.
5. Steriade, M., Coherent Oscillations and Short-Term Plasticity in Coricothalamic Networks, *Trends Neurosci.*, 1999, vol. 22, no. 8, pp. 337–345.

This paper was recommended for publication by Yu.S. Popkov, a member of the Editorial Board