

ON THE NUMBER OF STATES OF THE SOURCES DEFINED AS ENCODED NEURON RESPONSES IN THE VISUAL CORTEX

J.M. Amigó¹, J. Szczepański², E. Wajnryb², M.V. Sanchez-Vives³

¹Instituto de Investigación Operativa. Universidad Miguel Hernández, Elche (Spain)

²Institute of Fundamental Technological Research, PAS, Warsaw (Poland)

³Instituto de Neurociencias, Universidad Miguel Hernández-CSIC, Alicante (Spain)

jm.amigo@umh.es, jszczepa@ippt.gov.pl, ewajnryb@ippt.gov.pl, mavi.sanchez@umh.es

Summary

In a previous paper (Sanchez-Vives *et al.* 2001) the authors applied the Lempel-Ziv complexity to the study of neural signals from an information-theoretical point of view. Among other results, it is shown there that this concept of complexity allows to characterize the output sensory responses to random and periodic stimuli *in vivo* and *in vitro*. In this paper we take advantage of this data to analyze further consequences of our mathematical model, this time concerning the number of states of the corresponding information source. The number of states, assuming the source to be stationary, ergodic and Markovian, is directly related to the stochastic dependence lag of the (encoded) output signal at consecutive discrete times and provides a measure of its autocorrelation. Again, the number of states differs depending on the kind of stimuli and the preparation of the neuronal system.

Introduction

The mathematical modelling of nervous systems is so old as the first physical and chemical models of the action potentials (Hodgkin and Huxley 1952). In fact, only four years after Shannon's seminal work (Shannon 1948), MacKay and McCulloch estimated the entropy of spike trains (MacKay and McCulloch 1952) in what was probably the first application of information theory to neurosciences. Since then, the theory of information has become a major tool in the mathematical approach to the nervous phenomena and, in particular, to the communication among neurons.

The transmission of information about the stimulus features in the brain processes takes place in multiple stages. We study these processes at the stage of the information being transmitted among neurons by trains of action potentials (Rieke *et al.* 1998). A neuron responds to external stimuli by changing the polarity of its axon, what results in a train of spikes which can be monitored. It is still a challenging problem to find out

what kind of encoding and decoding methods are used by the neurons and which are the properties of such biological encoders (Borst and Theunissen 1999). Important for us is the fact that the relation between stimuli and neuron responses is not one-to-one: the same input s can generate different outputs r_k . Hence, *deterministic models* of the neural signal production require the introduction of internal *states* in the neuron, analogously to what occurs in the definition of encoders (Ziv and Lempel 1978); given s , the actual response r depends additionally on the current state of the neuron. In the *stochastic models*, the interest focuses on such quantities as conditional probabilities $p(r_k|s)$, correlations and mutual information.

An essential role in studying properties of information sources is played by the pattern matching approach (Wyner *et al.* 1998). In particular, the complexity, as defined in (Lempel and Ziv 1976), is based on the generation ratio of new patterns along a discrete signal. A related quantity, the relative complexity, provides a lower bound for the compression ratio of the signal by optimal coding (Ziv and Lempel 1978), so that the higher the relative complexity of a signal, the more information it conveys. Moreover, if the source is stationary and has "good" statistical properties (specifically: ergodicity, which allows to calculate mean values as time averages), the relative complexity of a single output gives with high probability a very good estimate of the source entropy, which is the average information in bits/symbol generated by the source. Observe that, whereas entropy is a property of sources, complexity is a property of individual sequences.

Relative complexity has been extensively applied by the authors in (Sanchez-Vives *et al.* 2001) to characterize the sensory responses *in vivo* and *in vitro* of the primary visual cortex to different kinds of stimuli (sinusoidal modulation of luminosity, injection of sinusoidal and random currents). To our knowledge, this concept of complexity had not been used before for the analysis of neural discharges. Our results there show, for example, that the relative complexity of the outputs *in vivo* are higher (and hence convey more information) than *in vitro* for the same kind of stimuli. This paper builds on such results on complexity of spike trains to analyze further information-theoretical aspects of neural systems.

In our approach, we treat the ensemble of stimuli, the neuron (eventually, neuron net) and the coding of the spike trains (to be explained below) as a source generating discrete signals which contain information about the stimuli. In order to quantify the number of states of such sources, we use a method which is a modification of others already known in the literature (Ziv 1990): we compare the entropy estimated by means of the Lempel-Ziv relative complexity with the entropy of the source, assumed to be stationary, ergodic and finite order Markovian. This technique is then applied to the same cases studied in (Sanchez-Vives *et al.* 2001). For convenience we will talk eventually of periodic or random *in vivo* and *in vitro* sources, depending on the kind of stimulus and preparation. Our results show that (i) the number of states of periodic *in vivo* sources differs significantly from the number of states of periodic *in vitro* sources (in particular, the first number is essentially smaller than the latter when each interspike time is encoded according to its duration) and (ii) the random *in vitro* sources have also a number of states significantly different from (in general smaller than) the number of states of periodic *in vitro* sources.

Methods

In an experimental setting, continuous input signals are sampled at discrete times, so that they actually come into the numerical analysis as time series. On the other hand, neurons have a limited reaction time to external stimuli, which effectively puts a minimal time scale onto the input. Including more sampling times in the input signal below this threshold only rescales the autocorrelation function of the signal without changing the neuron response. If, furthermore, one tracks the propagation of neural signals along a neural layer or net, the outputs (spike trains) of single neurons become the inputs of other neurons of the layer through the synapses. Again, neuron responses do not discriminate input clusters whose interspike times are too small.

In information theory one talks of signal sources and messages. At this point we would like to consider a given stimulus together with the neuron layer as the signal source and the spike train as the message sent by it. Unfortunately, a spike train is not quite a message in the sense of information theory. In fact, spike trains are first given by real numbers generally corresponding to the absolute times of the spike occurrences. In order to transform them in *bona fide* messages, one needs to translate these sequences of real numbers into sequences of symbols from a finite so-called alphabet. This step is called the *codification* of the signal and the procedure, the *(en)coding*. Henceforth, always when we talk about spike trains as messages we mean that the signal has been previously codified. Needless to say that this kind of codification has nothing to do with the codification neurons might employ to transmit information via spike trains.

Codification can be made in different ways. The following codings are among the most natural one can think of:

1. *Bin coding.* Let $\tau_{\min}$ and $\tau_{\max}$ be the minimal and maximal *interspike* times, respectively, in the signal. Divide the interval $[\tau_{\min}, \tau_{\max}]$ in α subintervals $\Delta\tau_k$ of the same length, $1 \leq k \leq \alpha$. If τ_i is the time between the spikes s_i and s_{i+1} and $\tau_i \in \Delta\tau_k$, then assign to the spike s_i the letter a_k from an alphabet $A = \{a_1, \dots, a_\alpha\}$ of α letters (Dan *et al.* 1996). In this way, we get an α -nary message which length is the number of spikes. In particular, if $\alpha = 2$ the message will be binary, i.e. a string of 0s and 1s.
2. *Interval coding.* Let the first spike of a train occur at time 0 and the last one T time units later. The time interval $[0, T]$ is then subdivided in β subintervals Δt_j , $1 \leq j \leq \beta$, of the same length. In each subinterval Δt_j will occur a number N_j of spikes. If we assign to the interval Δt_j the number N_j , then we get a message of length β with not more than β different letters. In particular, if β is so large (or, equivalently, the length of the subintervals Δt_j is so small) that only one spike at most occurs in each time slot Δt_j , the message will be binary (Rieke *et al.* 1998, Zador 1998). A different binary coding, independent of the number of subintervals β , can be achieved if the subinterval Δt_j is coded by 0 or 1 according to whether it contains no or at least one spike, respectively. Whenever necessary, this particular method will be called *binary interval coding* to distinguish it from the *general interval coding*.

3. *Median coding.* Let μ be the median of the interspike times $\Delta\tau_k$. If τ_i is the time between the spikes s_i and s_{i+1} , then assign 0 to the spike s_i if $\tau_i \leq \mu$ and 1 otherwise (Rapp *et al.* 1994). Hence, median coding is binary.

Once a spike train has been codified into a discrete sequence of symbols, it can be viewed as emitted by an information source. Because of their practical relevance, we will focus in this paper on the finite-state sources of Markovian type. Specifically, let $\mathbf{x} = x_1x_2\dots x_n$ be a sequence of observable random variables taking on values in a finite set or "alphabet" $A = \{a_1, \dots, a_\alpha\}$ of size $|A| = \alpha$. The elements of A are called letters and each realization of $\mathbf{x}$ is called a message of length n . Similarly, let $\mathbf{s} = s_1s_2\dots s_n$ be a sequence of observable random variables (states) corresponding to $\mathbf{x}$, which take on values in another finite set S of size $|S| = \sigma$. One can imagine the message $\mathbf{x}$ as being generated by a source with a set of internal states S , the letter x_i occurring at time i , when the source is in the state s_i . An information source is said to be *finite-state* (with σ states) if the joint probability of $\mathbf{x}$ and $\mathbf{s}$ is given by (Gallager 1968, Ziv 1990)

$$P(\mathbf{x}, \mathbf{s}) = \prod_{i=1}^n p(x_i, s_i | s_{i-1})$$

where the initial state $s_0 \in S$ is assumed fixed and known, and $p(x_i, s_i | s_{i-1})$ denotes the joint probability of a letter x_i and a state s_i at time instant i given the previous state s_{i-1} at time instant $i-1$.

An important subclass of the finite-state sources is built by all *stationary, ergodic*, discrete time Markovian sources of finite order. That is, if

$$x_m^n := x_mx_{m+1}\dots x_{n-1}x_n \quad \text{and} \quad x_{-\infty}^n := \dots x_{n-1}x_n$$

then

$$p(x_i | x_{-\infty}^{i-1}) = p(x_i | x_{i-k}^{i-1}), \quad i = 1, 2, \dots, n \quad (1)$$

for some integer $k \geq 1$ called the *order* of the source. Eq. (1) means that the probability for the letter x_i at instant i depends only on the previous k letters: $x_{i-1}, \dots, x_{i-k}$. For this reason, k is also called the *lag* of the source. In this case, the set of states can be identified with the set of all k -tuples of the form

$$x_{i-k}^{i-1} = (x_{i-1}, x_{i-2}, \dots, x_{i-k}) =: s_{i-1}$$

and hence the number of states of such a source is α^k . In other words: to estimate the number of states of a stationary, ergodic, discrete time, finite state Markovian source, only its order k is needed.

Let $s_{i-1} = x_{i-k}^{i-1} \in A \times \dots \times A =: A^k$ ($1 \leq i \leq n$) denote the state of a Markovian source at time i . We denote by $\delta(x_i, u; s_{i-1}, s)$ the indicator function for $x_i = u$ and $s_{i-1} = s$ ($u \in A, s \in A^k$), i.e.

$$\delta(x_i, u; s_{i-1}, s) = \begin{cases} 1 & \text{if } x_i = u \text{ and } s_{i-1} = s \\ 0 & \text{otherwise} \end{cases}$$

Now let define

$$\begin{aligned} q_{\mathbf{x}}^k(u, s) &= \frac{1}{n} \sum_{i=1}^n \delta(x_i, u; s_{i-1}, s) \\ q_{\mathbf{x}}^k(s) &= \sum_{u \in A} q_{\mathbf{x}}^k(u, s) \\ q_{\mathbf{x}}^k(u|s) &= \begin{cases} q_{\mathbf{x}}^k(u, s)/q_{\mathbf{x}}^k(s) & \text{if } q_{\mathbf{x}}^k(s) > 0 \\ 0 & \text{if } q_{\mathbf{x}}^k(s) = 0 \end{cases} \end{aligned}$$

The $\alpha^k \times \alpha$ matrix whose entries are $q_{\mathbf{x}}^k(u, s)$ will be referred to as the k -th order Markov-type of $\mathbf{x}$ and will be denoted by $q_{\mathbf{x}}^k$. It is easy to prove that $q_{\mathbf{x}}^k$ can be viewed as a k -th order Markovian probability measure P_k for any $\mathbf{x}$.

We next define the k -th order empirical entropy as

$$H(q_{\mathbf{x}}^k) = - \sum_{s \in A^k} q_{\mathbf{x}}^k(s) \sum_{u \in A} q_{\mathbf{x}}^k(u|s) \log_2 q_{\mathbf{x}}^k(u|s)$$

An *order estimator* is then given by (Ziv 1990)

$$k^* = \min \left\{ k : H(q_{\mathbf{x}}^k) - \frac{1}{n} U_{LZ}(\mathbf{x}) \leq \lambda \right\}, \quad (2)$$

where $U_{LZ}(\mathbf{x})$ is the Lempel-Ziv codeword length of $\mathbf{x}$ (Ziv and Lempel 1978). This estimator has the following intuitive interpretation. We seek the smallest model order k for encoding $\mathbf{x}$, such that the empirical entropy will be sufficiently close (difference less than λn) to the codeword length associated with the Lempel-Ziv algorithm, which in turn serves as an estimate of the source entropy. Then, for any positive integer k ,

$$P_k(k^* \geq k) \leq \frac{1}{2^{\lambda n}}$$

i.e. the k -th order Markovian probability that k^* overestimates k vanishes exponentially with λn .

In the case of stationary, ergodic finite-state sources, which are the ones considered here, numerical experiments show that the *relative complexity* of $\mathbf{x}$, $c_{rel}(\mathbf{x})$, converges faster (with increasing n) to the source entropy than $U_{LZ}(\mathbf{x})$, so that we propose to use

$$k^* = \min \left\{ k : H(q_{\mathbf{x}}^k) - c_{rel}(\mathbf{x}) \log_2 \alpha \leq \lambda \right\} \quad (3)$$

as the order estimator instead of (2).

The *relative complexity* $c_{rel}(\mathbf{x})$ of the sequence $\mathbf{x} = x_1 x_2 \dots x_n$ is defined as follows. A block B_l of length l is just a segment of $\mathbf{x}$ of length l ($1 \leq l \leq n$), i.e. a subsequence of l consecutive letters of $\mathbf{x}$, say $B_l = x_{k+1} x_{k+2} \dots x_{k+l}$ ($0 \leq k \leq n-l$). Set

$$B_1 = x_1$$

and suppose that

$$B_1 B_2 \dots B_k = x_1 x_2 \dots x_{n_k}$$

where $B_1 B_2 \dots B_k$ denotes the juxtaposition of the blocks $B_1, B_2, \dots, B_k$ and $2 \leq n_k < n$. Define

$$B_{k+1} = x_{n_k+1} \dots x_{n_{k+1}} \quad (n_k + 1 \leq n_{k+1} \leq n),$$

such that B_{k+1} is the block of minimal length which does not occur in the sequence $x_1 \dots x_{n_k} \dots x_{n_{k+1}-1}$. Proceeding in this way, we obtain a *decomposition* of $\mathbf{x}$ in blocks, say

$$\mathbf{x} = B_1 B_2 \dots B_p \tag{4}$$

in which only the last block B_p can eventually be the same as a previous one. The *complexity* $c(\mathbf{x})$ of $\mathbf{x}$ is then defined as the number of blocks in the decomposition (4):

$$c(\mathbf{x}) = p$$

To get the *relative complexity* of $\mathbf{x}$, normalize $c(\mathbf{x})$ in the following way:

$$c_{rel}(\mathbf{x}) = \frac{c(\mathbf{x})}{n / \log_{\alpha} n} = \frac{p}{n} \log_{\alpha} n$$

Observe that both $H(q_{\mathbf{x}}^k)$ and $c_{rel}(\mathbf{x})$ depend on α through $\mathbf{x}$, so that the order estimator k^* will depend also on α . For the bin and interval codings, the curves $c_{rel}(\mathbf{x})$ versus α turn out to be convex and concave, respectively, with horizontal asymptotes. We call the levels of such asymptotes the *saturation values* of the relative complexity. If c_0 is a saturation value of $c_{rel}(\mathbf{x})$ and H denotes the source entropy, then it can be shown (Sanchez-Vives *et al.* 2001) that

$$H = c_0 \log_2 \alpha$$

in the case of bin coding, while $H = c_0$ if the interval coding is used.

Results

Experimental data was obtained from primary cortex recordings both *in vivo* and in brain slice preparations (*in vitro*). Intracellular recordings *in vivo* were obtained from anesthetized adult cats. For the preparation of slices, 2-4 month old ferrets of either sex were used. Action potentials were detected with a window discriminator and the time of their occurrence was collected at 100 kHz with a 10 μ sec resolution. The resulting time series were used for analyzing the neurons spiking responses. Concerning the stimuli, they were of three kinds:

1. *Periodic current.* Intracellular sinusoidal currents were injected *in vivo* and *in vitro*. The frequency of the waveform was 2 Hz and the intensity ranged between 0.2 and 1.5 nA.
2. *Periodic visual stimulation.* The visual stimulus consisted of a sinusoidal drifting grating presented in a circular patch of 3-5 degrees diameter, centered on the receptive field center (*in vivo*).

3. *Random current.* Random currents with different degrees of correlations were injected during the intracellular recordings from cortical brain slices (*in vitro*).

For details see (Sanchez-Vives *et al.* 2000a, 2000b).

Consider in this framework a neuron or neuron layer firing a spike train (output) as response to a given stimulus (input). As explained above, in order to apply the methods of the information theory to the outputs, it is necessary to codify them. After codification, the original output becomes a message $\mathbf{x} = x_1x_2...x_n$, where both the length n and the number of letters α depend in general on the particular coding. Correspondingly, we identify the source of $\mathbf{x}$ with everything preceding it, namely, the ensemble of all possible stimuli, the neuronal system and the coding method. We assume furthermore that this source is stationary, ergodic and finite-state Markovian of order k . Out of the several codification techniques already seen, we will consider only the two sharpest ones for our present purposes, which happen to be the bin and the (general and binary) interval codings. The relative complexity of the outputs have been taken from the saturations values reported in (Sanchez-Vives *et al.* 2001). On calculating the k -th order empirical entropy $H(q_{\mathbf{x}}^k)$ of a given sequence $\mathbf{x}$ and inserting its relative complexity $c_{rel}(\mathbf{x})$ in (3), one gets typically the following order estimators k^* .

1. **Bin coding** (with 256 bins)

k^*	In vivo	In vitro
Periodic	0	2
Visual	0	—
Random	—	3

2. **General interval coding** (with 5000 intervals)

k^*	In vivo	In vitro
Periodic	3	10
Visual	3	—
Random	—	2

3. **Binary interval coding** (with 4048 intervals)

k^*	In vivo	In vitro
Periodic	5	10
Visual	4	—
Random	—	2

Discussion

The mathematical model we use to study the transmission of information among neurons via spike trains is a finite state information source. The neuron layers, which in our case belong to the primary visual cortex *in vivo* or *in vitro*, are exposed to several

external stimuli (sinusoidal modulation of luminosity, injection of sinusoidal and random currents). The properties of these stimuli (periodicity, randomness, etc.) go into the model as subsets of the parameter space, being each stimulus determined by fixed values of the parameters (e.g. frequency or correlation length). By observing different properties of the stimuli (what in fact amounts to exploring distinct domains of the parameter space), we can have different distribution functions on the same ensemble of stimuli. In the sense of Shannon (Shannon 1948), this means that the same subset of stimuli can contain different amounts of information depending on the properties we are interested in. The question arises as to (i) what is (are) the encoding(s) neurons use and (ii) whether the same encoding method is actually used to transmit different groups of properties of the stimuli. To gain some insight into these interesting and difficult problems we study information sources consisting of the neuron net together with the ensemble of stimuli and the encoding method. In our approach we further assume such sources to be stationary, ergodic and Markovian of order k . Hence the number of states of the source equals the number of encoding letters to the power k .

The results of the calculations show:

1. For periodic stimuli, the number of states *in vivo* are significantly smaller than the number of states *in vitro* in all encoding methods used. The restriction to periodic stimuli is of course due to the fact that output records both *in vivo* and *in vitro* were available only for this kind of stimuli.
2. The number of states *in vivo* is similar for periodic current and periodic visual stimulation within the same encoding method.
3. The number of states *in vitro* is significantly larger for periodic stimuli than for random stimuli if (binary or general) interval coding is used. For the bin coding the inequality reverses.

Multiplicity of coding methods poses the following question: If quantitative results depend on the coding used, are they meaningful in some sense? From a formal point of view there is no objection since, in our approach, the coding is part of the source being analyzed: quantities gained with different codings refer rather to different sources.

Other important issue concerns the choice of the coding. First of all, information tools are sometimes used only to discriminate neural signals produced under different conditions. In this case, one should choose of course those codings which lead to different behaviors or numerical values of the properties in question (entropy, complexity, ...) —the coding becomes part of the tag which characterizes the signal. So to say, two codings are equivalent as far as any of them can be used for this purpose, but the best coding will be the one which provides the broadest or sharpest gaps.

But many times one seeks quantitative rather than qualitative results. In these cases, the choice of the coding depends, generally speaking, on the information one wants to get. In principle, different codings target to different properties, although in a concrete situation it can be difficult to find out an appropriate coding. One could even try to tailor codings for specific properties!

Finally, the question about the choice of the coding can be also reversed: For a given coding of the spike trains, what kind of information can be obtained about the stimuli as a whole (e.g. distribution function of quantities characterizing stimuli) or about the individual stimuli that caused the spike train under consideration (Borst and Theunissen 1999)? This is something that, in general, can only be answered out of experience, although the very definition of the coding might help to figure out the relevant aspects involved.

Conclusions

From the results listed in the last section, we get to the following conclusions:

1. The periodic *in vivo* sources transmit more information than the periodic *in vitro* ones. This follows from the fact that the correlations of the output signals are shorter *in vivo*, what implies that they are more random and, consequently, more difficult to compress because they contain more new patterns.
2. It is possible to model periodic visual stimulation by using its electrical counterpart.
3. The choice of coding affects the results, or put in other words: the information obtained depends on the coding used and vice versa. We think that this interplay between properties and codings is more than a mathematical artifact of our model and somehow reflects what actually happens at neuronal level.

It seems therefore that the communication in the nervous system is based rather on multiple neural codes. We conjecture that different properties are transmitted by means of different codes.

Partially supported by Polish-Spanish Scientific Cooperation Program (PAS-CSIC), grant 20/2001-2002. L.G. Nowak collaborated in the *in vivo* experiments. Part of the experimental work included in this paper was performed at D.A. McCormick's laboratory.

References

- Borst A, Theunissen FE: Information theory and neural coding. *Nature Neurosci.* 1999; 2:947-957.
- Dan Y, Atick JJ, Reid RC: Efficient coding of natural scenes in the lateral geniculate nucleus: Experimental test of a computational theory. *J. Neurosci.* 1996; 16:3351-3362.
- Gallanger RC: *Information Theory and Reliable Communications*. 1968. John Wiley, New York.
- Hodgkin AL and Huxley AF: Currents carried by sodium and potassium ions through the membrane of the giant axon of *Loligo*. *J. Physiol.* 1952; 116:449-472.
- Lempel A, Ziv J: On the complexity of an individual sequence. *IEEE Trans. Inform. Theory* 1976; vol. IT-22: 75-88.

MacKay D, McCulloch WS: The limiting information capacity of a neuronal link. Bull. Math. Biophys. 1952; 14:127-135.

Rapp PE, Zimmerman ID, Vining EP, Cohen N, Albano AM, Jiménez-Montaño MA: The algorithmic complexity of neural spike trains increases during focal seizures. J. Neurosci. 1994; 14:4731-4739.

Rieke F, Warland D, de Ruyter van Steveninck R, Bialek W: *Spikes: Exploring the Natural Code*. 1998. MIT Press, Boston.

Sanchez-Vives MV, Nowak LG, McCormick DA: Membrane mechanisms underlying contrast adaptation in cat area 17 in vivo. J. Neurosci. 2000a; 20:4267-4285.

Sanchez-Vives MV, Nowak LG, McCormick DA: Cellular mechanisms of long-lasting adaptation in visual cortical neurons in vitro. J. Neurosci. 2000b; 20:4286-4299.

Sanchez-Vives MV, Szczepański J, Amigó JM, Wajnryb E: Measurement of the information content in the neuron responses of the visual cortex via Lempel-Ziv complexity. IX Congreso de la Sociedad Española de Neurociencias. Santiago de Compostela, September 9-11, 2001 (in Spanish. Accepted).

Shannon CE: A mathematical theory of communication. Bell Sys. Tech. Journal 1948; 27:379-423, 623-656.

Wyner AD, Ziv J, Wyner AJ: On the role of pattern matching in information theory. IEEE Trans. Inform. Theory 1998; 44:2045-2056.

Zador A: Impact of synaptic unreliability on the information transmitted by spiking neurons. J. Neurophysiol. 1998; 79:1219-1229.

Ziv J: Compression, tests for randomness and estimating the statistical model of an individual sequence. In Capocelli RM (editor): *Sequences*, 366-373, 1990. Springer Verlag, New York.

Ziv J, Lempel A: Compression of individual sequences via variable-rate coding. IEEE Trans. Inform. Theory 1978; vol. IT-24:530-536.