Review

CrossMark
click for updates

Author for correspondence:
Karsten Keller
e-mail: keller@math.uni-luebeck.de

Ordinal symbolic analysis and its application to biomedical recordings

José M. Amigó[1], Karsten Keller[2] and Valentina A. Unakafova[2,3]

[1]Centro de Investigación Operativa, Universidad Miguel Hernández, Avda. de la Universidad s/n, 03202 Elche, Spain
[2]Institut für Mathematik, and [3]Graduate School for Computing in Medicine and Life Science, Universität zu Lübeck, Ratzeburger Allee 160, 23562 Lübeck, Germany

Ordinal symbolic analysis opens an interesting and powerful perspective on time-series analysis. Here, we review this relatively new approach and highlight its relation to symbolic dynamics and representations. Our exposition reaches from the general ideas up to recent developments, with special emphasis on its applications to biomedical recordings. The latter will be illustrated with epilepsy data.

1. Introduction

Symbolic representation is an efficient technique in time-series analysis. For a symbolic representation to be useful, it must be simpler to analyse than the original time series, while still retaining the information one is interested in. This procedure is familiar from linear time-series analysis, where data binning is the first step in any statistical analysis. In this paper, we will focus on nonlinear time-series analysis [1,2] because of the kind of applications under investigation. Thus, our series will be sequences of real numbers obtained by sampling a complex dynamics and, more specifically, a biological process evolving in time. This is the realm of dynamical systems, the theory that models and studies time-varying phenomena to characterize different behaviours.

Consider a discrete-time dynamical system, i.e. a difference evolution law defined on a certain space whose points represent the states of our system. Given an initial state, the evolution law determines unambiguously the state of the system, step by step, at

any future time. The result is the trajectory of the initial state, with each state in the trajectory being reached from the previous one in one time step. At first glance, the time-discrete approach seems to be rather restrictive; however, the approach includes equidistant discretizations of time-continuous dynamical systems as given by autonomous differential equation systems, and Poincaré sections of time-continuous dynamical systems.

A working hypothesis of nonlinear time-series analysis is that the trajectory is unfolding in an attractor, possibly a chaotic one (at least, after an initial transient), so that stationarity, existence of an invariant measure and ergodicity may be taken for granted. Typical scopes of a nonlinear time-series analysis include the reconstruction of the attractor (in particular, the determination of fractal dimensions), and the computation of Lyapunov exponents.

Concurrently or alternatively, one may opt for a symbolic representation of the observations, depending on the scope and actual possibilities of the data analysis. In this approach, the state space is partitioned into a finite number of pieces, and the original trajectories are traded off for trajectories with respect to that partition. These so-called coarse-grained trajectories turn out to be realizations of a stationary random process with a finite alphabet. As a result, the analyst can now gain information from the underlying system in a more accessible way. In particular, recourse to symbol distributions and information-theoretical tools such as entropy and entropy-like quantities have a long tradition in biomedicine since MacKay and McCulloch [3] studied the information conveyed by neuronal signals. More recent contributions of symbolic representations are concerned with heart rate complexity and variability [4–8].

In this paper, we are going to concentrate on *ordinal symbolic representations*, i.e. representations whose symbols are ordinal patterns or, for that matter, permutations. Ordinal patterns were introduced by C. Bandt and B. Pompe in their foundational paper on permutation entropy [9]. An ordinal pattern of length $L \geq 2$ (or ordinal L-pattern) is a permutation displaying the rank order of L consecutive entries in a random or deterministic time series, while permutation entropy is the Shannon entropy of the resulting probability distribution. Therefore, ordinal patterns are not symbols *ad hoc* but they actually encapsulate qualitative information about the temporal structure of the underlying data. Ordinal symbolic analysis, i.e. the analysis of ordinal representations, has some practical advantages. First of all, it is in general acknowledged as being conceptually simple and computationally fast. Furthermore, ordinal patterns, being defined by inequalities, are relatively robust against observational noise. Let us mention in this regard that ordinal patterns of deterministic signals are more robust than those of random signals due to a mechanism called dynamic robustness [10, §9.1]. Another practical advantage is that the calculation of ordinal patterns does not require a knowledge of the data range, which is very useful in real-world data analysis. For practical and computational aspects of ordinal patterns, see [11,12].

Since its conception in 2002, ordinal symbolic representation has found a number of interesting applications in science and engineering. (For a large collection of related papers with further references, see [13].) Of this diversity, we are going to consider in the present paper only applications to biomedical recordings. Bearing in mind the scientific and social impact of biomedical research, and the constant need for ever better diagnostic tools, it is not surprising that the analysis of electroencephalograms (EEGs) of patients with epilepsy was among the very first applications of ordinal patterns and permutation entropy [14,15]. Since then ordinal symbolic analysis has remained a popular method in biology and medicine, especially when it comes to distinguishing abnormal from normal health conditions in real time. Specific examples include epilepsy [16–24], anaesthesia [25–35], cardiology [36–41], event-related potential (ERP) analysis [42,43], Alzheimer's disease [44], sleep research [45] and brain–computer interfaces [46]. For a review of biomedical (and econophysics) applications of permutation entropy, see also [47].

In the following review of ordinal symbolic dynamics, the interested reader will find an account of the concepts, methods and potentialities of the ordinal symbolic analysis of biomedical recordings. This is accompanied by illustrations using data from epilepsy research.

In order to present the material in a systematic and comprehensive way, the rest of this paper is organized as follows. In §2, we give an overview of dynamical systems, symbolic dynamics and symbolic representations. We intend hereby not only to set the mathematical

framework to symbolic representations, but also to integrate ordinal representations in the bigger picture. Section 3 presents some methods of visualizing the ordinal pattern distributions of a time series. Visualization is a useful way of displaying information so that interesting events can be easily recognized, as well as a complement to other, more quantitative tools. The latter, in the form of ordinal pattern-based complexity quantifiers, are considered in §4. They include permutation entropy along with other entropy-like measures such as symbolic transfer entropy, sorting information transfer, complexity coupling coefficients, transcript-based entropies and information directionality indices, regularity parameters, etc. The applicability and efficiency of most of them is still the subject of intense research, as evidenced by the many recent references given in the text. The paper ends with a summary in §5.

2. Symbolic dynamics

The beginning of symbolic dynamics is usually traced back to Hadamard's [48] coding of closed geodesics by symbolic words. From the time when Morse & Hedlund [49] first used the name 'symbolic dynamics' to present days, it has developed to a stand-alone field with a strongly formalized theory and widespread applications, in particular to time-series analysis. For the history of symbolic dynamics, see also [50], and for an overview of applications, see [51].

In the following subsections, we provide a somewhat formal account of symbolic representations via symbolic dynamics. The aim is to highlight the relationship between symbolic representations and their original time series in the setting of nonlinear dynamics, as well as to put ordinal symbolic representations in the right mathematical perspective.

(a) Symbolic representation of time series

To motivate symbolic dynamics, consider the dynamics generated by a self-map f of a subset Ω of a (low-dimensional) Euclidean space. The dynamics is introduced in the *state space* Ω via the repeated action of f on Ω. Given $\omega_0 \in \Omega$, the (forward) *orbit* or *trajectory* $\mathcal{O}_f^+(\omega_0)$ of ω_0 under f is defined as

$$\mathcal{O}_f^+(\omega_0) = (f^n(\omega_0))_{n \in \mathbb{N}_0} = (\omega_0, f(\omega_0), f^2(\omega_0), \ldots),$$

where $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$, $f^0(\omega) = \omega$ and $f^n(\omega) = f(f^{n-1}(\omega))$. In the particular case that f is invertible, one can also consider the full orbit $(\ldots, f^{-2}(\omega_0), f^{-1}(\omega_0), \omega_0, f(\omega_0), f^2(\omega_0), \ldots)$. Since in time-series analysis data acquisition starts necessarily at a finite time, we consider henceforth only the general case of forward orbits. The name 'orbit' clearly hints to the interpretation of the iteration index n as discrete time: each application of f on the point $\omega_n = f(\omega_{n-1}) \in \Omega$ updates the state of the system.

Furthermore, to obtain interesting pay-offs one often supposes that Ω is endowed with a probability measure μ, modelling the distribution of the states of a system. In this extended setting, the time evolution of the system should preserve the probability of the 'events' (measurable subsets of Ω), i.e. μ is supposed to be f-*invariant* in the sense that $\mu(f^{-1}B) = \mu(B)$ for any event B.

If a dynamics is complicated, we might content ourselves with a 'blurred' picture of the orbit behaviour. This can be done as follows. Divide Ω into a finite number of disjoint pieces A_i, $i = 0, 1, \ldots, k - 1$, and keep track of the trajectory of $\omega \in \Omega$ with the precision set by the *partition* $\alpha = \{A_0, \ldots, A_{k-1}\}$. Specifically, we assign to ω a sequence

$$\mathbf{S}(\omega) = (S_n(\omega))_{n \in \mathbb{N}_0} = (i_0, i_1, \ldots, i_n, \ldots),$$

the nth entry $i_n \in \{0, 1, \ldots, N - 1\}$ telling us in which element of α the iterate $f^n(\omega)$ is to be found. We call $\mathbf{S}(\omega)$ the *itinerary* (or address) of ω with respect to the partition α. Formally,

$$S_n(\omega) = i_n \quad \text{if and only if } f^n(\omega) \in A_{i_n},$$

for all $n \in \mathbb{N}_0$. Note that $\mathbf{S}(\omega)$ depends, in general, on the partition α chosen, so we write $\mathbf{S}^\alpha(\omega) = (S_n^\alpha(\omega))_{n \in \mathbb{N}_0}$ whenever convenient to stress that dependency. In principle, the relation between

points ω and their symbolic representations will be many-to-one, i.e. several orbits may have the same symbolic representation, unless α is a generating partition (or a generator): given an f-invariant probability measure μ, the partition α is said to be a generator of f if on a set of probability 1 the relation between states $\omega \in \Omega$ (hence, orbits $\mathcal{O}_f^+(\omega)$) and their itineraries $\mathbf{S}^\alpha(\omega)$ is 1-to-1.

In the case that the partition α is not generating, one can still use the symbolic representations for practical purposes. For example, one can use the statistics of the different symbols along single symbolic time series or samples thereof to characterize different dynamical behaviours or regimes of the data source. Another typical approach to the same objective is to calculate the entropy of the symbolic representation. Since this is a finite-alphabet message, one may resort to Shannon entropy or, of course, to other fancier proposals such as permutation entropy (see below).

(b) Entropy

It turns out then that the maps $S_n^\alpha : \Omega \to \{0, 1, \ldots, k-1\}$, sending ω to the nth component of its itinerary $\mathbf{S}^\alpha(\omega)$, define a stationary stochastic process $(S_n^\alpha)_{n \in \mathbb{N}_0}$, called the *symbolic dynamics* of f with respect to the partition α, for which elements A_i, $1 \le i \le N-1$, are supposed to be measurable. The joined probability functions are given by (or derived from)

$$\Pr(S_0^\alpha = i_0, S_1^\alpha = i_1, \ldots, S_n^\alpha = i_n) = \mu(A_{i_0} \cap f^{-1} A_{i_1} \cap \ldots \cap f^{-n} A_{i_n}). \tag{2.1}$$

We denote this random process by $\mathbf{S}^\alpha$. In this way, the symbolic representations $\mathbf{S}^\alpha(\omega) = (i_0, i_1, \ldots)$ become realizations of the stochastic process $\mathbf{S}^\alpha$.

Thus far, we set out from a (measure-preserving) dynamical system, possibly with a complex dynamics, and have coarse-grained the dynamics (via a finite partition of the state space) to obtain a stationary random process—a seemingly simpler environment.

In information theory, perhaps the most important quantity one can attach to a stationary random process (also called information source in this setting) is its entropy, which measures the average information per unit time conveyed by the outcomes of the process. For an account of entropy see, for example, [52].

Owing to (2.1), the *Kolmogorov–Sinai entropy* of f *with respect to the partition* α,

$$h_\mu(f, \alpha) = -\lim_{n \to \infty} \frac{1}{n} \sum \mu(A_{i_0} \cap \cdots \cap f^{-n+1} A_{i_{n-1}}) \log \mu(A_{i_0} \cap \cdots \cap f^{-n+1} A_{i_{n-1}}), \tag{2.2}$$

and the *Shannon entropy* of the random process $\mathbf{S}^\alpha$,

$$h(\mathbf{S}^\alpha) = -\lim_{n \to \infty} \frac{1}{n} \sum \Pr(S_0^\alpha = i_0, \ldots, S_{n-1}^\alpha = i_{n-1}) \log \Pr(S_0^\alpha = i_0, \ldots, S_{n-1}^\alpha = i_{n-1}),$$

coincide:

$$h_\mu(f, \alpha) = h(\mathbf{S}^\alpha). \tag{2.3}$$

As the logarithm base we take Euler's number e. The other usual choice is base 2. Clearly, the information we obtain from the original dynamical system via the random process $\mathbf{S}^\alpha$ depends on the partition α. The famous Kolmogorov–Sinai theorem, however, says that if the partition γ is a generator of f then

$$h_\mu(f, \gamma) = \max_\alpha h_\mu(f, \alpha),$$

where the maximum on the right-hand side is taken over all finite partitions of Ω. Consideration of the case with no generating partitions requires this maximum to be substituted by a supremum to get a partition-independent quantity, called the *Kolmogorov–Sinai* (KS) *entropy* of f,

$$h_\mu(f) = \sup_\alpha h_\mu(f, \alpha),$$

which turns out to be an invariant for the concept of equivalence in dynamical systems. As a result of (2.3), if γ is a generating partition, then

$$h(\mathbf{S}^\gamma) = h_\mu(f),$$

and, therefore, all the information derived from it will be intrinsic to the system (and only then!). The KS entropy $h_\mu(f)$ measures the (pseudo-)randomness of the dynamics generated by f.
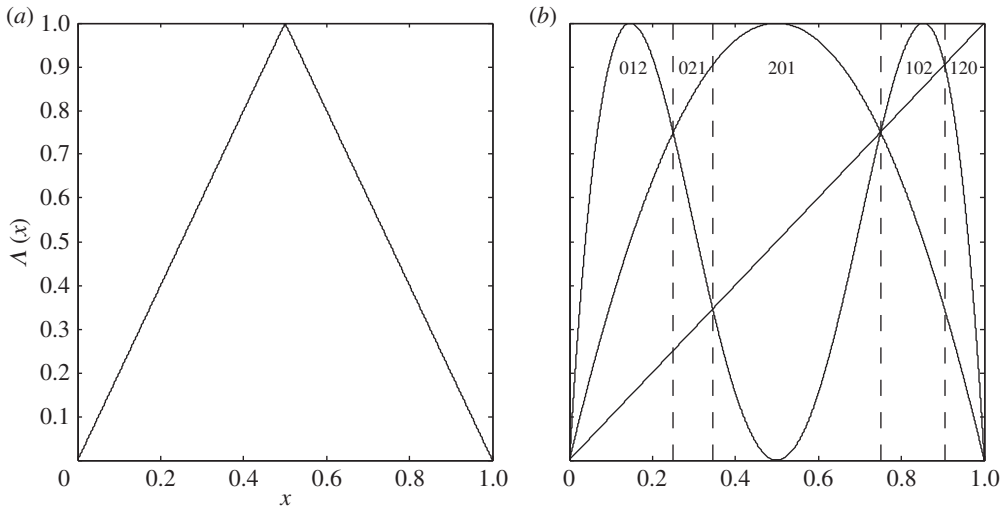
5

rsta.royalsocietypublishing.org *Phil. Trans. R. Soc. A* **373**: 20140091

Figure 1. (*a*) Graph of the tent map. (*b*) Graphs of the identity map and of f, f^2 for $f(x) = 4x(1 - x)$ determining the ordinal partition $\pi^{1,3}$ of $[0, 1]$, vertical dashed lines separate subintervals with ordinal patterns $(0, 1, 2)$, $(0, 2, 1)$, $(2, 0, 1)$, $(1, 0, 2)$ and $(1, 2, 0)$.

As a simple example, consider the tent map $\Lambda : [0, 1] \rightarrow [0, 1]$ defined as

$$\Lambda(\omega) = 1 - |1 - 2\omega| = \begin{cases} 2\omega & \text{if } 0 \leq \omega \leq \frac{1}{2}, \\ 2(1 - \omega) & \text{if } \frac{1}{2} < \omega \leq 1. \end{cases}$$

See figure 1*a* for its graph. It is known that the Lebesgue measure λ (i.e. the usual concept of length on the real axis) is invariant under Λ. If we choose the partition $\alpha = \{A_0, A_1\}$, where

$$A_0 = [0, \tfrac{1}{2}] \quad \text{and} \quad A_1 = (\tfrac{1}{2}, 1]$$

(to which subset the middle point belongs is irrelevant), then the symbolic representation of any orbit under Λ will be a binary sequence. Moreover, it can be proved that α is a generator of Λ, that is, the relationship between the initial conditions $\omega_0 \in [0, 1]$ and the binary sequences $\mathbf{S}^\alpha(\omega_0)$ is 1-to-1 (except for a countable subset of points). For instance, if $\mathbf{S}^\alpha(\omega_0) = (0, 1, \ldots)$, then $\omega_0 \in A_0 \cap \Lambda^{-1}A_1 = [0, \tfrac{1}{2}] \cap (\tfrac{1}{4}, \tfrac{3}{4}) = (\tfrac{1}{4}, \tfrac{1}{2}]$. More generally, the knowledge of k bits of $\mathbf{S}^\alpha(\omega_0)$ locates ω_0 within a dyadic interval $[m/2^k, m + 1/2^k]$ of length 2^{-k}, $0 \leq m \leq 2^k - 1$, and every further digit in $\mathbf{S}^\alpha(\omega_0)$ halves the previous interval, pinpointing ω_0 up to a precision $2^{-(k+1)}$. The corresponding KS entropy of the tent map can be calculated by means of the generator α,

$$h_\lambda(\Lambda) = -\lambda(A_0) \log \lambda(A_0) - \lambda(A_1) \log \lambda(A_1) = -\tfrac{1}{2} \log \tfrac{1}{2} - \tfrac{1}{2} \log \tfrac{1}{2} = \log 2.$$

This is the maximum randomness a variable (or an independent and identically distributed process) with two outcomes can have, the prototype being the tossing of a fair coin.

Generators happen to exist under rather general conditions, though they are explicitly known only for a few maps. In some cases, they can be constructed numerically. In any case, as already pointed out in the previous section, the scope of symbolic dynamics is usually more modest. In general, no invariant characterization of the underlying dynamics is sought, but a method to classify or filter out different dynamical behaviours. In such cases, the computation of partition-dependent complexity measures, such as $h(\mathbf{S}^\alpha)$, might be sufficient.

(c) The ordinal view

In the general setting of symbolic dynamics, we have considered a finite partition of the state space leading to a symbolization. Now we are going the other way around. The symbols will be

ordinal patterns of length $L \geq 2$, describing the order structure of vectors taken from a time series.

The main idea is to assign to a real vector $\mathbf{v} = (x_0, x_1, \ldots, x_{L-1})$ a permutation. We say that $\mathbf{v}$ has the *ordinal (L-)pattern* $o(\mathbf{v}) = \mathbf{r} = (r_0, \ldots, r_{L-1})$ if

$$x_{r_0} \leq x_{r_1} \leq \cdots \leq x_{r_{L-1}},$$

where, in case $x_i = x_j$, we agree to set $r_i < r_j$ if $i < j$. Since $\{r_0, \ldots, r_{L-1}\}$ is a reshuffling of the integers $\{0, \ldots, L-1\}$, the ordinal pattern $\mathbf{r}$ of a vector $\mathbf{v}$ can be identified with the permutation sending i to r_i, $0 \leq i \leq L - 1$.

Note that the $L!$ permutations of $\{0, 1, \ldots, L-1\}$ elements, endowed with the composition

$$\mathbf{r} \circ \mathbf{s} = (r_0, \ldots, r_{L-1}) \circ (s_0, \ldots, s_{L-1}) = (s_{r_0}, \ldots, s_{r_{L-1}}),$$

build an algebraic group called the *symmetric group of degree L*, which we denote by $\mathcal{S}_L$. This fact is used in the quantification of the coupling between time series, as we will discuss later on.

The ordinal representation of a time series is given via ordinal patterns of delay vectors. Given a real time series $(x_n)_{n=0}^{\infty}$ and some natural number $T \in \mathbb{N} = \{1, 2, 3, \ldots\}$ called *delay time*, to each $n \in \mathbb{N}_0$ one assigns the ordinal patterns $o(\mathbf{v}^{T,L}(n))$ of the *time delay vectors*

$$\mathbf{v}^{T,L}(n) = (x_n, x_{n+T}, \ldots, x_{n+(L-1)T}).$$

In order to relate this symbolization to measurements taken from a possibly more than one-dimensional state space, let X be an 'observable' defined on the state space as a measurable real-valued function. If $(\omega_n)_{n \in \mathbb{N}_0}$ is an orbit in Ω, one is interested in the 'measured' time series

$$(x_n)_{n \in \mathbb{N}_0} = (X(\omega_n))_{n \in \mathbb{N}_0}.$$

With

$$\mathbf{v}_X^{T,L}(\omega) = (X(\omega), X(f^T(\omega)), \ldots X(f^{(L-1)T}(\omega))),$$

one gets

$$\mathbf{v}^{T,L}(n) = \mathbf{v}_X^{T,L}(f^n(\omega)),$$

for all $n \in \mathbb{N}_0$. Thus, a so-called *ordinal partition* of Ω,

$$\pi_X^{T,L} = \{P_{\mathbf{r}} \neq \emptyset : \mathbf{r} \in \mathcal{S}_L\},$$

where

$$P_{\mathbf{r}} = \{\omega \in \Omega : o(\mathbf{v}_X^{T,L}(\omega)) = \mathbf{r}\}$$

is obtained by classifying points according to ordinal patterns. The symbols obtained from this partition can be considered as the ordinal patterns themselves.

In the situation of 'directly measuring' the states of Ω, the points ω_i and x_i are coinciding, which means that X is the identity map. Figure 1*b* shows how in such a situation the corresponding ordinal partition $\pi^{1,L}$ can be determined graphically for $L = 3$ (and likewise for other T and moderate values of L) for the logistic map $f(\omega) = 4\omega(1 - \omega)$, $0 \leq \omega \leq 1$. The function graphs of the identity map, of f and its second iterates are shown. Since ordinal patterns can only change at points where at least two of the three functions provide the same value, one has only to check such points in order to get the boundary of the partition pieces. In this concrete situation, $\pi^{1,3} = \{P_{(0,1,2)}, P_{(0,2,1)}, P_{(2,0,1)}, P_{(1,0,2)}, P_{(1,2,0)}\}$ with

$$P_{(0,1,2)} = \left(0, \frac{1}{4}\right), \quad P_{(0,2,1)} = \left(\frac{1}{4}, \frac{5-\sqrt{5}}{8}\right), \quad P_{(2,0,1)} = \left(\frac{5-\sqrt{5}}{8}, \frac{3}{4}\right),$$

$$P_{(1,0,2)} = \left(\frac{3}{4}, \frac{5+\sqrt{5}}{8}\right), \quad P_{(1,2,0)} = \left(\frac{5+\sqrt{5}}{8}, 1\right),$$

where the boundary points are neglected.

Note for further reference that no point has the pattern $(2, 1, 0)$. Such ordinal L-patterns, which cannot be materialized by a deterministic dynamics, are called *forbidden patterns* and, as shown

in [53,54], they are the rule rather than the exception. Forbidden patterns have been used to discriminate noisy deterministic signals from white noise [55,56].

Remark 2.1. When associating an ordinal pattern with a data vector $\mathbf{v}^{T,L}(n)$, the case of equal components was dealt with in a special way, having in mind continuously distributed data. In the case of data digitized with a low resolution, it was proposed by Bian *et al.* [57] to assign single symbols to vectors with equality of special components.

(d) Permutation entropy

As before, let μ be an invariant probability measure under f. For $L = 2, 3, \ldots$ and $T \in \mathbb{N}$ and some given 'observable' X, one is interested in the *L-permutation entropies* with respect to X defined by

$$h_{\mu,X}^{L,T} = -\frac{1}{L-1} \sum_{A \in \pi_X^{T,L}} \mu(A) \log \mu(A), \tag{2.4}$$

as complexity measures.

In the special situation that $T = 1$ all order relations of the elements of an orbit are coded by the ordinal L-patterns, $L = 2, 3, \ldots$. Here we call

$$h_{\mu,X}^* = \limsup_{L \to \infty} h_{\mu,X}^{1,L}$$

the *permutation entropy* of f with respect to X [58,59].

It can be proved that, under mild mathematical conditions (namely, that f is a piecewise, strictly monotonic self-map of a one-dimensional interval), $h_\mu^*(f) = h_{\mu,\mathrm{id}}^*(f)$ coincides with $h_\mu(f)$, the KS entropy of f [58], where id is the identity map, i.e. the states of Ω are directly measured. Moreover, when X separates orbits in a certain sense [60], the sequence of ordinal partitions $\{\pi_X^{1,L} : L = 2, 3, \ldots\}$ determines the KS entropy, i.e.

$$\lim_{L \to \infty} h(\mathbf{S}^{\pi_X^{1,L}}) = h_\mu(f) \tag{2.5}$$

and

$$h_\mu(f) \le h_{\mu,X}^*. \tag{2.6}$$

Let us remark that, in general, none of the ordinal partitions $\pi_X^{1,L}$ is generating for a given map f and that, by Taken's embedding theorem and variants of it, the set of X with the separation property mentioned above is large in a certain sense [59,60]. For a comprehensive discussion of the equality between the permutation and the KS entropies, see [59,60]. A different approach to permutation entropy, which turns out to be equivalent to the KS entropy, was studied in [61].

(e) Practical aspects

In the previous subsections, we have dealt with some formal underpinnings of symbolic representations in general, and ordinal symbolic representations in particular, and their connections to random processes and dynamical systems. In practice though, when analysing empirical data output by a, say, physical or biological system, much of the above framework is missing or hidden to the analyst. Let us mention some typical issues in this regard.

— Although the working hypothesis is that the system is deterministic, the evolution law of the observed dynamics is virtually always unknown.
— Real datasets are finite and very often too short for good statistical estimations of probability distributions. In biological systems, time series need to be short to guarantee stationarity.
— Other usual features of experimental data that one should take into account are noise, instability, and temporal multi-scalarity.

Note that these shortcomings are common to any sort of time-series analysis, whether linear or nonlinear. Additionally, symbolic representations introduce parameters in the analysis through the coarse-graining partition. This calls for a stability analysis of the results with respect to the parameters in some applications.

Ordinal time-series analysis is based on the ordinal pattern distribution obtained from the data. Moreover, on the model side, the data are supposed to be output by a nonlinear dynamical system. In doing so, we approximate probabilities by relative frequencies under the assumption that the considered orbits are 'typical'. These relative frequencies are called *empirical probabilities* in linear time-series analysis, and *natural measures* in nonlinear analysis. They are by construction ergodic (which means precisely that one may equate probabilities and time averages). Usually nonlinearity is checked by means of surrogate tests (e.g. [8]).

However, a purely descriptive data analysis is common as well. The above approach emphasizes (one-dimensional) time series which are the outputs of a possibly higher dimensional system, leading in particular to the problem of system reconstruction. The delay time T is chosen according to some criterion [2,62]. In the standard nonlinear time-series analysis, there are a number of criteria for choosing the parameter L (the so-called *embedding dimension*), perhaps the most popular being the false nearest neighbours method [63]. In ordinal time-series analysis, the usual criteria are computational cost and statistical significance in view of the amount of data available.

3. Ordinal pattern distributions of empirical data

(a) Ordinal patterns distributions

As noted above, the basic ingredients in ordinal time-series analysis are the ordinal patterns of a chosen length L for some delay time T. Given a finite time series

$$(x_n)_{n=0}^{N-1} = (x_0, x_1, \ldots, x_n, \ldots, x_{N-1})$$

of length N, a first important indicator is the distribution of relative frequencies

$$\nu(\mathbf{r}) = \frac{\#\{\mathbf{r} \in \mathcal{S}_L : o(\mathbf{v}^{T,L}(n)) = \mathbf{r}, \ 0 \leq n \leq N - 1 - (L-1)T\}}{N - (L-1)T} \tag{3.1}$$

of ordinal L-patterns $\mathbf{r}$ observed in a time series or in a (time-dependent) sample of it.

These frequencies have been employed in a variety of classification studies. For example, single ordinal pattern frequencies have been considered in electrocardiograms [37,38]. Schindler *et al.* [24] and Rummel *et al.* [64] have observed higher numbers of forbidden ordinal patterns during epileptic seizures, and Ouyang *et al.* [19] have shown a significant growth of forbidden ordinal patterns before absence seizures in rats.

Further insights into biomedical data have been obtained by considering ordinal pattern distributions or special properties of them. Particularly, frequencies of single ordinal patterns are of some interest. For example, Graff *et al.* [40] have discussed the recently found physiological phenomenon of heart rate asymmetry on the basis of ordinal pattern frequencies, providing new views on this topic. In particular, they present a study which shows significant frequency differences for some ordinal 4-patterns and their time reversals in the heart rate data of healthy subjects.

Taking into consideration the symbolic nature of the data obtained and the kind of problem considered, many ideas of nominal statistics and information theory can be applied. For example, Keller & Wittfeld [16] have employed a generalized correspondence analysis in order to visualize changes in the similarity between the ordinal pattern of EEG components, and Cammerota & Rogora [36] have introduced a symbolic independence test for investigating atrial fibrillation on the basis of heart rate data.

Below we will discuss commonly used quantifiers based on ordinal pattern distributions of whole time series or parts of them.

Table 1. Different representations of ordinal patterns for $L = 3$.

r	(2, 1, 0)	(2, 0, 1)	(1, 0, 2)	(1, 2, 0)	(0, 2, 1)	(0, 1, 2)
$(u_1(\mathbf{r}), u_2(\mathbf{r}))$	(0, 0)	(0, 1)	(0, 2)	(1, 0)	(1, 1)	(1, 2)
number representation $y(\mathbf{r})$	0	1	2	3	4	5

(b) Useful representations of ordinal patterns

Besides extracting information from given data via certain quantifiers, the visualization of ordinal pattern distributions can provide interesting insights into the dynamics of time series and systems behind them and is recommended as the first step of analysis. We want to use two visualization methods here. For this, we first describe a natural enumeration of ordinal patterns with some convenient properties.

Given an ordinal L-pattern $\mathbf{r} = (r_0, r_1, \ldots, r_{L-1})$, for $l = 1, \ldots, L - 1$ let

$$u_l(\mathbf{r}) = \#\{j < l : r_j < r_l\}.$$

It is easy to see that $\mathbf{r}$ can be reconstructed from $(u_1(\mathbf{r}), \ldots, u_{L-1}(\mathbf{r}))$ and that each finite sequence in $\{0, 1\} \times \{0, 1, 2\} \times \cdots \times \{0, \ldots, L - 1\}$ corresponds to an ordinal pattern. From this, similarly to Keller *et al.* [65], one obtains that the assignment

$$\mathbf{r} \rightarrow y(\mathbf{r}) := \sum_{l=1}^{L-1} u_l(\mathbf{r}) \frac{L!}{(l+1)!} \tag{3.2}$$

is a bijection from the set of ordinal L-patterns onto $\{0, 1, \ldots, L! - 1\}$. We call $y(\mathbf{r})$ the *number representation* of $\mathbf{r}$. Table 1 illustrates the bijection described for $L = 3$.

In order to see that the number representation of ordinal patterns obtained in this way is rather natural and well adapted to the analysis of a time series, consider a time series $(x_n)_{n \in \mathbb{N}_0}$ of real numbers. For $T \in \mathbb{N}$ and $n \in \mathbb{N}_0$, let a sequence

$$\mathbf{u}_T(n) = (u_{T,1}(n), u_{T,2}(n), \ldots) \in \{0, 1\} \times \{0, 1, 2\} \times \cdots$$

be defined by

$$u_{T,l}(n) = \#\{i : 0 \leq i < l, \, x_{n+Ti} \leq x_{n+Tl}\},$$

for $l \in \mathbb{N}$. Fixing some L, then by construction

$$u_{T,l}(n) = u_l(o(\mathbf{v}^{T,L}(n))),$$

for all $l = 1, 2, \ldots, L - 1$; hence $\sum_{l=1}^{L-1} u_{T,l}(n)(L!/(l+1)!) \in \{0, 1, \ldots, L! - 1\}$ is the number representation of the ordinal pattern of $\mathbf{v}^{T,L}(n)$ (compare (3.2)). Division by $L!$ provides the number

$$y_{T,L}(n) := \sum_{l=1}^{L-1} \frac{u_{T,l}(n)}{(l+1)!} \in [0, 1], \tag{3.3}$$

converging to

$$y_T(n) := \sum_{l=1}^{\infty} \frac{u_{T,l}(n)}{(l+1)!} \in [0, 1],$$

for $L \rightarrow \infty$. Both $y_T(n)$ and the sequence $\mathbf{u}_T(n)$ contain all information about the ordinal structure of the delay vectors $\mathbf{v}^{T,L}(n)$, $L \in \mathbb{N}$, and can therefore be considered as the 'infinite ordinal pattern' of the time series at time n for delay time T. Note that for white noise processes the 'infinite ordinal patterns' considered as numbers in $[0, 1]$ are theoretically equidistributed on $[0, 1]$.

Figure 2 shows the well-known *Feigenbaum diagram* (*a*) together with some related graphics (*b*) based on ordinal symbolic analysis. Recall that in the Feigenbaum diagram typical orbits of the quadratic map $x \mapsto rx(1 - x)$ on the interval $[0, 1]$ are drawn in the vertical direction
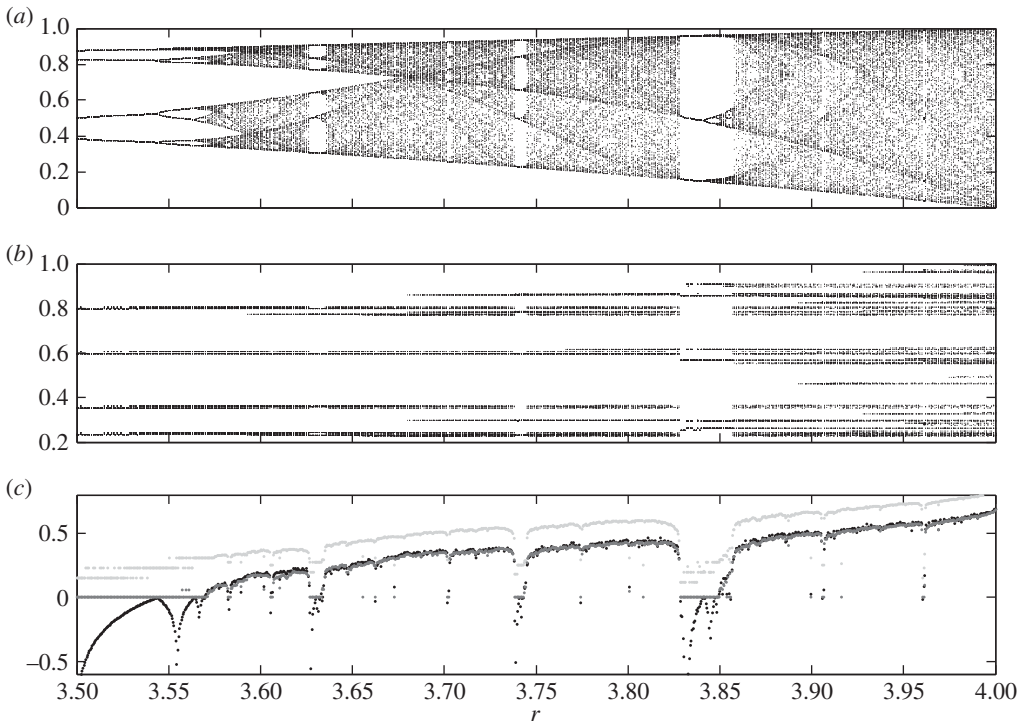
Figure 2. Complexity of the quadratic family: (*a*) the Feigenbaum diagram, (*b*) the ordinal version of the Feigenbaum diagram and (*c*) the corresponding Lyapunov exponents (black), permutation entropy $h^{1,10}$ (light grey) and the conditional entropies of ordinal patterns (dark grey).

in dependence on *r*. We took equidistant values of the parameter *r* between 3.5 and 4 and kept in each case the iterates $x_{1001}, x_{1002}, \ldots, x_{2000}$ of some randomly chosen initial point x_0. In the second graphic instead of the iterates themselves their ordinal patterns $y_{1,7}(1001)$, $y_{1,7}(1002), \ldots, y_{1,7}(2000)$ as numbers in [0, 1] are shown. Note that for many values of *r* one sees a very thin 'attractor', better unveiling determinism than in the Feigenbaum diagram.

Another interesting visualization of dynamical structure is provided by *recurrence plots*, which have ordinal analogues showing interesting details of the recurrence structure of a time series. Such order-recurrence plots have been considered by Groth [43] and Schinkel *et al.* [42] and applied to ERP data, particularly with the success of reducing the number of trials necessary.

(c) Illustration by two electroencephalographic examples

To illustrate ordinal pattern distributions in real-world data, we consider two datasets from the European Epilepsy Database [66] (figures 3 and 4). The upper plot (*a*) in both figures represents raw EEG data recorded at a sampling rate of 256 Hz; in each the duration of an epileptic seizure is marked in grey. Figure 3 illustrates scalp data of a 52-year-old female patient with epilepsy caused by inflammation. The data were recorded from electrode F4, and the seizure is a partial one registered during sleep stage I. Figure 4 illustrates scalp data of a 54-year-old female patient with epilepsy caused by hippocampal sclerosis. Here the data were recorded from electrode C4 and the seizure is a complex partial one registered during sleep stage II. The following plots are aimed at reflecting changes in EEG data before and during the seizure by visualization of ordinal pattern distributions.
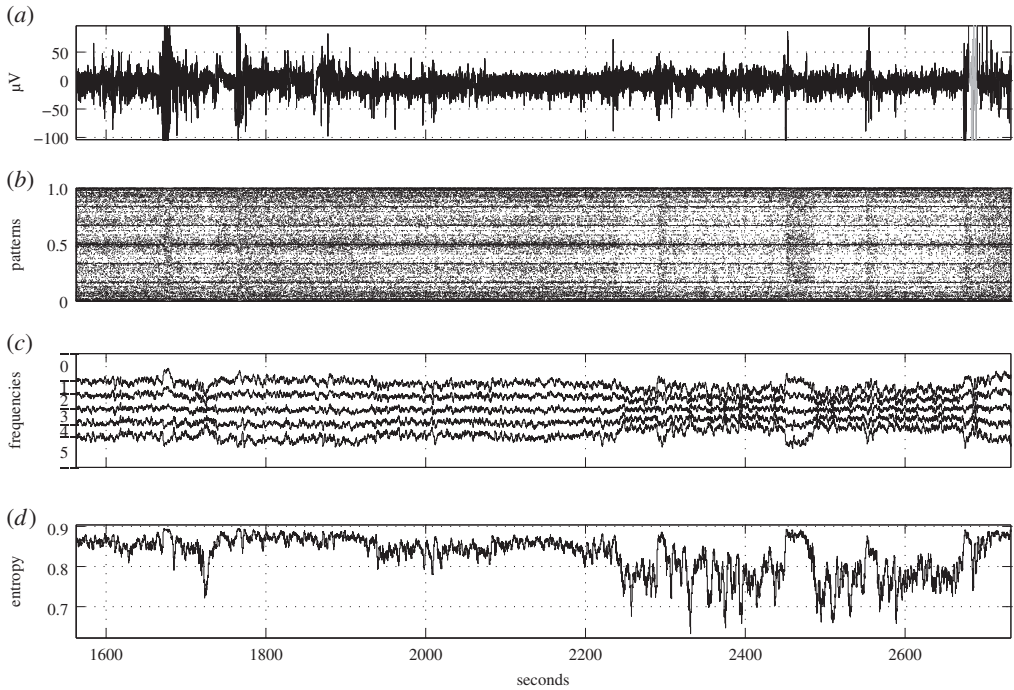
Figure 3. EEG example 1. (*a*) Original time series with epilepsy caused by inflammation (epileptic seizure is marked in grey), (*b*) corresponding time series of ordinal patterns ($T = 1$, $L = 7$) as numbers in [0, 1], (*c*) relative frequencies of the six ordinal 3-patterns (cf. table 1) in a sliding window of 2 s with partitioning in the vertical direction (ordinal pattern distribution, $T = 1$, $L = 3$) and (*d*) permutation entropy ($T = 1$, $L = 3$) for a sliding window of 2 s.
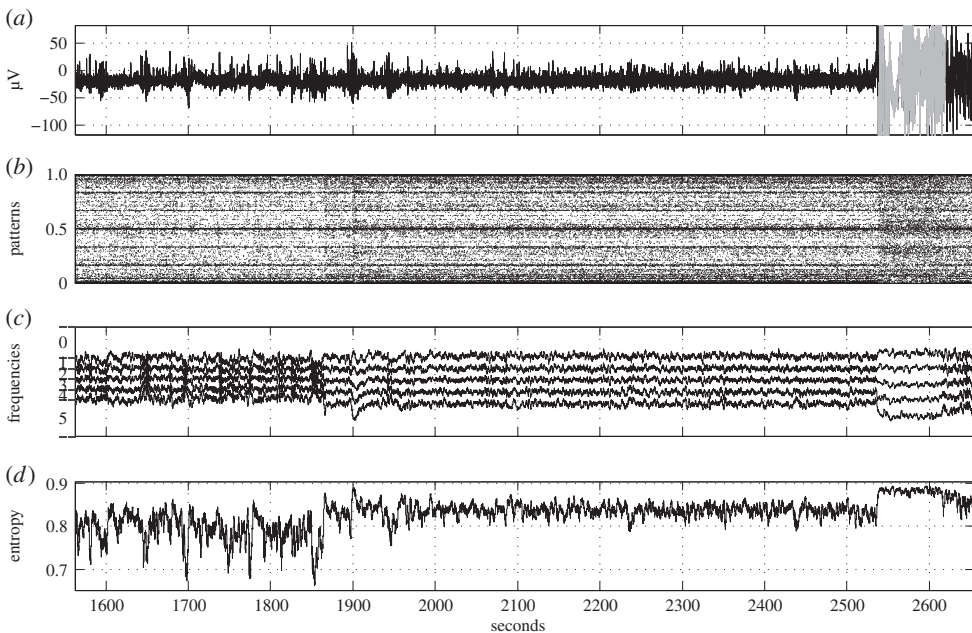

Figure 4. EEG example 2. (*a*) Original time series with epilepsy caused by hippocampal sclerosis (epileptic seizure is marked in grey), (*b*) corresponding time series of ordinal patterns ($T = 1$, $L = 7$) as numbers in [0, 1], (*c*) relative frequencies of the six ordinal 3-patterns (cf. table 1) in a sliding window of 2 s with partitioning in the vertical direction (ordinal pattern distribution, $T = 1$, $L = 3$) and (*d*) permutation entropy ($T = 1$, $L = 3$) for a sliding window of 2 s.

Plot (*b*) represents the numbers $y_{1,7}(n) \in [0,1]$ (see (3.3)) coding ordinal patterns in dependence on the time *n*. Since the representation is very truncated in the time direction, time-localized ordinal pattern distributions and changes in the distribution are well illustrated.

In plot (*c*), for $T = 1$ the relative frequencies of all ordinal 3-patterns in a sliding window of 2 s are represented. The plot has to be read as follows: enumerate the curves including the baselines at 1 and 0 from above to below by $0, 1, \ldots, 6$. Then, in the window starting at *n*, the relative frequency of the ordinal pattern with number $k = 0, 1, \ldots, 5$ according to table 1 is the vertical distance of the *k*th and $k + 1$th curve at time *n*.

Interesting aspects shown by figures 3 and 4 are the changes in ordinal pattern distribution at the beginning of epileptic seizures, and at 2240 s (figure 3*b*) and 1870 s (figure 4*b*) accompanied by changes (figure 3) in relative frequencies of up patterns (0) and down patterns (5), respectively. Although in figure 3 the EEG returns relatively quickly to the 'normal' level, in figure 4 the 'relaxation' seems to take much longer.

4. Ordinal complexity quantifiers

We proceed to present a few quantifiers or classifiers in ordinal symbolic representations that have been employed in the literature to characterize and/or discriminate different dynamical behaviours. Such quantities are also called indicators or markers.

(a) Permutation entropy

(i) The original concept

Historically, the first ordinal dynamic indicator was the permutation entropy of order *L* constructed with the frequencies (3.1) and to be considered as an estimator of (2.4),

$$h^{T,L} = -\frac{1}{L-1} \sum_{\mathbf{r} \in \mathcal{S}_{T,L}} \nu(\mathbf{r}) \log \nu(\mathbf{r}). \tag{4.1}$$

Indeed, (4.1) with $L = 2, 3, 4$ has already been used in the foundational paper of permutation entropy [9] to analyse speech signals.

It did not take long for further applications to appear. Indeed, just 2 years later Keller & Lauffer [14] discussed changes in brain dynamics related to vagus stimulation on the basis of permutation entropy of EEG data. About the same time, Cao *et al.* [15] used $h^{T,L}$ to find that the EEGs of patients with epilepsy are usually less complex during epileptic seizures than in the normal state. Since then permutation entropy has remained a popular biomarker of epileptic seizures. Thus, Veisi *et al.* [23] discussed the classification of normal and epileptic EEGs, Nicolaou & Georgiou [21] considered permutation entropy for automatic detection of epileptic seizures and Li *et al.* [17] even used it for predicting them. Also note that Ouyang *et al.* [20] applied permutation entropy and a special dissimilarity measure for detecting changes in the dynamics of data from rats with absence epilepsy, and Bruzzo *et al.* [18] included the problem of vigilance changes in the discussion of distinguishing epileptic and normal EEG epoches.

In anaesthesia, permutation entropy of EEG data is being used or tested increasingly as an anaesthetic depth measure for evaluating anaesthetic drug effects ([27,29,31]; for the concept of multi-scale permutation entropy, see below and [28,34,35]) or for distinguishing between consciousness and unconsciousness [26,33]. In particular, the performance of permutation entropy has been compared with that of other measures such as approximate entropy, sample entropy and index of consciousness with mainly good or best results [26,27,29–32,34,35]. Note that in [32] local field potentials were considered instead of EEGs, and that good results with permutation entropy were obtained after some burst suppression correction. Burst suppression patterns occur in EEGs in inactivated brain states.

Nicolaou & Georgiou [45] report that different sleep stages correspond to significantly different permutation entropies in the pertinent EEG segments, showing potential in automatic sleep stage

classification. Moreover, permutation entropy has been used in a brain–computer interface system for mental task classification [46] and in exploring heart rate variability [39,57]. In Cysarz *et al.* [39], three specially constructed symbolic representations and ordinal representations were considered in order to reflect changes in the cardiac autonomic nervous system during head-up tilt with nearly the same results.

It is not surprising that permutation entropy and similar measures are applied to physical data because such data are characterized by special underlying patterns often related to certain states of a system (e.g. spike-and-wave patterns related to epilepsy, special sleep patterns such as sleep spindles, and burst suppression patterns related to inactivated brain states, or Wolff–Parkinson–White patterns in abnormal electrocardiograms). It seems that ordinal patterns are appropriate for capturing structures containing such patterns and also abrupt changes in their distributions. Here because of the assumed nonlinear nature of the systems behind the data, the idea of pattern complexity is indicated.

Again in figures 3 and 4, note that plots (*d*) reflect changes in the permutation entropy $h^{1,3}$ computed in sliding windows of 2 s. Relations to the plots (*b*,*c*) are obvious.

(ii) Variants of permutation entropy

There are some quantifiers similar to the permutation entropy. Liu & Wang [67] have defined a *fine-grained permutation entropy*. In contrast to a modification of permutation entropy by Bian *et al.* [57] based on considering extra symbols in the case of equal components of the corresponding delay vector as mentioned above, they divide ordinal patterns into more symbols regarding distances between vector components.

The use of different delay times in permutation analysis gives some scale-dependent information on a time series. Another approach for gaining such information is the *multi-scale permutation entropy* invented by Li *et al.* [31] in anaesthesia. Here a given time series is averaged within non-overlapping windows of different lengths, and the permutation entropy is determined as a function of the window length. Then a special measure combining this information is compared with usual permutation entropy. The authors report that transitions between light and deep anaesthesia are better reflected by the new measure.

For the analysis of Alzheimer's disease, Morabito *et al.* [44] combine the new approach with the idea of pooling ordinal patterns from different EEG channels, first considered in [14]. Here the distribution of all ordinal patterns obtained, independent of their source, is investigated.

(b) Similar quantifiers

In the case that $h^{1,L}$ stabilizes for increasing L's one can even take $h^{1,L}$ for sufficiently large L to estimate the permutation entropy of the data source. Note in this connection that algebraic correction terms to finite size effects can be found in [68,69]. Numerical simulations show that, even in the simple case of logistic maps, where permutation entropy and KS entropy coincide, the stabilization mentioned is very slow. For this reason, Unakafov & Keller [70] (for $T = 1$) have considered the conditional entropy of successive ordinal L-patterns, providing better approximation convergency to KS entropy. The latter is illustrated for the logistic family by figure 2*c*, where the Lyapunov exponent, $h^{1,10}$ and the conditional entropy of successive ordinal 10-patterns are plotted in black, grey and dark grey, respectively. Having in mind that, by Pesin's theorem, the KS entropy is 0 if the Lyapunov exponent is negative and equal to the Lyapunov exponent (cf. [71]), these graphics show that the KS entropies, hence the permutation entropies, and considered conditional entropies are nearly the same. $h^{1,10}$ shows the same tendency as the other quantities but is significantly larger.

Finally, we refer to measures based on counting monotone changes in time series and applied to EEG data in Keller *et al.* [72]. Such measures can quantify 'roughness' of data and are, for example, used for a robust estimation of the Hurst exponent of a fractional Brownian motion.

(c) Ordinal coupling quantifiers

There are, of course, ordinal versions of other entropies and information-theoretical quantities (such as conditional entropies and mutual informations), mainly describing coupling. We do not want to go into detail here since this would fill a complete other paper.

One of the most notorious coupling measures is *symbolic transfer entropy* (actually a conditional mutual information), introduced by Staniek & Lehnertz [73], which is used to measure the information directionality in coupled time series. Symbolic transfer entropy, which can be considered as an ordinal variant of transfer entropy [74] has been used in Li & Ouyang [75] to study the interaction between epilepsy foci. It has been developed in different directions by Kugiumtzis [76,77] (transfer entropy on rank vectors, partial transfer entropy on rank vectors), Pompe & Runge [78] (momentary sorting information transfer) and Nakajima & Haruna [79] (symbolic local information transfer).

Along these lines, *coupling complexity coefficients* and information directionality indicators have also been proposed by means of the so-called 'transcripts' between two coupled ordinal symbolic representations [80,81]. Transcripts exploit the algebraic structure of the permutations [82,83] and are useful also for statistical reasons: whereas all combinations of two L-patterns provide $(L!)^2$ possible combinations, the corresponding transcript representation aggregates them to only $L!$ combinations. Biomedical applications of transcripts and coupling complexity coefficients can be found in Bunk *et al.* [84] and Amigó *et al.* [80] and Monetti *et al.* [83], respectively. Biomedical applications of the transcript-based directionality indicators is a topic of current research.

Finally, there are other ordinal indicators tailored to specific purposes. Such is the case for the *regularity parameters* introduced in Amigó *et al.* [85] to differentiate among different phases in coupled map lattices and also to apply to the analysis of multi-channel magnetoencephalograms recorded on human scalps.

5. Summary

Ordinal symbolic representation provides an interesting and promising approach to the analysis of complex time series, with a fast increasing number of biomedical applications and beyond. The potential of the method in biomedicine ranges from the detection of dynamical state changes of a system through the distinction of its states to state classification, mainly in epileptology, anaesthesia and cardiology. The applicability of ordinal time-series analysis has been demonstrated by many examples.

Let us summarize next some features of ordinal symbolic representations that distinguish them from other symbolic representations and make them special.

1. To find appropriate partitions in symbolic dynamics, one has to look for natural levels subdividing the range of measurements. In many practical situations, there is no clear indication of such levels. By contrast, ordinal patterns provide a kind of universal partitioning with standard symbols, moreover without knowing the range of measurements. Yet, partitions are not given *a priori* on the state space but based on the dynamics. The latter is not a practical problem, though, since one is only interested in ordinal patterns as the symbols obtained from the data in a simple way. There are situations, however, in which ordinal methods might be inapplicable without adjustments. For example, close delay vectors can provide very different ordinal patterns, this being a reason for the ordinal pattern modification proposed by Liu & Wang [67].

2. The symbols in ordinal symbolic dynamics are easily interpretable. For example, one has 'upward' and 'downward' ordinal patterns and ordinal patterns describing monotone changes. There are moreover symmetries in the ordinal pattern distributions related to time and data range reversal. Both can be exploited when considering special order-related aspects of data analysis; for example, for describing roughness by change statistics [86] and for quantifying the phenomenon of heart rate asymmetry [40] described above.

3. Finding convenient generating partitions in dynamical systems is an unsolvable problem in many cases. However, as noted above, the KS entropy can be approximated on the base of ordinal partitions for sufficiently large pattern length *L*. This is important at least from a conceptual view point, but in practice *L* can be too large to estimate the KS entropy accurately. Nevertheless, ordinal analysis with a small *L* is useful in descriptive data analysis. The fact that conditional variants of permutation entropy seem better estimators of the KS entropy than the original concepts is certainly worth further research [70].

4. As permutations, ordinal *L*-patterns are elements of the symmetric group of degree *L*. This additional algebraic structure of the symbols allows applications to the analysis of coupled dynamics [80,81].

Ordinal symbolic analysis is a relatively new field and many methods are not consolidated yet. Future research should particularly focus on the development of statistical models and methods aimed at showing the significance and reliability of the results obtained by ordinal analysis. There is moreover some need for systematic studies comparing different complexity and coupling measures, on both the ordinal and metric level.

References

1. Kantz H, Schreiber T. 2004 *Nonlinear time series analysis*. Cambridge, UK: Cambridge University Press.
2. Sprott JC. 2003 *Chaos and time-series analysis*. Oxford, UK: Oxford University Press.
3. MacKay D, McCulloch WS. 1952 The limiting information capacity of a neuronal link. *Bull. Math. Biophys.* **14**, 127–135. (doi:10.1007/BF02477711)
4. Kurths J, Voss A, Saparin P, Witt A, Kleiner HJ, Wessel N. 1995. Quantitative analysis of heart rate variability. *Chaos* **5**, 88–94. (doi:10.1063/1.166090)
5. Voss A, Kurths J, Kleiner HJ, Witt A, Wessel N. 1995 Improved analysis of heart rate variability by methods of nonlinear dynamics. *J. Electrocardiol.* **28**, 81–88. (doi:10.1016/S0022-0736(95)80021-2)
6. Porta A, Guzzetti S, Montano N, Furlan R, Pagani M, Malliani A, Cerutti S. 2001 Entropy, entropy rate, and pattern classification as tools to typify complexity in short heart period variability series. *IEEE Trans. Biomed. Eng.* **48**, 1282–1291. (doi:10.1109/10.959324)
7. Porta A *et al.* 2007 An integrated approach based on uniform quantization for the evaluation of complexity of short-term heart period variability: application to 24h Holter recordings in healthy and heart failure humans. *Chaos* **17**, 015117. (doi:10.1063/1.2404630)
8. Porta A, Guzzetti S, Furlan R, Gnecchi-Ruscone T, Montano N, Malliani A. 2007 Complexity and nonlinearity in short-term heart period variability: comparison of methods based on local nonlinear prediction. *IEEE Trans. Biomed. Eng.* **54**, 94–106. (doi:10.1109/TBME.2006.883789)
9. Bandt C, Pompe B. 2002 Permutation entropy: a natural complexity measure for time series. *Phys. Rev. Let.* **88**, 174102. (doi:10.1103/PhysRevLett.88.174102)
10. Amigó JM. 2010 *Permutation complexity in dynamical systems—ordinal patterns, permutation entropy, and all that*. Berlin, Germany: Springer.
11. Riedl M, Müller A, Wessel N. 2013 Practical considerations of permutation entropy. *Eur. Phys. J. Spec. Top.* **222**, 249–262. (doi:10.1140/epjst/e2013-01862-7)
12. Unakafova VA, Keller K. 2013 Efficiently measuring complexity on the basis of real-world data. *Entropy* **15**, 4392–4415. (doi:10.3390/e15104392)
13. Amigó JM, Keller K, Kurths J (eds). 2013 Recent progress in symbolic dynamics and permutation complexity. Ten years of permutation entropy. *Eur. Phys. J. Spec. Top.* **222**, 247–257.
14. Keller K, Lauffer H. 2003 Symbolic analysis of high-dimensional time series. *Int. J. Bifurcation Chaos* **13**, 2657–2668. (doi:10.1142/S0218127403008168)
15. Cao Y, Tung W, Gao JB, Protopopescu VA, Hively LM. 2004 Detecting dynamical changes in time series using the permutation entropy. *Phys. Rev. E* **70**, 046217. (doi:10.1103/PhysRevE.70.046217)

16. Keller K, Wittfeld K. 2004 Distances of time series components by means of symbolic dynamics. *Int. J. Bifurcation Chaos* **14**, 693–703. (doi:10.1142/S0218127404009387)
17. Li X, Ouyang G, Richards DA. 2007 Predictability analysis of absence seizures with permutation entropy. *Epilepsy Res.* **77**, 70–74. (doi:10.1016/j.eplepsyres.2007.08.002)
18. Bruzzo AA, Gesierich B, Santi M, Tassinari C, Birbaumer N, Rubboli G. 2008 Permutation entropy to detect vigilance changes and preictal states from scalp EEG in epileptic patients. A preliminary study. *Neurol. Sci.* **29**, 3–9. (doi:10.1007/s10072-008-0851-3)
19. Ouyang G, Li X, Dang C, Richards DA. 2009 Deterministic dynamics of neural activity during absence seizures in rats. *Phys. Rev. E* **79**, 041146. (doi:10.1103/PhysRevE.79.041146)
20. Ouyang G, Dang C, Richards DA, Li X. 2010 Ordinal pattern based similarity analysis for EEG recordings. *Clin. Neurophysiol.* **121**, 694–703. (doi:10.1016/j.clinph.2009.12.030)
21. Nicolaou N, Georgiou J. 2012 Detection of epileptic electroencephalogram based on permutation entropy and support vector machines. *Expert Syst. Appl.* **39**, 202–209. (doi:10.1016/j.eswa.2011.07.008)
22. Jouny CC, Bergey GK. 2012 Characterization of early partial seizure onset: frequency, complexity and entropy. *Clin. Neurophysiol.* **123**, 658–669. (doi:10.1016/j.clinph.2011.08.003)
23. Veisi I, Pariz N, Karimpour A. 2007 Fast and robust detection of epilepsy in noisy EEG signals using permutation entropy. In *Proc. 7th IEEE Int. Conf. on Bioinformatics and Bioengineering, Boston, MA, 14–17 October 2007*, pp. 200–203. Piscataway, NJ: IEEE.
24. Schindler K, Gast H, Stieglitz L, Stibal A, Hauf M, Wiest R, Mariani L, Rummel C. 2011 Forbidden ordinal patterns of periictal intracranial EEG indicate deterministic dynamics in human epileptic seizures. *Epilepsia* **52**, 1771–1780. (doi:10.1111/j.1528-1167.2011.03202.x)
25. Bretschneider M, Kreuzer M, Drexler B, Hentschke H, Antkowiak B, Schwarz C, Kochs E, Schneider G. 2006 Coherence of *in vitro* and *in vivo* field potential activity and EEG, detected by ordinal analysis. *J. Neurosurg. Anesthesiol.* **18**, 309. (doi:10.1097/00008506-200610000-00086)
26. Jordan D, Stockmanns G, Kochs EF, Pilge S, Schneider G. 2008 Electroencephalographic order pattern analysis for the separation of consciousness and unconsciousness: an analysis of approximate entropy, permutation entropy, recurrence rate, and phase coupling of order recurrence plots. *Anesthesiology* **109**, 1014–1022. (doi:10.1097/ALN.0b013e31818d6c55)
27. Li X, Cui S, Voss LJ. 2008 Using permutation entropy to measure the electroencephalographic effects of sevoflurane. *Anesthesiology* **109**, 448–456. (doi:10.1097/ALN.0b013e318182a91b)
28. Olofsen E, Sleigh JW, Dahan A. 2008 Permutation entropy of the electroencephalogram: a measure of anaesthetic drug effect. *Br. J. Anaesth.* **101**, 810–821. (doi:10.1093/bja/aen290)
29. Kortelainen J, Koskinen M, Mustola S, Seppanen T. 2009 Effect of remifentanil on the nonlinear electroencephalographic entropy parameters in propofol anesthesia. In *Proc. 31st Annu. Int. Conf. of the IEEE Engineering in Medicine and Biology Society, Minneapolis, MN, 2–6 September 2009*, pp. 4994–4997. Piscataway, NJ: IEEE.
30. Anier A, Lipping T, Jantti V, Puumala P, Huotari AM. 2010 Entropy of the EEG in transition to burst suppression in deep anesthesia: surrogate analysis. In *Proc. 32nd Annu. Int. Conf. of the IEEE Engineering in Medicine and Biology Society, Buenos Aires, 31 August–4 September 2010*, pp. 2790–2793. Piscataway, NJ: IEEE.
31. Li D, Li X, Liang Z, Voss LJ, Sleigh JW. 2010 Multiscale permutation entropy analysis of EEG recordings during sevoflurane anesthesia. *J. Neural. Eng.* **7**, 046010. (doi:10.1088/1741-2560/7/4/046010)
32. Silva A, Cardoso-Cruz H, Silva F, Galhardo V, Antunes L. 2010 Comparison of anesthetic depth indexes based on thalamocortical local field potentials in rats. *Anesthesiology* **112**, 355–363. (doi:10.1097/ALN.0b013e3181ca3196)
33. Nicolaou N, Houris S, Alexandrou P, Georgiou J. 2011 Entropy measures for discrimination of awake vs anaesthetized state in recovery from general anesthesia. In *Proc. 33rd Annu. Int. Conf. of the IEEE Engineering in Medicine and Biology Society, Boston, MA, 30 August–3 September 2011*, pp. 2598–2601. Piscataway, NJ: IEEE.
34. Silva A, Campos S, Monteiro J, Venâncio C, Costa B, Guedes de Pinho P, Antunes L. 2011 Performance of anesthetic depth indexes in rabbits under propofol anesthesia: prediction probabilities and concentration-effect relations. *Anesthesiology* **115**, 303–314. (doi:10.1097/ALN.0b013e318222ac02)
35. Silva A, Ferreira DA, Venâncio C, Souza AP, Antunes LM. 2011 Performance of electroencephalogram-derived parameters in prediction of depth of anaesthesia in a rabbit model. *Br. J. Anaesth.* **106**, 540–547. (doi:10.1093/bja/aeq407)

36. Cammarota C, Rogora E. 2005 Independence and symbolic independence of heartbeat intervals during atrial fibrillation. *Physica A* **353**, 323–335. (doi:10.1016/j.physa.2005.01.030)
37. Berg S, Luther S, Lehnart SE, Hellenkamp K, Bauernschmitt R, Kurths J, Wessel N, Parlitz U. 2010 Comparison of features characterizing beat-to-beat time series. In *Proc. of Biosignal 2010, Berlin, Germany*, *14–16 July 2010*, pp. 1–4.
38. Parlitz U, Berg S, Luther S, Schirdewan A, Kurths J, Wessel N. 2012 Classifying cardial biosignals using ordinal pattern statistics and symbolic dynamics. *Comp. Biol. Med.* **42**, 319–327. (doi:10.1016/j.compbiomed.2011.03.017)
39. Cysarz D, Porta A, Montano N, Leeuwen PV, Kurths J, Wessel N. 2013 Quantifying heart rate dynamics using different approaches of symbolic dynamics. *Eur. Phys. J. Spec. Top.* **222**, 487–500. (doi:10.1140/epjst/e2013-01854-7)
40. Graff G, Graff B, Kaczkowska A, Makowiec D, Amigó JM, Piskorski J, Narkiewicz K, Guzik P. 2013 Ordinal pattern statistics for the assessment of heart rate variability. *Eur. Phys. J. Spec. Top.* **222**, 525–534. (doi:10.1140/epjst/e2013-01857-4)
41. Frank B, Pompe B, Schneider U, Hoyer D. 2006 Permutation entropy improves fetal behavioural state classification based on heart rate analysis from biomagnetic recordings in near term fetuses. *Med. Biol. Eng. Comput.* **44**, 179–187. (doi:10.1007/s11517-005-0015-z)
42. Schinkel S, Marwan N, Kurths J. 2007 Order patterns recurrence plots in the analysis of ERP data. *Cogn. Neurodyn.* **1**, 317–325. (doi:10.1007/s11571-007-9023-z)
43. Groth A. 2005 Visualization of coupling in time series by order recurrence plots. *Phys. Rev. E* **72**, 046220. (doi:10.1103/PhysRevE.72.046220)
44. Morabito FC, Labate D, La Foresta F, Bramanti A, Morabito G, Palamara I. 2012 Multivariate multi-scale permutation entropy for complexity analysis of Alzheimer's disease EEG. *Entropy* **14**, 1186–1202. (doi:10.3390/e14071186)
45. Nicolaou N, Georgiou J. 2011 The use of permutation entropy to characterize sleep electroencephalograms. *Clin. EEG Neurosci.* **42**, 24–28. (doi:10.1177/155005941104200107)
46. Nicolaou N, Georgiou J. 2010 Permutation entropy: a new feature for brain-computer interfaces. In *Conf. Proc. IEEE Biomedical Circuits and Systems Conference 2010, Paphos, Cyprus*, *3–5 November 2010*, pp. 49–52. Piscataway, NJ: IEEE.
47. Zanin M, Zunino L, Rosso AO, Papo D. 2012 Permutation entropy and its main biomedical and econophysics applications: a review. *Entropy* **14**, 1553–1577. (doi:10.3390/e14081553)
48. Hadamard J. 1898 Les surfaces à courbures opposées et leurs lignes géodésiques. *J. Maths. Pures Appliqués* **4**, 27–73.
49. Morse M, Hedlund GA. 1938 Symbolic dynamics. *Am. J. Math.* **60**, 815–866. (doi:10.2307/2371264)
50. Coven EM, Nitecki ZH. 2008 On the genesis of symbolic dynamics as we know it. *Colloq. Math.* **110**, 227–242. (doi:10.4064/cm110-2-1)
51. Daw CS, Finney CEA, Tracy ER. 2003 A review of symbolic analysis of experimental data. *Rev. Sci. Instrum.* **74**, 915–930. (doi:10.1063/1.1531823)
52. Walters P 2000 *An introduction to ergodic theory*. New York, NY: Springer.
53. Amigó JM, Kocarev L, Szczepanski J. 2006 Order patterns and chaos. *Phys. Lett. A* **355**, 27–31. (doi:10.1016/j.physleta.2006.01.093)
54. Amigó JM, Zambrano S, Sanjuán MAF. 2007 True and false forbidden patterns in deterministic and random dynamics. *Europhys. Lett.* **79**, 50001. (doi:10.1209/0295-5075/79/50001)
55. Amigó JM, Zambrano S, Sanjuán MAF. 2008 Combinatorial detection of determinism in noisy time series. *Europhys. Lett.* **83**, 60005. (doi:10.1209/0295-5075/83/60005)
56. Amigó JM, Zambrano S, Sanjuán MAF. 2010 Detecting determinism with ordinal patterns: a comparative study. *Int. J. Bifurcation Chaos* **20**, 2915–2924. (doi:10.1142/S021812741002 7453)
57. Bian C, Qin C, Ma QD, Shen Q. 2012 Modified permutation-entropy analysis of heartbeat dynamics. *Phys. Rev. E* **85**, 021906. (doi:10.1103/PhysRevE.85.021906)
58. Bandt C, Keller G, Pompe B. 2002 Entropy of interval maps via permutations. *Nonlinearity* **15**, 1595–1602. (doi:10.1088/0951-7715/15/5/312)
59. Keller K. 2011 Permutations and the Kolmogorov–Sinai entropy. *Discrete Contin. Dyn. Syst. A* **32**, 891–900. (doi:10.3934/dcds.2012.32.891)
60. Antoniouk A, Keller K, Maksymenko S. 2014 Kolmogorov–Sinai entropy via separation properties of order-generated σ-algebras. *Discrete Contin. Dyn. Syst. A* **34**, 1793–1809.
61. Amigó JM. 2012 The equality of Kolmogorov–Sinai entropy and metric permutation entropy generalized. *Physica D* **241**, 789–793. (doi:10.1016/j.physd.2012.01.004)

62. Fraser AM, Swinney HL. 1986 Independent coordinates for strange attractors from mutual information. *Phys. Rev. A* **33**, 1134–1140. (doi:10.1103/PhysRevA.33.1134)
63. Kennel MB, Brown R, Abarbanel HDI. 1992 Determining minimum embedding dimension using a geometric construction. *Phys. Rev. A* **45**, 3403–3411. (doi:10.1103/PhysRevA.45.3403)
64. Rummel C, Abela E, Hauf M, Wiest R, Schindler K. 2013 Ordinal patterns in epileptic brains: analysis of intercranial EEG and simultaneous EEG-fMRI. *Eur. Phys. J. Spec. Top.* **222**, 569–588. (doi:10.1140/epjst/e2013-01860-9)
65. Keller K, Emonds J, Sinn M. 2007 Time series from the ordinal viewpoint. *Stoch. Dyn.* **2**, 247–272. (doi:10.1142/S0219493707002025)
66. The European Epilepsy Database. 2014 See http://epilepsy-database.eu/ (accessed on 8 April 2014).
67. Liu X-F, Wang Y. 2009 Fine-grained permutation entropy as a measure of natural complexity for time series. *Chin. Phys. B* **18**, 2690–2695. (doi:10.1088/1674-1056/18/7/011)
68. Grassberger P. 1988 Finite sample corrections to entropy and dimension estimates. *Phys. Lett. A* **128**, 369–373. (doi:10.1016/0375-9601(88)90193-4)
69. Paninski L. 2003 Estimation of entropy and mutual information. *Neural Comput.* **15**, 1191–1253. (doi:10.1162/089976603321780272)
70. Unakafov AM, Keller K. 2014 Conditional entropy of ordinal patterns. *Physica D* **269**, 94–102. (doi:10.1016/j.physd.2013.11.015)
71. Choe GH. 2005 *Computational ergodic theory*. Berlin, Germany: Springer.
72. Keller K, Lauffer H, Sinn M. 2007 Ordinal analysis of EEG time series. *Chaos Complexity Lett.* **2**, 247–258.
73. Staniek M, Lehnertz K. 2008 Symbolic transfer entropy. *Phys. Rev. Lett.* **100**, 158101. (doi:10.1103/PhysRevLett.100.158101)
74. Schreiber T. 2000 Measuring information transfer. *Phys. Rev. Lett.* **85**, 461–464. (doi:10.1103/PhysRevLett.85.461)
75. Li X, Ouyang G. 2010 Estimating coupling direction between neuronal populations with permutation conditional mutual information. *NeuroImage* **52**, 497–507. (doi:10.1016/j.neuroimage.2010.05.003)
76. Kugiumtzis D. 2012 Transfer entropy on rank vectors *J. Nonlin. Syst. Appl.* **3**, 73–81.
77. Kugiumtzis D. 2013 Partial transfer entropy on rank vectors. *Eur. Phys. J. Spec. Top.* **222**, 385–404. (doi:10.1140/epjst/e2013-01849-4)
78. Pompe B, Runge J. 2011 Momentary information transfer as a coupling measure of time series. *Phys. Rev. E* **83**, 051122. (doi:10.1103/PhysRevE.83.051122)
79. Nakajima K, Haruna T. 2013 Symbolic local information transfer. *Eur. Phys. J. Spec. Top.* **222**, 421–439. (doi:10.1140/epjst/e2013-01851-x)
80. Amigó JM, Monetti R, Aschenbrenner T, Bunk W. 2012 Transcripts: an algebraic approach to coupled time series. *Chaos* **22**, 013105. (doi:10.1063/1.3673238)
81. Monetti R, Bunk W, Aschenbrenner T, Springer S, Amigó JM. 2013 Information directionality in coupled time series using transcripts. *Phys. Rev. E* **88**, 022911. (doi:10.1103/PhysRevE.88.022911)
82. Monetti R, Bunk W, Aschenbrenner T, Jamitzky F. 2009 Characterizing synchronization in time series using information measures extracted from symbolic representations. *Phys. Rev. E* **79**, 046207. (doi:10.1103/PhysRevE.79.046207)
83. Monetti R, Amigó JM, Aschenbrenner T, Bunk W. 2013 Permutation complexity of interacting dynamical systems. *Eur. Phys. J. Spec. Top.* **222**, 405–420. (doi:10.1140/epjst/e2013-01850-y)
84. Bunk W, Amigó JM, Aschenbrenner T, Monetti R. 2013 A new perspective on transcripts by means of the matrix representation. *Eur. Phys. J. Special Topics* **222**, 363–381. (doi:10.1140/epjst/e2013-01847-6)
85. Amigó JM, Zambrano S, Sanjuán MAF. 2010 Permutation complexity of spatiotemporal dynamics. *Europhys. Lett.* **90**, 10007. (doi:10.1209/0295-5075/90/10007)
86. Sinn M, Keller K. 2011 Estimation of ordinal pattern probabilities in Gaussian processes with stationary increments. *Comp. Stat. Data Anal.* **55**, 1781–1790. (doi:10.1016/j.csdo.2010.11.009)