Symbolic Dynamics from Chaotic Time Series

A. Destexhe, G. Nicolis and C. Nicolis

Service de Chimie Physique, Université Libre de Bruxelles, CP231 Campus Plaine, Blvd. Triomphe, B-1050 Bruxelles, Belgium

INTRODUCTION

Following the ideas of Ruelle and others [1,2], an embedding phase space can be reconstructed from experimental systems on the basis of time series data. The introduction of numerical methods for calculating dimensions, entropies, Lyapunov exponents and other related properties, has permitted extensive investigations of chaotic experimental systems these last years. However, severe restrictions about the applicability of these methods were noticed, especially for high dimensional systems [3-6].

In this paper, we use the symbolic description as an alternative approach to analyze chaotic dynamical systems algorithms. The idea is...
Starting from an initial condition \( (X=Y=Z=1) \) the system reaches a chaotic attractor \((a=0.3, b=0.3, c=4.5)\) after \( t \approx 300 \) units of time.

The mapping into symbolic dynamics will be performed as follows [11]: once a variable crosses a prescribed threshold \( L \), a symbol specific to that variable is produced. The sequence of these symbols will then provide the image of the various thresholds encountered successively by the dynamics. The symbols are also an image of what happens in phase space. For example, an orbit rotating around the origin near the \( Z=0 \) plane will produce the symbols \( \text{XY} \). The symbol \( Z \) will be produced at each reinsertion in the plane. The sequence obtained for thresholds \( L=3 \) reads

\[
\text{ZYX ZZXY ZZXY ZXZX XYY ZYX ZXZ ZYX ZXX ZYX ZXZ ZYX} \ldots
\]

This sequence has a remarkable property: it can be entirely reformulated by introducing the hypersymbols \( \alpha=\text{ZYY}, \beta=\text{ZXY}, \gamma=\text{ZX} \). The result reads

\[
\alpha \beta \beta \alpha \beta \alpha \gamma \alpha \beta \alpha \ldots
\]

The second sequence is obviously less regular than the first one. In other words, we can say that the dynamics is made from the irregular repetition of three typical orbits around the origin, each one of which is preceded by a reinsertion (\( Z \) begins all hypersymbols).

To check more quantitatively the variability of the sequence, let us consider these sequences as a Markov process of arbitrarily high order [12]. The order of the process is the number of symbols over which the system's output can be evaluated as follows [11,13]. Let \( X_1, \ldots, X_L \) be the sequence of \( N \) symbols. The probability of observing in the sequence the word \( X_1 \ldots X_L \) is \( P(X_1 \ldots X_L) = P_i \) where \( i=1 \ldots n^L \). Here the subscript \( i \) indicates that \( X_1 \ldots X_L \) is the \( i \)th word among the \( n^L \) possible words of length \( L \) (the words are sorted in alphabetic order). For a Markov process of order \( k \), the conditional probability of observing \( X_L \) as the \( i \)th symbol in the word obeys to the relation [12]

\[
P(X_L \mid X_1 \ldots X_{L-1}) = P(X_L \mid X_{L-k} \ldots X_{L-1})
\]

The above two quantities can be calculated from the sequence and compared for different values of \( k \) until the values coincide for all possible words. It is also useful to compare \( P(X_1 \ldots X_L) \) with \( P(k)(X_1 \ldots X_L) \) which is the probability of observing the sequence \( X_1 \ldots X_L \) deduced by assuming a \( k \)th order Markov chain.

The values \( P(X_1 \ldots X_L) \) deduced from numerical counting are compared with the computed values of \( P(k)(X_1 \ldots X_L) \) by using a test statistic such as

\[
\chi^2 = \frac{1}{n^L} \sum_{i=1}^{n^L} \left[ \frac{P_i - P_i(k)}{P_i} \right]^2
\]

**Figure 1.** Test statistic \( \chi^2 \) vs. order \( k \) for the Rössler model. Symbols (triangles, \( N=5.1 \) 106) and hypersymbols (circles, \( N=2.1 \) 16) are shown. See text for other parameters.

The procedure is repeated for increasing values of \( k \) and \( \chi^2 \) must converge to zero for \( k \geq k_0 \), where \( k_0 \) is the order of the Markov process.

Figure 1 shows \( \chi^2 \) vs. \( k \) for the above described sequences. \( \chi^2 \) converges to zero at a value near 5 for the sequence of symbols. A more accurate approach leads to a value of five [13], showing that the sequence of thresholds produced in the Rössler model has a memory which extends over 5 symbols. On the other hand, the sequence of hypersymbols displays saturating behavior of \( \chi^2 \) at lower values of \( k \) (near 1). From these values, we can say that the Rössler model is made from the aperiodic repetition of three cycles.
Figure 2 depicts the time evolution of two biological signals for which a characteristic event is produced irregularly. The first signal (Fig 2.a) represents the electrocardiogram (ECG) of a normal human heart and is measured at the level of the thorax. The second signal (Fig 2.b) shows the electrical activity of the human brain (electroencephalogram or EEG) recorded during a pathological state near deep coma. The recording of the electrical potentials and the digitization were performed techniques which will not be described here (more details can be found in [4] for 2.a and in [4] for 2.b). The reconstruction of phase portraits and the evaluation of the correlation dimension [4,5,14] or Lyapunov exponents [14] all point to the same conclusion, namely that the electrical activity of the heart as well as the brain waves may follow chaotic dynamics. We will show here that the symbolic description provides an alternative approach which can be very useful to the analysis of this kind of systems [13].

![Electrical activity of the heart (ECG) and EEG](image)

**Figure 2.** (a) Electrical activity of the heart (ECG) recorded at the level of the thorax from a normal human subject. (b) Electrical activity of the human brain (EEG) during a pathological state.

We will use here a different type of mapping into symbols, which takes temporal information into account (temporal mapping). From Fig. 2, it is seen that both for ECG and EEG recordings, potential spikes are produced at irregular time intervals. In the case of EEG, the distribution of inter-spike time intervals may be bimodal. In this case, associating a symbol to each peak of the distribution is equivalent to studying the succession of “short” and “long” intervals. Therefore, in this case, it is natural to associate a symbol to the duration of the interval. In the following paragraphs, the intervals will be partitioned into subsets depending on the duration and each subset will be associated with a prescribed symbol.

![Temporal mapping](image)

**Figure 3.** Test statistic $\chi^2$ vs. order $k$ for the sequences of symbols extracted from the signals of Fig. 2: (a) ECG (N=2132), (b) EEG (N=1756). Several different partitions of the interval in $n$ symbols have been considered: $n=2$ (+), $n=3$ (X), $n=4$ (triangles) and $n=5$ (squares). (c) Third order Markov process obtained numerically on a random number generator.

Figure 3 displays the values of $\chi^2$ obtained respectively from the sequences of ECG (Fig. 2a) and EEG intervals (Fig. 2b). The order of the corresponding Markov process
deduced from the convergence of $\chi^2$ is of about three for ECG intervals (see fig 3c for comparison). In the case of the EEG intervals, $\chi^2$ converges to zero at $k = 6$ (Fig. 3b). The convergence of $\chi^2$ to zero does not depend significantly on the number $n$ of different symbols used in the partition of the interval. Actually, different values of $n$ were tested (from $n=2$ to $n=5$) and give similar results. Two results are to be retained: a) the succession of time intervals is far from being uncorrelated and b) the order of the process does not depend on the number of symbols $n$.

One of the most remarkable facts is that these sequences produce only few words with high frequency while some other words never appear. Among all the $2^9$ possible words of length $L=9$ ($n=2$), only 279 were realized by the dynamics of ECG intervals and only 151 words were seen in EEG sequences (for comparison, the Rössler model produced only 21 sequences among $3^7$ [11]). This finding shows that despite the apparent irregularity of these symbol sequences, strong "grammatical rules" may exist. These rules probably reflect the deterministic nature of the underlying dynamical system.

**CONCLUSION**

The mapping of the Rössler model into symbolic dynamics reveals highly correlated sequences of symbols, according to the strong regularities inherent in the structure of the attractor. From the reformulation of the sequence of symbols into hypersymbols, it appears that the dynamics reduces to a stochastic succession of three characteristic orbits.

Biological time series where the same pattern of activity is repeated at irregular time intervals can be studied through symbolic dynamics. The duration of the time interval is partitioned into several typical lengths, each corresponding to a defined symbol. The evaluation of the order of the sequence obtained from heart and brain recordings leads to unexpected high values. Despite the apparent randomness of these two rhythms, the intervals between spikes are remarkably correlated. A simple model of a noise-driven limit cycle is therefore insufficient to describe this activity. Following the previous work where chaotic dynamics were observed from the evaluation of the attractors dimension, such strong correlations provide further evidence that the non-linear temporal evolution of these important physiological states stems from chaotic dynamics.

The evaluation of the order of the Markov process obtained from the mapping of a chaotic system provides interesting insights into the system's dynamics, which are complementary to the information gained by analyzing the topological properties of the attractor. Moreover, this approach depends on relatively simple algorithms and can be used independently of any specific hypothesis concerning the reconstruction of phase space.

**REFERENCES**