Time reversal, symbolic series and irreversibility of human heartbeat

Camillo Cammarota * Enrico Rogora †

Abstract

We study the time reversal properties of time series by means of a ternary coding of the differentiated series. For the symbolic series obtained in this way we show that suitable pairs of ternary words have the same probability if the time series is reversible. This provides tests in which time reversibility is rejected if the estimated probabilities are significantly different. We apply one of these tests to the human heartbeat series extracted from 24-hours Holter recordings of 19 healthy subjects. Data analysis shows a highly significant prevalence of irreversibility. Our symbolic approach to time reversal gives further support to the suitability of non linear modeling of the normal heartbeat.

Keywords: Time reversal, heartbeat, symbolic dynamics, nonlinearity, time series, ternary word, RR interval.
Mathematics Subject Classification 2000: 62P10, 92C55

1 Introduction

The heartbeat time series is defined as the sequence of time intervals between consecutive R peaks (RR intervals) in the ECG. This series reflects the physiological control mechanisms of the heart rate, which act on several time scales. The main mechanism is the autonomous nervous system, which controls heart rate via the sympathetic and vagal terminations; in particular vagal control acts on a time scale of a few seconds, hence it is related to the high frequency spectral components of the RR sequence [17].
In time series extracted from physiological measurements, one often uses a method of analysis based on some kind of symbolic coding. For details of these methods

*Dept. of Mathematics, University of Rome ”La Sapienza”, cammar@mat.uniroma1.it
†Dept. of Mathematics, University of Rome ”La Sapienza”, rogora@mat.uniroma1.it

and some applications to the analysis of the heartbeat time series we refer for instance to [1, 2, 3, 4, 5, 6, 7, 10, 11, 14, 15, 18, 20].

The basic idea is to partition the range of the series into intervals and to associate a symbol to each interval. If X_i, $i = 1,\ldots,n$ denotes the series the coding used here is based on the differentiated series $D_i = X_{i+1} - X_i$, $i = 1,\ldots,n-1$. Given a positive threshold a, the symbolic series S_i is defined as

$$S_i = \begin{cases} 1 & \text{if} \quad D_i > a \\ -1 & \text{if} \quad D_i < -a \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

If the series X_i, $i = 1,\ldots,n$ is modeled as a stationary sequence of random variables, one is interested in the joint distribution of the variables $X_1,\ldots,X_k$, where k is a small integer. The estimate of this distribution cannot be fully accomplished with the available data; however one can still get useful information from the distribution of the symbolic variables $S_1,\ldots,S_{k-1}$. In the present paper we show how to use the distribution of $S_1,\ldots,S_{k-1}$ to provide a test for the hypothesis of time reversibility of the series X_i, $i = 1,\ldots,n$.

Our approach, based on symbolic series, is different from the usual ones reviewed for example in [16]. See also [12] for further results.

Basic models used in the analysis of time series are linear Gaussian ARMA models. It is known that for stationary linear models the time reversibility is equivalent to the normality of the innovations [19]. For this reason, time irreversibility has been used to test non linearity [8, 9].

Non linear properties of human heartbeat are widely accepted (see for instance [13]). Time irreversibility of this series has been investigated for the first time in [7], using a multi scale index of asymmetry. In the present paper we approach the problem of time irreversibility of human heartbeat using a short scale analysis motivated by the fact that the vagal control acts on a time scale of a few seconds.

We have tested the hypothesis of time reversibility of 19 RR series of healthy subjects. This hypothesis is rejected in most cases at high level of significance (see Table 1). This provides further support to the suitability of nonlinear modeling of these series.

The content of the paper is the following. In the next section we prove that the probabilities of conjugate words are equal for time reversible series. In the third section we show how this result provides a test for reversibility and, in the last section, we apply this test to the heartbeat series.

2 Time reversibility properties of the symbolic series

We consider a random vector $X_1,\dots,X_k$ having probability distribution P and the vector of the consecutive differences $D_i = X_{i+1} - X_i$, $i = 1,\dots,k-1$.

Definition 1 *A random vector $X_1,\dots,X_k$ is time reversible if*

$$P(X_1 \in I_1,\dots,X_k \in I_k) = P(X_k \in I_1,\dots,X_1 \in I_k)$$

for any k-tuple $I_1,\dots,I_k$ of intervals of $\mathbb{R}$.

We say that a stationary sequence $X_1,\dots,X_n$ is time reversible if for any $k \le n$, the vector $X_1,\dots,X_k$ is time reversible.
In this paper we assume that the vector $X_1,\dots,X_k$ has a k-variables density function $f(x_1,\dots,x_k)$; hence reversibility is equivalent to

$$f(x_1,\dots,x_k) = f(x_k,\dots,x_1) \tag{2}$$

Let $a > 0$ be a fixed positive number and let A_i be an event of one of the following three possible types

$$\{D_i > a\}, \qquad \{D_i < -a\}, \qquad \{|D_i| \le a\}, \qquad i = 1,\dots,k-1 \tag{3}$$

We define

$$A_i^* = \begin{cases} \{D_{k-i} < -a\} & \text{if} \quad A_i = \{D_i > a\} \\ \{D_{k-i} > a\} & \text{if} \quad A_i = \{D_i < -a\} \\ \{|D_{k-i}| \le a\} & \text{if} \quad A_i = \{|D_i| \le a\} \end{cases}$$

If $E = A_1 \cap \cdots \cap A_{k-1}$, we define $E^* = A_1^* \cap \cdots \cap A_{k-1}^*$. The events E and E^* are said to be *conjugate*.
The test of time reversibility which we use in data analysis is based on the following result.

Theorem 1 *If the random vector $X_1,\dots,X_k$ is time reversible then, for each event $E = A_1 \cap \cdots \cap A_{k-1}$ as above, one has*

$$P(E) = P(E^*)$$

Proof Given

$$E = A_1 \cap \cdots \cap A_{k-1} \tag{4}$$

let us choose, for each of the three types of events (3), an event A of this form in the sequence $(A_1, A_2,\dots,A_{k-1})$, if the type occurs. Let

$$A_i = \{D_i > a\}, \qquad A_j = \{D_j < -a\}, \qquad A_h = \{|D_h| \le a\} \tag{5}$$

be the chosen events which, for some E's, may be less than three (for instance if $E = \{D_1 > a\} \cap \{D_2 > a\} \cap \cdots \cap \{D_{k-1} > a\}$). Since each of the $A_1, \dots, A_{k-1}$ is of one of the types (3), the following computation indicates what to do for each E.
One has

$$P(E) = P(\cdots \cap \{D_i > a\} \cap \cdots \cap \{D_j < -a\} \cap \cdots \cap |\{D_h| \le a\} \cap \dots) =$$
$$P(\cdots \cap \{X_{i+1} - X_i > a\} \cap \cdots \cap \{X_{j+1} - X_j < -a\} \cap \cdots \cap \{|X_{h+1} - X_h| \le a\} \cap \dots) \tag{6}$$

For computing $P(E)$ we fi rst integrate on a domain which is in normal form w.r.t. the variable x_1.

$$P(E) = \int_{-\infty}^{+\infty} \mathrm{d}x_1 \cdots \int_{x_i+a}^{+\infty} \mathrm{d}x_{i+1} \cdots \int_{-\infty}^{x_j-a} \mathrm{d}x_{j+1} \cdots \int_{x_h-a}^{x_h+a} \mathrm{d}x_{h+1} \cdots f(x_1, \dots, x_k) \tag{7}$$

Making the change of variables $\xi_l = x_{k+1-l}$, $l = 1, \dots, k$ in the integral (7) and using time reversibility, we get

$$P(E) = \int_{-\infty}^{+\infty} \mathrm{d}\xi_k \cdots \int_{\xi_{k-i+1}+a}^{+\infty} \mathrm{d}\xi_{k-i} \cdots \int_{-\infty}^{\xi_{k-j+1}-a} \mathrm{d}\xi_{k-j} \cdots \int_{\xi_{k-h+1}-a}^{\xi_{k-h+1}+a} \mathrm{d}\xi_{k-h} \cdots f(\xi_1, \dots, \xi_k) \tag{8}$$

Note that

1. $\xi_{k-i} > \xi_{k-i+1} + a$ is equivalent to $\xi_{k-i+1} < \xi_{k-i} - a$;
2. $\xi_{k-j} < \xi_{k-j+1} - a$ is equivalent to $\xi_{k-j+1} > \xi_{k-j} + a$;
3. $\xi_{k-h+1} - a < \xi_{k-h} < \xi_{k-h+1} + a$ is equivalent to $\xi_{k-h} - a < \xi_{k-h+1} < \xi_{k-h} + a$;

hence, by writing the domain of integration in normal form w.r.t. the variable ξ_1, integral (8) can be written as

$$P(E) = \int_{-\infty}^{+\infty} \mathrm{d}\xi_1 \cdots \int_{-\infty}^{\xi_{k-i}-a} \mathrm{d}\xi_{k-i+1} \cdots \int_{\xi_{k-j}+a}^{+\infty} \mathrm{d}\xi_{k-j+1} \cdots$$
$$\cdots \int_{\xi_{k-h}-a}^{x_{k-h}+a} \mathrm{d}\xi_{k-h+1} \cdots f(\xi_1, \dots, \xi_k) =$$
$$P(\cdots \cap \{D_{k-i} < -a\} \cap \cdots \cap \{D_{k-j} > a\} \cap \cdots \cap \{|D_{k-h}| \le a\} \cap \dots) = P(E^*) \tag{9}$$

□

3 Testing time reversibility

Let $X_1, \ldots, X_n$ be a stationary time series and let $S_1, \ldots, S_{n-1}$ be the symbolic series defined in (1). We consider the pair of events

$$W_{k-1}^+ = \bigcap_{i=1}^{k-1} \{D_i > a\}; \qquad W_{k-1}^- = \bigcap_{i=1}^{k-1} \{D_i < -a\}.$$

Note that these events are conjugate, i.e.

$$(W_{k-1}^+)^* = W_{k-1}^-.$$

The consideration of these events has already revealed fruitful in the analysis of heartbeat series, allowing for example to discriminate between groups of age [11]. Our test is based on the fact that if the vector $X_1, \ldots, X_k$ is time reversible then

$$P(W_{k-1}^+) = P(W_{k-1}^-). \tag{10}$$

Unbiased estimators of the above probabilities are given by the frequency of occurrence of the corresponding words in the symbolic series. At this aim we define

$$N_k^- = \sum_{i=1}^{n-k+1} \chi_{\{S_i = S_{i+1} = \cdots = S_{i+k-2} = -1\}}; \quad N_k^+ = \sum_{i=1}^{n-k+1} \chi_{\{S_i = S_{i+1} = \cdots = S_{i+k-2} = 1\}} \tag{11}$$

where χ denotes the characteristic function of the event in its argument. In other words we count the strings such that all the differences in them are greater than the threshold a. The estimators of the probabilities are

$$\hat{p}_k^- = \frac{1}{n-k+1} N_k^-; \quad \hat{p}_k^+ = \frac{1}{n-k+1} N_k^+. \tag{12}$$

In the hypothesis of reversibility the two probabilities are equal and so our test is based on the following: the hypothesis is rejected if the estimated $\hat{p}_k^+, \hat{p}_k^-$ are significantly different.

4 Data analysis

In this section we apply the above results to test time reversibility of heartbeat series.

We consider the RR time series extracted from the 24 hours Holter recording of a group of 19 healthy subjects. The software (ELA Medical) provided a classification of the beats in several categories: normal, artifact and so on. We select the

$\mathbf{p < 0.001}$	$\mathbf{0.001 < p < 0.05}$	$\mathbf{p > 0.05}$
16	2	1

Table 1: In the first row the significance intervals for the p-values for the sign test. In the second row the number of cases in each interval out of 19.

normal words which are formed only by normal beats. In our data analysis we have chosen $a = 10$ *ms* as the value for the threshold and $k = 3$ for the length of words. The number of normal words range from 80 % to 99% of the total, which is of the order of 10^5 for each subject.

It is well known that the RR series is non stationary on the scale of the 24 hours: this is evident for example in the first panel of Figure 1 where the night segment is markedly shifted up-words. Hence we divide the series into m adjacent segments of length $n = 1000$. This value has been chosen for two reasons. First one can assume that the RR series, and a fortiori the symbolic series, is stationary over this time scale. Second, a segment of length 1000 contains a number of normal words which is sufficiently large for providing an accurate estimate of the probability of each word; of course this is possible since we have chosen to perform a short scale analysis.

We denote $N_k^+(i)$ and $N_k^-(i), \quad i = 1,\dots,m$ the values of N_k^+ and N_k^- computed on the $i-$th segment. This defines two paired sequences

$$N_k^+(i), \quad i = 1,\dots,m; \quad N_k^-(i) \quad i = 1,\dots,m$$

that were computed for the RR time series of each subject in our 19 cases database for $k = 3$. We performed a sign test for the differences

$$\Delta_3(i) = N_3^+(i) - N_3^-(i)$$

assuming that $\Delta_3(i)$, $i = 1,\dots,m$ are independent and equally distributed. The results are summarized in Table 1.

This test clearly suggests that normal RR time series are generally time irreversible.

In order to investigate further the features of this irreversibility we say that N_k^+ (resp. N_k^-) *dominates* N_k^- (resp. N_k^+) if the median of N_k^+ (resp. N_k^-) is greater than the median of N_k^- (resp. N_k^+) and the sign test is significant. In our data analysis we have fixed a significance level of 0.05 and we have found that in 15 cases out of 19 N_3^+ dominates N_3^- and in 3 cases out of 19 N_3^- dominates N_3^+.

A typical case of the first group is shown in Figure 1. The two main qualitative features are shown in the second panel. First the two profiles N_3^+ and N_3^- are not stationary on a scale larger than 1000 beats and this reflects the non-stationarity

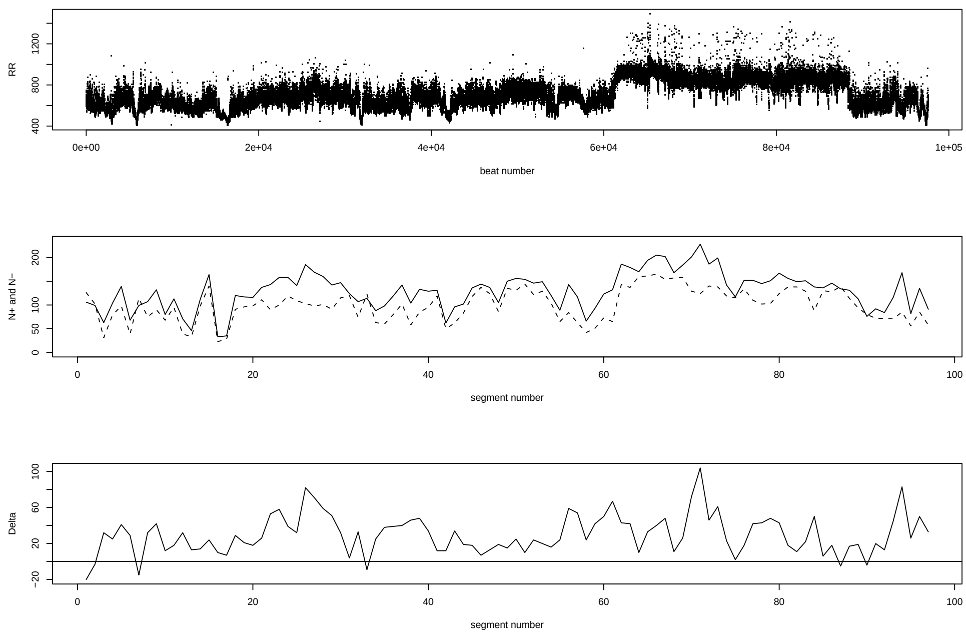

Figure 1: In the first panel the 24 hours RR sequence of a normal subject. In the second panel: the values of $N_3^+(i)$ are in solid lines and the values of $N_3^-(i)$ are in dashed lines, both referring to segments of 1000 beats. In the third panel the sequence of delta indexes.

of the RR sequence itself. Second the relationship of domination between N_3^+ and N_3^- comes from highly correlated profiles.

Acknowledgments We thank Prof. G. Guarini for useful discussion and Dr. M.Ambrosini of Policlinico Umberto I of our University for providing us the data.

References

[1] Albert, C.-C. Yang, Shu-Shya Hseu, Huey-Wen, Yien, Ary L. Goldberger, C.-K. Peng "Linguistic Analysis of the Human Heartbeat Using Frequency and Rank Order Statistics", *Phys. Rev. Lett.*, **90** 108103-1,108103-4 (2003).

[2] Ashkenazy Y., Yvanov P.Ch., Havlin S.,C-K. Peng, Goldberger A. L., Stanley H.E. "Magnitude and Sign Correlations in Heartbeat Fluctuations", *Phys. Rev. Lett.*, **86**,1900-1903 (2001).

[3] Bandt C., Pompe B. "Permutation Entropy: A natural Complexity Measure for Time Series", *Phys. Rev. Lett.* **88** 174102-1:174102-4 (2002).

[4] Cammarota C., Guarini G., Ambrosini M. "Analysis of stationary periods of heart rate via symbolic dynamics", *Medical Data Analysis*, Lecture Notes in Computer Science vol. 2526, Colosimo, Giuliani and Sirabella eds. Springer 2002

[5] Cammarota C. Rogora E. "Testing independence in time series via universal distributions of permutations and words" *Int. J. Bif Chaos*, **15**, 1757-1765, (2005)

[6] Cammarota C. Rogora E. "Independence and symbolic independence of nonstationary heartbeat series during atrial fibrillation" *Physica A*, **353**, 323-335, (2005)

[7] Costa M., Goldberger A.L., C.-K. Peng "Broken asymmetry of the human heartbeat: loss of time irreversibility in aging and disease", *Phys. Rev. Lett.*, **95**, 198102, (2005).

[8] Diks C., van Houwelingen J.C., Takens F., DeGoede J. "Reversibility as a criterion for discriminating time series" *Physics Letters A* **201**, 221-228 (1995).

[9] Diks C. *Nonlinear Time Series Analysis*, World Scientific, 1999.

[10] Garcia-Gonzales M. A., Ramos-Castro J., Fernandez-Jimeno M. “A new index for the analysis of heart rate variability dynamics: characterization and application”, *Physiol. Meas.* **24** 819-832 (2003).

[11] Germanò G., Piccirillo G., Cammarota C., Guarini G., Rogora E., Cacciafesta M. “A new measure of acceleration of heart rate: dependence on age and comparison with time domain conventional heart rate variability measures”, available on line at the Internet location `http://w3.uniroma1.it/CISB/Cisb/fr_documents.html`

[12] Giannakis, G.B., Tsatsanis, M.K. “Time-Domain Tests for Gaussianity and Time-Reversibility” *IEEE Trans. Sign. Proc.* **42**,3460-3472 (1994).

[13] Irurzun I. M., Bergero P., Cordero M. C., Defeo M. M., Vicente J. L., Mola E. E., “Non-linear properties of R-R distributions as a measure of heart rate variability”, *Chaos, Soliton & Fractals*, **16** 699-708 (2003).

[14] June-Soo Kim et al. “Decreased entropy of symbolic heart rate dynamics during daily activity as a predictor of positive head up tilt test in patients with alleged neurocardiogenic syncope”, *Phys. Med Biol.* **45** 3403-3412 (2000).

[15] Keller K., Lauffer H., “Symbolic analysis of high dimensional time series”, *Int. J. Bifurcation and Chaos* **13** 2657-2668 (2003).

[16] Lawrance, A. J. “Directionality and Reversibility in Time Series” *Int. Stat. Rev.* **59** 67-79 (1991).

[17] Malliani A, Pagani M, Lombardi F, Cerutti S. ”Cardiovascular neural regulation explored in the frequency domain” *Circulation* **84**, 482-492 (1991).

[18] Meyerfeldt U. et al. “Heart rate variability before the onset of ventricular tachycardia: differences between slow and fast arrhythmias”, *International journal of Cardiology* **84** 141-151, (2002).

[19] Weiss G. “Time reversibility of linear stochastic processes” *J. Appl. Prob.* **12**, 831-836 (1975).

[20] Zebrowski J.J., Poplawska W., Baranowski R., Buchner T. “Symbolic dynamics and complexity in a physiological time series” *Chaos, solitons & Fractals* **11** 1061-1075 (2000).