

# Entropies, Partitionings and Heart Rate Variability

Milan Paluš<sup>1</sup> and Jan J. Żebrowski<sup>2</sup>

<sup>1</sup> Institute of Computer Science, Academy of Sciences of the Czech Republic  
Pod vodárenskou věží 2, 182 07 Prague 8, Czech Republic

E-mail: mp@uivt.cas.cz, mp@santafe.edu

<sup>2</sup> Institute of Physics, Warsaw University of Technology  
ul. Koszykowa 75, 00-662 Warszawa, Poland

E-mail: zebra@if.pw.edu.pl

**Abstract.** Several definitions of static entropies and coarse-grained entropy rates CER's were used to analyze 24-hour RR interval sequences of three sex, age and disease matched control pairs. The matching was done with the assessment of the risk of cardiac arrest in mind. Details of estimating probability densities (histograms), used to compute the static entropies, have been found crucial to obtain the proper discrimination between the members of control pairs. The effect of signal variance on the entropy estimation is discussed. In a preliminary result, CER's have been found to yield additional information which may be helpful in discriminating health and pathology.

## 1 Introduction

This study was inspired by the basic question, which is often asked by all time-series analysts dealing with physiological data, namely, which property of the data discriminates health from pathology and how this property can be described and measured, so that a reliable diagnostic tool is obtained. The data investigated were RR interval series, forming 3 sex, age and disease matched control pairs. These data pairs were chosen to test whether any given technique of nonlinear dynamical analysis is able to correctly distinguish between a person at risk of a cardiac arrest (CA) and his/her control. Below the term “health, healthy” will mean control, while the term “pathology” will signify a person with a high risk of the cardiac arrest. Note, that a sex, age and disease matched control medically resembles the target person in as many respects as possible, except for the single feature investigated – in this case the risk of the cardiac arrest.

It has been questioned, whether the health-pathology differences are encoded in the dynamics of the RR interval series, characterized by their entropy rates, or in the geometry of “phase portraits”, obtained by time-lagging the series and characterized by 2- and 3-dimensional Shannon entropies, or even in the static marginal distributions (histograms) of the series, characterized

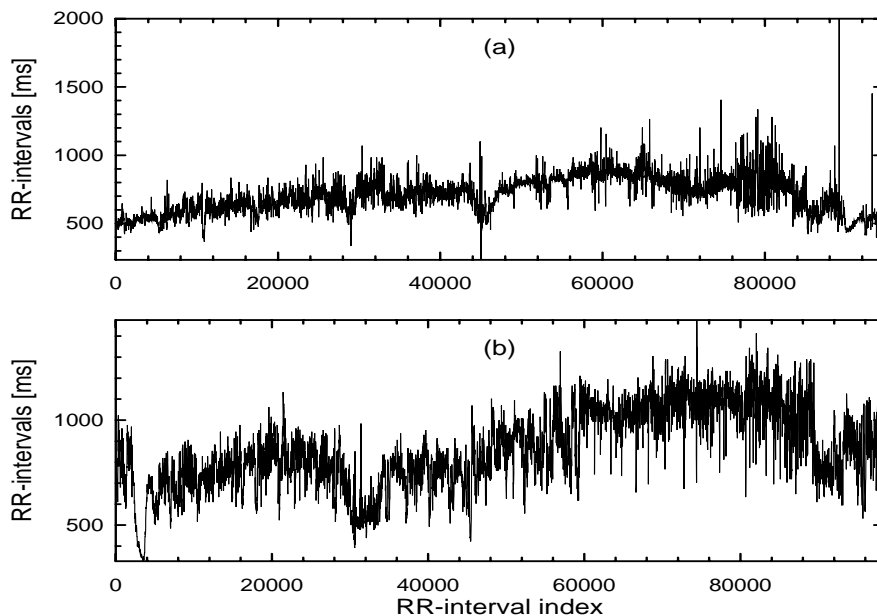
by 1-dimensional Shannon entropies, and eventually in the higher-order properties of the latter. The higher-order properties of the probability distributions are characterized by Rényi entropies and apparently also by complexity measures called pattern entropies as introduced by Żebrowski et al. (1994). Considering the amount of processed recordings and the short time of the investigation, it is not surprising, that we are not able to give the final answer to the question asked here. However, we present a report on the work in progress, which was initiated at the NTPTSA workshop and which contains some interesting findings related to practical estimations of entropies and entropy-related quantities from experimental data. It was found that the results were much more influenced by a particular method of estimating the entropies (more precisely – by a particular method of computing the histograms as estimates of probability distributions used to calculate the entropies themselves) than by the choice of a particular entropy measure. All the static entropies considered below brought consistent results when the same histogram estimation method was used. The number of data series processed was unfortunately not sufficient to decide which of the estimation methods is the best one for the medical problem from which the data studied here were taken.

Dynamic entropies – the entropy rates, estimated by yet another approach, showed to be consistent in some features related to the discrimination between health and pathology. Their actual diagnostic efficacy, however, should be established by a more extended study.

## 2 Data

Holter ECG 24-hour tapes of both healthy individuals and of cases of heart disease with the highest risk of sudden death were analyzed by commercial software (Del Mar Avionics 563 Strata Scan) at the National Institute of Cardiology in Warszawa, Poland. The data were sampled at 128 Hz and the time distance between consecutive R peaks (the RR intervals) were extracted. No special filtering was used, but RR intervals larger than 2500 ms were treated as artifacts and ignored. Each of the studied 24-hour sequences of RR intervals was 80000 to 125000 data points long and began at about 8 o'clock in the morning. The data were used to create medically matched control pairs.

At the time of the conference there were 3 sex, age and disease matched control pairs available: all target members of the pairs had had at least a single event of cardiac arrest (CA) in their history. No arrhythmia filters were used. It should be stressed here that the method used to find the appropriate control pair to a given target patient was based on time honoured medical procedures alone. This means that the choice was based only on general medical (clinical) knowledge. Neither nonlinear methods were used to form the sex, age and disease matched pairs, nor time domain and frequency domain

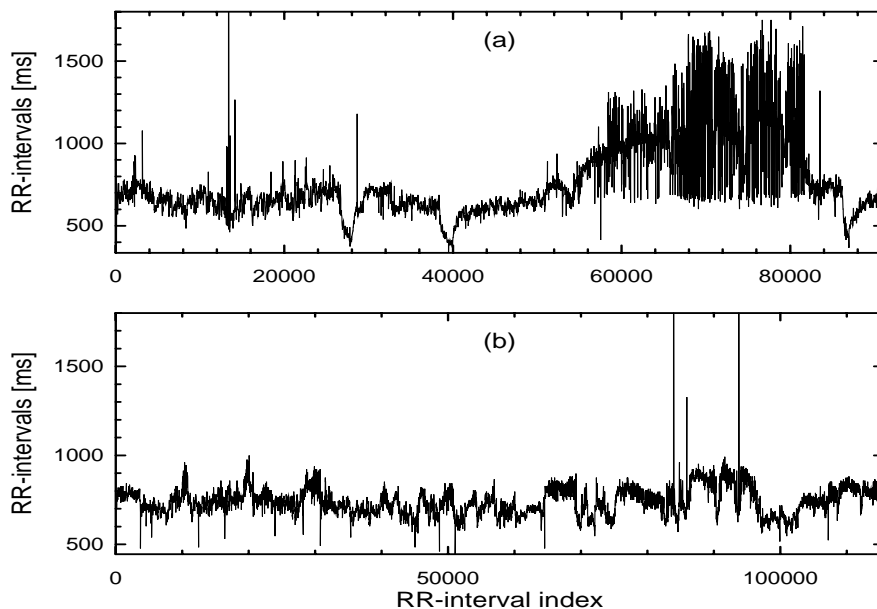


**Fig. 1.** Plot of 24-hour RR-interval series *dp1a* (a) – the 27 year old male CA patient and *b*) his control *dp1b*.

measures of ECG were applied, and the RR intervals as functions of the time were not inspected. The members of pair 1 and pair 3 who had undergone CA were labeled “a”, their controls were labeled “b”, while the labels of pair 2 were reversed. This was done in order to make the comparison of nonlinear techniques resembling a blind test.

In pair 1, the RR interval series labelled *dp1a* (Fig. 1a) was taken from a 27 year old male who had had CA twice and now wears an implanted defibrillator (which, however, did not activate itself during the Holter recording analysed). All known methods of the ECG analysis in the time and frequency domains, as well as the biopsy of the heart muscle and biochemical analysis, all pronounced him to be in exemplary health. However, pattern entropy (Zebrowski et al., 1994) was found to be at an elevated level which distinguished him well from his control pair *dp1b* (Fig. 1b).

In pair 2, the RR interval series labelled *dp2b* (Fig. 2b) was taken from a 55 year old male who had undergone a myocardial infarction some time ago, complicated by a postinfarction aneurism. He died several months after having been included in the 15 control pairs analysed by pattern entropy and was pronounced at a high risk of CA on the basis of the behavior of this



**Fig. 2.** Plot of 24-hour RR-interval series *dp2a* (a) – the control for *dp2b*; b) *dp2b* – 55-year old male CA patient.

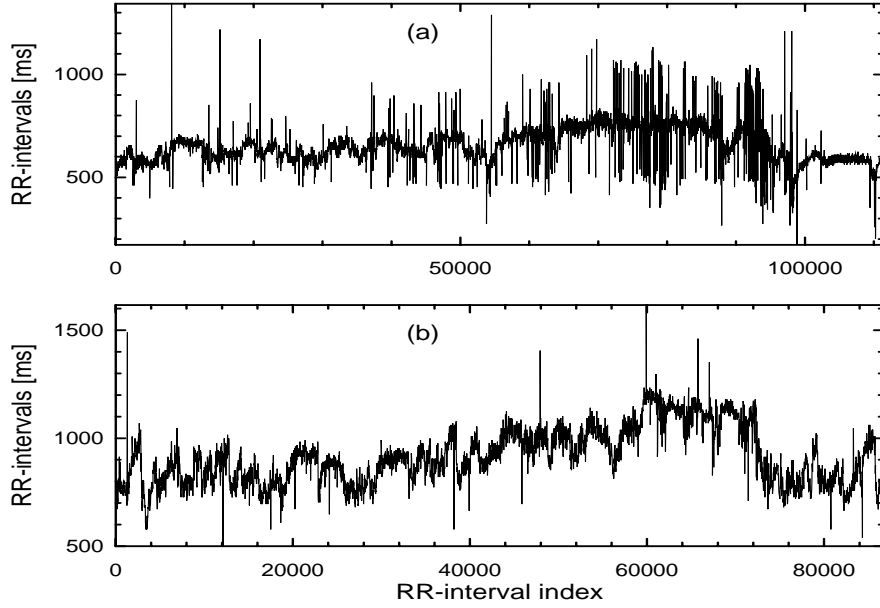
complexity measure. The control for this CA patient is depicted in Fig. 2a.

Pair 3 is composed of two recordings of the same person - a 67 year old male also some time after a myocardial infarction. The recording of *dp3a* (Fig. 3a) contains data taken prior and during a CA (when a ventricular fibrillation occurred), subsequent reanimation and ends 1 hour after the CA event. The series *dp3b* (Fig. 3b) was recorded 9 months afterwards and shows the results of medication. Thus, this person became his own sex, age and disease matched control pair.

### 3 Entropies and Entropy Rates

In this section basic definitions of entropies and entropy rates are briefly reviewed. More details can be found in Refs. (Cover and Thomas, 1991), (Fraser, 1989), (Sinai, 1976), (Petersen, 1983), (Paluš, 1993) and references therein.

Let  $X$  be a discrete random variable with a set of values  $\Xi$  and probability mass function  $p(x) = \Pr\{X = x\}$ ,  $x \in \Xi$ . We denote the probability mass function by  $p(x)$ , rather than  $p_X(x)$ , for convenience.



**Fig. 3.** Plot of 24-hour RR-interval series dp3a (a) – the 67 year old male CA patient and b) his control dp3b – the same person after 9 months of medication.

The *Shannon entropy*  $H(X)$  of a discrete random variable  $X$  is defined by

$$H(X) = - \sum_{x \in \mathcal{E}} p(x) \log p(x). \quad (1)$$

For  $n$  variables  $X_1, \dots, X_n$  with the joint distribution  $p(x_1, \dots, x_n)$  the joint ( $n$ -dimensional) Shannon entropy is defined as

$$H(X_1, \dots, X_n) = - \sum_{x_1 \in \mathcal{E}_1} \dots \sum_{x_n \in \mathcal{E}_n} p(x_1, \dots, x_n) \log p(x_1, \dots, x_n). \quad (2)$$

The *marginal redundancy*  $\varrho(X_1, \dots, X_{n-1}; X_n)$  is an information-theoretic measure which quantifies average amount of information about the variable  $X_n$ , contained in the variables  $X_1, \dots, X_{n-1}$ , and is defined as

$$\varrho(X_1, \dots, X_{n-1}; X_n) = H(X_1, \dots, X_{n-1}) + H(X_n) - H(X_1, \dots, X_n). \quad (3)$$

In the case of two variables  $X_1, X_2$ , the redundancy  $\varrho(X_1; X_2)$  is also known as mutual information  $I(X_1; X_2)$ .

Now, let  $\{X_i\}$  be a stochastic process, i.e., an indexed sequence of random variables, characterized by the joint probability mass function  $p(x_1, \dots, x_n) = \Pr\{(X_1, \dots, X_n) = (x_1, \dots, x_n)\}$ ,  $(x_1, \dots, x_n) \in \Xi_1 \times \dots \times \Xi_n$ . The entropy rate of  $\{X_i\}$  is defined as

$$h = \lim_{n \rightarrow \infty} \frac{1}{n} H(X_1, \dots, X_n), \quad (4)$$

and can be expressed using the marginal redundancy  $\varrho(X_1, \dots, X_{n-1}; X_n)$  as

$$h = \lim_{n \rightarrow \infty} [H(X_n) - \varrho(X_1, \dots, X_{n-1}; X_n)]. \quad (5)$$

In practical applications a time series  $\{y(t)\}$  is considered as a realization of a stationary and ergodic stochastic process  $\{Y(t)\}$ . Due to ergodicity the entropies and redundancies can be estimated using time averages instead of ensemble averages, and, the variables  $X_i$  are substituted as

$$X_i = y(t + (i - 1)\tau). \quad (6)$$

Due to stationarity, entropies

$$H^n(\tau) \equiv H(y(t), y(t + \tau), \dots, y(t + (n - 1)\tau)) \quad (7)$$

and marginal redundancies

$$\varrho^n(\tau) \equiv \varrho(y(t), y(t + \tau), \dots, y(t + (n - 2)\tau); y(t + (n - 1)\tau)) \quad (8)$$

are functions of  $n$  and  $\tau$ , independent of  $t$ .

Then, the entropy rate of  $\{y(t)\}$  can be written as

$$h_\tau = \varrho^n(0) - \lim_{n \rightarrow \infty} \varrho^n(\tau). \quad (9)$$

For the entropy rate  $h_1$  per a time unit the following equation holds (Sinai, 1976)

$$h_\tau = |\tau| h_1. \quad (10)$$

The possibility to compute the entropy rates from data is limited to a few exceptional cases: for stochastic processes it is possible, e.g., for finite-state Markov chains (Cover and Thomas, 1991). If a time series was generated by a dynamical system, its entropy rate, known as Kolmogorov-Sinai entropy (KSE), can, in principle, be estimated from the asymptotic dependence of the marginal redundancy on time delay (Fraser, 1989):

$$\lim_{n \rightarrow \infty} \varrho^n(\tau) = H^1 - |\tau| h. \quad (11)$$

It was shown (Fraser, 1989), (Paluš, 1993), (Paluš, 1996b) that the asymptotic behavior  $\varrho^n(\tau) \approx H^1 - |\tau| h$  could be observed in  $\varrho^n(\tau)$  estimated from time series, however, the data requirements (series length, precision, negligible

noise) are not realistic for usual experimental situations such as those in the analysis of physiological time series.

Instead of unsuccessful attempts to estimate the exact entropy rates from real data, Paluš (1996b) has proposed “coarse-grained entropy rates” (CER’s), computed from the (coarse-grained) marginal redundancy or mutual information. The simplest CER  $h^{(0)}$  is defined as

$$h^{(0)} = \frac{\varrho^n(\tau_0) - \varrho^n(\tau_1)}{\tau_1 - \tau_0}, \quad (12)$$

where  $\tau_0$  and  $\tau_1$  are parameters of the method to be chosen by a user.

For an alternative CER  $h^{(1)}$ , which has better numerical properties than  $h^{(0)}$ , one should compute the marginal redundancies  $\varrho^n(\tau)$  for all analyzed datasets and find such  $\tau_{max}$  that for  $\tau' \geq \tau_{max}$ :  $\varrho^n(\tau') \approx 0$  for all the datasets. Then a norm of the marginal redundancy is defined

$$\|\varrho^n\| = \frac{\sum_{\tau=\tau_0}^{\tau_{max}} \varrho^n(\tau)}{\tau_{max} - \tau_0}. \quad (13)$$

Here the lag  $\tau_{max}$  is derived from the data, the lag  $\tau_0$  is usually set to zero.

Having defined the norm  $\|\varrho^n\|$ , the difference  $\varrho^n(\tau_0) - \|\varrho^n\|$  can be considered as the alternative definition of the CER. It was found, however, that the definition of the CER, which does not depend on the absolute values of  $\varrho^n(\tau)$ , has better numerical properties, namely the estimates are more stable and less influenced by noise. Thus, the CER  $h^{(1)}$  is defined as

$$h^{(1)} = \frac{\varrho^n(\tau_0) - \|\varrho^n\|}{\|\varrho^n\|}. \quad (14)$$

Some aspects of the estimation of the CER’s will be discussed in the next section, for more details and description of properties of the CER’s see Ref. (Paluš, 1996b).

The Shannon entropy  $H(X) \equiv H^1$  evaluates properties of the probability function  $p(x)$ . Higher-order properties of  $p(x)$  can be characterized by *Rényi entropies*. The Rényi entropy  $H_q(X)$  of order  $q$  for a discrete random variable  $X$  is defined as

$$H_q(X) = \frac{1}{1-q} \log \sum_{x \in \mathcal{E}} p(x)^q. \quad (15)$$

Żebrowski et al. (1994) have defined “pattern entropy” as

$$Z_q(\tau) = - \sum_{x \in \mathcal{E}} P_q(\tau) \log P_q(\tau), \quad (16)$$

where

$$P_q(\tau) = p(x_1)p(x_2) \dots p(x_q). \quad (17)$$

The variables  $X_i$  are given as in Eq. (6). The quantity  $Z_q(\tau)$  is not an entropy in the strict mathematical sense. We will, however, use the term pattern

entropy, as it has been introduced by Żebrowski et al. (1994). The product  $P_q(\tau)$  can be considered as a joint probability distribution  $p(x_1, \dots, x_q)$  only if the variables  $X_1, \dots, X_q$  are independent. Then  $Z_q(\tau) = H(X_1, \dots, X_q) = H(X_1) + \dots + H(X_q)$  and for a stationary series  $Z_q = qH(X_1) \equiv qH^1$ . For general stationary series  $P_q = p(x)^q$  and

$$Z_q = - \sum_{x \in \Xi} p(x)^q \log p(x)^q,$$

which does not have usual properties of the entropy:  $Z_q$  is the larger the more ordered the analyzed signal is and the additivity properties of the usual entropy are lost. In spite of these peculiarities it has been shown that  $Z_q$  is able to distinguish between a high risk and a low risk of cardiac arrest CA (Żebrowski et al., 1994), (Żebrowski et al., 1995a), (Żebrowski et al., 1995b).  $Z_q$  is also a quantity very well correlated with the level of the hormone noreprephanine as measured early in the morning once a day: the correlation coefficient is -0.74 (Popławska et al., 1994). Recent, as yet unpublished results also show a good relation of the pattern entropy to classical quantitative measures of heart rate variability like SDNN, RMSSD and PNN50 showing that it may be possible to use the pattern entropy to monitor both the sympathetic and the parasympathetic nervous system and avoid the well known problems of the sensitivity of the classical measures to extrasystoles and artifacts.

## 4 Estimating Entropies from Experimental Data

Estimating entropies and other functionals of probability functions from continuous variables is always problematic and none of various approaches is generally accepted as the best one. Here we will consider the simplest “box-counting” approach, when the probability distribution functions  $p(x_1, \dots, x_n)$  are estimated as time-averaged histograms. A state-space partition  $\Xi_1 \times \dots \times \Xi_n$  must be defined and the probability distribution function  $p(x_1, \dots, x_n)$  is estimated as the relative frequencies of occurrence of time-series samples in particular bins. Consider first the one-dimensional case and probability mass function  $p(x) = \Pr\{X = x\}$ ,  $x \in \Xi$ . The elements of  $\Xi$ , called bins, must be defined on a part of a straight line of length  $R$ , given by the extent of the data studied,  $R = \text{maximum} - \text{minimum}$ , found in the data. The simplest way to define  $\Xi$  is using bins of the same width, called *equidistant partition*. The equidistant partitions were used here for estimations of all of the static entropies. There are still two ways how to define an equidistant partition: Given a number  $Q$  of bins, the bin width  $W$  is defined as  $W = R/Q$ ; or, given the bin width  $W$ , the number of bins is defined as  $Q = R/W$ . These approaches are not equivalent, if several time series with different extents  $R$  are evaluated. In multidimensional cases, i.e., when estimating entropies or



redundancies of  $n$  variables (which may be defined by time-lagging one variable), the  $n$ -dimensional partition  $\mathcal{E}_1 \times \dots \times \mathcal{E}_n$  will be defined by defining each marginal partition  $\mathcal{E}_i$ . Here we will consider using the same  $Q$  ( $W$ ) for each  $\mathcal{E}_i$ .

Estimating entropies and redundancies so that the number  $Q$  of bins is given and the bin width  $W$  is derived according to the range  $R$  of the data (henceforth referred to as “fixed- $Q$ ” approach) is desirable, considering the critical effect of  $Q$  on the entropy/redundancy estimates. The naive approach to estimate entropies/redundancies of continuous variables should be the use of the finest possible partition given, e.g., by an available space in computer memory or by measurement precision. We must remember, however, that we usually have a finite number  $N$  of data samples. Hence when using a too fine partition the estimation of entropies and redundancies can be heavily biased: Estimating the joint entropy of  $n$  variables using  $Q$  marginal bins one obtains  $Q^n$  boxes covering the state space. If  $Q^n$  approaches the number  $N$  of data samples, or even  $Q > N$ , the estimate of  $H(X_1, \dots, X_n)$  can be equal to  $\log(N)$ , or, in any case, it can be determined mainly by the number of the data samples and/or by the number of the distinct data values and not by the data structure, i.e. by the dynamical properties of the system under study. Thus even “natural” partitions of experimental data given by an A/D converter is usually too fine for a reliable estimation of entropies and redundancies.

Having a large number  $N$  of data samples, the estimated entropy increases with  $Q$  as  $\log Q$ . Thus, classifying time series according to their entropies, using fixed  $Q$  is necessary to obtain consistent results. In some cases, however, the fixed- $Q$  approach can have some disadvantages. For instance, consider a set of time series having their basic ranges more or less constant, however, some of the series contain a few outliers, i.e., values far outside the basic range. Then using  $W = R/Q$  with a fixed  $Q$  can lead to an underestimation of the actual entropy of the series, considering  $0 \log 0 = 0$ .

Having defined the bin width  $W$  and deriving the number of bins as  $Q = R/W$  (henceforth referred to as “fixed- $W$ ” approach) can help in the above case with the outliers. Ranges of the time series studied, however, should be approximately constant or an increased variance should be equivalent to “richer dynamics”, i.e., to an actual increase of entropy. Otherwise, two series would be classified as having different entropies, though one of them is just the other series multiplied by a constant. In their original work, Żebrowski et al. (1994) have applied the fixed- $W$  approach, considering the argument that in some cases such a linear rescaling of the data series may be of importance (i.e., we would have then RR interval series with a different variance) and the fact that entropy is larger for one of them may be used as an indication of the increase of the variance. This is why — in the case of the pattern entropy where the fixed- $W$  approach is used — the physical interpre-

tation of the changes of the value of this complexity measure are somewhat ambiguous. They may be caused both by an increase of the ordering of the series analyzed (e.g. a drop of the number of frequencies present in the signal) and by a change in variance. So in spite of the fact that the pattern entropy appears to be good at yielding a statistical differentiation between high risk and low risk CA cases, other complexity measures are sought (Żebrowski et al., 1995c).

An approach, different from the above two, was used to estimate the coarse-grained entropy rates CER's. The CER's were computed from mutual information, which was estimated by a fixed- $Q$  box-counting algorithm, however, marginal partitions were not equidistant, but "equiquantal". This means that the marginal bins were defined in such a way that they contained approximately the same number of samples, while their widths were not constant. I.e., effectively the data were transformed into a uniform distribution with marginal entropy  $H^1 = \log Q$  and all dynamical information was obtained from the joint probability functions/histograms. This approach eliminates the effects of outliers. For more details see Ref. (Paluš, this volume) and references therein.

## 5 Numerical Details

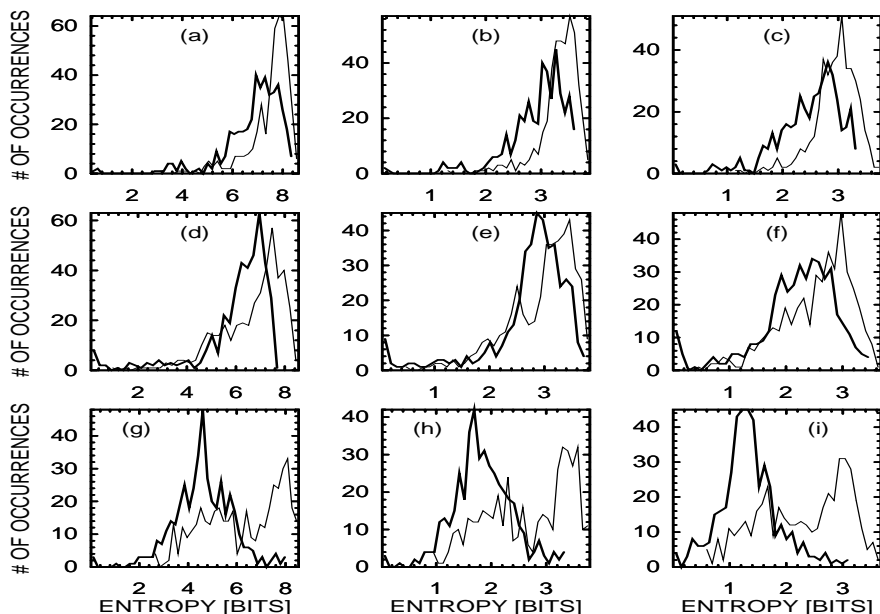
The 24-hour time series of RR intervals are nonstationary, while the entropy measures are defined for stationary processes. Therefore in practical processing we do not estimate the entropy measures from the whole records, but from a short window, which "slides" through the time series in an overlapping way. It is believed, that the violation of the stationarity condition inside a window is not so critical as compared to the whole recording. Thus, in choosing the window length, one needs to compromise between the basic need to use a long enough part of the data series to obtain mathematically meaningful results and the necessity to make this series as short as possible to be able to analyze nonstationary states. Initially, the approach adopted for pattern entropy calculations by Żebrowski et al. (1994,1995a) was to use stretches of time such as a doctor analyzing an ECG would use. It turned out that usually 5 min. epochs of ECG are used for power spectral density analysis and that a doctor analyzes an ECG visually from recordings of lengths from 30 s to 5 min. However, using real time creates ambiguities as heart variability changes within a single recording greatly so that the number of data points in a time epoch of predefined length is far from constant. Instead of a real time window, a window with a fixed number of evolutions was adopted. In the initial study a window of 400 beats was used (the equivalent of 3-7 min. of real time) and found to give good statistical predictions on the risk of CA. It has also been shown (Żebrowski et al., 1995b) that a statistically reliable result

in distinguishing between low risk and high risk groups of CA is obtained for  $Z_q$  when an apparently too short time window is used (e.g. 50 heart beats or about 30-50 s). For the fixed bin width approach this means that some of the bins of the histogram for a given window may be empty or contain only very few data points. For an estimate of a probability distribution this seems to be inadequate. However, when the window was slid along the data series with a step of every single beat (Żebrowski et al., 1994), it was found that the pattern entropy as a function of time is a reasonably smooth curve. This result indicates that, although some of the information is lost when static entropies are calculated within a single window, sliding the window every single evolution of the system helps alleviate this problem. The mean, maximum and minimum of pattern entropy calculated in such a way allows to distinguish between a low and a high risk of CA both for 24-hour Holter ECG (Żebrowski et al., 1994), (Żebrowski et al., 1995b) and for short Holter (14 min.) ECG recordings (Baranowski et al., 1994). 24-hour histograms of pattern entropy of RR intervals together with the above mentioned statistical data are currently under extensive study but already have been shown to carry medically interesting information (Żebrowski et al., 1995b).

In this paper, in order to obtain more stable results when different complexity measures were calculated, a relatively long time window of 512 beats (6-9 min. of real time) was used throughout. To make the calculations somewhat faster the window was slid every 200 beats and the 24-hour histograms made from the resultant entropy values.

For the fixed-Q approach the number of bins  $Q=16$  was used, while for the fixed-W the bin width  $W=25$  was applied. Wherever applicable the time lag was 2 beats. Note that for a 512 window the time lag of 2 beats is almost negligible. However, this may not necessarily be the case when a short window of 50 or even 100 beats is used, especially when the heart rate variability is extremely low as in deep pathology or using animal data (Hoyer et al., this volume).

The coarse-grained entropy rates CER's were computed from  $n$ -dimensional probability distributions (histograms),  $n = 2, \dots$ . Therefore the problem of the number of data samples and the number  $Q$  of marginal bins is more critical than in estimations of the static 1-dimensional entropies. As a preliminary result we present here the CER's obtained using the window length equal to 4096 beats (an equivalent of 35 min. to 70 min. of real time) and  $Q = 4$ ,  $n = 2$ . The CER  $h^{(0)}$  was computed using  $\tau_0 = 0$   $\tau_1 = 1$ , for the CER  $h^{(1)}$  the lags  $\tau_0 = 0$  and  $\tau_{max} = 35$  beats were applied. The window step was also 200 beats and the obtained values were processed into histograms of the CER's values in the same way as the static entropies.

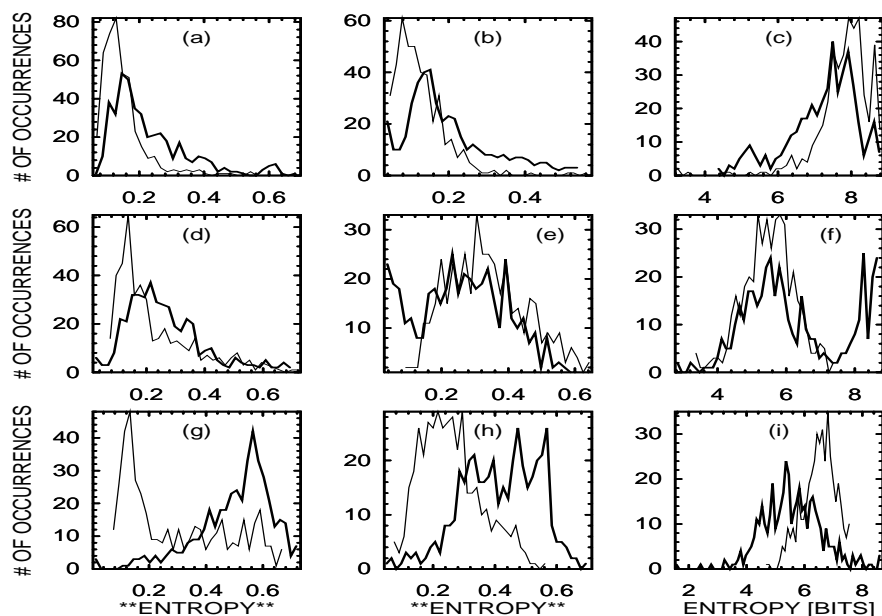


**Fig. 4.** Histograms of the fixed- $Q$  static entropies: the Shannon entropy  $H^3(2)$  (a,d,g), the Shannon entropy  $H^1$  (b,e,h) and the Rényi entropy  $H_3$  (c,f,i); for the pairs dp1 (a,b,c), dp2 (d,e,f) and dp3 (g,h,i). The histograms of the “a”-files and the “b”-files are plotted using the thick and thin lines, respectively.

## 6 Results

The results for the Shannon and Rényi entropies using fixed- $Q$  partitions are presented in Fig. 4, the first row displays the histograms for the pair dp1, the second for dp2, the third for dp3. The histograms for a-files are plotted using thick lines, thin lines are used for b-files. The first column in Fig. 4 contains histograms of 3-dimensional Shannon entropies  $H^3(2)$ , the second column contains histograms of 1-dimensional Shannon entropies  $H^1$ , while the histograms of the third-order Rényi entropies  $H_3$  are displayed in the third column. Also the second-order Rényi entropies  $H_2$  were computed, but their histograms were the same as related histograms of  $H_3$ . The results of all entropies in Fig. 4 are consistent – histograms of a-files (thick lines) are located on the left-sides of the histograms of the b-files. Thus the Shannon and Rényi entropies estimated using the fixed- $Q$  approach sort all the a-files into one group and all the b-files into the other group. The same sorting is obtained from the 3rd order pattern entropy  $Z_3(2)$ , if it was also estimated using fixed- $Q$  partitions (first column in Fig. 5).

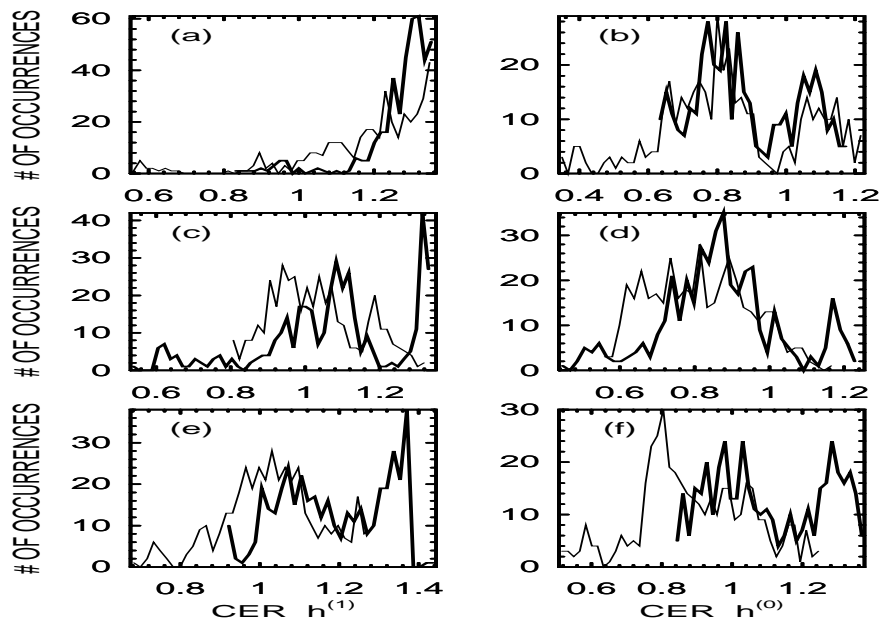
When the 3-rd order pattern entropy  $Z_3(2)$  is estimated using the fixed- $W$



**Fig. 5.** Histograms of the static entropies: the fixed-Q pattern entropy  $Z_3(2)$  (a,d,g), the fixed-W pattern entropy  $Z_3(2)$  (b,e,h), and the fixed-W Shannon entropy  $H^3(2)$  (c,f,i); for the pairs dp1 (a,b,c), dp2 (d,e,f) and dp3 (g,h,i). The histograms of the “a”-files and the “b”-files are plotted using the thick and thin lines, respectively.

partition (the second column in Fig. 5), the results for dp1 and dp3 are more or less consistent with the fixed-Q approach, however, this is not the case for dp2. It seems that now the sorting is a-b-a and b-a-b, unlike the a-a-a and b-b-b sorting found above. This result is not typical for the pattern entropy, but can be obtained from other static entropies, when fixed-W partitions are used. An example for the fixed-W Shannon entropy  $H^3(2)$  is presented in the third column of Fig. 5.

Considering the original health-pathology (patient-control) coding, does this mean, that the fixed-W partition is appropriate for detection of pathology in this case? We will see, that this example is not appropriate to decide this problem. First, however, we report results from dynamical entropies - coarse grained entropy rates CER. CER  $h^{(1)}$  – the first column in Fig. 6, CER  $h^{(0)}$  – the second column in Fig. 6. At the first sight it seems, that the sorting given by CER’s is even worse than that given by the static entropies, and giving a fixed-Q-like (i.e. incorrect) sorting. There is one feature, however, in the CER’s histograms, which in this case provide the correct a-b-a/b-a-b sorting. It is the existence of small values (0.6 – 0.8 in CER  $h^{(1)}$ , and 0.4 –

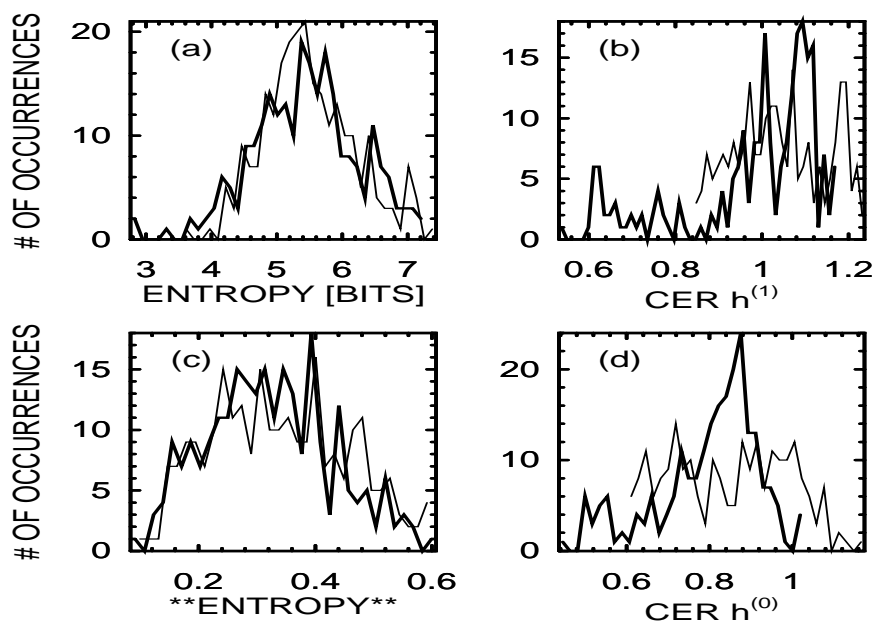


**Fig. 6.** Histograms of the dynamic entropies – the coarse-grained entropy rates CER  $h^{(1)}$  (a,c,e) and CER  $h^{(0)}$  (b,d,f); for the pairs dp1 (a,b), dp2 (c,d) and dp3 (e,f). The histograms of the “a”-files and the “b”-files are plotted using the thick and thin lines, respectively.

0.6 in CER  $h^{(0)}$ ), which occurs only in the case of the controls.

Let us concentrate on the dp2 pair and results which discriminate the a-file here correctly as the control. In the case of the fixed-W static entropies, the results can be influenced just by different variance of the data. And this is the case of dp2a (Fig. 2a), where the last third of the data has much larger variance than the rest of the record. Let us rerun the computations for the dp2 pair, but now only for the first 50,000 samples, for the Shannon and pattern entropies (fixed-W) and for the CER’s. The results are presented in Fig. 7. The static entropies (a – Shannon, c – pattern) lost the part that distinguished a from b, while the low-CER part of the histogram is still present in the histograms of the CERs.

From the point of view of characterizing “complexity of dynamics” it seems that the dp2a–dp2b distinction by the static entropies was obtained due to an artifact – the larger variance seen in the last third of Fig. 2a. From the point of view of practical medical diagnostics, however, the variance itself can be an important factor. The fixed-W static entropies are sensitive both to a larger variance and to a possibly accompanying change in the level of



**Fig. 7.** Histograms of entropies and entropy rates for the pair *dp2*, computed from the first 50,000 samples (i.e., excluding the high variance part of *dp1a*, see Fig. 2), the fixed-*W* Shannon entropy  $H^3(2)$  (a), the fixed-*W* pattern entropy  $Z_3(2)$  (c), CER  $h^{(1)}$  (b) and CER  $h^{(0)}$  (d). The histograms of the “a”-files and the “b”-files are plotted using the thick and thin lines, respectively.

ordering in the signal. The first 50,000 points in the series represent heart rate variability during the day. It is well known that there are people whose heart rate variability during the day may be very different from that during the night. Often dangerous symptoms occur only during the night and not during the activity periods (or vice versa). In the case of the patient *dp2b*, the relatively uniform variance throughout the 24 hours was a danger signal for his doctors. On the other hand, the large increase of variance in the control *dp2a* in the night hours was considered normal. In general, however, the value of the variance of the RR intervals itself is not a sufficient indicator of a risk of impending cardiac arrest. The fixed-*W* pattern entropy successfully classified the target member *dp2b* to be in the high risk class – he died due to CA several months afterwards. The question of discriminating power of dynamics versus variance, however, still needs further investigation.

The interesting finding here is that the difference between the members of the pair 2, as detected by the CER’s, persists in the data in general, not

only in the visibly high variance part of the data. Thus it is possible, that the distinction between health and pathology may be better detectable in the complexity of dynamics as measured by (coarse-grained) entropy rates. In other words, there are features of the RR series detected by the CER's, which cannot be found in the static complexity of the marginal distributions. This result, however, is very preliminary, and a much larger database of RR recordings need to be processed to confirm/refuse validity of this finding.

## 7 Acknowledgement

Milan Paluš was supported by the Academy of Sciences of the Czech Republic (grant No. 230404), and Jan Żebrowski was supported by Polish Research Council KBN Grant No. 2 P30B 20 09. The authors thank Dr. Wanda Popławska for very important discussions and for selecting and providing the RR-interval data.

## References

- Baranowski R., Żebrowski J.J., Popławska W., Rywik S. (1994): Nonlinear dynamical methods in heart rate variability analysis from short Holter recordings. *J. Ambulatory Monitoring* **7** 81
- Cover T.M., Thomas J.A. (1991): *Elements of Information Theory*. (J. Wiley & Sons, New York)
- Fraser A.M. (1989): Information and entropy in strange attractors. *IEEE Transactions on Information Theory* **35** 245-262
- Hoyer D. et al., this volume.
- Paluš M. (1993): Identifying and quantifying chaos by using information-theoretic functionals. In: Weigend A.S., Gershenfeld N.A. (eds.): *Time Series Prediction: Forecasting the Future and Understanding the Past*. Santa Fe Institute Studies in the Sciences of Complexity, Proc. Vol. XV (Addison-Wesley, Reading, Mass.) 387-413
- Paluš M. (1996b): Coarse-grained entropy rates for characterization of complex time series, *Physica D*, in press.  
Electronic preprint: <ftp://ftp.santafe.edu/pub/mp/eq/eq3pr.ps>.
- Paluš M.: Probing nonlinear relations in multivariate time series, this volume.
- Petersen K. (1983): *Ergodic Theory* (Cambridge University Press, Cambridge)
- Popławska W., Grzybowski J., Żebrowski J.J., Baranowski R., Bilinska Z.T., Pszona B. (1994): Pattern Entropy - nonlinear noninvasive estimation of adrenergic drive in patients with dilated cardiomyopathy. *J.Amb. Monitoring* **7** 80
- Sinai Ya.G. (1976): *Introduction to Ergodic Theory* (Princeton University Press, Princeton)
- Żebrowski J.J., Popławska W., Baranowski R. (1994): Entropy, pattern entropy and related methods for the analysis of data on the time intervals between heart beats from 24 hour electrocardiograms. *Phys.Rev.E* **50** 4187-4205



- Żebrowski J.J., Popławska W., Baranowski R. (1995a): Pattern Entropy a Tool for Nonlinear Dynamical Analysis of a Biological Nonstationary System: The Human Heart. *Acta Phys. Pol.* **26** 1055-1064
- Żebrowski J.J., Popławska W., Baranowski R., Buchner T. (1995b): Local Time Properties of 24- hour Holter recordings of RR-intervals: a nonlinear quantitative method through pattern entropy of 3-D return maps. *Computers in Cardiology 1995* 177-180
- Żebrowski J.J., Popławska W., Baranowski R., Buchner T. (1995c): Phase space analysis of heart rate variability, to appear in *Acta Phys. Pol. A* (Proceedings of the International Conference on Nonlinear, Chaotic and Complex Systems, Zakopane 1995)