Block Entropy Analysis of Heart Rate Variability Signals

K. Karamanos\textsuperscript{a}, S. Nikolopoulos\textsuperscript{b}, G. Manis\textsuperscript{c}, A. Alexandridi\textsuperscript{d}  
K. Hizanidis\textsuperscript{b} and S. Nikolakeas\textsuperscript{e}

\textsuperscript{a} Centre for Nonlinear Phenomena and Complex Systems  
Universite Libre de Bruxelles  
CP 213, Campus Plaine  
B-1050 Brussels  
Belgium

\textsuperscript{b} National Technical University of Athens  
Dept. of Electrical and Computer Engineering  
Iroon Polytechniou 9  
GR-15773, Zografou Campus  
Greece

\textsuperscript{c} University Of Ioannina  
Dept. of Computer Science  
P.O. Box. 1186, GR-45110, Ioannina  
Greece

\textsuperscript{d} Foundation of Biomedical Research of the Academy of Athens  
4 Soronou Efesiou Str 11527 Athens,  
Greece

\textsuperscript{e} Athens Air Force and Veterans Hospital  
Dept. of Cardiology  
P. Kanellopoulos St, Athens,  
Greece

Corresponding Author: S. Nikolopoulos  
kyriakos@central.ntua.gr
Abstract

In this paper we present a novel approach to the analysis of Heat Rate Variability (HRV) data, that of coarse-graining analysis using the estimation of Block Entropies with the technique of lumping. HRV time series are generated from long recordings of Electrocardiograms (ECGs) and are then filtered in order to produce a coarse-grained symbolic dynamics. Block Entropy analysis is applied to these dynamics in order to examine its coarse-grained statistics. Our data set is comprised of two subsets, one of healthy subjects and another of Coronary Artery Disease (CAD) patients. It is found that Entropy analysis provides a quick and efficient tool for the differentiation of these series according to subject category. Healthy subjects provided more complex statistics compared to patients; specifically, the healthy data files provided higher values of block Entropies compared to patient ones. We also compare these results with the Correlation Dimension Estimation in order to establish coherency. We believe that this analysis may provide a useful statistical method towards the better understanding of the human cardiac system.

Keywords: Symbolic dynamics, Block Entropies, lumping, Electrocardiograms, Heart Rate Variability

Introduction

The analysis of electrocardiograms (ECG) provides the most common view to cardiac function. It is a widespread technique, as it constitutes the simplest non-invasive and reliable method for cardiac pathology diagnosis and investigation. Since the development of the digital electrocardiograph and its extended use, the acquisition of long ECG recordings has become popular in HRV analysis. HRV refers to the beat-to-beat alterations in heart rate, which is found through the measurement of the variability of successive R-wave peaks in the ECG. The methods that have been applied are usually separated into time domain methods, frequency domain methods, rhythm pattern analysis methods and the more recently popular nonlinear methods. Although HRV has been the subject of numerous clinical studies investigating a wide
spectrum of biosignals and their respective diseases and clinical conditions, the practical use of HRV in medicine has been widely accepted only in two clinical settings: Depressed HRV can be used as a predictor of risk after acute myocardial infarction and as an early warning sign of diabetic neuropathy.

It has been established that nonlinear phenomena are key factors in the genesis of HRV. They are determined by complex interactions of haemodynamic and electrophysiological variables, as well as by autonomic and central nervous system regulation. It has been speculated that HRV analysis that based on the methods of nonlinear dynamics might elicit valuable information for the physiological interpretation of HRV and for the assessment of the risk of sudden death [Goldberger et al., 1987], [Babloyantz et al., 1988]. The parameters which have been used to measure nonlinear properties of HRV include 1/f scaling of Fourier spectra, the H scaling exponent and coarse-graining spectral analysis [Malik et al. 1996]. Furthermore, Poincare sections, low dimensional attractor plots, singular value decompositions, the D2 Correlation Dimension estimation, Lyapunov exponents and the Kolmogorov Entropy have been used among other in order to extract a deterministic complex dynamical process which could represents the cardiac system’s dynamics. Although in principle these techniques have been shown to be powerful tools for characterization of various complex systems, no major breakthrough has yet been achieved by their application to biomedical data including HRV time series analysis [Malik et al. 1996].

In this paper, we perform an Entropy analysis of the coarse-grained dynamics of these two categories of subjects: Healthy subjects and Coronary Artery Disease (CAD) patients. As we will show in detail, the computation of Block Entropies by lumping allows one to distinguish between healthy subjects and CAD patients. Therefore, Entropy analysis provides a useful and quick tool for the analysis of long-recording ECGs, yielding results coherent to other nonlinear dynamics methods such as the Correlation Dimension estimation [Nicolis J.S., 1991], [Nicolis G., 1995].

In order to apply this Entropy analysis we generate a new coarse-grained time series, taking values on a finite alphabet, with a modified method of level crossing [Nicolis J.S., 1991], [Nicolis G., 1995]. We examine the scaling of the Shannon-like Block Entropy H, as a function of the length of the words, applying a novelty called lumping [Karamanos, 2001]. The examination of the Entropy statistics allows one to draw some general conclusions.
We subsequently calculate the Correlation Dimension so that we may combine the results of the nonlinear methods and the Block Entropy analysis. These seem to be qualitatively coherent even though the nonlinear methods are computationally more consuming and more sensitive than the Block Entropy analysis by lumping.

**Data Collection and Stationarity**

The healthy subject data set used in this study is made up of continuous ECG recordings (5000-6000 R-R intervals) derived from normal young males 25-29 years old, with a clean medical history and normal physical examination. All healthy subjects were non-smokers, received no drugs and abstained from caffeine for 24 hours prior to acquisition. All recordings were performed in a quiet room, between the hours of 15.00-17.00 (3:00-5:00pm), in the supine position, with a steady respiratory rate (12/min) under continuous monitoring by a cardiologist who confirmed the absence of any cardiac rhythm disturbance throughout the recording. Continuous ECG recordings (an average of 5000-6000 R-R intervals) were also acquired from a group of hospitalized CAD patients, under similar conditions. All unhealthy subjects had 1 or 2 vessel coronary disease, angiographically confirmed, without infarction or arrhythmia and they were not taking any β-blockers. None of them had been subjected to coronary angioplasty or coronary by-pass grafting and showed no significant signs of myocardial ischemia during the rest ECGs.

A time series was constructed from continuous R-R samples, which were extracted from the long-term ECG recordings. Long ECG recordings were preferred for more accurate analysis and since large amounts of data ensure higher precision of the results; especially in nonlinear methods, the amount of data is a crucial issue. An alternative solution would have been to concatenate short recordings, but this would forcibly result in non-stationary data. Stationarity is an essential property of a time series since it ensures that all statistical quantities of a process are independent of absolute time.

We should note that we did not follow Holter recordings even though they are large enough, because they are obviously non stationary; the subject may be moving, sleeping, eating, running, etc, thus varying the external conditions significantly. In addition Holter recordings introduce a lot of noise artefacts in the signal, especially
due to the muscular movements of the subject. In contrast, long recordings with the electrocardiograph are more stationary since the subject is restricted to the hospital bed during the entire acquisition.

Apart from all the above, the recordings were examined for stationarity using theoretical methods. The most common method for investigating the stationarity of a process is by examining the mean and the Autocorrelation function for the entire time series and subparts. Then stationary subparts were those that were used for analysis.

Another data set we used for the comparison with our own is the MIT fantasia database [Iyengar et al., 1996]. This database contains long ECG recordings of young healthy subjects and recordings of elderly healthy subjects.

**Generation of the coarse-grained signal**

The coarse-grained time series were constructed using a modified level crossing method. According to this method, certain thresholds are applied and if a data value is between these thresholds (greater or less than), then this value is replaced with a specific linguistic value or a letter. Our modification is quite trivial and consisted of the consideration of a non equivalent discretization of the range of the time series, based in the distribution of the values. Under this discretization the time series is modified to another one, which contains the “linguistic” dynamics characteristics.

An example of the above is illustrated in Fig. 1. The HRV time series of a healthy subject is coarse-grained for a 10 letter alphabet. Fig. 1 depicts the original time series, the coarse-graining of this time series and the corresponding letters substituting the values of the coarse-graining time series. By definition, the coarse-graining transformation of the original time series eliminates the high frequency section of the spectrum as it smoothens the time series. Moreover, the values of the mean and the standard deviation will be changed in proportion to the linguistic values. However, by plotting the Autocorrelation functions of the original and the coarse-grained time series we observed that the Decorrelation time remained unaffected in all cases, as shown in Fig. 2.
Entropy analysis by lumping

Entropy-like quantities are a very useful tool for the analysis of arbitrary symbolic sequences. Of special interest are the block entropies, extending Shannon’s classical definition of the Entropy of a single state to the Entropy of a succession of states [Nicolis and Gaspard, 1994]. In particular, it has been shown that the scaling of the block entropies with length sometimes gives interesting information on the structure of a sequence [Ebeling and Nicolis, 1991, 1992]. In light of these results, we have come up with certain criteria that could illuminate the structure of a symbolic sequence and give us some more specific information beyond that provided by block entropies computed by gliding. In fact, we have included automaticity, a specific yet important algorithmic property of the sequence, as entropy criterion.

Lumping is the reading of the symbolic sequence by ‘taking portions’, as opposed to gliding, where one has essentially a ‘moving frame’. We note that gliding is the standard convention in the literature. Reading the symbolic sequence in a specific way is also called decimation of the sequence.

In [Karamanos and Nicolis 1999], we examined how the estimation of the block entropies may actually depend on the method of reading, meaning the observer. This has immediate consequences on the “decoding” procedure, as different values of the block entropies mean different kinds and amounts of information extracted by the symbolic sequence. By using lumping, we have established a new decimation scheme for the symbolic dynamics of the Feigenbaum attractors of unimodal maps. The coarse-grained statistical properties of the attractors have been subsequently derived, with emphasis on the behavior of the block entropies. Moreover, in [Berthe, 1995] it has been shown that the estimation of the (conditional) block entropies with the usual prescription of gliding cannot help us distinguish between sequences with different spectral properties and different levels of complexity.

Let us consider a subsequence of length N selected out of a very long (theoretically infinite) symbolic sequence. We stipulate that this subsequence is to be read in terms of distinct ‘blocks’ of length \( n \),

\[
A_1, \ldots, A_n / A_{n+1}, \ldots, A_{2n} / \ldots, / A_{(j+1)n}, A_{(j+1)n}.
\] (1)
As mentioned, we call this reading procedure lumping. We shall apply this method in all the following description.

The following quantities characterize the information content of the sequence [Ebeling and Nicolis 1991] [Khinchin, 1957]

(i) The dynamical (Shannon-like) Block Entropy for blocks of length $n$

$$H(n) = - \sum_{(A_1, \ldots, A_n)} p^{(n)}(A_1, \ldots, A_n) \cdot \ln p^{(n)}(A_1, \ldots, A_n), \quad (2)$$

where the probability of occurrence of a block, denoted by $p^{(n)}$, is defined (when it exists) in the statistical limit as:

$$p^{(n)}(A_1, \ldots, A_n) = \frac{\text{No of blocks of the form } A_1, \ldots, A_n \text{ encountered when lumping}}{\text{total No of blocks encountered when lumping}} \quad (3)$$

starting from the beginning of the sequence and the associate Entropy per letter

$$h^{(n)} = \frac{H(n)}{n}. \quad (4)$$

(ii) The conditional Entropy or Entropy excess associated with the addition of a symbol to the right of an $n$-block

$$h_{(n)} = H(n+1) - H(n). \quad (5)$$

(iii) The Entropy of the source (a topological invariant), defined as the limit (if it exists)

$$h = \lim_{n \to \infty} h^{(n)} = \lim_{n \to \infty} h_{(n)}, \quad (6)$$

which is the discrete analogue of the metric or Kolmogorov Entropy.

The next phase is the selection problem, that is, the possibility of the emergence of some preferred configurations (blocks) out of the complete set of different possibilities. The number of all possible symbolic sequences of length $n$ (complexions in the sense of Boltzmann) in a $K$-letter alphabet is:

$$N_k = K^n. \quad (7)$$

Yet not all of these configurations are necessarily realized by the dynamics, nor are they equiprobable. A remarkable theorem by McMillan [Khinchin, 1957] gives a partial answer to the selection problem asserting that for stationary and ergodic sources the probability of occurrence of a block is:

$$p^{(n)}(A_1, \ldots, A_n) \approx e^{-H(n)}. \quad (8)$$

At this point, one can state an important theorem, first proved in [Karamanos, 2001], connecting the world of the machines with Block Entropy:
Theorem 1: If the symbolic sequence \((u_n)_{n\in\mathbb{N}}\) is \(m\)-automatic, then

\[ H(m^k) = H(m), \forall k \geq 1, \tag{9} \]

when lumping starts from the beginning of the sequence.

As we have already mentioned, the Fourier spectrum or the standard convention of the Entropy analysis by gliding, do not help us to distinguish between symbolic sequences with completely different levels of complexity and spectra. Unlike the previous methods, the novelty of the Entropy analysis by lumping gives results that can be connected with algorithmic aspects of the sequences, in particular with the property of the sequence to be generated by deterministic or stochastic automata [Karamanos, 2001].

It is also important to note that Entropy analysis by lumping of weakly chaotic systems gives a rather characteristic Entropy spectrum, as explained in [Karamanos, 2003]. This shows that the Entropy analysis by lumping is much more sensitive in algorithmic and ergodic properties of (weakly) chaotic systems than the classical conventional Entropy analysis by gliding. Therefore, this is what we use here.

Thus this method of analysis is attractive for use in coarse-grained HRV signals. Indeed, these signals exhibit complex dynamics when analyzed with classical nonlinear methods and the associated block Entropies analysis completes the detailed of the cardiac system dynamics [Karamanos et al. 2004].

**Correlation Dimension Estimation**

The Correlation Dimension estimation is essential in nonlinear time series analysis. It provides both a qualitative and quantitative measure of the complexity of a time series, was first established by Grassberger & Procaccia [Grassberger and Procacca 1983] and is based on the Takens Theorem [Takens 1980].

According to the Grassberger and Procaccia algorithm, the time series \(X(1), X(2), X(3), \ldots, X(N)\), is a measure of a single coordinate of an \(m\)-dimensional system’s underlying dynamics. Assuming \(m\) is the embedding dimension (the
dimension of space in which the assumed system’s trajectory is unfolded) and \( \tau \) is the time lag, then phase space reconstruction (described below) is performed with time delays and the following \( m \)-dimensional vectors are constructed:

\[
\tilde{x}(t) = [X(t), X(t + \tau), X(t + 2\tau), ..., X(t + (m-1)\tau)].
\]  

(13)

which represent the trajectory from \( \tilde{x}(0) \) up to and including \( \tilde{x}(t) \) within the reconstructed phase space. If the reconstruction is accurate, then \( A' \) is the topological conjugate of the original attractor, \( A \). The selection of the window length \( wl = (m-1)\tau \) is critical, thus it has to be done with the proper way [Kugiumtzis 1996].

After the phase space construction the correlation integral is:

\[
C(m,r,\tau) = \frac{2}{N-1} \sum_{i=1}^{N} \sum_{j=1}^{N} \Theta \left[r - \|\tilde{x}_i - \tilde{x}_j\|\right],
\]  

(14)

where \( N \) is the time series length, and \( r \) a distance magnitude which represents a radius of a spherical area in the \( m \)-dimensional phase space. \( \Theta \) is the Heavyside function:

\[
\Theta(i) = \begin{cases} 
1, & \text{if } i \geq 0 \\
0, & \text{if } i < 0 
\end{cases}
\]  

(15)

We could interpret (14) as: find all the pairs \( \tilde{x}_i, \tilde{x}_j \) in \( m \)-dimensional state space whose distance \( \|\tilde{x}_i - \tilde{x}_j\| \) is less than \( r \).

The Euclidean norm used in the above equation states that the difference between \( \tilde{x}_i \) and \( \tilde{x}_j \) is the maximum difference among their coordinates:

\[
\|\tilde{x}_i - \tilde{x}_j\| = \left| \left[ \left| X(i) - X(j) \right| + \left| X(i + \tau) - X(j + \tau) \right| + \ldots \right|^{\frac{1}{2}} \right|
\]  

(16)

Once we have calculated the \( C(m,r,\tau) \) for a range of \( r \) and for various \( m \), we construct the plot \( \text{Log}(C(m,r,\tau)) \) versus \( \text{log}(r) \) for each \( m \). From each \( \text{Log}(C(m,r,\tau)) \) versus \( \text{log}(r) \) plot, we select a scaling region for which we calculate the slope. If, by increasing the dimension of the phase space reconstruction, the slope \( m \) converges to a specific value, then we set this value as the Correlation Dimension:

\[
D_2 = \lim_{N \to \infty} \frac{\text{Log}(C(m,r,\tau))}{\text{Log}(r)}
\]  

(17)
Results

*Entropy Analysis by lumping*: The block entropies per letter by lumping are depicted in Fig. 3 for a sample healthy subject. Similar results hold for all data sets. From the diagram it follows that we have good statistics for the determination of the Block Entropy up to $n=10$ approximately. The determination of the region of good statistics is one additional advantage of the method of the Block Entropy by lumping.

The normalized associated block Entropies per letter by lumping are depicted in Fig. 4 (where we have good statistics) for a bipartition. The partitioning is done through the level crossing method, where the median of the time series has been selected as the crossing level. If the signal bypasses this level, we call the coarse-grained signal as U (for upper) and if not, we call it L (for lower). With this technique the initial time series is coarse-grained to a new discretized linguistic time series and one can use standard methods of mathematical linguistics to analyze the final signal.

As a first step we analyze the HRV signals of 5 healthy subjects and 5 CAD patients by the method of block Entropies, as discussed in previous sections. The resulting entropies are shown in Fig. 4. It seems that there is a continuous transition between healthy subjects and coronary patients.

From the Fig. 4 it is evident that the healthy subjects present higher values of the block entropies compared to the patients. In particular, healthy subjects present a normalized associated Entropy per letter in the range from 70% for small word length to 35% for large word length. Furthermore the coronary patients present a corresponding $h(n)$ from 50% for small word length to 18% for large word length. The curves are descending in a monotonic manner.

In Fig. 5 the monotonic descent is more pronounced and steeper. However, the distinction between healthy subjects and CAD patients is clearer. Indeed there is an obvious empty zone in the transition region. This could be supported by a theoretical argument, which is that the healthy subjects exhibit an increased complexity compared to the patients. This complexity is enforced by the more detailed coarse-grained description by many letters.

Another important issue is the distinction between healthy and young subjects and healthy elderly subjects. The data used for this analysis were also taken from the
Fantasia database. The corresponding normalized associated Block Entropy per letter, for a 10 letter alphabet, is depicted in Fig. 6. A monotonic decay is present here as well. The values for the healthy and young subjects are analogous to those of our data. As was the case in Fig. 5, these healthy and young subjects present a normalized associated Entropy per letter in the range from 70% for small word length to 35%. The healthy elderly subjects present a corresponding $h(n)$ from 55% for small word length to 25% for large word length. There are two zones, one for younger and one for older subjects and, once again, there is a continuous transition. Based on these encouraging observations, it is apparent that his method could be applied in the clinical setting.

An important point that should be made here is the qualitative explanation of the Entropy values. The values of the Entropy for coarse-grained HRV time series coming from healthy subjects seem to be increased compared to those of the patients. This can be qualitatively been explained by the fact that:

1) The time series of healthy subjects possess high dimensionality dynamics and thus increased block entropies in all word lengths.

2) The increased complexity of the signals is unambiguously coherent to the statistics of the signals, thus the Entropy values of complex signals are increased compared to non-complex ones. Apparently pure noise would theoretically present the highest Entropy for all word lengths.

The power of the method of the block entropies by lumping is manifested in the fact that the method works sufficiently even after only 2-symbols linguistics.

**Correlation Dimension Estimation:** Through the use of the Autocorrelation and mutual information functions, a value was determined for the time lag, $\tau$, that is most suitable for this study and that was $\tau=5$. The dimensions chosen for the phase space reconstruction started at $m=3$ and went to $m=20$, based on the fact that after several trials these values yielded the best reconstruction and thus led to more accurate results and subject discrimination. The correlation integral was then calculated for an extended range of $r$ (up to $10^8$, experimentally determined). This correlation integral was used to estimate the Correlation Dimension values.
The distinction between the two groups of subjects is clear, as shown in Fig. 7. The signals from the healthy subjects have higher Correlation Dimensions depicting a more complicated dynamic system with more degrees of freedom than those from the CAD patients. This is expected, according to our initial hypothesis. The dynamic system of the CAD patients tends to behave more normally and linearly, represented by models with 4 to 5 degrees of freedom for non-severe cases or 3 to 4 for worse ones. In contrast, the healthy subject signals depict a dynamic system with strong nonlinearity, which requires 6 to 7 degrees of freedom for describing cardiac behaviour. The mean values for the estimated dimensions are: The mean Correlation Dimension, $D_2 = 8.4261$ for healthy subjects with standard deviation of $0.480169$ and for the patients, $D_2=4.2533$ and STD= 0.827082. The respective values for the Fantasia data are $D_2=8.2827$ for the young subjects and $D_2=4.9872$ for the elderly subjects.

Generally, the Delay Times method is very sensitive and arrives at a close value of a realistic Correlation Dimension. In this application, certain assumptions were made in order to estimate the $D_2$. The first assumption is that we used the Euclidean norm Eq. (16) instead of the commonly used infinity norm. The second assumption is that we did not take consideration of the Theiler window [Theiler 1986]. These two assumptions improved the method’s effectiveness and provided some results that yield a good and clear discrimination as can be seen in Fig. 7.

Discussion

The implementation of Block Entropy analysis of HRV signals attempts to introduce an innovative statistical analysis ECG time series. The coarse-graining of the original time series with the method of level crossing generates a second series that projects the cardiac dynamics in a linguistic context. Linguistic dynamics have been present in cybernetics for more than two decades. The scope of the specific analysis of HRV signals offered here is both practical and theoretical.

The theoretical scope is examining the system’s information content. The coarse-graining of the time series provides a projection of the system’s information into a more normative sequence. This technique allows for Entropy analysis to be performed under a linguistic concept.
Also, one of the goals of our investigation was to examine whether this analysis provides an effective tool for distinguishing healthy subjects and CAD patients, which was proven successfully. This could be developed into a useful tool for clinical practice. Indeed, a long recording ECG from a subject in conjunction with other methods could provide an efficient way of characterizing cardiac health.

The method’s results are coherent to the nonlinear dynamics analysis of similar signals with other methods. The correlation of the analysis of HRV with nonlinear dynamical methods manifest that youth and health are related to high dimensional dynamics or strong complexity, while age and disease are related to weak complexity and low dimensionality in the dynamics.

However, some authors claim that these methods are sensitive and require the careful observation of specialists, as they contain a lot of “weak” parameters that must be accounted for. They have also proven more useful for the classification of ECG (and especially HRV) data rather than for providing absolute and precise values of nonlinear parameters. This is especially evident from the different Correlation Dimension values found in the literature for the characterization of the same categories of subjects [Bogaert et al. 2001]. However, we support that this does not indicate that nonlinear methods are unfit for such analysis, but rather that they provide measures that describe the complexity of such dynamical time series. Furthermore, the method presented in this manuscript enforces the results of nonlinear methods, since it presents coherent conclusions. The two categories of subjects are clearly differentiated quickly and without the need for further processing.

An important point that we wish to pose once more is the length of the time series. For many nonlinear methods an adequate time series length is $N = 10^{2.04D_2}$, which demands $N=1000000$ points for a $D_2 \approx 10$ estimation. This corresponds to an ECG recording of about 16000 hours, which is, of course, non-realistic for experimental data acquisition. Thus some assumptions have been made so that shorter data lengths may be suitable for analysis. For a sufficient Entropy analysis by lumping, a 6000 points time series can be statistically analysed for a maximum word length of eight ($n \leq 8$). However, 6000 points are too few for accurate Correlation Dimension estimation. Indeed, the large amount of data necessary is a drawback of nonlinear methods, when the goal is to describe the dynamics of the system under question. Therefore, we used this method only for the differentiation between healthy subjects and CAD patients.
and were not as concerned with obtaining a qualitative characteristic. We selected an average length of 7000 points for both methods, which corresponds to about 115 min recording at about time.
References


FIGURE 1

K. Karamanos\textsuperscript{a}, S. Nikolopoulos\textsuperscript{b}, G. Manis\textsuperscript{c}, A. Alexandridi\textsuperscript{d}, K. Hizanidis\textsuperscript{b} and S. Nikolakeas\textsuperscript{e}
FIGURE 2

K. Karamanos\textsuperscript{a}, S. Nikolopoulos\textsuperscript{b}, G. Manis\textsuperscript{c}, A. Alexandridi\textsuperscript{d}
K. Hizanidis\textsuperscript{b} and S. Nikolakeas\textsuperscript{e}
FIGURE 3

K. Karamanos\textsuperscript{a}, S. Nikolopoulos\textsuperscript{b}, G. Manis\textsuperscript{c}, A. Alexandridi\textsuperscript{d}
K. Hizanidis\textsuperscript{b} and S. Nikolakeas\textsuperscript{e}
FIGURE 4

K. Karamanos\textsuperscript{a}, S. Nikolopoulos\textsuperscript{b}, G. Manis\textsuperscript{c}, A. Alexandridi\textsuperscript{d}
K. Hizanidis\textsuperscript{b} and S. Nikolakeas\textsuperscript{e}
FIGURE 5

K. Karamanos\textsuperscript{a}, S. Nikolopoulos\textsuperscript{b}, G. Manis\textsuperscript{c}, A. Alexandridi\textsuperscript{d}
K. Hizanidis\textsuperscript{b} and S. Nikolakeas\textsuperscript{e}
FIGURE 6

K. Karamanos\textsuperscript{a}, S. Nikolopoulos\textsuperscript{b}, G. Manis\textsuperscript{c}, A. Alexandridi\textsuperscript{d}
K. Hizanidis\textsuperscript{b} and S. Nikolakeas\textsuperscript{e}
FIGURE 7

K. Karamanosa, S. Nikolopoulosb, G. Manis, A. Alexandridid
K. Hizanidisb and S. Nikolakease
FIGURE 1: Part “A” depicts an RR HRV time series of a healthy subject. Part “B” is the coarse-graining of the time series for a 10 letters alphabet. Part “C” represents the coarse-graining values and the corresponding letters of a time series’ snapshot. The first 10 values are: D, F, E, J, I, I, G, F, E, D.

FIGURE 2: The Autocorrelation function of pure and coarse-grained time series. The initial data is depicted with a solid line, the dashed line depicts the 2-letter symbolic coarse-grained time series and the dotted line represents the 10-letter data.

FIGURE 3: Shannon-like Entropy statistics for a sample healthy subject, as a function of word length. Circles depict the coarse-grained time series for 10-letter alphabet and pluses represent the corresponding results for a 2-letter alphabet.

FIGURE 4: Normalized Block Entropy per letter for healthy subjects and CAD patients. The linguistic time series produced from the original one was made of two letters.

FIGURE 5: Normalized Block Entropy per letter for healthy subjects and coronary patients. The linguistic time series produced from the original one was made of ten letters.

FIGURE 6: Normalized Block Entropy per letter for healthy young subjects and healthy elderly subjects from the Fantasia database. The linguistic time series produced from the original was made of ten letters.

FIGURE 7: Delay Times method for estimating the Correlation Dimension. Solid lines represent Group A subjects, dotted lines represent Group B subjects. The separation between the two groups of subjects is clear.