Block Entropy analysis of long recorded Electrocardiograms as a good way for discrimination between Normal subjects and Coronary patients

K. Karamanos* , S. Nikolopoulos** and K. Hizanidis**

*Centre for Nonlinear Phenomena and Complex Systems
Universite Libre de Bruxelles
CP 213, Campus Plaine
B-1050 Brussels
Belgium

**National Technical University of Athens
Dept. of Electrical and Computer Engineering
Iroon Polytechniou 19
GR-15773, Zografou Campus,

Abstract
Heart Rate Variability timeseries of long recordings of Electrocardiograms are filtered in order to generate a coarse grained linguistic dynamics. Then the Block Entropy analysis is applied in order to analyze the coarse grained statistics. The set of the timeseries is separated in two categories: Normal subjects, and Coronary patients. The entropy analysis provides a quick and efficient tool for discrimination of the timeseries according to these categories. Healthy subjects provided more complex statistics compared to patient ones. In particular the healthy data files provided higher values of block Entropies compared to patient ones. Furthermore all this analysis may provide a useful statistical way to the understanding of the human cardiac system.
Keywords: Symbolic dynamics, Block Entropies, lumping, Electrocardiograms, Heart Rate Variability

1 Introduction

During the last years various methods have been applied to the analysis of ECG signals, besides the Fourier transform and spectral methods. Non-linear techniques have given quite satisfactory results in many cases and have attracted the interest of many authors, especially to ECG data. Dealing with ECG analysis we examine the R-R Heart Rate Variability signal (R-R HRV), which is strongly coupled with the human heart's dynamical system. Heart Rate Variability (HRV) refers to the beat-to-beat alterations in heart rate. R-R refers to the time interval between R-wave peaks in the ECG diagram. RR HRV signals have been studied in detail during the last decades and they serve as a means for exposing cardiac dysfunction.

Many authors have applied techniques of non linear dynamics to Electrocardiograms (ECG) Heart Rate Variability (HRV) signals. Since Goldberger [1] and Babloyantz [2] up to nowadays works as Costa et al. [2], the researchers have been arguing about the usefulness of examining ECG signals with non linear methods, and whether or not the signal of ECG HRV is random or not. Yet, all of them agree about the differences between Healthy and Patients in the complexity of the signals. In general, healthy and young subjects exhibit stronger complexity behavior compared to patients or old subjects. (By the term patient in this work, we mean a coronary patient, whose the ECG seems to be normal with no abnormalities).

In this work we perform an entropy analysis of the coarse grained dynamics of these two categories of subjects: healthy and patient. As we will anticipate in the sequel, the Calculation of Block Entropies by lumping, allows one to distinguish in principle between healthy subjects and coronary patients. Thereby, entropy analysis provides a useful and quick tool for long recorded ECG analysis, yielding results coherent to other non linear dynamics' methods such as the Correlation dimension estimation [20, 22, 24].

In order to apply this Entropy analysis we generate a new coarse grained timeseries, taking values on a finite alphabet, with the method of level crossing [22, 26]. 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 [17]. The examination of the Entropy statistics allows one to draw some general conclusions.

The article is articulated as follows:

In section 2 we present the data collection and the aspect of stationarity. In section 3 we demonstrate the coarse graining technique of level crossing, as applied to the present manuscript. In section 4 the entropy analysis by lumping is considered in some detail. In section 5 we depict the results of the entropy analysis by lumping, as applied to different categories of subjects. Finally, in the last section we draw some general conclusions.

2 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 aged 25-29 yrs, 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 coronary artery disease 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 ECGs recordings were preferred for more accurate analysis and since large amounts of data ensure higher precision of the results; especially in non-linear 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.

There are two ways of collecting large amounts of data. The first is the Holter recording, which is non-stationary by nature and the second is to use a digital Electrocardiograph and acquire ECGs of forty minutes about. Although Holter recordings are large enough, they are not stationary; the subject may be moving, sleeping, eating, running, etc, thus varying the external conditions significantly. In contrast, long recordings with the electrocardiograph are 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 as well. The most common method for investigating the stationarity of a process is by examining the mean and the Autocorrelation function for the entire timeseries and subparts. Then stationary subparts were taken and forwarded for been analyzed.

Another data set we used was found in the MIT fantasia database, which is available on http://physionet.caregroup.harvard.edu/physiobank/database/fantasia/subset/.

This database contains long ECG recordings of young healthy subjects and old still healthy subjects.

3 Generation of the coarse grained signal

The coarse grained timeseries were constructed using the level crossing method. According to this method, thresholds are supposed, and if a value of the data is between these thresholds, (greater or less than specific values), then this value is replaced with a specific linguistic value or a letter. Under this discretization the timeseries is modified to another one, which contains "linguistic" dynamics' characteristics.

As an example we can look at the FIGURE 1. A healthy subject RR HRV timeseries is coarse grained for a 10 letters alphabet. FIGURE 1 depicts the original timeseries, the coarse graining of this timeseries, and the corresponding letters substituting the values of the coarse graining timeseries.

The first query is whether the main statistical properties of the timeseries are remaining invariant via this transformation. Of course the values of the mean and the standard deviation will be changed in proportion to the linguistic values.

The part of the timeseries' spectrum, which is related to high frequencies, is expected to be lost. This happens because the coarse graining is a filter actually which smoothens the timeseries. This can be evidently noticed in the coherence function estimate (cfe) plots in FIGURE 2.

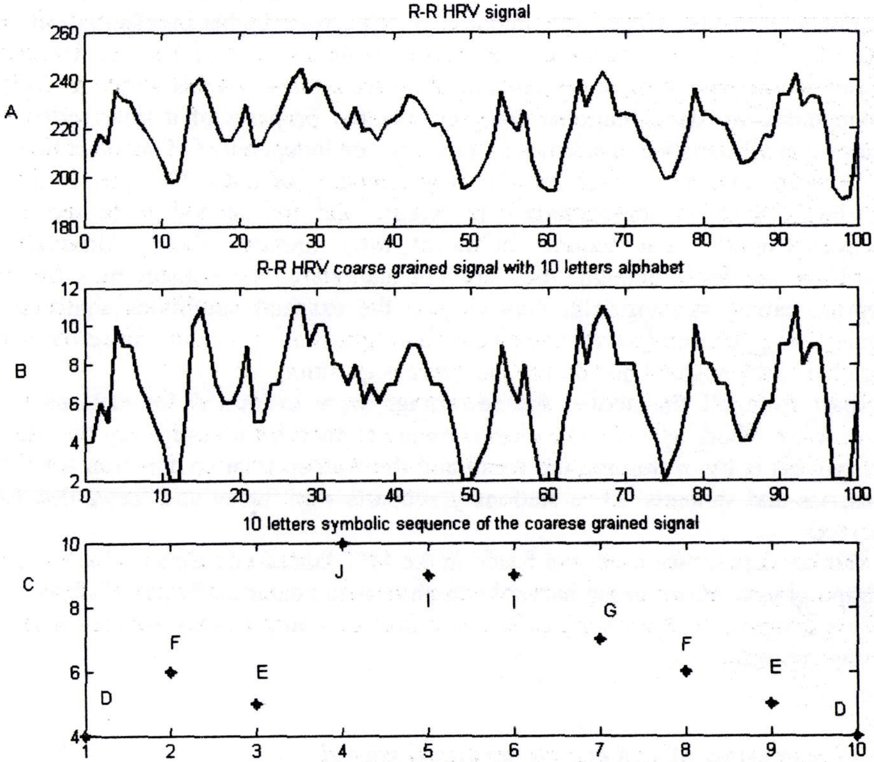

FIGURE 1: Part "A" depicts an RR HRV timeseries of a healthy subject. Part "B" is the coarse graining of the timeseries for a 10 letters alphabet. Part "C" represents the coarse graining values and the corresponding letters of a timeseries' snapshot. The first 10 values are: D, F, E, J, I, I, G, F, E, D.

The less of the number of letters to the linguistic timeseries, the less coherence in high frequencies will occur. From the scope of time domain this can be hardly identified in the autocorrelation's plot in FIGURE 3.

However the decorrelation time remains unaffected, as it can be easily noticed. This is quit important, as the autocorrelation function is strongly related to the timeseries statistics.

Therefore, we are expecting a losing of the information related to high frequencies of the spectrum only.

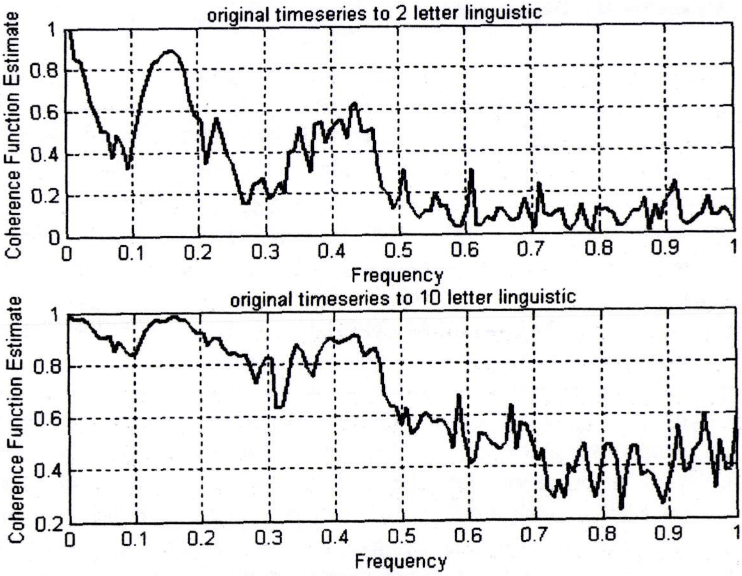

FIGURE2: Coherence function estimate plots between pure and coarse grained timeseries.

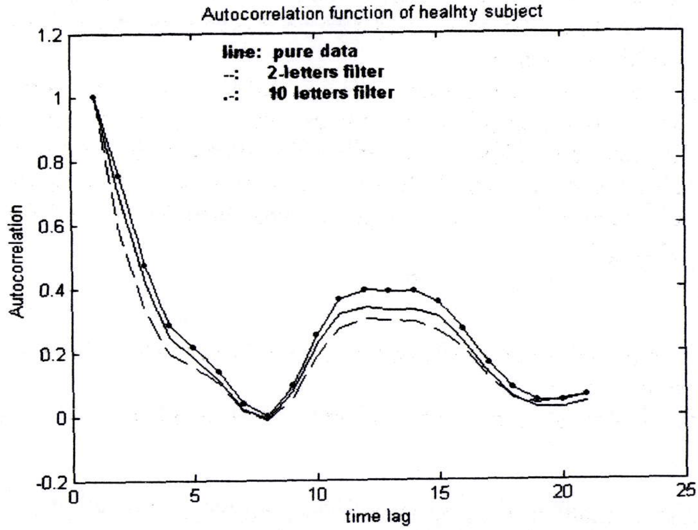

FIGURE3: Autocorrelation function of pure and coarse grained timeseries. With a continuous line is depicted the initial data. With a dashed line the is depicted the 2-letter symbolic coarse grained timeseries. The dotted line represents the corresponding with 10-letter.

4 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 [21]. In particular, it has been shown that the scaling of the block entropies with length gives sometimes interesting information on the structure of a sequence [10,11].

In the light of these results, we have tried to invent some criteria, which could illuminate the structure of a symbolic sequence and give us some more specific information beyond the one provided by block entropies computed by gliding. In fact, we have derived as entropy criterion for the specific, yet quite important algorithmic property of automaticity of the sequence.

In [13], we have examined how the estimation of the block entropies may actually depend on the way of reading, that is on 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 [19]. The coarse grained statistical properties of the attractors have been subsequently derived, with emphasis on the behavior of the block entropies.

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

Moreover, in [7, 8] 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.

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,

$$A_1,\dots,A_n \,/\, A_{n+1},\dots,A_{2n} \,/,\dots,/\, A_{jn+1},\dots,A_{(j+1)n} \qquad (1)$$

As we have mentioned already, we call this reading procedure lumping. We shall follow lumping in the sequel.

The following quantities characterize the information content of the sequence [10, 18]

(i) The dynamical (Shannon-like) block entropy for blocks of length n

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

where the probability of occurrence of a block , denoted by ,
is defined (when it exists) in the statistical limit as:

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

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.

We now turn to the selection problem, that is, to 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 [18] gives a partial answer to the selection problem asserting that for stationary and ergodic sources the probability of occurrence of a block is:

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

At this point, one can state an important theorem, first proved in [17], connecting the world of the machines with block entropy:

Theorem 1: *If the symbolic sequence* $(u_n)_{n\in N}$ *is m-automatic, then*

$$H\left(m^k\right) = H\left(m\right), \forall k \geq 1, \quad (9)$$

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 [7, 23]. Unlike the previous methods, the novelty of the entropy analysis by lumping gives results, which 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, see [17].

A standard example of substitutive sequence is the Thue-Morse sequence defined in the alphabet $\{0,1\}$ as the fixed point (i.e. the infinite iteration of the Transformation σ^T, starting with 0). The transformation is defined as follows:

$\sigma^T(0) = 01,\ \sigma^T(1) = 10$ (10)

and can be in an equivalent manner generated by a finite automaton of two states. (We state this without proof, for more details see [17]).

The sequence reads:

$(\sigma^T)^\infty(0) = 0110100110010110...$ (11)

This is a useful toy model, in which one can demonstrate the technique of *lumping*. Indeed for n=2 the successive blocks that one finds following the convention of lumping are:

01, 10, 10, 01, 10, 01, 01, 10, ..., in opposition with the convention of gliding where one finds:

01, 11, 10, 01, 10, 00, 01, 11, ...

In the same manner, for n=3, one finds with lumping the following blocks:

011, 010, 011, 001, 011, ..., in opposition with the convention of gliding where one finds:

011, 110, 101, 010, 100, 001, ...

It is evident in the previous example, that different ways of readings might appear, leading yielding different statistics for the entropies. This may depend on the kind of sequence.

In particular, for the Thue-Morse sequence, the convention of lumping leads obviously to the following invariance property for the block entropies:

$H^T(2^k) = H^T(2) = H^T(1) = \ln 2$, (12)

according to the theorem 1 above.

Also, the entropy analysis by lumping of some weakly chaotic systems, gives a rather characteristic entropy spectrum, as explained in [27]. 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, or the correlation functions, and this is what we use here.

Thus this way of analysis is attractive for being used in coarse grained ECG HRV signals. Indeed these signals exhibit complex dynamics when analyzed with classical non-linear methods. The associated block Entropies analysis comes to complete the detailed research of the human cardiac system's dynamics.

5 Results

In FIGURE 4, the block entropies per letter by lumping are depicted for a healthy subject.

Similar results hold for all of the 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.

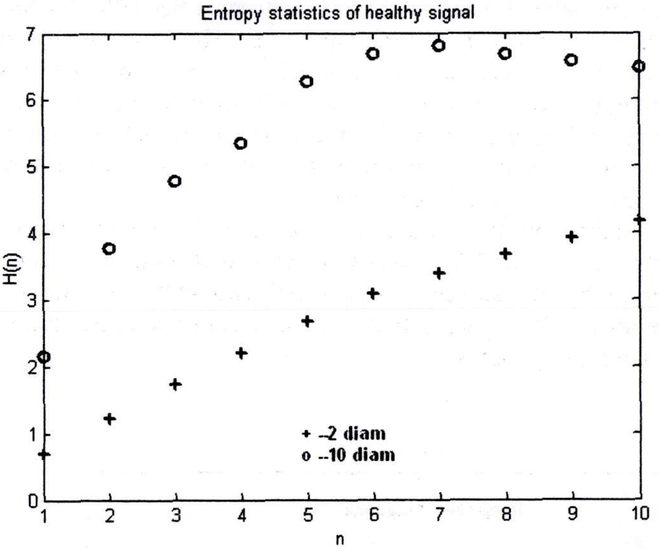

FIGURE4: Shannon like Entropy statistics for a Healthy subject, as a function of word length. With circles is depicted the coarse grained timeseries for 10-letter alphabet. With crosses is represented the corresponding for a 2-letter alphabet.

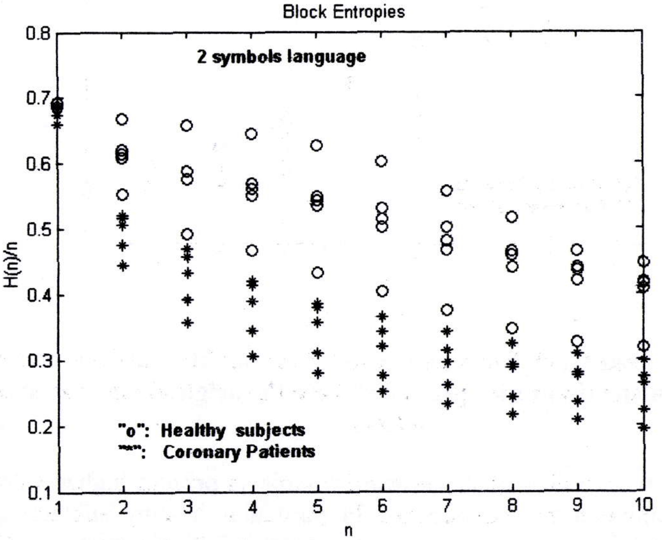

FIGURE5: Normalized block Entropy per letter for healthy subjects and coronary patients. The linguistic timeseries produced from the original one was made of two letters.

In FIGURE 5, the normalized associated block Entropies per letter by lumping are depicted (where we have good statistics) for a bipartition. The partitioning is done with the method of level crossing, where as crossing level, the median of the timeseries has been selected. If the signal bypasses this level, we call the coarse grained signal as U (upper) and if no we call it L (lower). With this technique the initial timeseries is coarse grained to a new discretized linguistic timeseries, and one can use standard methods of mathematical linguistics to analyze the final signal.

As a first step we analyze by the method of block Entropies analysis the ECG HRV signals from 5 healthy subjects and 5 coronary patients as discussed in section 2. The resulted entropies for the healthy subjects are depicted with an "o". For the patients the results are marked with a "*". It seems that there is a continuous transition between healthy subjects and coronary patients.

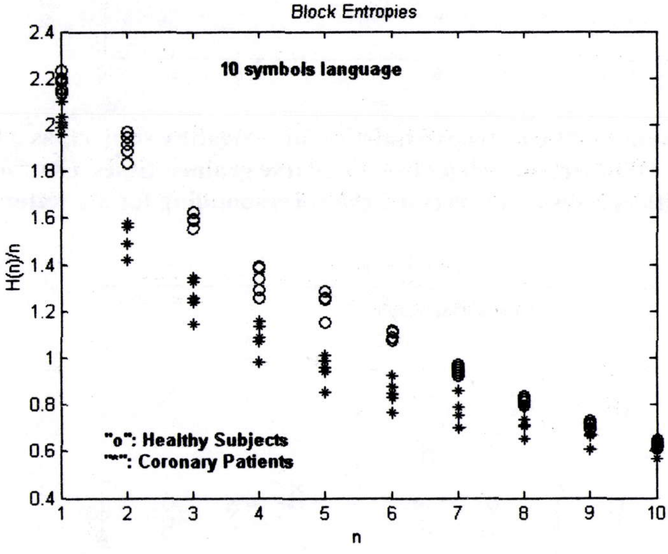

FIGURE 6 Normalized block Entropy per letter for healthy subjects and coronary patients. The linguistic timeseries produced from the original one was made of ten letters.

From the diagram 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.

Inverting this argument, one could imagine that with this method one can in principle distinguish in some cases between these two main categories of subjects. Clearly more data are needed for proving the advantage of the clinical practice of this method.

In FIGURE 6 the monotonic descent is more pronounced and steepest. However the distinction between healthy subjects and coronary patients is clearer. Indeed there is an obvious empty zone in the region of transition. This could be supported by a theoretical argument, which is that the healthy subject exhibit an increased complexity compared to the patient ones. This complexity is enforced by the more detailed coarse grained description by many letters.

We have tested also the value of this method for fantasia MIT database and as we will see below the results are encouraging.

The corresponding diagram for a 10 letter statistics is presented In FIGURE 7. In general the results are similar.

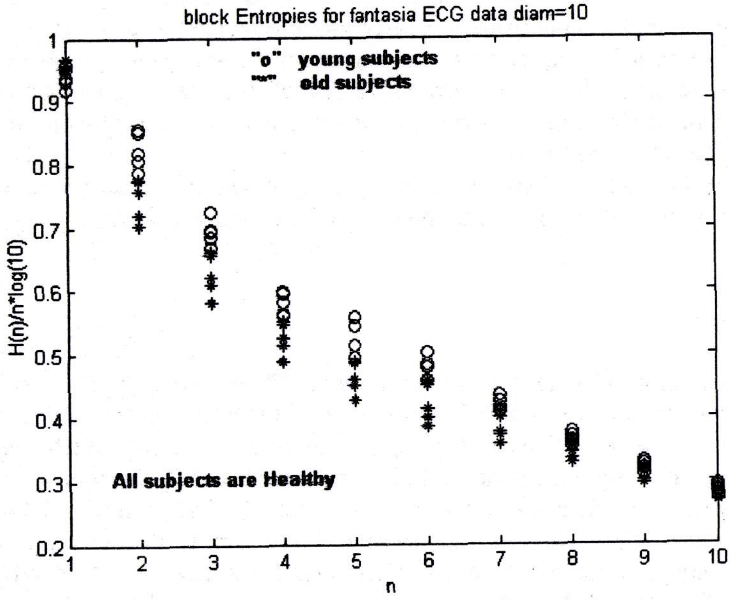

FIGURE 7: Normalized block Entropy per letter for healthy subjects and coronary patients, from fantasia MIT Database. The linguistic timeseries produced from the original one was made of ten letters.

Another important issue is the distinction between Healthy and young subjects and non patient but old ones. The data used for this analysis were taken from the fantasia database in MIT. The corresponding normalized associated block entropy per letter, for a 10 letter alphabet, is depicted in FIGURE 7. A monotonic decay can be also noticed again. The values for the healthy and young subjects are analogous to ones were taken by ours, for our data. Indeed healthy and young subjects present a normalized associated entropy per letter in the range from 70% for small word length, to 35% again. There are two zones, one for young and one for old subjects and again there is a continuous transition.

The old but non patient subjects present a corresponding h(n) from 55% for small word length, to 25% for large word length.

Based on these encouraging observations one could imagine clinical applications discussed also in the conclusions.

Retrospectively, an important point in connection with the previous results, is the qualitative explanation of the entropy values. The values of the entropy for coarse grained ECG HRV timeseries coming from healthy subjects seem to be increased, compared to the corresponding for timeseries coming from patients. This can be qualitatively been explained by:

1) The timeseries coming from healthy subjects possess high dimensionality dynamics and one can thus anticipate, 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 2-symbols linguistics only.

6 Discussion

The implementation of block entropy analysis of RR HRV signals attempts to apply an innovative statistical analysis in ECG timeseries. The coarse graining of the original timeseries with the method of level crossing generates another one, which projects the cardiac dynamics in a linguistic context. This is the «cardiac language» now, which can be either garrulous or prudent, except for non-intellectual. The linguistic dynamics are present in cybernetics for more than two decades. The scope of the specific analysis of ECG HRV signals is both practical and theoretical. Practical for its computational convenience. Indeed the whole analysis takes no more than several seconds.

The theoretical scope is the system's information content understanding. The coarse graining of the timeseries provides a projection of the system's information into a more normative sequence. This technique allows the entropy analysis under a linguistic concept.

Apart from that and remaining in this manuscript's scope, all the analysis provides an effective tool for distinguishing healthy subjects and coronary patients. 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 the Subject's heart robustness.

The method's results are coherent to the ECG non linear dynamics analysis' corollaries. The corollaries of the analysis of ECG HRV with non linear dynamical methods, manifest that youngness and health are related to high dimensional dynamics or strong complexity, while oldness 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. Also they have rather been useful for classification of the patient and healthy subjects, than providing absolute and precise values for the non linear parameters. Indeed over the literature, there are different correlation dimension values characterizing the same categories of subjects [25]. This does not indicate a drawback of the non linear methods, but it rather describes the difficulty of the analysis of complex dynamical timeseries.

Furthermore the method presented in this manuscript, enforces the non linear method's results as it presents coherent conclusions. Indeed it distinguishes the two categories of subjects clearly, quickly and without any other method's contribution.

The question is whether this method either itself or cooperating with others, can provide a coronary patient's detection tool.

The answer is rather philosophical. All these methods, help in suspecting diseases, or at least weakness. The prove of a disease is rather a medical business. The subject should make medical check ups like exhausting test, angiography etc. However a systematic recording of a person of 2 hours digital ECG's could reflex the general condition and robustness of his (her) heart. This method along with others such as non linear dynamics, could become useful in clinical practice.

7 Conclusion

In this manuscript we applied the innovative Block Entropies analysis by lumping in Hear Rate Variability timeseries which were generated from long recording Electrocardiograms. The set of the timeseries is separated in two categories: Normal subjects, and Coronary patients. These timeseries were filtered in order to generate a coarse grained linguistic dynamics. Then the Block Entropy analysis was applied in order to analyze the coarse grained statistics of the timeseries and finally provided a quick and efficient tool for discrimination of the timeseries according to these categories. Healthy subjects provided more complex statistics compared to patient ones. In particular the healthy data files provided higher values of block Entropies compared to patient ones. Apart from the Entropies values, all this analysis provided a useful statistical way to understand and confirm the strong complexity of the human cardiac system.

Acknowledgements

We would like to thank professors G. Nicolis and J. Nicolis for their invaluable guidance and support.

References

[1] A.L. Goldberger and B.J. West, *"Applications of nonlinear dynamics to clinical cardiology", Ann. NY Acad. Sci.* **504,** 195 (1987).

[2] A. Babloyantz and A. Destexhe, *"Is the normal heart a periodic oscillator?" Biol. Cybern.* **58,** 203-211 (1988).

[3] Costa et al., *J. of Cardiovascular Electrophysiology* **10(10),** 1350-1357 (1999).

[4] J.-P. Allouche, "On the complexity of infinite sequences", *Bull. Belg. Math. Soc.* **1(2),** 133-143 (1994).

[5] J.-P. Allouche, *"Algebraic and analytic randomness",* in *Noise, Oscillators and Algebraic Randomness,* M. Planat (Ed.), *Lecture Notes in Physics, Springer* **550,** 345-356 (2000).

[6] H. Bai-Lin, *Chaos,* World Scientific, Singapore (1994).

[7] V. Berthé, *"Conditional entropy of some automatic sequences", J. Phys. A: Math. Gen.* **27,** 7993- 8006 (1994); and references therein.

[8] V. Berthé, *"Entropy in deterministic and random systems",* in *Beyond quasicrystals, Les Houches,* F. Axel and D. Gratias (Eds.), Les Editions de Physique, Springer, 441-463 (1995); and references therein.

[9] J. Cassaigne, "*Complexity and special factors", Bull. Belg. Math. Soc.* **4,** 67-88 (1997).

[10] W. Ebeling and G. Nicolis, *Europhys. Lett.* **14(3),** 191-196 (1991).

[11] W. Ebeling and G. Nicolis, *"Word frequency and entropy of symbolic sequences: a dynamical perspective", Chaos, Solitons & Fractals* **2(6),** 635-650 (1992).

[12] P. Grassberger, *"Toward a quantitative theory of self-generated complexity", Int. J. Theor. Phys.* **25(9),** 907-938 (1986).

[13] K. Karamanos and G. Nicolis, *"Symbolic dynamics and entropy analysis of Feigen- baum limit sets", Chaos, Solitons & Fractals* **10(7),** 1135-1150 (1999).

[14] K. Karamanos, *"Entropy analysis of automatic sequences revisited: an entropy diagnostic for automaticity",* Proceedings of Computing Anticipatory Systems 2000, CASYS2000, *AIP Conference Proceedings* **573,** D. Dubois (Ed.), 278-284 (2001).

[15] K. Karamanos, *"From symbolic dynamics to a digital approach: Chaos and transcendence",* in *Noise, Oscillators and Algebraic Randomness,* M. Planat (Ed.), *Lecture Notes in Physics, Springer* **550,** 357-371 (2000).

[16] K. Karamanos, *"From symbolic dynamics to a digital approach", Int. J. Bif. Chaos* **11(6),** 1683-1694 (2001).

[17] K. Karamanos, *"Entropy analysis of substitutive sequences revisited", J. Phys. A: Math. Gen.* **34,** 9231-9241 (2001).

[18] A.I. Khinchin, *Mathematical Foundations of Information Theory,* Dover, New York (1957).

[19] N. Metropolis, M.L. Stein and P.R. Stein, *"On finite limit sets for transformations on the unit interval", J. Comb. Theory Ser. A* **15**, 25-44 (1973).

[20] G. Nicolis, *Introduction to Nonlinear Science,* Cambridge University Press, Cambridge (1995).

[21] G. Nicolis and P. Gaspard, *"Toward α probabilistic approach to complex systems"*, *Chaos, Solitons & Fractals* **4(1),** 41-57 (1994).
[22] J.S. Nicolis, *Chaos and Information Processing,* World Scientific, Singapore (1991).
[23] M. Queffélec, *"Spectral study of automatic and substitutive sequences",* in *Beyond quasicrystals, Les Houches,* F. Axel and D. Gratias (Eds.), Les Editions de Physique, Springer, 369-414 (1995); and references therein.
[24] M. Schröder, *Fractals, Chaos, Power Laws,* Freeman, New York (1991).
[25] C Bogaert, F Beckers, D Ramaekers, A Aubert, *Analysis of heart rate variability with correlation dimension method in a normal population and in heart transplant patients*. Autonomic Neuroscience: Basic and Clinical, **90,** 142-147 (2001).
[26] G. Nicolis, C. Nicolis, and J.S. Nicolis, *Chaotic Dynamics, Markov Partitions, and Zipf's Law, Journal of Statistical Physics* **54,** 915 (1989).
[27] K. Karamanos, *On the regognizability of Cantorian stochastic automata by generalized entropy-like quantities*, to appear in *Kybernetes* (2003)