ECG signals, computational intelligence, neurocomputing, fuzzy sets, information granules, granular computing, interpretation, classification, interpretability

Adam GACEK[1]

NEW FRONTIERS OF ANALYSIS, INTERPRETATION AND CLASSIFICATION OF BIOMEDICAL SIGNALS: A COMPUTATIONAL INTELLIGENCE FRAMEWORK

The methods of Computational Intelligence (CI) including a framework of Granular Computing, open promising research avenues in the realm of processing, analysis and interpretation of biomedical signals. Similarly, they augment the existing plethora of “classic” techniques of signal processing. CI comes as a highly synergistic environment in which learning abilities, knowledge representation, and global optimization mechanisms and this essential feature is of paramount interest when processing biomedical signals. We discuss the main technologies of Computational Intelligence (namely, neural networks, fuzzy sets, and evolutionary optimization), identify their focal points and elaborate on possible limitations, and stress an overall synergistic character, which ultimately gives rise to the highly symbiotic CI environment.

The direct impact of the CI technology on ECG signal processing and classification is studied with a discussion on the main directions present in the literature. The design of information granules is elaborated on; their design realized on a basis of numeric data as well as pieces of domain knowledge is considered. Examples of the CI-based ECG signal processing problems are presented. We show how the concepts and algorithms of CI augment the existing classification methods used so far in the domain of ECG signal processing. A detailed construction of granular prototypes of ECG signals being more in rapport with the diversity of signals analyzed is discussed as well.

ECG signals, Computational Intelligence, neurocomputing, fuzzy sets, information granules, Granular Computing, interpretation, classification, interpretability.

1. INTRODUCTION

We have been witnessing a significant number of various information technologies applied to ECG signal analysis, interpretation, and classification. Along with the steady progress of hardware platforms, new more advanced algorithmic developments have been reported and made practically relevant. There are several compelling reasons behind this progress, which mainly results from the exposure to the ongoing challenges inherently associated with the domain of ECG signal processing, analysis, and interpretation:

- ECG signals are one of the most important sources of diagnostic information. Their proper acquisition and processing provide an indispensible vehicle to support medical diagnosis. Acquired signals are affected by noise and call for advanced filtering techniques,
- A description and classification of ECG signals call for nonlinear mechanisms producing a suitable set of features (descriptors) of the signal so that the ensuing classifiers come with significant discriminatory capabilities. We observe a great deal of various ways used to describe ECG signals followed by the use of numerous classifiers,
- It is expected that any computerized interpretation of ECG signals has to be user-friendly, meaning that the results of classification/interpretation could be easily comprehended (perceived) by a human user. This requirement calls for an effective way of dealing with knowledge acquisition and knowledge manipulation when working with plain numeric signals.

[1] Institute of Medical Technology and Equipment, 118 Roosevelt St, 41-800 Zabrze, Poland.

Quite often these problems are intertwined and need to be dealt with in a holistic manner. We notice that some of them (preprocessing, filtering) require advanced nonlinear processing techniques while the others (interpretation) call for knowledge-oriented techniques. Altogether, a comprehensive methodological and algorithmic environment, which is offered through Computational Intelligence, comes as a viable alternative.

In this study, we discuss the main conceptual, methodological and algorithmic pillars of Computational Intelligence (CI), identify their main features and elaborate on their role in biomedical signal processing (Sections 2 and 3). In Section 4, several linkages between the main phases of classification problems of ECG signals and the contributing technologies of CI are discussed and supplied with a suite of representative examples encountered in the literature. A study on a granular representation of prototypes of ECG signals where the granularity quantifies the variability of signals being represented is discussed in Section 5. Concluding comments are offered in Section 6.

2. THE KEY TECHNOLOGIES OF COMPUTATIONAL INTELLIGENCE

In this section, we take a close look at the underlying technologies of neurocomputing, evolutionary optimization, and computing with information granules. We highlight their main features, contrast the associated research agendas and then elaborate on the emergence of the paradigm of Computational Intelligence viewed as an inherently synergistic setting, which dwells upon the strengths of its key components.

2.1. NEURAL NETWORKS AND NEUROCOMPUTING

There exists an immensely vast body of literature on neural networks. Neural networks are viewed as highly versatile distributed architectures realizing a concept of universal approximation [37,16], which offers a very much attractive feature of approximating nonlinear (continuous) mappings to any desired level of accuracy and in this way supporting various classification and mapping tasks.

The two main taxonomies commonly encountered in neurocomputing concern: (a) topologies of networks, and (b) a variety of ways of their development (training) schemes. With regard to the first coordinate of the taxonomy, one looks at a way in which individual neurons are arranged together into successive layers and a way in which processing is realized by the network, namely if this is of feedforward nature or there are some feedback loops within the structure. Typically, within the spectrum of learning scenarios one distinguishes between supervised learning and unsupervised learning however there are a number of interesting learning schemes, which fall in-between these two extreme positions (say, learning with partial supervision, proximity-based learning, etc.).

One needs to be aware of some limitations of neural networks that start manifesting in practical scenarios (those drawbacks might be alleviated to some extent but it is unlikely they will vanish completely). From the perspective of practice of neural networks, in Table 1 we compiled a list of advantages and shortcomings of neurocomputing.

Table 1. Neurocomputing: main advantages and limitations.

Advantages	Universal approximation capabilities, Significant learning abilities, a large repository of algorithms, well –developed and validated training methods, Distributed processing , Potential for significant fault tolerance, Efficient realizations of networks .
Limitations	Black-box architectures (require effort to interpret constructed networks), Mostly gradient-based learning with all limitations associated with this type of learning, Non-repetitive results of learning of the networks (depending upon initial learning condition, parameters of the learning algorithm, etc.), Slow, inefficient learning in presence of high-dimensional and large data sets.

From the perspective of applications, we should be aware that neural networks could offer a highly competitive solution however one has to proceed very prudently with the learning process. Most importantly, the learning results might not be repetitive: running the same method while starting from a slightly different initial configuration (say, a different random initialization of the connections of the neurons) may result in quite substantial differences in the performance of the constructed network. Likewise setting different numeric values of the learning environment (say, a learning rate) could lead to a different solution. A formation of the input space, which becomes of a genuine challenge, when dealing with highly dimensional data and a large number of data themselves, requires attention. Ignoring this problem may result in a highly inefficient learning producing quite poor, non-competitive results lacking generalization abilities.

We should stress that by no means neural networks can be sought as a plug-and-play technology. To the contrary: its successful usage does require careful planning, data organization and data preprocessing, a prudent validation and a careful accommodation of any prior domain knowledge being available. The black box nature of neural networks can bring some hesitation and reluctance to use the neural network solution and one has to be prepared for further critical evaluation of the obtained results.

2.2. GRANULAR COMPUTING: INFORMATION GRANULES AND THEIR PROCESSING

Information granules permeate numerous human endeavors [3,2,31,38]. No matter what problem is taken into consideration, we usually express it in a certain conceptual framework of basic entities, which we regard to be of relevance to the problem formulation and problem solving. This becomes a framework in which we formulate generic concepts adhering to some level of abstraction, carry out processing, and communicate the results to the external environment.

This remarkable and unchallenged ability of humans dwells on our effortless ability to construct information granules, manipulate them and arrive at sound conclusions. As another example, consider a collection of time series. From our perspective we can describe them in a semi-qualitative manner by pointing at specific regions of such signals. Specialists can effortlessly interpret various diagnostic biomedical signals including ECG recordings. They distinguish some segments of such signals and interpret their combinations. Experts can interpret temporal readings of sensors and assess the status of the monitored system.

Being convinced of the qualitative underpinnings of the problem, the challenge is to develop a computing framework within which all these representation and processing endeavors could be formally realized. The common platform emerging within this context comes under the name of Granular Computing. In essence, it is an emerging paradigm of information processing. It brings together the existing formalisms of set theory (interval analysis) [23], fuzzy sets [38,40], rough sets [27,28,29] under the same roof by clearly visualizing that in spite of their visibly distinct underpinnings (and ensuing processing), they exhibit some fundamental commonalities. In this sense, Granular Computing establishes a stimulating environment of synergy between the individual approaches.

Granular Computing forms a unified conceptual and computing platform. Yet, it directly benefits from the already existing and well-established concepts of information granules formed in the setting of set theory, fuzzy sets, rough sets and others. While Granular Computing offers a unique ability to conveniently translate the problem in the language of information granules, it is not free from limitations, refer to Table 2.

Table 2. Granular Computing: a list of main advantages and limitations.

Advantages	Efficient knowledge representation in the form of information granules and granular models, Transparency and high interpretability of resulting constructs, Diversity of formal schemes of representation of information granules.
Limitations	Lack of effective learning abilities, Prescriptive nature of granular constructs, Scalability issues.

2.3. EVOLUTIONARY AND POPULATION-BASED OPTIMIZATION

The attractiveness of this paradigm of computing stems from the fact that all pursuits are realized by a population of individual –potential solutions so that this offers a very much appealing opportunity of exploring or exploiting a search space in a holistic manner [15]. The search is realized by a population – a collection of individuals, which at each iteration (generation) carry out search on their own and then are subject to some processes of interaction.

In case of genetic algorithms, evolutionary methods, and population-based methods (say, genetic algorithms, evolutionary strategies, particle swarm optimization), in general, a population undergoes evolution; the best individuals are retained, they form a new population through recombination. They are subject to mutation. Each operator present in the search process realizes some mechanism of exploration or exploitation of the search space. A general processing scheme can be schematically outlined as follows

{evaluate population (individuals)
select mating individuals (selection process)
recombination
mutation}

The above generic sequence of processing steps is repeated (iterated).

In contrast to evolutionary methods, in the swarm-based methods [9], we encounter an interesting way of sharing experience. Each particle relies on its own experience accumulated so far but it is also affected by the cognitive component where one looks at the performance of other members of the population as well as an overall behavior of the population.

The essential phase of any evolutionary and population-based method (directly affecting its performance) is a representation problem. It is concerned about a way how to represent the problem in the language of the search strategy so that (a) the resulting search space is made compact enough (to make the search less time consuming) and (b) is well reflective of the properties of the fitness function to be optimized. By forming a suitable search space we pay attention to avoid forming extended regions of the search space where the fitness function does not change its values.

The key advantage of the methods falling under the rubric of these population-based optimization techniques is the genuine flexibility of the fitness function – there is a great deal of possibilities on how it can be formulated to capture the essence of the optimization problem. This translates into an ability to arrive at a suitable solution to the real-world task.

The inevitable challenges come with the need to assess how good the obtained solution really is and a formation of the efficient feature space itself.

Overall, the advantages and limitations of this paradigm of computing and optimization are collected in Table 3.

Table 3. Evolutionary and biologically inspired Computing; an overview.

Advantages	Mechanisms of global search, General form of fitness function, Abilities to deal with a wide range of structural and parametric optimization.
Limitations	Construction of search space (encoding and decoding mechanisms), Selection/adjustments of control parameters (e.g., crossover rate, mutation rate, recombination parameters), Assurance of optimality of solutions.

3. COMPUTATIONAL INTELLIGENCE: EMERGENCE OF SYNERGY

Computational Intelligence can be defined in many different ways. Let us start by recalling two definitions or descriptions, which are commonly encountered in the literature:

> *A system is computationally intelligent when it: deals with only numerical (low-level) data, has pattern recognition components, does not use knowledge in the AI sense; and additionally when it (begins to) exhibit (1) computational adaptivity; (2) computational fault tolerance, (3) speed approaching human-like turnaround, and (4) error rates that approximate human performance* [6,7]

The description provided by W. Karplus comes as follows:

> *CI substitutes intensive computation for insight how the system works. Neural networks, fuzzy systems and evolutionary Computation were all shunned by classical system and control theorists. CI umbrellas and unifies these and other revolutionary methods*

The first description captures the essence of the area. Perhaps today such a definition becomes slightly extended by allowing for some new trends and technologies, which are visible in the design of intelligent systems. Nevertheless the essence of CI is well-captured.

The comprehensive monograph on CI [33] emphasizes the importance of synergy of the contributing and very much complementary technologies of fuzzy sets, neurocomputing and evolutionary optimization. In a nutshell, CI is about effective and omnipresent mechanisms of synergy exploited in a variety of tasks of analysis and design of intelligent systems. The reader may refer to [12] and [24], which serve as comprehensive sources of updated material on Computational Intelligence.

The emergence of CI is justifiable and one would say, in some sense, unavoidable. Over time, being faced with more advanced problems, increased dimensionality and complexity of systems one has to deal with, neural networks, fuzzy sets and evolutionary computing started to exhibit some clear limitations. This is not startling at all as their research agendas are very much distinct (as we highlighted in the previous sections) and they focus on different aspects of the design of intelligent systems. The synergistic environment, in which knowledge representation, learning and global optimization go hand in hand, becomes highly desirable.

One may emphasize an important and enlightening linkage between Computational Intelligence and Artificial Intelligence (AI). To a significant extent, AI is a synonym of symbol-driven processing facilities. CI effectively exploits numeric data however owing to the technology of Granular Computing, it may invoke computing based on information described at various levels of granularity by inherently associating such granules with their underlying semantics described in a numeric or semi-numeric fashion (such as e.g., membership functions, characteristic functions or interval-valued mappings). The granularity of results supports the user-friendly nature of CI models. They can also form an important construct to be further used in facilitating interaction with the user as well as forming linkages with symbolic processing of AI constructs.

4. ECG SIGNAL ANALYSIS, CLASSIFICATION, AND INTERPRETATION: A CI REALIZATION

Alluding to the analysis of biomedical signals, especially ECG ones, we can see an important mapping between the fundamental quests existing in the area and the conceptual and computing faculties being effectively offered by the individual technologies of CI.

Formation of feature space A design of a suitable feature space in which patterns (ECG signals) are described is crucial to the effective performance of any classifier constructed at a later stage of signal classification. Here we witness a significant role played by Evolutionary Computing given the fact that the optimization problems are of combinatorial nature.

Classification of ECG signals Classification procedures (classifiers) realize mappings from the given feature space (in which individual signals are described) to the space of class membership (which could be Boolean, fuzzy, or probabilistic). We distinguish between linear and nonlinear classifiers. Here the role of neural networks is profoundly visible in the realization of nonlinear classifiers both in terms of realizing various forms of nonlinearities and ways in which learning of the ensuing neural network is being realized. Evolutionary computing arises as a significant contributor to the realization of a structural optimization of pattern classifiers.

Interpretation of ECG signals Here the role of information granules is critical as they help interpret the classification results (e.g., in terms of membership values of patterns to the corresponding classes) as well as the classifiers themselves. The role of information granules is also present when it comes to unsupervised learning – clusters come as a visible manifestation of a structure in a collection of ECG signals.

In what follows, let us elaborate in more detail on how some of these main classes of problems are supported with the use of the technology of fuzzy sets. Here the original data space, typically, an n-dimensional space of real number vectors, $\mathbf{R}^n$, is transformed via a finite collection of information granules (fuzzy sets), say, $A_1, A_2, \ldots, A_c$. We say that the input space (feature space) has been granulated. Each input $\mathbf{x}$ is perceived by the following classifier/analyzer through the "eyes" of the information granules, meaning that the following relationship is satisfied,

$$\mathcal{G}: \ \mathbf{R}^n \rightarrow [0,1]^c \tag{1}$$

where $\mathcal{G}$ stands for the mapping realized in terms of the information granules. Note that the result of the mapping is a c-dimensional vector positioned in the [0,1] hypercube.

There are at least three important and practically advantageous aspects of the mapping realized by information granules:

nonlinear mapping of the data space with an intent of forming information granules in such a way that the transformed data $\mathcal{G}(\mathbf{x}_1)$, $\mathcal{G}(\mathbf{x}_2)$,,…, $\mathcal{G}(\mathbf{x}_N)$ are more suitable to construct an effective classifier. We rely on the nonlinearity effect that can be carefully exploited to boost the discriminatory properties of the transformed feature space.

The tangible advantage results directly from the nonlinear nature of membership functions. A properly adjusted nonlinearity could move apart patterns belonging to different classes and bring closer those regions in which the patterns belong to the same category. For instance, patterns belonging to two classes and distributed uniformly in a one-dimensional space become well separated when transformed through a sigmoid membership function A, $A(x) = 1/(1+exp(-(x-2)))$ and described in terms of the corresponding membership grades. In essence, fuzzy sets play a role of a nonlinear transformation of the original feature space. While the patterns in the original space are distributed uniformly, their distribution in the space of membership degrees [0,1], $u=A(x)$ results in a very distinct distribution: two groups of patterns are located at the opposite ends of the unit interval with a large separation gap in-between.

reduction of the dimensionality of the feature space. While the dimensionality of the original feature space could be quite high (which is common in many classification problems), the dimensionality of the space of information granules is far lower, $c<<n$. This supports the developments of the classifiers, especially neural networks and reduces a risk of memorization resulting in poor generalization capabilities. We often witness this role of information granules in the construction of neuro-fuzzy systems.

information granules as essential constructs supporting the development of interpretable models. For instance, in rule-based systems (classifiers, analyzers), the condition parts (as well as conclusions) comprise information granules – interpretable entities, which make rules meaningful. A compelling example is displayed in Figure 1. Information granules are formed in the feature space. They are *logically* associated with classes in the sense that for each class its degree of class membership is a logic expression of the activation levels (matching degrees) of the individual information granules. The flexibility of the logic mapping is offered through the use of the collection of logic neurons (fuzzy neurons) whose connections are optimized during the design of the classifier.

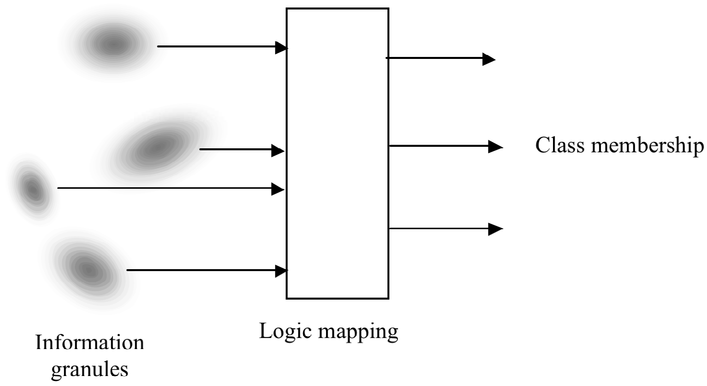


Fig. 1. An overall scheme of logic mapping between information granules – fuzzy sets formed in the feature space and the class membership grades.

We note that the ongoing intensive research in the area is reflective of the tendencies outlined above. Table 4 offers a snapshot of the main directions present in the literature.

Table 4. The technology of Computational Intelligence in ECG signal classification and interpretation: a collection of selected examples.

Study	Technologies of CI	Category of problem
Mitra et al 2006	Rough sets	Classification
Ozbay et al 2011	Type-2 fuzzy sets	Clustering of signals and classification
Yeh et al 2010	Fuzzy clustering	Classification
Chua et al to appear, 2011	Genetic algorithms, fuzzy sets	arrhythmia classification
Yeh et al 2010	Fuzzy sets	Feature selection (signal description)
Lee and Wang 2008	Fuzzy sets (ontology)	Signal description
Meau et al 2006	Neural networks and fuzzy sets (neurofuzzy system)	Signal classification
Engin 2004	Neural networks and fuzzy sets	Signal classification
Acharya et al 2003	Fuzzy sets and neural networks	Classification of heart rate
Kundu et al 2000	Fuzzy sets, neural networks , knowledge-based systems	Signal interpretation
Presedo et al 1996	Fuzzy sets	Ischemia detection
Gacek and Pedrycz 2003	Genetic segmentation of signals	Preprocessing of ECG signals
Gacek and Pedrycz 2006	Granulation of signals	Representation (compactification) of ECG signals
Pedrycz and Gacek 2001	Fuzzy automata	Classification of ECG signals
Barro et al 1991	Fuzzy sets (fuzzy grammars)	arrhythmia classification
Barro et al 1990	Fuzzy sets (rule-based systems)	Classification (beat labeling)
Korurek et al 2010	Particle swarm optimization and neural networks	Beat classification
Fei 2010	Particle swarm optimization and neural networks (support vector machines)	Arrhytmia detection
Moavenian and Khorrami 2010	Neural networks	Arrhytmia classification
Osowski et al 2008	Neural networks	Arrhytmia classification

The collected summaries of publications returned by Google Scholar (search done in August 2011)

ECG & fuzzy sets 5,510
ECG & rough sets 9,740
ECG & neural networks 16,400
ECG & particle swarm optimization 3,780
ECG & genetic algorithms 7,860

are also convincing stressing the visibility of the CI technologies in ECG signal analysis and classification.

5. DESIGNING GRANULAR REPRESENTATIVES OF ECG SIGNALS –A STUDY IN INFORMATION GRANULARITY

In this section, we present an interesting concept of granular representatives of a certain set of ECG signals; denote it by $\mathcal{X}$. To determine the best representative of $\mathcal{X}$, we can consider a mean (average) of the signals or some other statistical representative, like a median or medoid. The representatives of this nature are a result of solving an underlying optimization problem. For instance, the mean is a result of minimizing a Euclidean distance between the signals and their representative. The median results as a solution to the same problem in which the distance is specified as the Hamming one. In spite of the genuine diversity of possibilities of choosing the representative of $\mathcal{X}$, all of these variants share a striking resemblance. The obtained representative is just an element in the same space (feature space) in which the original signals were expressed. Thus if $\mathcal{X}$ is expressed in $\boldsymbol{X} \subset \boldsymbol{R}^n$, so is the space in which the representative of $\mathcal{X}$ becomes formed. Formally speaking, the representation problem gives rise to the formation of the representative of $\mathcal{X}$, call it $\mathbf{v}$, such that it represents (or approximates) all elements in $\mathcal{X}$. We can capture the essence of this category of the signal representation problems in the following manner

$$\mathcal{X} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\}, \ \mathbf{x}_k \in \mathbf{R}^n \rightarrow \mathbf{v} \in \mathbf{R}^n \tag{2}$$

As intuition suggests, by noting an inherent many-to-one nature of the mapping (many elements in $\mathcal{X}$ and a single representative) and in order to accommodate the diversity of the signals to be represented, one could envision that the structural complexity (a level of abstraction) of $\mathbf{v}$ is supposed to be higher than the original signals it has to represent. This entails that rather than being a vector of numeric entities, one may anticipate that the representative can be sought as a certain information granule being of non-numeric character. For instance, we may envision that such representative could be a collection of intervals or a family of fuzzy sets formed over $\mathbf{X}$. The granularity of information, which is inherently associated with the representative is fully reflective of the many-to-one nature of the mapping of the elements of $\mathcal{X}$ to a single representative. More generally, we can envision the representative to be realized as any granular construct, say $\boldsymbol{G}(\mathbf{R}^n)$ where $\boldsymbol{G}(.)$ stands for a family of information granules as discussed earlier. Alluding to the concise notation of the numeric prototype, we capture the granular counterpart by using the following expression

$$\mathcal{X} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\}, \ \mathbf{x}_k \in \mathbf{R}^n \rightarrow \mathbf{V} \in \boldsymbol{G}(\mathbf{R}^n) \tag{3}$$

The essence of the development of granular representatives can be viewed as an optimization problem of distribution of the available granularity of information where the granularity itself is treated as an important knowledge-based modeling resource. In a nutshell, given a predefined level of information granularity, we allocate it to the elements of the universe of discourse $\mathbf{X}$ in such a way the resulting granular representative captures most of the signals (viz. the signals are “contained” within the bounds of the information granules of the representative). The higher the admitted level of granularity ε^* is, the less specific (detailed) the granular representative becomes. This tendency is not encouraging. At the same time, with the lowered values of granularity, more signals are being “covered” (which is evidently advantageous). These two conflicting requirements need to be carefully reconciled during the design of the granular representatives.

It is worth noting that granular representatives are of particular interest in interpreting, analyzing, and comparing biomedical signals (e.g., ECG complexes) where a concept of a representative (norm) and its variation explicitly associated with the norm becomes of relevance. For instance, one could envision a template (representative) of normal ECG, where its *granular* nature reflects the variability of the existing signals and a way in which it is distributed over time. Any comparison of a new ECG complex is done with the granular template and on this basis a certain classification could be carried out.

The expressive power of the information granule **V** articulated with respect to $\mathcal{X}$ is higher than the original **v** in the sense that **V** "covers" (includes) some entries of $\mathbf{x}_k$. The broader the interval built around the numeric representative **v**, the more data points falls within the bounds of the information granule **V**. Note that the length of the interval of **V** may vary over the entire space when moving from one coordinate of $\mathbf{x}_k$ to another. There are no particular restrictions on the distribution (allocation) of granularity.

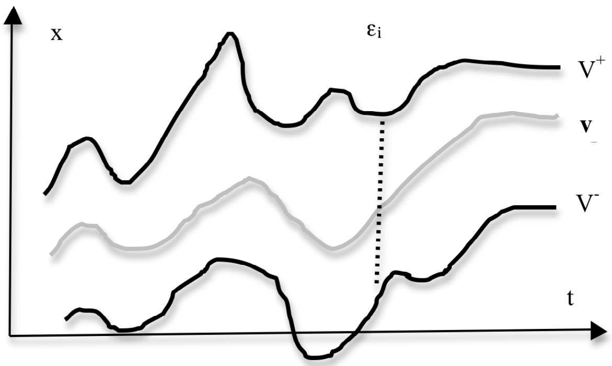


Fig. 2. Formation of granular prototype around the numeric representative.

We regard the level of granularity ε to be a useful source of knowledge representation whose distribution over **X** is instrumental in the maximization of the coverage requirement. More formally, we can translate the problem into the corresponding optimization task with the objective to allocate granularity along the universe of discourse **X** in such a way so that as many coordinates of $\mathbf{x}_k$ are included within the bounds of **V**, see Figure 2. The bounds of **V** are described in the interval-like form, that is **V**= $[[V_1^-,V_1^+]\,[V_2^-,V_2^+]..[V_n^-,V_n^+]\,\}]$ that is we are concerned with the interval-type of granularity of the representative.

The optimized performance index reads as follows

$$\text{Maximize card } \{(i, k)\,|x_{ki} \in [V_i^-,V_i^+]\} \tag{4}$$

with the maximization realized with respect to the vector of information granularity $\boldsymbol{\varepsilon} = [\varepsilon_1\ \varepsilon_2 \ldots\ \varepsilon_n]^T$ with the constraint imposed by the assumed cumulative level of overall granularity ε^*, that is:

$$\varepsilon^* = \sum_{i=1}^{n} \varepsilon_i \tag{5}$$

In this sense, we are faced with the constraint-based optimization problem. The optimization of the vector of information granularities **ε** is quite demanding as the values of the coordinates of the vector are linked in a quite indirect manner with the minimized performance index. Clearly, the problem we are facing here does not fall within the realm of gradient-based optimization. In contrast, some techniques of Evolutionary Computing could be of relevance here.

An overall flow of determining the granular prototypes comprises of two steps: (a) we pick up a certain element of $\mathcal{X}$, and (b) for some given value of ε*, we optimize a distribution of granularity so that the performance index becomes maximized. The visualization of this two-phase process is shown in Figure 3. Depending upon the element of $\mathcal{X}$ around which the granular representation is being constructed, we arrive at different values of the performance index. Likewise these values depend on some predetermined level of the granularity ε*.

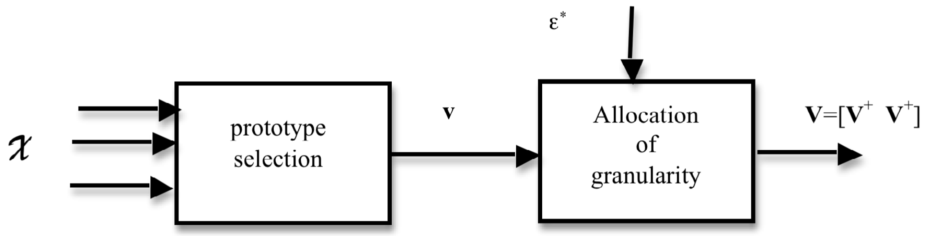


Fig. 3. The optimization (distribution) of granularity in the development of granular representatives of the signals: a two-phase development process.

Considering the overall flow of optimization, it is apparent that the selection procedure depends upon a choice of a certain value of the overall granularity ε*. To avoid being potentially affected by any particular selection, we introduce a global way of expressing the quality of a certain granular representative **V**. What is intuitively straightforward, is an observation that the relationship Q = Q(ε*) is non-decreasing function of ε* meaning that higher level of granularity available for distribution can result in covering more data points in **X**. Instead of admitting a particular value of ε* (whose choice is usually biased to some extent and implied by some design performance), we sweep through a range of values of ε^* starting from zero (in which case Q is typically close to zero) and moving to some upper bound, say ε_{max}. At the same time, we record the corresponding values of Q (those are optimized values of Q for the specific value of ε*). The resulting area under curve (AUC) serves as a viable global indicator of the suitability of the granular representative **V** (formed via the formation of the granular representation of **v**),

$$\text{AUC}=\int_0^{\varepsilon_{max}} Q(\varepsilon)d\varepsilon \tag{5}$$

The higher the value of AUC, the better the granular representative **V** is and this quantification is of general character independent from the required level of granularity. In this manner any choice of **v**, we have started with can be quantified in terms of the AUC.

As an illustration, we consider a collection of six normal ECG complexes coming from the MIT-BIH arrhythmia database shown in Figure 4. While these signals exhibit some similarities, there is a certain level of variability present among them. The evaluation of each of the signals in terms of the AUC measure, see Figure 5, indicates that the fifth one is the most suitable as a granular representative and returns the highest value of this measure.

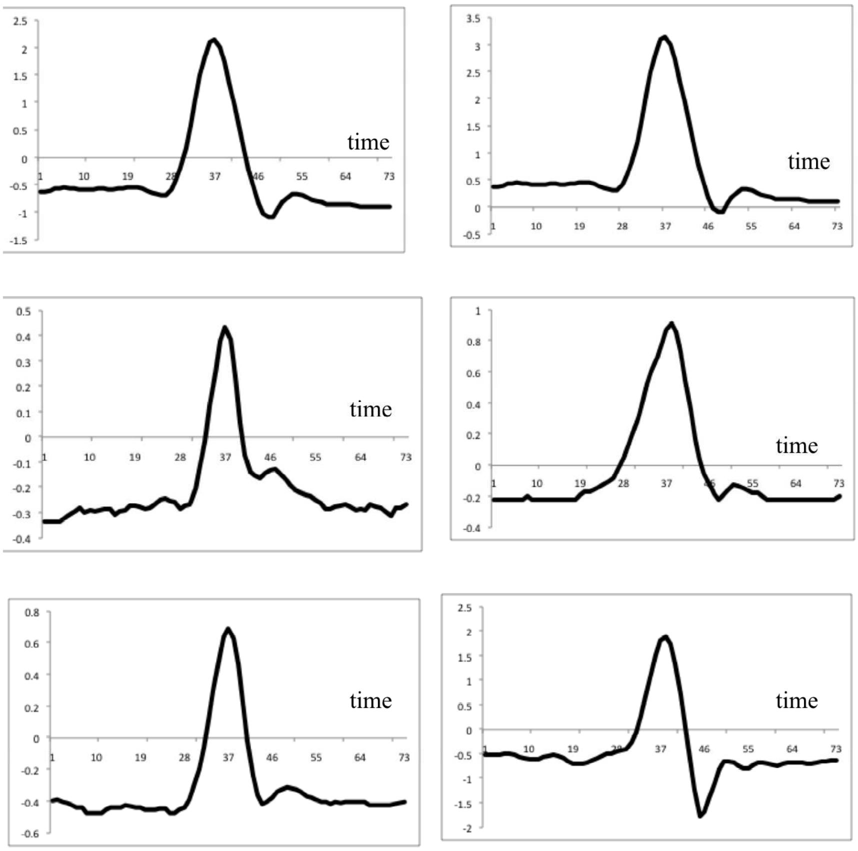


Fig. 4. A collection of 6 normal ECG signals.

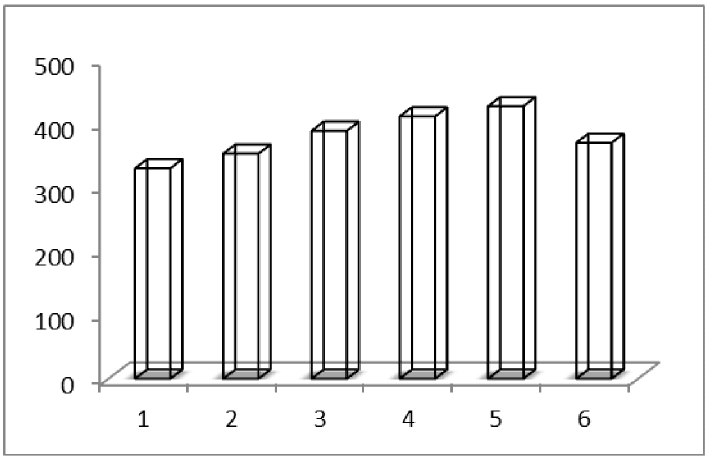
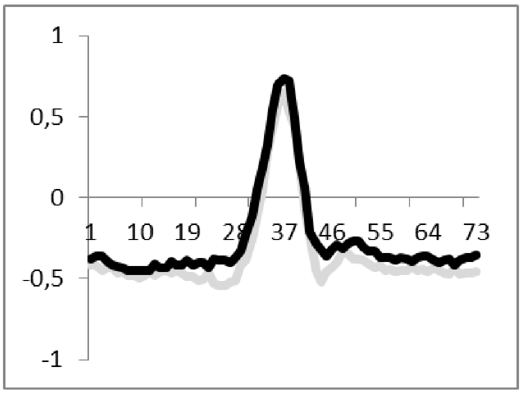
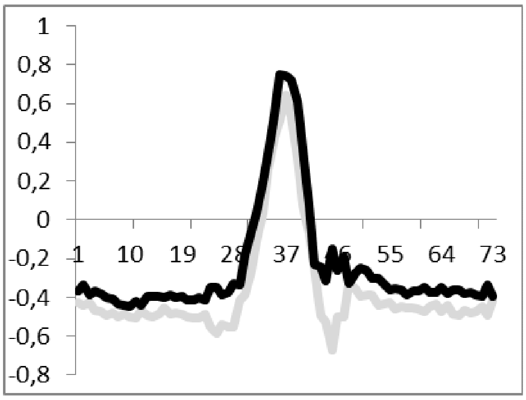


Fig. 5. The AUC values computed for the ECG signals.

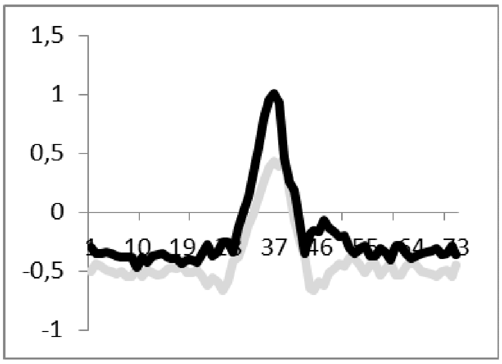
Considering some selected values of ε* that is 0.05, 0.10, and 0.20 (those values are picked up for illustrative purposes), the lower and upper bounds of the granular representations of the signal are illustrated in Figure 6.



(a)



(b)



(c)

Fig 6. Granular realization of the representative for selected values of ε*: (a) 0.05, (b) 0.10, and (c) 0.20.

We note that the granularity of the representative becomes more apparent with the increase of the allowed granularity level. Furthermore the distribution of granularity (the lengths of the intervals) differs quite substantially: it is non-existent in the neighborood of the R peak while it shows up in other regions of the QRS complex.

6. CONCLUSIONS

A wealth of problems of signal processing (filtering, discrimination, interpretation) can be effectively formulated and solved in the setting of Computational Intelligence. CI provided new, attractive opportunities by bringing a facet of nonlinear processing (supported by neural networks) and deliver a realization of a variable perspective at the problem description through information granularity. Furthermore evolutionary computing helps approach the system design from the perspective of structural optimization – a unique opportunity not commonly available when dealing with the “standard” methods of signal processing or classification.

We outlined the fundamentals of Computational Intelligence showing that the synergy of the technologies of fuzzy sets becomes a vital component of the design of intelligent systems. With this regard, fuzzy sets or being more general, information granules, form an important front- and back-end of constructs of CI. By forming the front end, they help develop a suitable view at ECG data, incorporate available domain knowledge and come up with a feature space that supports the effectiveness of ensuing processing, quite commonly engaging various schemes of neurocomputing or evolutionary neurocomputing. Equally important role is played by fuzzy sets in the realization of the back end of the overall processing scheme: they strengthen the interpretability of classification results as well as provide useful interpretation faculties to neural networks or help develop logic mappings in the form fuzzy logic neural networks.

Our intention was to highlight the main ideas and the principles of research agenda of Computational Intelligence as well as show that they are well aligned with the challenges we witness in ECG signal processing and interpretation. There have been a number of promising studies at the junction of CI and ECG classifier; they form a solid starting point for further progression.

7. ACKNOWLEDGEMENT

This study was in part financed by the Polish Ministry of Science and Higher Education and by the Polish National Science Center.

BIBLIOGRAPHY

[1] ACHARYA U.R., BHAT P.S., IYENGAR S.S., RAO A., DUA S., Classification of heart rate data using artificial neural network and fuzzy equivalence relation, Pattern Recognition, Vol. 36, 2003, pp. 61–68.

[2] BARGIELA A., PEDRYCZ W., Recursive information granulation: aggregation and interpretation issues, IEEE Trans. Systems, Man and Cybernetics-B, Vol. 33, No. 1, 2003, pp. 96-112.

[3] BARGIELA A., PEDRYCZ W., Granular Computing: An Introduction, Kluwer Academic Publishers, Dordrecht, 2002.

[4] BARRO S., RUIZ R., PRESEDO J., MIRAI J., Grammatic representation of beat sequences for fuzzy arrhytmia diagnosis, Int J Biomed Comput, Vol. 21, 1991, pp. 245-259.

[5] BARRO S., RUIZ R., MIRAI J., Fuzzy beat labeling for intelligent arrhythmia monitoring, Computers and Biomedical Research, Vol. 2, 1981, pp. 240-258.

[6] BEZDEK J.C., On the relationship between neural networks, pattern recognition and intelligence, Int. J. of Approximate Reasoning, Vol. 6, No. 2, 1992, pp. 85-107.

[7] BEZDEK J.C., What is computational intelligence, In: Zurada JM, RJ Marks II, Robinson CJ (eds.), Computational Intelligence Imitating Life, IEEE Press, Piscataway, NJ, 1994, pp. 1-12.

[8] CHUA T.W., TAN W., Non-singleton genetic fuzzy logic system for arrhythmias classification, Engineering Applications of Artificial Intelligence, 2011, (to appear).

[9] ENGELBRECHT A.P., Fundamentals of Computational Swarm Intelligence, Wiley, London, UK, 2005.

[10] ENGIN M., ECG beat classification using neuro-fuzzy network, Pattern Recognition Letters 25, 2004, pp. 1715–1722.

[11] FEI S.W., Diagnostic study on arrhythmia cordis based on particle swarm optimization-based support vector machine, Expert Systems with Applications, Vol. 37, 2010, pp. 6748–6752.

[12] FULCHER J., JAIN L.C. (eds), Computational Intelligence: A Compendium, Springer, Berlin, 2008.

[13] GACEK A., PEDRYCZ W., A genetic segmentation of ECG signals, IEEE Transactions on Biomedical Engineering, Vol. 50, No. 10 , 2003, pp. 1203–1208.

[14] GACEK A., PEDRYCZ W., A Granular description of ECG Signals, IEEE Transactions on Biomedical Engineering, Vol. 53, No. 10 , 2006, pp. 1972–1982.

[15] GOLDBERG D.E., Genetic Algorithms in Search, Optimization, and Machine Learning, Addison Wesley, Reading, MA, 1989.

[16] HAYKIN S., Neural Networks: A Comprehensive Foundation, 2nd ed. Prentice Hall Upper Saddle River, NJ, 1999.
[17] KORUREK M., DOGAN B., ECG beat classification using particle swarm optimization and radial basis function neural network, Expert Systems with Applications, Vol. 37, 2010, pp. 7563–7569.
[18] KUNDU M, NASIPURI M, BASU D.K., Knowledge-based ECG interpretation: a critical review, Pattern Recognition, Vol. 33, 2000, pp. 351-373.
[19] LEE C.S., WANG M.H., Ontological fuzzy agent for electrocardiogram application, Expert Systems with Applications, Vol. 35, 2008, pp.1223–1236.
[20] MEAU Y.P. et al., Intelligent classification of electrocardiogram (ECG) signal using extended Kalman Filter (EKF) based neuro fuzzy system, Computer Methods and Programs in Biomedicine Vol. 8, No. 2, 2006, pp. 157–168.
[21] MITRA S, MITRA M, CHAUDHURI B.B., A rough-set-based inference engine for ECG classification, IEEE Trans. on Instrumentation and Measurement, Vol. 55, No. 6, 2006, pp. 2198-2206.
[22] MOAVENIAN M., KHORRAMI H, A qualitative comparison of artificial neural networks and support vector machines in ECG arrhythmias classification, Expert Systems with Applications, Vol. 37, 2010, pp. 3088–3093.
[23] MOORE R., Interval analysis, Englewood Cliffs, NJ: Prentice-Hall, 1966.
[24] MUMFORD C.L., JAIN L.C. Eds., Computational Intelligence, Springer, Berlin, 2009.
[25] OSOWSKI S, MARKIEWICZ T, TRAN HOAI L., Recognition and classification system of arrhythmia using ensemble of neural networks, Measurement, Vol. 41, 2008, pp. 610–617.
[26] ÖZBAY Y, CEYLAN R., KARLIK B., Integration of type-2 fuzzy clustering and wavelet transform in a neural network based ECG classifier, Expert Systems with Applications, Vol. 38, 2011, pp. 1004–1010.
[27] PAWLAK Z, Rough sets, Int. J. Comput. Inform. Sci., Vol. 11, 1982, pp. 341–356.
[28] PAWLAK Z., Rough Sets. Theoretical Aspects of Reasoning About Data, Kluwer Academic Publishers, Dordercht, 1991.
[29] PAWLAK Z., SKOWRON A., Rough sets and Boolean reasoning, Information Sciences, Vol. 177, No. 1, 2007, pp. 41-73.
[30] PEDRYCZ W., GACEK A., Learning of fuzzy automata, Int. J. Computational Intelligence and Applications, Vol. 1, 2001, pp. 19-33.
[31] PEDRYCZ W., BARGIELA A., Granular clustering: a granular signature of data, IEEE Trans. Systems, Man and Cybernetics, Vol. 32, No. 2, 2002, pp. 212-224.
[32] PEDRYCZ W, GACEK A., Temporal granulation and its application to signal analysis Information Sciences, 143, 1-4, 2002, pp. 47-71.
[33] PEDRYCZ W., Computational Intelligence: An Introduction, CRC Press, Boca Raton, Fl, 1997.
[34] PRESEDO J. et al., Fuzzy modelling of the expert's knowledge in ECG-based ischaemia detection, Fuzzy Sets and Systems, Vol. 77, 1996, pp. 63-75.
[35] YEH Y.C., WANG W.J., CHIOU C.W., A novel fuzzy c-means method for classifying heartbeat cases from ECG signals, Measurement, Vol. 43, 2010, pp. 1542–1555.
[36] YEH Y.C., WANG W.J., CHIOU C.W., Feature selection algorithm for ECG signals using range-overlaps method, Expert Systems with Applications, Vol. 37, 2010, pp. 3499–3512.
[37] WASSERMANN P.D., Neural Computing: Theory and Practice, Van Nostrand, Reinhold, New York, NY, 1989.
[38] ZADEH L.A., Fuzzy sets, Information & Control, Vol. 8, 1965, pp. 338-353.
[39] ZADEH L.A., Towards a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic, Fuzzy Sets and Systems, Vol. 90, 1997, pp. 111-117.
[40] ZADEH L.A., Toward a generalized theory of uncertainty (GTU)–an outline, Information Sciences, 172, 200.