Automated Ischemic Beat Classification Using Genetic Algorithms and Multicriteria Decision Analysis

Yorgos Goletsis, *Member, IEEE*, Costas Papaloukas, Dimitrios I. Fotiadis*, *Member, IEEE*, Aristidis Likas, *Senior Member, IEEE*, and Lampros K. Michalis

***Abstract*—Cardiac beat classification is a key process in the detection of myocardial ischemic episodes in the electrocardiographic signal. In the present study, we propose a multicriteria sorting method for classifying the cardiac beats as ischemic or not. Through a supervised learning procedure, each beat is compared to preclassified category prototypes under five criteria. These criteria refer to ST segment changes, T wave alterations, and the patient's age. The difficulty in applying the above criteria is the determination of the required method parameters, namely the thresholds and weight values. To overcome this problem, we employed a genetic algorithm, which, after proper training, automatically calculates the optimum values for the above parameters. A task-specific cardiac beat database was developed for training and testing the proposed method using data from the European Society of Cardiology ST-T database. Various experimental tests were carried out in order to adjust each module of the classification system. The obtained performance was 91% in terms of both sensitivity and specificity and compares favorably to other beat classification approaches proposed in the literature.**

***Index Terms*—Automated ischemia detection, genetic algorithms (GAs), multicriteria analysis.**

I. Introduction

MYOCARDIAL ischemia is caused by insufficient blood flow to the muscle tissue of the heart. This reduced blood supply may be due to narrowing of the coronary arteries, obstruction by a thrombus, or, less commonly, due to diffuse narrowing of arterioles and other small vessels within the heart. Ischemia is one of the leading causes of death in modern societies and, as a consequence, its early diagnosis and treatment is of great importance [1]–[3]. In the electrocardiographic (ECG) signal, ischemia is expressed as slow dynamic changes of the ST segment and/or the T wave [4], [5]. Long duration electrocardiography, like Holter recordings or continuous ECG monitoring in the coronary care unit, is a simple and noninvasive method to observe such alterations. The development of suitable automated analysis techniques can make this type of ECG recording very effective in supporting the physician's diagnosis and guide patient management in clinics and clinical applications.

The accurate ischemic episode detection in the recorded ECG is based on the correct classification of the ischemic cardiac beats. Several techniques have been proposed for ischemic beat classification, which evaluate the ST segment changes and the T-wave alterations with different methodologies. Their evaluation process is based upon parametric modeling [6], [7], wavelet transform [8], set of rules [9], [10], and artificial neural networks [11]–[15]. Although these methods have achieved satisfactory results, improvements can be made. None of the above techniques combines high performance with the ability to interpret the classification results.

In the present study, multicriteria decision analysis (MCDA) is utilized for ischemic beat classification. Our analysis is based on the comparison between prototypes and beats, to be classified using a similarity system [16], [17], according to the outranking concept [18]. This MCDA approach avoids using distance measures and is able to utilize data expressed in different units for each criterion. MCDA tools have been used for a series of applications ranging from urban management and transportation [19] to project [20] or financial evaluation [21]. Although MCDA has been used in several classification problems [22], only a few cases have been reported concerning medical applications [23], [24]. A possible reason for this could be the difficulty in modeling the doctors' preference process and value system through MCDA. Doctors' knowledge often needs to be transformed to threshold and weight values, a task that requires a considerable effort from the medical expert. Moreover, it is rather unlikely to achieve similar values from different experts, since each one exhibits differences in his/her value system.

Aiming at overcoming these difficulties and at supporting automated diagnosis, we employ a genetic algorithm (GA) in order to properly adjust all necessary parameters of the MCDA model. The simplicity in the development of GAs and their ability to work equally well in a series of optimization problems have led to several successful applications of GAs for medical tasks, such as medical imaging and medical data mining [25]. GAs have been broadly used as learning tools for fuzzy systems, mainly

Manuscript received May 28, 2003; revised October 19, 2003. *Asterisk indicates corresponding author.*

Y. Goletsis is with the Unit of Medical Technology and Intelligent Information Systems, Department of Computer Science, University of Ioannina, GR 45110 Ioannina, Greece (e-mail: goletsis@cc.uoi.gr).

C. Papaloukas was with the Unit of Medical Technology and Intelligent Information Systems, Department of Computer Science, University of Ioannina, GR 45110 Ioannina, Greece. He is now with the Department of Biological Applications and Technology, University of Ioannina, GR 45110 Ioannina, Greece (e-mail: papalouk@cc.uoi.gr).

*D. I. Fotiadis is with the Unit of Medical Technology and Intelligent Information Systems, Department of Computer Science, University of Ioannina, and also with the Biomedical Research Institute, FORTH and Michaelideion Cardiology Center, GR 45110 Ioannina, Greece (e-mail: fotiadis@cs.uoi.gr).

A. Likas is with the Department of Computer Science, University of Ioannina, GR 45110 Ioannina, Greece (e-mail: arly@cs.uoi.gr).

L. K. Michalis is with the Department of Cardiology, Medical School, University of Ioannina, GR 45110 Ioannina, Greece (e-mail: lmihalis@cc.uoi.gr).

Digital Object Identifier 10.1109/TBME.2004.828033

in the field of genetic fuzzy rule-based systems [26]. In the multiple criteria area, GAs have been mostly used in multiobjective programming, for identifying nondominated (Pareto optimal) solutions [27]. Several cases can be found in the literature where a GA is used for training (e.g., a neural network), including a case related to cardiac beat classification [28]. In our application, the required parameters for the MCDA model were defined after applying a GA in a training set. A task-specific cardiac beat database was developed especially for training and testing the beat classification system.

In Section II, the multicriteria method is described. A background overview follows, concerning GA implementation. In Section IV, the integrated model for ECG beat classification is described. In the same section, the ECG signal processing procedure is presented as well as the details of GA application and MCDA utilization. The tests that were carried out in order to evaluate the classification system, are given in Section V. In addition, the performance of the proposed method is compared to other similar approaches. In Section VI, the implementation of the integrated technique is discussed; its advantages and disadvantages compared to other systems are given, while some possible further improvements are mentioned as well.

II. Multicriteria Decision Analysis

The multicriteria classification problem (known in multicriteria decision aid as a sorting problematic—or problematic β [16]) deals with assigning patterns to one or several categories. In our approach, the assignment is achieved through the examination of the intrinsic value of the pattern by referring to preestablished norms. A category can be defined either by one reference pattern—prototype—(being a monoprofile category) or by a set of patterns (a multiprofile category). The second is harder to treat, but it offers greater flexibility as far as categories modeling is concerned. This implies that it is possible for a pattern to be assigned in different manners (through different prototypes) to a category [29]. The assignment procedure associates a pattern with a category. The category is represented by reference patterns. The test pattern is compared with the reference pattern and is assigned to the category where the most similar reference pattern belongs.

More specifically, our case deals with assigning cardiac beats to predefined multiprofile categories (ischemic or not), which are not ordered. The latter implies that there are no boundaries defining each category (in operations research, this problem is called nominal sorting). The two categories are defined by a set of prototypes, which correspond to already diagnosed cardiac beats. Pairwise comparison between each beat and every prototype beat is applied for the calculation of their similarity according to [29].

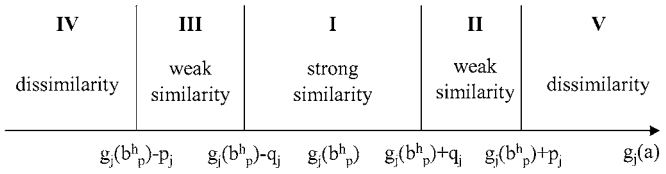

Fig. 1. Five-zone similarity system.

In order to use a flexible model and to avoid strict thresholding, pairwise comparison takes place with the use of a five-zone similarity system (see Fig. 1). This allows for intermediate evaluations of similarity and dissimilarity. The intermediate zone of weak similarity is used to express the ambiguity of whether the two patterns are similar or dissimilar.

More analytically, let A be the finite set of beats, F the set of n criteria (features), with $n \geq 1$, w_j the weight of each criterion, $\sum_j w_j = 1$, $C = \{C^1, \ldots, C^K\}$ be the set of the categories $(K > 1)$, and $B^h = \{b^h_p \,|\, 1, \ldots, L^h \text{ and } h = 1, \ldots, K\}$ the set of reference patterns (prototypes) of the category C^h, where b^h_p represents the p prototype of the category C^h and L^h the number of the prototypes of this category. Each beat in A and B is characterized by a feature vector $\bar{g}$ containing its feature values for all n criteria of F (i.e., $\forall a \in A, g(a) = (g_1(a), \ldots, g_n(a))$ and $\forall b^h_p \in B, g(b^h_p) = (g_1(b^h_p), \ldots, g_n(b^h_p))$). Every beat a is compared to each prototype beat b^h_p under each criterion j. A similarity index $(\mathrm{SI}_j(a, b^h_p))$ is computed for every criterion. In order to model the five-zone similarity system, the following two thresholds must be specified.

1) A *similarity threshold* q_j,[1] representing the maximum allowed difference $|g_j(a) - g_j(b^h_p)|$ between a and b^h_p, so as these beats to be judged similar under the jth criterion (region I, Fig. 1).
2) A *dissimilarity threshold* p_j, representing the minimum allowed difference between beats a and b^h_p, so that these beats are to be considered totally dissimilar under the jth criterion (regions IV and V, Fig. 1).

The similarity index $\mathrm{SI}_j(a, b^h_p)$ is then computed according to (1), shown at the bottom of the page (Fig. 2 indicates the different values of a similarity index). By aggregating the similarity indexes, we compute a concordance index $\mathrm{CI}(0 \leq \mathrm{CI} \leq 1)$

[1]Threshold values could vary according to the category in question or even the prototype with which each pattern is compared, having $q_j = q_j(b^h_p), p_j = p_j(b^h_p)$. In our application, we followed the simplified approach of constant values, which are easier to be provided manually or even computed automatically.

$$\mathrm{SI}_j\left(a, b^h_p\right) = \begin{cases} 1, & \left|g_j(a) - g_j\left(b^h_p\right)\right| \leq q_j \\ \dfrac{\left|g_j(a) - g_j\left(b^h_p\right)\right| - p_j}{q_j - p_j}, & q_j < \left|g_j(a) - g_j\left(b^h_p\right)\right| < p_j \\ 0, & \left|g_j(a) - g_j\left(b^h_p\right)\right| \geq p_j \end{cases} \quad (1)$$

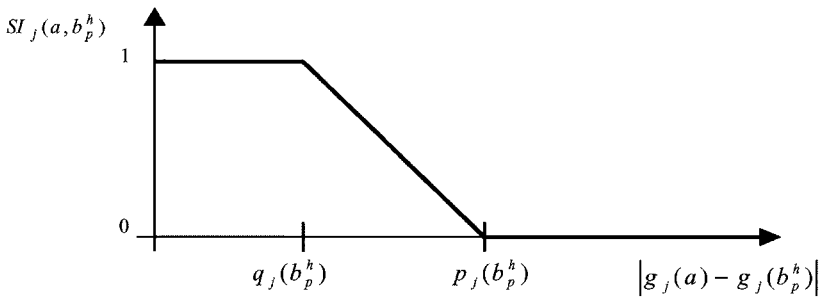

Fig. 2. SI values.

which indicates the concordance to the hypothesis where beat a is similar to beat b_p^h as follows:

$$\mathrm{CI}\left(a, b_p^h\right) = \sum_j w_j \mathrm{SI}_j(a, b_p^h). \tag{2}$$

Having calculated the CI for all of the prototypes of category C^h, we can compute the membership degree of the beat a to category h as follows:

$$d(a, C^h) = \max\left\{\mathrm{CI}\left(a, b_1^h\right), \mathrm{CI}\left(a, b_2^h\right), \ldots, \mathrm{CI}\left(a, b_{Lh}^h\right)\right\}. \tag{3}$$

Finally, using the above membership degree, we can classify a beat to category $C(a)$ with the maximum membership degree as follows:[2]

$$C(a) = \arg\max_h d(a, C^h). \tag{4}$$

The pseudocode for the multicriteria algorithm is presented below:

```
get thresholds
get weights
for a=1 to ⟨Number_of_beats⟩
  for h=1 to ⟨Number_of_categories⟩
     for p=1 to ⟨
        Number_of_prototypes_in_category_h⟩
          for j=1 to ⟨Number_of_criteria⟩
           calculate SIj(a, bhp)
         end
     calculate CI(a, bhp)
     end
  calculate d(a, Ch)
  end
  classify beat (a)
end
```

The main difficulty encountered when applying the proposed multicriteria method is the specification of weights w_j and thresholds q_j and p_j. These parameter values can be defined using medical knowledge according to the doctor's value system. However, sensitivity analysis has revealed that there could be sets of parameters providing better performance compared to that achieved when the parameters are set by medical experts [30]. Aiming at automated beat classification, we incorporated GAs for the adjustment of the parameters of the multicriteria method.

[2] $\arg\max_h f(h)$ provides the h that maximizes $f(h)$.

III. Using GAs for Parameter Estimation

GAs, first developed by Holland [31], are general-purpose search algorithms that use principles inspired by natural population genetics to evolve solutions to problems. GAs have been successfully applied in a variety of problems, including routing and scheduling, design optimization, curve fitting, and machine learning. The basic idea of a GA is to maintain a population of knowledge structures (called chromosomes), each one representing a candidate solution to the problem. An initial population is first created, most of the time in a random way (step 1). This population evolves over time through competition and controlled variation with the application of genetic operators. In an iterative process (see Fig. 3), each chromosome is evaluated according to the quality of the solution that it represents (according to a fitness function) (step 2). Chromosomes are selected according to their fitness by the selection operator (step 3) and are combined (mated) in order to produce new chromosomes (by the crossover operator), hopefully combining the "good characteristics" of the parent chromosomes (step 4). Some alterations in chromosomes are allowed (by the mutation operator) in order to ensure that all parts of the search space will be reached. The whole process ends, if either a certain fitness value is achieved or a maximum number of iterations is reached [32].

GAs can be used in parameter optimization problems by encoding a set of k parameters in a chromosome of the population. In that case, each chromosome represents a possible set of parameters of the system to be optimized as follows:

$$\text{chromosome} = \{x_1, x_2, x_3, \ldots, x_{k-1}, x_k\}.$$

The chromosome with the highest fitness, as found by the GA provides the optimum set of parameters.

The structure of a chromosome depends on the type of encoding that is used. Binary, real, "gray," and other encodings have been used, with the binary encoding being closer to the theory of Holland and better investigated. In this case, each real parameter is mapped into a binary number (substring). Since in this way a continuous variable is mapped into a discrete variable, the length of the binary string is defined according to the desired precision.

As far as the genetic operators are concerned, the roulette wheel selection and the two-point crossover have been selected.

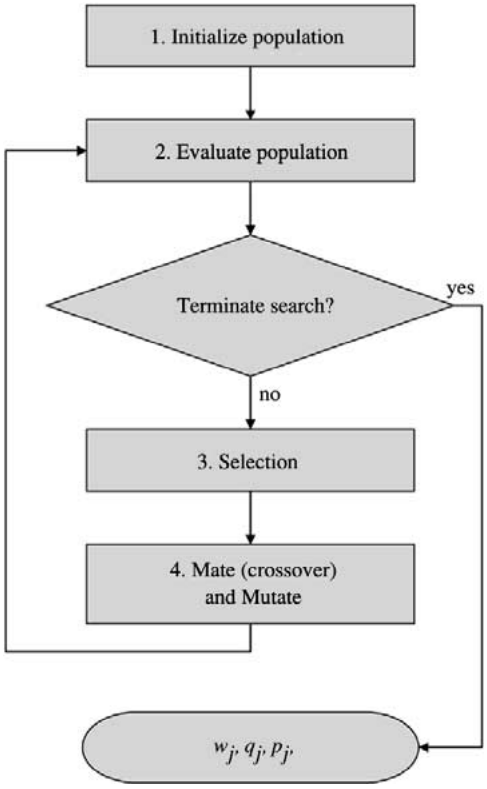

Fig. 3. A simple GA.

- In *roulette wheel selection* (or fitness-proportional selection), the probability of selection is proportional to individual's fitness, i.e., the probability for selecting the ith chromosome is $f_i / \sum_{z=1}^{N} f_z$, where N is the population size.
- In *two-point crossover*, two sites are randomly chosen in both chromosomes. The middle part of the chromosomes, between these two points, is then swapped (see Fig. 4).

In order to avoid premature convergence, two more advanced genetic operators are also applied, linear fitness scaling and elitism.

- In *linear fitness scaling*, the fitness of the population members is pivoted about the average population fitness. Scaling has a dual effect in convergence: in the first steps, scaling ensures that the well-performing individuals will not dominate the selection process, while in the last steps, when the population is largely converged, it ensures that best performers will be rewarded, since small differences will become meaningful.
- In *elitism*, the best individual of a population is always kept in the next generation. In this way, the best solution is never lost during the search.

IV. ECG Beat Classification

In order to classify each cardiac beat as ischemic or normal, all of the relevant ECG features (isoelectric line, ST segment, and T wave) need to be defined and measured. The automated feature extraction procedure starts with the QRS complex detection, which is realized with an already reported algorithm that has been proven to be very effective [33]. According to this algorithm, the ECG signal is bandpassed, differentiated, squared, and integrated. In the transformed signal, a set of rules is applied to detect the QRS complexes. To make this algorithm faster (30% approximately), we propose a modification where the 300-ms interval that follows each detected QRS is discarded from the searching procedure [9]. Once the QRS for each beat has been recognized, the isoelectric line and the ST segment are detected next, using an edge detection algorithm [34]. Due to the presence of noise in the ECG recordings (such as power line interference, electromyographic contamination, and baseline wandering), a filtering module is applied [9], [34] to remove baseline wandering effects and to accurately detect the isoelectric line and the J point (start of the ST segment). Briefly, the technique implements a least square procedure and models the baseline wandering, which is a slow type of noise, with a low-order polynomial. Subtracting this polynomial from the recorded signal results in correction of the ECG baseline without altering the ECG waveform patterns. As for the other two types of noise, A/C interference and electromyographic noise, these are handled properly by applying a 20-ms moving averaging window. It should be clarified that the averaging technique is used as a rule of thumb for the definition of the isoelectric line and the J point and not for removing the noise from the ECG signal.

Concerning the T wave, only its peak is needed and is computed as the point with the maximum difference in amplitude with respect to the J80 point or the J60 point in cases where the heart rate is faster than 120 beats/min (J80 and J60 are the points that lie 80 ms or 60 ms, respectively, after the J point). It should be noted that cardiac beats containing a substantial amount of noise (when the J point cannot be defined) are rejected at this stage.

For proper adjustment and evaluation of the proposed classification system, a task-specific ECG database was developed based on the recordings of the European Society of Cardiology ST-T (ESC ST-T) database [36]. This was developed especially for the evaluation of algorithms designed to analyze ST and T-wave changes. It consists of 90 continuous two-channel recordings, two hours each, taken from 79 different ambulatory ECGs. Eleven hours of continuous two-channel recordings were extracted from 10 representative files of the ESC ST-T database (whole e0104 recording and the first hour of the e0103, e0105, e0108, e0113, e0114, e0147, e0159, e0162, and e0206 recordings). The selected ECG signals contain 20 ischemic ST segment episodes and 20 ischemic T-wave episodes. The excerpts were diagnosed in beat-by-beat mode from three experienced cardiologists. This accumulated experience yielded a dataset of 86 384 cardiac beats characterized as normal, ischemic, or artifact. Thus, each record in the database contains the values of the criteria used by the MCDA algorithm and the classification of the corresponding beat. From the initial number of beats, those that are not detected by the QRS detector [33] as well as the artifacts (6 754 beats) were rejected, leaving a total of 76 989 beats. The 37 663 beats (48.92%) from these were characterized as ischemic and the rest as normal. The training set for the multicriteria method (i.e., the set of prototypes) was constructed after selecting the first beat out of

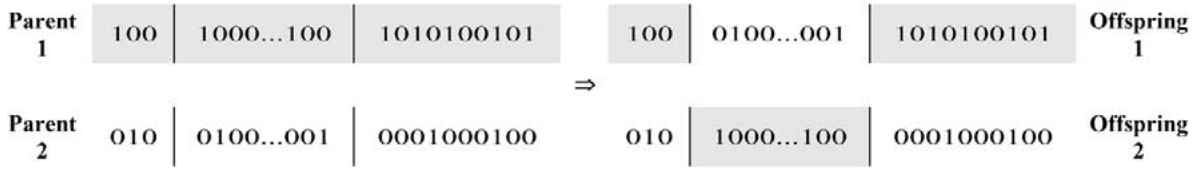

Fig. 4. Two-point crossover function.

a sequence of 40 (2.5% of the final dataset). This training set contained 954 ischemic beats and 982 normal, leaving 75 053 beats for training the GA and testing the overall system. The large number of prototypes used (1 936) ensured a consistent training. Approximately one third of these beats (i.e., 25 000) was selected randomly for training the GA while the remaining 50 053 were used as a test set.

The multicriteria decision analysis algorithm uses five criteria:

C1: the ST segment deviation;
C2: the ST segment slope;
C3: the T-wave amplitude;
C4: the T-wave normal amplitude and polarity;
C5: the patient's age.

The ST deviation is measured at the J80 (or J60) point and refers to the amplitude deviation of the ST segment from the isoelectric line, which is the line that defines the zero amplitude level. The ST segment slope is defined as the slope of the line connecting the J and J80 (or J60) points. The T-wave amplitude is defined as the amplitude deviation of the T-wave peak from the isoelectric line (Fig. 5). Finally, the T-wave normal amplitude and polarity expresses the amplitude and polarity of normal beats for a specific ECG recording. It is estimated using the first 30 s of each recording and is computed by the mean value of the T-wave amplitudes at this ECG interval.

As already noted, the three types of parameters of the multicriteria method that must be specified through the application of the GA were the similarity thresholds q_j, the dissimilarity thresholds p_j, and the criterion weights w_j for all n criteria $(j = 1 \ldots n)$. All 15 parameters were encoded in the chromosome as follows:

$$\text{chromosome} = \{q_1, \ldots, q_5, p_1, \ldots, p_5, w_1, \ldots, w_5\}.$$

In order to limit the search space only to reasonable values, the lower and upper bounds for each parameter were set in collaboration with medical experts. The desired precision e (also set by medical experts) was used to specify the minimum number l of digits (length) for each parameter (Table I). Consequently, a binary chromosome[3] of at least 81 bits should be used for the adequate representation of the MCDA parameters.

The encoding ensures that all values produced by the GAs are positive. In order to ensure that $p_j \geq q_j$, a transformation $p_j = a_j + q_j$ was applied and the a_j values were specified by the GA instead of p_j. Moreover, weight values were normalized before used in the MCDA model to ensure that $\sum_j w_j = 1$. Aiming at

[3]The formula used for decoding (from a binary encoding A to a real parameter c) was $c = L + (A/2^l - 1) \times (U - L)$, where c is the real parameter value, U and L are the upper and the lower bounds of the parameter, A is the binary number in decimal form, and l is the number of digits (length) of this part of the chromosome.

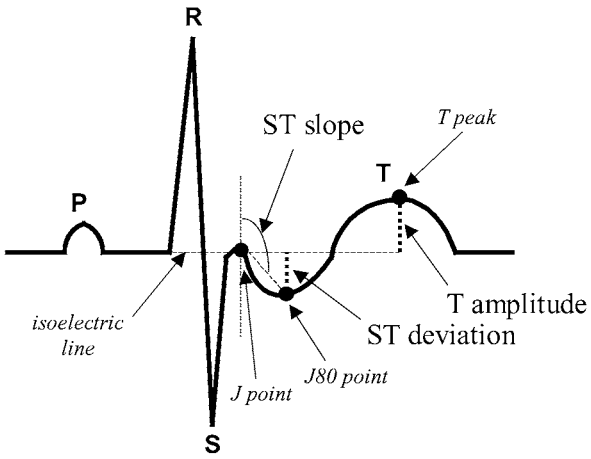

Fig. 5. ECG features of a cardiac beat used by the MCDA algorithm.

maximizing both the sensitivity (Se) and the specificity (Sp)[4] provided by the multicriteria algorithm, the following fitness function was employed:

$$f = \sqrt{\left(\frac{\text{Se} + \text{Sp}}{2}\right)^2 - (\text{Se} - \text{Sp})^2}. \tag{5}$$

The first part of the fitness function ensures the maximization of both Se and Sp, while the second part ensures that *both* metrics will be maximized, forcing a penalty in cases when only one metric is improved.

Following a series of experiments, the values of the parameters applied were the following: the population size was set to 30, the maximum number of generations to 400, $p_c = 0.95$ (crossover probability), and $p_m = 0.02$ (mutation probability). These values are also proposed in [37]. Fig. 6 presents an overview of the automated classification procedure.

V. Results

Table II summarizes the set of parameters as discovered by the genetic algorithm. The corresponding fitness value was $f = 8295.5$[5] corresponding to $\text{Se} = 91.4\%$ and $\text{Sp} = 90.8\%$ for the GA training set. Using the parameters of Table II and applying the multicriteria method for the test set data resulted in high values of both sensitivity (91.2%) and specificity (90.9%). As seen in Table III, out of the 50 053 beats, the developed classifier misses 2 151 ischemic and 2 317 normal beats. In Table IV the results of the proposed beat classification system are compared to those of other approaches to beat classification, such as

[4]Se measures the ability to detect the ischemic beats while Sp measures the ability to detect the normal beats.

[5]In the hypothetical case where $\text{Se} = \text{Sp} = 100\%$, 10 000 would be maximum fitness value.

TABLE I
SEARCH SPACE AND CODING LENGTHS

Parameter	q				p				w			
Criterion	lower	upper	e	l	lower	upper	e	l	lower	upper	e	l
C1.	0	0.01	0.001	4	0	0.02	0.001	5	0.04	0.4	0.01	6
C2.	0	10	1	4	0	10	1	4	0.04	0.4	0.01	6
C3.	0	0.025	0.001	5	0	0.2	0.001	8	0.04	0.4	0.01	6
C4.	0	0.025	0.001	5	0	0.2	0.001	8	0.04	0.4	0.01	6
C5.	0	10	1	4	0	15	1	4	0.04	0.4	0.01	6

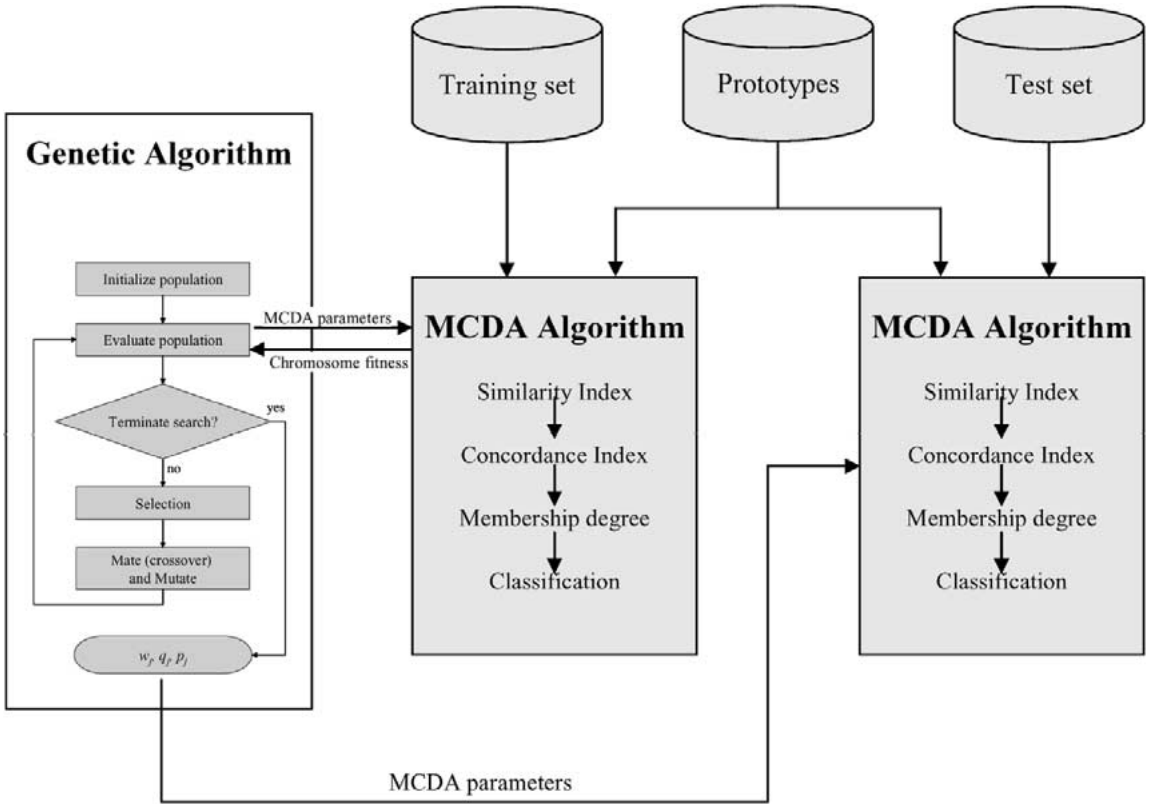

Fig. 6. Integrated classifier.

TABLE II
q_j, p_j, w_j THRESHOLDS DERIVED BY THE GA

Criterion	q	p	w
C1. ST segment deviation	0.0080	0.0190	0.1991
C2. ST segment slope	9.3333	18.6777	0.0597
C3. T wave amplitude – polarity	0.0000	0.1906	0.1792
C4. T wave normal amplitude - polarity	0.0081	0.0253	0.4181
C5. Patient's age	3.3333	5.3333	0.1438

a combination of an auto-associative nonlinear artificial neural network (ANN) that performs nonlinear principal component analysis (PCA) with a radial basis function network [11], set of rules [9], PCA and feed-forward neural networks [15], and parametric modeling combined with feed-forward neural networks [7]. All of the methods mentioned in Table IV are automated

TABLE III
CLASSIFICATION RESULTS OF THE PROPOSED METHOD

		GA & MCDA classifier	
		Ischemic	Normal
Cardiac beat database	Ischemic	22,306	2,151
	Normal	2,317	23,279

TABLE IV
PERFORMANCE MEASURES OF ISCHEMIC BEAT CLASSIFICATION SYSTEMS

System Description	Se[1] (%)	Sp[2] (%)
ANN & nonlinear PCA [11]	79	75
Rule-based [9]	70	63
ANN & PCA [15]	90	90
ANN & parametric modeling [7]	81	84
Current work	**91**	**91**

[1] Se: Sensitivity

[2] Sp: Specificity

and can operate in real time. However, their computational effort cannot be compared directly since it is either not reported in all of the works or it has been estimated in different computer systems. In our classification system, approximately 10 ms are needed for processing 1 s for each recorded lead (using an Intel Pentium IV at 1.5 GHz and 512 MB RAM). In addition to the above, it should be mentioned that other beat classification systems have also been proposed, but their performance cannot be judged against the above mentioned methods, since either they have been evaluated with other test sets [6], [8] or they employed different performance measures [6], [13], [14].

VI. DISCUSSION

In the current study, a multicriteria approach was developed for the classification of recorded cardiac beats in long duration ECGs. All steps in the proposed method (filtering, feature extraction, training, classification) are realized automatically. In our method, we define each category by a set of reference patterns. The fuzzy modelization eliminates the problem of strict thresholding (and the need for physical explanation), while the multicriteria approach employs a comparison between alternative beats through the use of values in different criteria. The application of thresholds and the pseudocriterion modelization provides a more realistic similarity system, since it is not realistic to set a strict value under which there is similarity between two beats and over which strict dissimilarity.

The parameters of the multicriteria method are defined by a GA. Compared to the results obtained in test cases with manual entry of the model parameters (described in [30]), the automated method performs better. Moreover, its major advantage is that it eliminates the need for the definition of parameter values from medical experts, which according to our experience is a rather hard task. Nevertheless, the application of a GA requires the definition of various parameters such as population size, selection type, crossover type and rate, and mutation rate. There has long been discussion and research on the (automated) definition of these parameters forming a meta-decision problem (a review can be found in [38]). Some researchers argue that this defines a separate problem that could be solved by a meta-genetic algorithm (a GA to optimize the GA parameters) [37], [39]. Two sets of parameters, one proposed in [40] and the other in [37], seem to work adequately well in most cases. In our application, following a series of tests with different parameters, the second one was chosen, since it provided better results. In general, the multicriteria method has been proven to be not very sensitive in changes of its parameters [30].

Having a more in-depth view of the GA application and following a long series of runs, we can claim that elitism and scaling provided better results. Due to the length of the chromosome, two-point crossover outperformed one-point crossover (in one-point crossover there is only one random site where the exchange of chromosome parts is done). The fact that it also outperformed uniform crossover (in uniform crossover, each bit is selected from the corresponding bit values of one of the two parents, with equal probability) indicates that the latter destroyed the parts of the chromosomes contributing to a high fitness value (known as building blocks [32]). Gray encoding (an alteration of binary encoding that ensures that any pair of adjacent points in the problem space differs only by a single bit in the representation space) did not seem to add to the performance of the GA. Both scaling and elitism seem to significantly improve GA performance.

The obtained results are better than those of other similar approaches [7], [9], [11], [15], [30] in terms of both sensitivity and specificity. All of the above methods were tested using data from the ESC ST-T database, which is used as a standard reference for myocardial ischemia. It should be noted that in [11] a different subset from the ESC ST-T database was employed to evaluate the beat classifier. Specifically, it was considered that every annotated episode in the database contained only ischemic beats. Furthermore, the multicriteria method could, if necessary, provide an interpretation of the results, following a postprocessing [identification of the closest prototype(s)]. This cannot be achieved easily in approaches based on ANNs that are considered, most commonly, as a "black box" and need tedious postprocessing to interpret the classification. This is of great importance when designing a device for medical decision support and will assist doctors to reach a diagnosis faster and safer.

The method's performance can be further improved by a more accurate definition of the employed ECG features. The presence of noise in the ECG recordings of the ESC ST-T database causes poor detection of the ST segment, the isoelectric line, and the T wave. However, modern ECG recorders provide less noisy signals than those included in the ESC ST-T database, and, consequently, the proposed method is expected to perform better. Furthermore, in order for the described methodology to be used in clinical practice, the developed cardiac beat database should be extended and additional types of ischemic and normal ECG waveform patterns must be included. This can be implemented

easily in the MCDA approach since only the beats used as prototypes need to be modified, which is a simple procedure. Additional (not provided by the employed database) risk factors often taken into consideration in clinical examination (e.g., LDL cholesterol) can be included in a real-life application in the form of new criteria. Moreover, our system could be adapted to address other cardiac abnormalities, such as arrhythmias. As far as the GA parameters are concerned, future research could include a more advanced, self-adapted GA that modifies parameters of the GA in real time to the ongoing search. However, based on our experiments, we believe that such advanced techniques will not significantly improve the GA performance.

VII. Conclusion

We presented a novel technique for detecting the ischemic cardiac patterns in the ECG signal based on MCDA and GA. The application of the MCDA algorithm was successful and modeled suitably the dynamic characteristics of myocardial ischemia. On the other hand, the employment of a GA assisted in automatically defining the values of the parameters (weights and thresholds) used by the multicriteria method. Furthermore, the developed cardiac beat database contributed significantly to the proper training and testing of the whole system.

The proposed methodology classified the cardiac beats sufficiently well, indicating that it could be part of an integrated system for ischemia diagnosis. Still, further testing is needed using data extracted from real clinical conditions. The overall ischemic episode detection system should be tried and assessed with several ECGs extracted from ambulatory recordings or continuous monitoring in the coronary care unit. Such testing will fully reveal the potential of the proposed ischemic beat classifier.

References

[1] R. Silipo, A. Taddei, and C. Marchesi, "Continuous monitoring and detection of ST-T changes in ischemic patients," in *Proc. IEEE Comput. Cardiol.*, 1994, pp. 225–228.

[2] M. Emdin, A. Taddei, M. Varanini, M. Raciti, S. Pola, C. Marchesi, and A. L'Abbate, "Electrocardiographic and signal monitoring in ischemic heart disease: State of the art and perspective," *J. Med. Eng. Technol.*, vol. 21, pp. 162–165, 1997.

[3] L. T. Proctor and C. P. Kingsley, "Con: ST-segment analysis-who needs it?," *J. Cardiothorac. Vasc. Anesth.*, vol. 10, pp. 681–682, 1996.

[4] D. J. Rowlands D. J., *Understanding the Electrocardiogram (Section 2: Morphological Abnormalities)*. Cheshire, U.K.: Imperial Chemical, 1982.

[5] M. J. Goldman, *Principles of Clinical Electrocardiography*, 11th ed. Los Altos, CA: LANGE Medical, 1982.

[6] I. Pitas, M. G. Strintzis, S. Grippas, and C. Xerostylides, "Machine classification of ischemic electrocardiograms," presented at the *IEEE Mediterranean Electrotechnical Conf.*, Athens, Greece, 1983.

[7] C. Papaloukas, D. I. Fotiadis, A. Likas, and L. K. Michalis, "An expert system for ischemia detection based on parametric modeling and artificial neural networks," in *Proc. Eur. Med. Biol. Eng. Conf.*, 2002, pp. 742–743.

[8] L. Senhadji, G. Carrault, J. J. Bellanger, and G. Passariello, "Comparing wavelet transforms for recognizing cardiac patterns," *IEEE Eng. Med. Biol. Mag.*, vol. 14, pp. 167–173, Mar./Apr. 1995.

[9] C. Papaloukas, D. I. Fotiadis, A. Likas, A. P. Liavas, and L. K. Michalis, "A knowledge-based technique for automated detection of ischemic episodes in long duration electrocardiograms," *Med. Biol. Eng. Comput.*, vol. 39, pp. 105–112, 2001.

[10] C. Papaloukas, D. I. Fotiadis, A. Likas, C. S. Stroumbis, and L. K. Michalis, "Use of a novel rule-based expert system in the detection of changes in the ST segment and the T wave in long duration ECGs," *J. Electrocardiol.*, vol. 35, pp. 27–34, 2002.

[11] T. Stamkopoulos, K. Diamantaras, N. Maglaveras, and M. Strintzis, "ECG analysis using nonlinear PCA neural networks for ischemia beat detection," *IEEE Trans. Signal Processing*, vol. 46, pp. 3058–3067, Nov. 1998.

[12] N. Maglaveras, T. Stamkopoulos, K. Diamantaras, C. Pappas, and M. Strintzis, "ECG pattern recognition and classification using nonlinear transformations and neural networks: A review," *Int. J. Med. Inform.*, vol. 52, pp. 191–208, 1998.

[13] N. Maglaveras, T. Stamkopoulos, C. Pappas, and M. Strintzis, "ECG processing techniques based on neural networks and bidirectional associative memories," *J. Med. Eng. Technol.*, vol. 22, pp. 106–111, 1998.

[14] S. Papadimitriou, S. Mavroudi, L. Vladutu, and A. Bezerianos, "Ischemia detection with a self-organizing map supplemented by supervised learning," *IEEE Trans. Neural Networks*, vol. 12, pp. 503–515, May 2001.

[15] C. Papaloukas, D. I. Fotiadis, A. Likas, and L. K. Michalis, "An ischemia detection method based on artificial neural networks," *Artif. Intell. Med.*, vol. 24, pp. 167–178, 2002.

[16] B. Roy, *Multicriteria Methodology for Decision Aiding*. Dodrecht, The Netherlands: Kluwer, 1996.

[17] Ph. Vinke, *Multicriteria Decision Aid*. New York: Wiley, 1992.

[18] B. Roy, *Méthodologie Multicritère d' Aide à la Decision*. Paris, France: Economica, 1985.

[19] B. Roy and B. Hugonard, "Ranking of suburban line extension projects on the Paris metro system by a multicriteria method," *Transport. Res. Rec.*, vol. 16, pp. 301–312, 1982.

[20] Y. Goletsis, J. Psarras, and J.-E. Samouilidis, "Project ranking in the Armenian energy sector using a multicriteria method for groups," *Ann. Oper. Res.*, vol. 120, pp. 135–157, 2003.

[21] R. Slowinski and C. Zopounidis, "Application of the rough set approach to evaluation of bankruptcy risk," *Int. J. Intell. Syst. Account. Financ. Manag.*, vol. 4, pp. 27–41, 1995.

[22] C. Zopounidis and M. Doumpos, "Multicriteria classification and sorting methods: A literature review," *Eur. J. Oper. Res.*, vol. 138, pp. 229–246, 2002.

[23] Ph. Du Bois, J. P. Brans, F. Cantraine, and B. Mareschal, "MEDICIS: An expert system for computer-aided diagnosis using the PROMETHEE multicriteria method," *Eur. J. Oper. Res.*, vol. 39, pp. 268–273, 1989.

[24] N. Belacel, Ph. Vincke, J. M. Scheiff, and M. R. Boulassel, "Acute leukaemia diagnosis aid using multicriteria fuzzy assignment methodology," *Comp. Meth. Progr. Biomed.*, vol. 64, pp. 145–151, 2001.

[25] C. A. Pena-Reyes and M. Sipper, "Evolutionary computation in medicine: An overview," *Artif. Intell. Med.*, vol. 19, pp. 1–23, 2000.

[26] O. Cordón, F. Gomide, F. Herrera, F. Hoffman, and L. Magdalena, "Ten years of genetic fuzzy systems: Current framework and new trends," *Fuzzy Sets Syst.*, vol. 141, pp. 5–31, 2004.

[27] C. A. Coello Coello, "Handling preferences in evolutionary multiobjective optimization: A survey," in *Proc. 2000 Congr. Evolutionary Computation*, 2000, pp. 30–37.

[28] Z. Dokur and T. Olmez, "ECG beat classification by a novel hybrid neural network," *Comp. Meth. Progr. Biomed.*, vol. 66, pp. 167–181, 2001.

[29] J. Léger and J.-M. Martel, "A multicriteria assignement procedure for a nominal sorting problematic," *Eur. J. Oper. Res.*, vol. 138, pp. 349–364, 2002.

[30] Y. Goletsis, C. Papaloukas, D. I. Fotiadis, A. Likas, and L. K. Michalis, "A multicriteria decision based approach for ischemia detection in long duration ECGs," in *Proc. IEEE EMBS 4th Int. Conf. Information Technology Applications in Biomedicine (ITAB 2003)*, 2003, pp. 230–233.

[31] J. M. Holland, *Adaptation in Natural and Artificial Systems*. Cambridge, MA: MIT Press, 1975.

[32] D. E. Goldberg D. E., *Genetic Algorithms in Search, Optimization and Machine Learning*. Reading, MA: Addison-Wesley, 1989.

[33] W. J. Tompkins, *Biomedical Digital Signal Processing (C-Language Examples and Laboratory Experiments for the IBM® PC)*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[34] I. K. Daskalov, I. A. Dotsinsky, and I. I. Christov, "Developments in ECG acquisition, preprocessing, parameter measurement, and recording," *IEEE Eng. Med. Biol. Mag.*, vol. 17, pp. 50–58, Jan./Feb. 1998.

[35] C. Papaloukas, D. I. Fotiadis, A. Likas, A. P. Liavas, and L. K. Michalis, "A robust knowledge-based technique for ischemia detection in noisy ECGs," in *Proc. 4th Int. Conf. Knowledge-Based Intelligent Engineering Systems & Allied Technologies (KES2000)*, 2000, pp. 768–771.

[36] "European ST-T Database Directory," Eur. Soc. Cardiol., S.T.A.R., Pisa, Italy, 1991.
[37] J. J. Grefenstette, "Optimazation of control parameters for genetic algorithms," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-16, pp. 122–128, Jan./Feb. 1986.
[38] M. Mitchel, *An Introduction to Genetic Algorithms*. London, U.K.: MIT Press, 1996.
[39] M. F. Bramlette, "Initialization, mutation and selection methods in genetic algorithms," in *Proc. 4th Int. Conf. Genetic Algorithms and Their Applications*, 1991, pp. 100–107.
[40] K. A. DeJong and W. M. Spears, "An analysis of the interacting roles of population size and crossover in genetic algorithms," in *Proc. 1st Workshop Parallel Problem Solving from Nature*, 1990, pp. 38–47.

Yorgos Goletsis (M'03) was born in Ioannina, Greece, in 1972. He received the Diploma degree in electrical engineering and the Ph.D. degree in electrical and computer engineering both from the National Technical University of Athens, Athens, Greece.

Since 2002, he has been with the Department of Computer Science, University of Ioannina, Ioannina, as a Senior Researcher. Currently, he is also a Visiting Professor with the Department of Economics, University of Ioannina. His research interests include operational research, decision support systems, and evolutionary computation.

Costas Papaloukas was born in Ioannina, Greece, in 1974. He received the diploma degree in computer science and the Ph.D. degree in biomedical technology from the University of Ioannina, Ioannina, Greece, in 1997 and 2001, respectively.

He is a Lecturer of Bioinformatics with the Department of Biological Applications and Technology, University of Ioannina. His research interests include biomedical engineering and bioinformatics.

Dimitrios I. Fotiadis (M'01) was born in Ioannina, Greece, in 1961. He received the Diploma degree in chemical engineering from National Technical University of Athens, Athens, Greece, and the Ph.D. degree in chemical engineering from the University of Minnesota, Minneapolis.

Since 1995, he has been with the Department of Computer Science, University of Ioannina, Ioannina, Greece, where he currently is an Associate Professor. He is the Director of the Unit of Medical Technology and Intelligent Information Systems. His research interests include biomedical technology, biomechanics, scientific computing, and intelligent information systems.

Aristidis C. Likas (S'91–M'96–SM'03) was born in Athens, Greece, in 1968. He received the Diploma degree in electrical engineering and the Ph.D. degree in electrical and computer engineering both from the National Technical University of Athens, Athens, Greece.

Since 1996, he has been with the Department of Computer Science, University of Ioannina, Ioannina, Greece, where he is currently an Assistant Professor. His research interests include neural networks, machine learning, pattern recognition, and intelligent systems for biomedical engineering.

Lampros K. Michalis was born in Arta, Greece, in 1960. He graduated with distinction from the Medical School, University of Athens, Athens, Greece in 1984. He received the M.D. degree (with distinction) from Athens Medical School, Athens, Greece, in 1989.

Since 1995, he has been with the Medical School, University of Ioannina, Ioannina, Greece, where he is currently an Associate Professor of Cardiology. He is in charge of the Coronary Care Unit and the Catheter Laboratory, University Hospital, Medical School, University of Ioannina. His research interests focus on bioengineering and interventional cardiology.