# Investigating electrocardiographic features in fuzzy models for cardiac arrhythmia classification

Rosaria Silipo[1]

International Computer Science Institute, Berkeley, USA

**Abstract.** Simple and composed measures of the electrocardiographic (ECG) waveshape are generally used to detect and classify different kinds of cardiac arrhythmic beats. In the last years, the development of computer technique allowed an increasing number of electrocardiographic measurements to be extracted and stored for each cardiac beat. As a consequence, analysis systems of any kind result to be more complex and the interpretation of the decision process less transparent. This happens for traditional "black-box" systems, such as neural networks, as well as for techniques supposed to be easy to interpret, such as fuzzy systems. A method is presented that analyzes high-dimensional fuzzy classifiers "a posteriori", in order to give insights about the decision process and possibly to reduce the input space dimension. This method is applied to an arrhythmia classification task, considering twelve ECG measures for each beat and three arrhythmic classes on the MIT-BIH database records. A possibilistic information gain is defined for every ECG measure with respect to the adopted fuzzy model. Such information gain supplies a measure of the discriminative power of the corresponding ECG feature and offers insights about its impact on the decision process.

## 1 Introduction

In the last years it has become more and more common to collect and store large amounts of data from different sources. Also in the automatic analysis of the electrocardiographic (ECG) signal, the parameters extracted from each ECG beat waveshape are growing dramatically, including more simple ECG measures and more of their linear and nonlinear combinations. An automatic classification system [1] is usually applied, to detect similarities among the ECG beats and/or to cluster the input space, with performances depending on the adopted method.

The high dimension of the input space, however, can make the task of the classifier very complicated, even more if such a high input dimension includes uninformative or unreliable measures. In addition some input features can be redundant, because the carried information heavily overlap with that of other input features. The structure of the resulting classifier will be then very complicated and the corresponding decision process not easy to interpret.

Some data analysis techniques are complicated "per se ", such as neural networks, because they do not implement a transparent decision process to the user.

Other data analysis algorithms are by now famous for their easy interpretability, such as fuzzy systems [2]. On the other hand, the high dimension of the input space and the strong fragmentation of the output clusters can still produce a quite complex set of fuzzy rules. The easy interpretability of the fuzzy models may fail. Thus automatic tools for the analysis of fuzzy systems become necessary, as a support for the following interpretation performed by human experts.

A possibilistic information gain, introduced in [3], is used to characterize the discriminability power of the input features in a fuzzy model, that implements a cardiac arrhythmia classification task. Such information gain indirectly measures the impact of each ECG feature on the decision process, constructed with a given automatic algorithm on a particular training set. This analysis of the implemented fuzzy model produces useful insights on the effectiveness of the ECG features in terms of arrhythmia characterization and gives hints for dimension reduction.

## 2 The problem of cardiac arrhythmia classification

A very suitable area for fuzzy decision systems is represented by medical data. Medical reasoning is quite often a qualitative and approximative process, so that the definition of precise diagnostic classes with crisp membership functions can sometimes lead to inappropriate conclusions. The intra-patient variability as well as the amount of uncertainty in the symptoms description quite often require only a qualitative reasoning, where exceptions benefit of individual decision rules. One of the most investigated fields in medical reasoning is the automatic analysis of the electrocardiogram (ECG), and inside that the arrhythmia classification.

Some cells (the sino-atrial node) of the upper chambers (the atria) in the cardiac muscle (the myocardium) spontaneously and periodically change their polarization, which progressively extends to the whole cardiac muscle. This periodic and progressive electric depolarization of the myocardium is recorded as small potential differences between two different locations of the human body or with respect to a reference electrode. An almost periodic electrical signal, the ECG, that describes the electrical activity of the myocardium in time, is the result. Each time period, varying between 700 and 1000ms ca in healthy subjects, consists of a basic waveshape, where some time components or waves are easily recognizable and generally marked with alphabet letters: P, Q, R, S, T (Fig. 1).

The P wave describes the depolarization process of the two upper myocardium chambers, the atria; the QRS complex all together the depolarization of the two lower myocardium chambers, the ventricula; and the T wave the repolarization process at the end of each cycle. The muscle contraction follows the myocardium depolarization phase. Anomalies in the PQRST waveshape are often connected to a misfunction of the electrical impulse conduction on the myocardium.

A big family of electrical misfunctions of the myocardium consists of arrhythmia. Arrhythmic heart beats generally refer to anomalous (ectopic) origin of the depolarization in the myocardium. If the depolarization does not originate in the sino-atrial node, a different path is followed by the depolarizing wavefront

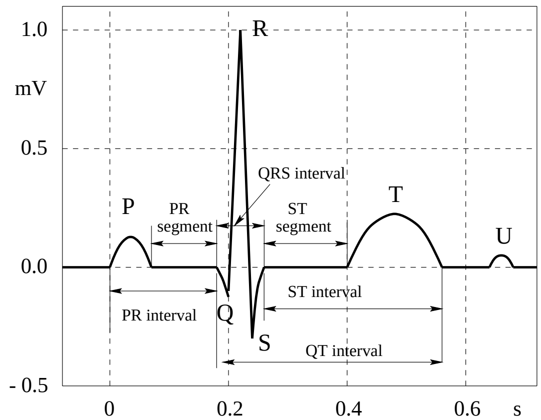


**Fig. 1.** The ECG waveshape

and a different waveshape appears in the ECG signal. There are several classes of arrhythmic beats, the most common of which refer to an anomalous origin in the atria (SupraVentricular Premature Beats, SVPB) or in the ventricula (Ventricular Premature Beats, VPB) and occur randomly in time. Supraventricular arrhythmia originate in the atria, therefore not far from the sino-atrial node. The morphology of the resulting ECG waveshape is very close to the one of normal beats, except for the P wave and a shorter preceding beat-to-beat interval (the RR interval). Ventricular arrhythmia originate in the ventricula and the corresponding ECG waveshape presents generally a very altered QRS shape and sometimes a shorter preceding RR interval.

With the development of automatic systems for the detection of QRS complexes and the extraction of their quantitative measurements, large sets of data can be generated by extracting beat measures from hours of ECG signal. A higher number of measures though does not guarantee better performances of the upcoming beat classifier, because of the quality of each measure as well as of its significance. It becomes important then a first screening of the collected measures, in order to keep only the most significant ones for the analysis. This has the double advantage of making easier the classification task and possibly of improving the classifier's performance if poor quality measures are discarded.

The MIT-BIH database [4] represents by now a standard in the evaluation of methods for the automatic classification of the ECG signal, because of the wide set of examples of arrhythmic events provided. The MIT-BIH ECG records are two-channel, 30 minutes long, sampled at 360 samples/s and annotated by trained cardiologists. A subset of 39 ECG records is selected for this work on the basis of the presence of only arrhythmic classes. All files with pace-maker beats are not considered. QRS complexes are detected and for each beat waveshape a set of 12 measures [5] is extracted by using the first of the two channels in the ECG record (Tab. 1).

**Table 1.** Set of measures characterizing each QRS complex

| | |
|---|---|
| RR | RR interval/average of the previous 10 RR intervals |
| QRSw | QRS width (ms) |
| pAmp | Positive amplitude of the QRS ($\mu$V) |
| nAmp | Negative amplitude of the QRS ($\mu$V) |
| pQRS | Positive area of the QRS ($\mu$V * ms) |
| nQRS | Negative area of the QRS ($\mu$V * ms) |
| T | area of the T wave ($\mu$V * ms) |
| IVR | Inverted ventricular repolarization: $IVR = \frac{pQRS+nQRS}{T}$ |
| ST | ST segment level ($\mu$V) |
| STsl | slope of the ST segment ($\mu$V/ms) |
| P | P exist (yes 0.5, no -0.5) |
| PR | PR interval (ms) |

# 3 Impact of input features on a fuzzy decision process

## 3.1 The use of fuzzy logic in medical reasoning

As it was already observed in section 2, fuzzy logic has always been particularly appealing to physicians and medical experts for two main reasons. The first one is that fuzzy logic allows the definition of qualitative rules with a certain amount of uncertainty, which is very similar to the medical reasoning process. The second reason of fuzzy logic popularity for clinical applications consists of its easy interpretability, so that the decision process is extremely easy to understand by human experts. Moreover, such easy interpretability allows external changes by experts on the decision process.

In the past, because of this similarity of fuzzy logic with medical reasoning, several attempts have been made to translate medical knowledge into a set of fuzzy rules [5]. In general, however, such translation process was not accurate enough, lacking of the description of many particular cases. Consequently, automatic fuzzy classifiers quite frequently outperformed systems with fuzzy rules derived from medical knowledge. On the other side, to cover all possible particular cases of the training set, a very high number of fuzzy rules is usually generated by the automatic classifiers. This high number of fuzzy rules and the high dimension of the input space complicate the work of medical experts in interpreting and comparing different fuzzy systems for the same analysis task. The definition of automatic tools for the description of particular properties of fuzzy models would supply a support for the interpretation work of experts.

In this paper, we apply a modification of a fuzzy feature merit measure, introduced in [3], for the analysis of fuzzy systems that implement the three-class discrimination task described in section 2. The fuzzy feature merit measure consists of a measure of the information gain that follows the use of a given input feature $x_j$ for the classification process.

### 3.2 Possibilistic information gain

Given a number $m$ of output classes $C_i$, $i = 1, \ldots, m$, and an $n$-dimensional input space, numerous algorithms exist, which automatically derive a set of $N_R$ fuzzy rules $\{R_k\}$, $k = 1, \ldots, N_R$, mapping the $n$-dimensional input into the $m$-dimensional output space. This set of rules models the relationships between the input data $\boldsymbol{x} \in \mathcal{R}^n$ and the output classes $C_i \in \mathcal{R}^m$.

Each input pattern $\boldsymbol{x} = \{x_1, \ldots, x_n\}$ is associated to each output class $C_i$ with a membership value $\mu_{C_i}(\boldsymbol{x})$ resulting from the set of rules $\{R_k\}$.

$$V(C_i) = \int_{D \supset R^n} \mu_{C_i}(\boldsymbol{x}) \; d\boldsymbol{x} \tag{1}$$

$$v(C_i) \; = \; \frac{V(C_i)}{\sum_{j=1}^{m} \; V(C_j)} \tag{2}$$

The volume of the membership function $\mu_{C_i}(\boldsymbol{x})$ over the domain $D \supset R^n$ (eq. 1) represents the smallest tile of information to use to span the whole decision process.

An information measure can then be applied, to distinguish fuzzy models with only one membership function (no information) from fuzzy models with a very high number of membership functions (high information). The usual information measures from information theory could be applied, such as the entropy or the Gini function. However such information functions require the variable to sum up to 1, to correctly describe the different contribution of each variable. The sum of volume $V(C_i)$ across output classes $i = 1, \ldots, m$ is not necessarily one, but actually grows with the number of output classes.

A solution to this problem was found in [3] by applying the cited information measures to the relative volume $v(C_i)$ of class $C_i$ (eq. 2), as defined in equation 3 and 4. Such information measures are 0 when the set of fuzzy rules is empty or describes only one output class. The maximum of information is reached with a number $m$ of output classes with similar relative volumes. If the membership functions range between 0 and 1, similar relative volumes means volumes extending on similar portions of domain $D$.

$$I(C) \; = H(C) \; = \; - \sum_{i=1}^{m} \; v(C_i) \; log_2(v(C_i)) \tag{3}$$

$$I(C) \; = G(C) \; = \; 1 \; - \; \sum_{i=1}^{m} \; (v(C_i))^2 \tag{4}$$

The measures of information defined in 3 and 4 do not take into account how representative a membership function is. For example, membership function $\mu_{C_i}(\boldsymbol{x})$ of class $C_i$ could represent only a few outliers of the training set. On the opposite $\mu_{C_j}(\boldsymbol{x})$ of class $C_j$ could represent a relevant portion of data of the training set. More information is then associated with $\mu_{C_j}(\boldsymbol{x})$ than to $\mu_{C_i}(\boldsymbol{x})$.

This observation is reported into the definition of the information measure by using a different volume from $V(C_i)$ in eq. 1. In eq. 5 the number $N(C_i)$ of

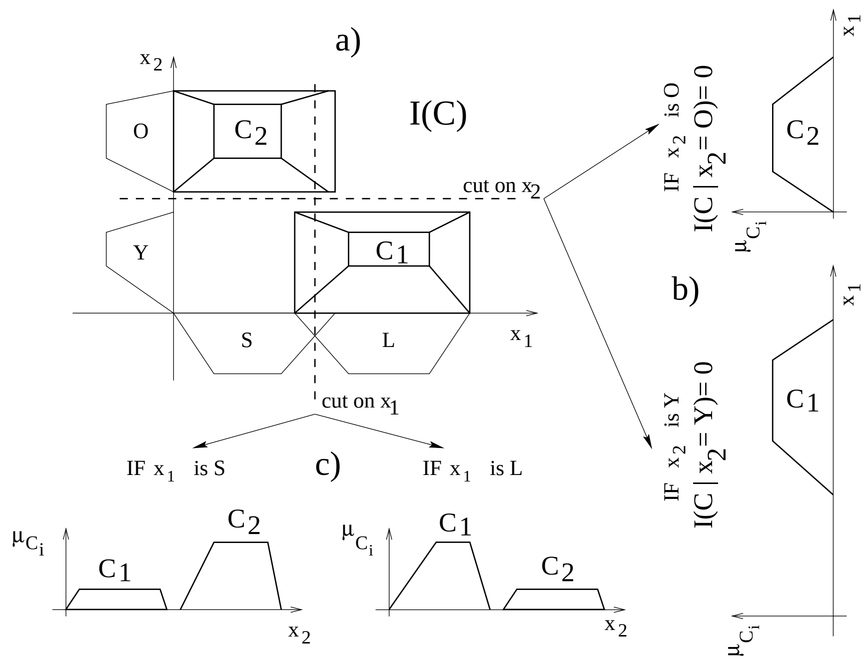


**Fig. 2.** Stripes $c_k$ generated by cutting along variable $x_2$ (b)and $x_1$ (c)

training data represented by membership functions $\mu_{C_i}(\boldsymbol{x})$ is introduced as a weight for the volume of $\mu_{C_i}(\boldsymbol{x})$. Membership functions that represent a higher number of data will have more influence in the definition of the information measure. Membership functions covering just a small number of outliers will influence less the information measure value.

$$\hat{V}(C_i) = \left( \int_{D \supset R^n} \mu_{C_i}(\boldsymbol{x})\ d\boldsymbol{x} \right)\ N(C_i) = V(C_i)\ N(C_i) \tag{5}$$

$I(C)$, as in eq. 3 and 4, represents the amount of possibilistic information intrinsically available in the fuzzy model, reflecting the information contents in the training set. At this point, we want to characterize the discriminative power of the input features of the fuzzy model by means of such defined information measure.

Given a fuzzy description of the input space, the use for classification of input variable $x_j$ consists of an appropriate threshold system definition along input dimension $j$. The optimal classification thresholds on a given dimension $j$ are located at the intersection points of contiguous membership functions of different output classes [6]. If the input parameter $x_j$ is employed to classify the input space, a set of thresholds is created to separate the $F_j \leq N_R$ contiguous trapezoids on the $j$ input dimension representing different output classes. Between two consecutive thresholds a linguistic value $L_k$, $k = 1, \ldots, F_j$, can be defined for variable $x_j$ (figure 2).

Considering $x_j = L_k$ corresponds to intersect the original membership functions with the segment $x_j = L_k$, that is to isolate one stripe $c_k$ on the input space. Each stripe $c_k$ is characterized by a local information $I(C|x_j = L_k)$ as defined in eq. 3 or 4. The average possibilistic information $I(C|x_j)$ of the fuzzy model across all linguistic values $L_k$ corresponds to the averaged sum of the local information of stripes $c_k$ (eq. 6) and represents the information left into the system after variable $x_j$ was used for the analysis.

$$I(C|x_j) = \frac{1}{F_j} \sum_{k=1}^{F_j} I(C|x_j = L_k) \tag{6}$$

Finally, the relative information gain derived from the use of variable $x_j$ can be defined as:

$$g(C|x_j) = \frac{I(C) - I(C|x_j)}{I(C)} \tag{7}$$

The $x_j$ input features producing the highest information gain are the ones mainly used by the adopted model to express the possibilistic information intrinsic to the input space, and therefore the most effective for the proposed analysis.

The defined possibilistic information gain was developed taking as a model the decision trees approach [7]. Instead of the a priori probability of class $C_i$, $p(C_i)$, the relative volume of class $C_i$ is considered. The main advantages of using a fuzzy instead of a probabilistic model are the low computational expenses and the interpretability in terms of medical knowledge.

Low computational expenses are due to the simple nature of fuzzy logic. Because of that, the definition of the membership functions and the calculation of their volumes do not require any complex mathematical operations, even more if particularly simple membership functions as trapezoids are chosen. The representation of decision process by means of qualitative rules represents the second advantage of the use of fuzzy models. In this way, a precise quantization of the input features and a sure attribution of each pattern to an exclusive output class are not required.

In the future a comparison between the defined fuzzy information gain and probabilistic feature merit measures should be performed. In this work, we limit our investigation to the use of fuzzy models.

## 4 ECG measures impact on cardiac arrhythmia classification

### 4.1 Arrhythmia classification

A three-class problem – Normal (N) vs. Ventricular Premature (VPB) vs. SupraVentricular Premature Beats (SVPB) – is considered. Two thirds of the MIT-BIH

**Table 2.** System performance on the test set

| N | VPB | SVPB | average | unc. |
|---|---|---|---|---|
| .92 | .78 | .71 | .80 | .04 |

database records are used as training set and the remaining one third as test set. Twelve ECG beat measures, as described in table 1, are used as input vector to characterize the ECG beats.

At this point, an automatic fuzzy rules generator is implemented to classify the input space. Several algorithms are available to perform this task. We adopted the algorithm described in [8], because it has already shown good performances on other classification tasks [8]. Trapezoidal membership functions are used, because of their computational properties. The algorithm automatically constructs a set of fuzzy rules based on a set of training examples, by introducing new fuzzy points when necessary and adjusting the existing ones to cover new training examples. For classification purpose, the membership functions of the fuzzy model were weighted by the number of covered training patterns $N(C_i)$. This helps in solving conflicts between very representative membership functions towards membership functions representing outliers of the training set. With this strategy performance usually improves for all the output classes and particularly for the correct recognition of SVPBs [6].

As a first attempt, all the twelve ECG measures are used for the classification. The proportions of correctly classified beats (N, VPB and SVPB) of the test set are reported in table 2. Beats are labeled as uncertain (unc.) if they are not covered by any trapezoids of the fuzzy model. The proportion of uncertain beats (unc.) is defined with respect to the number of beats in the whole test set, as in eq. 8 where $n_{uncovered}$ is the number of uncovered ECG beats and $n_{patterns}$ the total number of ECG beats in the test set.

$$unc = \frac{n_{uncovered}}{n_{patterns}} \tag{8}$$

We have decided to show the system performance by means of the proportion of correctly classified beats for each output class separately, in order to have an idea of the system performance towards each output class. More global indices, such as the average of correctly classified beats across output classes reported in the fourth column of table 2, do not depict a sufficiently detailed frame of the system performance for the purpose of this paper.

The system performance is quite good in comparison with the state of the art [5]. It remains to ascertain whether all twelve ECG measures are necessary for such performance and, if not, which are mainly responsible for the correct classification and which are useless.

### 4.2 Ranking ECG measures

The information gain, as defined in section 3.2, is now calculated for each one of the input features of the fuzzy model. The obtained values are reported in the first row of table 3 in the first twelve columns. The last four columns report the proportions of correctly classified and uncertain beats, that, at least in the first row, are the same as in table 2. The highest information gains are marked in bold.

Subsequently, the ECG measures with smallest information gain are progressively removed from the input vector. A new set of fuzzy rules is then obtained as a projection of the previous system on the remaining input features and without any re-training process. The new information gains and the new system performance are calculated and reported in the following rows of table 3.

**Table 3.** Information gains for different ECG measures.

| RR | QRSw | pAmp | nAmp | pQRS | nQRS | T | IVR | ST | STsl | P | PR | N | VPB | SVPB | unc. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| .09 | .16 | **.35** | .02 | **.44** | .09 | .01 | .05 | .00 | .04 | .00 | **.36** | .92 | .78 | .71 | .04 |
| .09 | .15 | **.36** | .02 | **.44** | .05 | .01 | .04 | .00 | .03 | - | **.37** | .91 | .77 | .78 | .02 |
| .06 | .14 | **.36** | .01 | **.45** | .02 | .01 | .05 | - | .03 | - | **.36** | .94 | .77 | .76 | .00 |
| .07 | .15 | **.31** | - | **.25** | .05 | .01 | .01 | - | .04 | - | .15 | .97 | .76 | .80 | .00 |
| .04 | .16 | **.30** | - | **.24** | .04 | - | .01 | - | .04 | - | .15 | .97 | .74 | .82 | .00 |
| .02 | .14 | **.26** | - | **.21** | .02 | - | - | - | .03 | - | .18 | .96 | .76 | .84 | .00 |
| .03 | .14 | **.25** | - | **.21** | - | - | - | - | .03 | - | **.19** | .89 | .81 | .68 | .00 |
| .01 | **.13** | **.14** | - | **.14** | - | - | - | - | - | - | **.11** | .93 | .81 | .69 | .00 |
| - | .07 | .04 | - | .04 | - | - | - | - | - | - | **.15** | .72 | .81 | .44 | .00 |
| - | .05 | - | - | .06 | - | - | - | - | - | - | **.14** | .68 | .80 | .64 | .00 |
| - | - | - | - | **.11** | - | - | - | - | - | - | **.18** | .72 | .16 | .65 | .00 |
| - | - | - | - | - | - | - | - | - | - | - | .05 | .73 | .00 | .67 | .00 |
| - | - | - | - | .13 | - | - | - | - | - | - | - | .71 | .47 | .19 | .00 |
| - | - | .15 | - | - | - | - | - | - | - | - | - | .80 | .28 | .00 | .00 |
| - | .10 | - | - | - | - | - | - | - | - | - | - | .78 | .75 | .04 | .00 |

In the first row of table 3, where all twelve ECG input parameters are used, only three ECG measures present a remarkably high information gain: the positive amplitude of the QRS, the positive area of the QRS and the PR interval. The first two are related with each other and likely share the task of distinguishing VPBs from normal beats. The PR interval on the contrary should be responsible for the classification of normal vs. SVPB beats. Smaller, but still not negligible, information lies on the width of the QRS complex, the RR interval and the negative area of the QRS complex.

Of the original twelve ECG input measures only six appear to be necessary to solve the arrhythmia classification problem on the adopted subset of the MIT-BIH database. The remaining six ECG measures produce zero (P wave and ST amplitude) or close to zero (negative amplitude of the QRS complex, Inverted Ventricular Repolarization, ST slope and area of the T wave) information gain.

The removal of the input parameters with zero information gain does not make worse the performance of the fuzzy classifier, but actually improves the normal vs. SVPB beats classification of some point, as shown in second and third row of table 3 after removing the P wave existence and the ST amplitude respectively.

Even the removal of IVR, T wave area and negative amplitude of the QRS complex – all with information gain close to zero in the original system – does not degrade, but improves the system performance up to .96 normal beats, .76 PVBs and .84 SVPBs correctly recognized. Such ECG measures with circa zero information gain are not only redundant but lower the system performance, because very likely they focus on the classification of a few exceptions of the training set.

Until now, only the input features with very low information gain in the first row of table 3 were removed and the system performance was not damaged by that. At this point only the six input features with not negligible information gain in the original set of fuzzy rules are left in the input vector. Let's see what happens if the ones with lowest information gain in the derived fuzzy model are removed. Removing the ST slope and the negative amplitude of the QRS complex reduces the system performance of some points, particularly the normal vs. SVPB classification.

However the system performance is still acceptable (seventh and eighth row of table 3). It drops down dramatically only when one of the remaining ECG measures is removed from the input vector. Indeed the four ECG measures left at this point in the input vector were the ones with highest information gain in the original model and because of that they are expected to be quite influential on the overall decision process.

As first, the RR parameter is removed, because it has the lowest information gain at this point of the analysis. While this does not affect the VPBs classification rate (still .81), the discrimination of SVPBs (only .44) vs. normal beats (.72) becomes less reliable. The information gain, as defined in section 3.2, does not only describe the separability of the output classes, but also the fragmentation of the membership functions on a given input dimension. A very fragmented description, that is many small linguistic values, of input dimension $x_j$ will produce many stripes $c_k$ each one covering a limited number of training examples. In each stripe $c_k$ all relative volumes will be similar, because of the reduced size of the stripe and because of the low value of weight $N(C_i)$. $I(C|x_j)$ will be high and consequently $g(C|x_j)$ will be low. The RR parameter is then penalized by its very fragmented set of linguistic values.

In the following row, a more fair discrimination between normal beats and SVPBs is re-established, by the removal of the positive amplitude of the QRS complex. This shows that the pAmp parameter is informative only if together with the prematurity degree of the beat. The system performance is still relatively high, but the configuration of the input vector is quite minimal. The removal of the next input feature (QRS width) will damage further the system performance.

In the bottom part of the table, the most important input features are analyzed alone. The PR interval results to be important for the discrimination between SVPBs and normal beats exclusively. The QRS width and up to some extent the QRS positive amplitude provide the discrimination between VPBs and normal beats. Finally the positive area of the QRS complex distinguishes part of the VPBs and part of the SVPBs from the normal beats.

Alone these ECG input measures can not perform as well as when they are considered together. This is expressed by the drop in information gain, occurring when one of the other ECG measures is missing. For example, the RR interval itself shows quite a low information gain, due to a too fragmented classification along this dimension. On the other side, when the RR interval is removed from the input patterns the information gains of QRS width, positive amplitude and positive area decrease dramatically. Only the PR interval keeps its information gain constant, showing its independence on the RR interval in the decision system.

Such considerations are confirmed by medical knowledge, where the QRS complex features are usually employed for ventricular arrhythmia classification and the prematurity degree, the PR interval and the P wave existence are used mainly for the SVPBs and partially for the VPBs classification vs. normal beats. The fact that the prematurity degree parameter (RR) is so fragmented is somewhat surprising, but that could be the key of the different behaviour of fuzzy systems derived from medical knowledge and automatic fuzzy classifiers. Indeed the clusterization of the output classes based on the RR parameter could be more fragmented than what assumed by medical knowledge, because different side conditions may give different meaning to the prematurity degree of a beat. On the other side, it could also be that the prematurity degree as defined in table 1 is inadequate for the implementation of a reduced set of fuzzy rules.

A similar set of experiments was performed on some of the considered ECG records [3], particularly on records 200 and 233. In this case, a re-training phase of the system was allowed after the removal of the input features with lowest information gain. Generally, after re-training the information gain is differently distributed among the ECG measures describing the same part of the ECG waveshape. For example, new QRS complex measures, different from the previous ones, can carry the main information for VPBs classification. Since all QRS measures are related with each other and carry similar amounts of information, the algorithm chooses every time one of those measures according to the sequence of training pattern presentation and to the more or less informative combinations of input features.

## 5 Conclusions

A method is presented to analyze input features merit in fuzzy systems. After a fuzzy model is available for a given training set, its input features are analyzed in terms of information for the final task implemented by the fuzzy model. In particular, a cardiac arrhythmia classification task has been considered for this paper.

Twelve simple and composed ECG measures are used to characterize cardiac beats of electrocardiographic records from the MIT-BIH database. A fuzzy model is built and the input features merit analyzed. Six main ECG measures are detected: the prematurity degree (RR), the QRS width, the positive QRS amplitude and area, the negative QRS area and the PR interval. Among them, the PR interval was mainly responsible for normal vs. SVPBs classification, while the QRS width and positive area and amplitude characterized several sub-clusters of VPBs. The main advantage of the presented method consists of the low computational expenses, inheritated from the characteristics of fuzzy systems.

In the future, a comparison with probabilistic feature merit measures has to be performed and the information gain associated with combinations of ECG measures for arrhythmia classification has to be investigated.

## 6 Acknowledgments

The author would like to thank Wei Zong, George Moody, and prof. Roger G. Mark from Harvard-MIT Division of Health Sciences and Technology M.I.T. (USA) for the ECG measures.

## References

1. V. Cherkassky and F. Mulier, "Learning from data", John Wiley and Sons Inc., 1998.
2. M. Berthold and D. Hand eds, "Intelligent Data Analysis: An Introduction", Springer-Verlag, Berlin-Heidelberg, 1999.
3. R. Silipo and M. Berthold, "Discriminative Power of Input Features in a Fuzzy Model", Proc. of Intelligent Data Analysis, Lecture Notes in Computer Science LNCS1642, Springer-Verlag, pp. 85-96 1999.
4. MIT-BIH database distributor, Beth Israel Hospital, Biomedical Engineering, Division KB-26, 330 Brookline Avenue, Boston, MA 02215, USA.
5. W. Zong, D. Jiang. "Automated ECG rhythm analysis using fuzzy reasoning", Proc. of Computers in Cardiology, 1998.
6. R. Silipo and M. Berthold, "Analyzing fuzzy models: Discriminative Power of Input Features", Technical Report, University of California at Berkeley, to appear, 1999.
7. Quinlan,J.R., "Induction of Decision Trees", in: *Machine Learning*, Vol. 1, Kluwar Academic Publishers, 1986.
8. K.P. Huber, M.R. Berthold, "Building precise classifiers with automatic rule extraction", in *IEEE Int. Conf. on Neural Networks* 3: 177-184, 1995.