Heartbeat classification using abstract features from the abductive interpretation of the ECG

Tomás Teijeiro*, Paulo Félix, Jesús Presedo, Daniel Castro

Abstract—*Objective:* **This paper aims to prove that automatic beat classification on ECG signals can be effectively solved with a pure knowledge-based approach, using an appropriate set of abstract features obtained from the interpretation of the physiological processes underlying the signal.** *Methods:* **A set of qualitative morphological and rhythm features are obtained for each heartbeat as a result of the abductive interpretation of the ECG. Then, a QRS clustering algorithm is applied in order to reduce the effect of possible errors in the interpretation. Finally, a rule-based classifier assigns a tag to each cluster.** *Results:* **The method has been tested with the MIT-BIH Arrhythmia Database records, showing a significantly better performance than any other automatic approach in the state-of-the-art, and even improving most of the assisted approaches that require the intervention of an expert in the process.** *Conclusion:* **The most relevant issues in ECG classification, related to a large extent to the variability of the signal patterns between different subjects and even in the same subject over time, will be overcome by changing the reasoning paradigm.** *Significance:* **This work demonstrates the power of an abductive framework for time series interpretation to make a qualitative leap in the significance of the information extracted from the ECG by automatic methods.**

Index Terms—**Heartbeat classification, Abductive reasoning, Knowledge based systems, Biomedical signal processing.**

I. INTRODUCTION

HEARTBEAT classification from electrocardiogram signals is a valuable tool for the study of the cardiac arrhythmia, and it is one of the challenges that has raised more efforts in the field of biosignal analysis [1]. Despite new proposals appear incessantly, it is still considered an open problem, and it seems we are far from providing sufficiently satisfactory solutions to be transferred to clinical routine, integrated in the bedside instrumentation or in the emergent home monitoring. A number of difficulties can be identified: 1) the variability of the physiological and pathophysiological processes underlying the ECG tracing between different patients, or even in the same patient over time; 2) the stochastic nature of these processes; 3) the simultaneous occurrence of multiple physiological processes that can interact in different ways; 4) the presence of noise and artifacts in the signal which mask the physiological processes; or 5) the absence of an accurate heart model and the tacit, subjective, and hardly formalizable knowledge that constitutes the experience of the cardiologist.

Heartbeat classifiers have traditionally failed to apply to new patients those categories learned from a training set, no matter how large this training set may be. Some sort of automatic adaptation to the specific characteristics of each new patient could be thought of as the solution, but results have not been satisfactory enough. In fact, recent bibliography shows that most of the state-of-the-art approaches rely on some form of expert assistance for processing every new patient: some of them by combining a general classifier, previously trained with large collections of ECG recordings, with a local one, trained specifically with an annotated fragment of the ECG record of the new patient [2], [3], [4]; others by following a clustering strategy that requires from the expert the final assignment of a class label to each beat morphology [5], [6], [7].

The present paper aims to return to the challenge of performing an autonomous classification, that is, without requiring expert intervention after the classifier has been designed. To this end, we take inspiration from human interpretation of ECG, which is based on the construction of an explanation for the behavior of the heart that can account for the ECG tracing of a given recording. Our proposal provides an interpretation of the ECG recording as a sequence of processes that are conjectured over time. This requires a new formalism for representing expert knowledge and a new reasoning paradigm, under the following principles:

- *The knowledge will be explicitly represented*, the kind of knowledge that can be found in an ECG handbook, so that it can be understood and validated by experts; the feature set is limited to those used by experts in clinical practice, such as wave durations, intervals, amplitudes, etc. Interpretability of the classification model is therefore a natural consequence.

- *Ad-hoc thresholds and values will be avoided*, since they may be suspicious of being overfitted to a training database. For this reason, all classification features will be qualitative, and the separation between values will be performed either by purely electrocardiographic criteria, or by generic criteria if no quantifiable knowledge is available.

- *Ignorance will be admitted*, in case of a beat cannot be successfully assigned to any predefined class. Ignorance has the ability to clearly show those weaknesses in the knowledge base, providing a direction for further improvement. Moreover, the credibility of the results is enhanced, as long as it comes together with a good performance.
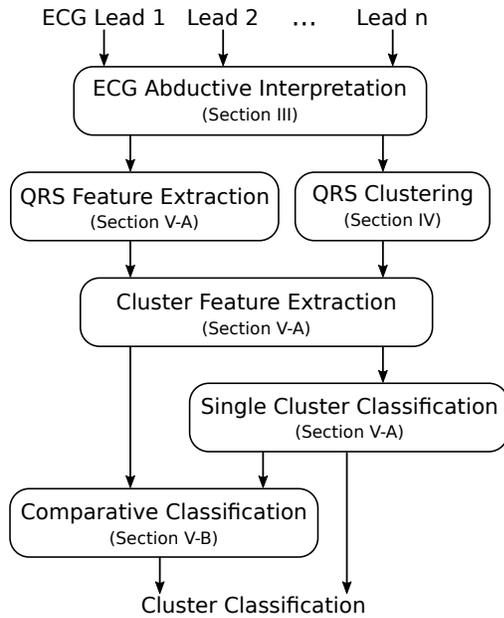
Fig. 1. Overview of the proposed classification method

- *Robustness to variability will be demanded*, even to the kind of noise present in the ECG. The interpretation cannot involve any single piece of evidence in isolation, but it should take into account contextual information, combining bottom-up and top-down processing in order to provide a more informed result.

The present proposal is based on the notion of temporal abstraction pattern, representing a set of constraints that must be satisfied by some evidence from the ECG for being interpreted as the hypothetical observation of a certain physiological process, together with an observation procedure providing a set of measurements for the features of the conjectured observation. A set of algorithms is used in order to achieve the best explanation through a process of successive abstraction from raw data, by means of a hypothesize-and-test strategy [8]. The interpretation results in a set of observations describing the myocardial behavior at the conduction and rhythm abstraction levels, and providing the same signal features used by cardiologists in ECG analysis [9]. A context-based adaptive QRS clustering method is then applied to cope with within-patient variability, obtaining a reduced number of groups representing the cardiac activity [10]. Finally, a simple knowledge-based classification procedure assigns a label to each cluster. Figure 1 shows an overview of the present proposal.

The rest of this paper is outlined as follows: Section II describes the database and methodology used for the validation of the proposal. Section III introduces the abductive method for ECG interpretation and the knowledge base used. Section IV describes the algorithm for clustering the QRS complexes resulting from the interpretation. Afterwards, section V details the feature extraction and classification stages. Section VI presents the classification results, compared with other state-of-the-art approaches. Finally, section VII discusses the advantages and drawbacks of this method, as well as the future work and evolution perspectives.

II. VALIDATION DATABASE

The validation of the proposed approach was performed using the MIT-BIH Arrhythmia Database [11] from the Physionet initiative [12]. This database can be considered the gold standard for arrhythmia and heartbeat classification, and it has been used in most of the published research. It contains 48 ECG recordings sampled at 360 Hz for 30 minutes and from 47 different patients. Each recording comprises two ECG leads, mainly a modification of lead II (electrodes on the chest) and V1, but in some records they are replaced by V2, V5 or V4.

All beats present in the database were annotated by at least two expert cardiologists, and assigned a class label using a 17 label set. This label set was converted to the five standard beat classes defined by the AAMI, following the recommendations of the American National Standard for Ambulatory ECGs (ANSI/AAMI EC38:2007) [13], detailed in table I. All the comparisons were performed using the `bxb` application from Physiotools [12]. It is worth mentioning that many works in the bibliography, particularly those following the methodology introduced by de Chazal *et al.* [14], are not fully compliant with this standard, since MIT-BIH 'j' and 'e' classes are grouped in the AAMI 'N' class instead of 'S' class. This incorrect labeling affects 245 beats in the database, representing 8% of the total supraventricular beats.

TABLE I
EQUIVALENCE BETWEEN BEAT CLASSES IN THE MIT-BIH ARRHYTHMIA
DATABASE AND THE STANDARD AAMI CLASSES

AAMI	MIT-BIH equivalent classes
N	N, L, R, B
S	a, J, A, S, j, e, n
V	V, E
F	F
Q	/, f, Q

III. ABDUCTIVE INTERPRETATION OF THE ECG

The interpretation of the ECG aims to identify and characterize the physiological processes underlying signal behavior by building a representation of the cardiac phenomena in multiple abstraction levels. To this end, an abductive framework for time series interpretation is adopted [8]. This framework demands the definition of a domain-dependent knowledge base that is used to build an explanation of what is observed.

Domain entities are represented by *observables*. Formally, an observable is a tuple $q = \langle \psi, \mathbf{A}, T^b, T^e \rangle$, where ψ is a name representing the process being observable, $\mathbf{A} = \{A_1, ..., A_{n_q}\}$ is a set of attributes, and T^b and T^e are two temporal variables representing the beginning and the end of the observable. An observable may be observed in multiple instances called *observations*, that are obtained by assigning a specific value to each attribute and to the temporal variables.

For example, the observable $q_{Twave} = \langle \texttt{ventricular\_re-covery}, \{\texttt{amplitude}\}, T^b, T^e \rangle$ represents the T wave resulting from the ventricular electrical recovery process, and the observation $o_i = \langle q_{Twave}, \texttt{amplitude} = (V1 : 0.09mV,$

$MLII : 0.33mV$), $T^b = 00{:}00.403, T^e = 00{:}00.606\rangle$ represents the first specific T wave that can be observed in the record 101 of the MIT-BIH arrhythmia database.

The full set of observables $\mathcal{Q}$ defined for the ECG interpretation problem is shown in table II. This set is organized into three abstraction levels: The first one corresponds to signal deviation phenomena, and includes a set of intervals or time points labeled as *deflections*, consistent with the presence of electrical activity from the cardiac muscle. Observations from this level can be obtained from any of the available ECG leads, and constitute the initial evidence for the abductive interpretation process. The second abstraction level corresponds to intracardiac conduction phenomena, and defines the *P wave*, *QRS complex*, and *T wave* observables. They can be conjectured taking as evidence observations from the first abstraction level, providing a description of the ECG as a sequence of waves corresponding to the atrial activation, ventricular activation and ventricular recovery processes, respectively. These observables integrate information from multiple leads, as the observation o_i shows in the above example. In particular, the observation of a *Deflection* in a single channel consistent with the definition of a *T wave* is enough to consider that the ventricular recovery process has been observed, regardless the presence of evidence in other leads. Finally, the third abstraction level collects the patterns characterizing the *cardiac rhythm* under different physiological conditions [9], including normal rhythm and arrhythmias.

TABLE II
OBSERVABLES FOR THE ECG INTERPRETATION PROBLEM, SPLIT INTO THE THREE CONSIDERED ABSTRACTION LEVELS

Name	Underlying Process
Deflection	Signal deviation consistent with the electrical activity of the cardiac muscle fibers.
R-Deflection	Signal deviation consistent with the electrical activity generated in the ventricular activation.
P wave	Atrial electrical activation.
QRS complex	Ventricular electrical activation.
T wave	Ventricular electrical recovery.
Sinus Rhythm	Normal rhythm originated in the sinoatrial node.
Bradycardia	Regular rhythm with a low frequency heart rate (under 60 beats per minute).
Tachycardia	Regular rhythm with a high frequency heart rate (over 100 beats per minute).
Extrasystole	Premature contraction of the heart.
Bigeminy	Rhythm pattern in which every sinus beat is followed by a premature beat.
Trigeminy	Rhythm pattern in which every second sinus beat is followed by a premature beat.
Couplet	Concurrence of two consecutive extrasystoles.
Rhythm Block	One-time elongation of the cardiac rhythm.
Atrial Fibrillation	Arrhythmia caused by the independent and erratic contractions of the atrial muscle fibers, characterized by an irregularly irregular heart rhythm.
Ventricular Fibrillation	Rapid ventricular activity without discernible QRS complexes or T waves in the ECG.
Asystole	Interval of absence of ventricular activity.

An abstraction relation between the aforementioned observables is defined by a set of abstraction patterns. An *abstraction pattern* is a knowledge representation primitive that allows us to hypothesize the presence of an observable from a set of other observables, playing the role of findings, when they appear with a distinctive temporal arrangement. Formally, an abstraction pattern is a tuple $P = \langle q_h, M_P, C_P, \pi_P \rangle$, where q_h is the observable being hypothesized, $M_P = \{m_1, \ldots, m_n\}$ is the set of findings abstracted by the hypothesis q_h, $C_P(\mathbf{A}_h, T_h^b, T_h^e, \mathbf{A}_1, T_1^b, T_1^e, \ldots, \mathbf{A}_n, T_n^b, T_n^e)$ is a set of constraints involving the temporal location and the attributes of q_h, and of the findings from M_P, and $\pi_P(\mathbf{A}_1, T_1^b, T_1^e, \ldots, \mathbf{A}_n, T_n^b, T_n^e) \in O(q_h)$ is an observation procedure to compute the specific attribute and temporal location values of a newly generated observation $o_h \in O(q_h)$, where we denote by $O(q)$ the set of observations of the observable q, and we denote by $\mathcal{O} = O(\mathcal{Q})$ the full set of observations for a given recording.

To define the temporal constraints in abstraction patterns the Simple Temporal Problem (STP) formalism [15] has been adopted. A STP is a network defining a constraint $\tau c(T_i, T_j)$ between every two temporal variables T_i and T_j as a closed interval $\tau c_{ij} = [a_{ij}, b_{ij}]$, where a_{ij} and b_{ij} are integer numbers constraining the possible values of the duration of the interval between both temporal variables, so that $a_{ij} \leq T_j - T_i \leq b_{ij}$.

As an example, the following abstraction pattern defines the conditions to conjecture the observation of a *QRS complex* (q_{QRS}) from the observation of a *R-Deflection* (q_{RDef}):

$$P = \langle q_{QRS}, \{m_1^{RDef}\}, \{\tau c(T_{RDef}, T_{QRS}^b) = [-80ms, 80ms],$$
$$\tau c(T_{RDef}, T_{QRS}^e) = [15ms, 400ms]\}, \texttt{qrsdel}() \rangle$$

R-Deflections are instantaneous observables, so they can be represented with a single temporal variable T_{RDef}. C_P defines two temporal constraints between the time point of the *R-Deflection* and the beginning and the end of the *QRS complex*, implicitly limiting its duration. The observation procedure $\texttt{qrsdel}()$ refers to any QRS delineation method used to find the specific temporal limits of the complex. The multi-lead method described in [16] has been used in this work.

In general, the set of findings M_P is divided into two sets: A_P and E_P, with $A_P \cap E_P = \varnothing$. A_P collects the set of findings that is abstracted by q_h, while E_P is the set of findings comprising the environment of q_h, that is, the set of findings that can condition the hypothesis of q_h, but which are not a constituent part of q_h. For example, to conjecture a *T wave* from the observation of a signal deviation it is necessary an earlier observation of a *QRS complex*, but only the signal deviation is abstracted by the *T wave*, being the *QRS complex* the contextual evidence.

By definition, an abstraction pattern is built upon a fixed set of evidence findings M_P, but usually the notion of abstraction involves an undetermined number of evidence observations. An example is the normal rhythm abstraction, which is built upon an indefinite number of heartbeats. To support this type of abstractions, the interpretation framework provides a method based on the formal language theory for the dynamic generation of abstraction patterns. This method considers the set of observables $\mathcal{Q}$ as an alphabet, and defines a class

of grammars G^{ap} for the generation of abstraction patterns. An *abstraction grammar* $G \in G^{ap}$ is defined as a tuple (V_N, V_T, H, R), where the production rules in R take one of the following general forms:

$$H = q_h \rightarrow q[c]D$$
$$D \rightarrow q[c]F \mid q[c] \mid \lambda$$

$H = q_h$ is the initial symbol of the grammar, corresponding to the hypothesis guessed by the patterns generated by G. Thus, a single grammar describes all the different ways of abstracting an observable q_h. V_N is the set of non-terminal symbols of the grammar (here D and F). V_T is the set of terminal symbols, including a set of observables $\mathcal{Q}_G \subseteq \mathcal{Q}$ that can be abstracted by the hypothesis, a set of optional constraint descriptions $[c]$ between the observables, and the membership of each finding to the set A_P or E_P. λ stands for the empty string.

Given a grammar $G \in G^{ap}$, the abstraction patterns that allow us to hypothesize q_h are built iteratively by the application of the production rules in R. Thus, each production adds a new observable as a finding, and a set of constraints among this finding and those generated previously, according to the following method:

1) The start symbol H initializes an abstraction pattern with the hypothesis and no evidence findings:

$$P := \langle q_h, M_P = \varnothing, C_P = C(T_h^b, T_h^e, \mathbf{A}_h), \pi_P() \rangle$$

2) Productions of the form $H = q_h \rightarrow q[c]D$ include the first finding of the pattern:

$$P := \langle q_h, M_P = \{m_1^q\}, C_P \cup C(T_1^b, T_1^e, \mathbf{A}_1),$$
$$\pi_P(\mathbf{A}_1, T_1^b, T_1^e) \rangle$$

3) Productions of the form $D \rightarrow q[c]F \mid q[c]$ entail:

$$P := \langle q_h, M_P \cup \{m_k^q\}, C_P \cup C(T_k^b, T_k^e, \mathbf{A}_k),$$
$$\pi_P(\mathbf{A}_1, T_1^b, T_1^e, \ldots, \mathbf{A}_k, T_k^b, T_k^e) \rangle$$

4) Productions of the form $D \rightarrow \lambda$ finish the pattern without altering it.

This method generates a potentially infinite set of abstraction patterns from a single grammar, describing all the possible ways to conjecture an observable q_h from observations of lower abstraction levels. Moreover, it is possible to define adaptive observation procedures, since $\pi_P(\mathbf{A}_1, T_1^b, T_1^e, \ldots, \mathbf{A}_k, T_k^b, T_k^e)$ can be different at each step.

The following example shows an abstraction grammar to hypothesize a *heartbeat* (q_{hb}) from *P wave*, *QRS complex*, and *T wave* observations, by using common knowledge from an ECG manual[9]:

$$H = q_{hb} \rightarrow q_{Pw}[c_1]A \mid q_{QRS}[c_3, c_4]D \mid q_{QRS}[c_3, c_4, c_5]$$
$$A \rightarrow q_{QRS}[c_2, c_4]B$$
$$B \rightarrow q_{Tw}[c_6, c_7]$$
$$D \rightarrow q_{Tw}[c_7]$$
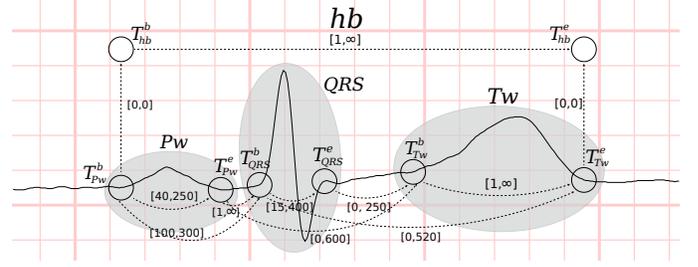
Fig. 2. Example of a *heartbeat* abstraction pattern

with the following constraints (all units are milliseconds):

$c_1 = \{\tau c(T_{Pw}^b, T_{Pw}^e) = [40, 250], \tau c(T_{Pw}^b, T_{hb}^b) = [0, 0]\}$

$c_2 = \{\tau c(T_{Pw}^b, T_{QRS}^b) = [100, 300], \tau c(T_{Pw}^e, T_{QRS}^b) = [1, \infty]\}$

$c_3 = \{\tau c(T_{hb}^b, T_{QRS}^b) = [0, 0]\}$

$c_4 = \{\tau c(T_{QRS}^b, T_{QRS}^e) = [15, 400]\}$

$c_5 = \{\tau c(T_{hb}^e, T_{QRS}^e) = [0, 0]\}$

$c_6 = \{\tau c(T_{Pw}^e, T_{Tw}^b) = [0, 600]\}$

$c_7 = \{\tau c(T_{QRS}^e, T_{Tw}^b) = [0, 250], \tau c(T_{QRS}^b, T_{Tw}^e) = [0, 520],$
$\quad\quad \tau c(T_{Tw}^b, T_{Tw}^e) = [1, \infty], \tau c(T_{Tw}^e, T_{hb}^e) = [0, 0]\}$

This grammar supports the generation of three abstraction patterns: The first one assembles the common observables of a normal heartbeat, i.e. a *P wave*, a subsequent *QRS complex*, and a final *T wave*, completing the full atrial and ventricular depolarization/repolarization cycle. The second pattern avoids the observation of a *P wave*, which may occur if the heartbeat has an activation focus outside the sinoatrial node. Finally, the third pattern only requires the observation of a *QRS complex*, which might be useful if the quality of the signal makes it difficult to properly observe the subsequent *T wave*.

In the set of constraints, c_1 represents the *P wave* duration limits, and sets the beginning of the heartbeat as the beginning of the *P wave*; c_2 represents the PR interval limits; c_3 sets the beginning of the heartbeat as the beginning of the *QRS complex* in absence of a preceding *P wave*; c_4 represents the *QRS complex* duration constraints; c_5 sets the end of the heartbeat as the end of the *QRS complex* in absence of a posterior *T wave*; c_6 limits the combined duration of the PR segment, the *QRS complex*, and the ST segment; c_7 sets the limits of the ST segment and the QT interval, and sets the end of the heartbeat as the end of the abstracted *T wave*. Figure 2 shows the most complete pattern that can be generated by this grammar.

For the ECG interpretation problem, we have defined an abstraction grammar for each one of the observables in table II, except for the signal deviation phenomena. This set of grammars $\mathcal{G}$ induces an *abstraction relation* on the set of observables $\mathcal{Q}$, such that we write $q_i \Vdash q_j$ if the observation of q_i allows us to conjecture the presence of q_j, that is, if there exists an abstraction pattern P, generated by some $G \in \mathcal{G}$, such that $q_j = q_{h_P}$ and $M_P^{q_i} \cap A_P \neq \varnothing$, where $M_P^{q_i}$ is the set of findings in P corresponding to the observable q_i. Moreover, to ensure the consistency of the knowledge base it is required that $q_i \nVdash^+ q_i$, where $\Vdash^+$ is the transitive closure of $\Vdash$.

The abstraction relation organizes the whole set of observables in a set of abstraction levels, as it is shown in table II. It is worth noting that, for any pattern P, the abstraction relation only involves the hypothesis q_h and the observables from the A_P set, so we can find in the set E_P different observables from the same or even higher abstraction levels than q_h.

Once the presence of a certain observable is hypothesized, it could provide an explanation for a set of observations from lower levels. Such a set of observations, behaving as an evidence set, is a matter of choice from an abductive point of view; think about deciding on the end of an electrocardiographic wave and the beginning of the next one. The result of this choice is represented by an *abstraction hypothesis*, which is defined as a tuple $h = \langle o_h, P_h, \leftharpoondown_h \rangle$, where o_h is an observation hypothesis conjectured by the abstraction pattern P_h through a *matching relation* $\leftharpoondown_h$. This matching relation injectively assigns a subset of existing observations from $\mathcal{O}$ to the findings in M_{P_h}. If this assignment is consistent with the constraints in C_{P_h}, the procedure π_{P_h} calculates the attribute and temporal location values for o_h. Once a hypothesis is found to be consistent, it may become part of the evidence for a new hypothesis of a higher abstraction level. Thus, all observations from a level above the first are generated as hypotheses, by abstracting a set of observations satisfying the constraints of an abstraction pattern generated by the grammar of the hypothesized observable. Finally, an *interpretation* is defined as a consistent set of abstraction hypotheses $I = \{h_1, ..., h_m\}$.

Solving an interpretation problem is posed as a heuristic search on the space of all consistent interpretations. This space has a tree structure, being the root the *trivial interpretation* I_0, containing no hypotheses. The CONSTRUE algorithm explores this space by expanding the most promising nodes according to four heuristic principles:

1) A *coverage principle*, which states the preference for interpretations explaining more initial observations.
2) A *simplicity principle*, which states the preference for interpretations with fewer abstraction hypotheses.
3) An *abstraction principle*, which states the preference for interpretations involving higher abstraction levels.
4) A *predictability principle*, which states the preference for interpretations that properly predict future evidence.

The expansion of a particular node is performed in the GET_DESCENDANTS() procedure, reproduced in algorithm 1. This procedure relies in different reasoning modes guided by a focus of attention f. Intuitively, the focus of attention points to the next observation or finding to be processed, and determines the next step in a hypothesize-and-test cycle. Lines 4-8 generate the descendants of an interpretation I when the attention is on an observation. These descendants are the result of two possible reasoning modes: the deduction of new findings, performed by the DEDUCE() function provided that the observation being focused on is an abstraction hypothesis; and the abduction of a new hypothesis explaining the observation being focused on, performed by the ABDUCE() function. A last descendant is obtained using the ADVANCE() function, which simply moves on the interpretation process by

declaring the focused observation as unintelligible according to the available knowledge.

On the other hand, if the focus points to a finding, then algorithm 1 obtains the descendants of the interpretation from the SUBSUME() and PREDICT() functions (line 10). The first one looks for an existing observation satisfying the constraints on the finding being focused on, while the second makes predictions about observables that have not yet been observed. All these reasoning modes, as long as the CONSTRUE algorithm implementation are detailed in [8].

Algorithm 1 Method for obtaining the descendants of an interpretation using different reasoning modes and guided by a focus of attention.

```
 1: function GET_DESCENDANTS(I)
 2:     var f = GET_FOCUS(I).TOP()
 3:     var desc = ∅
 4:     if IS_OBSERVATION(f) then
 5:         if f = o_h | h ∈ I then
 6:             desc = DEDUCE(I, f)
 7:         end if
 8:         desc = desc ∪ ABDUCE(I, f) ∪ ADVANCE(I, f)
 9:     else if IS_FINDING(f) then
10:         desc = SUBSUME(I, f) ∪ PREDICT(I, f)
11:     end if
12:     return desc
13: end function
```

For the ECG interpretation, the base evidence is a set of initial *Deflection* and *R-Deflection* observations. *R-Deflections* can be provided by any QRS detection method, like [17]. Regarding *Deflections*, the observation procedure operates on the raw ECG signal, and calculates a set of relevant intervals by thresholding the energy of the signal resulting from a multilevel 1-D wavelet decomposition/reconstruction of the ECG [18] using the 'haar' wavelet. The wavelet transform operates on fragments of 256 samples of the original signal, giving 8 decomposition levels, and performs a reconstruction by considering only the detail coefficients of levels 2 and 4, calculating the energy signal E by squaring the reconstructed signal. This energy signal focuses on frequency bands [11.25Hz, 22.5Hz] and [45Hz, 90Hz], matching the typically used spectrum for the detection of P waves, T waves, and QRS complexes [19]. Once the full energy signal has been calculated, the observation procedure π_{Def} provides a set of relevant intervals by filtering those time points t such that $E(t) > Thresh = quantile(E, 0.95^p)$, where $p \in [1, 20]$ is a parameter to control the sensitivity of the filter. This parameter is automatically modified during the interpretation. For example, if the PREDICT() procedure predicts an unobserved *Deflection*, p is increased until the *Deflection* is observed, or $p = 20$ is reached. If the distance between two time instants exceeding the threshold is lower than 20 ms, they are integrated in a single *Deflection* observation.

Figure 3 shows an example of an ECG fragment, its corresponding energy signal, the threshold for the selection of relevant points with $p = 2$, and the resultant *Deflections*. These intervals are the supporting evidence for a predicted

QRS complex or a P or T wave, but the specific time points of the beginning and the end of each one of them could be subsequently corrected along the interpretation procedure.
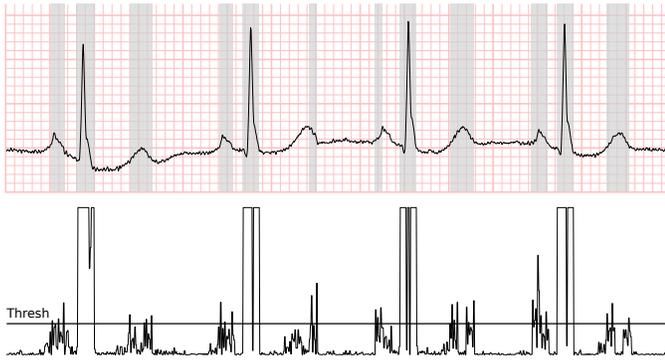
Fig. 3. Observation procedure to generate *Deflection* observations. At the bottom, the energy signal is shown, along with the threshold for relevant points detection with $p = 2$. [Source: MIT-BIH Arrhythmia DB, record: 101, between 00:16.000 and 00:19.500]

When the interpretation procedure begins, the first *R-Deflection* observation is focused, and the hypothesize-and-test cycle starts by trying to abduce it in a *QRS complex* observation. Eventually, this new hypothesis may lead to a consistent rhythm hypothesis, which will allow us to deduce new *QRS complexes* as long as the *P wave* and *T wave* observations corresponding to each cardiac cycle. The final target is to obtain a set of rhythm hypotheses explaining all the initial *R-Deflections* as *QRS complexes*, and maximizing the ECG *coverage* with the associated *P wave* and *T wave* observations abstracting the observed *Deflections*.

The reason to focus first on *R-Deflections* is that it gives more importance to the interpretation of relevant signal deviations. Furthermore, since the reasoning paradigm is non-monotonic, the initial set of *R-Deflections* set can be modified during the interpretation by including new predicted beats or by discarding the annotations that are considered false positives as they are declared unintelligible according to the available knowledge [20]. Hence, QRS detection errors do not necessarily lead to interpretation errors. Figure 4 shows an example of an ECG trace in which the initial set of *R-Deflections* has both false positive and false negative detections. The interpretation at rhythm level concludes that best hypothesis explaining the fragment is normal sinus rhythm, and therefore the third annotation is discarded as a false positive, and the actual third and fourth QRS complexes are predicted from the smaller signal deviations that appear consistently with the rhythm temporal constraints.

Figure 5 shows the full interpretation of another ECG fragment following the procedure described above. At the rhythm level, the fragment is explained as a *Sinus Rhythm* episode interrupted by an *Extrasystole* in the fourth beat, a return to *Sinus Rhythm* during five beats and a final *Trigeminy* episode during five beats. At the conduction level, all *QRS complexes* and *T waves* are properly delineated. With respect to *P waves*, the detection is accurate in all cases except the third beat in the trigeminy pattern in which it is missed. As can be seen, the final result of the interpretation procedure is
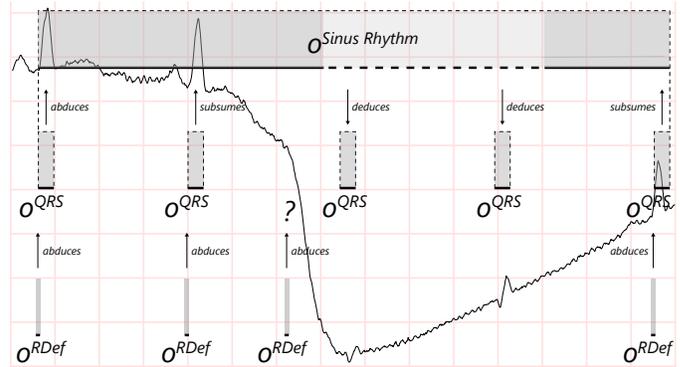
Fig. 4. How abductive interpretation can fix errors in the initial evidence [Source: MIT-BIH Arrhythmia DB, record: 121, lead: MLII, between 27:32.500 and 27:36.000. *R-Deflections* obtained with the `wqrs` application [17]]

an explanation of the physiological processes observed in the ECG record in the same terms used by experts. An abductive interpretation in abstraction levels will make it possible to adapt the sort of knowledge that can be found in any ECG handbook to a reduced set of classification rules, as described below.

IV. QRS CLUSTERING

Classification relies on a previous clustering task in order to exploit the high similarity between large number of beats that is invariably observed in ECG signals. A clustering algorithm $\psi : O(q_{QRS}) \rightarrow \mathbf{P}(O(q_{QRS}))$ should find a partition of the set of QRS observations $O(q_{QRS}) = \{o_1^{QRS}, \ldots, o_n^{QRS}\}$, satisfying the following requirements:

1) Maximum cluster purity. This will be achieved if no beats from different classes belong to the same cluster. All purity reductions below 100% will be directly translated to classification errors.
2) Minimum number of clusters. Ideally, the number of clusters should be equal to the number of beat classes present in the ECG recording. A higher number of clusters increase the uncertainty of the classification features and the number of decisions the classifier has to make.

In this work, we use the clustering method presented in [10], which performs an adaptive multi-lead context-based clustering of QRS complexes in real-time. This method creates a dynamic number of clusters represented by evolving templates to obtain sets of QRS complexes with similar rhythm and morphology. A high robustness to noise and morphological variability is achieved by means of a segment-based QRS complex characterization inspired by dominant points detection, a similarity measure that performs non-linear alignment using Dynamic Time Warping, and a noise-cluster proliferation control mechanism. As a result, the method provides good cluster purity while minimizing the number of clusters.

V. BEAT CLASSIFICATION

As a result of a QRS clustering we obtain a set $\psi(O(q_{QRS})) = \{Q_1, \ldots, Q_k\}$ of clusters. In the following
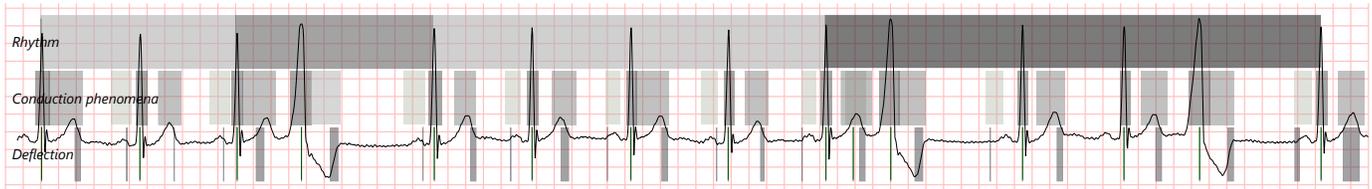
Fig. 5. Result of the abductive interpretation of an ECG fragment. [Source: MIT-BIH Arrhythmia, record: 116, lead: MLII, between 09:24.000 and 09:34.500]

we explain how to assign a label to each Q_i, classifying all the observations $o_j^{QRS} \in Q_i$ with the information resulting from the abductive interpretation of the ECG.

Heartbeat classification actually involves the determination of two different properties of the beat nature: 1) The physiological origin, i.e. the cardiac muscle area where the electrical activation begins; and 2) the temporal location of the beat with respect to the underlying cardiac rhythm. Usually, the origin distinguishes between supraventricular beats (generated at the atria or the atrioventricular node) and ventricular beats (generated at the ventricles), and the distinction is made based on morphological criteria. On the other hand, the temporal location distinguishes premature beats, escape beats and regular beats, that can be identified from rhythm analysis.

In general, practical electrocardiography makes a distinction between: 1) the interpretation of the "normal" electrocardiogram, and 2) the characterization of the possible transient anomalies [9]. The goal of the first task is to examine the most common situation observed in the subject's ECG, identifying the possible presence of permanent disorders. The goal of the second task is to study punctual or temporary changes with respect to the "normality" characterized in the first task.

This proposal reproduces this strategy by splitting the label assignment procedure in two stages. The first one, detailed in section V-A, considers a reduced set of individual features of each cluster Q_i and applies a set of general rules to determine the class of the beats in Q_i. But if the value of the individual features is not significant enough to decide the origin, the set of features is extended with additional features that are calculated by comparing the Q_i cluster with an already classified cluster Q_N that is assumed to represent a "normality" situation. This second classification stage is described in section V-B.

A. Single cluster classification

The first stage for cluster classification considers a reduced set of five qualitative features obtained from the interpretation results: *Heart Rate, QRS Duration, Heart Axis, P Wave* and *Rhythm*. Each QRS observation in the cluster has a specific value for each feature, and the values for the cluster are obtained by aggregating the values of all the QRS in the cluster for rhythm-related and P Wave features, and by taking the value of a QRS considered the representative of the cluster for morphological features. This representative is selected as the QRS observation with minimum distance to the mean duration, amplitude and heart axis of the cluster. Following we describe these features individually, and table III details the possible qualitative values each feature may take.

1) *Heart Rate*: This feature is calculated from the time distance of a QRS observation with respect to the previous QRS observation in the interpretation, what is called the instantaneous RR interval. For a cluster, this feature is calculated as the mean heart rate of all beats in the cluster.
2) *QRS Duration*: Represents the time interval between the onset and the offset of a QRS observation. If the delineation information is available in more than one ECG lead, the duration is taken as the distance between the earliest onset and the latest offset. For a cluster, the duration is calculated as the duration of the representative QRS of the cluster.
3) *Heart Axis*: Represents the mean direction of cardiac depolarization. To be accurately determined, it is necessary the information of all limb leads, specially I, II, and AVF. In this work, since the validation database only has one limb lead available (lead II), we obtain an approximation of the heart axis using only that lead, by calculating the amplitude relation between positive and negative waves in the QRS. The value of this approximation is in the range [-90º, +90º]. Fig. 6 shows an example of each possible qualitative value for this feature, and its corresponding numerical value. For a cluster, the heart axis is taken as the axis of the representative QRS. In the absence of lead II information, a balanced heart axis is assumed.

(a) Normal (+70º)　　(b) Balanced (+5º)　　(c) Inverted (-56º)
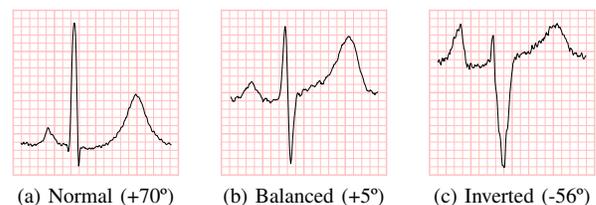
Fig. 6. Examples of heartbeats with different heart axis

4) *P Wave*: This feature indicates whether a QRS complex is preceded by a P wave in the interpretation result, and therefore has a sinus origin. To determine the presence or absence of a P wave in a cluster, for each available lead an amplitude histogram of all *P wave* observations is calculated, with the bins set in 0 µV, 50 µV, 100 µV, 500 µV and 1 mV. Then, the peak of the histogram is taken as the P wave amplitude in the lead, and finally it is required the mean amplitude of all leads to be higher than 50 µV. For example, consider the interpretation in Fig. 5 and suppose a cluster will all but the 4th, 10th, and 13th heartbeats. The amplitude of the P waves associated to the clustered heartbeats are 0µV (absent P wave), 42µV, 58µV, 56µV, 49µV, 54µV, 54µV, 58µV, 58µV, 0µV, 49µV. If we

calculate the amplitude histogram, the bin [0μV, 50μV] has 5 elements, and the bin [50μV, 100μV] has 6 elements, so in the absence of another lead the *P Wave* feature has a value of 1 for this cluster.

5) *Rhythm*: This feature is obtained from the highest abstraction level of the interpretation procedure: the rhythm hypothesis in which a QRS complex is enclosed. The *Regular* value is assigned to all beats interpreted in a regular rhythm (Bradycardia, Sinus rhythm or Tachycardia), to the odd beats in Bigeminisms, to the first and second beats of each triple in Trigeminisms, and to the beat after a compensatory pause in Extrasystoles and Couplets. The *Advanced* value is assigned to beats interpreted as Extrasystoles or Couplets, as long as the even beats in Bigeminisms and the third beat of each triple in Trigeminisms. The *Atrial fibrillation* value is assigned to all beats interpreted in an atrial fibrillation episode. Finally, the *Delayed* value is assigned to the beats interpreted as Asystole or as Rhythm Block. For a cluster, this feature is calculated as the most repeated rhythm. If the *Advanced* or *Delayed* tags have a significant frequency in the cluster (over 20%), a correction factor is applied based on the number of *Regular* beats with a higher heart rate (lower in the case of *Delayed*) than the mean heart rate of the cluster. By default, it is assumed that 50% of *Regular* beats have a higher heart rate than the mean. The number of beats with higher heart rate exceeding this 50% is then added to the number of *Advanced* beats, and if this count exceeds the number of *Regular* beats, the cluster is tagged as *Advanced*. For example, for the interpretation shown in Fig. 5 and according to these rules, the 4th, 10th and 13th heartbeats have the *Advanced* value for the *Rhythm* feature, while the others have the *Regular* value.

TABLE III
FEATURE VALUES FOR SINGLE CLUSTER CLASSIFICATION

Feature	Qualitative Values
Heart Rate	-1: Lower than 60 beats per minute (Bradycardia)
	0: Between 60 and 100 bpm (Normal rhythm)
	1: Between 100 and 150 bpm (Asymptomatic tachycardia)
	2: Over 150 bpm (Symptomatic tachycardia)
QRS Duration	-1: Less than 80 milliseconds (Narrow)
	0: Between 80 ms and 100 ms (Normal)
	1: Between 100 ms and 120 ms (Abnormal)
	2: More than 120 ms (Wide)
Heart Axis	-1: Angle lower than -45º (Inverted axis)
	0: Angle between -45º and 45º (Balanced axis)
	1: Angle higher than 45º (Normal axis)
P Wave	0: Absent P wave, or mean amplitude lower than 50 μV
	1: Mean amplitude over 50 μV
Rhythm	0: Regular
	1: Atrial fibrillation
	2: Advanced beat
	3: Delayed beat

Once these qualitative features have been obtained for each cluster, a simple set of general rules is applied in order to decide the physiological origin of the beats in the cluster. To ensure robustness in this step, which determines the reliability of the comparative classification, only the clusters with a significant number of members are considered. In our case, we require as a rule of thumb a minimum of 30 QRS to ensure a stable mean value for the classification features. Table IV shows the four classification rules considering only the individual features of each cluster. The rules are applied in order, so if the antecedent of one rule is met no further rules are evaluated.

TABLE IV
SINGLE CLUSTER CLASSIFICATION RULES

$Rhythm = Regular$ and $Pwave$	$\rightarrow$	NORMAL
$Rhythm = Atrial\ fib.$ and $Rate \geq 0$	$\rightarrow$	AFIB
$Rhythm = Advanced$ and $Duration = Narrow$	$\rightarrow$	SVEB
$Pwave$ and $Duration = Narrow$	$\rightarrow$	NORMAL

Clusters classified as NORMAL by rules in table IV are further analyzed to identify possible intraventricular conduction abnormalities such as right bundle branch block (RBBB) and left bundle branch block (LBBB) [9]. These conditions are detected from the representative QRS morphology, with RBBB requiring the QRS duration to be higher than 100 ms and the QRS morphology in lead V1 to finish with a positive wave with amplitude higher than 0.5 mV. LBBB requires the QRS duration to be higher than 120 ms, and the QRS morphology in lead V1 to be tagged as *QS* or *rS*, according to [16].

In the MIT-BIH Arrhythmia Database used for the validation of the proposal, the single cluster classification is able to provide a label for more than 84% of the heartbeats.

B. Comparative classification

In some records, the variability of the cardiac conduction and rhythm makes it impossible to distinguish a predominant "normality" situation satisfying the rules described in section V-A. But even in such circumstances, establishing a baseline cardiac behavior will permit to discriminate the possible transient anomalies by comparison. In our case, if no clusters were classified by the single cluster classification rules, the reference cluster Q_N is selected as follows:

- If there are clusters with more than 30 QRS, with a P wave, and with a duration lower than 120 ms, then the largest one is selected as Q_N.
- If there are clusters with a predominant *Regular* or *Atrial fibrillation* rhythm label, the largest cluster in this set is selected.
- If no clusters match any of these two criteria, the cluster with more complexes interpreted as *Regular* or *Atrial fibrillation* is selected.

The comparative classification stage considers an extended set of features, calculated by comparison between the target unclassified cluster Q_i and the reference cluster Q_N, which are described below. The qualitative values for these features are detailed in table V.

1) *Heart rate difference (Heart Rate$'$)*: Measures the heart rate ($Rate$) difference between the target cluster Q_i and a reference cluster Q_N.

2) *Duration difference (QRS Duration$'$)*: Measures the duration (Dur) difference between the representative QRS complexes of Q_i and Q_N.

3) *Axis difference (Heart Axis$'$)*: Measures the heart axis ($Axis$) difference between the representative QRS complexes of Q_i and Q_N.

4) *Amplitude difference (QRS Amplitude$'$)*: This feature measures the relative amplitude difference between the representative QRS of Q_i and Q_N. For this, the lead in which Q_i has a higher amplitude is selected, and the ratio of both amplitudes is calculated. The amplitude is calculated as the difference between the maximum and minimum ECG values inside the region delineated as QRS complex.

5) *Morphological Similarity*: Provides a measure of the signal similarity between the QRS onset and offset in all available leads for the representative QRS of Q_i and Q_N. Similarity is obtained by cross-correlation (*xcorr*), and the result is the average value of the maximum for each lead.

TABLE V
FEATURE VALUES FOR COMPARATIVE CLASSIFICATION

Feature	Qualitative Values
Heart Rate$'$	-1: $Rate_N - Rate_i \geq 20bpm$ (Slower rate)
	0: $-20bpm < Rate_N - Rate_i < 20bpm$ (Equal rate)
	1: $Rate_i - Rate_N \geq 20bpm$ (Faster rate)
QRS Duration$'$	-1: $Dur_N - Dur_i \geq 20ms$ (Narrower complex)
	0: $-20ms < Dur_i - Dur_N < 20ms$ (Equal duration)
	1: $20ms \leq Dur_i - Dur_N < 40ms$ (Wider complex)
	2: $Dur_i - Dur_N \geq 40ms$ (Much wider complex)
Heart Axis$'$	0: $|Axis_N - Axis_i| < 45°$ (Equal axis)
	1: $45° \leq |Axis_N - Axis_i| < 90°$ (Deviated axis)
	2: $90° \leq |Axis_N - Axis_i| < 135°$ (Far deviated axis)
	3: $|Axis_N - Axis_i| \geq 135°$ (Opposite axis)
QRS Amplitude$'$	-1: $Amp_i/Amp_N < 0.75$ (Lower amplitude)
	0: $0.75 \leq Amp_i/Amp_N \leq 1.25$ (Equal amplitude)
	1: $Amp_i/Amp_N > 1.25$ (Higher amplitude)
Morphological Similarity	0: $xcorr(Q_N, Q_i) < 0.25$ (Very different)
	1: $0.25 \leq xcorr(Q_N, Q_i) < 0.5$ (Different)
	2: $0.5 \leq xcorr(Q_N, Q_i) < 0.75$ (Similar)
	3: $0.75 \leq xcorr(Q_N, Q_i) < 0.9$ (Very similar)
	4: $xcorr(Q_N, Q_i) \geq 0.9$ (Identical)

On the basis of the attributes of the clusters labeled in the basic classification stage, two contexts requiring some specific rules for the classification of Q_i are identified:

a) *Wide QRS context:* This context is set on the presence of an artificial pacemaker or if a bundle branch block (left or right) was identified in some cluster. The presence of an artificial pacemaker is set if for some cluster more than 20% of beats are classified as paced by the QRS delineation algorithm [16].

b) *Atrial fibrillation context:* This context is set if some cluster was classified as *Atrial fibrillation*.

If none of these specific contexts are recognized, then a Normal Sinus Rhythm context is assumed, and the reference cluster Q_N is selected as the largest cluster classified as normal.

The comparative classification tries first to account for the physiological origin of the cluster, considering the morphological features, and then the temporal location using the rhythm features. If there is an already classified cluster Q_j with identical similarity and equal amplitude than Q_i, then the supraventricular/ventricular origin of Q_i is assumed to be the same of Q_j. If no such Q_j exists, then the origin is determined by considering typical values for the seven morphological features and checking which type is best matched with the feature values of Q_i through a simple majority vote. In the unusual case of tie vote, a supraventricular origin is assumed. Table VI shows these characteristic values for each origin.

TABLE VI
CHARACTERISTIC VALUES FOR SUPRAVENTRICULAR/VENTRICULAR DISCRIMINATION

S	$Axis \geq 0$	$Duration < 2$	$Pwave \geq 0$	$Similarity > 2$
	$Axis' = 0$	$Duration' < 2$	$Amplitude' \leq 0$	
V	$Axis \leq 0$	$Duration > 0$	$Pwave = 0$	$Similarity < 3$
	$Axis' > 0$	$Duration' > 0$	$Amplitude' \neq 0$	

Table VII shows the classification rules used for each final resulting class, considering the physiological origin and rhythm-related features. As can be seen, each specific context modulates the rules according to the special characteristics of "normality" they represent. For example, in the *Atrial fibrillation* context classification rules avoid the use of $Rate$ and $Rate'$ features, due to the highly erratic behavior of the RR interval [9]. On the other hand, in the *Wide QRS* context, clusters with regular rhythm and a QRS morphology similar to Q_N (high morphological similarity or low axis deviation) are classified as normal, and clusters with advanced rhythm and appreciable morphological differences with respect to Q_N (in duration or similarity) are classified as ectopic ventricular beats. Finally, the *Normal sinus rhythm* context includes a rule for the classification of fusion beats in clusters with an identified supraventricular origin. The reason not to include this rule in the *Atrial fibrillation* and *Wide QRS* contexts is its dependence with the *Duration* and *Rate* features, that are not representative in these contexts.

VI. RESULTS

In the bibliography, the commonly used methodology for heartbeat classification evaluation is the presented by de Chazal *et al.* [14], focused on the MIT-BIH Arrhythmia database. This methodology excludes the records of patients using pacemakers and divides the remaining into two datasets, DS1 (training) and DS2 (test). In our case, a training set is not necessary and the validation was performed with all records, including those with paced beats. In this manner, we can evaluate the interpretation and classification algorithms without making any a priori assumptions about the characteristics of the ECG signal. Like other works following this methodology, the beat annotations in the .atr files included in the database were used as the initial evidence for the interpretation stage, considering them as *R-Deflection* observations. It should be noted that these annotations contain perfect knowledge about

TABLE VII
RULES FOR COMPARATIVE CLASSIFICATION

Wide QRS and Normal sinus rhythm contexts

$Origin = S$	
$Rate = 0$ and $Duration = 0$ and $Similarity = 4$	$\rightarrow$ NORMAL
$Rhythm = 2$ or $Rate' = 1$ or $Rate = 2$	$\rightarrow$ SVEB
$Rhythm = 3$ and $Rate' = -1$	$\rightarrow$ SVESC
$Default$	$\rightarrow$ NORMAL
$Origin = V$	
$Rhythm = 2$ or $Rate' = 1$ or $Rate = 2$	$\rightarrow$ VEB
$Rhythm = 3$ or $Rate \leq 0$	$\rightarrow$ VESC
$Rhythm = 0$ and $Duration = 1$ and $Rate = 1$	$\rightarrow$ FUSION
$Default$	$\rightarrow$ VEB

Atrial fibrillation context

$Origin = S$	
$Rhythm = 0$ and $Duration = 0$ and $Similarity = 4$	$\rightarrow$ NORMAL
$Rhythm = 2$	$\rightarrow$ SVEB
$Rhythm = 3$	$\rightarrow$ SVESC
$Default$	$\rightarrow$ NORMAL
$Origin = V$	
$Rhythm = 2$	$\rightarrow$ VEB
$Rhythm = 3$ or $Rate \leq 0$	$\rightarrow$ VESC
$Default$	$\rightarrow$ VEB

Wide QRS context

$Rhythm = 0$ and $Duration' = 0$ and $Similarity > 2$	$\rightarrow$ NORMAL
$Rhythm = 0$ and $Duration' = 0$ and $Axis' < 2$	$\rightarrow$ NORMAL
$Rhythm = 2$ and $Duration > 0$ and $Duration' > 0$	$\rightarrow$ VEB
$Rhythm = 2$ and $Duration > 0$ and $Similarity < 4$	$\rightarrow$ VEB

Normal sinus rhythm context

$Origin = S$	
$Rhythm = 0$ and $Duration = 1$ and $Rate = 1$	$\rightarrow$ FUSION

Fig. 7. False negative VEB detection due to an oversimplified interpretation [Record: 202, lead: MLII, between 27:58.600 and 28:02.100. *R-Deflections* are the reference `.atr` annotations]

TABLE VIII
CONFUSION MATRICES OF THE BEAT CLASSIFICATION RESULTS FOR ALL 48 RECORDS IN THE MIT-BIH ARRHYTHMIA DATABASE (TOP), FOR THE 44 RECORDS IN THE DS1 AND DS2 SETS (MIDDLE), AND FOR THE 22 RECORDS IN THE DS2 TEST SET (BOTTOM)

Ground Truth		Classification Result				
		O	N	S	V	F	Q
	O	0	19	6	49	0	0
	N	115	89459	355	446	9	2
	S	15	330	2575	106	0	0
Full database	V	85	275	116	6705	41	14
	F	1	273	5	80	444	0
	Q	2	713	0	143	1	7184

Ground Truth		Classification Result				
		O	N	S	V	F	Q
	O	0	19	6	49	0	0
	N	115	88960	355	441	9	0
	S	15	330	2575	106	0	0
DS1+DS2	V	73	273	116	6506	41	0
	F	1	273	5	80	444	0
	Q	0	13	0	1	1	0

Ground Truth		Classification Result				
		O	N	S	V	F	Q
	O	0	2	0	0	0	0
	N	84	43634	277	47	4	0
	S	10	199	1787	54	0	0
DS2	V	37	50	64	3048	22	0
	F	0	140	5	42	201	0
	Q	0	7	0	0	0	0

QRS locations, but the interpretation algorithm treats them like any other initial evidence that might be modified during the hypothesize-and-test cycle. For example, Fig. 7 shows an ECG fragment that is interpreted as an *Atrial fibrillation* episode, and the 4th *R-Deflection* is incorrectly abstracted in a *T wave* instead of a *QRS complex*, leading to a false negative in the 'V' class. Even if it is detrimental to the results, we believe this is the most fair validation strategy, since the main distinctive feature of the adopted reasoning paradigm is to not assume any previous conclusion as unfailing.

Table VIII shows the confusion matrices of the classification results for all the 48 records in the database, for the 44 records in the DS1 and DS2 sets, and for the records in the DS2 dataset commonly used in comparison studies. These matrices were obtained as described in section II, fixing the incorrect label association for MIT-BIH classes 'j' and 'e'.

In addition to the five standard beat classes defined by the AAMI, the **O** class has been included to represent false positives and false negatives in QRS detection due to the modifications introduced by the abductive interpretation process in the initial set of beat annotations. The **Q** class represents beats with unknown origin, but the standard tools also use it to represent paced beats and the fusion of paced and normal beats. This is the reason for the great proportional difference

of this class in the first matrix with respect to the other.

Table IX shows a comparison of the classification performance with the most relevant algorithms of the state-of-the-art, using sensitivity and positive predictive value of the ventricular and supraventricular ectopic beat classes. It is worth noting that most of the proposals are assisted, meaning that they require an expert to manually annotate some beats. In this sense, some of the assisted methods are closer to a *clustering* approach rather than a *classification* approach, since once the expert has labeled one of the beats in each group the remaining beats are assigned the same label. In contrast, the present

TABLE IX
VEB AND SVEB CLASSIFICATION PERFORMANCE AND COMPARISON WITH THE MOST RELEVANT AUTOMATIC AND ASSISTED METHODS OF THE STATE-OF-THE-ART

Dataset	Method	VEB		SVEB	
		Se	P^+	Se	P^+
DS1+DS2	This work - Automatic	92.82	92.23	85.10	84.51
	Llamedo *et al.* [5] - Assisted	90±1	97±0	89±2	88±3
	Kiranyaz *et al.* [4] - Assisted	93.9	90.6	60.3	63.5
	Ince *et al.* [21] - Assisted	84.6	87.4	63.5	53.7
	Llamedo *et al.* [5] - Automatic	80±2	82±3	76±2	43±2
DS2	This work - Automatic	94.63	96.79	87.17	83.98
	Llamedo *et al.* [5] - Assisted	93±1	97±1	92±1	90±3
	Kiranyaz *et al.* [4] - Assisted	95.0	89.5	64.6	62.1
	Oster *et al.* [7] - Assisted	92.7	96.2	NA	NA
	Chazal *et al.* [6] - Assisted	93.4	97.0	94.0	62.5
	Chazal *et al.* [2] - Assisted	94.3	96.2	87.7	47.0
	Zhang *et al.* [22] - Automatic	85.48	92.75	79.06	35.98
	Llamedo *et al.* [5] - Automatic	89±1	87±1	79±2	46±2
	Chazal *et al.* [14] - Automatic	77.7	81.9	75.9	38.5

proposal provides a method that autonomously assigns a label to each and every beat in a record.

Results show that this proposal outperforms any other automatic method in the bibliography, and even most of the assisted methods. The only algorithm with better performance is the assisted version of [5], although it would be desirable to perform a statistical test with the record-by-record results to determine whether the differences are relevant.

In general, the most remarkable improvement with respect to the state-of-the-art is the classification of supraventricular ectopic beats (S class). The main difficulty with this type of beat is that the morphology is usually very similar to the normal morphology, so class separation has to be made according to rhythm-related features. In this sense, the information provided by the rhythm level of the abductive interpretation, and in particular the *Rhythm* feature detailed in section V-A, goes far beyond than the classical analysis of the RR intervals around the beat, which is severely affected by phenomena such as atrial fibrillation. Also, it should be noted that results of other approaches in table IX are affected by the incorrect label association discussed in section II, so the actual sensitivity in the classification of the S class is expected to be lower.

VII. CONCLUSION

We propose a novel knowledge-based approach to the heartbeat classification problem, grounded on the abductive interpretation of the signal. This interpretation produces a set of observations in multiple abstraction levels, as a result of conjecturing the set of physiological processes underlying ECG signal, and expressed in the sort of terms used by experts, enabling a subsequent classification by mimicking the common rules from any electrocardiography handbook.

The proposal has been validated following the AAMI recommendations and compared with the most relevant works of the state-of-the-art. With respect to automatic approaches, results show an average improvement of around 10% in the sensitivity and positive predictive value of ventricular ectopic beats and in the sensitivity of supraventricular ectopic beats. This improvement is increased to around 40% in the positive predictive value of supraventricular ectopic beats. With respect to assisted approaches, the improvement in the sensitivity and positive predictive value of ventricular ectopic beats is around 1%. For supraventricular ectopic beats, the average improvement in sensitivity and positive predictive value is around 8% and 20%, respectively. The statistical significance of these differences cannot be assessed due to the lack of detailed validation results for the other methods, but the magnitude of the differences, specially in the specificity of the supraventricular class, is notable beyond any doubt.

The key factor behind these results is the non-monotonic nature of the hypothesize-and-test cycle, making it possible to exploit the complementariness between bottom-up and top-down processing to find the best explanation consistent with the evidence. As in perception, ECG interpretation is assumed to be mostly bottom up, although top-down processing has proved to be decisive to cope with noise, artifacts or ambiguities in the signal. However, as a pure knowledge-based approach it is highly dependent on the quality of the abstraction grammars describing the domain knowledge, and in some cases the heuristic principles guiding the search may lead to incorrect interpretations. For example, the oversimplified result shown in Fig. 7 is due to the predominance of the *simplicity* principle in a scenario requiring a more complex description.

This work demonstrates that there is still room for improvement in automatic ECG heartbeat classification. This is essential to achieve an effective transfer of these techniques to the clinical routine, integrated in the bedside instrumentation or in the emergent home monitoring. To this end, the evolution of the present work will focus on three main objectives: On the one hand, beat labeling should be more strongly integrated in the hypothesize-and-test cycle. In the present proposal, classification is performed at the end, after the interpretation makes the necessary features available. A better result could be expected if every class label for each QRS observation is considered as a conjecture, enabling to correct it on the basis of posterior evidence, at the expense of a greater complexity of reasoning. On the other hand, the scenarios in which a fully automatic processing of the ECG makes a difference are mainly related with continuous monitoring. In general, these scenarios require an online interpretation strategy, capable of providing results while data acquisition is ongoing. This poses a deep change in our approach, since notions like "normality context" for comparative classification are no longer static, and they have to be updated during the interpretation process. Finally, we aim to meet real-time constraints in the execution of the full interpretation cycle. For this, it is necessary to sacrifice completeness in the traverse of the search space performed by the CONSTRUE algorithm, and to improve the computing performance of the implementation. At this moment we are exploring a pruning strategy following the principles of the K-Beam algorithm [23], with promising results. This will allow us to assess the applicability of the method to more complex problems, including ECG databases with more than two leads and multi-modal signal databases.

IMPLEMENTATION

With the aim of supporting reproducible research, the source code of the interpretation and classification algorithms presented in this paper has been published under an open source license[1], along with the ECG knowledge base including the abstraction grammars for the observables in table II, and all the necessary files and instructions to reproduce the results. Also, in this repository there are available some representative ECG fragments to reproduce the interpretation procedure interactively, allowing to explore the interpretation tree space and the reasoning steps leading to the final solution.

REFERENCES

[1] E. J. S. Luz, W. R. Schwartz, G. Cámara-Chávez, and D. Menotti, "ECG-based Heartbeat Classification for Arrhythmia Detection: A Survey," *Computer Methods and Programs in Biomedicine*, 2016.

[2] P. de Chazal and R. B. Reilly, "A Patient-Adapting Heartbeat Classifier Using ECG Morphology and Heartbeat Interval Features," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 12, pp. 2535–2543, Dec 2006.

[3] H. Hu, S. Palreddy, and W. Tompkins, "A Patient-Adaptable ECG Beat Classifier Using a Mixture of Experts Approach," *IEEE Transactions on Biomedical Engineering*, vol. 44, no. 9, pp. 891–900, 1997.

[4] S. Kiranyaz, T. Ince, and M. Gabbouj, "Real-Time Patient-Specific ECG Classification by 1-D Convolutional Neural Networks," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 3, pp. 664–675, March 2016.

[5] M. Llamedo and J. P. Martinez, "An Automatic Patient-Adapted ECG Heartbeat Classifier Allowing Expert Assistance," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 8, pp. 2312–2320, Aug 2012.

[6] P. de Chazal, "An adapting system for heartbeat classification minimising user input," in *36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Aug 2014, pp. 82–85.

[7] J. Oster, J. Behar, O. Sayadi, S. Nemati, A. E. W. Johnson, and G. D. Clifford, "Semisupervised ECG Ventricular Beat Classification With Novelty Detection Based on Switching Kalman Filters," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 9, pp. 2125–2134, Sept 2015.

[8] T. Teijeiro and P. Félix, "On the adoption of abductive reasoning for time series interpretation," 2016, arXiv:1609.05632.

[9] G. S. Wagner, *Marriott's Practical Electrocardiography*, 11th ed. Wolters Kluwer Health/Lippincott Williams & Wilkins, 2008.

[10] D. Castro, P. Félix, and J. Presedo, "A Method for Context-Based Adaptive QRS Clustering in Real Time," *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 5, pp. 1660–1671, Sept 2015.

[11] G. B. Moody and R. G. Mark, "The impact of the MIT-BIH Arrhythmia Database," *IEEE Engineering in Medicine and Biology Magazine*, vol. 20, no. 3, pp. 45–50, May 2001.

[12] A. L. Goldberger *et al.*, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals," *Circulation*, vol. 101, no. 23, pp. 215–220, Jun. 2000.

[13] AAMI, "American National Standard for Ambulatory Electrocardiographs, publication ANSI/AAMI EC38-2007," 2007.

[14] P. D. Chazal, M. O'Dwyer, and R. B. Reilly, "Automatic classification of heartbeats using ECG morphology and heartbeat interval features," *IEEE Transactions on Biomedical Engineering*, 2004.

[15] M. Dechter, J. Meiri, and J. Pearl, "Temporal constraint networks," *Artificial Intelligence*, vol. 49, pp. 61–95, 1991.

[16] T. Teijeiro, P. Félix, and J. Presedo, "A noise robust QRS delineation method based on path simplification," in *2015 Computing in Cardiology Conference (CinC)*, Sept 2015, pp. 209–212.

[17] W. Zong, G. Moody, and D. Jiang, "A robust open-source algorithm to detect onset and duration of QRS complexes," in *Computers in Cardiology*, 2003, pp. 737–740.

[18] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674–693, Jul 1989.

[19] J. P. Martínez *et al.*, "A wavelet-based ECG delineator: evaluation on standard databases." *IEEE transactions on Biomedical Engineering*, vol. 51, no. 4, pp. 570–81, Apr. 2004.

[20] T. Teijeiro, P. Felix, and J. Presedo, "Using Temporal Abduction for Biosignal Interpretation: A Case Study on QRS Detection," in *2014 IEEE International Conference on Healthcare Informatics*. IEEE, Sep. 2014, pp. 334–339.

[21] T. Ince, S. Kiranyaz, and M. Gabbouj, "A Generic and Robust System for Automated Patient-Specific Classification of ECG Signals," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 5, pp. 1415–1426, May 2009.

[22] Z. Zhang, J. Dong, X. Luo, K.-S. Choi, and X. Wu, "Heartbeat classification using disease-specific feature selection," *Computers in Biology and Medicine*, vol. 46, pp. 79–89, 2014.

[23] S. Edelkamp and S. Schrödl, *Heuristic Search: Theory and Applications*. Morgan Kaufmann, 2011.

Tomás Teijeiro received the B.Sc. degree in computer engineering and the M.Sc. degree in research in information technologies from the University of Santiago de Compostela, Spain, in 2010 and 2011, respectively. He is a Researcher and currently working toward the Ph.D. degree at the Centro Singular de Investigacion en Tecnoloxías da Información, University of Santiago de Compostela. His research interests include knowledge representation, non-monotonic temporal reasoning, and their application to time series abstraction and interpretation.

Paulo Félix received the B.Sc. and Ph.D. degrees in physics from the University of Santiago de Compostela, Santiago de Compostela, Spain, in 1993 and 1999, respectively. He is currently a Researcher at the Centro Singular de Investigacion en Tecnoloxías da Información, University of Santiago de Compostela. His research interests include temporal reasoning, machine learning, and signal processing.

Jesús Presedo received the B.Sc. and Ph.D. degrees in physics from the University of Santiago de Compostela, Santiago de Compostela, Spain, in 1989 and 1994, respectively. He is currently a Researcher at the Centro Singular de Investigacion en Tecnoloxías da Información, University of Santiago de Compostela. His research interests include biomedical digital signal processing, heart rate variability, nonlinear dynamics, soft computing, and the development of ubiquitous healthcare systems.

Daniel Castro received the B.Sc. and M.Sc. degrees in physics from the University of Santiago de Compostela, Santiago de Compostela, Spain, in 1998 and 1999, respectively. He is a Researcher and currently working toward the Ph.D. degree at the Centro Singular de Investigacion en Tecnoloxías da Información, University of Santiago de Compostela. His research interests include signal processing and its application to biomedical signals.

[1]https://github.com/citiususc/construe