An APN model for Arrhythmic beat classification

Hsiu-Sen Chiang¹, Dong-Her Shih²,⁎, Binshan Lin³ and Ming-Hung Shih⁴
¹Department of Information Management, National Taichung University of Science and Technology, 129, Section 3, Sanmin Road, Taichung City 404, Taiwan, ²Department of Information Management, National Yunlin University of Science and Technology, 123, Section 3, University Road, Douliu City, Yunlin County, Taiwan, ³College of Business Administration, BE321, Louisiana State University in Shreveport, Shreveport, LA 71115, USA and ⁴Department of Electrical and Computer Engineering, Iowa State University, Ames, IA 50011, USA

Associate Editor: John Hancock

ABSTRACT
Motivation: Changes in the normal rhythm of a human heart may result in different cardiac arrhythmias, which may be immediately fatal or cause irreparable damage to the heart sustained over long periods of time. Therefore, the ability to automatically identify arrhythmias from ECG recordings is important for clinical diagnosis and treatment. In this article, classification by using associative Petri net (APN) for personalized ECG-arrhythmia-pattern identification is proposed for the first time in literature.

Results: A rule-based classification model and reasoning algorithm of APN are created for ECG arrhythmias classification. The performance evaluation using MIT-BIH arrhythmia database shows that our approach compares well with other reported studies.

Contact: shihdh@yuntech.edu.tw

Received on October 8, 2013; revised on January 22, 2014; accepted on February 6, 2014

1 INTRODUCTION
Heart disease is an umbrella term for a number of different diseases which affect the heart such as arrhythmia, myocardial ischemia and myocardial infarction (Minino et al., 2007). As of 2007, it is also one of the leading causes of death in the world, and especially according to the American Heart Association, One American is dying in every 34 s because of this disease (American Heart Association, 2008). Changes in the normal rhythm of a human heart may result in different cardiac arrhythmias, which may be immediately fatal or cause irreparable damage to the heart when sustained over long periods of time. Electrocardiogram (ECG) is an important routine clinical practice for continuous monitoring of cardiac activities. Since ECG signals vary greatly for different individuals and within patient groups (Hoekema et al., 2001), ECG-pattern recognition is a difficult problem even with the aid of a computer. While normal sinus rhythm originates from the sinus node of heart, arrhythmias have various origins and indicate a wide variety of heart problems. Same symptoms of arrhythmia produce different morphologies due to their origins such as premature ventricular contraction (Minami et al., 1999). For effective diagnostics, the study of ECG pattern and heart-rate-variability signal may have to be carried out over several hours. Thus, computer-based analysis and classification of diseases can be very helpful in diagnostics (Güler and Übeyli, 2005; Saxena et al., 2002). A number of methods have been proposed to classify ECG-heartbeat patterns based on the features extracted from ECG signals. APNs have knowledge representation ability and could be translated into production rule systems. In addition, the APN formalism also is a graphical and mathematical tool for design, specification, simulation and verification of systems. Therefore, APN can be widely used in computer, knowledge-based systems, process control, decision making as well as other kinds of engineering applications. In this article, classification of ECG arrhythmias by using associative Petri net (APN) is proposed for the first time in the literature and is compared with several pioneering studies for arrhythmias detection.

2 RELATED WORK
An ECG is a graphic produced by an electrocardiograph. A single normal cycle of the ECG represents the successive arterial depolarization and ventricular repolarization, and can be approximately associated with the peaks and other ECG waveforms, which labeled P, Q, R, S and T (McSharry et al., 2003). Recently, different characteristic, such as P, T, U waves and PQ, QRS, ST segments, are also used for diagnostics (Singh and Tiwari, 2006). Automated classification of heartbeats has been previously reported by other investigators using a variety of features to represent the ECG and a number of classification methods. Features include RR-Interval features, heartbeat-interval features and ECG-morphology features (Philip et al., 2004), ST-segment deviation, -segment slope (STS), -segment area (STA), T-normal amplitude (Exarchos et al., 2006), STA, R–S interval (RSI), STS, R–T interval (RTI), QRS area (QRSA), Q–T interval (QTI), R-wave amplitude (RWA), heart-rate beat (HBR), statistical features QRS energy (QRSE), mean of the power spectral density (MPSD), auto-correlation coefficient (ACC), signal histogram (SH) (Gholam et al., 2006), discrete Fourier transform coefficients (DFT) (Dokur and Ölmez, 2001), statistical features of the QRS complexes (Osowski and Linn, 2001), Hermite coefficients (Linh et al., 2003), shift-invariant (Yeong et al., 2006), continuous wavelet transform coefficients (Andreao et al., 2006), QRS complex wave width (Homaiezhad et al., 2012; Martits et al., 2013; Yeha et al., 2012; Zhou et al., 2005), amplitude value, DCT coefficients, DWT coefficients (Acir, 2006) and DWT coefficients (Abawajy et al., 2001).
et al., 2013; Engin, 2004; Inan and Übeyl, 2005). In addition, the classification of heartbeats employed include linear discriminants (Philip et al., 2004), association rules (Exarchos et al., 2006), neural network (1 Dokur and Ölmex, 2001; Gholam et al., 2006; Inan and Übeyl, 2005), fuzzy neural network (Engin, 2004; Lín et al., 2003; Osowski and Linh, 2001; Yeong et al., 2006), hidden Markov models (Andreao et al., 2006), mirrored Gauss model (Zhou et al., 2005), artificial immune-recognition system (Polat et al., 2006), support vector machine (Acir, 2006) and dynamic time warping (Zhang et al., 2009), cluster analysis (Yeha et al., 2012), Neuro-SVM–KNN fusion (Homaeezehad et al., 2012), LibLINEAR, LibSVM (Abawajy et al., 2013) and Principal component analysis, Linear discriminant analysis, Independent component analysis (Martis et al., 2013).

Recent studies developed new feature-extracted and classification methods to detect various heart-related diseases and other complications by analyzing ECG characteristics. Ghaffari et al. (2010) use two innovative modified Hilbert transform-based algorithms to extracted QRS complexes and end-systolic end-diastolic pulses for detecting acute hypotensive episodes and mean arterial pressure dropping regimes. A new classification tree based on conditions combinations competition (T-3C) is proposed for the accurate diagnosis of cardiac ischemia and five measurements were considered in their study (Fayn, 2011). Vullings et al. (2011) use Bayesian framework to develop an adaptive Kalman filter. The adaptive estimation of the process and measurement noise covariance is performed by maximizing the Bayesian evidence function of the sequential ECG estimation. Sufi and Khalil (2011) use data mining techniques, to perform a real-time classification of cardiovascular disease. Lin et al. (2010) use Bayesian algorithm combined with a Markov chain Monte Carlo method to detection and delineation of ECG signals. Recently, association rules have been utilized for the extraction of knowledge from medical history and for the analysis of medical signals (Bourien et al., 2001; Gholam et al., 2006; Zastrow et al., 2009) in reasoning and classification of ECG arrhythmias is straightforward in this article.

3 MATERIALS AND METHODS

This section is a brief description of methods used in this research such as a statistical sampling method, membership function generation by Minimize Entropy Principle Approach (MEPA) and reasoning process of ECG signals by proposed APN.

3.1 Dataset

PhysioNet’s MIT-BIH Arrhythmia Database (PhysioBank, 2006) contains 48 half-hour excerpts of two-channel ambulatory ECG recordings, obtained from 47 subjects (records 201 and 202 are from the same subject) studied by the BIH Arrhythmia Laboratory between 1975 and 1979. Of these, 23 recordings were chosen at random from a collection of >4000 Holter tapes, and the other 25 recordings were selected to include examples of uncommon but clinically important arrhythmias that would not be well represented in a small random sample.

These groups comprise men and women between the ages of 23 and 89, and are analyzed by two independent cardiologists who use nomenclature to classify them by the types of beats and rhythms. Our selected heartbeat clusters are classified into the eleven heartbeat types including one normal and various abnormal types as shown in Table 1.

3.2 MEPA

In this section, an MEPA is adopted to provide the persuasiveness of determining the length of intervals and membership functions in heartbeat patterns. Assume that a threshold value is sought for a sample in the range between \(x_1\) and \(x_2\). An entropy equation with each value of \(x\) is written for the regions \([x_1, x_1 + x]\) and \([x_1 + x, x_2]\) which are defined as the first region \(f\) and the second region \(g\), respectively. Entropy with each value of \(x\) in the region between \(x_1\) and \(x_2\) is expressed as (Christensen, 1980):\

\[
E(x) = f(x)E_f(x) + g(x)E_g(x)
\]

Where

\[
E_f(x) = -E[f_1(x) + \ln f_1(x) + f_2(x) + \ln f_2(x)]
\]

\[
E_g(x) = -E[g_1(x) + \ln g_1(x) + g_2(x) + \ln g_2(x)]
\]

\[
f(x) + g(x) = 1
\]

and \(f_1(x)\) and \(g_1(x)\) are conditional probabilities that class \(k\) sample in the region \([x_1, x_1 + x]\) and \([x_1 + x, x_2]\) respectively. A value of \(x\) that gives the minimum entropy is the optimum threshold value. The value estimates of \(f(x)\) and \(g(x)\), \(f(x)\) and \(g(x)\), are calculated as follows:

\[
f(x) = \frac{n_1(x) + 1}{n(x) + 1}
\]

\[
g(x) = \frac{n_2(x) + 1}{n(x) + 1}
\]

and \(f(x)\) and \(g(x)\), \(f(x)\) and \(g(x)\), are calculated as follows:

\[
f(x) = \frac{n_1(x)}{n(x)}
\]

\[
g(x) = 1 - f(x)
\]

where \(n_1(x)\) = number of class \(k\) samples located in \([x_1, x_1 + x]\), \(n(x)\) = the total number of samples located in \([x_1, x_1 + x]\), \(n_2(x)\) = number of class \(k\) samples located in \([x_1 + x, x_2]\), \(N(x)\) = the total number of samples located in \([x_1 + x, x_2]\), \(n\) = total number of samples in \([x_1, x_2]\), \(z\) = a general length along the interval \([x_1, x_2]\).

Table 1. Type of heartbeat

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>C1</td>
</tr>
<tr>
<td>Abnormal</td>
<td>C2</td>
</tr>
<tr>
<td>Abnormal</td>
<td>C3</td>
</tr>
<tr>
<td>Abnormal</td>
<td>C4</td>
</tr>
<tr>
<td>Abnormal</td>
<td>C5</td>
</tr>
<tr>
<td>Abnormal</td>
<td>C6</td>
</tr>
<tr>
<td>Abnormal</td>
<td>C7</td>
</tr>
<tr>
<td>Abnormal</td>
<td>C8</td>
</tr>
<tr>
<td>Abnormal</td>
<td>C9</td>
</tr>
<tr>
<td>Abnormal</td>
<td>C10</td>
</tr>
<tr>
<td>Abnormal</td>
<td>C11</td>
</tr>
</tbody>
</table>
While moving $x$ in the region $[x_1, x_2]$, we can determine the values of entropy for each position of $x$, as in Figure 1. The value in the region that holds the minimum entropy is called the primary threshold (PRI) value. If we denote a secondary threshold in one area such as SEC1 and the other secondary threshold in another area will be SEC2. In order to develop seven partitions, we need tertiary threshold values and denoted as TER1, TER2, TER3 and TER4. The induction is performed by the entropy minimization principle, which clusters most optimally the parameters corresponding to the output classes (Ross, 2000). By minimizing the entropy, we can find intervals in which the distribution of samples of any class is as relatively uniform as possible.

3.3 APN model
Classification using APN was applied for the detection of arrhythmia heats. APN is defined as follows (Shih et al., 2007).

3.3.1 Definition of APN
An APN is a directed graph, which contains three types of nodes: places, squares and transitions. Where circles represent places, squares represent thresholds of association degree and bars represent transitions. Each place may contain a token associated with a truth-value between zero and one. Each transition is associated with a certainty factor (CF) value between zero and one. Directed arcs represent the relationships between places. A generalized APN structure can be defined as a 13-tuple and the APN can be mathematically and graphically illustrated in Figures 2 and 3.

Let $A$ be a set of directed arcs. If $p_j \in R(t_i)$, then there exists a directed arc $a_{ij}, a_{ij} \in A$, from the place $p_j$ to the transition $t_i$. If $p_j \in O(t_i)$, then there exists a directed arc $a_{kj}, a_{kj} \in A$, from the transition $t_j$ to the place $p_k$. If $W(s_m) = w_m, w_m \in [0,1]$, then the support $s_m$ is said to be associated with a real value $w_m$. If $G(c) = c_1, c_2 \in [0,1]$, then the transition $t_c$ is said to be associated with a real value $c_1$. If $\beta (p_j) = d_j, d_j \in D$, then the place $p_j$ is said to be associated with the proposition $d_j$. An APN with some places containing tokens is called a marked APN. In a marked APN, the token in a place is represented by a labeled dot $\text{dot}(a_{ij})$. The token value in a place $p_j$, $p_j \in P$, is denoted by $a(p_j)$, where $a(p_j) \in [0,1]$. If $a(p_j) = y_j$ and $\beta (p_j) = d_j$, then it indicates that the proposition $d_j$ is $y_j$. According to the antecedent portion or consequence portion of an APR contains ‘and’ or ‘or’ connectors, the APRs have five rule types (Shih et al., 2007).

3.3.2 Associative transition function
Let $IT = \{i_1, i_2, ..., i_m\}$ be a set of items and $DT$ be a set of database transactions. An association rule is an implication of the form $A \Rightarrow B$, where $A \subseteq IT$, $B \subseteq IT$ and $A \cap B = \phi$. The association rule $A \Rightarrow B$ holds in the transaction set DT with support $s$, where $s$ is the percentage of transactions in DT that contain $A \cup B$. This is taken to be the probability $P(A \cup B)$. The association rule $A \Rightarrow B$ has confidence $c$ in the transaction set DT if $c$ is the percentage of transactions in DT containing $A$ that also contain $B$. This is taken to be the conditional probability $P(B|A)$ (Agrawal et al. 1993). That is,

\[
\text{Support} (A \Rightarrow B) = P(A \cup B) \tag{9}
\]

\[
\text{Confidence} (A \Rightarrow B) = P(B|A) \tag{10}
\]

Typically, association rules are considered interesting if they satisfy both a minimum support threshold $\tau_s$, $\tau_s \in [0,1]$, and a minimum confidence threshold $\gamma$, $\gamma \in [0,1]$. If the values of support and confidence are higher than their threshold $\tau$ and $\gamma$, the transition is enabled and the $\text{CF value of transition is corresponding to its confidence value (CF = c)}$ else the relationship does not exist (CF = 0). A generalized formulation of the CF is shown in Equation (11):

\[
G(t_i) = \begin{cases} 
\bar{c}_i & \text{if } s_m \geq \tau_s \text{ and } c_i \geq \gamma_i \\
0 & \text{Otherwise} 
\end{cases} \tag{11}
\]

where $c_i$ is the confidence value, $\gamma_i$ is a threshold of confidence, $s_m$ is the support value and $\tau_s$ is a threshold of support.

3.3.3 Reasoning process
An APN model can be expressed as a network structure. Each place in the network is denoted by a triple ($p_i, a(p_i)$).

\[
APN = \{P, T, S, C, D, A, \Gamma, I, O, \alpha, \beta, G, W\}
\]

where

- $P = \{p_1, p_2, ..., p_N\}$ is a finite set of places,
- $T = \{t_1, t_2, ..., t_M\}$ is a finite set of transitions,
- $S = \{s_1, s_2, ..., s_n\}$ is a finite set of supports,
- $C = \{c_1, c_2, ..., c_m\}$ is a finite set of confidences,
- $D = \{d_1, d_2, ..., d_l\}$ is a finite set of propositions,
- $A = \{a_1, a_2, ..., a_k\}$ is a finite set of thresholds of the supports,
- $\Gamma = \{\gamma_1, \gamma_2, ..., \gamma_m\}$ is a finite set of thresholds of the confidences,
- $P \cap T \cap D = \emptyset, |P| = |T|$
- $I : T \rightarrow P^*$ is an input function, a mapping from transitions to bags of places,
- $O : T \rightarrow P^*$ is an output function, a mapping from transitions to bags of places,
- $\alpha : P \rightarrow [0,1]$ is an association function, a mapping from places to real values between zero and one,
- $\beta : P \rightarrow D$ is an association function, a bijective mapping from places to propositions,
- $G : T \rightarrow [0,1]$ is an association function which assigns a real value between zero to one to each transition,
- $W : S \rightarrow [0,1]$ is an association function which assigns a real value between zero to one to each support.

![Fig. 1. Partitioning process of MEPA](https://academic.oup.com/bioinformatics/article-abstract/30/12/1739/2748137/365x93-521x209)

![Fig. 2. Definition of APN](https://academic.oup.com/bioinformatics/article-abstract/30/12/1739/2748137/365x93-521x209)

![Fig. 3. A generalized APN structure](https://academic.oup.com/bioinformatics/article-abstract/30/12/1739/2748137/365x93-521x209)
IRs(πk), where πk ∈ P and IRs(πk) is the immediate reachability set of πk. RS(πk) is the set of places which is reachable from place πk and AP is a set of antecedent places of πk. Let sxy denote the support degree and cxy denote the CF value associated with transition between place πx and πy. The degree of truth of proposition dx is defined as a(dx). The threshold of degree of truth of each proposition is given as λ. If a(dx) ≥ λ, then the proposition dx is asserted. Assume that the thresholds of support and confidence degree are given with τ and γ. If sxy ≥ τxy and cxy ≥ γxy, then the transition is fired. A reasoning algorithm of APN can generate reasoning paths from starting place p0, to goal place pν, as described below.

Reasoning Algorithm of APN

**Definition.** Let V be a set of successful reasoning paths and ν is a successful reasoning path, ν ∈ V. The proposition of goal place pν is dν(ν) and a(pν(ν)) is the degree of truth of pν(ν) if and only if there exist a successful reasoning paths from starting place to goal place.

**Input.** P = {p1, p2, ..., pν}, D = {d1, d2, ..., dν}, S = {s1, s2, ..., sν}, C = {c1, c2, ..., cν} THEN T = {t1, t2, ..., tν}, Λ = {τ1, τ2, ..., τν}, Γ = {γ1, γ2, ..., γν} where P, D, S, C, T, Λ and Γ are a finite set of places, propositions, transitions, supports, confidences, thresholds of support and thresholds of confidence.

**Output.** The proposition p(πν) = dν and the degree of truth of a(πν) in a goal place.

**Step 1:** From starting place p0, The starting place (p0, a(p0), IRs(p0)) is a non-terminal place, where
1. p(π0) = d0
2. a(p0) ≥ λ, a transition is enabled to fire.
3. IRs(p0) ≠ φ

**Step 2:** At descendant place p(0) (from starting place p0)
For each place πk in IRs(p0),
IF πk = pν, a(πk) ≥ λ, CFν = cνς, σν = γνς and νς = γνς THEN
An arc, labeled aνς, is directed from (πk, a(πk), IRs(πk)) to (πν, a(πν), IRs(πν)), where a(πν) = a(πk) * cνς. The path is called a successful reasoning path.
ELSE IF πk ≠ pν, a(πk) ≥ λ, CFν = cνς, σν ≥ γνς and νς ≥ γνς THEN
An arc, labeled aνς, is directed from (πk, a(πk), IRs(πk)) to (πν, a(πν), IRs(πν)), where a(πν) = a(πk) * cνς.
ELSE IF APν = {πν, πν, ..., πν} and pν ∈ RS(πν) THEN
Let g = Min {a(πν) * cνς, a(πν) * cνς, ..., a(πν) * cνς}.
IF g ≥ λ, CFν = cνς, σν ≥ γνς and νς ≥ γνς THEN
An arc, labeled aνς, is directed from (πk, a(πν), IRs(πν)) to (πν, a(πν), IRs(πν)), where a(πν) = g * cνς.
END IF
ELSE
Mark place πν as a terminal place.
END IF

**Step 3:** At descendant place p(0)
For each place πk in RS(πν) and πk ∈ IRs(pν),
IF πk = pν, a(πν) ≥ λ, CFν = cνς, σν ≥ γνς and νς ≥ γνς THEN
An arc, labeled aνς is directed from (πk, a(πν), IRs(πν)) to (πν, a(πν), IRs(πν)), where a(πν) = a(πk) * cνς. The path is called a successful reasoning path.
ELSE IF πk ≠ pν, πk ∈ RS(πν) and a(πk) ≥ λ, CFρ = cρς, σρ ≥ γρς and ρς ≥ γρς THEN
An arc, labeled aρς is directed from (πk, a(πk), IRs(πk)) to (πν, a(πν), IRs(πν)), where a(πν) = a(πk) * cρς.
ELSE IF APν = {πν, πν, ..., πν} and pν ∈ RS(πν) THEN
Let g = Min {a(πk) * cρς, a(πk) * cρς, ..., a(πk) * cρς}.
IF g ≥ λ, CFρ = cρς, σρ ≥ γρς and ρς ≥ γρς THEN

4 EXPERIMENT

The experimental procedure for arrhythmia-beat detection is divided into ECG filtering and sampling, feature extraction, transfer function building and classification.

4.1 ECG filtering and sampling

First, we use a non-linear bilateral filter (Elad, 2002) to remove noise. The sampling size of our dataset has been estimated by the population ratio (Mendenhall and Beaver, 1994):

\[ n = \frac{z^2 \cdot \alpha^2 \cdot P(1 - P)}{\epsilon^2} \]

where n is the sampling size, P is the estimator, ε is the sampling error, α is the confidence coefficient and z is the standard normal probability distribution. Since there is no further study in the estimation of an ECG beat will be the abnormal beat’s probability. Therefore, we use non-constant number sampling to obtain the abnormal beat’s proportion of population in ten time's trial, which is shown in Table 2. From Table 2, we can obtain an abnormal beat’s estimator (population proportion expected value) as follow:

\[ P = \sum_{i=1}^{10} C_i * P_i = 0.0611844 \]

<table>
<thead>
<tr>
<th>Beat number</th>
<th>Abnormal number</th>
<th>Abnormal ratio (P_i)</th>
<th>Weight (C_i)</th>
</tr>
</thead>
<tbody>
<tr>
<td>6628</td>
<td>4936</td>
<td>0.7447</td>
<td>0.0428</td>
</tr>
<tr>
<td>13588</td>
<td>10615</td>
<td>0.5000</td>
<td>0.0308</td>
</tr>
<tr>
<td>14321</td>
<td>10860</td>
<td>0.7583</td>
<td>0.0777</td>
</tr>
<tr>
<td>8886</td>
<td>5429</td>
<td>0.6109</td>
<td>0.0817</td>
</tr>
<tr>
<td>11056</td>
<td>9417</td>
<td>0.3000</td>
<td>0.0578</td>
</tr>
<tr>
<td>9967</td>
<td>8584</td>
<td>0.8612</td>
<td>0.0668</td>
</tr>
<tr>
<td>15476</td>
<td>12867</td>
<td>0.8314</td>
<td>0.1605</td>
</tr>
<tr>
<td>10392</td>
<td>8109</td>
<td>0.0900</td>
<td>0.0358</td>
</tr>
<tr>
<td>13596</td>
<td>7680</td>
<td>0.5649</td>
<td>0.1425</td>
</tr>
<tr>
<td>9490</td>
<td>7537</td>
<td>0.7941</td>
<td>0.2831</td>
</tr>
</tbody>
</table>

Table 2. Ten times non-constant sampling
where $C_i$ is beat weight that is the sample proportion of sampling $i$ time, $P_i$ is abnormal ratio of sampling $i$ time. $C_i$ and $P_i$ are shown in Table 2.

Substitute $P$-value, with 0.95 confidence levels and 0.02 sampling error in y sampling beat in different group, as shown in Table 3, we obtained 553 beat (with 120 normal and 433 abnormal) in our sampling data set. By randomly select, about two-thirds (67.5%) of the data set are used to training our proposed APN knowledge model and one-third (32.5%) of the data set are used to evaluate the proposed APN model.

### 4.2 Feature extraction

Selected features are important in enhancing the performance of ECG arrhythmia-beat detection. After detailed survey, we used the features listed in Table 4 for our prototyping.

The steps of detecting peaks are (i) calculating the moving average from original signal, and usually using the past ten records to calculate it; (ii) subtracting moving average from original signal, and getting a new signal; (iii) finding the peak of the signal; and (iv) setting the threshold to decide the detected peaks. As soon as the peak ‘R’ determined, we can obtain peaks P, Q, S and T through their relative position. In order to prevent possible error, peaks’ amplitude would be measured from $k$ line as shown in Figure 4 (Shih et al., 2010). The definition of $k$ line is in Equation (13).

$$k = \text{Max}(\theta_i, i = 1, 2, \ldots, 11) + c$$  \hspace{1cm} (13)

where $k$ is a baseline, $\theta$ is the greatest amplitude of all peak, $i$ is type of heartbeat and $c$ is a constant. Therefore, the amplitude of peaks P, Q, R, S and T are all homogeneous negative.

### 4.3 Transfer function building

Membership functions and thresholds are determined by MEPA (Christensen, 1980) in Section 3.2. Each of the features is partitioned into three states which are high, median, and low. For example, Figure 5 is the membership function of TH and the membership functions of some other features are shown in Appendix A.

### 4.4 APN model construction

The structure of our APN model for arrhythmia beat detection is constructed by using the construction algorithm from Shih et al. (2013) and is shown in Figure 6. Our APN model for the symptoms of heart-disease reasoning have eight variables, 11 middle states (one normal heartbeat and 10 symptoms of heart disease), and one final state. All the degree of truth and states of the input places in the APN are gained by the membership functions calculated by Equation (11).

### 4.5 An example of arrhythmia-beat classification

An APN reasoning algorithm, in Section 3.3, is used to classify heart-beat symptoms into the following category: normal/abnormal heartbeat or 11 different types of arrhythmia beat. APN

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Beat number</th>
<th>Ratio (%)</th>
<th>Sample number</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>75041</td>
<td>21</td>
<td>120</td>
</tr>
<tr>
<td>C2</td>
<td>8123</td>
<td>13</td>
<td>72</td>
</tr>
<tr>
<td>C3</td>
<td>7336</td>
<td>4</td>
<td>21</td>
</tr>
<tr>
<td>C4</td>
<td>2566</td>
<td>11</td>
<td>60</td>
</tr>
<tr>
<td>C5</td>
<td>151</td>
<td>4</td>
<td>23</td>
</tr>
<tr>
<td>C6</td>
<td>7171</td>
<td>16</td>
<td>88</td>
</tr>
<tr>
<td>C7</td>
<td>886</td>
<td>6</td>
<td>32</td>
</tr>
<tr>
<td>C8</td>
<td>259</td>
<td>3</td>
<td>18</td>
</tr>
<tr>
<td>C9</td>
<td>224</td>
<td>3</td>
<td>19</td>
</tr>
<tr>
<td>C10</td>
<td>7028</td>
<td>13</td>
<td>74</td>
</tr>
<tr>
<td>C11</td>
<td>982</td>
<td>5</td>
<td>26</td>
</tr>
</tbody>
</table>

**Table 3. Sampling number in clusters**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Feature</th>
<th>Description</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_1$</td>
<td>RH</td>
<td>R wave amplitude.</td>
<td>Hosseini et al. (2006)</td>
</tr>
<tr>
<td>$x_2$</td>
<td>PH</td>
<td>P wave amplitude.</td>
<td>Acir (2006)</td>
</tr>
<tr>
<td>$x_3$</td>
<td>QH</td>
<td>Q wave amplitude.</td>
<td>Acir (2006)</td>
</tr>
<tr>
<td>$x_4$</td>
<td>SH</td>
<td>S wave amplitude.</td>
<td>Acir (2006)</td>
</tr>
<tr>
<td>$x_5$</td>
<td>TH</td>
<td>T wave amplitude.</td>
<td>Acir (2006), Exarchos et al. (2006)</td>
</tr>
<tr>
<td>$x_6$</td>
<td>PR</td>
<td>PR Interval.</td>
<td>Singh and Tiwari (2006)</td>
</tr>
<tr>
<td>$x_7$</td>
<td>QS</td>
<td>QRS wave duration.</td>
<td>Singh and Tiwari (2006), Zhou et al. (2005)</td>
</tr>
<tr>
<td>$x_8$</td>
<td>ST</td>
<td>ST Interval.</td>
<td>Singh and Tiwari (2006), Hosseini et al. (2006)</td>
</tr>
</tbody>
</table>

**Table 4. Features description**

**Fig. 4. A single ECG wave with $k$ baseline**
model in Figure 6 is used for reasoning of arrhythmia-beat classification and has five inputs, three middle states, 11 outputs and one final state. For example, assume that five inputs characteristics \(x_1, x_2, \ldots, x_5\) of an ECG-extracted signal from ECG wave are \([-2.7162, -2.3648, -1.2564, -0.5810, -0.213]\). From Appendix A, we can obtain the membership degree and the linguistic values of this ECG-extracted signal are \([M, L, M, M, H]\) and \([0.708, 1.0, 0.970, 0.791, 1.0]\). Then \(\alpha(p_0) = 0.708, \alpha(p_2) = 1.0, \alpha(p_3) = 1.0\) are a mapping from places \(p_0, p_2, p_3\) to its real values by using membership function. The \(p_1, p_2, \ldots, p_9\) are starting places, where \(\lambda > 0\) shows that the proposition exists and \(\lambda = 0\) shows that the proposition does not exist. Let \(\beta(p_1) = d_1, \beta(p_2) = d_2, \ldots, \beta(p_9) = d_8\) and \(d_1, d_2, d_3, \ldots, d_8\) are eight propositions of ECG signals is shown in Table 4. The propositions \(d_9, d_{10}, \ldots, d_{19}\) denote 11 clusters of heartbeat. Let place \(p_{20}\) be a goal place, \(d_{20}(v)\) is the proposition of goal place \(p_{20}\) and \(\alpha(p_{20}(v))\) is the degree of truth of \(d_{20}(v)\). Assume that all the threshold values of \(\tau_d\) and \(\gamma_d\) are set to 0.05 and 0.02, respectively. The reasoning example of our APN for the symptoms of heart-disease reasoning is as follow:

Step 1: From starting places \(p_1, p_2, \ldots, p_5\): places \(p_1, p_2, \ldots, p_5\) are starting places and \(\beta(p_1) = d_1, \beta(p_2) = d_2, \ldots, \beta(p_9) = d_8\) are 5 propositions of starting places. IRS\(p_s) = \{p_0, p_1, p_2, p_3\}\n
Step 2: In descendant place \(p_4\), \(p_4 \in IRS(p_s), s = 1, 2, \ldots, 5; i = 6, 7, 8\).

\[
\begin{align*}
\text{AP}_{s,i} = \{p_1, p_2, \ldots, p_5\} \\
\text{and all transitions are enabled, places } p_1, p_2, \ldots, p_5 \text{ are exist since } (\alpha(p_s) \geq \lambda) \text{ and all the values of support and confidence are greater than threshold } \tau_d \text{ and } \gamma_d, \text{ except } c_{99} \text{ (i.e. } s_{99} \geq \tau_{99} = 0.05 \text{ and } c_{99} \geq \gamma_{99} = 0.02, s = 1, 2, \ldots, 5, i = 6, 7, 8) \text{. Tokens are moved from } p_1, p_2, \ldots, p_5 \text{ to } p_0, p_0, p_7, p_8. \text{ Since the antecedent portion or consequence portion of APRs between } p_1, p_2, \ldots, p_5 \text{ and } p_6, p_7, p_8 \text{ belongs to type 3, the reasoning function is ‘Max’ operator (Shih et al., 2007). The degree of truth of proposition of place } p_6, p_7, p_8 \text{ are}
\end{align*}
\]

\[
\begin{align*}
g = \alpha(p_6) &= \max(\alpha(p_s) * c_{s1}, \alpha(p_s) * c_{s9}, \alpha(p_s) * c_{s8}) \\
&= \max(0.708 \times 0.970, 0.970 \times 0.791, 0.708 \times 1.0) \\
&= 0.7077 \\

\alpha(p_6) &= \max(\alpha(p_s) * c_{s1}, \alpha(p_s) * c_{s9}, \alpha(p_s) * c_{s8}) \\
&= \max(0.708 \times 0.970, 0.970 \times 0.791, 0.708 \times 1.0) \\
&= 0.7077 \\
\end{align*}
\]

Step 3: At descendant place \(p_4, p_4 \in IRS(p_s), i = 6, 7, 8; j = 9, 10, \ldots, 19\).

\[
\begin{align*}
\text{AP}_{s,i} = \{p_1, p_2, p_3\} \\
\text{and all transitions are enabled, places } p_6, p_7, p_8 \text{ are exist since } (\alpha(p_s) \geq \lambda). \text{ Some paths are not discarded, since the values of support and confidence are less than threshold } \tau_d \text{ and } \gamma_d, \text{ such as } p_0 \rightarrow p_{10}, p_{14}, p_{15} \rightarrow p_{11}, \text{ etc. (i.e. } s_{99} \geq \tau_{99} = 0.05 \text{ and } c_{99} \geq \gamma_{99} = 0.02, j = 6, 7, 8; j = 9, 10, \ldots, 19) \text{. Tokens are moved from } p_6, p_7, p_8. \text{ Since the antecedent portion or consequence portion of APRs between } p_6, p_7, p_8 \text{ and } p_9, p_{10}, \ldots, p_{19} \text{ belongs to type 1, the reasoning function is ‘Min’ operator (Shih et al., 2007). The degree of truth of proposition at place } p_9, p_{10}, \ldots, p_{19} \text{ are}
\end{align*}
\]

\[
\begin{align*}
g = \alpha(p_9) &= \min(\alpha(p_s) * c_{s1}, \alpha(p_s) * c_{s9}, \alpha(p_s) * c_{s8}) \\
&= \min(0.708 \times 0.2077, 0.2077 \times 0.0827, 0.1182 \times 0.0889) \\
&= 0.0105 \\

\alpha(p_9) &= \min(\alpha(p_s) * c_{s1}, \alpha(p_s) * c_{s9}, \alpha(p_s) * c_{s8}) \\
&= \min(0.708 \times 0.2077, 0.2077 \times 0.0827, 0.1182 \times 0.0889) \\
&= 0.0105 \\
\end{align*}
\]

Step 4: At descendant place \(p_{10}\), since \(p_k = p_{10} = p_{20}\), places \(p_0, p_{10}, \ldots, p_{19}\) are exist (\(\alpha(p_s) \geq \lambda\)) and all the values of support and confidence are greater than threshold \(\tau_d\) and \(\gamma_d\), all transitions are enabled (i.e. \(s_{99} \geq \tau_{99} = 0.05\) and \(c_{99} \geq \gamma_{99} = 0.02, j = 9, 10, \ldots, 19; k = 20\)). Assume that all the CFs linked to \(p_{20}\) are 0.9, and let \(c_{j,20} = 0.9,\)
Ventricular Premature Contraction) and degree of truth is rate

\[ \alpha(p_2(v)) = \alpha(p_3) \ast c_{20} \quad (j = 9, 10, \ldots, 19) \]  

\[ \alpha(p_2(1)) = \alpha(p_3) \ast c_{10} = 0.0105 \ast 0.9 = 0.0095 \]  

\[
\vdots
\]

\[ \alpha(p_2(11)) = \alpha(p_3) \ast c_{19,20} = 0.0387 \ast 0.9 = 0.0348 \]

There are eleven successful reasoning paths, \( v = 1, 2, \ldots, 11 \), \( v \in V \).

Step 5: There are 11 successful reasoning paths into goal place \( p_{20} \), \( V \neq \phi \). Since the APRs between \( p_{19}, p_{10}, \ldots, p_{19} \) and \( p_{20} \) belongs to type 3, the reasoning function is ‘Max’ operator, we can select proposition \( d_{20}(v) \) and the degree of truth of \( d_{20}(v) \) at the place \( p_{20} \) as:

\[
\begin{align*}
V &= \{(p_{20}, d_{20}(1), \alpha(p_{20}(1))), (p_{20}, d_{20}(2), \alpha(p_{20}(2))), \ldots, (p_{20}, d_{20}(11), \alpha(p_{20}(11)))\} \\
\]  
\[
\begin{align*}
z &= \text{Max} \{\alpha(p_{20}(1)), \alpha(p_{20}(2)), \ldots, \alpha(p_{20}(11))\} \\
&= \text{Max}\{0.0095, 0.01, 0.0287, 0.0995, 0.042, 0.0077, 0.0077, 0.0059, 0.0056, 0.0065, 0.0348\} \\
&= 0.0348 \text{ (at } p_{19} = p_{20})
\end{align*}
\]

Step 6: Obtain proposition \( d_{20}(v) \) and the degree of truth \( \alpha(p_{20}(v)) \) at place \( p_{20} \).

Finally, we can obtain the belonged cluster (Cluster 11: Ventricular Premature Contraction) and degree of truth is 0.0348 in this example.

5 EXPERIMENTAL RESULTS AND DISCUSSION

The MIT-BIH arrhythmia database is used in our experiment. We use 368 cases in training and 185 cases for testing from sampled data in Table 3. After APN model construction, the testing result in confusion matrix of our proposed APN model is summarized in Table 5. Table 6 summarized the other reported studies. The experimental results strongly suggest that our approach compares well with other data-mining methods in sensitivity, F-measure and accuracy measurement. And, other measurements are great too. Therefore, we can conclude that our proposed APN model have a good result on classification of ECG arrhythmias.

<table>
<thead>
<tr>
<th>Heartbeat</th>
<th>APN (%)</th>
<th>Shih (%)</th>
<th>Acir (%)</th>
<th>Engin (%)</th>
<th>Osowski (%)</th>
<th>Zhou (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>93.51</td>
<td>92.9</td>
<td>89.8</td>
<td>93</td>
<td>98.1</td>
<td>93.9</td>
</tr>
<tr>
<td>C2</td>
<td>98.92</td>
<td>90.8</td>
<td>85.8</td>
<td>-</td>
<td>97</td>
<td>-</td>
</tr>
<tr>
<td>C3</td>
<td>96.76</td>
<td>96.7</td>
<td>-</td>
<td>99</td>
<td>94</td>
<td>-</td>
</tr>
<tr>
<td>C4</td>
<td>99.46</td>
<td>92.4</td>
<td>-</td>
<td>-</td>
<td>91.3</td>
<td>-</td>
</tr>
<tr>
<td>C5</td>
<td>100.00</td>
<td>97.8</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>C6</td>
<td>100.00</td>
<td>97.3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>C7</td>
<td>100.00</td>
<td>95.6</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>C8</td>
<td>98.92</td>
<td>94.5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>C9</td>
<td>98.92</td>
<td>95.1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>C10</td>
<td>100.00</td>
<td>95.6</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>C11</td>
<td>100.00</td>
<td>93.5</td>
<td>90.3</td>
<td>100</td>
<td>96.5</td>
<td>93.9</td>
</tr>
</tbody>
</table>

\(< \ast \) Did not discriminate this heartbeat class.

Table 6. Comparison with other research

<table>
<thead>
<tr>
<th>Methods</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>AUC</th>
<th>F-measure</th>
<th>G-mean</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>NB</td>
<td>85.0</td>
<td>83.4</td>
<td>0.842</td>
<td>0.694</td>
<td>0.842</td>
<td>0.838</td>
</tr>
<tr>
<td>C4.5</td>
<td>87.5</td>
<td>94.5</td>
<td>0.910</td>
<td>0.843</td>
<td>0.909</td>
<td>0.930</td>
</tr>
<tr>
<td>SVM</td>
<td>92.5</td>
<td>87.6</td>
<td>0.900</td>
<td>0.779</td>
<td>0.900</td>
<td>0.886</td>
</tr>
<tr>
<td>Bayes Net</td>
<td>85.0</td>
<td>92.4</td>
<td>0.887</td>
<td>0.800</td>
<td>0.886</td>
<td>0.908</td>
</tr>
<tr>
<td>APN</td>
<td>85.0</td>
<td>95.9</td>
<td>0.908</td>
<td>0.850</td>
<td>0.903</td>
<td>0.935</td>
</tr>
<tr>
<td>FPN</td>
<td>92.5</td>
<td>83.4</td>
<td>0.880</td>
<td>0.733</td>
<td>0.879</td>
<td>0.854</td>
</tr>
</tbody>
</table>

AUC indicate a classifier has a better classification performance. Other data mining methods such as Na"ïve Bayes (NB), Decision tree (C4.5), Bayes Net, Support vector machine (SVM) and Fuzzy Petri net (FPN) running with Weka and MATLAB 2012 are shown in Table 7 for comparison. APNs and C4.5 have a higher TN-rate in detecting abnormal heartbeat. SVM and FPN have a good TP-rate in detecting normal heartbeat in our experiment. Generally, in Table 7, our APN have the best performance than other data-mining methods in sensitivity, F-measure and accuracy measurement. And, other measurements are great too. Therefore, we can conclude that our proposed APN model have a good result on classification of ECG arrhythmias.

6 CONCLUSION

We have proposed a methodology for the detection of ECG arrhythmia using a generalized APN model in this article. The performance evaluation using MIT-BIH arrhythmia database in Table 7 shows that our approach compares well with other reported studies. The experimental results strongly suggest that our proposed APN model can help doctors in the diagnosis of ECG Arrhythmia and may applied to other symptom decision field. In the future, an applicable decision system for the clinical diagnosis of ECG arrhythmias can be developed based on our provided APN model. And, the health care practitioners can be aware of types
of cardiac arrhythmias and give early treatment to reduce deterioration.

APN is a powerful method for representing knowledge and logic reasoning in the domain of decision support systems. Unlike other Petri net model, APN model has a systematic procedure in model construction. The reasoning path of expert systems can be reduced to simple sprouting trees if APN-based reasoning algorithms can be applied. However, there still drawbacks exist in APN such as state explosion and no hierarchy concepts. We shall overcome all these drawbacks in the near future.

Funding: National Science Council of Taiwan (grants NSC 98-2218-E-212-001 and NSC-97-2218-E-224-008).

Conflict of Interest: none declared.

REFERENCES


