

# A methodology to analyze heart data using fuzzy automata<sup>1</sup>

Iván Calvo<sup>a</sup>, Mercedes G. Merayo<sup>a</sup> and Manuel Núñez<sup>a,\*</sup>

<sup>a</sup> *Universidad Complutense de Madrid, Spain*

*E-mail: {ivcalvo,mgmerayo,manuelnu}@ucm.es*

**Abstract.** Uncertainty and imprecision play an important role in the specification and analysis of complex systems. Therefore, it is important to provide methodologies and tools to support the correct development of these systems. In this paper we present a new formalism, based on *fuzzy automata*, to facilitate the different phases involved in the development of a system where information is *fuzzy*. The formal syntax and semantics of our formalism are based on previous work, which has been adapted to be easily implemented and automated. We introduce a methodology to analyze systems modelled with one of our fuzzy automata. Finally, we show how our framework can be used to define a model of the heart based on electrocardiograms (ECGs) and use this model to analyze data of real patients.

Keywords: Fuzzy automata, Formal specification, ECG

## 1. Introduction

The development of complex computer systems has benefited from the application of formal methods. In particular, with the aid of these methods it is possible to increase the reliability of the developed systems. The main downside of formal methods is often associated with their own complexity and the lack of good tools to support them. The situation is slowly improving and relevant disciplines in the development of systems, such as software testing, are providing a myriad of methodologies [5] supported by tools [21]. Another important drawback is that *generalist* formalisms may not be well suited for specific fields. For example, the application considered in this paper, the automated analysis of heart recordings, is characterized by uncertainty and imprecision in the data. This common issue in biological and medical systems can be seen in several approaches proposing a formal model of the

heart [11,12,16]. The control over these inaccuracies is crucial for the suitability of any model, since no useful result can be obtained from an unrealistic model. Actually, there is a long tradition on using intelligence to deal with big amounts of information to analyze these systems [3].

A possible way to cope with uncertainty and imprecision is given by formalisms using fuzzy logic [24], in particular, formalisms based on automata theory [22, 18,7,1,4]. In this paper, we propose a version of *fuzzy automata*, based on our recent work [2], and show how it can be applied to model medical and biological systems taking into account their inherent imprecision. We use our formalism to propose a model of the behavior of the heart that takes values obtained from ECGs (electrocardiograms) recordings as input data. These values are compared with *normal* levels in a *fuzzy* way. We have used a recent list of these levels [19] but these values are very similar to the levels that we used in our previous work [20]. It is worth to mention that there are studies where normal limits are determined for specific populations [6,13,14,23] and these values can be used if we want to focus on a specific group.

This paper represents a revised, extended and enhanced version of our previous work [2]. Specifically,

---

<sup>1</sup>Research partially supported by the Spanish projects DARDOS (TIN2015-65845-C3-1-R) and FAME (RTI2018-093608-B-C31) and the Comunidad de Madrid project FORTE-CM (S2018/TCS-4314).

\*Corresponding author. E-mail: manuelnu@ucm.es.

in this paper we have given an extensive revision of the formal aspects underlying the proposed framework. After a detailed introduction of all the required concepts, we give a revised formal definition of the syntax and semantics of our proposed *fuzzy automata*. The new semantics no longer depends on the application of an implicit variable transformation. The two categories of actions (inputs and outputs) have been slightly re-defined, putting extra emphasis on their different roles. In the definition of the model of the heart, an improved naming convention has allowed us to easily adapt the proposed model to different age ranges and normal limits coming from diverse studies. Finally, the most significant advance appears in the definition of a general analysis methodology based on the proposed framework. This methodology includes a method to produce grades of confidence of certain traces. The flexibility of this method allows the analyst to formally define the events whose likelihood is relevant to study. The multiple grades of confidence obtained this way may then be presented as tables and plots so that domain experts are assisted in their decision-making process.

The rest of the paper is structured as follows. Section 2 introduces the syntax of our formalism and its operational semantics. Section 3 gives our model of the heart with the proposed formalism. In Section 4 we generalize the steps followed in the study of real patients data. We propose an extensible methodology for the analysis of data, which may be used in a wide range of contexts. Finally, in Section 5 we present our conclusions and some lines for future work.

## 2. Formalism of fuzzy automata

In this section we introduce our new formalism. Its syntax and semantics are based on previous work [2]. Before we present the definition of our fuzzy automata, we will introduce some auxiliary concepts.

### 2.1. Fuzzy relations

The first concept that we need to define a fuzzy version of the classical automata model is the notion of *fuzzy relation*. These relations are similar to *boolean relations* but instead of returning true or false, they return a real value in the interval  $[0, 1]$ . Intuitively, the returned value will provide a confidence on the likelihood of the conclusion that we obtain. Therefore, if we have that a computation returns confidence equal to 1,

then we assume that the result holds. On the contrary, if we obtain a confidence equal to 0 then we know that the result does not hold. Finally, intermediate values indicate a relative confidence on the validity of the result. These relations will be parameterized by a value, usually denoted by  $\delta$ , that indicates the maximum allowed *distance* to the correct value. For example, if we are analyzing whether a certain value  $x$  belongs to an interval  $[\alpha, \beta]$  and  $x$  is such that  $\alpha \leq x \leq \beta$ , then we claim that the relation holds with confidence 1. If the distance from  $x$  to  $\alpha$  or  $\beta$  is less than  $\delta$  then we have a positive confidence (the confidence diminishes when the distance increases). Finally, if  $x \notin [\alpha, \beta]$  and it is *far* from the interval then we have confidence 0 on  $x$  belonging to the interval.

Next, we give the *fuzzy* versions of the classical less/greater than and equal operators.

$$\overline{x \leq y}^\delta \equiv \begin{cases} 1 & \text{if } x < y \\ \frac{\delta + y - x}{\delta} & \text{if } y \leq x \leq y + \delta \\ 0 & \text{if } y + \delta < x \end{cases}$$

$$\overline{x \geq y}^\delta \equiv \begin{cases} 1 & \text{if } x > y \\ \frac{\delta + x - y}{\delta} & \text{if } y \geq x \geq y - \delta \\ 0 & \text{if } y - \delta > x \end{cases}$$

$$\overline{x = y}^\delta \equiv \begin{cases} 0 & \text{if } x \leq y - \delta \\ \frac{x - y + \delta}{\delta} & \text{if } y - \delta < x \leq y \\ \frac{-x + y + \delta}{\delta} & \text{if } y < x \leq y + \delta \\ 0 & \text{if } y + \delta < x \end{cases}$$

$$\overline{x \leq y \leq z}^\delta \equiv \begin{cases} 0 & \text{if } z < x \\ 0 & \text{if } y \leq x - \delta \\ \frac{y - x + \delta}{\delta} & \text{if } x - \delta \leq y \leq x \leq z \\ 1 & \text{if } x < y < z \\ \frac{-y + z + \delta}{\delta} & \text{if } x \leq z \leq y \leq z + \delta \\ 0 & \text{if } z + \delta < y \end{cases}$$

Note that every fuzzy relation reduces to its boolean version when  $\delta$  tends to 0. In Figures 1, 2 and 3 we give plots of examples of the use of these relations. In the  $x$ -axis we have the parameter while in the  $y$ -axis we have the confidence.

The previous fuzzy relations include a certain notion of *symmetry*. Specifically, if we switch the roles of  $x$  and  $y$  in the definition of  $\overline{x \geq y}^\delta$ , then we obtain the definition of  $\overline{x \leq y}^\delta$ . This means that, for each  $x$  and

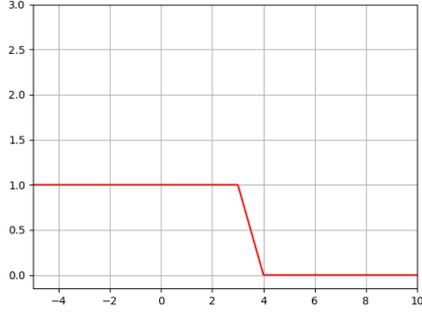


Fig. 1. Plot of  $\overline{\leq 3}^1$ .

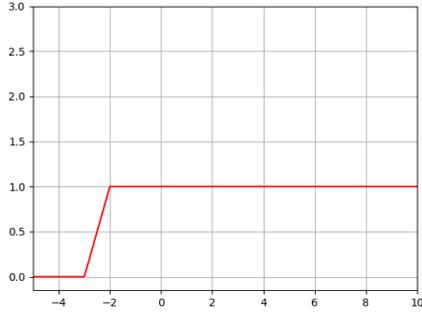


Fig. 2. Plot of  $\overline{\geq -2}^1$ .

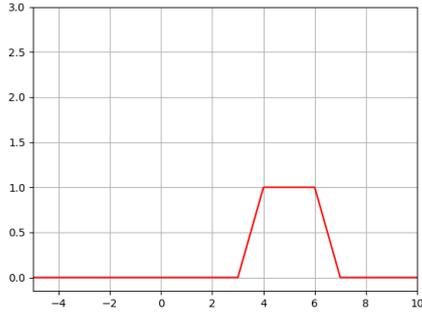


Fig. 3. Plot of  $\overline{4 \leq \cdot \leq 6}^1$ .

$y$ , we have  $\overline{x \leq y}^\delta = \overline{y \geq x}^\delta$ , which generalizes the fact that  $x \leq y$  if and only if  $y \geq x$ .

### 2.2. T-norms

In order to combine grades of confidence we use  $t$ -norms.

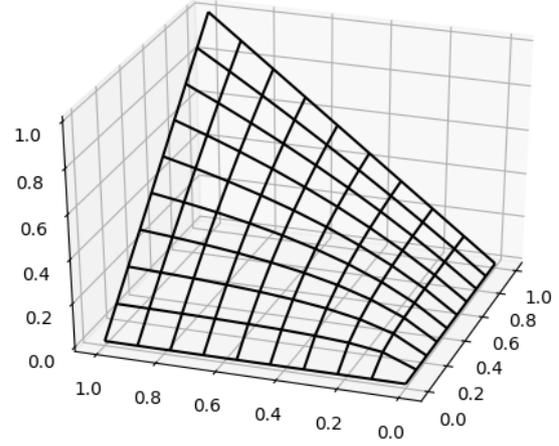


Fig. 4. Plot of the Hamacher T-norm.

**Definition 1** A  $t$ -norm is a function

$$\Delta : [0, 1] \times [0, 1] \longrightarrow [0, 1]$$

satisfying the following properties:

- *Commutativity*:  $x \Delta y = y \Delta x$  for all  $x, y \in [0, 1]$ .
- *Associativity*:  $(x \Delta y) \Delta z = x \Delta (y \Delta z)$  for all  $x, y, z \in [0, 1]$ .
- *Monotonicity*: if  $x_1 \leq y_1$  and  $x_2 \leq y_2$  then  $x_1 \Delta x_2 \leq y_1 \Delta y_2$  for all  $x_1, x_2, y_1, y_2 \in [0, 1]$ .
- *Identity element*:  $1 \Delta a = a$  for all  $a \in [0, 1]$ .

Associativity allows us to extend the definition to finite sequences of values. In addition, monotonicity will play an important role in our formalism: it will ensure that once a small value is obtained (in particular, if we reach a grade of confidence equal to 0), further applications of a  $t$ -norm will not be able to increase the result.

We will use the Hamacher  $t$ -norm, which is defined as  $x * y = \frac{x \cdot y}{x + y - x \cdot y}$  for all  $x, y \in [0, 1]$ . We will use the notation  $\text{Hama}(x, y)$  when we cannot use the symbol notation (for example, this will be the notation in our tool). It can be seen that this function satisfies all the required properties. In particular, by associativity, we will be able to apply this  $t$ -norm to more than two values. Figure 4 shows a plot of this  $t$ -norm.

### 2.3. Fuzzy constraints

Fuzzy constraints will combine relations by using  $t$ -norms to conform complex expressions.

**Definition 2** A fuzzy constraint is a formula consisting of fuzzy relations, possibly combined with  $t$ -norms, which may have free variables. The set of all fuzzy constraints is denoted by  $\mathcal{FC}$ .

Given a mapping from all the free variables to actual values, we can assign a *satisfaction degree* or *GoC* (*Grade of Confidence*) to the constraint. Grades of confidence will be used in the definition of the semantics of fuzzy automata transitions.

**Example 1** An example of a fuzzy constraint is

$$\overline{60 \leq \text{age} \leq 69}^{12}$$

We will use this type of constraints in our model of the heart. This fuzzy constraint has one free variable, *age*, mapped to the age of a patient. If this age falls within the interval  $[60, 69]$ , then the corresponding grade of confidence will be equal to 1. Otherwise, the grade of confidence will linearly decrease with respect to the distance to the interval, until reaching a value equal to 0.

### 2.4. Variables and expressions

In order to generate results from observed data, our fuzzy automata include real valued variables. Automata states will be decorated with mappings from variables to values and automata transitions will be decorated with *variable transformations*.

**Definition 3** Let  $x_1 \dots x_m$  be variables and  $t_1 \dots t_m$  be real valued expressions. A variables transformation, denoted by  $[t_1/x_1, \dots, t_m/x_m]$ , assigns to each variable  $x_i$  the value obtained after evaluating the term  $t_i$ . The evaluations of all the terms are computed before any assignment is performed. The set of variable transformations is denoted by  $\mathcal{VT}$ .

Expressions are formed by variable evaluations, clamped  $t$ -norms (functions that are  $t$ -norms on the interval  $[0, 1]$  and identical to 0 outside the interval), and the reserved variable  $\mu$ . This last reserved variable takes as value the grade of confidence of the fuzzy constraint from the corresponding automata transition (this will be formally defined in Definition 5).

### 2.5. Actions

Every model will define a finite set of actions. This set will be the alphabet of the automaton defined in the model.

**Definition 4** We consider two types of actions: inputs and outputs. On the one hand, inputs consist of their name, followed by a tuple of variable names. On the other hand, outputs consist of their name, followed by a tuple of expressions. Input and output actions are distinguished by their name, with inputs being preceded by ? and outputs being preceded by !.

### 2.6. Fuzzy automata definition

Now we can give a formal definition of our version of fuzzy automata.

**Definition 5** A fuzzy automaton is a tuple

$$(S, \text{Acts}, X, X_0, s_0, T)$$

where its components denote:

- $S$  is a finite set of states.
- $\text{Acts}$  is a finite set of actions, partitioned into a set of inputs  $I$  and a set of outputs  $O$ .
- $X$  is a set of variable names.
- $X_0 : X \rightarrow \mathbb{R}_+$  is the initial mapping from variable names to real values. We call these mappings *variable states*.
- $s_0 \in S$  is the initial state.
- $T \subseteq S \times \text{Acts} \times \mathcal{FC} \times \mathcal{VT} \times S$  is a set of transitions satisfying the following conditions:
  - \* Output transitions have the constant *True* as constraint.
  - \* Input transitions have a constraint whose free variables are a subset of the variables of the input action.

Intuitively, fuzzy automata are directed graphs with transitions decorated with constraints and variable transformations. Before an input action is performed, the associated constraint is evaluated with the variable values mapped by the action. If the grade of confidence is higher than 0, then the transition can be performed and the variable transformations may be applied. When an output action is performed, its expressions are evaluated from the current value of the automaton variables. These ideas will be formalized in the following section when we define an operational semantics.

## 2.7. Automata semantics

Next, we are going to define the operational behavior of fuzzy automata.

**Definition 6** Let  $A = (S, \text{Acts}, X, X_0, s_0, T)$  be a fuzzy automaton and  $\Delta$  be a  $t$ -norm. Let  $s_1, s_2 \in S$  be states and  $X_1, X_2 : X \rightarrow \mathbb{R}$  be variable states.

We have a transition from  $(s_1, X_1)$  to  $(s_2, X_2)$ , after performing the input action  $i \in I$  for the real valued tuple  $\alpha$  with confidence  $\epsilon$ , denoted by

$$(s_1, X_1) \xrightarrow{(i(\alpha))}_\epsilon (s_2, X_2)$$

if the following conditions hold:

- There exists a fuzzy constraint  $C \in \mathcal{FC}$  such that  $(s_1, i, C, VT, s_2) \in T$ .
- The grade of confidence of  $C$ , with the corresponding variables mapped to the values in  $\alpha$ , is higher than 0. We denote it by  $\epsilon$ .
- $X_2$  is the result of applying  $VT$  to  $X_1$  considering:
  - \* The reserved variable  $\mu$  takes the value  $\epsilon$ .
  - \* The values of the variables in the expressions of  $VT$  are taken from  $\alpha$  when the corresponding variable belongs to the input  $i$ . The values for the rest of variables in the expressions of  $VT$  are taken from  $X_1$ .

We have a transition from  $(s_1, X_1)$  to  $(s_2, X_2)$ , after performing the output action  $o \in O$ , denoted by

$$(s_1, X_1) \xrightarrow{(o(\alpha))}_1 (s_2, X_2)$$

if the following conditions hold:

- There exists  $(s_1, o, \text{True}, VT, s_2) \in T$ .
- $X_2$  is the result of applying  $VT$  to  $X_1$  and  $\alpha$  is the tuple of expressions from the output  $o$ , considering:
  - \* The reserved variable  $\mu$  takes the value 1.
  - \* The values of the variables in the expressions are taken from  $X_1$ .

We say that a sequence

$$(s_0, X_0) \xrightarrow{a_1(\alpha_1)}_{\epsilon_1} (s_1, X_1) \cdots \xrightarrow{a_n(\alpha_n)}_{\epsilon_n} (s_n, X_n)$$

of consecutive transitions starting in the initial state of the automaton is a  $\Delta$ -trace of  $A$  if  $\epsilon = \Delta\{\epsilon_1, \dots, \epsilon_n\}$

is greater than zero. In this case, we write

$$(s_0, X_0) \xrightarrow{a_1(\alpha_1), \dots, a_n(\alpha_n)}_\epsilon (s_n, X_n)$$

Next we illustrate the previous notions with some examples.

**Example 2** Consider the automaton depicted in Figure 5, a given  $t$ -norm  $\Delta$  and the sequence of actions:

$$?checkGender(1), ?checkAge(58), !ageRange60()$$

The first two actions are inputs while the last action is an output. We have that

$$(s_0, x_0) \xrightarrow{?checkGender(1)}_1 (s_1, x_1)$$

with  $s_0$  and  $s_1$  being, respectively, the states initial and femaleAgeBranch of the automaton. Note that there is no other possible transition starting from  $s_0$  with that input action because  $\overline{1} = \overline{0} = 0$ . In addition, we have that  $x_0 = x_1$  since that particular transition performs no variable transformation. Next, we have that

$$(s_1, x_1) \xrightarrow{?checkAge(58)}_{0.83} (f60s1, x_2)$$

since  $\overline{60} \leq \overline{58} \leq \overline{69}^{12} = 0.83$ . Since this particular transition performs a variable transformation assigning the reserved variable  $\mu$  to the variable branchGoC, we have that  $x_2 = [\text{branchGoC} \mapsto 0.83]$ . Note that in this case, there are other possible transitions. For example,  $\overline{70} \leq \overline{58} \leq \overline{79}^{14} = 0.14 > 0$ . Now, starting in the state  $f60s1$  we have the transition

$$(f60s1, x_2) \xrightarrow{!ageRange60()}_1 (f60s2, x_3)$$

with  $x_3$  being the variable state given by the two assignments  $[\text{GoC} \mapsto 0.83, \text{branchGoC} \mapsto 0.83]$ . Summarizing, we have that

$$(s_0, x_0) \xrightarrow{?checkGender(1), ?checkAge(58), !ageRange60()}_1 (f60s2, x_3)$$

since  $\Delta(1, 0.83, 1) = 0.83$  for every  $t$ -norm  $\Delta$  (due to the identity property of  $t$ -norms).

**Example 3** Consider again the automaton given in Figure 5. If we consider that  $x_0$  is the variable state  $[\text{min} \mapsto 5]$ , then we have

$$(f60s2, x_0) \xrightarrow{?minute(5)}_1 (f60s3, x_1)$$

Assuming that  $y_0$  is the variable state [ $GoC \mapsto 0.66, branchGoC \mapsto 0.8, min \mapsto 7$ ] and  $y_1$  is [ $GoC \mapsto 0.8, branchGoC \mapsto 0.8, min \mapsto 7$ ], we have

$$(f60s6, y_0) \xrightarrow{?recordAlarm(7,0.66)}_1 (f60s2, y_1)$$

### 3. A formal model of the heart

In this section we apply our theoretical framework to formally specify a model of the heart. This model will be used as an auxiliary tool when checking data of *real* hearts with the final goal of detecting and predicting heart problems. In order to produce a diagnosis, we use the available information and physical evidence collected from electrocardiograms (ECGs).

#### 3.1. Information and data

The information managed by the automaton is the following:

- Gender. We have 2 groups: men and women.
- Age. We have 8 groups of age.
- Heartbeats. The range of correct heartbeats per minute (BPM) for healthy patients, according to their gender and age.
- RR intervals. The range of correct RR interval duration (measured in milliseconds) for healthy patients, according to their gender, age and BPM.

The action set of our automaton is the following:

- $?checkGender(gen)$ . It reads the gender of the patient.
- $?checkAge(age)$ . It reads the age of the patient.
- $?minute(m)$ . It reads the current minute of the recording.
- $?readBPM(bpm)$ . It reads the amount of beats in the current minute.
- $?readRR(rr)$ . It reads the next RR interval in the current minute.
- $?noMorePendingRR()$ . It receives a notification that there are no more RR intervals in the current minute.
- $?endOfRecord()$ . It receives a notification to denote that there are no more minutes in the analyzed record.
- $!recordAlarm(min, GoC)$ . It indicates, with a grade of confidence equal to  $GoC$ , that an alarm will be raised in the  $min$  minute.

- $!ok(min, GoC)$ . It indicates, with a grade of confidence equal to  $GoC$ , no alarm will be recorded for the  $min$  minute.
- $!ageRange20()$ . It indicates that the automaton is entering the 20 – 29 age branch.
- ...
- $!ageRange80()$ . It indicates that the automaton is entering the 80 – 89 age branch.
- $!ageRange90()$ . It indicates that the automaton is entering the > 90 age branch.

#### 3.2. Automaton topology

Next, we describe the main lines of the topology of our automaton. This fuzzy automaton has a total of 132 states and 274 transitions. Initially, our automaton has two transitions: one per gender. After that, each branch has 8 transitions, one per age group. Each of these age groups has an associated sub-automaton with 8 states and a transition to a common *final* state to denote that all the data for this patient have been processed. Each minute, the automaton checks if the number of beats falls within the normal amount of beats per minute in the age and gender range. If it does, then the automaton does not take into account the RR intervals in that minute and signals that minute to be *ok*. Otherwise, the automaton processes each RR interval. Every RR interval out of range will decrease the grade of confidence in the trace signaling an *ok*. If there is at least one interval out of range, an alarm is raised with the grade of confidence obtained at that point. If a state having as suffix  $s6$  (e.g. the state  $f60s6$ ) is reached, then the automaton will record the current state of the patient as a case in which he or she suffers the risk of having a heart problem (formally, an output action indicating that an alarm is raised will be performed). Therefore, for each patient, the number of alarms that the automaton can raise is bounded by the number of minutes in the recording (our typical samples have a duration of around 30 minutes).

#### 3.3. Fuzzy relation metaparameters

The values used to define our fuzzy constraints are taken from previous work. Normal ranges for heartbeats per minute, classified by gender, have been gathered from the work of Palhares et al. [19] because they are the most recent that we were able to find but, as we said in the introduction of the paper, other ranges are available although the obtained results are very similar.

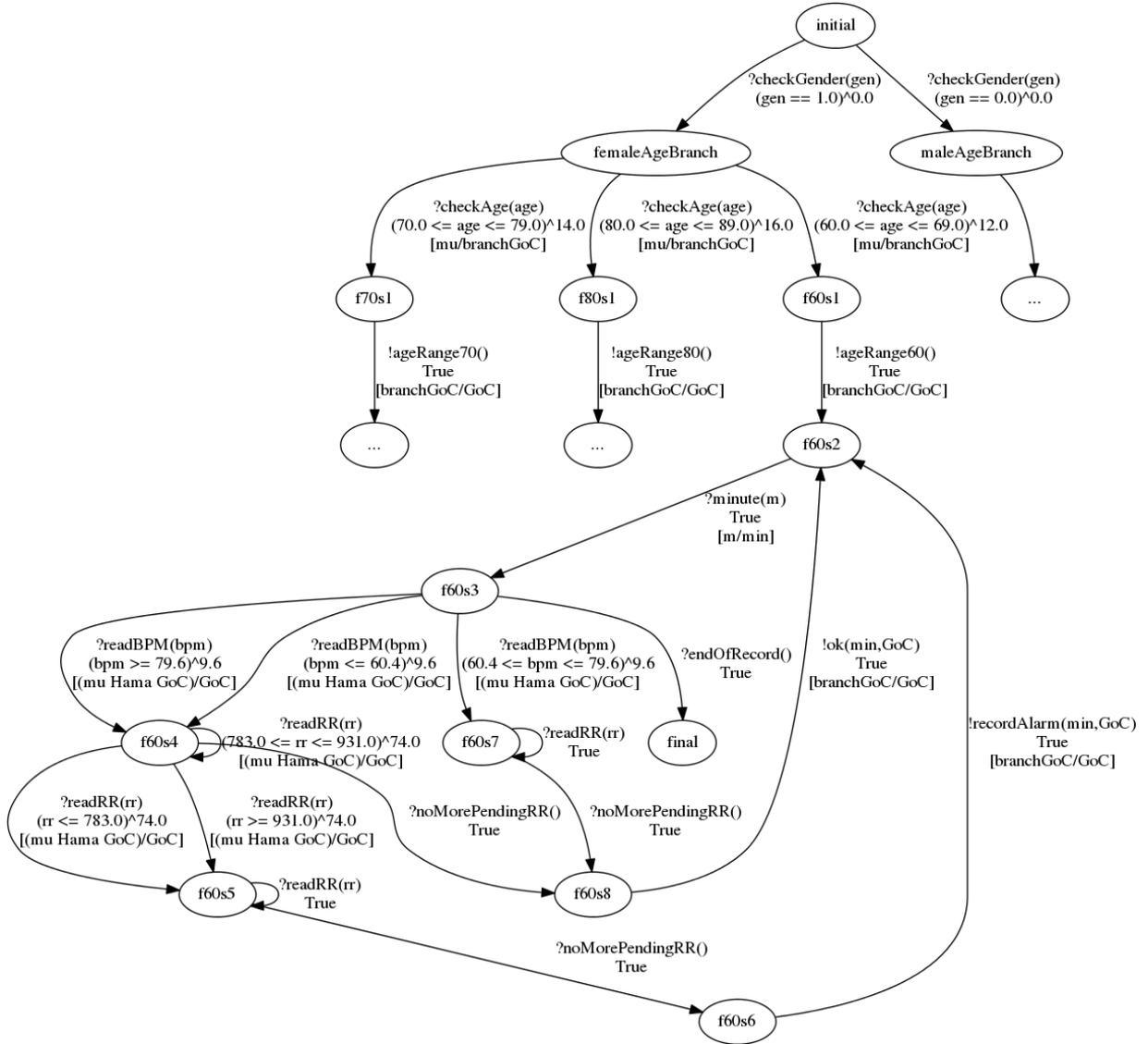


Fig. 5. Branch of the automaton corresponding to female patients between 60 and 69 years old.

In the case of the age, the  $\delta$  value is obtained from the 20% of the lowest value of each age range. The idea is that it is possible to wrongly classify a patient according to their age. For example, if we have a 53 years old male patient then we should classify him in the age group between 50 and 59. In addition, it may happen that the patient has a very healthy life style and, therefore, his/her heart is *younger*. Therefore, we should also classify these patients in the previous age group and decide whether their recorded data fits better in their *real* age group or in the *closest* to the real one. We have the following considerations:

- In the case of heartbeats, we had the median, 1st percentile and 99th percentile from the database.
- It is shown in Hozo et al [10] that  $range/6$  is a good estimator for the standard deviation when the sample is big enough.
- The median is also shown to be a good estimator for the mean.
- For these reasons, the midpoint in our fuzzy intervals is located in the median and the constant  $\delta$  is obtained by the following formula

$$\delta = \frac{b - a}{6}$$

where  $a$  is the 1st percentile and  $b$  is the 99th percentile.

- Concerning RR interval lengths, in order to determine the constant  $\delta$  of the *fuzzy intervals* we have used the standard deviation shown in the work of Haarmark et al. [8].
- The midpoint from these *fuzzy intervals* is fixed to be

$$RR = \frac{60000}{bpm}$$

with bpm being the BPM medians obtained from the already mentioned work of Palhares et al. [19].

#### 4. A methodology to use fuzzy automata and their associated data

In this section we explain the methodology that we use to analyze the results obtained by feeding data to a fuzzy automaton. We will show how the methodology is applied to our case study, that is, the analysis of BPMs and RR wave durations in our model of the heart, but it can be applied to any fuzzy automaton and a coherent set of traces.

##### 4.1. The methodology

The proposed methodology consists of the following steps:

1. A fuzzy automaton is defined.
2. The data source is formatted as a sequence of input actions.
3. The output actions will be inserted into the sequence in the appropriate places.
4. The complete sequence is fed to the automaton. If at a certain point the grade of confidence falls to 0, then the automaton rejects the sequence. Otherwise, the automaton records the actual values taken by the expressions associated to the output actions. The automaton also records the grade of confidence obtained after each step.
5. The resulting values are conveniently recorded in a tabular format and rendered in some graphical format.
6. The evidence obtained from the obtained tables and graphs is used to guide further analysis, possibly with the intervention of an expert.

```
?checkGender(1),?checkAge(77),!ageRange60,
?minute(1), ?readBPM(50),
?readRR(277),?readRR(397),?readRR(1250),?readRR(1183),
?readRR(1202),?readRR(1213),?readRR(1263),?readRR(1225),
?readRR(1172),?readRR(1213),?readRR(1144),?readRR(1216),
?readRR(1172),?readRR(1224),?readRR(1169),?readRR(1247),
?readRR(1191),?readRR(1233),?readRR(1202),?readRR(1219),
?readRR(1188),?readRR(1211),?readRR(1166),?readRR(1216),
?readRR(1183),?readRR(1224),?readRR(1183),?readRR(1194),
?readRR(1208),?readRR(1205),?readRR(1183),?readRR(1177),
?readRR(1208),?readRR(1191),?readRR(1188),?readRR(1258),
?readRR(1216),?readRR(1238),?readRR(1252),?readRR(1255),
?readRR(1200),?readRR(1250),?readRR(1219),?readRR(1247),
?readRR(1236),?readRR(1202),?readRR(1258),?readRR(1247),
?readRR(1261),?readRR(1250),
?noMorePendingRR(),
!ok,
?minute(2), ?readBPM(49),
?readRR(1191),?readRR(1227),?readRR(1219),?readRR(1266),
...
```

Fig. 6. Aspect of a sequence with the *!ageRange60* action before being processed by the automaton.

```
?checkGender(1), 1; ?checkAge(77), 0.33; !ageRange60(), 0.33;
?minute(1), 0.33; ?readBPM(50), 0.33;
?readRR(277), 0;
...
...
```

Fig. 7. Aspect of a sequence with the *!ageRange60* action after being processed by the automaton.

```
?checkGender(1),?checkAge(77),!,
?minute(1), ?readBPM(50),
?readRR(277),?readRR(397),?readRR(1250),?readRR(1183),
?readRR(1202),?readRR(1213),?readRR(1263),?readRR(1225),
?readRR(1172),?readRR(1213),?readRR(1144),?readRR(1216),
?readRR(1172),?readRR(1224),?readRR(1169),?readRR(1247),
?readRR(1191),?readRR(1233),?readRR(1202),?readRR(1219),
?readRR(1188),?readRR(1211),?readRR(1166),?readRR(1216),
?readRR(1183),?readRR(1224),?readRR(1183),?readRR(1194),
?readRR(1208),?readRR(1205),?readRR(1183),?readRR(1177),
?readRR(1208),?readRR(1191),?readRR(1188),?readRR(1258),
?readRR(1216),?readRR(1238),?readRR(1252),?readRR(1255),
?readRR(1200),?readRR(1250),?readRR(1219),?readRR(1247),
?readRR(1236),?readRR(1202),?readRR(1258),?readRR(1247),
?readRR(1261),?readRR(1250),
?noMorePendingRR(),
!ok,
?minute(2), ?readBPM(49),
?readRR(1191),?readRR(1227),?readRR(1219),?readRR(1266),
...
```

Fig. 8. Aspect of a sequence with the *!* action before being processed by the automaton.

```
?checkGender(1), 1; ?checkAge(77), 1; !ageRange70(), 1;
?minute(1), 1; ?readBPM(50), 1;
?readRR(277), 0;
```

Fig. 9. Aspect of a sequence with the *!* action after being processed by the automaton.

##### 4.2. Input sequence generation

In this section we discuss how sequences fed to an automaton are produced and processed.

In order to assess the usefulness of our model of the heart, we used the MIT/BIH Arrhythmia Database Di-

rectory [17].<sup>1</sup> This study includes 48 ECG recordings with a duration of 30 minutes from the Massachusetts Institute of Technology - Beth Israel Hospital arrhythmia database. All of them present some heart pathology. Specifically, 48% of the samples have been annotated in the database as representative cases of routine clinical recordings while the remaining 52% reflect uncommon cases of arrhythmias. The access and processing of this data has been done with the aid of the *Python WFDB package*.<sup>2</sup>

When a sequence is generated, the values associated with input actions are obtained from the ECG recording and the output actions are chosen depending on which particular event we are trying to check. For example, if we add the *!ageRange60* output then the automaton is forced to enter the 60–69 branch. An example of this situation can be seen in Figures 6 and 7. On the contrary, if the *!ageRange60* action is substituted by the unnamed action *!*, then the automaton will enter the branch with the highest grade of confidence. This situation is shown in Figures 8 and 9. Let us note that the *!ok* action discards all executions not ending in an *ok* signal for the minute 1. The sequence asking for the confidence of an alarm being raised on minute 1 would therefore be obtained by substituting the *!ok* action by an *!alarm* action.

#### 4.3. Patient analysis

In this section we analyze the data corresponding to two different patients. Table 1 contains all the generated grades of confidence. Each cell contains two numbers: the first one is the Hamacher GoC of sending an *ok* signal while the second one is the Hamacher GoC of raising an *alarm*, both referring to that minute and age/gender branch. This table clearly shows why our new approach represents a big step forward with respect to our previous work: we are able to produce and appropriately process several dozens of *ok*/*alarm* signals.

Concerning these two chosen patients, we have that patient 124 is a 77 years old female patient while patient 108 is a 87 years old female patient. According to the age range, there are four branches where we initially obtain a positive grade of confidence for patient 124 while there are 3 positive branches for patient 108.

Each line appearing in Figure 10 represents, at each given minute, the highest possible  $\epsilon$  such that the global execution of the automaton forms a trace decorated with that particular  $\epsilon$ , that is, the  $t$ -norm of each GoC value from all transitions in the trace. While Table 1 can be interpreted as a measure of the most likely diagnostic at a certain minute, Figure 10 is meant to indicate how well patient 108 does fit into each particular age branch. This figure shows that the branch with the highest GoC for patient 108 is the 70–79 branch (that is, one that does not correspond to its real age). This figure also shows how each branch starts with an initial grade of confidence depending on the distance from the actual age to the age interval and we can see how the 70–79 branch becomes higher than the branch corresponding to the actual age of patient 108. This fact may indicate that the patient has a younger heart tissue than expected according to her age. We can also observe that minute 2 represents the steepest drop in the grade of confidence for all *!ok* branches of patient 108.

Finally, Figure 11 shows the second minute of the ECG recording. This plot has been obtained using the *plot\_wfdb* subroutine from the *Python WFDB package*, applied to patient 108 from *mitdb*.

## 5. Conclusions and future work

In this paper we have introduced a new variant of fuzzy automata. The main goal of the formalism is to provide a framework to formally specify and analyze systems where information must be considered to be imprecise, that is, fuzzy. We have defined a methodology to analyze systems modelled as fuzzy automata. We have applied our framework to model the heart and analyze the data corresponding to different patients.

This paper represents a notable extension of our previous work [2]. Although we expect that the syntax and semantics are already stable, after the refinement introduced in this paper, there are several lines to continue our work. First, we would like to produce a user-friendly tool to support our framework. Obviously, the management of data has been automatized via a software developed by us, but this software does not provide an easy-to-use interface. In this line, we are currently working on an expert system that should be able to synthesize all the computed data into a small amount of (useful) information. We are taking as initial step our previous work on passive testing [9,15], in particular concerning the development of tools, because what we do when we feed a sequence of actions to an automaton can be seen as a form of passive testing.

<sup>1</sup>The data can be accessed at <https://physionet.org/physiobank/database/mitdb/>.

<sup>2</sup>The package can be found at <https://github.com/MIT-LCP/wfdb-python>.

	#124				#108		
	(60-69)	(70-79)	(80-89)	(>90)	(70-79)	(80-89)	(>90)
min. 1	0.00/0.33	0.00/1.00	0.00/0.81	0.00/0.28	0.43/0.39	1.00/0.95	0.75/0.83
min. 2	0.00/0.33	0.00/1.00	0.00/0.81	0.00/0.28	0.23/0.43	0.29/1.00	0.18/0.83
min. 3	0.00/0.33	0.00/1.00	0.00/0.81	0.00/0.28	0.40/0.43	0.76/1.00	0.55/0.83
min. 4	0.00/0.33	0.00/1.00	0.00/0.81	0.00/0.28	0.31/0.43	0.48/1.00	0.34/0.83
min. 5	0.23/0.33	0.33/1.00	0.27/0.81	0.13/0.28	0.38/0.43	0.67/1.00	0.48/0.83
min. 6	0.33/0.33	0.85/1.00	0.65/0.81	0.24/0.28	0.31/0.43	0.48/1.00	0.34/0.83
min. 7	0.20/0.33	0.23/1.00	0.18/0.81	0.08/0.28	0.23/0.43	0.29/1.00	0.18/0.83
min. 8	0.16/0.33	0.13/1.00	0.09/0.81	0.02/0.28	0.40/0.43	0.76/1.00	0.55/0.83
min. 9	0.16/0.33	0.13/1.00	0.09/0.81	0.02/0.28	0.40/0.43	0.76/1.00	0.55/0.83
min. 10	0.28/0.33	0.54/1.00	0.43/0.81	0.19/0.28	0.28/0.43	0.38/1.00	0.26/0.83
min. 11	0.28/0.33	0.54/1.00	0.43/0.81	0.19/0.28	0.31/0.43	0.48/1.00	0.34/0.83
min. 12	0.16/0.33	0.13/1.00	0.09/0.81	0.02/0.28	0.35/0.43	0.57/1.00	0.41/0.83
min. 13	0.20/0.33	0.23/1.00	0.18/0.81	0.08/0.28	0.38/0.43	0.67/1.00	0.48/0.83
min. 14	0.10/0.33	0.02/1.00	0.00/0.81	0.00/0.28	0.42/0.43	0.86/1.00	0.62/0.83
min. 15	0.10/0.33	0.02/1.00	0.00/0.81	0.00/0.28	0.35/0.43	0.57/1.00	0.41/0.83
min. 16	0.16/0.33	0.13/1.00	0.09/0.81	0.02/0.28	0.40/0.43	0.76/1.00	0.55/0.83
min. 17	0.02/0.33	0.00/1.00	0.00/0.81	0.00/0.28	0.35/0.43	0.57/1.00	0.41/0.83
min. 18	0.28/0.33	0.54/1.00	0.43/0.81	0.19/0.28	0.43/0.42	0.95/1.00	0.68/0.83
min. 19	0.02/0.33	0.00/1.00	0.00/0.81	0.00/0.28	0.43/0.22	1.00/0.48	0.83/0.61
min. 20	0.02/0.33	0.00/1.00	0.00/0.81	0.00/0.28	0.35/0.43	0.57/1.00	0.41/0.83
min. 21	0.10/0.33	0.02/1.00	0.00/0.81	0.00/0.28	0.35/0.43	0.57/1.00	0.41/0.83
min. 22	0.16/0.33	0.13/1.00	0.09/0.81	0.02/0.28	0.43/0.31	1.00/0.67	0.83/0.74
min. 23	0.23/0.33	0.33/1.00	0.27/0.81	0.13/0.28	0.43/0.31	1.00/0.67	0.83/0.74
min. 24	0.30/0.33	0.65/1.00	0.50/0.81	0.21/0.28	0.43/0.39	1.00/0.95	0.75/0.83
min. 25	0.28/0.33	0.54/1.00	0.43/0.81	0.19/0.28	0.42/0.43	0.86/1.00	0.62/0.83
min. 26	0.28/0.33	0.54/1.00	0.43/0.81	0.19/0.28	0.31/0.43	0.48/1.00	0.34/0.83
min. 27	0.26/0.33	0.44/1.00	0.35/0.81	0.16/0.28	0.40/0.43	0.76/1.00	0.55/0.83
min. 28	0.33/0.33	0.85/1.00	0.65/0.81	0.24/0.28	0.43/0.39	1.00/0.95	0.75/0.83
min. 29	0.33/0.31	1.00/0.94	0.78/0.81	0.26/0.28	0.43/0.00	1.00/0.19	0.83/0.40
min. 30	0.26/0.33	0.44/1.00	0.35/0.81	0.16/0.28	0.43/0.09	1.00/0.10	0.83/0.32

Table 1

Results corresponding to patients #124 and #108. Each cell includes a pair *GoC ok/GoC alarm* computed from the data observed during that minute.

## References

- [1] C. Andrés, L. Llana, and M. Núñez. Self-adaptive fuzzy-timed systems. In *13th IEEE Congress on Evolutionary Computation, CEC'11*, pages 115–122. IEEE Computer Society, 2011.
- [2] I. Calvo, M. G. Merayo, and M. Núñez. An improved and tool-supported fuzzy automata framework to analyze heart data. In *10th Asian Conference on Intelligent Information and Database Systems, ACIIDS'18, LNAI 10751*, pages 694–704. Springer, 2018.
- [3] A. Camacho, M. G. Merayo, and M. Núñez. Collective intelligence and databases in eHealth: A survey. *Journal of Intelligent & Fuzzy Systems*, 32(2):1485–1496, 2017.
- [4] A. Camacho, M. G. Merayo, and M. Núñez. Using fuzzy automata to diagnose and predict heart problems. In *19th IEEE Congress on Evolutionary Computation, CEC'17*, pages 846–853. IEEE Computer Society, 2017.
- [5] A. R. Cavalli, T. Higashino, and M. Núñez. A survey on formal active and passive testing with applications to the cloud. *Annales of Telecommunications*, 70(3-4):85–93, 2015.
- [6] M. J. Dewhurst, L. Y. Di Marco, F. Dewhurst, P. C. Adams, A. Murray, G. P. Orega, J. C. Mwita, R. W. Walker, and P. Langley. Electrocardiographic reference values for a population of older adults in sub-Saharan Africa. *Annals of Noninvasive Electrocardiology*, 19(1):34–42, 2014.
- [7] M. Doostfatemeleh and S. C. Kremer. New directions in fuzzy automata. *International Journal of Approximate Reasoning*, 38(2):175–214, 2005.
- [8] C. Haarmark, C. Graff, M. Andersen, T. Hardahl, J. Struijk, E. Toft, J. Xue, G. Rowlandson, P. Hansen, and J. Kanters. Reference values of electrocardiogram repolarization variables in a healthy population. *Journal of electrocardiology*, 43(1):31–39, 2010.
- [9] R. M. Hierons, M. G. Merayo, and M. Núñez. An extended framework for passive asynchronous testing. *Journal of Logical and Algebraic Methods in Programming*, 86(1):408–424, 2017.
- [10] S. Hozo, B. Djulbegovic, and I. Hozo. Estimating the mean

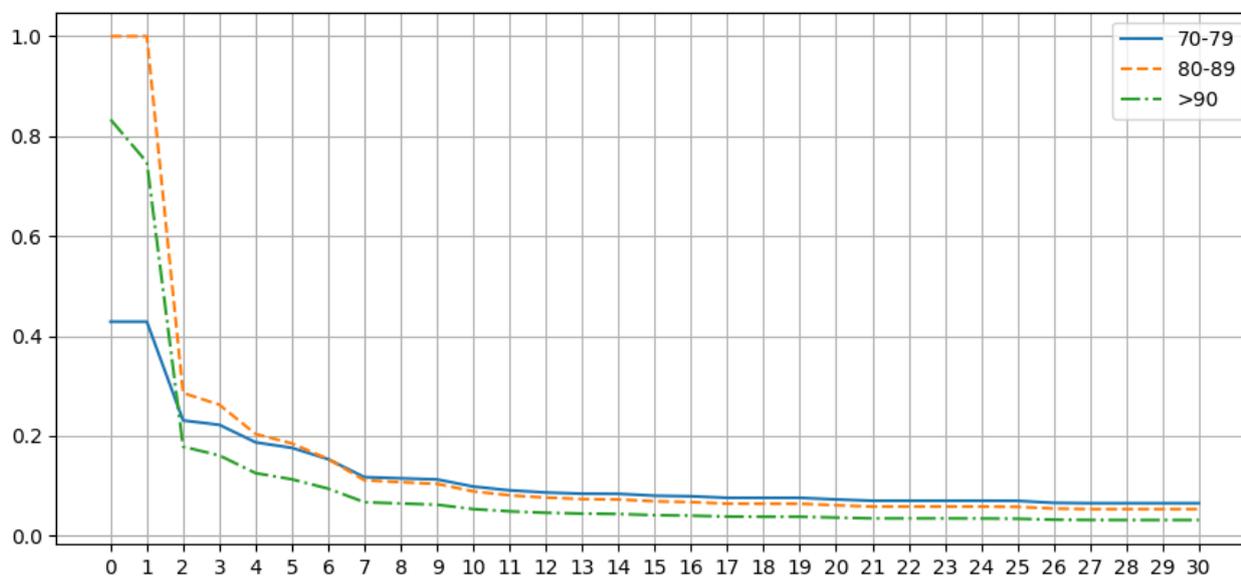


Fig. 10. Progress of the GoC for each alternative branch (patient 108).

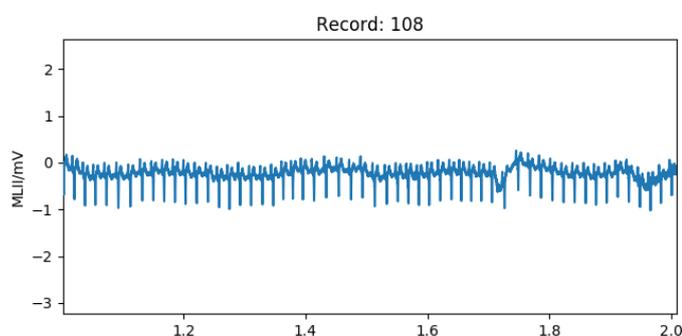


Fig. 11. Electrocardiogram from the second minute of the patient 108 recording.

and variance from the median, range, and the size of a sample. *BMC medical research methodology*, 5(1):1, 2005.

- [11] P. Hunter, A. Pullan, and B. Smaill. Modeling total heart function. *Annual review of biomedical engineering*, 5(1):147–177, 2003.
- [12] Z. Jiang, A. Connolly, and R. Mangharam. Using the virtual heart model to validate the mode-switch pacemaker operation. In *32nd Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society, EMBC'10*, pages 6690–6693. IEEE Computer Society, 2010.

- [13] I. Katibi, E. N. Clark, B. Devine, S. M. Lloyd, and P. W. Macfarlane. Normal limits of the electrocardiogram in Nigerians. *Journal of Electrocardiology*, 46(4):289–295, 2013.
- [14] P. W. Macfarlane, S. M. Lloyd, D. Singh, S. Hamde, E. N. Clark, B. Devine, B. G. Francq, and V. Kumar. Normal limits of the electrocardiogram in Indians. *Journal of Electrocardiology*, 48(4):652–668, 2015.
- [15] M. G. Merayo, R. M. Hierons, and M. Núñez. A tool supported methodology to passively test asynchronous systems with multiple users. *Information & Software Technology*, 104:162–178, 2018.
- [16] D. Méry and N. Singh. Formalization of heart models based on the conduction of electrical impulses and cellular automata. In *1st Int. Symp. on Foundations of Health Informatics Engineering and Systems, FHIES'11, LNCS 7151*. Springer, 2012.
- [17] G. Moody and R. Mark. The impact of the MIT-BIH arrhythmia database. *IEEE Engineering in Medicine and Biology*, 20(3):45–50, 2001.
- [18] J. N. Mordeson and D. S. Malik. *Fuzzy automata and languages: theory and applications*. Chapman & Hall/CRC, 2002.
- [19] D. M. F. Palhares, M. S. Marcolino, T. M. M. Santos, J. L. P. da Silva, P. R. Gomes, L. B. Ribeiro, P. W. Macfarlane, and A. L. P. Ribeiro. Normal limits of the electrocardiogram derived from a large database of Brazilian primary care patients. *BMC Cardiovascular Disorders*, 17:article 152, 2017.
- [20] P. Rijnbeek, G. van Herpen, M. Bots, S. Man, N. Verweij, A. Hofman, H. Hillege, M. Numans, C. Swenne, J. Witteman, et al. Normal values of the electrocardiogram for ages 16–90 years. *Journal of electrocardiology*, 47(6):914–921, 2014.
- [21] M. Shafique and Y. Labiche. A systematic review of state-

- based test tools. *International Journal on Software Tools for Technology Transfer*, 17(1):59–76, 2015.
- [22] W. G. Wee and K. S. Fu. A formulation of fuzzy automata and its application as a model of learning systems. *IEEE Transactions on Systems Science and Cybernetics*, 5(3):215–223, 1969.
- [23] J. Wu, J. A. Kors, P. R. Rijnbeek, G. van Herpen, Z. Lu, and C. Xu. Normal limits of the electrocardiogram in Chinese subjects. *International Journal of Cardiology*, 87(1):37–51, 2003.
- [24] L. A. Zadeh. *Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems*. Advances in Fuzzy Systems - Applications and Theory: vol. 6. World Scientific Press, 1996.