Towards Designing A Hierarchical Fuzzy System for Early Diagnosis of Heart Disease

Tajul Rosli Razak1*, Nurika Filzah Mahadi2, Iman Hazwam Abd Halim3, Muhammad Nabil Fikri Jamaluddin4, Mohammad Hafiz Ismail5, Shukor Sanim Mohd Fauzi6, Ray Adderley JM Gining7
1,2,3,4,5,6,7 Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Perlis Branch, Malaysia

Corresponding author: *tajulrosli@uitm.edu.my
Received Date: 29 October 2019
Accepted Date: 11 November 2019

ABSTRACT

Heart disease may represent a range of conditions that affect our heart. Disease under heart diseases umbrella include coronary heart disease, heart attack, congestive heart failure, and congenital heart disease, is the leading cause of death. Moreover, heart disease not only attacks the elderly. In the present day, lots of younger people might be getting affected by the number of heart diseases. In order to decrease the mortality rate caused by heart disease, it is necessary for the disease, to be diagnosed at an early stage. In this paper, we have proposed the use of hierarchical fuzzy systems (HFSs) for early diagnosis of heart disease. However, to design the HFSs is challenging, especially for the complex system. Therefore, in this paper, we focus on designing a hierarchical fuzzy system to handle the complex medical application. The designed HFS consists of six key main steps implemented on heart disease. The input variables of heart disease includes shortness of breath, discomfort, pressure, heaviness, or pain in the chest, arm, or below the breastbone, fatigue, nausea, difficulties in climbing stairs, swelling in ankles, difficulty to sleep at night, irregular heartbeats, fullness, sweating, take frequent break during the day, dizzy and depressed. Additionally, the output of heart disease is to classify whether the patient is healthy or suspecting with heart disease. The study contributes to providing insight into a way of designing the HFSs, particularly for the complex medical application.

Keywords: Heart disease, early diagnosis, hierarchical fuzzy systems

INTRODUCTION

Heart disease relates to a health problem of the heart and vascular system. The heart is an essential organ in our body because of its purpose as blood blower. There are numerous factors behind the heart problems, which include behaviours of lifestyle, congenital malformations at birth, as well as unhealthy eating behaviours.

According to the American Heart Association (2007), America’s number one killer is heart attack which resulting from shut down of coronary heart disease. Moreover, heart disease not only attacks the elderly. In the present day, lots of younger people might be getting affected by the number of heart diseases. Furthermore, newborn babies might encounter abnormalities of the heart. There are fifty more various types of heart problems eyeing the ones with a pattern of living and following a healthy food plan. There is no such existing system or tool that sufferers of heart disease can diagnose and detect the disease by themselves yet. It mostly depends on the medical institution to diagnose it – diagnosis by a medical doctor. Therefore, there is a need to have a system that can perform an early diagnosis of heart disease.
A Fuzzy logic system (FLS) becomes a popular method in modelling with uncertainty and imprecision information. Also, one of its strength is “interpretability” particularly in applications such as knowledge extraction and decision support (Nauck & Kruse, 1999). Interpretability refers to the capability of FLSs to express the behaviour of the system in an understandable way (Casillas, Cordón, Herrera, & Magdalena, 2003). This is due to the fact that the FLS use linguistic variable and rules that are very close to human language. Substantial research has shown that FLS is proven reliable to be utilised in the medical domain (Geman, Turcu, & Graur, 2013; Mostafa, Mustapha, Mohammed, Ahmad, & Mahmoud, 2018; Razak, Wahab, & Ramli, 2013). However, recently, FLS face the issue with the problem of the “curse of dimensionality” that is the number of rules is increasing exponentially with the number of the input variables, particularly in the complex system having many input variables.

A Hierarchical Fuzzy System (HFS) was introduced to overcome the problem of the curse of dimensionality that appears in the conventional FLS when dealing with a complex system (Raju, Zhou, & Kisner, 1991). The HFS is a special type of FLS that has a unique property, that is the number of rules is increased linearly with the number of input variables. Consequently, it may reduce the number of rules in the FLS and thus avoid the problem of the curse of dimensionality – rule explosion. However, to design the HFS is not an easy task because of its structure that is having multiple subsystems, layers, intermediate variables and different topology (Torra, 2002).

In this paper, the use of HFS is explored to be adopted in the complicated medical application. Specifically, we put forward an approach to guide the design of HFS for early diagnosis of heart disease, consisting of 6 key steps. Note that the heart disease consists of 13 symptoms (input variables) and one primary output, that is to analyse whether the patient has the risk of heart disease at the early stage.

The rest of this paper is organised as follows. The second section discusses the background to heart disease, fuzzy logic systems and hierarchical fuzzy systems. This is followed by the third section that introduces an approach to design a hierarchical fuzzy system for early diagnosis of heart disease, consisting of six essential steps. In the fourth section, the discussion of the design steps is carefully explained. Finally, the fifth section presents the conclusions and future works.

**BACKGROUND**

In this section, the background in respect to the heart disease, fuzzy logic system, and hierarchical fuzzy system are briefly described.

1. **Heart Disease**

Heart disease is a kind of health problem that involve the heart as well as blood vessels. Heart disease is the top cause of mortality in the United States. Over 600,000 people in America die of heart disease annually. That is one in every four deaths in this particular country (Kochanek, Miniño, Murphy, Xu, & Kung, 2011).

Heart disease points to any health problem that risks the cardiovascular system. The factors that cause heart disease are numerous, but atherosclerosis and hypertension are the most common. In addition, with ageing come some physiological and morphological effects that affect cardiovascular performance and then result in a consequently increased risk of cardiovascular disease, even in healthy asymptomatic people.

There are various kinds of heart diseases. Some of the common heart diseases are coronary heart disease. It contributes to severe chest pain and also discomforts specifically named as angina. Second, myocardial
Infarction or also known heart attack develops when heart muscle cells die due to blood circulation to the heart is interrupted. Coronary heart disease is a cause of heart disease. On top of that, heart failure is also due to a decrease in heart pumping. The last but not least is cardiac arrhythmia is an uncertain heartbeat (Portal Rasmi Jabatan Kesihatan Pulau Pinang, 2009).

Signs or symptoms differ depending on the types of heart disease. For many individuals, chest discomfort or heart attack is the first warning sign. Anyone having a heart attack may suffer from a number of symptoms, such as chest pain or discomfort which does not be gone after several minutes, the pain or discomfort in the jaw, neck, or even backside, fatigue, light-headedness, nausea or vomiting (feeling sick to your stomach), or cold sweat, pain or distress in the arms as well as shoulders and also difficulty of breathing. Therefore, diagnosis may be made by looking at the symptoms experienced by the patient.

ii. Fuzzy Logic Systems

FLSs are one of the currently used techniques for modelling non-linear, uncertain and complex systems. An essential characteristic of FLSs is the partitioning of the space of system variables into fuzzy regions using fuzzy sets (Zadeh, 1965). In each region, the characteristics of the system can be described merely using a rule. Generally, an FLS consists of a rule base with rules associated with particular regions, where the information available is transparent and easily readable. This characteristic of fuzzy systems has been employed in many fields including medical (Bárdossy et al., 2014; Razak et al., 2013), engineering (Gad & Farooq, 2001), decision support (O. W. Samuel, Omisore, & Ojokoh, 2013), pattern recognition (Pedrycz, 1990) and others.

In medical domain, FLS has been successfully applied in many areas such as Breast Cancer (Khan, Choi, Shin, & Kim, 2008; Peña-Reyes & Sipper, 1999; Phillips et al., 2006), Heart Disease (Anooj, 2012; Bhatla & Jyoti, 2012; Stetson, 2003), Lung cancer (Phillips et al., 2007; C. C. Samuel, Saravanan, & Devi, 2007; Schneider et al., 2002), Liver (Badawi, Derbala, & Youssef, 1999; Polat, Şahan, Kodaz, & Güneş, 2007; Shioni, Kuroki, Kuriyama, Morikawa, & Masaki, 1997) and Diabetes (Ganji & Abadeh, 2010, 2011; Lee & Wang, 2011). From the published literature, it may be concluded that fuzzy logic has proven reliable and useful in many medical domains since it can address the specific problem of disease classification in the presence of uncertain or vague knowledge of a linguistic nature (John & Innocent, 2005).

However, because of the rapid development in a complex system problem with a large set of input variables, the conventional FLS cannot cope with the increase of the total number of rules concerning the computational time in fuzzy inference (Raju et al., 1991). At present, the crucial issues in the conventional FLS are how to reduce the number of rules involved and their corresponding computation requirement. One effective way to deal with this problem is through the use of a special type of FLS, namely hierarchical fuzzy systems (HFSs) (Wang, 1998).

iii. Hierarchical Fuzzy Systems

Hierarchical fuzzy systems (HFSs) were introduced and played an extraordinary approach to overcome the problem of the curse of dimensionality arise in FLSs (Raju et al., 1991). The idea of HFS is to put the input variables into a collection of low-dimensional fuzzy logic systems, instead of creating a single high dimensional rule base for a fuzzy logic system. Each low-dimensional fuzzy logic control system constitutes a level in the HFS. Thus, by implementing this approach, it may prevent the problem of rule explosion.
HFS has a significant property that the total number of rules increases linearly rather than exponentially as in the conventional fuzzy system and consist of multiple low-dimensional fuzzy systems in hierarchical form. In HFS, if we define $m$ fuzzy sets for each variable, then each low-dimensional fuzzy system consist of $m^2$ rules; therefore, the total number of rules is $(n - 1)m^2$ which is a linear function of the number of input variables $n$ as shown in Figure 2.

The first level hierarchy (FLS1) gives an approximate output $y_1$ obtained from input variables $(x_1, x_2)$, which then modified by the second-level hierarchy (FLS2). The inputs to the latter are the output of the first level hierarchy and other system variables. This process, as shown in Figure 2, is repeated at succeeding levels of the hierarchy to eventually produce a control output (Raju et al., 1991). From these Figures 1 and 2, it can be concluded that HFS is capable of overcoming the current problem in conventional FLS such that it can reduce the number of rules in their inference system in comparison to FLS. Also, HFS can give a similar output in the conventional FLS with less time-consuming.

**AN APPROACH FOR DESIGNING A HIERARCHICAL FUZZY SYSTEM**

Despite the fact the HFS has advantages over standard FLS in term of reducing the model complexity, it is challenging to design an HFS, especially for a complicated medical applying having many input variables. Besides, (Torra, 2002) has claimed that building an HFS is a challenging task because of the need to define the architecture of the HFS (the subsystems, the input variables of each subsystem, and the interactions between each subsystem), as well as the rules of each subsystem. Therefore, as an introductory approach, we are proposing a way of designing an HFS for early diagnosis of heart disease consist of 6 key steps, as shown in Figure 3.
Figure 3: Stepwise for designing HFS

Step 1: Interviewing Expert

Step 2: Classify inputs and output variables

Step 3: Construct membership functions for the input and output variables

Step 4: Form the hierarchical structure – topology

Step 5: Compose the rule bases of the subsystems

Step 6: Link-up the subsystem in the hierarchical fuzzy systems.

\[\text{i. Step 1 – Interviewing Experts}\]

In Step 1, information from a domain expert, which is a doctor will be obtained. In this phase, an interview technique will be used to obtain the information from a doctor. The doctor will be interviewed to get the correct information about heart disease, particularly on the symptoms of this disease. Figure 1 shows the process of interviewing the medical expert, Dr Rosmawati, who is a medical doctor at UiTM Perlis Branch. The information from this interview is essential as it will ease the process of constructing membership functions, HFS topology and rules.

Figure 4: Interviewing a medical expert
ii. **Step 2 – Classify inputs and output variables**

Following the information obtained from Step 1, all inputs and output are classified in Step 2. In this study, the 13 symptoms of heart disease are listing and named as the input variable for this HFS, as shown in Table 1. Additionally, Table 1 also includes the linguistic term for each symptom, namely *mild, moderate* and *severe*.

<table>
<thead>
<tr>
<th>No.</th>
<th>Symptoms (Input variables)</th>
<th>Linguistics Term</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>Shortness of breath</td>
<td>{Mild, Moderate, Severe}</td>
</tr>
<tr>
<td>S2</td>
<td>Discomfort, pressure, heaviness, or pain in the chest, arm, or below the breastbone</td>
<td>{Mild, Moderate, Severe}</td>
</tr>
<tr>
<td>S3</td>
<td>Fatigue</td>
<td>{Mild, Moderate, Severe}</td>
</tr>
<tr>
<td>S4</td>
<td>Nausea</td>
<td>{Mild, Moderate, Severe}</td>
</tr>
<tr>
<td>S5</td>
<td>Difficulties in walking or climbing stairs</td>
<td>{Mild, Moderate, Severe}</td>
</tr>
<tr>
<td>S6</td>
<td>Swelling in ankles or legs</td>
<td>{Mild, Moderate, Severe}</td>
</tr>
<tr>
<td>S7</td>
<td>Difficulty to sleep with one pillow</td>
<td>{Mild, Moderate, Severe}</td>
</tr>
<tr>
<td>S8</td>
<td>Rapid or irregular heartbeats</td>
<td>{Mild, Moderate, Severe}</td>
</tr>
<tr>
<td>S9</td>
<td>Fullness, indigestion, or choking feeling (may feel like heartburn)</td>
<td>{Mild, Moderate, Severe}</td>
</tr>
<tr>
<td>S10</td>
<td>Sweating</td>
<td>{Mild, Moderate, Severe}</td>
</tr>
<tr>
<td>S11</td>
<td>Take frequent breaks</td>
<td>{Mild, Moderate, Severe}</td>
</tr>
<tr>
<td>S12</td>
<td>Dizziness</td>
<td>{Mild, Moderate, Severe}</td>
</tr>
<tr>
<td>S13</td>
<td>Depressed</td>
<td>{Mild, Moderate, Severe}</td>
</tr>
</tbody>
</table>

Table 2 shows the output of the HFS that is the risk of having heart disease, whether *healthy, not healthy with moderate probability* and *not healthy with high probability*. These risks are then termed with *A, B*, and *C*, respectively, as shown in Table 2.

<table>
<thead>
<tr>
<th>No.</th>
<th>Output variable</th>
<th>Linguistics Term</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Risk of heart disease:</td>
<td>{A, B, C}</td>
</tr>
<tr>
<td></td>
<td>• Healthy - (A)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Not healthy (Moderate probability) - (B)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Not healthy (High probability) - (C)</td>
<td></td>
</tr>
</tbody>
</table>
### iii. Step 3 – Construct membership functions for the input and output variables

For simplicity, the number of linguistic terms for all input variables of HFS must not be higher than seven. Based on a study in cognitive psychology (Miller, 1956), the number of different entities efficiently stored in the short-term memory should not exceed the limit of $7 \pm 2$. Therefore, in this step, we assign three linguistic terms for all inputs and output variable in constructing their membership functions. Figure 4 shows the example of membership functions that built for input variable S4 - Nausea.

![Figure 4: Membership functions for the input of S4 - Nausea](image)

### iv. Step 4 – Form the hierarchical structure – topology

As mentioned earlier, HFSs are produced by decomposing the input variables in FLSs into multiple low-dimensional FLSs. By doing this, several layers are generated in HFSs. Based on the same input variables, HFSs may be produced using different topologies, e.g., serial and parallel. In this step, we adopt a serial topology to construct for the complex medical problem that is for early diagnosis of heart disease, as shown in Figure 5. As can be seen in Figure 5, this topology uses strictly one FLS per layer.

![Figure 5: Topology of HFS for Heart disease](image)
For instance, in Figure 5, symptoms S1 and S2 served as input for subsystem FLS_1 and produced an intermediate output $y_1$. Then, $y_1$ will act as input together with symptom S3 for subsystem FLS_2 in the next layer of HFS. This process will continue until the last subsystem, that is FLS_{12}.

v. **Step 5 – Compose the rule bases of the subsystems**

Following the results of the interview obtained earlier, in this step, the information is transformed into a meaningful table of the rule base for all subsystem. For example, Tables 3, 4 and 5 show the rule base created for subsystem FLS_1, FLS_2 and FLS_{12}, respectively.

### Table 3: Rule base of a subsystem – FLS_1

<table>
<thead>
<tr>
<th>S1</th>
<th>S2</th>
<th>Mild</th>
<th>Moderate</th>
<th>Severe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mild</td>
<td>mild</td>
<td>mild</td>
<td>mild</td>
<td>severe</td>
</tr>
<tr>
<td>Moderate</td>
<td></td>
<td>moderate</td>
<td>severe</td>
<td></td>
</tr>
<tr>
<td>Severe</td>
<td>moderate</td>
<td>severe</td>
<td>severe</td>
<td></td>
</tr>
</tbody>
</table>

The rule base of a subsystem FLS_1 can also be view as bellow:

- IF $S_1$ is *mild* AND $S_2$ is *mild* THEN $y_1$ is *mild*
- IF $S_1$ is *mild* AND $S_2$ is *moderate* THEN $y_1$ is *mild*
- IF $S_1$ is *mild* AND $S_2$ is *severe* THEN $y_1$ is *moderate*
- IF $S_1$ is *moderate* AND $S_2$ is *mild* THEN $y_1$ is *mild*
- IF $S_1$ is *moderate* AND $S_2$ is *moderate* THEN $y_1$ is *moderate*
- IF $S_1$ is *moderate* AND $S_2$ is *severe* THEN $y_1$ is *severe*
- IF $S_1$ is *severe* AND $S_2$ is *mild* THEN $y_1$ is *moderate*
- IF $S_1$ is *severe* AND $S_2$ is *moderate* THEN $y_1$ is *severe*
- IF $S_1$ is *severe* AND $S_2$ is *severe* THEN $y_1$ is *severe"

### Table 4: Rule base of a subsystem – FLS_2

<table>
<thead>
<tr>
<th>y_1</th>
<th>S3</th>
<th>Mild</th>
<th>Moderate</th>
<th>Severe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mild</td>
<td>mild</td>
<td>mild</td>
<td>mild</td>
<td>moderate</td>
</tr>
<tr>
<td>Moderate</td>
<td></td>
<td>moderate</td>
<td>severe</td>
<td></td>
</tr>
<tr>
<td>Severe</td>
<td>moderate</td>
<td>severe</td>
<td>severe</td>
<td></td>
</tr>
</tbody>
</table>

The rule base of a subsystem FLS_2 can also be view as bellow:

- IF $y_1$ is *mild* AND $S_3$ is *mild* THEN $y_2$ is *mild*
- IF $y_1$ is *mild* AND $S_3$ is *moderate* THEN $y_2$ is *mild*
- IF $y_1$ is *mild* AND $S_3$ is *severe* THEN $y_2$ is *moderate*
- IF $y_1$ is *moderate* AND $S_3$ is *mild* THEN $y_2$ is *mild*
- IF $y_1$ is *moderate* AND $S_3$ is *moderate* THEN $y_2$ is *moderate*
- IF $y_1$ is *moderate* AND $S_3$ is *severe* THEN $y_2$ is *severe*
- IF $y_1$ is *severe* AND $S_3$ is *mild* THEN $y_2$ is *moderate*
- IF $y_1$ is *severe* AND $S_3$ is *moderate* THEN $y_2$ is *severe*
- IF $y_1$ is *severe* AND $S_3$ is *severe* THEN $y_2$ is *severe*
Table 5: Rule base of a subsystem – FLS_{12}

<table>
<thead>
<tr>
<th>( y_{12} )</th>
<th>Mild</th>
<th>Moderate</th>
<th>Severe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mild</td>
<td>A</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>Moderate</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>Severe</td>
<td>B</td>
<td>C</td>
<td>C</td>
</tr>
</tbody>
</table>

The rule base of a subsystem FLS_{12} can also be viewed as follows:

- \( \text{IF } y_{12} \text{ is mild AND S13 is mild } \Rightarrow \text{Heart Disease is A} \)
- \( \text{IF } y_{12} \text{ is mild AND S13 is moderate } \Rightarrow \text{Heart Disease is A} \)
- \( \text{IF } y_{12} \text{ is mild AND S13 is severe } \Rightarrow \text{Heart Disease is B} \)
- \( \text{IF } y_{12} \text{ is moderate AND S13 is mild } \Rightarrow \text{Heart Disease is A} \)
- \( \text{IF } y_{12} \text{ is moderate AND S13 is moderate } \Rightarrow \text{Heart Disease is B} \)
- \( \text{IF } y_{12} \text{ is moderate AND S13 is severe } \Rightarrow \text{Heart Disease is C} \)
- \( \text{IF } y_{12} \text{ is severe AND S13 is mild } \Rightarrow \text{Heart Disease is B} \)
- \( \text{IF } y_{12} \text{ is severe AND S13 is moderate } \Rightarrow \text{Heart Disease is C} \)
- \( \text{IF } y_{12} \text{ is severe AND S13 is severe } \Rightarrow \text{Heart Disease is C} \)

vi. Step 6 – Link-up the subsystem in the hierarchical fuzzy system

As can be seen in Figure 5, the HFS produced several intermediate variables, namely \( y_1 \) to \( y_{11} \). In this step, these intermediate variables play an essential role in HFS structure as a connection between subsystems FLS_1 and FLS_2, FLS_3 and FLS_4, until FLS_{11} and FLS_{12}, as shown in Figure 5. For example, in Figure 5, subsystems FLS_1 and FLS_2 are connected using the intermediate variable \( y_1 \). Likewise, subsystems FLS_2 and FLS_3 are connected using the intermediate variable \( y_2 \).

DISCUSSION

The study in this paper was conducted as an initial approach to guide the process of designing HFS for a complex medical application. Specifically, it was focused on designing HFS for early diagnosis of heart disease and consisted of six main steps.

For the first step, we interviewed a medical expert to obtain useful information regarding the early diagnosis of heart disease. This is important to capture the information on the disease from an expert medical view. Then, the symptoms and output of heart disease were classified and described in detail in the second step. The finding has shown the 13 of symptoms were needed in order to identify an early diagnosis of heart disease.

For the third step, the information obtained in steps 1 and 3 were utilised to construct membership functions for all inputs and output variables. In this case, the trapezoidal membership functions with three linguistic terms were used to model all the inputs and output variables. Then, the selection of HFS topology was determined in the fourth step. For the early diagnosis of heart disease example, the serial topology was selected in order to represent the hierarchical structure that consists of 13 input variables, output variable, 12 subsystems and 12 layers.

For the fifth step, the rule base for all subsystems in HFS were constructed. Since all subsystems consist of 2 input and one input – multi inputs single-output (MISO), the rule base was constructed using a basic matrix table as shown in Tables 3, 4 and 5 based on the information obtained from the interview with medical experts (as in Step 1). For the last step, the intermediate variables were used to connect all the subsystems in
HFS, as shown in Figure 5. The intermediate variables not only serve as input to another subsystem in the next layer, but they also play an important role in order to link-up all the systems.

While the proposed approach seems promising to guide the process of designing HFS, however, this is just an initial approach. There is still more research on this topic that needs to be undertaken in future. The proposed design yet needs to be validated in term of showing how it works in practice. Therefore, in future work, we will focus on developing the HFS system using the proposed design in any programming language, e.g. MATLAB or R programming.

CONCLUSION

In conclusion, we have proposed a new approach to guide the process of designing an HFS, particularly for early diagnosis of heart disease, consisting of six key main steps. Although the current study only focuses on designing HFS for early diagnosis of heart disease, the proposed design is encouraging to ease the process of designing the HFS.

For the future work, we will focus on applying the proposed design in order to develop the HFS in practice, i.e. developing the HFS using a MATLAB or R programming.

ACKNOWLEDGEMENTS

We would like to thank Dr Rosmawati, Medical Doctor, Health Unit, UiTM Perlis Branch, who provided insight and expertise that greatly assisted the research.

REFERENCES


