Study of Fuzzy Cognitive Maps for Modeling Clinical Support Systems

1S. Sujamol, 2Sreeja Ashok and 3U. Krishna Kumar
1Department of Computer Science & I.T,
School of Arts & Sciences,
Amrita University,
Kochi.

2Department of Computer Science & I.T,
School of Arts & Sciences,
Amrita University,
Kochi.

3Department of Computer Science & I.T,
School of Arts & Sciences,
Amrita University,
Kochi.

Abstract

A biological network comprises of highly complex entities that are interrelated and depended. It is indeed a challenge to deal with these massively different data forms. Modeling the relationship between these data elements require efficient techniques. Recently, computational solutions find a promising approach in both medical and biological data analysis. Powerful methods with computational and mathematical features provide an efficient way of representing and analyzing biological data. Fuzzy Cognitive Maps, a combination of fuzzy logic and neural networks seems to be a promising approach for visualizing and manipulating data there by answering theoretical and experimental questions in biology. This work proposes the application of various inference rules in Fuzzy Cognitive Maps depending upon the nature of the problem and level of inference required.

Keywords: Fuzzy Cognitive Maps, Medical Decision Support, Inference rules.
1. Introduction

With the increase of research and development in biological and medical field, large amounts of data were produced day by day that needs to be organized and transformed into suitable scientific format. A disease is often characterized by a group of genes/proteins which will become active at different stages of the disease. The interaction of these malfunctioning genes forms a complex network with a sequence of genetic activities. Due to its complex and heterogenic nature, these data are not at all well served. These data often require efficient techniques for modeling, storage, analysis and interpretation so that useful information can be extracted. Various algorithms and computational methods have been developed for analyzing protein–protein interaction, cell and molecular level functioning, genetic information, disease prediction, metabolic pathway simulation etc. Similarly a single disease often has a number of causes. If it is possible to identify these disease causing factors, suitable pharmaceutical interventions and effective treatment strategies can be developed. Available methods that convert these heterogeneous data into biological knowledge include neural networks, fuzzy systems, machine learning algorithms and optimization methods such as evolutionary computing, swarm intelligence, immune computing and simulated annealing. Models based on these techniques were used for performing classification, clustering, feature selection and visualization tasks [25, 26]. Existing methods lack stability to integrate different types of biological data. Also all these methods doesn’t deal with imprecision and uncertainty which are indispensible in medical and biological data. Diagnosing the disorder is an important challenge and it requires a cost effective, exact and easy to use system. In order to deal with inaccuracies resulting from uncertainties, a strong firm work is necessary. The learning capacity of neural networks and generalization property of fuzzy logic can be combined to provide more reliable outputs. Fuzzy logic overcomes the difficulties in developing complex systems using mathematical models by enabling computational solutions to behave like human reasoning. Fuzziness is often experienced when dealing with naturally vague and complex systems controlled by human observations. The combination of neural networks and fuzzy logic provides a distributed representation of knowledge and has the ability to handle uncertain and imprecise data. The Fuzzy Cognitive Maps (FCM) are an effective soft computing technique that combine fuzzy and neural networks resulting in human-like reasoning by producing dynamic and parallel processing systems that estimate input output functions. The FCM consists of conventional fuzzy system components where computation at each stage is performed by n hidden layer neurons and the system knowledge is enhanced by the learning capacity of neural network.

Fuzzy Cognitive Maps

FCM originated from the concept of cognitive maps proposed by Robert Axelrod in 1976. Cognitive map provides a graphical representation for the problem domain with a set of concepts and relationship among them.
Relationships can be positive, negative and neutral. Positive relationships have a promoting influence between pair of concepts and are denoted by a positive sign whereas inhibiting influence between concepts pairs are represented by a negative sign. Independent concepts have no connection between them. But these cognitive maps only showed the type of relationship and failed to represent the extent of relationships and were not feasible for modeling complex systems. After ten years, Kosko in 1986 put forward the idea of Fuzzy Cognitive Maps. FCMs are graphical structures where each node stands for concepts and relationship between these concepts are represented by signed and weighted arcs [1]. Each concept depicts a state, variable or feature of a particular system and their values are usually normalized in the range \([0, 1]\). If a concept \(c_i\) causes an increase in other concept \(c_j\), then there is a positive relationship between \(c_i\) and \(c_j\). If concept \(c_i\) causes a decrease in \(c_j\), then a negative relation is observed. Finally, if there is no relationship between these concepts, the relationship is denoted as 0. Each connection has both sign and a numerical value, known as weight usually between \(-1\) and \(+1\) showing the extent of relationship between pair of concepts often described as weak, medium, strong or very strong etc. The entire knowledge extracted in the form of concepts and weighted relationship are visualized in the form of a graph structure with nodes and edges denoted as a paired vector \(<C, W>\) where \(C\) is a set of concepts and \(W\) is the weight matrix. FCMs are created either manually or by computational methods [2, 3]. In manual creation, the key concepts and their relations can be identified with the help of expert in each domain. FCMs can also be automatically generated without the involvement of experts by using available historic data [9]. Learning algorithms are used to frame out the weight matrix in order to improve the accuracy of the system. Fig 1 shows a simple FCM and the corresponding connection matrix. Thus, the structural anatomy of FCM comprises of fuzzy digraph that shows the intelligent behavior of a system by letting the relevant concepts to interact each other thereby producing the required output using an inference rule [4].

![FCM Diagram](image)

Let \(A_i\) denote value of concept \(c_i\). Its value is computed by the combined influence of rest of the concepts and weight matrix, \(W\) by using the following equation.

\[
A_i^{t+1} = f(\sum_{j=1}^{M} w_{ji} A_j^t) \quad i \neq j(1)
\]
This equation is known as Kosko inference rule. The rule is applied iteratively until the stopping criteria are satisfied. After a specific number of iteration, the FCM will be in a) Equilibrium point b) Limited cycle c) Chaotic behavior [21]. When the FCM reaches the fixed points, it is said to be converged. Function \( f(\cdot) \), in equation 1 denote the activation function and it can be binary, trivalent, sigmoid and hyperbolic tangent. Table 1 shows the commonly used activation function for FCM along with their equations and significance. The other inference rules applicable to FCM are Modified Kosko inference rule and Rescale inference rule.

### Table 1: Activation Rules Used in FCMs

<table>
<thead>
<tr>
<th>Function</th>
<th>Equation</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bivalent</td>
<td>( f(x) = \begin{cases} 0, &amp; x \leq 0 \ 1, &amp; x &gt; 0 \end{cases} )</td>
<td>Allows 0 and 1 values for each concepts there by representing whether a concept is activated or not</td>
</tr>
<tr>
<td>Trivalent</td>
<td>( f(x) = \begin{cases} 1, &amp; x &gt; 0 \ 0, &amp; x = 0 \ -1, &amp; x &lt; 0 \end{cases} )</td>
<td>Can have three activation levels, 1 means concept is increasing, -1 means concept is decreasing and 0 means concept is stable</td>
</tr>
<tr>
<td>Logistic Sigmoid</td>
<td>( f(x) = \frac{1}{1 + e^{-ax}} )</td>
<td>Calculated value of each concept belong to the range ([0,1])</td>
</tr>
<tr>
<td>Hyperbolic Tangent</td>
<td>( f(x) = \tanh(\lambda \times x) )</td>
<td>Concept values can have negative values and falls in the interval ([-1,1])</td>
</tr>
</tbody>
</table>

### 2. Related Work

FCM theory, based on both computational and manual methods has been widely used in decision support but very few works has been done in exploring its inference capabilities. Most of the work is done in finding the relationship between concept values in a particular domain. For manually developed FCMs, expert suggest the important concepts and they give the relationship between these concepts in the form of fuzzy rules and linguistic variables. For example

"IF the value of concept \( A_i \) is \( x \) then value of concept \( A_j \) is \( y \) and the linguistic weight is \( w_{ij} \)."

If more number of experts are present, their entire responses in fuzzy form are collected and added using SUM method and it is defuzzified to get corresponding numerical weight[2,3]. Weight values are created using manual, semi-automated and completed automated methods such as Hebbian based, Population based and Hybrid [19]. Chrysostomos D. Stylios et al. 2004 proposed the mathematical description of a manually created FCM using a modified inference rule where the previous values of concepts were also considered for finding their next value[8]. Wojciech Stach et al. 2005 put forward the difficulties of manual FCMs as they rely entirely on human knowledge which may produce
inaccurate results and proposed genetic algorithms for learning connection matrix[9] in which the available historic data is used for finding weight matrix. Later Elpiniki I. Papageorgiou 2005 suggested particle swarm optimization for formulating weight matrix[10]. Athanasios K. Tsadiras 2008 analysed inference capabilities of three types of FCMs that uses bivalent, trivalent and sigmoid transfer function respectively. Prediction capability depends on the nature and required capability for a problem. He concluded that binary FCMs only show an increase of concept or a stable concept and failed to show its decrease. Trivalent FCMs can represent increase, decrease and stability of a concept but does not show the degree. Sigmoid FCMs are more suitable as it can represent increase, decrease and stability of a concept [14]. Elpiniki I. Papageorgiou 2011 proposed an FCM for medical decision support based on Fuzzy Rule Extraction methods. The rescale inference rule was used here and the knowledge extraction methods in the form of fuzzy rules were applied to FCMs [17]. Gonzalo Napoles et.al 2017 addressed challenges and advances of FCM and applied FCM in creating models for pattern classification. The paper investigated FCM based classifiers and discussed about both supervised and unsupervised learning algorithms along with common inference rules used for FCM convergence [22].

**Proposed Approach**

Concepts and weight values of FCM represent specific variables of a problem. So when compared with Artificial Neural Networks, FCMs have superior representation capabilities. The decision making problem in predicting various diseases is a complex process due to the presence of numerous disease symptoms and signs. FCM approach provides an effective representation of cause-effect relationship between these medical data [5, 7]. Due to its feedback mechanism, it is regarded as a dynamic tool well suited in medical application for assisting doctors in disease prediction and diagnosis [12,13].

Inference procedure of FCM returns concept values of each state thereby modeling complex decision problems with numerous interrelated factors. In addition to the basic Kosko inference mechanism, modified Kosko inference rule was also used when the current situation requires the updating of activation value of concepts that are not influenced by other concepts. For avoiding difficulties caused by inactive concepts, rescale inference rule was proposed [17]. This rule was used when initial values of concepts are not known and it helps to avoid saturation problem. The table 2 below shows various inference rules in FCM along with the equations.

<table>
<thead>
<tr>
<th>Inference Rule</th>
<th>Equation</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kosko’s Inference Rule</td>
<td>$A_i^{t+1} = f \left( \sum_{j=1}^{N} w_{ij} A_j^{t} \right)$, $i \neq j$</td>
<td>B.Kosko,1986</td>
</tr>
<tr>
<td>Modified Kosko’s Inference Rule</td>
<td>$A_i^{t+1} = f \left( \sum_{j=1}^{N} w_{ij} A_j^{t} + A_i^{t} \right)$, $i \neq j$</td>
<td>Stylios and Groumpos,2004</td>
</tr>
<tr>
<td>Rescale Inference Rule</td>
<td>$A_i^{t+1} = f \left( \sum_{j=1}^{N} w_{ij} (2A_j^{t} - 1) + (2A_i^{t} - 1) \right)$, $i \neq j$</td>
<td>Elpiniki Papageorgiou 2009</td>
</tr>
</tbody>
</table>

Table 2: Inference Rules in FCMs
Other than these, clamped versions of these rules are also available. Here, till the output is stabilized, initially activated concepts are activated in all states. The current work investigates all the inference rules by applying them on medical data and simulating the result. The complex medical data was chosen as it contains more number of uncertain factors for decision making.

3. Results and Discussion

The behavior of various inference rules are analyzed here using autism data downloaded from UCI repository. Autism, a developmental disorder has both neurological and behavioral causes\cite{6,23,24}. Early clinical diagnosis can reduce the risk and help clinical practitioners to perform appropriate treatment strategy. The downloaded data set are created through a screening process on children between the age group 4-11 and have 293 cases with binary, nominal and categorical values. After performing the data preprocessing steps of cleaning and normalization, 16 concept values are taken. The 15 symptoms are factor concepts responsible for autism and the final concept is considered as the decision concept which indicate whether autism is present or not\cite{0,1}\cite{14}. The updating rule is repeated until the system converges to a fixed point because at this stage same output is produced at each iteration step. The fixed point attractor ensures that activation degrees of neurons remain unchanged\cite{20}. For the consistency of modeled system, fixed point attractors are preferred here. The overall system was implemented using R software with a randomly created weight matrix shown in Table 3. The activation vector contains the initial concept values and is represented as a 1 x m matrix as below.

\[
A_0 = [1,1,0,0,1,0,1,0,0,0,3,1,0,0,0,5,0]
\]

The input values and weight matrix shown in Table 3 is used in the inference rule to study the convergence state of the proposed model.

<table>
<thead>
<tr>
<th>Table 3: Randomly generated Weight Matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
</tr>
<tr>
<td>-----</td>
</tr>
<tr>
<td>0.32</td>
</tr>
<tr>
<td>0.18</td>
</tr>
<tr>
<td>0.26</td>
</tr>
<tr>
<td>-0.28</td>
</tr>
<tr>
<td>0.49</td>
</tr>
<tr>
<td>0.09</td>
</tr>
<tr>
<td>0.08</td>
</tr>
<tr>
<td>0.82</td>
</tr>
<tr>
<td>0.51</td>
</tr>
<tr>
<td>0.22</td>
</tr>
<tr>
<td>0.07</td>
</tr>
<tr>
<td>0.82</td>
</tr>
<tr>
<td>-0.54</td>
</tr>
</tbody>
</table>
When Kosko inference rule is applied with sigmoid function, the output as shown in fig 2 was obtained. The entire system converged at 15th stage.

![Figure 2](image1.png)

Fig 2. After applying kosko rule the system converged at 15th state

When modified Kosko rule was applied, it gives the following result as shown in fig 3 and the system converged at 16th state.

![Figure 3](image2.png)

Fig 3: Simulation after applying Modified Kosko rule

Both Kosko rule and Modified Kosko rule can be enhanced by applying the clamping function where key concepts are activated at each step. Kosko-clamped rule when applied on autistic patient data gives following result. Here the outputs are converged at 11th state.
Clamped version of modified Kosko was also applied where the states converged at 10th state and it is found to be more suitable for the current situation.

The system showed no convergence when rescale and rescaled clamped rule was applied.
The table 4 below shows an overall summary of convergence state when various inference rule was applied along with values of each concepts during the convergence

<table>
<thead>
<tr>
<th>Rule</th>
<th>Converged state</th>
<th>Concept I</th>
<th>Concept II</th>
<th>Concept III</th>
<th>Concept IV</th>
<th>Concept V</th>
<th>Concept VI</th>
<th>Concept VII</th>
<th>Concept VIII</th>
<th>Concept IX</th>
<th>Concept X</th>
<th>Concept XI</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Kosko</td>
<td>0.10</td>
<td>0.30</td>
<td>0.25</td>
<td>0.35</td>
<td>0.50</td>
<td>0.75</td>
<td>0.85</td>
<td>0.05</td>
<td>0.04</td>
<td>0.02</td>
<td>0.04</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Modified</td>
<td>0.70</td>
<td>0.50</td>
<td>0.30</td>
<td>0.50</td>
<td>0.60</td>
<td>0.70</td>
<td>0.80</td>
<td>1.00</td>
<td>0.10</td>
<td>0.09</td>
<td>0.08</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>Kosko</td>
<td>0.11</td>
<td>1.00</td>
<td>0.50</td>
<td>0.20</td>
<td>0.30</td>
<td>0.50</td>
<td>0.60</td>
<td>0.70</td>
<td>0.80</td>
<td>0.90</td>
<td>1.00</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Modified</td>
<td>0.50</td>
<td>0.50</td>
<td>0.30</td>
<td>0.50</td>
<td>0.60</td>
<td>0.70</td>
<td>0.80</td>
<td>0.90</td>
<td>1.00</td>
<td>0.10</td>
<td>0.09</td>
<td>0.08</td>
<td></td>
</tr>
</tbody>
</table>

4. Conclusion

Fuzzy Cognitive Maps are widely used for modeling complex real life problems. The entire process of decision making can be done with a graphical structure that can model existing uncertainties. The main objective of this work was to bring out various inference rules used in FCMs. Fully automated FCM was studied without any expert intervention. The capabilities of various inference rules were analysed. Present study used a randomly generated weight matrix for inference process. As a future work, convergence capabilities can be analysed after finding the weight matrix using various learning algorithms like hebbian learning, population based learning and hybrid algorithms.

References


