This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2019.2961350, IEEE Transactions on Fuzzy Systems

A New Design of Mamdani Complex Fuzzy Inference System for Multi-attribute Decision Making Problems

Ganeshsree Selvachandran, Shio Gai Quek, Luong Thi Hong Lan, Le Hoang Son, Nguyen Long Giang, Weiping Ding, Senior Member, IEEE, Mohamed Abdel-Basset, and Victor Hugo C. de Albuquerque, Senior Member, IEEE

Abstract**—**This paper proposes the Mamdani complex fuzzy inference system (Mamdani CFIS) to improve performance of the classical FIS and complex FIS. The applicability of the proposed CFIS is demonstrated by applying it to six commonly available datasets from UCI Machine Learning under the comparison with Mamdani FIS and the Adaptive Neuro Complex Fuzzy Inference System (ANCFIS). It is successfully proven that the proposed Mamdani CFIS is computationally less expensive, and presents a more efficient method to handle time-series data and time-periodic phenomena, among all the fuzzy Inference System found thus far in the literature. Furthermore, the novelty of CFIS mainly lies in its implementation of the complex number throughout the entire procedures of computation, this gives much greater flexibility of implementing unexpected, non-linear fluctuations.

Index Terms**—Complex fuzzy inference system; complex fuzzy logic; Mamdani fuzzy inference system; decision making.**

I. Introduction

FUZZY logic is an offshoot of the theory of fuzzy sets that imitates human thinking and reasoning to increase efficiency of the decision making process when handling uncertain or vague data [1]. A *fuzzy inference system* (FIS) can be defined as a nonlinear mapping that derives its output based on fuzzy reasoning and a set of fuzzy IF-THEN rules [1]. The fast-paced development of fuzzy set has led to the development of FISs including the most commonly used FISs namely the Mamdani, Sugeno and Tsukamoto systems. Since its inception, FISs have been successfully applied in many real-life situations. These include the construction of a subway system in Japan, the design of low-cost microcontrollers based on fuzzy logic operations by Intel, automated space docking by NASA, and the design and manufacturing of various electrical appliances such as washing machines, air-conditioners, and television.

Despite the wide use of FIS, most if not all of them are examples of Mamdani FIS, Sugeno FIS and Tsukamoto FIS, all of which implements only real number in their computation. Whereas on the other hand, many of the results in reality can have unexpected fluctuations, such as the shrinkage of volume when water mixed with alcohol, for which there are no known formula to quantitatively calculate the shrinkage in volume. Moreover, there are also many scenarios that involves a phase term, which is encountered in data with a periodic trend, such as rainfall recorded in a region, or the sound waves produced by a musical instrument. It is therefore evident, that complex numbers must be given a place in the literature of fuzzy inference system as well. This is therefore the main motive of this paper.

The study of *complex fuzzy sets* (CFSs) began with Ramot's pioneer paper on CFSs [2] and complex fuzzy logic (CFL) [3]. In [2, 3], the authors asserted that many real-life phenomena are periodic or display recurring behavior. Examples of these include solar activity, the effects of financial indicators on one another, stock trading on the NYSE, and signal processing. Dick [4] concurred with Ramot's reasoning in [2, 3] and smart city [4]. Specifically, traffic jams in the city never exactly repeat themselves, but there is an obvious pattern [4]. The jam is very heavy in the morning and evening when people are going to work and going back home from work, and lighter at other times of the day. To handle such phenomena, Ramot et al. [2] added another dimension called the phase term to capture the information related to time in time-periodic phenomena, which motivates the conceptualization of the CFS model.

In the CFS, codomain of membership function belongs to the unit disc of the complex plane $z \in \mathbb{C}$, with $|z| \leq 1$. The membership function of CFSs consists of two components: an amplitude term and a phase term. The amplitude term plays a similar role to the membership grade in ordinary fuzzy sets with its values describing the degree of belongingness of an element to an attribute, whereas the phase term captures information pertaining to the periodicity or seasonality of time-series data in time-periodic phenomena. This combination of the amplitude and phase terms enables a precise representation of recurring behaviors in phenomena with approximately periodic or seasonal behavior, compactly in a single set. Although the study of CFSs is still at its infancy, it has been steadily gaining

G. Selvachandran is with the Department of Actuarial Science and Applied Statistics, Faculty of Business & Information Science, UCSI University, Jalan Menara Gading, 56000 Cheras, Kuala Lumpur, MALAYSIA (Email: ganeshsree86@yahoo.com).

S. G. Quek is with the Department of Actuarial Science and Applied Statistics, Faculty of Business & Information Science, UCSI University, Jalan Menara Gading, 56000 Cheras, Kuala Lumpur, Malaysia (e-mail: queksg@ucsiuniversity.edu.my).

L. H. Lan is with the Faculty of Computer Science and Engineering, Thuyloi University, Hanoi, Vietnam; and Institute of Information Technology, Vietnam Academy of Science and Technology, Hanoi, Vietnam (e-mail: lanlhbk@gmail.com).

L. H. Son is with the VNU Information Technology Institute, Vietnam National University, Hanoi, Vietnam (e-mail: sonlh@vnu.edu.vn).

N. L. Giang is with the Institute of Information Technology (IoIT), Vietnam Academy of Science and Technology (VAST), Vietnam (e-mail: nlgiang@ioit.ac.vn).

W. Ding* is with School of Information Science and Technology, Nantong University, Nantong 226019, China. (E-mail: dwp9988@163.com).

M. Abdel-Basset* is with Zagazig university, 44519, Zagazig, Egypt (e-mail: analyst_mohamed@zu.edu.eg).

V. H. C. de Albuquerque is with the University of Fortaleza (UNIFOR) (email: victor.albuquerque@unifor.br).

*: Corresponding authors: **W. Ding** and **M. Abdel-Basset**

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2019.2961350, IEEE Transactions on Fuzzy Systems

attention [2-7]. Tamir and his collaborators have been actively advancing the study of complex fuzzy logic via their works in [8-11] with a review for developments and shortcomings in this area [11].

As a **summary of the related works regarding CFISs** which will be presented in Section 2, all of these early stages of CFIS model ignore the phase terms in the decision-making process; thus, the valued inputs are still mostly real numbers. Therefore, these systems are unable to capture the fluctuations in time (information related to phase terms) or the interaction between substances (ranging from total accentuation to total cancellation of each other's effects) which effectively reduces these systems to ordinary Adaptive-Network Based Fuzzy Inference System (ANFIS).

In the other words, most of the so-called CFISs that have been proposed in the literature are not truly complex systems. Although these systems use CFSs as their inputs but they only used the amplitude terms in the decision-making process, while ignoring the phase terms. For instance, Ramot's CFLS ignores the phase term of the output during the process of defuzzification; thereby making it insufficient to handle time-series data of recurring or periodic phenomena, and effectively reduces the CFIS to ordinary FIS. The ANCFIS models by Man [12] and Chen [13] used vector dot-product for the aggregation stage and treated the complex-valued inputs as real values, thereby enabling them to obtain scalar values for the dot product. This would not be possible if the inputs are indeed treated as complex values, as the dot product of two complex numbers is a complex number and not a scalar value. The ANCFIS system is therefore not truly complex as the outputs of the system will not be representative of the periodicity of the elements.

The ANCFIS model were then further improvised by Yazdanbakhsh and Dick [32], into Randomized Adaptive-Network Based Fuzzy Inference System (RANCFIS) and Fast Adaptive-Network Based Fuzzy Inference System (FANCFIS). However, it was observed that such generalizations do nothing to improve the actual role of complex numbers in pre-existing ANCFIS.

The ordinary fuzzy inference systems such as the Mamdani, Sugeno and Tsukamoto systems and various versions of the ANFIS architectures are only able to handle phenomena that are not periodic or seasonal. In order to handle time-series data in time-periodic phenomena, FISs and ANFISs employ two general strategies: 1) ignore the information related to the phase term; 2) represent the amplitude and phase terms separately using two fuzzy sets. This would cause loss of information and produce unreliable results (if information related to the phase terms are ignored), distortion of information, and a reduction in computational efficiency (if information related to the amplitude and phase are represented separately) as it becomes more time-consuming due to the increased number of sets that need to be dealt with.

For instance, when dealing with the mixing of concrete, it is well understood that different chemical may interact unexpectedly in a non-linear way. Such interaction of chemicals can range from total accentuation of each other effect, to total cancellation of other effect. In ordinary fuzzy inference systems, such implementation proves extremely difficult and time consuming as real numbers only possess magnitude. Thus, only in our CFIS model that we are capable of assigning a phase term to each individual parameter in the entire process, allowing the interference of chemicals to take place, for which identical phase result in accentuation of effects, whereas opposite phase result in cancellation of effects.

The deficiencies present in the existing systems that were explained above served as the main **motivation** that led us to introduce a truly complex FIS in this paper. Therefore, this paper proposes the Mamdani complex fuzzy inference system (Mamdani CFIS) to improve performance of the classical and complex FISs. Specifically, the **contributions** and **novelty** of are demonstrated as follows:

(i) We develop a robust and fully-functioning framework of a Mamdani based complex fuzzy inference system (CFIS) that is truly complex, i.e. a system takes into consideration the information related to the phase term of the membership function throughout the entire decision-making process, which was ignored in previous studies. Each and every parameter involved is a genuinely complex number, as shown in Fig. 1. This is in contrast with the other CFIS model which incorporates complex numbers to a small fraction of the parameters only, as in ANCFIS by Chen [13].

(ii) The proposed Mamdani CFIS is applied to six commonly available benchmark datasets from UCI, namely Concrete Slump, Breast Cancer Wisconsin, Image Segmentation, Energy Efficiency, Chronic Kidney Disease and Wine Quality datasets. The proposed Mamdani CFIS is run side by side with the ordinary Mamdani FIS and the ANCFIS model by Chen [13] to have their performances compared.

(iii) It is subsequently proven that the results obtained via our Mamdani CFIS are more accurate and reliable compared to the results obtained via ordinary Mamdani FIS and the ANCFIS model by Chen [13], in all of the six experiments aforementioned.

(iv) It is then proven further that our proposed Mamdani CFIS is able to produce results with a higher level of reliability and accuracy, and is able to do so in lesser computation time and a lower level of computational complexity.

In what follows, we will present the related researches with background concepts followed by the new methodology and experimental results in the sub-sequent sections.

II. Related works

A. *Fuzzy Inference Systems on Type-1 Fuzzy Set*

Here, we present a brief review of an advanced FIS called ANFIS. Jang [14] introduced the Adaptive-Neuro Fuzzy Inference System (ANFIS), which is a hybrid of artificial neural network based on the Sugeno FIS. Since the pioneering work by Jang in [14] and [15], ANFIS has garnered a lot of attention from the fuzzy research community, and great progress has been made in the research related to ANFIS. FIS and ANFIS frameworks were applied to a myriad of areas including mining, geotechnical engineering, modelling traffic accidents, forecasting dam inflows, reliability estimation of software systems, breast cancer risk detection and recurrence risk, forecasting voltage losses in networks, and early diagnosis of dengue fever.

Several researchers have also developed **Mamdani based FIS** and ANFIS architectures and proceeded to apply these to various real-life problems. Camastra et al. [16] applied the Mamdani FIS for plant environmental risk assessment. Gayathri and Sumathi [17] applied their Mamdani FIS breast cancer risk detection, whereas Erturk and Sezer [18] proposed several versions of the Mamdani FIS model and applied these to software fault prediction problems with good results. Mamoria and Raj [19] proposed a Mamdani based FIS with multiple membership functions, applied it in a contrast enhancement problem and evaluated the results obtained using different membership functions, Mohanraj et al. [20] constructed a fuzzy logic controller for air-conditioners using a Mamdani and Sugeno based FIS, whereas Ruzic, Skenderovic and Lesic [21] applied the Mamdani FIS in HRM performance evaluation problem and tested this model to measure the HRM performance in middle-sized hotel companies. Thakur et al. [22] used the Mamdani FIS model to determine the severity of Thalassemia disease in patients, and Ahamed et al. [23] applied the Mamdani FIS model in the process of recognizing and identifying running conditions using data collected with a triaxial accelerometer.

Mamdani based ANFIS models and their applications were introduced in [24, 25]. Chai, Jia and Zhang [25] proposed a Mamdani based ANFIS framework and applied this in computing traffic Level-of-Service evaluation. The authors proved that this Mamdani based ANFIS is superior to ANFIS with a lesser computation time and testing errors. Borkar et al. [24] also applied the Mamdani ANFIS model in developing a performance monitoring system for shell and tube heat exchanger, and used experimental results to also prove that the Mamdani ANFIS is computationally less expensive and more efficient compared to the ordinary ANFIS model.

B. Fuzzy Inference Systems on Complex Fuzzy Set

Early attempts at creating fuzzy inference systems based on complex fuzzy sets (CFSs) were due to Li and Jang [26] and Malekzadeh and Akbarzadeh [27]. Li and Jang [26] proposed a FIS based on CFSs called the Complex Adaptive Fuzzy Inference System (CAFIS), which is a straightforward extension of ANFIS, but one which accepts complex-valued inputs and outputs. This system is however not truly complex in nature, as the real and imaginary parts of the input membership functions are dealt with separately using two type-1 fuzzy sets. Separating the real and imaginary parts also leads to an increased number of rules which makes this system computationally expensive.

Malekzadeh and Akbarzadeh [27] proposed another inference system based on complex fuzzy sets, called the Complex-valued Adaptive Neuro-fuzzy Inference System (CANFIS), which is a hybrid of CFIS and fuzzy neural networks. However, the authors did not present any method to deal with the defuzzification of the complex-valued outputs to crisp values, and chose to only consider the real part of the output. Deshmukh et al. [28] introduced a complex fuzzy logic module and applied this to the design process of a fuzzy microprocessor using the VLSI approach. Similar to the systems introduced in [26, 27], the authors did not implement rule interference and did not provide a valid defuzzification module.

In [12], Man, Chen and Dick defined an inductive learning algorithm for CFL called ANCFIS, which is a hybrid of CFIS and ANFIS. This led to the improvement of many other ANCFIS systems and other hybrid architecture which were built using this ANCFIS system. Although the ANCFIS model was introduced in 2007, research related to CFIS, CFL and ANCFIS are still at the beginning stages.

III. Preliminaries

Herein, we discuss few concepts leading to the formation of a *truly complex model of a complex fuzzy inference system* (CFIS).

A. Mamdani Fuzzy Inference System

The general steps involved in a FIS are as below [1]:

(i) *Fuzzification*: the process of fuzzification of the input variables with the membership functions for each linguistic label.

(ii) *Aggregation*: obtain the firing strength of each rule from the membership values.

(iii) *Consequence*: use the rule firing strength to generate values of the consequent for each fuzzy rule.

(iv) *Defuzzification*: process of converting the fuzzy outputs obtained in the previous step into crisp values.

B. Complex Fuzzy Set

Definition 3.1 [2]**.** A *complex fuzzy set* A on U is characterized by $\mu_A(x)$, that assigns to any $x \in U$, a complex-valued grade of membership in A which could lie within the unit circle in the complex plane: $\mu_A(x) = r_A(x)e^{i\,\omega_A(x)}$, where $i = \sqrt{-1}$, each of $r_A(x)$ and $\omega_A(x)$ are both real-valued with $r_A(x) \in [0,1]$ and $\omega_A(x) \in (0, 2\pi]$.

$$A = \{(x, \mu_A(x)) : x \in U\} = \{(x, r_A(x)e^{i\,\omega_A(x)}) : x \in U\} \quad (1)$$

Definition 3.2 [2]**.** Let $\mu_A(x) = r_A(x)e^{i\,\omega_A(x)}$ and $\mu_B(x) = r_B(x)e^{i\,\omega_B(x)}$ be the membership functions of A and B, respectively.

(i) The *complement* of A is

$$\bar{A} = \{(x, \mu_{\bar{A}}(x)) : x \in U\} = \{(x, r_{\bar{A}}(x)e^{i\,\omega_{\bar{A}}(x)}) : x \in U\} \quad (2)$$

where $r_{\bar{A}}(x) = 1 - r_A(x)$ and $\omega_{\bar{A}}(x) = 2\pi - \omega_A(x)$.

(ii)The *union* of A and B is

$$A \cup B = \{(x, \mu_{A\cup B}(x)) : x \in U\} = \{(x, r_{A\cup B}(x)e^{i\,\omega_{A\cup B}(x)}) : x \in U\}, \quad (3)$$

where $r_{A\cup B}(x) = \max\{r_A(x), r_B(x)\}$ and $\omega_{A\cup B}(x) = \max(\omega_A(x), \omega_B(x))$.

(iii) The *intersection* of A and B is

$$A \cap B = \{(x, \mu_{A\cap B}(x)) : x \in U\} = \{(x, r_{A\cap B}(x)e^{i\,\omega_{A\cap B}(x)}) : x \in U\}, \quad (4)$$

where $r_{A\cap B}(x) = \min\{r_A(x), r_B(x)\}$ and $\omega_{A\cap B}(x) = \min(\omega_A(x), \omega_B(x))$.

C. Ramot's complex fuzzy inference system

Definition 3.3 [3]. A *complex fuzzy logic system* (CFLS) consists of a complex fuzzy rule base in the form of IF-THEN statements and has three components, namely fuzzification, fuzzy inference and finally defuzzification. The three stages of the CFLS are summarized below:

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2019.2961350, IEEE Transactions on Fuzzy Systems

(i) *The fuzzification module:* the process of converting the crisp inputs into complex fuzzy inputs. The choice of complex fuzzy inputs in [3] is the classical form of complex fuzzy sets of the form $\mu_A(x) = r_A(x).e^{\omega_A(x)}$.

(ii) *The fuzzy inference stage*: the process of using complex fuzzy rules to map the complex fuzzy inputs into complex fuzzy outputs through the complex fuzzy implication. Outputs of separate rules are combined to produce a single complex fuzzy output set using any vector aggregation operation.

(iii) *The defuzzification process*: involves defuzzification of the complex fuzzy output set obtained in the second stage to produce a crisp output. In [3], Ramot et al. did not outline any specific method of defuzzification to reduce the complex fuzzy outputs into crisp outputs. Their chosen approach of defuzzification is to consider only the amplitude terms of the complex fuzzy output set. Hence the authors recommended the use of any suitable defuzzification methods used in traditional FLSs for this purpose. A thorough study of the rather limited literature in this area showed there are still no defuzzification methods in literature that is able to convert complex fuzzy sets into crisp outputs by taking the phase information into consideration.

Definition 3.4 [3]. *Complex fuzzy logic* employs rules in the form of IF-THEN statements constructed with complex fuzzy sets to create a CFLS.

Definition 3.5 [3]. A *complex fuzzy rule* represents a *complex fuzzy implication* relation between complex fuzzy propositions p and q, where p ~ "X is A", and q ~ "Y is B", respectively. A complex fuzzy implication is then defined as

$$\mu_{A\to B}(x,y) = r_{A\to B}(x,y).e^{\omega_{A\to B}(x,y)}, \quad (5)$$

where the $r_{A\to B}(x,y)$ is the amplitude term and phase term of the complex fuzzy implication. The term $r_{A\to B}(x,y)$ is equivalent to the real-valued grade of membership while $\omega_{A\to B}(x,y)$ indicates the phase associated with the implication and its significance becomes prominent when it is considered together with the amplitude term or when several complex fuzzy implications are considered together, as occurs in most CFLSs.

IV. THE MAMDANI COMPLEX FUZZY INFERENCE SYSTEM

The traditional Mamdani FIS is adapted to a complex fuzzy setting using operations and functions that specifically cater to periodic data and CFSs. The general structure of our proposed CFIS is similar to the CFLS introduced in Ramot et al. [3], but with significant differences in the choice of aggregation operator, and defuzzification methods. Fig. 1 shows the general framework of the proposed Mamdani CFIS.

Before we outline the different stages involved in our CFIS, we discuss our choice of membership function, methods of determining the rule firing strength and aggregation operator in the subsequent subsections.

A. Complex fuzzy membership function

So far, in the study of CFLSs and CFISs, the sinusoidal membership function and Gaussian membership function have been used. Sinusoidal membership function over the unit disc codomain introduced by Chen [13] is given in Eq. (6). Most of the other work related to ANCFIS that were introduced subsequently also used this class of complex fuzzy membership function.

$$\mu(x) = r(\theta(x)).e^{i\,\theta(x)},$$
$$r(\theta(x)) = d\sin\big(a(\theta(x)) + b\big) + c \quad (6)$$

such that $r(\theta)$ is the amplitude and $\theta(x)$ is the phase of the membership grade of $x \in U$.

On the other hand, we intend for our proposed Mamdani CFIS to be a faithful generalization of the pre-existing Mamdani FIS. In particular, it must be able to handle any $\mu(x)$ in exactly the same way as the Mamdani FIS whenever $\mu(x) \in [0,1]$. We are thus motivated to define our complex fuzzy membership function, for our Mamdani CFIS as:

$$\mu(x) = r(x).e^{i\,\theta(x)},$$
$$\theta(x) \in (0,2\pi] \text{ and } r(x) \in [0,1] \quad (7)$$

This is the classic complex membership function defined by Ramot et al. [2], with $r(x)$ and $\theta(x)$ representing the amplitude and phase terms of the elements, respectively. Moreover, since this is an initial attempt to develop a Mamdani based CFIS, we intend to focus on the implementation process of the system, and thus chose to use this classic complex membership function to keep things simple.

B. Operations used in the Mamdani CFIS

In this research, the operations that will be used in our Mamdani CFIS are given below:

(i) **The minimum T-norm** is used for calculating the firing strength of a complex fuzzy rule with AND connecting the antecedents.

(ii) **The maximum T-conorm** is used for calculating the firing strength of a complex fuzzy rule with OR connecting the antecedents.

(iii) **The Mamdani implication rule** for complex fuzzy sets given in Eq. (8) is used to calculate the values of the consequent of each complex fuzzy rule.

$$\mu_{A\to B}(x,y) = \big(r_A(x).r_B(y)\big).e^{i\,2\pi\left(\frac{\omega_A(x)}{2\pi}.\frac{\omega_B(y)}{2\pi}\right)} \quad (8)$$

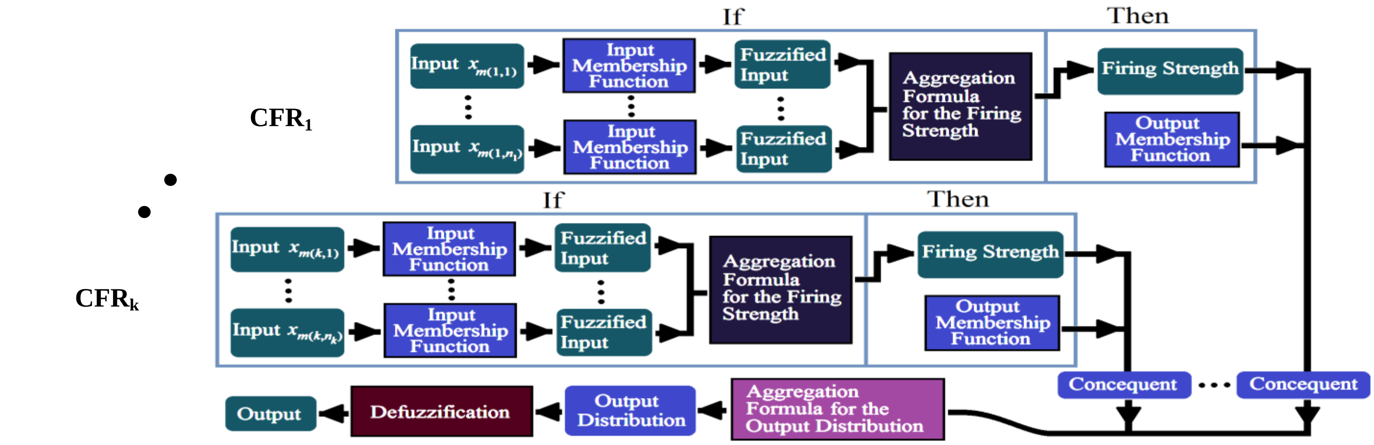
The standard definition of dot product for complex-valued vectors will be used to calculate the values of the consequent of each complex fuzzy rule. Note that dot product for complex vectors is generalization of the dot product of real vectors. Therefore, it is also a generalization of the product T-norm that is used in the context of ordinary Mamdani FISs.

Remark: As this is a CFIS, the firing strengths and the consequents of the complex fuzzy rules are also complex values.

C. Vector aggregation for CFSs

Ramot et al. [3] introduced a novel aggregation function called vector aggregation to allow multiple complex fuzzy rules to be combined together by taking into consideration their phase terms. The output is based on the principle of rule interference: If all of the arguments of the phase terms are aligned, the resulting amplitude is maximized; if the arguments of the phase terms are not aligned, the resulting amplitude may be lesser than that of the original input.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2019.2961350, IEEE Transactions on Fuzzy Systems



an element of $\mathbb{C}$; a function of the form $f: \mathbb{C} \longrightarrow \mathbb{C}$; a function of the form $g: \mathbb{C}^d \longrightarrow \mathbb{C}$ for some $d \in \mathbb{N}$;

a function of the form $\Psi: (\mathcal{F}(\mathbb{C},\mathbb{C}))^d \longrightarrow \mathcal{F}(\mathbb{C},\mathbb{C})$ for some $d \in \mathbb{N}$; a function of the form $\Phi: \mathcal{F}(\mathbb{C},\mathbb{C}) \longrightarrow \mathbb{C}$

Fig. 1. The framework of our CFIS (CFR: Complex Fuzzy Rule)

In the ANCFIS structure proposed by Man et al. [12], vector aggregation is accomplished via the dot product operation ".", with $w_p.\mu_{A_p}(x)$ denoting the dot product of the rule firing strength w_i with the complex fuzzy membership function of each complex fuzzy set $A_i \in U$. However, "." is a real-valued dot product and differs significantly from the conventional definition of dot product between complex-valued vectors which is as given in Eq. (9) below.

$$w_p.\mu_{A_p}(x) = w_p\,\overline{\mu_{A_p}(x)} = \left(r'_p e^{i\,\omega'_p}\right)\left(r_{A_p}(x)e^{-i\omega_{A_p}(x)}\right)$$
$$= r'_p r_{A_p}(x) e^{i\left(\omega'_p - \omega_{A_p}(x)\right)}, \qquad (9)$$

which in turn equals $r' r_{A_p}(x)\left(\cos\left(\omega'_p - \omega_{A_p}(x)\right) + \right.$ $i\ \sin \omega' p - \omega A p x$ and therefore not necessarily real. In the case of the proposed Mamdani CFIS, we aim to develop a CFIS that is truly complex i.e. one that implements rule interference and at the same time takes both the amplitude and phase terms into consideration throughout the decision-making process. As such, we opted to remain faithful to the conventional definition of the dot product between complex vectors, and therefore choose to use Eq. (9) as our aggregation operator.

D. Aggregation of the output distribution

To ensure that the Mamdani CFIS is a faithful generalization of the Mamdani FIS, it needs to be able to handle complex numbers exactly in the way that the Mamdani FIS handles real numbers. Furthermore, in the case where all of the membership functions are real, CFIS must behave identical to an FIS. To achieve this, we have designed the system in such a way that the weighting would be done during the implication step (step 4), similar to the pre-existing Mamdani FIS. Therefore, we do not consider w_p during the aggregation of the output distribution (step 5). In this regard, we define the output distribution $D(y)$ as follows

$$D(y) = \Gamma_1(y) + \Gamma_2(y) + \dots + \Gamma_k(y) \qquad (10)$$

In this definition, $D(y) = \Gamma_1(y) + \Gamma_2(y) + \dots + \Gamma_k(y)$ if $\Gamma_p(y)$ are complex functions. This way, we can be sure of obtaining a truly complex CFIS in which the information pertaining to the phase are not disregarded but taken into consideration in every step of the decision-making process.

E. Structure of the Mamdani CFIS

As a generalization to Mamdani FIS, the proposed Mamdani CFIS consists of six stages which must be completed before an output is obtained. Each of these individual stages are as given below.

Let $x_1, x_2, \dots, x_n \in \mathbb{C}$ be the inputs.

Stage 1: Determine a set of complex fuzzy rules.

Establish a set of complex fuzzy rules of the form:

CFR$_1$: If $x_{m(1,1)}$ is $A_{1,1}$ $\mathbf{O}_{1,1}$ $x_{m(1,2)}$ is $A_{1,2}$ $\mathbf{O}_{1,2}$ … $\mathbf{O}_{1,n_1-1}$ $x_{m(1,n_1)}$ is A_{1,n_1} , then y is C_1

CFR$_2$: If $x_{m(2,1)}$ is $A_{2,1}$ $\mathbf{O}_{2,1}$ $x_{m(2,2)}$ is $A_{2,2}$ $\mathbf{O}_{2,2}$ … $\mathbf{O}_{2,n_2-1}$ $x_{m(2,n_2)}$ is A_{2,n_2} , then y is C_2

… … … …

… …

CFR$_k$: If $x_{m(k,1)}$ is $A_{k,1}$ $\mathbf{O}_{k,1}$ $x_{m(k,2)}$ is $A_{k,2}$ $\mathbf{O}_{k,2}$ … $\mathbf{O}_{k,n_k-1}$ $x_{m(k,n_k)}$ is A_{k,n_k} , then y is C_k

in which for all p, q, the following holds:

(a) $m(p,q) \in \{1,2, \dots, n\}$, with $1 \le m(p,1) < m(p,2) < \dots < m(p,n_p) \le n$

Remark: In other words, we do not need all the *n* inputs of a CFIS to appear in one CFR of a CFIS. For each p, n_p signifies the number of inputs in CFR$_p$, the *p*-th rule of a CFIS. The input that is involved can be any n_p inputs out of $x_1, x_2, \dots, x_n$, and need not be all of $x_1, x_2, \dots, x_{n_p}$. It is for this reason we write $x_{m(p,1)}, x_{m(p,2)}, \dots, x_{m(p,n_p)}$ to denote the way of picking n_p inputs out of all the *n* inputs that can be chosen freely and distinctively among all

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2019.2961350, IEEE Transactions on Fuzzy Systems

CFRs.

(b) $\mu_{A_{p,q}}(x_{m(p,q)}) = r_{A_{p,q}}(x_{m(p,q)})\ \mathrm{e}^{i\,\omega_{A_{p,q}}(x_{m(p,q)})}$, for which $r_{A_{p,q}}: \mathbb{C} \to [0,1]$ and $\omega_{A_{p,q}}: \mathbb{C} \to (0,2\pi]$.

(c) $\mu_{C_p}(y) = r_{C_p}(y)\ \mathrm{e}^{i\,\omega_{C_p}(y)}$, for which $r_{C_p}: \mathbb{C} \to [0,1]$ and $\omega_{C_p}: \mathbb{C} \to (0,2\pi]$.

(d) T_0 is a T-norm, and S_0 is the S-norm (i.e. the T-conorm) that corresponds to T_0.

(e) $f_p\ : (0,2\pi]^{n_p} \to (0,2\pi]$, with $f_p(2\pi, 2\pi, \dots, 2\pi) = 2\pi$

(f) $w_p = \tau_p\ \mathrm{e}^{i\,\psi_p}$, where

$$\tau_p = N_{p,n_p-1}\left(\dots N_{p,2}\left(N_{p,1}\left(r_{A_{p,1}}(x_{m(p,1)}), r_{A_{p,2}}(x_{m(p,2)})\right), r_{A_{p,3}}(x_{m(p,3)})\right)\dots, r_{A_{p,n_p}}\left(x_{m(p,n_p)}\right)\right)$$

$$\psi_p = f_p\begin{pmatrix}\omega_{A_{p,1}}(x_{m(p,1)}), \omega_{A_{p,2}}(x_{m(p,2)}), \omega_{A_{p,3}}(x_{m(p,3)}), \\ \dots, \omega_{A_{p,n_p}}(x_{m(p,n_p)})\end{pmatrix}.$$

(i) $\mathbf{O}_{p,q}$ = <u>*and*</u> iff $N_{p,q} = T_0$

(ii) $\mathbf{O}_{p,q}$ = <u>*or*</u> iff $N_{p,q} = S_0$

Remark: Similar to the case of FISs, the choice of T_0 would differ depending on the situation and problem that is being studied. Furthermore, in order to remain a faithful generalization of the Mamdani FIS, whenever $\mu_{A_{p,q}}(x_{m(p,q)})$ are real $\forall q$, w_p must be real. The restriction of $f_p(2\pi, 2\pi, \dots, 2\pi) = 2\pi$ is thus imposed for this reason.

Stage 2: Fuzzification.

This stage involves finding the fuzzified input membership function values:

$\mu_{A_{p,q}}(a_{m(p,q)}) = r_{A_{p,q}}(a_{m(p,q)}).\, e^{i\,\omega_{A_{p,q}}(a_{m(p,q)})}$ for all p, q,

Stage 3: Establishing the rule firing strength.

Compute the firing strengths for each complex fuzzy rule. Find the value of w_p, which yields $\tau_p\ \mathrm{e}^{i\,\psi_p}$, where

$$\tau_p = N_{p,n_p-1}\left(\dots N_{p,2}\left(N_{p,1}\left(r_{A_{p,1}}(a_{m(p,1)}), r_{A_{p,2}}(a_{m(p,2)})\right), r_{A_{p,3}}(a_{m(p,3)})\right)\dots, r_{A_{p,n_p}}\left(a_{m(p,n_p)}\right)\right)$$

$$\psi_p = f_p\begin{pmatrix}\omega_{A_{p,1}}(a_{m(p,1)}), \omega_{A_{p,2}}(a_{m(p,2)}), \omega_{A_{p,3}}(a_{m(p,3)}), \\ \dots, \omega_{A_{p,n_p}}(a_{m(p,n_p)})\end{pmatrix}.$$

Stage 4: Calculating the consequence of the complex fuzzy rules.

Choose a function $U_0: [0,1]^2 \to [0,1]$, with $U_0(1,1) = 1$, and a function $g_0: (0,2\pi]^2 \to (0,2\pi]$, with $g_0(2\pi, 2\pi) = 2\pi$. We form the consequent of CFR$_p$ for each p:

$$\Gamma_p(y) = U_0\left(\tau_p, r_{C_p}(y)\right)\mathrm{e}^{i\,g_0\left(\psi_p, \omega_{C_p}(y)\right)}.$$

Remark: The choices of U_0 and g_0 would also depend on the situation and problem that is being studied. For instance, it can be chosen as

$$\Gamma_p(y) = \left(\tau_p r_{C_p}(y)\right)\mathrm{e}^{i\left(\psi_p - \omega_{C_p}(y)\right)} = w_p \cdot \mu_{C_p}(y),$$

where "." denotes the conventional definition of the complex dot product.

Stage 5: Aggregation for the output distribution

Find the output distribution $D(y)$ which is defined as: $\Gamma_1(y) + \Gamma_2(y) + \dots + \Gamma_k(y)$.

Remark 1: $D \in \mathcal{F}(\mathbb{C}, \mathbb{C})$.

Remark 2: The motivation for using such as aggregation is such that the mutual interaction between the inputs in terms of the phase can be taken into account properly. For example, as shown in the CFIS for some of the datasets, the phase term alone take place of the entire pair of mutually opposite literature descriptions, say "Big" and "Small", both for the input and the output, this significantly reduces the number and complexity of the CFR that will be needed in a CFIS as one CFR will take care of both "Big" and "Small" (and any possible description depending on the context) depending on the phase of the input.

Stage 6: Defuzzification.

Choose a function $\Phi: \mathcal{F}(\mathbb{C}, \mathbb{C}) \to \mathbb{C}$. Determine the value of the output $y_{\mathrm{op}} = \Phi(D)$. For instance, we may choose the approximation of $\Phi(D) = \frac{\int_{-\infty}^{\infty} y|D(y)|\mathrm{d}y}{\int_{-\infty}^{\infty}|D(y)|\,\mathrm{d}y}$ using the trapezoidal rule, for all $f \in \mathcal{F}(\mathbb{R}, \mathbb{R})$.

V. Application of Mamdani CFIS to Real Datasets

A. Preliminaries

In order to justify the superiority of our CFIS over the existing model of FIS and ANCFIS [67], an objective way of comparison must be used. This is because, from experience, the performance of the algorithm produced depends on the knowledge, skills, and the machine's limitation. Moreover, it is evident that the longer one invests his time on searching & fine-tuning a FIS/CFIS/ANCFIS, the better the chances of obtaining a better algorithm.

It is for these reasons that we deployed 3 sets of structurally similar programming codes for each data-set, respectively, to *search* for an algorithm for FIS, ANCFIS, and our CFIS for each of the 6 datasets in a similar manner. Particularly in the FIS version, the program does not have to consider any complex numbers, so it runs at least 25 times the speed of the CFIS version but does not necessarily produce better results, as evident from the later sections in the sense that only 5^2 values are tried in each cycle, as compared to 5^4 values for CFIS.

When comparing the results among the three models, namely FIS, ANCFIS and our CFIS model, each of the three versions of the programming code was copied into three executable interfaces on SAGE, running parallel on three different threads of a core i7-4770 CPU. Each one of the three threads will be running the FIS, the ANCFIS, and our CFIS. Not only will our program compute the result of a system, but more importantly, our program contains the AI script *in search of* such systems, thereby simulating what humans will do under similar circumstances. Hence,

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2019.2961350, IEEE Transactions on Fuzzy Systems

all three versions of the program for a given dataset were started simultaneously with the same initial values to perform the search of the best FIS, ANCFIS and CFIS, respectively. Such initial values were all trivial integers like 1 or 0, simulating the worst-case scenario where neither the user not the AI has any prior knowledge on the subject of study. The three programs will then compute the stopped simultaneously after an appropriate interval of time (the total time for all the cycles are between 1 to 8 hours on a same computer, and run using the same program SAGE, version 7.3).

On top of these, it would not be fair if one model (e.g. FIS) is restricted to use linear membership functions, whereas the other models (e.g. our CFIS) can use all kinds of complicated membership functions. So, to ensure fair competition, the FIS, ANCFIS, and CFIS modes of the program will yield a string of numbers suggesting a FIS/ANCFIS/CFIS with the general format as mentioned in section B.

B. The general format yielded from our search algorithm

Throughout the entire section B, let $L \in \mathbb{R}^+$ be a sufficiently large number, so that no input membership values can reach the modulus of 1.

B1. Input and output format

Input : $x_1, x_2, \dots x_m \in \mathbb{R}$

Output : $k \in \mathbb{R}$

The particulars for the inputs and the output involved in each of the 6 datasets are discussed in section D.

B2. Format of the fuzzy rules

B2.1. Format of the fuzzy rules for the 5 datasets besides concrete mixing.

The format is fixed to be as follows:

FIS: up to $4m$+1 rules in the following format:

FR_0 : If x_1 is stable, then k is P
FR_1 : If x_1 is big, then k is $\mathrm{D}_{1,\mathrm{big}}$
FR_2 : If x_2 is big, then k is $\mathrm{D}_{2,\mathrm{big}}$
⋮
FR_m : If x_m is big, then k is $\mathrm{D}_{n,\mathrm{big}}$
FR_{m+1} : If x_1 is small, then k is $\mathrm{D}_{1,\mathrm{small}}$
FR_{m+2} : If x_2 is small, then k is $\mathrm{D}_{2,\mathrm{small}}$
⋮
FR_{2m} : If x_n is small, then k is $\mathrm{D}_{n,\mathrm{small}}$
FR_{2m+1} : If x_1 is very_big, then k is $\mathrm{D}_{1,\mathrm{very\_big}}$
FR_{2m+2} : If x_2 is very_big, then k is $\mathrm{D}_{2,\mathrm{very\_big}}$
⋮
FR_{3m} : If x_n is very_big, then k is $\mathrm{D}_{n,\mathrm{very\_big}}$
FR_{3m+1} : If x_1 is very_small, then k is $\mathrm{D}_{1,\mathrm{very\_small}}$
FR_{3m+2} : If x_2 is very_small, then k is $\mathrm{D}_{2,\mathrm{very\_small}}$
⋮
FR_{4m} : If x_n is small, then k is $\mathrm{D}_{n,\mathrm{very\_small}}$

where $\{\mathrm{D}_{i,\mathrm{big}}, \mathrm{D}_{i,\mathrm{small}}\} = \{\mathrm{D}_{i,\mathrm{very\_big}}, \mathrm{D}_{i,\mathrm{very\_small}}\} = \{\mathrm{P}, \mathrm{N}\}$ for all i.

ANCFIS/CFIS: up to m+1 rules in the following format:

FR_0/CFR_0 : If x_1 is stable, then k is P
FR_1/CFR_1 : If x_1 is big, then k is P
FR_2/CFR_2 : If x_2 is big, then k is P
⋮
FR_n/CFR_n : If x_n is big, then k is P

where n is given for a data-set

The two literature descriptions for : P, N, are the opposite of one another. Their particulars however depend on the context of the data-set (strong-weak, high-low, red-blue etc)

Thus, the advantage is given to FIS as it is allowed to use twice the number of fuzzy rules.

B2.2. Format of the fuzzy rules for concrete mixing dataset.

Based on the background knowledge of concrete mixing, which is well understood by all personnel working in the relevant field, we allow the following structure for all the three models for the dataset of concrete.

FIS/ANCFIS/CFIS:

$CFR_1/FR_1/ANCFR_1$: If x_1 is big **and** x_6 is big, then k is P
$CFR_2/FR_2/ANCFR_2$: If x_2 is big **and** x_6 is big, then k is P
$CFR_3/FR_3/ANCFR_3$: If x_3 is big **and** x_6 is big, then k is P
$CFR_4/FR_4/ANCFR_4$: If x_4 is big, then k is P
$CFR_5/FR_5/ANCFR_5$: If x_5 is big, then k is P

The literature description, **and** , is taken to be the usual multiplication of complex numbers (and hence real numbers).

B3. Input membership functions

FIS:

Four for each x_i (including x_1) : "very big", "big", "small", "very small", "stable".

ANCFIS/CFIS:

Only 1 for each x_i (including x_1): "big" ; "stable" for x_1 only.

The input membership function is fixed to be one among the following type of format, and the same type of format must be used for a particular input x_i in all the three models. Again, the advantage is given to FIS as it is allowed to use four input membership functions instead of one for ANCFIS and CFIS.

B3.1. Input membership functions for "stable"

FIS:

Let $c \in \mathbb{R}$. $\mu_{\mathrm{big}}(x_1) = \left|\frac{c}{L}\right|$. $\mu_{\mathrm{small}}(x_1) = \left|\frac{c}{L}\right|$.

ANCFIS/CFIS:

Let $c \in \mathbb{C}$. $\mu(x_1) = \frac{c}{L}$.

For all the three models, the value of c is to be determined by the program.

B3.2. Input membership functions for the other literature values

(a) Type 1.

FIS:

Let $a_i \in \mathbb{R}$.

If a_i is positive then: i) $(\mathrm{D}_{i,\mathrm{big}}, \mathrm{D}_{i,\mathrm{small}}) = (\mathrm{P}, \mathrm{N})$,

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2019.2961350, IEEE Transactions on Fuzzy Systems

ii) $\mu_{\text{big}}(x_i) = \text{maximum}\left\{\frac{a_i x_i}{L}, 0\right\}$.

$\mu_{\text{small}}(x_i) = -\text{minimum}\left\{\frac{a_i x_i}{L}, 0\right\}$

Otherwise: iii) $(\mathrm{D}_{i,\text{big}}, \mathrm{D}_{i,\text{small}}) = (\mathrm{N}, \mathrm{P})$,

iv) $\mu_{\text{big}}(x_i) = -\text{minimum}\left\{\frac{a_i x_i}{L}, 0\right\}$.

$\mu_{\text{small}}(x_i) = \text{maximum}\left\{\frac{a_i x_i}{L}, 0\right\}$

ANCFIS/CFIS:

Let $a_i \in \mathbb{C}$. $\mu(x_i) = \frac{a_i x_i}{L}$.

For all three models, the value of a_i is to be determined by the program.

(b) Type 2.

FIS:

Let $a_{i,1}, a_{i,2} \in \mathbb{R}$.

If $a_{i,1}$ is positive, then: $(\mathrm{D}_{i,\text{big}}, \mathrm{D}_{i,\text{small}}) = (\mathrm{P}, \mathrm{N})$.

$\mu_{\text{big}}(x_i) = \text{maximum}\left\{\frac{a_{i,1} x_i}{L}, 0\right\}$.

$\mu_{\text{small}}(x_i) = -\text{minimum}\left\{\frac{a_{i,1} x_i}{L}, 0\right\}$.

Otherwise: $(\mathrm{D}_{i,\text{big}}, \mathrm{D}_{i,\text{small}}) = (\mathrm{N}, \mathrm{P})$.

$\mu_{\text{big}}(x_i) = -\text{minimum}\left\{\frac{a_{i,1} x_i}{L}, 0\right\}$.

$\mu_{\text{small}}(x_i) = \text{maximum}\left\{\frac{a_{i,1} x_i}{L}, 0\right\}$.

If $a_{i,2}$ is positive, then: $(\mathrm{D}_{i,\text{very\_big}}, \mathrm{D}_{i,\text{very\_small}}) = (\mathrm{P}, \mathrm{N})$.

$\mu_{\text{very\_big}}(x_i) = \text{maximum}\left\{\frac{a_{i,2} {x_i}^2}{L}, 0\right\}$.

$\mu_{\text{very\_small}}(x_i) = -\text{minimum}\left\{\frac{a_{i,2} {x_i}^2}{L}, 0\right\}$

Otherwise: $(\mathrm{D}_{i,\text{very\_big}}, \mathrm{D}_{i,\text{very\_small}}) = (\mathrm{N}, \mathrm{P})$.

$\mu_{\text{very\_big}}(x_i) = -\text{minimum}\left\{\frac{a_{i,2} {x_i}^2}{L}, 0\right\}$.

$\mu_{\text{very\_small}}(x_i) = \text{maximum}\left\{\frac{a_{i,2} {x_i}^2}{L}, 0\right\}$.

ANCFIS/CFIS:

Let $a_{i,1}, a_{i,2} \in \mathbb{C}$. $\mu(x_i) = \frac{a_{i,1} x_i + a_{i,2} {x_i}^2}{L}$.

For all three models, the value of $a_{i,j}$ is to be determined by the program.

Moreover, for ANCFIS, the normalizing of the firing strength, as well as the dot product, applies to all the firing strengths. Whereas for CFIS, the firing strengths remains as complex numbers.

B3.3. Input membership functions of the input for superplasticizer in the concrete data-set

Based on the background knowledge about concrete, x_6 in the concrete dataset, which is the input for the superplasticizer, is given the following membership function.

FIS:

Let $a_{6,0}, a_{6,1}, a_{6,2} \in \mathbb{R}$.

If $a_{6,0}$ is positive, then: $(\mathrm{D}_{6,\text{stable}}, \mathrm{D}_{6,\text{unstable}}) = (\mathrm{P}, \mathrm{N})$.

$\mu_{\text{stable}}(x_6) = \text{maximum}\left\{\frac{a_{6,0} x_6}{L}, 0\right\}$.

$\mu_{\text{unstable}}(x_6) = -\text{minimum}\left\{\frac{a_{6,0} x_6}{L}, 0\right\}$.

Otherwise: $(\mathrm{D}_{6,\text{stable}}, \mathrm{D}_{6,\text{unstable}}) = (\mathrm{N}, \mathrm{P})$.

$\mu_{\text{stable}}(x_6) = -\text{minimum}\left\{\frac{a_{6,0} x_6}{L}, 0\right\}$.

$\mu_{\text{unstable}}(x_6) = \text{maximum}\left\{\frac{a_{6,0} x_6}{L}, 0\right\}$.

The membership function of $a_{6,1}$ and $a_{6,2}$ follows the usual definition as Type 2 in section B3.2. (b).

ANCFIS/CFIS:

Let $a_{6,0}, a_{6,1}, a_{6,2} \in \mathbb{C}$. $\mu(x_6) = \frac{a_{6,0} + a_{6,1} x_6 + a_{6,2} {x_6}^2}{L}$.

For all three systems, the value of $a_{6,j}$ is to be determined by the program.

B4. Output membership functions

The output membership functions are always as follows:

$$\mu_{\mathrm{P}}(k) = \begin{cases} \frac{k}{1+k}, k \geq 0 \\ 0, k < 0 \end{cases} ; \quad \mu_{\mathrm{N}}(k) = \begin{cases} \frac{-k}{1-k}, k \geq 0 \\ 0, k < 0 \end{cases}$$

Remark: μ_{N} and μ_{P} are monotone functions.

B5. Aggregation

The usual multiplication is used (even for the case of CFIS).

B6. Calculating the consequent

The usual way of multiplication is used.

B7. Aggregation for the consequent functions

The usual way of addition is used.

B8. Defuzzification

Take $k_0 = L\left(\lim_{k\to\infty} D(k) - \lim_{k\to-\infty} D(k)\right)$.

FIS/ANCFIS:

If a real output is desired, then k_0 itself is the defuzzified output value of k. If a Boolean output (i.e. True or False) is desired, then take: $k_1 = \begin{cases} \alpha, & k_0 \leq \xi \\ \beta, & k_0 > \xi \end{cases}$, where $\{\alpha, \beta\} = \{\text{True, False}\}$ and ξ is to be determined by the program to yield the best possible result.

CFIS:

If a real output is desired, then take $k_1 = |k_0|$ as the defuzzified output value of k. If a Boolean output (i.e. True or False) is desired, then take: $k_1 = \begin{cases} \alpha, & k_0 \in L \\ \beta, & k_0 \notin L \end{cases}$, where $\{\alpha, \beta\} = \{\text{yes, no}\}$ and $L \subseteq \mathbb{C}$ is to be determined by the program to yield the best possible result.

In our testing, L is restricted to be the region enclosed by a loop, $\mathcal{O}_L$, on the Argand diagram, and is of the form:

$\mathcal{O}_L : r = f(\theta)$,

where $r = |k| \geq 0$ and $\theta = \arg(k) \in (-\pi, \pi]$,

and with: $\frac{\mathrm{d}r}{\mathrm{d}\theta}$ defined for *all* $\theta \in (-\pi, \pi]$,

such as (but not limiting to) a circle centered at 0. This is again another advantage given to FIS and ANCFIS as the choice of loops greatly affect the versatility of CFIS.

Under all our setups, it was observed that k_0 can be explicitly expressed in all the previously defined terms such as , $a_i x_i$ and $a_{i,1} x_i + a_{i,2} {x_i}^2$, It is for this reason that we put such explicit expressions into our SAGE codes to speed up the computation all the three models. In fact, this is again another advantage given to FIS as it needs a more complicated structure to produce the same result.

C. *The use of SAGE in our work*

C1. The way SAGE simulates how a human would search for the best CFIS model.

Suppose a human were to come up with a FIS/CFIS model, $\mathfrak{C}$, and is looking for a better one. He now looks for

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2019.2961350, IEEE Transactions on Fuzzy Systems

a better one by altering certain part of $\mathfrak{C}$ (say, the 3rd fuzzy rule) following a certain pattern (say linearly shifting the fuzzy membership function of the 5th input) to several intensities which are quantitative in nature (say, shifting that function to the left by either 1, 2, 3 or 4). Depending on his capability, and the complexity of the FIS/CFIS model, he may consider *all the combinations of several independent ways* of changing $\mathfrak{C}$ at a time. For example:

i) linearly shifting the fuzzy membership function of the 5th input in the 3rd fuzzy rule,
ii) linearly shifting the fuzzy membership function of the 2nd input in the 4th fuzzy rule,
iii) linearly shifting the fuzzy membership function of the output in the 5th fuzzy rule,
iv) scaling up/down the fuzzy membership function of the 7th input in the 8th fuzzy rule,

thereby forming a set of ways of changing $\mathfrak{C}$, denoted as W.

If the manages to find a better FIS/CFIS model, $w(\mathfrak{C})$, for some $w \in W$ which is different from his initial model $\mathfrak{C}$, then one of the most likely scenarios is that he will start from $w(\mathfrak{C})$, and then continue to apply the same kind of modification *and* the same intensity of variation as he had done before to see if "there is an even better one if he keeps on going in the same direction".

On the other hand, if after all the efforts, $\mathfrak{C}$ is still deemed the best, then there are two scenario that may happen.

Firstly, he may "narrow down" the search, by using the same kind of modification, but to a lesser intensity (say, instead of shifting a membership function by 1, 2, or 3 units; he may now shift it by 0.1, 0.2, or 0.3 units).

Secondly, he may abandon all the ways of changing $\mathfrak{C}$ that he had chosen to deal with previously, and try out another set of totally different ways of changing $\mathfrak{C}$, denoted as W', which is mutually exclusive with W. When this is done, it is very unlikely that the shifting intensity will be small right away, as he still has no idea on the effect of the new way of changing $\mathfrak{C}$. Thus, he will likely begin with several appropriately large intensities, and then "narrow down the search" again.

The SAGE program that we have designed was developed/programmed to simulate such decision making procedures that are normally done by humans but in a much more efficient manner compared to what a human would do, thereby producing better results in a shorter period of time.

C2. The reasons of using SAGE instead of Fuzzy Toolbox from Matlab.

There has been a proprietary, graphical user interface dedicated for Fuzzy Computing called the Fuzzy Toolbox from Matlab. This interface allows users to enter fuzzy membership functions and input, and the program calculates the output in return. Apparently, it is the most widely used software, but there are several major disadvantages in using the Fuzzy Toolbox from Matlab and these are highlighted below.

Firstly, on Fuzzy Toolbox from Matlab, the choice of input membership function needs to be entered by the user. In other words, this program is only meant to run a FIS, ANCFIS or CFIS based on the inputs by the user, but it cannot generate an FIS. As the determination of a fuzzy inference system depends on the time spent by the person, and also the skill they possess, such comparison between different systems is fair, only if each system is handled by a team with an equal number of people in each team, and the members of each team possess equal skills.

It is primarily for this reason that we deploy an AI script that is able to search for the best FIS, instead of relying on different teams working on different projects to provide the necessary input to run the program.

On top of these, most of the features on Fuzzy Toolbox do not have the capability to search for complex fuzzy systems, such as our CFIS. Even on FIS, Fuzzy Toolbox only allows us to choose the input membership function from a list of preset entities. It is therefore biased to judge a fuzzy system simply by its capability to run on an end-user GUI like platform. Moreover, most devices who incorporate fuzzy system rely on OEM circuits to do the fuzzy computing. It is therefore more convincing by using a lower level (say, pure programming script instead of GUI) computing procedure, so that we can more accurately mimic what is happening in an electronic circuit.

Not to mention that, the Fuzzy Toolbox from Matlab uses fuzzy systems that are pre-defined, and due to the proprietary nature of this software, the user may never know the underlying algorithms of the fuzzy system he uses. Thus, there is no way such interface can itself reflect on the credibility of such fuzzy system, whether in terms of correlation coefficient, or the percentage of correctness. This is made worse by the fact that each fuzzy system established must be tested on hundreds, if not thousands, of different *sets* of input, from the dataset itself.

C3. A summary with two diagrams.

Here we represent two diagrams that clearly shows the advantages of using SAGE over the Fuzzy Toolbox on Matlab (see Fig. 2 and 3). In particular, each of our steps in Matlab is given more elaboration as necessary in the next section.

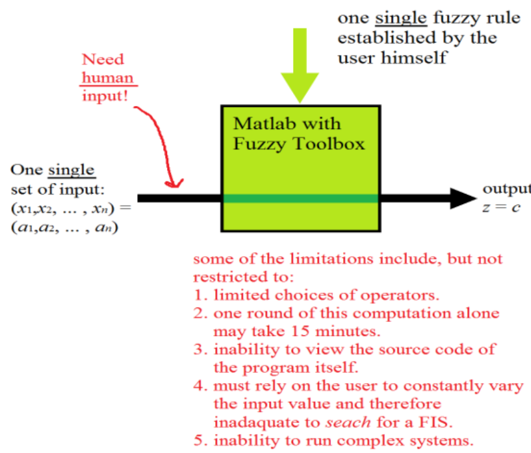


Fig. 2. The workings of the fuzzy toolbox on Matlab and its disadvantages

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2019.2961350, IEEE Transactions on Fuzzy Systems

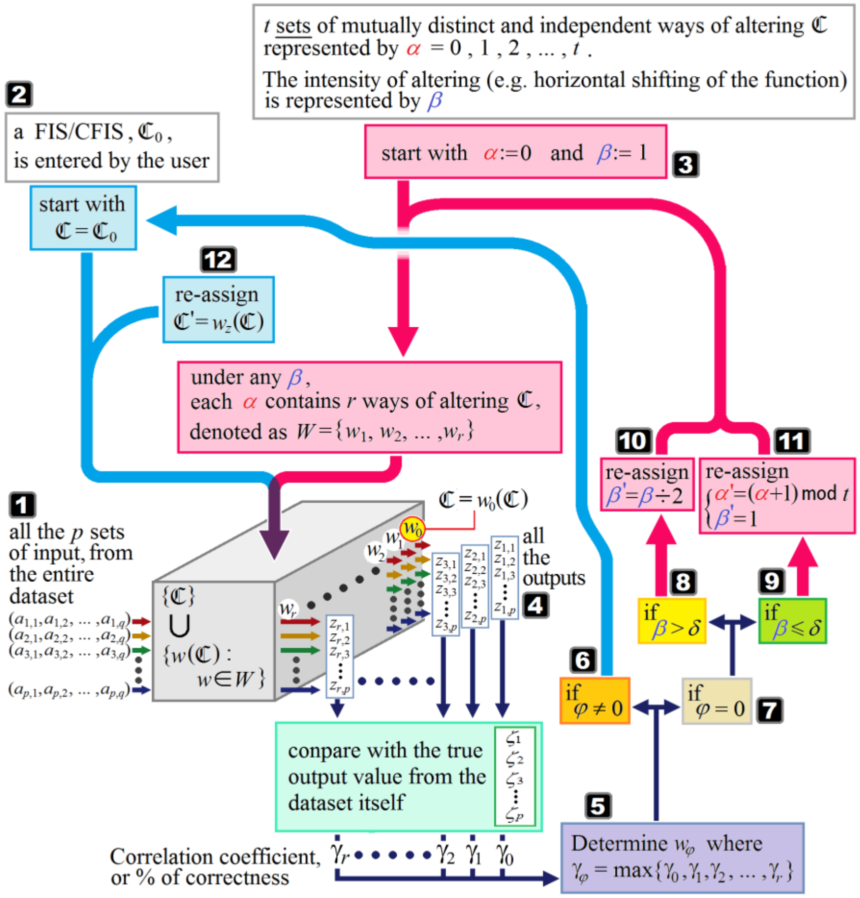


Fig. 3. The framework of our programming model using SAGE, each of the steps marked with a circled number will be further elaborated in section B4

C4. The further elaborations on our framework

With regards to each of the steps marked with a circled number, we now give further elaboration on each of them. Each of the numbers below corresponds to the steps labeled with that same number.

1. All at once, the entire datasets are put into consideration by the program. This is much more efficient than the fuzzy toolbox which require the user to manually enter every single set of input.
2. FIS (Fuzzy Interference System) and CFIS (Complex Fuzzy Interference System), that we started with, can be any trivial ones. For example, in the detection of breast cancer, it is well understood that the larger the tumor. Thus, upon choosing the starting FIS/CFIS, users are given considerable latitude to decide their FR/CFR, simply based on the qualitative relationship between the variables, it is the AI script contend in the SAGE code itself that search for a better and better FIS/CFIS after the initial input. This is contrary with the fuzzy toolbox from Matlab, which is just a tool to test a manually input FIS/CFIS, and does nothing to improve from there. Moreover, the Fuzzy Toolbox of Matlab severely limits the choice of membership function mainly to those consist of only straight lines, and with the formula not displayed; where as in SAGE, each of the membership function can be fully customized with an explicit formula, and such formula can be clearly seen in the programming code.
3. In real life datasets, there could be numerous inputs, all of which must be taken into account. Thus, in the ever-going search for a better FIS/CFIS, The FR/CFR on each of the inputs will be taken into account and thus modified. In the fuzzy toolbox of Matlab, the user himself have to decide which FR/CFR to alter, and the he will have to manually take down the results each time he tries a new set of FR/CFR. On the other hand, our SAGE framework autonomously generated several independent ways (t in Fig. 3) of altering the FR/CFR. As shown in the diagram, each of $\alpha = 0,1, \ldots , t-1$ denote a set of different ways of altering a particular set of membership function in the FIS/CFIS. During each α, one or several membership functions in the FIS/CFIS will be shifted by an amount of $k\beta$, where k belongs to a particular finite subset of real number S_α. In the beginning, β=1. Thus, when β is assigned a lower value, $k\beta$ will clearly decrease for all k in S_α. Again, in case of the Matlab's Fuzzy Toolbox, there is no way a user can systematically alter the fuzzy membership function, as the functions are usually presented only graphically through a graphical user interface. Furthermore, all the adjustments on the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2019.2961350, IEEE Transactions on Fuzzy Systems

characteristic of a fuzzy membership functions must be done by hand. Whereas in our program, with just a choice of α and β provided, the program handles all the tedious work hereafter.

4. The original FIS/CFIS chosen by the user in step 2, is now subjected to the α^{th}-set of way of modification, and with the shifting factor β as described in step 3. This yields a collection of new FIS/CFIS's. Then, for each of the FIS/CFIS in hand (whether new or old), the output values are calculated for each of the set of input values from the dataset. For instance, with 3 FIS/CFIS's in hand, and with a dataset having 5 sets of input values, there will be a total of 15 output values obtained, each correspond to a particular set of input under a particular kind of FIS/CFIS. It is again clear that the fuzzy toolbox in Matlab clearly incapable of doing this.
5. The set of outputs calculated from each of the FIS/CFIS are compared with the true values from the dataset. Each set of output is assessed on how close it is from the true values, by calculating a quantitative value of representative, such as the correlation coefficient. The FIS/CFIS that yields the value closest to the true values is identified (referred as the *best FIS/CFIS* in all the following procedures). Such feature is again absent from Matlab, not to mention that we can customize our own way of assessment only in SAGE.
6. If the best FIS/CFIS is not the original FIS/CFIS from step 2, then proceed to step 12.
7. If the best FIS/CFIS is still the original FIS/CFIS from step 2, Then the value of β will be given attention.
8. If β is not small enough (the threshold δ is decided by the user), proceed to step 10.
9. If β is already small enough, proceed to step 11.
10. Halve the value of β (i.e. narrow down the search), and is subject to the same α. Presume the search further by going to step 4, and with the original FIS/CFIS.
11. Choose a different α (i.e. modifying FIS/CFIS in different way), and the shifting factor β will be reset to 1. Presume the search further by going to step 4, and with the original FIS/CFIS.
12. Replace the original FIS/CFIS with the best FIS/CFIS and will be used for the next round of search. The same way of altering the FIS/CFIS, α will be used again, and the shifting factor β will be reset to 1. Presume the search further by going to step 4, but with the original FIS/CFIS replaced by the best FIS/CFIS found.

Likewise, step 6 to 12 can only be manually done by the user for each cycle if using the Fuzzy Toolbox from Matlab.

D. Results obtained from application on six real-life datasets

In this section, we present the application of our Mamdani CFIS to six real-life datasets sourced from the UCI Machine Learning Repository. The six datasets are Concrete Slump, Breast Cancer Wisconsin, Image Segmentation, Energy Efficiency, Chronic Kidney Disease and Wine Quality. A brief explanation consisting of the intention of the CFIS, the results obtained via our CFIS, and the comparison between the results obtained via the Mamdani FIS and ANCFIS [13] methods are provided for each of the six datasets.

i. The compressive strength of concrete:
https://archive.ics.uci.edu/ml/datasets/Concrete+Slump+Test on "cslump data.txt" with 103 samples.

-Description on the scenario

The mixing of concrete involves adjusting the quantities of ingredients, to which there are no known explicit mathematical formula allowing one to precisely compute the output strength from the amount of each of the component. Moreover, as different chemicals are mixed together, some unexpected deviation can result, like the volume shrinkage when water mixed with ethanol. It is for these reasons that we choose this dataset to demonstrate our CFIS, and to compare with the existing FIS and ANCFIS.

-Intention of the CFIS

Construct the CFIS that generates the best approximation of concrete compressive strength.

-Name of the input and output

Input:
x_1=cement to water ratio; x_2=slag to water ratio;
x_3=fly ash to water ratio; x_4=coarse aggregate $\times 0.001$;
x_5=fine aggregate $\times 0.001$; x_6=superplasticizer

Output: k = compressive strength (numerical)
Literature description – P = Strong, N = Weak

-Results from the program

Multipliers: $c = 2.200$.
$a_{1,1} = 1.000$, $a_{1,2} = 0.600\text{i}$.
$a_{2,1} = 0.100\text{i}$, $a_{2,2} = 0.100 - 0.500\text{i}$.
$a_{3,1} = 0.800 - 0.600\text{i}$, $a_{3,2} = -0.100$
$a_{4,1} = 3.200 - 1.400\text{i}$, $a_{4,2} = -2.200 + 1.600\text{i}$,
$a_{5,1} = 13.100 + 7.000\text{i}$, $a_{5,2} = -8.800 - 9.600\text{i}$,
$a_{6,0} = 0.600 + 0.100\text{i}$, $a_{6,1} = 0.000$, $a_{6,2} = 0.000$

i.e. It is deemed that the superplasticizer does not affect the compressive strength of the concrete at all, which is consistent with its intended purpose: to make the concrete more free-flowing.

Duration of computation: 2.5 hours.

-Comparison of our CFIS with FIS and ANCFIS

	Correlation coefficient of k with the actual output for all the 103 samples: (the closer to 1, the better)
Our CFIS	**0.971477**
ANCFIS	0.790867
Mamdani FIS	0.967225

ii. The detection of breast cancer:
https://archive.ics.uci.edu/ml/machine-learning-databases/breast-cancer-wisconsin/ on "wdbc data.txt" with 569 samples.

-Description on the scenario

In this case, although there is a common qualitative understanding that "the larger the tumor, the more likely it is malignant", again there are no known explicit mathematical formula allowing one to precisely determine the nature of the tumor.

-Intention of the CFIS

Construct the CFIS that produce the fewest misdiagnosis of breast cancer.

-Name of the input and output

Input:

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2019.2961350, IEEE Transactions on Fuzzy Systems

x_1=mean reading of radius,
x_2=mean reading of concavity,
x_3=mean reading of concave points,
x_4=mean reading of smoothness,
x_5=mean reading of compactness,
x_6=variance reading of concavity,
x_7=variance reading of concave points,
x_8=variance reading of smoothness,
x_9=variance reading of compactness,
Output: $k \in$ {Benign, Malignant} (Boolean)
Literature description - Dangerous

-*Results from the program*

$c = 0$ (i.e. FR$_0$/CFR$_0$ not needed)
$a_1 = 1.0$, $a_2 = 4.0 - 16.0\text{i}$, $a_3 = 60.0 - 56.0\text{i}$,
$a_4 = 4.0 + 52.0\text{i}$, $a_5 = 8.0 - 28.0\text{i}$,
$a_6 = -4.0 + 28.0\text{i}$, $a_7 = 48.0 + 8.0\text{i}$,
$a_8 = -32.0 - 120.0\text{i}$, $a_9 = -76.0 - 64.0\text{i}$
Radius of the best circle to enclose the region, r_0 = 17.610513.
Duration of computation: 1.0 hour

-*Comparison of our CFIS with FIS and ANCFIS*

For our CFIS: $\mathcal{O}_L$ is chosen to be $r = r_0$ for all $\theta \in (-\pi, \pi]$ (i.e. using the circle as suggested).

	Number of misdiagnosed cases out of 569
Our CFIS	**33**
ANCFIS	45
Mamdani FIS	42

iii. Object recognition

https://archive.ics.uci.edu/ml/datasets/Image+Segmentation
on “seg test.txt” with 2100 samples.

-*Description on the scenario*

Object recognition is another sector that is very important in crime investigation. In this scenario, brick face is used as the target, but such method can be free adopted to the recognition of any objects or even a particular person’s face.

-*Intention of the CFIS*

Construct the CFIS that recognizes an object (brick face in this example) to the highest success rate.

-*Name of the Input and Output*

Input:
x_1=hue-mean; x_2=region-centroid-row;
x_3=saturatoin-mean; x_4=rawblue-mean;
x_5=exred-mean; x_6=vedge-mean;
x_7=vegde-sd; x_8=hedge-mean;
x_9=hedge-sd; x_{10}=exgreen-mean
Output: $k \in$ {Object, Background} (Boolean)
Literature description - Background

-*Results from the program*

$c = 0$ (i.e. FR$_0$/CFR$_0$ not needed), a_1= 20.0
a_2= 0.1 ($-$0.750000000000 + 1.50000000000 i) ,
a_3= 20.0 (2.00000000000 $-$3.62500000000 i)
a_4= 0.1 (4.87500000000 + 7.37500000000 i) ,
a_5= 0.5 (2.75000000000 + 0.500000000000 i)
a_6= 1.0 ($-$4.58105468750 $-$3.51855468750 i) ,
a_7= 0.05 (1.08398437500 + 21.3076171875 i)
a_8= 1.0 (0.750000000000 + 0.830078125000 i) ,
a_9= 0.05 ($-$2.25195312500 + 23.2187500000 i)
a_{10}= 1.0 (0.000000000000 + 0.000000000000 i)

Radius of the best circle to enclose the region: r_0 = 31.187714.
Duration of computation: 5.0 hours.

-*Comparison of our CFIS with FIS and ANCFIS*

$\mathcal{O}_L$ is chosen to be

$$r = \begin{cases} r_0 + 48000\big((x+3.00)(x+2.62)\big)^2(x+2.83) \\ \quad \text{for } -3.00 < x \le -2.95 \\ r_0 + 48000\big((x+3.00)(x+2.62)\big)^2(x+2.83) \\ \quad +50000\big((x+2.95)(x+2.83)\big)^2 \\ \quad \text{for } -2.95 < x \le -2.83 \\ r_0 + 48000\big((x+3.00)(x+2.62)\big)^2(x+2.83) \\ \quad +10000\big((x+2.62)(x+2.83)\big)^2, \\ \quad \text{for } -2.83 < x < -2.62 \\ r_0, \ \text{otherwise} \end{cases}$$

	Number of mistakes out of 2100 objects
Our CFIS	**0**
ANCFIS	2
Mamdani FIS	5

iv. Thermal efficiency of building

https://archive.ics.uci.edu/ml/datasets/Energy+efficiency
on “ENB2012_data.xslx” with 768 samples.

-*Description on the scenario*

With the issue of energy conservation, the cooling capacity of premises is of major concern, especially among countries of not climate. Theoretically, a house cooling capacity can be computed with formula, but practically that is usually a tedious process as it involves a thorough understanding of the detailed chemical composition of *each* of the house building materials, including the concrete and the re-bar within (which cannot be seen as it is embedded in the concrete). In fact, the specific heat capacity of many common building material can still differ greatly among different production n batches, whose value are usually unknown even to the producer of those building materials. As a result, we deploy our CFIS to see how well it can predict the results just based on the extremely limited information provided in the dataset.

-*Intention of the CFIS*

Construct the CFIS that generate the best approximation of a house’s cooling load based on its composition.

-*Name of the Input and Output*

Input:
x_1=relative compactness; x_2=surface area;
x_3=wall area; x_4=roof area;
x_5=overall height; x_6=orientation;
x_7=glazing area; x_8=glazing area distribution
Output: k= Cooling Load
Literature description - High

-*Results from the program*

Multipliers $c = 0.0000$, $a_1 = 1.0000$,
$a_2 = -0.0300 + 0.0900\text{ i}$, $a_3 = 0.0900 - 0.1140\text{ i}$,
$a4 = 0.0200 + 0.1000\text{ i}$, $a_5 = 8.9000 - 0.5000\text{ i}$,
$a_6 = 0.0000$, $a_7 = 20.0000$, $a_8 = 0.0000$

-*Comparison of our CFIS with FIS and ANCFIS*

	Correlation coefficient of k with the actual output for all the 768 samples: (the closer to 1, the better)
Our CFIS	**0.943386**

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2019.2961350, IEEE Transactions on Fuzzy Systems

ANCFIS	0.942791
Mamdani FIS	0.938078

v. The detection of chronic kidney disease:
https://archive.ics.uci.edu/ml/datasets/Chronic_Kidney_Disease on "chronic_kidney_disease A.txt" with 400 samples.
-*Description on the scenario*
Again, like the example of breast cancer, there are no known explicit mathematical formula allowing one to precisely determine if one suffer from the disease just from the readings alone. Moreover, the unusual reading from urine may not necessarily indicate the presence of the disease as there could be other courses, such as taking certain medication, or could be due to external/internal injury of the patient. Our CFIS thus must be able to detect the disease even under this circumstance.
-*Intention of the CFIS*
Construct the CFIS that produces the fewest misdiagnosis of chronic kidney disease.
-*Name of the Input and Output*
All numerical values are normalized by the reading mean and the variance of that from a healthy person. In this testing, those means and the variances are taken from all the normal patients from the dataset itself.
Input:
x_1=Blood Urea; x_2=Serum Creatinin;
x_3=Hemoglobin; x_4=Packed Cell Volume;
x_5=Specific Gravity; x_6=Blood Glucose Random;
x_7=Sodium; x_8=Red Blood Cell Count;
x_9=Albumin; x_{10}=Sugar;
x_{11}=Red Blood Cells Status (0=normal , 1=abnormal);
x_{12}=Pus Cell (0=normal , 1=abnormal);
x_{13}=Pus Cell clumps (0=not present , 1=present);
x_{14}=Bacteria (0=not present , 1=present);
x_{15}=Blood Pressure; x_{16}=Potassium;
x_{17}=White Blood Cell Count
Output: $k \in$ {Normal, Disease} (Boolean)
Literature description - Dangerous
-*Results from the program*
$c = 12.40 - 12.20\mathrm{i}$, $a_1 = 1.000$
, $a_2 = 6.600 + 1.200\mathrm{i}$,
$a_3 = -6.800 - 3.100\mathrm{i}$,
$a_4 = 0.000 + 3.300\mathrm{i}$, $a_5 = -7.300 - 0.900$
$a_6 = -0.600 - 3.700\mathrm{i}$, $a_7 = 0.300 - 2.000\mathrm{i}$
$a_8 = -2.300 - 4.800\mathrm{i}$, $a_9 = -25.10 - 33.10\mathrm{i}$
$a_{10} = -30.60 - 32.40\mathrm{i}$, $a_{11} = -26.60 - 32.40\mathrm{i}$,
$a_{12} = -29.00 - 31.70\mathrm{i}$, $a_{13} = -32.00 - 32.00\mathrm{i}$,
$a_{14} = -32.60 - 32.60\mathrm{i}$, $a_{15} = -0.200 - 2.400\mathrm{i}$,
$a_{16} = 1.400 + 1.000\mathrm{i}$, $a_{17} = -3.300 - 4.700\mathrm{i}$,
Radius of the best circle to enclose the region, r_0 = 40.093495
Duration of computation: 8.0 hours
-*Comparison of our CFIS with FIS and ANCFIS*
For our CFIS: $\mathcal{O}_L$ is chosen to be $r = r_0$ for all $\theta \in (-\pi, \pi]$ (i.e. using the circle as suggested).

	Number of misdiagnosed cases out of 569
Our CFIS	**2**
ANCFIS	21
Mamdani FIS	4

vi. Wine Quality
https://archive.ics.uci.edu/ml/datasets/wine+quality on 'winequality-white.csv' with 4898 samples.
-*Description on the scenario*
In this scenario, the wine quality score is assessed by having people tasting the wine. It is therefore regarded as a very subjective matter. We nonetheless use this example to demonstrate how our CFIS can detect a possible objective trend that influence the flavor of the wine.
-*Intention of the CFIS*
Construct the CFIS that generates the best approximation of a wine's quality load based on its composition.
-*Name of the Input and Output*
Input:
x_1=fixed acidity ; x_2=volatile acidity;
x_3=citric acid; x_4=residual sugar;
x_5=chlorides; x_6=free sulfur dioxide;
x_7=total sulfur dioxide; x_8=density;
x_9=pH; x_{10}=sulfates; x_{11}=alcohol
Output: k= Quality (numerical, but need *not* be an integer)
Literature description - Good
-*Results from the program*
Multipliers
c = 30.50 + 30.50i, $a_1 = -$ 1.000, a_2 = 1.000−5.000i,
a_3 = 1.500−1.000i, a_4 = 3.500−1.000i,
a_5 = −2.500 + 1.000i, a_6 = 1.500 + 0.000i ,
a_7 = −0.500 + 0.000i, a_8 = −6.000 + 4.000i,
a_9 = 2.500−2.000i, a_{10} = 1.000 + 0.000i,
a_{11} = 6.000 + 0.000i,
-*Comparison of our CFIS with FIS and ANCFIS*

	Correlation coefficient of k with the actual output for all the 4898 samples: (the closer to 1, the better)
Our CFIS	**0.542536**
ANCFIS	*unable to produce any results*
Mamdani FIS	0.494140

Remark: The quality score of the wine is usually regarded as a subjective matter. This is because it is measured by a person according to his/her taste. Hence, a correlation coefficient of 0.542536 obtained from this event is considered very significant as it indicates a possible trend of preference over a matter that is generally regarded as being subjective.
It can be clearly seen from all the six applications presented above that our proposed Mamdani CFIS outperforms the Mamdani FIS and ANCFIS every time in terms of the accuracy and reliability of the results obtained, as our CFIS is the only model that takes the phase term into consideration throughout the decision-making process.

VI. Conclusion

In this paper, we proposed a novel Mamdani based complex fuzzy inference system (Mamdani CFIS) that integrates the concept of a complex fuzzy inference system (CFIS) and the classical Mamdani FIS. Our Mamdani CFIS is a faithful generalization of the original Mamdani FIS that takes into consideration the information pertaining to the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2019.2961350, IEEE Transactions on Fuzzy Systems

phase term throughout the decision-making process. This makes it a highly suitable and efficient tool to handle time-periodic data. To prove this, we applied our proposed Mamdani CFIS to six real-life datasets representing a variety of situations. To further validate our Mamdani CFIS, we verified the results obtained via our model with the actual datasets, and also compared the results obtained via other methods in literature. We were able to successfully prove that our proposed Mamdani CFIS is not only able to produce more reliable and accurate results compared to the other methods, but also outperforms the Mamdani FIS and ANCFIS methods in terms of computational time, computational complexity and accuracy. This proves the reliability, validity and superiority of our proposed Mamdani CFIS framework.

Further research since this work will involve the investigation of new complex fuzzy measures and operations to be adapted to the proposed Mamdani CFIS. In addition, we will also implement dynamic complex fuzzy rules as a way to further improve the accuracy, reliability and flexible of the model.

Moreover, the search algorithm will also be improved to incorporate more different types of input and output membership functions, particularly those that can only be implemented using our proposed Mamdani CFIS.

ACKNOWLEDGEMENT

The authors (Ganeshsree Selvachandran and Quek Shio) would like to gratefully acknowledge the financial assistance received from the Ministry of Education, Malaysia under Grant FRGS/1/2017/STG06/UCSI/03/1.

The authors (Nguyen Long Giang, Le Hoang Son, Luong Thi Hong Lan) would like to acknowledge Vietnam Academy of Science and Technology for funding by the Research Project, VAST 01.05/19-20.

The author (Weiping Ding) would like to acknowledge the National Natural Science Foundation of China under Grant 61976120, the Natural Science Foundation of Jiangsu Province under Grant BK20191445, the Six Talent Peaks Project of Jiangsu Province under Grant XYDXXJS-048, and sponsored by Qing Lan Project of Jiangsu Province.

REFERENCES

[1] L.A. Zadeh, “Fuzzy sets.” *Information and Control*, vol. *8*, pp. 338-353, 1965.

[2] D. Ramot, R. Milo, M. Friedman, & A. Kandel, “Complex fuzzy sets.", *IEEE Transactions on Fuzzy Systems*, vol. 10(2), pp. 171-186, 2002.

[3] D. Ramot, M. Friedman, G. Langholz, & A. Kandel, "Complex fuzzy logic.", *IEEE Transactions on Fuzzy Systems*, vol 11(4), pp. 450-461, 2003.

[4] S. Dick, Scott, "Toward complex fuzzy logic.", *IEEE Transactions on Fuzzy Systems*, vol. 13(3), pp. 405-414, 2005.

[5] J. Ma, G. Zhang, and J. Lu, "A method for multiple periodic factor prediction problems using complex fuzzy sets.", *IEEE Transactions on Fuzzy Systems*, vol. 20(1), pp. 32-45, 2012.

[6] O. Yazdanbakhsh, and S. Dick, "Multi-variate timeseries forecasting using complex fuzzy logic.", *Fuzzy Information Processing Society (NAFIPS) held jointly with 2015 5th World Conference on Soft Computing (WConSC), 2015 Annual Conference of the North American*. IEEE, 2015.

[7] O. Yazdanbaksh, A. Krahn, and S. Dick, "Predicting solar power output using complex fuzzy logic.", *IFSA World Congress and NAFIPS Annual Meeting (IFSA/NAFIPS)*, 2013.

[8] D. E. Tamir, and A. Kandel, "Axiomatic theory of complex fuzzy logic and complex fuzzy classes.", *International Journal of Computers Communications & Control*, vol. 6(3), pp. 562-576, 2011.

[9] D. E. Tamir, M. Last, and A. Kandel, "The theory and applications of generalized complex fuzzy propositional logic.", *Soft Computing: State of the Art Theory and Novel Applications,* pp. 177-192, 2013.

[10] D. E. Tamir, N. D. Rishe, and A. Kandel, "Complex Fuzzy Sets and Complex Fuzzy Logic an Overview of Theory and Applications.", *Fifty Years of Fuzzy Logic and its Applications*, Springer International Publishing, pp. 661-681, 2015.

[11] O. Yazdanbakhsh, and S. Dick, "A systematic review of complex fuzzy sets and logic.", *Fuzzy Sets and Systems,* 2017. [Online] Doi: 10.1016/j.fss.2017.01.010.

[12] J. Y. Man, Z. Chen, and S. Dick, "Towards inductive learning of complex fuzzy inference systems.", *Fuzzy Information Processing Society, NAFIPS'07. Annual Meeting of the North American*. IEEE, 2007, pp. 415-420.

[13] Z. Chen, S. Aghakhani, J. Man & S. Dick, "ANCFIS: A neurofuzzy architecture employing complex fuzzy sets.", *IEEE Transactions on Fuzzy Systems*, vol. 19(2), pp. 305-322, 2011.

[14] J. S. R. Jang, "ANFIS: adaptive-network-based fuzzy inference system.", *IEEE transactions on systems, man, and cybernetics, vol* 23(3), pp. 665-685, 1993.

[15] J. S. R. Jang, “Fuzzy modelling using generalized neural networks and Kahman filter algorithm.”, *Proceedings of the Ninth National Conference on Artificial Intelligence* (AAAI-91), pp. 762-767, 1991.

[16] F. Camastra, A. Ciaramella, V. Giovannelli, M. Lener, V. Rastelli, A. Staiano, G. Staiano, & A. Starace, "A fuzzy decision system for genetically modified plant environmental risk assessment using Mamdani inference.", *Expert Systems with Applications,* vol. 42(3), pp. 1710-1716, 2015.

[17] B. M. Gayathri, and C. P. Sumathi, "Mamdani fuzzy inference system for breast cancer risk detection.", *Proceedings of the IEEE International Conference on Computational Intelligence and Computing Research*, Madurai, India, 2015.

[18] E. Erturk, and E. A. Sezer, "Software fault prediction using Mamdani type fuzzy inference system.", *International Journal of Data Analysis Techniques and Strategies,* vol. 8(1), pp. 14-28, 2016.

[19] P. Mamoria, and D. Raj, "Comparison of Mamdani Fuzzy Inference System for Multiple Membership Functions.", *International Journal of Image, Graphics and Signal Processing,* vol. 9, pp. 26-30, 2016.

[20] M. R. Mohanraj, M. Balamurugan, V.P. Suresh, & R. Gobu, “Design of air-conditioning controller by using Mamdani and Sugeno fuzzy inference systems.”, *South Asian Journal of Engineering and Technology*, vol. 2(16), pp. 7-16, 2016.

[21] M. D. Ruzic, J. Skenderovic, & K.T. Lesic, “Application of the Mamdani fuzzy inference system to measuring HRM performance in hotel companies – A pilot study.”, *Theory in Practice*, vol. *53*(4), pp. 976-999, 2016.

[22] S. Thakur, S.N. Raw, R. Sharma, & P. Mishra, "Detection of type of thalassemia disease in patients: A fuzzy logic approach.", *International Journal of Applied Pharmaceutical Sciences and Research,* vol. 1(2), pp. 88-95, 2016.

[23] N. U. Ahamed, L. Benson, C. Clermont, S.T. Osis, & R. Ferber, "Fuzzy Inference System-based Recognition of Slow, Medium and Fast Running Conditions using a Triaxial Accelerometer.", *Procedia Computer Science,* vol. 114, pp. 401-407, 2017.

[24] P. K. Borkar, M. Jha, M.F. Qureshi,. & G.K. Agrawal, “Performance assessment of heat exchanger using Mamdani based adaptive neuro-fuzzy inference system (M-ANFIS) and dynamic fuzzy reliability modeling.”, *International Journal of Innovative Research in Science, Engineering and Technology*, vol. *3*(9), pp. 16083-16096, 2014.

[25] Y. Chai, L. Jia, and Z. Zhang, "Mamdani model based adaptive neural fuzzy inference system and its application.", *International Journal of Computational Intelligence,* vol 5(1), pp. 22-29, 2009

[26] W. Ding, C.-T. Lin, and Z. Cao. “Deep neuro-cognitive co-evolution for fuzzy attribute reduction by quantum leaping PSO with nearest-neighbor memeplexes”. *IEEE Transactions on Cybernetics*, vol. 49, no. 7, pp. 2744-2757, July 2019.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TFUZZ.2019.2961350, IEEE Transactions on Fuzzy Systems

[27] A. Malekzadeh, and M. R. Akbarzadeh, "Complex-valueo adaptive neuro fuzzy inference system-canfis.", *Proceedings of the World Automation Congress, Seville, Spain*, Vol. 17, pp. 477-482, 2004.

[28] Y. Deshmukh, A.B. Bavaskar, P.R. Bajaj, & A.G. Keskar, "Implementation of complex fuzzy logic modules with VLSI approach." , *International Journal on Computer Science and Network Security*, vol. 8(9), pp. 172-178, 2008.

[29] A.A. El-Harby & E. Algrafy, "Efficient Neuro-Fuzzy Inference System (ANFIS) and Neural Networks Systems for Different Beams Collisions with Light Nuclei". *Asian Journal of Scientific Research*, vol. 12(1), pp. 71-78, 2019.

[30] H. Moayedi, M. Mehrabi, B. Kalantar, M. Abdullahi Mu'azu, A. Safuan A. Rashid & L. K. Foong, "Novel hybrids of adaptive neuro-fuzzy inference system (ANFIS) with several metaheuristic algorithms for spatial susceptibility assessment of seismic-induced landslide", *Geomatics, Natural Hazards and Risk*, vol. 10(1), pp. 1879-1911, 2018.

[31] S. Sremac, E. K. Zavadskas, B. Matić, M. Kopić & Ž. Stević, "Neuro-fuzzy inference systems approach to decision support system for economic order quantity", *Economic Research-Ekonomska Istraživanja*, vol. 32(1), pp. 1114-1137, 2019.

[32] O. Yazdanbakhsh, S. Dick, FANCFIS: "Fast adaptive neuro-complex fuzzy inference system", *International Journal of Approximate Reasoning*, vol. 105, pp. 417-430, 2019.

Ganeshsree Selvachandran is currently an Associate Professor in the Faculty of Business and Information Science at UCSI University, Kuala Lumpur, Malaysia. She holds a PhD (Pure Mathematics), MSc (Pure Mathematics) and a BSc. (Hons) (Mathematics) from the National University of Malaysia (UKM) and has nine years of teaching experience. She has published close to 50 research articles in various international journals and has presented many papers at national and international conferences. Her areas of interest include fuzzy computing, multi-criteria decision making (MCDM), soft computing, neutrosophic set theory, applications of fuzzy and soft computing in engineering, economics and pattern recognition, as well as fuzzy algebra and hyperalgebra

Shio Gai Quek is with the Department of Actuarial Science and Applied Statistics, Faculty of Business & Information Science, UCSI University, Malaysia. He holds a PhD (Pure Mathematics) and BEdu (Mathematics) from University Malaya (UM), Malaysia. He has 4 years of teaching experience and has thus far published several research articles in journals and conferences. His areas of interest include fuzzy algebraic and hyperalgebraic theory, fuzzy computing, soft computing and applications of fuzzy computing in engineering, decision-making and robotics.

Luong Thi Hong Lan received the Bachelor on Applied Mathematics and Informatics in 2002 and Master degree on Computer Science at Hanoi University of Technology in 2007. In 2017, she enrolled as PhD student in Institute of Information Technology (IoIT), Vietnam Academy of Science and Technology, Ha Noi, Viet Nam. Now, she is working at Faculty of Computer Science and Engineering, Thuyloi University. Her interests include Optimization, Machine learning, Data mining, Artificial Intelligence, Soft Computing, and Fuzzy Computing.

Le Hoang Son is a Associate Professor at the Department of Multimedia and Virtual Reality, VNU Information Technology Institute. His major fields include Artificial Intelligence, Data Mining, Soft Computing, Fuzzy Computing, Fuzzy Recommender Systems, Geographic Information System. He is an Associate Editor of Journal of Intelligent & Fuzzy Systems (SCIE), IEEE Access (SCIE), Data Technologies and Applications (SCIE), International Journal of Data Warehousing and Mining (SCIE), Neutrosophic Sets and Systems (ESCI), Vietnam Research and Development on Information and Communication Technology, VNU Journal of Science: Computer Science and Communication Engineering, Frontiers in Artificial Intelligence. Dr. Son serves as Editorial Board of Applied Soft Computing (SCIE), PLOS ONE (SCIE), International Journal of Web and Grid Services (SCIE), International Journal of Ambient Computing and Intelligence (ESCI), and Vietnam Journal of Computer Science and Cybernetics. He is the Guest Editor of several Special Issues at International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems (SCIE), and Journal of Ambient Intelligence and Humanized Computing.

Nguyen Long Giang obtained the phd degree on math fundamentals for informatics at institute of information technology (ioit), vietnam academy of science and technology (vast) in 2012. He has been promoted to associate professor in information technology since 2017. Dr. Giang worked as senior researcher and head of department at the ioit, vast during 2009 - 2019. His major field includes artificial intelligence, data mining, soft computing, fuzzy machine learning, rough sets, and fuzzy rough sets.

Weiping Ding (M'16-SM'19) received the Ph.D. degree in Computation Application, Nanjing University of Aeronautics and Astronautics (NUAA), Nanjing, China, in 2013. He was a Visiting Scholar at University of Lethbridge(UL), Alberta, Canada, in 2011. From 2014 to 2015, He is a Postdoctoral Researcher at the Brain Research Center, National Chiao Tung University (NCTU), Hsinchu, Taiwan. In 2016, He was a Visiting Scholar at National University of Singapore (NUS), Singapore. From 2017 to 2018, he was a Visiting Professor at University of Technology Sydney (UTS), Ultimo, NSW, Australia. His main research directions involve deep learning, data mining, evolutionary computing, granular computing, machine learning and big data analytics. He has published over 50 papers in flagship journals and conference proceedings as the first author.Dr. Ding currently serves on the Editorial Advisory Board of Knowledge-Based Systems, and Editorial Board of Information Fusion. He serves/served as an Associate Editor of several prestigious journals, including IEEE Transactions on Fuzzy Systems, Information Sciences, Swarm and Evolutionary Computation, IEEE Access and Journal of Intelligent & Fuzzy Systems.

Mohamed Abdel-Basset is an Associate Professor at Zagaig University, Department of Computer Science, Egypt. Received his B.Sc., M.Sc and the Ph.D in operations research from Faculty of Computers and Informatics, Zagazig University, Egypt. His current research interests are Optimization, Data Mining, Computational Intelligence, Robust Optimization, Engineering Optimization, Multi-objective Optimization, Swarm Intelligence, Evolutionary Algorithms, and Soft Computing. He is working on the application of multi-objective and robust meta-heuristic optimization techniques. He is also an/a Editor/reviewer in different international journals and conferences. He has published more than 170 articles in international journals and conference proceedings.

Victor Hugo C. de Albuquerque (M'17- SM'19) is a professor and senior researcher at the University of Fortaleza, UNIFOR, Brazil. He has a Ph.D in Mechanical Engineering from the Federal University of Para ba, an MSc in Teleinformatics Engineering from the Federal University of Cear á, and he graduated in Mechatronics Engineering at the Federal Center of Technological Education of Cear á. He is currently an Associate Pro-fessor of the Graduate Program in Applied Informatics of UNIFOR and leader of the Industrial Informatics, Electronics and Health Research Group (CNPq). He is a specialist, mainly, in IoT, Machine/Deep Learning, Patter Recognition, and Robotic.