CHAPTER 2 LITERATURE REVIEW

Fuzzy logic have a widespread application in industry, some of the common applications of fuzzy logic lies in the field of automobiles, consumer electronics, image processing, machine learning, and non-linear control systems to name a few. Due to this fuzzy logic has seen a lot of advancements since its introduction by Zadeh. These developments can be categorized as:

1. Theoretical developments: these improve the applicability of fuzzy systems and can be adapted for any application.
2. Application based: these are specific for particular applications and may not work effectively for other applications.

This chapter summarizes the key developments in fuzzy logic with reference to the proposed entropy based membership function optimization. The literature review comprises of two basic parts. First theoretical membership function developments are investigated, which do not include optimization techniques, but are based on development of new membership functions. Thereafter optimization based membership function developments are surveyed.

2. 1 FUZZY LOGIC BASED CONTROL SYSTEMS

Professor Mamdani developed the first fuzzy controller in 1974 [40]. The primary philosophy of a fuzzy logic based control system is to utilize the human operator experience by creating a controller which emulates behavior of a human expert. The current popular controller structures for fuzzy controller are discussed below. This categorization is based on the type of fuzzy rules utilized by the system.

A.) **Mamdani type controller** – In Mamdani type of FLC the first step is fuzzification of input crisp value into appropriate FS membership value. In second step results obtained by all the rules are summoned into a precise single value for output. The if-then rules used are defined using FS. Further the reshaping of the FS associated with fuzzy rule is done using a matching number, after aggregation of these FSs defuzzification is carried out [41], [42]. A simple scenario for a practical control system example can be described as:

IF error is positive AND rate of change of error is zero THEN controller output is small positive,

Here "error", "rate of change of error" are input variables and "controller output" is output variable. "positive", "zero", "small positive" are fuzzy sets.

If the input/output variables are represented as linguistic term x_1, x_2, y and the fuzzy sets are represented $\tilde{P}_1$, $\tilde{P}_2$, $\tilde{Q}$ respectively. A simple example of a Mamdani fuzzy rule can be described as:

If x_1 is $\tilde{P}_1^k$ and x_2 is $\tilde{P}_2^k$ THEN y^k is Q^k

Advantages of Mamdani controller are:

1. Applicable to almost every system.
2. Well suited for linguistic variables.
3. Works well for uncertain systems.
4. Works well for multiple input multiple output (MIMO) systems

B.) **Takagi-Sugeno type controller** – In T-S fuzzy controller the output of the rule base is a polynomial related to input variables, thus the output of each rule is a single number. For example a nonlinear continuous time system T-S fuzzy model is represented as [41] [42]:

$$\boldsymbol{x} = \sum_{j=1}^{r} \boldsymbol{\sigma_j} \left(\boldsymbol{A_j x} + \boldsymbol{B_j x} \right) \tag{2-1}$$

A typical rule for a T-S model can be expressed as:

IF x_1 is $\tilde{P}_1^k$ and x_2 is $\tilde{P}_2^k$ THEN z is $y^k = f(x_1, x_2)$

It is to be noted that the consequent function in T-S rule is a crisp function. Advantage of T-S controller is:

1. It is computationally efficient as it does not require defuzzification.
2. Applicable for linear techniques.

The controller used for the current research is Mamdani type, due to following key factors:

1. Systems to be controlled are non-linear.
2. TRMS is a MIMO system thereby necessitating use of only Mamdani controller.
3. Uncertainty in the systems encourages use of a Mamdani controller.

Fuzzy PD type controller - The simplest fuzzy PD type controller uses two input variables and one output variable. The scaling factor is employed for scaling of input and output variables. These inputs are generally: "error" and "rate of change of error". The output for this kind of FLC is generally the "control output. The input crisp values are fuzzified to FS; these fuzzy inputs are given to the fuzzy inference engine which processes the inputs according to the rule base and output is inferred. The fuzzy output from inference engine is then processed by defuzzifier to obtain a crisp output. Figure 2-1 depicts the architecture of a PD type FLC [3], [20].

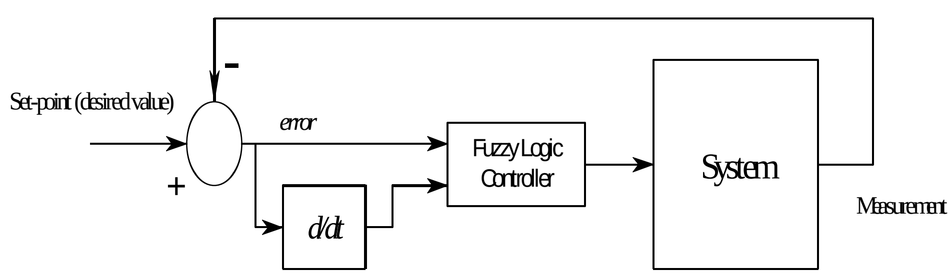


Figure 2-1 PD type fuzzy controller

Fuzzy PID type controller - Fuzzy PID type controller uses three input variables and one output variable. The scaling factor is employed for scaling of input and output variables. These inputs are generally: "error", accumulation of error", and "rate of change of error". The output is of this FLC is "control output". The input crisp values are fuzzified to FS; these fuzzy inputs are given to the fuzzy inference engine which processes the inputs according to the rule base and output is inferred. The fuzzy output from inference engine is then processed by defuzzifier to obtain a crisp output. Figure 2-2 depicts the architecture of a PID type FLC [3], [20].

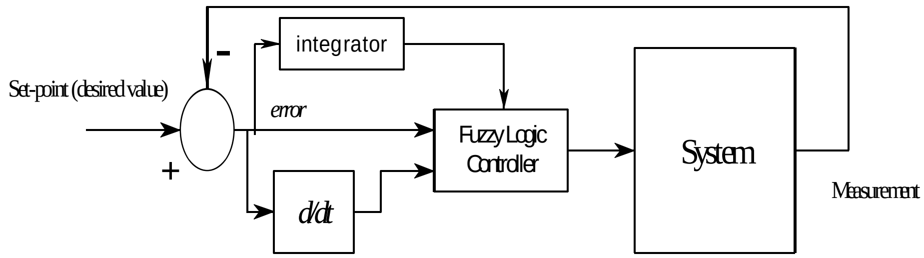


Figure 2-2 PID type fuzzy controller

2. 2 MEMBERSHIP FUNCTION

FSs can be simply defined as a set with a vague (ambiguous) boundary as compared to a crisp boundary of classical sets. For a discrete finite universe of discourse, x FS can be represented as follows [21]:

$$A = \left\{\frac{\mu_A(x_1)}{x_1} + \frac{\mu_a(x_2)}{x_2} + \ \ldots.\right\} = \left\{\sum_i \frac{\mu_a(x_i)}{x_i}\right\} \tag{2-2}$$

here, $\mu_A(\boldsymbol{x_1})$ is a value between 0 and 1.

When the universe of discourse, x, is continuous and ∞, FS is denoted by:

$$A = \left\{\int \frac{\mu_a(x)}{x}\right\} \tag{2-3}$$

In equations (2-2) and (2-3), the horizontal bar does not represent quotient but is rather a delimiter. The numerator represents the membership value of set A associated with the element x of the universe represented in the denominator. In equation (2-2), the summation symbol denotes the collection of each element and does not signify algebraic summation. Similarly in equation (2-3) the integral sign does not signify algebraic integral operator but is a continuous function-theoretic aggregation operator for continuous variables.

Figure 2-3 (a) illustrates a fuzzy set where a comparison between classical (crisp) set and a FS is depicted. In crisp sets there is no ambiguity of membership grade for an element and its membership value is either 0 (not an element of set) or 1 (an element which belongs to the set), whereas in a FS membership grade may lie anywhere between 0 and 1 (both inclusive). Here 1 is the highest membership value and 0 is the lowest membership value [20], [21]. Figure 2-3 (b) depicts a normal convex fuzzy set.

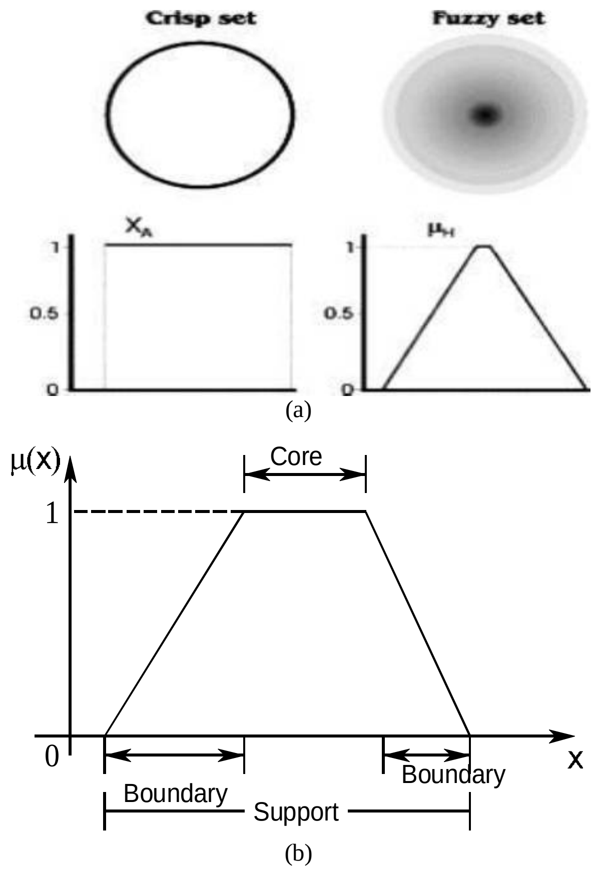


Figure 2-3 (a) Crisp set vs. FS (b) A typical FS with set parameters **[20]**

According to standard definitions, the parameters for MF are defined as [20]:

- Support – it is defined as the region of the universe having a non-zero membership in a given fuzzy set $\tilde{A}[\mu_{\tilde{A}}(x) > 0]$.
- Core – it is defined as the region of the universe which has a membership value of 1 in a given fuzzy set $\tilde{A}$ $[\mu_{\tilde{A}}(x) = 1]$.
- Boundary – it is defined as the region of universe which has a membership value between 0 and 1 in a given fuzzy set $\tilde{A}$ $[0 < \mu_{\tilde{A}}(x) < 1]$.
- Normal FS – it is a FS having a membership value of 1 for at least one element.
- Convex FS – it is a MF having membership values: (a) strictly monotonically increasing, or strictly monotonically decreasing, or (b) whose membership values are strictly monotonically increasing then strictly monotonically decreasing with increasing values for elements in the universe.

2. 2. 1 TYPES OF MEMBERSHIP FUNCTION

Some common mathematical functions which are used to MF are depicted below [20], [35], [43], [44], [45]:

Triangular MF – Considered as universal MFs for designing FLS. The fundamental attributes of triangular MFs which differentiate the triangular MF from others are:

- Here the boundary changes linearly from highest to lowest membership grade
- There's only 1 discrete element having highest membership grade.

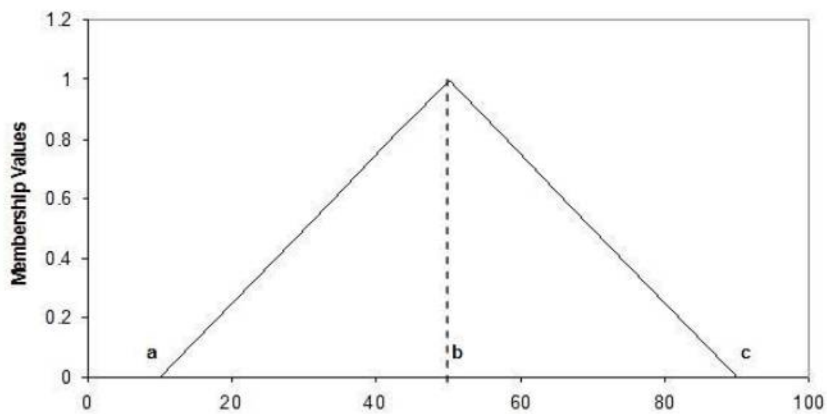


Figure 2-4 Triangular fuzzy MF [20]

A triangular MF function can be written as:

$$\mu_A(x) = \begin{cases} 0, & x \le a \\ (x-a)/(b-a), & a < x \le b \\ (x-c)/(b-c), & b < x < c \\ 0, & x \ge c \end{cases} \quad (2\text{-}4)$$

here points "a" and "c" denotes the FS support and "b" denotes discrete value with highest membership value. The FS used for the proposed optimization method is normal convex triangular MFs.

Trapezoidal MF – Like triangular, the trapezoidal also has a linear boundary for FS; however a trapezoidal MF is characterized by a range of elements having maximum membership grade. Trapezoidal MF is represented by:

$$\mu_A(x) = \left\{ \begin{array}{cc} 0, & x < a \\ (x-a)/(b-a), & a < x \le b \\ 1 & b < x < c \\ (d-x)/(d-c), & c \le x \le d \\ 0, & x \ge d \end{array} \right\} \quad (2\text{-}5)$$

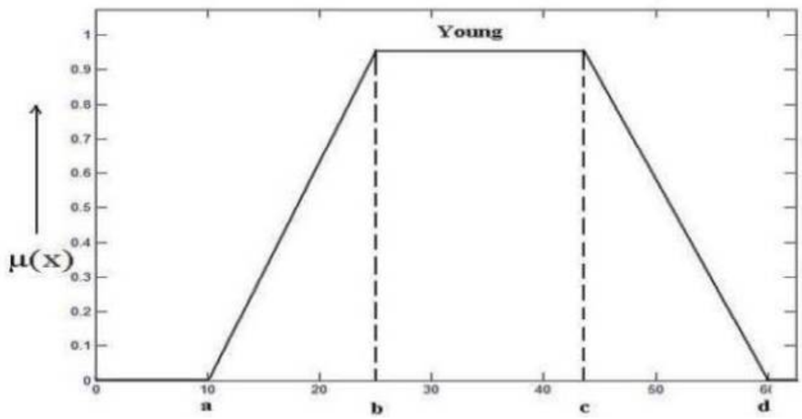


Figure 2-5 Trapezoidal MF **[20]**

Gaussian MF – Gaussian MFs are characterized by non-linear boundaries for MF variation & is defined as follows:

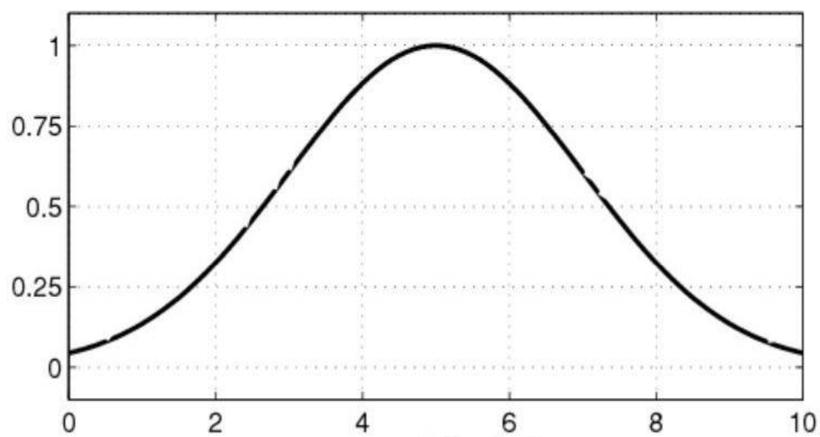


Figure 2-6 Gaussian MF [20]

$$\boldsymbol{\mu_A(x) = e^{-(x-c)^2/\sqrt{(2\sigma)}}} \qquad \textbf{(2-6)}$$

here, σ denotes the standard deviation and c is curve fitting constant, which governs the shape of MF.

Probability Density Function – This FS is estimated using probability density function for a variable. The MF function is depicted as:

$$\mu(x) = 1 - \frac{x}{a}, \; x \in [0, a] \qquad (2\text{-}7)$$

For a nonlinear function, the linearized form is given as

$$\mu(x) = \alpha\left(1 - e^{(b-x)(b-a)}\right), \; x \in [a, b] \qquad (2\text{-}8)$$

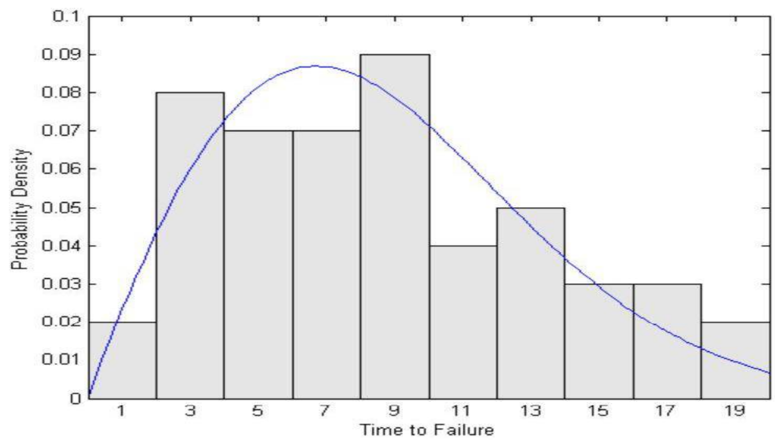


Figure 2-7 Probability density function based MF **[35]**

S function – These sets are attributed to its non-linear boundaries for MF variation. These are generally used for analytical approximation, and are depicted as:

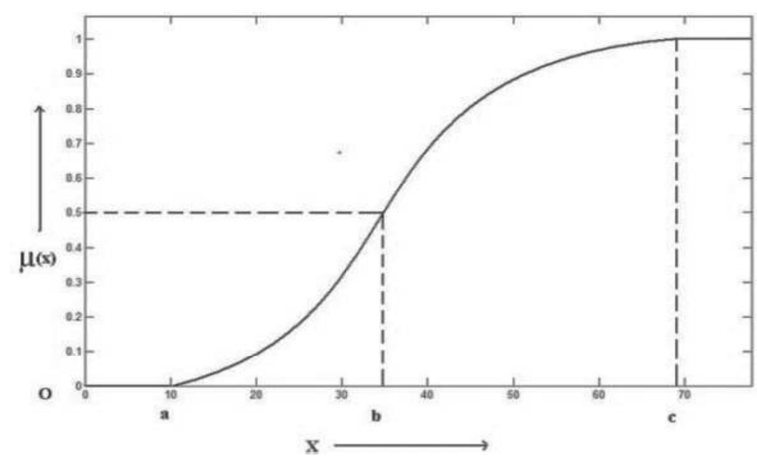


Figure 2-8 S-function fuzzy membership function [43]

$$S(x) = \begin{cases} 0, & x \leq a, \\ 2\left[\frac{(x-c)}{(c-a)}\right]^2 & a \leq x \leq b, \\ 1-2\left[\frac{(x-c)}{(c-a)}\right]^2, & b \leq x \leq c, \\ 1, & x \geq c, \end{cases} \quad (2\text{-}9)$$

here a, and c are parameters which are used to change shape the MF curve & b is the mid-point.

2. 2. 2 MEMBERSHIP FUNCTION DEVELOPMENT

Fuzziness of a system can be depicted in various ways; similarly, there are numerous ways to graphically depict the MFs that describe this fuzziness. Few such procedures to develop MF are described here. Since MF embodies all fuzziness for a particular FS, its representation is the extract of fuzzy property or operation. Intuition includes semantic and contextual knowledge about an issue; it may also involve linguistic truth values about this knowledge. The most imperative factors involve approximate placement of the curve on variable's universe of discourse, number of curves used, and the overlapping character. Inference employs the knowledge to perform deductive reasoning. Rank of ordering assess preferences by a single individual, a committee, a poll, and other opinion methods are used to assign the membership values to a fuzzy variable [20]. Primitive methods for MF estimation were reliant on statistical methods rather than optimization. Some milestone literatures with reference to design problem employed for proposed research are:

Devi and Sarma [46] proposed a technique to estimate MF form statistically defined probability density function (pdf) of data which is obtained from histogram generated from finite no. of samples. probability mass function is depicted by:

$$f(x) = \frac{1}{2Nh}\sum_{j=1}^{N}\varphi_x(x_j) \tag{2-10}$$

Where,

$$\varphi_x(x_j) = \begin{Bmatrix} 1, & x_j \in [x-h, x+h] \\ 0, & otherwise \end{Bmatrix} \tag{2-11}$$

The pdf can be approximated as:

$$\hat{f}(x) = \frac{M_l(x)}{N_p(x)} = \frac{P(y)}{Q(y)} = \frac{p_0 + p_1 y^1 + \cdots \ldots + p_l y^l}{q_0 + q_1 y^1 + \cdots \ldots + q_p y^p} \tag{2-12}$$

Now, the fuzzy membership is formulated by:

$$\mu(x) = \frac{\hat{f}(x)}{\hat{f}(a)} \tag{2-13}$$

Where,

$$\hat{f}(a) = \max_x\{\hat{f}(x)\} \tag{2-14}$$

In this method "arithmetic mean" is employed to compute the shape of MF. The "standard deviation" is utilized to establish the support for FS. The formulated MF obtained were deployed for identification of Fisher Iris data [47] with a yield of 83.33% to 97.33% accuracy. The proposed technique is an explicit motivation for current research work. As it directly utilizes the standard deviation for estimating the support of a FS.

Civanlar and Trussell [35] proposed a method to obtain FSs using pdf and Zadeh's possibility-probability consistency principle [48]. Civanlar formulated the consistency principle for real line sets as:

$$\max_{x \in D} \mu(x) / max\mu(x) \geq \int_D p(x)dx \tag{2-15}$$

Consistency principle states that degree of possibility for any event is always greater than or equal to the degree ofprobability, the subsequent MF is evaluated to satisfy possibility probability principle. The developed MF is estimated to be equitable and is extended to various practical applications. The proposed method was employed to eliminate noise occurring in a set of linear equations. The method utilized probability density function (pdf) obtained from the statistical data. The optimal MF is depicted as:

$$\mu(x) = \begin{cases} \alpha p(x) & if\ \alpha p(x) < 1 \\ 1 & if\ \alpha p(x) \geq 1 \end{cases} \tag{2-16}$$

Where p(x) is pdf derived from the histogram of used feature.

Optimal function is calculated by solving:

$$\boldsymbol{minimize\ f}(\mu) = \mathbf{1/2} \int_{-\infty}^{\infty} \mu^2(\boldsymbol{x})\boldsymbol{dx} \tag{2-17}$$

$$\begin{gathered} such\ that \quad G(\mu) = c - E\{\mu\} = c - \int_{-\infty}^{\infty} \mu(x)p(x)dx \leq 0, \\ and\ \mu \in \Omega = \{\mu(x), 0 \leq \mu(x) \leq 1\}, c < 1 \end{gathered} \tag{2-18}$$

The Lagrange expression for the above function can be depicted as:

$$\boldsymbol{L}(\mu, \boldsymbol{\alpha}) = \mathbf{1/2} \int_{-\infty}^{\infty} \mu^2(\boldsymbol{x})\, \boldsymbol{dx} + \boldsymbol{\alpha}\{\boldsymbol{c} - \int_{-\infty}^{\infty} \boldsymbol{\mu}(\boldsymbol{x})\boldsymbol{p}(\boldsymbol{x})\boldsymbol{dx}\} \tag{2-19}$$

Where α is taken as Lagrange multiplier [∝≥ 0]. Remark: this method does not employ any optimization technique nor is entropy calculated for obtained FSs.

FS are obtained using statistical "confidence levels" and are illustrated as:

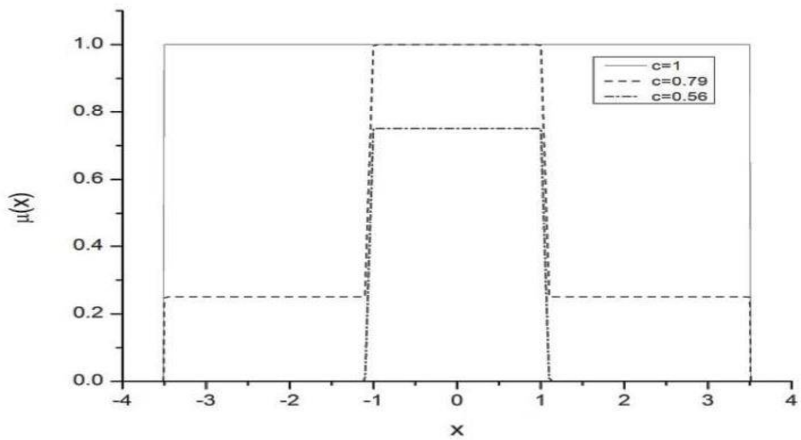


Figure 2-9 MFs obtained for different confidence levels [35]

Dombi [44] compared various MFs based on their mathematical functions. These chosen MFs were based on various applications suggested by the author. The comparison was focused on following parameters:

- Support for membership function, [a,b].
- The sharpness coefficient, λ.
- The decision level, ν.

Dombi [44] classified the MFs into following categories:

- Heuristically based MFs.
- MFs based on reliability parameters.
- MFs based on theoretical decision making.
- MFs specifically utilized for designing control systems.
- Linguistic models based MFs.

A thorough analysis of the MFs yields some common propertied amongst various MFs:

- All MFs were continuous and linear/piecewise linear (linearized if the intrinsic mathematical function is non-linear).

- The MFs map a crisp interval of [a,b] of a crisp set to a FS $\mu[a,b] \rightarrow [0,1]$.
- MFs are either monotonically increasing or monotonically decreasing or a combination of both.

Combining all the MFs a new MF proposed by Dombi consists of two parts, a monotonically increasing:

$$\mu(x) = \frac{(1-\nu)^{\lambda-1}(x-a)^{\lambda}}{(1-\nu)^{\lambda-1}(x-a)^{\lambda} + \nu^{\lambda-1}(b-x)^{\lambda}} \quad (2\text{-}20)$$

And a monotonically decreasing:

$$\mu(x) = \frac{(1-\nu)^{\lambda-1}(b-x)^{\lambda}}{(1-\nu)^{\lambda-1}(b-x)^{\lambda} + \nu^{\lambda-1}(x-a)^{\lambda}} \quad (2\text{-}21)$$

Where $x \in [a,b]$.

For linear case, $\lambda = 1$, we obtain:

$$\mu(x) = \frac{x-a}{b-a} \quad (2\text{-}22)$$

Chen and Otto [49] proposed an interpolation based curve fitting method to estimate the MF from measurement data. The interpolation method employs Bernstein polynomial:

$$B(g)(x) = \frac{g(x_i)(t-x_i)^2 + 2b(x-x_i)(t_i-x) + g(t_i)(x-t_i)^2}{(t_i-x_i)^2} \quad (2\text{-}23)$$

FS obtained are then utilized to assess the stress relations for computer-aided design engineering application. Results are compared for (a) 3 points, (b) 6 points, and (c) 9 points measurement samples.

Hong, Lee [50] in their paper proposed a proficient method for learning through an algorithm based on training examples. This approach is employed to develop an expert system. System knowledge, predefined MF (triangular type) library are predetermined in the knowledge base and desired MF and decision rules are obtained using learning algorithm. The first stage of the learning algorithm is searching for relevant attributes within the training data. Step-by-step procedure for this is:

- The first step is to sort various attribute values in increasing order.
- For all attributes determine the no. of instances which belong to the same class.
- Evaluation of the fitness degree of each attribute, using: $f_i = t_i/n$, where n is the total number of training instances and t is the ith training instance.
- Now sort all attributes in ascending order, according to the fitness degree calculated in above steps.
- Finally, choose the most relevant attribute.

Next stage is to determine MF. Step-by-step procedure for this method is:

- Assign initially a default group to each selected attribute, $G = 1 + 3.3logn$
- Determine range of each attribute, $R_i = \max(A_i') - \text{main}(A_i')$.
- Determine the group interval for each attribute, $H_i = R_i/(G - 1)$.
- Extend the possible minimum attribute value of attribute A_i' to$V_i = \min(A_i') - H_i/2$.
- Divide the range for each of the attribute into G groups.
- Find point $b_{ij}$for each initial MF, $b_{ij} = \sum_{s=1}^{r_{ij}} A_{ij}'(I_{ijs})/r_{ij}$, where $A_{ij}'$is the jth interval for A_i', and r_{ij} is the total no. of instances in A_{ij}'.
- Finally, determine the points a_{ij} and c_{ij} using, $a_{ij} = b_{i(j-1)}$and $c_{ij} = b_{i(j+1)}$.

The proposed method is utilized for identification of Fishers Iris data problem [47] and the results attained 100% accuracy for Setosa, 94% for versicolour and 92.72% for Virgininca, hence the proposed technique gave good performance and rational result. The overall classification efficiency achieved is 96.67%.

Tamaki et. al. [51] proposed a pdf based fuzzy observation constrained method to determine MF. For any function $f(x)$ if pdf is determined by:

$$P_i = \int_{-\infty}^{\infty} \chi_i(x) f(x) dx \qquad (2\text{-}24)$$

Assuming a total of "t" FSs, $S_1, S_2, \ldots\ldots S_t$, the following constraint holds on:

$$\forall x \in S_i \cap S_{i+1};\ \chi_i(x) + \chi_{i+1}(x) = 1 \qquad (2\text{-}25)$$

In this method, the FS domain is partitioned into several subsets consisting of "t" FSs. Required FSs are then chosen from predefined sets is governed by their pdf. The MF assortment criterion is based on the "discrimination degree", and is defined as:

$$\xi_{i-1,i} = \frac{\int_{\theta_1} f(x) dx}{P_{(i-1)a}} \qquad (2\text{-}26)$$

where, $\theta_i = \{x | \chi_{(i-1)b}\}$

Hence the conditional probability of selecting S_i becomes:

$$\chi_i(x) = Prob\{S_i\ selected | x\} \qquad (2\text{-}27)$$

Medasani et. al. [52] provided an overview of several methods used for determination of MF particularly for pattern recognition application. All the techniques for MF generation were covered which include; heuristics, the probability to possibility transformations, histograms, nearest neighbor techniques, feed-forward neural networks, clustering, and mixture decomposition. Following MFs are discussed:

- Triangular
- Piecewise linear
- Gaussian
- S-function
- Monotonic functions
- π-functions
- pdf based methods
- Neural network based methods
- Clustering based methods

Concluding the comparison author states there's not a best method, and every technique has its relevant benefit based on application.

Medaglia et al [43] introduced a method based on Bézier curves [53] to fit a MF from given data points. This method allowed the system designer to regulate the shape of FS depending on the selection of control points. The curve passing from the vicinity of these control points is represented by mathematical formulation proposed by Bezier, and MF for this method is depicted as:

$$\mu_A(x(t)) = \begin{cases} 0, & x(t) < m_L - \gamma, \\ \mu_{A_L}(x(t)), & m_L - \gamma \le x(t) \le m_L, \\ 1, & m_L < x(t) < m_R, \\ \mu_{A_R}(x(t)), & m_R \le x(t) \le m_R + \beta, \\ 0, & x(t) > m_R + \beta, \end{cases} \tag{2-28}$$

where γ and β are crossover points and are placed on the LHS and RHS of MF respectively; $m_L, m_R \in X$ are values of the lowest and the highest membership respectively. $\mu_{A_L}(x(t))$ *and* $\mu_{A_R}(x(t))$ are left and right membership values. Further in order to compute the MF the collection of membership value required is calculated through data driven estimation method, where the number of control points along with their location is now determined. The sum of the squared error between data and MF on left and right sided MF is minimized using the the mathematical formulation:

$$min \sum_{i=2}^{M_R-1} \left(y_{R,i} - \sum_{k=0}^{n_R} \binom{n_R}{k} t_i^k (1-t_i)^{(n_R-k)} y_{R,k} \right)^2 \tag{2-29}$$

$$\text{Subject to } \sum_{k=0}^{n_R} \binom{n_R}{k} t_i^k (1-t_i)^{n_R-k} \chi_{R,k} = \check{\chi}_{R,i} \tag{2-30}$$

where M_R is number of data control points, $y_{R,i}$ is the membership determined by a direct approach, n_R is illustrated as the degree of the polynomial on RHS. M_L is no. of data points, $y_{L,i}$ is MF determined from the direct approach, n_L is degree of a polynomial on LHS. The proposed technique is used to produce the MF of any imprecise concept. Through properties of Bezier curve, it is noted that the variation is significant near the control points and MF deviates completely. Figure 2-10 illustrates the change in the MF when a single control point

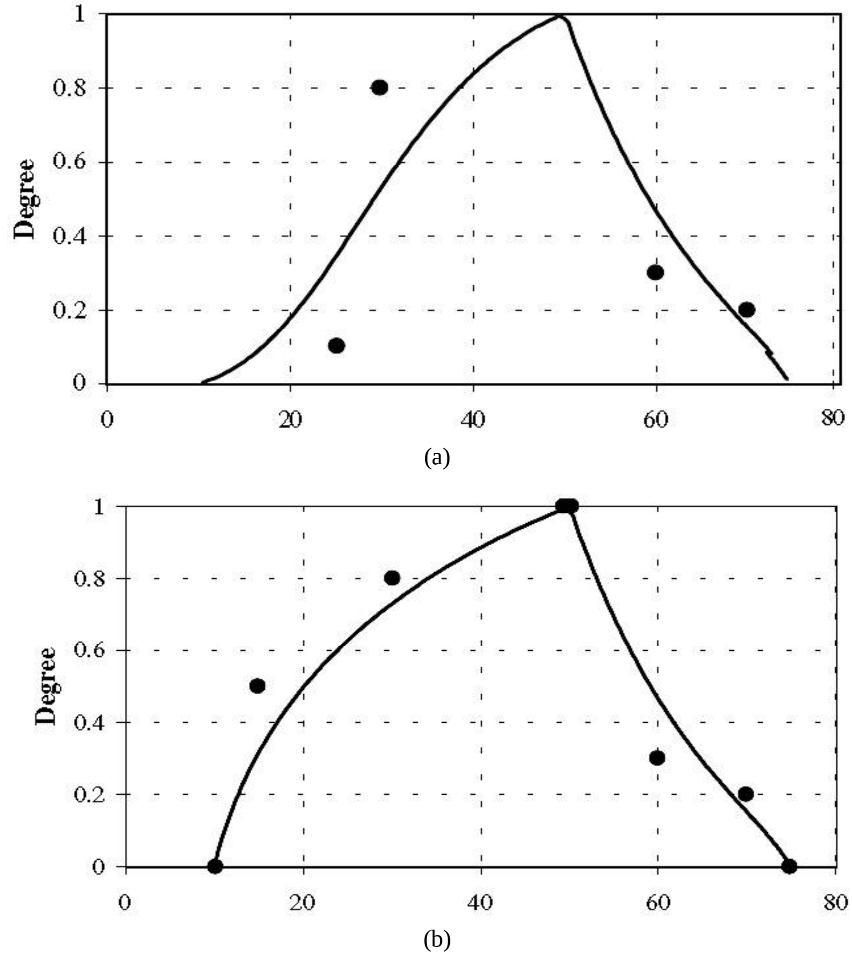


Figure 2-10 Effect of changing single control point. (a) before change (b) after change **[43]**.

In the area of pattern recognition, automatic generation of fuzzy MF is required. Yang, Bose [54] proposed a method of generating MF based on unsupervised learning through self-organizing feature map (SOFM) employing feed forward neural network. The robustness of developed system was tested for identification of Fisher's Iris data. MF using SOFM is generated from two-step method, the first step is clustering generation, with every feature input vector input weighted neurons are associated. the feature input vectors have a task of finding a winning neuron, The combination of two input features generate a new input feature for SOFM and SOFM becomes MF generator. However in retrieving stage, only one input feature is present, and after finding for winning neuron the o/p from SOFM is subvector of weight with associated labeling information. The winning neuron has information of both learning and retrieves phase which can be retrieved. The testing carried on the iris data sheet found to be applicable for multidimensional input feature with labeled information. The results showed MF to be consistent with the pattern distribution in free space. In proposed method neurons in o/p layer of SOFM are inadequate but the amount of rules in a rule-based system is reduced. This however, does not disturbs the location and form of MF by the occurrence of elements. MF is estimated as the function of the object and for universe of discourse

McCloskey et. al. [55] proposed a dynamic method to obtain MFs using data analysis method. Proposed method is employed for financial planning and decision assisting for corporate merger process. This method employs user-scalable triangular MFs.

$$\mu_A(x) = \begin{cases} 0, & x \le a \\ (x-a)/(b-a), & a < x \le b \\ (x-c)/(b-c), & b < x < c \\ 0, & x \ge c \end{cases} \quad (2\text{-}31)$$

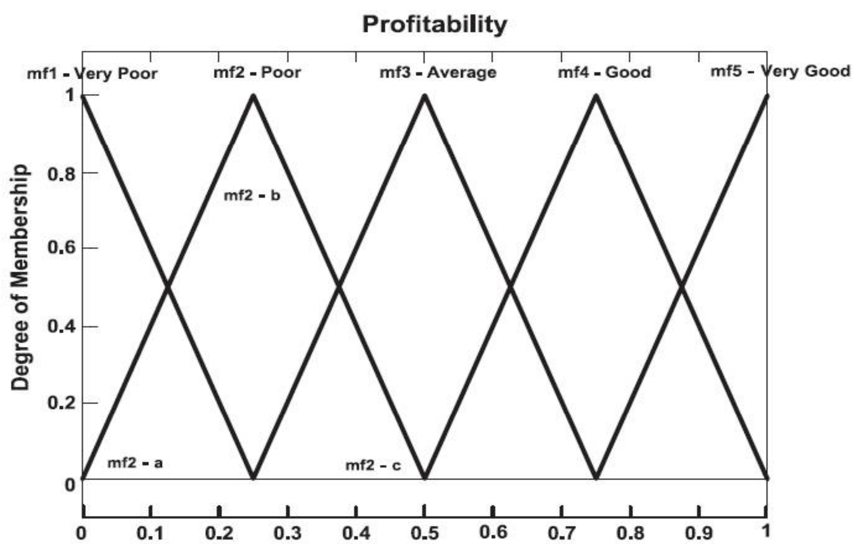


Figure 2-11 Predefined MFs **[55]**

Figure 2-11shows predefined MFs used for scaling purpose. Parameter b for MF 1 and MF 5 are set as 0 and 1, respectively. And for MFs 2, 3, and4 b is defined as:

$$mf2(b) = \frac{\sum_{i=1}^{n} X_i}{2(n)} \quad (2\text{-}32)$$

$$mf3(b) = \frac{\sum_{i=1}^{n} X_i}{n} \quad (2\text{-}33)$$

$$mf4(b) = \left(\frac{1 - \frac{\sum_{i=1}^{n} X_i}{n}}{2}\right) + \frac{\sum_{i=1}^{n} X_i}{n} \quad (2\text{-}34)$$

here, n = number of inputs for X.

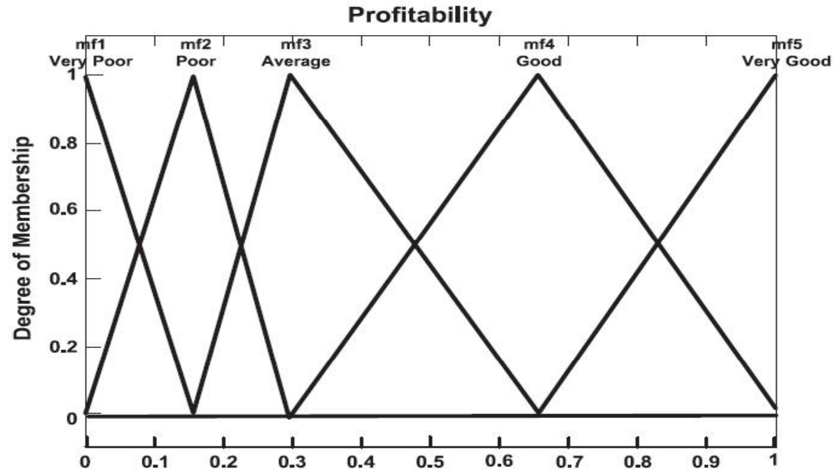


Figure 2-12 optimized MFs **[55]**

The proposed method is employed to predict acquisition decision for 50 corporate entities. The results obtained exhibit 100% correspondence when compared to financial expert's panel decision. However, a major limitation of this technique is requirement of a huge data set for MF determination. This method, however, doesn't employ optimization technique nor is entropy calculated for the obtained FSs.

Zhou, Khotanzad [56] proposed a genetic algorithm (GA) based method to determine the: fuzzy MFs, size, and structure for rule base using training data. Consider input vector to be X, and number of MFs required are predetermined. Let the jth MF be a Gaussian MF depicted by$[\mu_j, \sigma_j]$. Let $[R_{mini}, R_{maxi}]$ denote the minimum and maximum values of x_i and the training data be expanded by both sides by 20%. The real valued $\mu_j$can be determined by:

$$\mu_j = \mu_{defaultj} + (M_j - 128)/(128 \times delta) \quad (2\text{-}35)$$

$$\text{Where } delta = (R_{maxi} - R_{mini)})/(\text{n} + 1) \quad (2\text{-}36)$$

$$\text{And } \mu_{defaultj} = R_{mini} + \text{delta} \times \text{j}, 1 < \text{j} \leq \text{n} \quad (2\text{-}37)$$

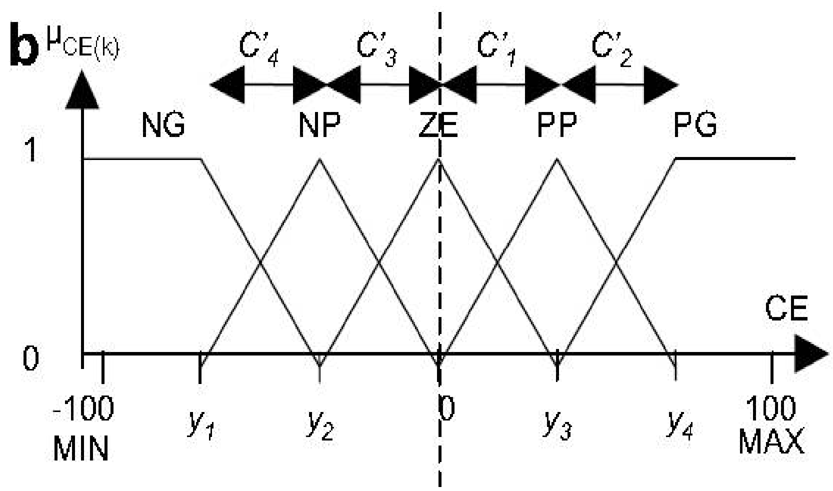
Hence the possible range covered by x_i is divided equally among n MFs and the location of mean for each MF is than rearranged depending on the computed M_j value. Variance conversion is obtained through:

$$\sigma_j = c_1 \times V_j + c_2 \qquad (2\text{-}38)$$

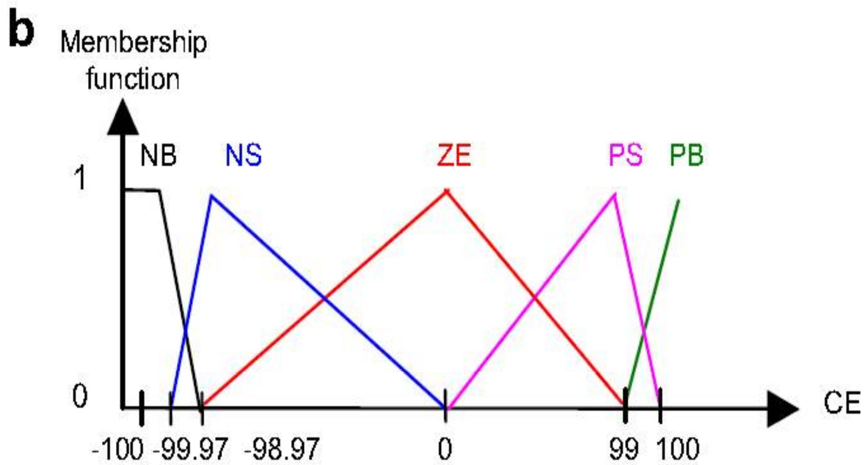
Where c_1, c_2 are constants selected such that a V_j value of 64 covers 80% of *delta*.

The proposed method is employed to classify patterns from 178 samples based on Fishers Iris data [47]. Results acquired exhibit only 2 misclassified samples. However, the main success of the proposed technique is a reduction in number of rules. The major limitation of the proposed technique is the limited application to pattern recognition problems.

Larbes et. al. [57] proposed a GA optimization based FLC for tracking maximum power point in PV system. Utilizing model-free approach exhibited by FLC the designer has implemented FLC using knowledge available from experts. FLC performance is further improved with help of GA based online optimization. Initially normalized FSs are used which are uniformly allocated thru the set's domain. Designed FLC is then implemented for controlling the system. Online optimization is done to modify the predesigned MFs.



(a)



(b)

Figure 2-13 Membership function (a) before optimization, (b) after optimization [57]

The objective function for online optimization is: minimization of ISE:

$$J = \int e^2 dt \qquad (2\text{-}39)$$

Where, $e = P_{max} - P$ (error). (Maximum power obtained from PV cell − desired power). Major limitation of the online optimization method is the mandate for a continual operating system to estimate the optimization results.

Nohe et. al. [58] proposed a GA based optimization method for the o/p regulation for a servomechanism with backlash. The optimization was proposed for type-1 and type-2 FLS. Optimization methodology followed is:

- Designing the system layout for FLS: I/O selection, MFs to be used, initialize rule base.
- Identify the parameters which are manipulated for obtaining optimized MF.
- Initialize GA chromosome by sorting parameters of each variable respectively.
- Define and evaluate the optimization objective function.
- Select the various GA operations and desired optimization parameters.
- Obtain optimized MF by executing GA and interchange the previous MF parameters with new optimized parameters.

The proposed research uses triangular MF for type-1 and tyoe-2 FLS. The objective function for optimization is:

$$fitness_j = \min(mean|error|) \qquad (2\text{-}40)$$

The results obtained indicate faster settling times and fewer error indices as compared with non-optimized system performance. The major drawback of proposed method is it's long execution time (350 hours) for achieving optimization results. Being an online optimization method a limitation exhibited by the system is that it has to be working continually to achieve optimization results.

Acilar and Arslan [59] proposed a clonal selection algorithm (CLONALG) for optimization triangular MF support. Proposed CLONALG is inspired from the immune system adaptability for various antigenic stimulus. The algorithm is inspired from biological phenomena: the cells which are capable of identifying an antigenic stimulus that thrives and separates into effectors cells. The main features of this algorithm are affinity proportional reproduction and mutation. The higher affinity, the higher is the no. of generated offsprings. The mutation is writhed by each immune cell during reproduction and is inversely proportional to the affinity of cell receptor to that of the antigen. The method comprises of following steps [60]:

1. Generate a population os *n* antibodies randomly.
2. Repeat the following *m* times.
    a. Compute affinity (objective function) of each antibody.
    b. Select *j* highest affinity antibodies.
    c. Selected *j* antibodies are cloned according to their affinity function, thereby producing *C* clones.
    d. Clones are than subjected to hypermutation process, where higher affinity means a lower hypermutation & vice versa.
    e. Compute the affinity for generated clones *C*.
    f. From this generation of clones and antibodies, select *n* highest affinity cells to generate a new population.
    g. Replace the *l* lowest affinity population by new generation generated randomly.
3. Repeat till *m*.

The affinity function defined in the proposed method is:

$$Affinity = maxerror - totalerror \tag{2-41}$$

Here,

$$totalerror = \sum_{i=1}^{j} (y_{CLONALGi} - y_i)^2 \tag{2-42}$$

$$maxerror = \sum_{i=1}^{j} (y_i - y_{max})^2 \tag{2-43}$$

Where, y_i is the ith reference i/p of o/p. $y_{CLONALGi}$ is the o/p obtained by CLONALG & is defines as o /p for ith reference & n is number of i/o data. Performance of this method is compared to GA & BPSO algorithms. Results indicate faster convergence for CLONALG as compared to BPSO or GA. Being an online optimization method a limitation exhibited is: system has to be in continuous operation to estimate the results.

Ang and Quek [61] proposed a bio-inspired algorithm based method for generation of MFs. Supervised Pseudo Self-Evolving Cerebellar (SPSEC) algorithm is motivated from the development method of human nervous system whereby the basic architecture is first laid out without any activity-dependent processes and is further refined based on activity-dependent ways. The SPSEC algorithm is then used to produce MF by reconciling with semantics of human interpretable linguistic terms. The generated MFs align with the human-interpretable linguistic variables hence no further optimization is required [62]. Authors tested the MF generated using this method for conformity of decision boundaries and results indicate. Decision boundaries for generated MF were found to be consistent with the inherent probabilistic decision boundaries corresponding to training data set.

Kao et. al. [63] proposed a hybrid density based clustering method for MF generation in fuzzy controller for a ball mill pulverizing system. Proposed method is two stage MF generation process where it first employs shared nearest neighbour similarity is based on fuzzy weighted distance considered as the distance metric between objects. Secondly, a density based spatial clustering of applications with noise (DBSCAN) like clustering process & the divisive clustering procedure is employed to detect the clusters. Finally, the clusters are projected in the variables domain in order to determine the MFs. The fuzzy controller developed from the generated MF is employed in real-time of ball-pulverizing process control and the results confirm the adeptness of the proposed technique.

2. 3 Entropy function and its significance

Shannon laid the concepts of communication systems by formulating the fundamentals for information theory [64] [65]. The uncertainty associated to random variables is measured by entropy function or simply “entropy determines the expected value of information available in a corresponding message”. However fuzzy entropy is defined as a measure of the fuzziness in a given FS or system as degree of variable randomness, & uncertainty associated is used as information measure. FLS are amongst systems which employs Shannon’s findings to various fields like: fuzzy logic based clustering, decision making, optimization techniques to name a few. Fuzzy entropy is defined as the measure of fuzzy information available obtained through a FS.

$$H(A) = -\int_{-\infty}^{\infty}\{\mu_i \log \mu_i + (1-\mu_i)\log(1-\mu_i)\} \qquad (2\text{-}44)$$

Here $\mu_i$is membership value of element i in a FS A. This expression is also written as:

$$H(A) = \int_{-\infty}^{\infty} f\big(\mu(x)\big)dx \qquad (2\text{-}45)$$

Here $f(x) = -xlnx - (1-x)\ln(1-x)$

2. 3. 1 Optimization of MF using Fuzzy Entropy

Being a measure of fuzzy information encompassed in a FS, computation of entropy is carried to examine applicability of given fuzzy system. Determination of MF is an important facet for designing a fuzzy system, numerous entropy optimization based techniques have been proposed for MF estimation. Cheng and Chen [66] used the notion of maximum entropy for estimating MF. The set defined represent the membership value for brightness of gray level in a given image. The brightness function is given by:

$$bright = \sum_{r_k \epsilon \Omega} \mu_{bright}(r_k)/r_k \qquad (2\text{-}46)$$

The MF μ_{bright} express the brightness of a gray level r_k.. $\Omega = \{r_0, r_1, \ldots r_{l-1}\}$.
Probability of this particular fuzzy event is given as:

$$P(bright) = \sum_{k=0}^{L-1} \mu_{bright}(r_k) P(r_k) \qquad (2\text{-}47)$$

Here P is probability measure for incident of gray level in an image.
MF for brightness of gray level is expressed by following s-function:

$$S(x) = \begin{cases} 0, & x \le a, \\ \dfrac{(x-a)^2}{(b-a)(c-a)}, & a \le x \le b, \\ 1 - \dfrac{(x-c)^2}{(c-b)(c-a)}, & b \le x \le c, \\ 1, & x \ge c, \end{cases} \qquad (2\text{-}48)$$

where x is independent variable, b as the center of MF, a and c as crossover points. The range of fuzzy event is regulated by a and c, parameter b defines the increasing rate, that widely deviates the shape of s-function.
Since brightness for gray levels of any image is contextual dependent, obtaining optimized value for parameters is often challenging. For choosing s-function, parameters defining the brightness are formulated as:

$$(bright; a_{opt}, b_{opt}. c_{opt}) = \max\{H(bright; a, b, c) | r_0 \le a < b < c \le r_{l-1.}\} \qquad (2\text{-}49)$$

To obtain the optimized solution for brightness function formulated in equation (2-49) simulated annealing algorithm is used. Experimental results yield a rapid decrease in iterations required to reach the optimum result. Simulated annealing algorithm converged to a solution in about 300 iterations, which is 1/900th number of iterations undertook by exhaustive search algorithm.

Hasuike et al [67] proposed an optimization method to integrate fuzzy entropy and pdf to acquire optimal FSs. This method is also compared to modified S-curve based mf optimization and obtained results were verified. But due to complexity of optimization indices the proposed method was computationally intensive as found out by the author. The technique for MF generation based on fuzzy Shannon entropy and human's interval estimation integrating pdf with fuzzy Shannon entropy. Proposed optimization is based on modified s-function formulated by Peidro et. al. [68]. The improved s-curve functions are inspired from human cognitive behavior thereby creates a MF which fits close to human's subjectivity. The membership function (left side) is depicted in Figure 2-14 for which the mathematical equation is:

$$S_l(x) = \begin{cases} 1, & x^b < x \\ w^b, & x^b = x \\ 1 - \dfrac{B_1}{1 + C_1 e^{\alpha_1 x}}, & x^a < x < x^b \\ w^a, & x^a = x \\ 0 & x < x^a \end{cases} \tag{2-50}$$

Here $B_1, C_1, and \ \propto_1 > 0$ are the parameters which determine the shape of the s-curve. w^a, w^b are the constants which varies form 1 and 0 & are set decided by the expert.

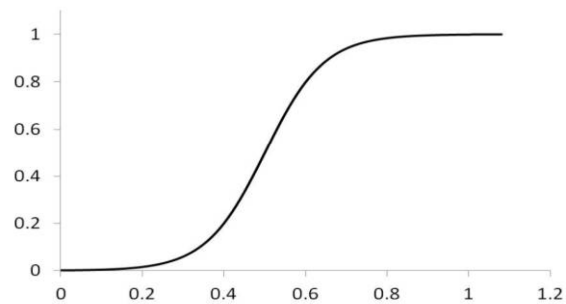


Figure 2-14 left-hand s-curve membership function

Similarly, right side of s-curve membership function is depicted by:

$$S_r(x) = \begin{cases} 1, & x < x^c \\ w^c, & x^c = x \\ \dfrac{B_2}{1 + C_2 e^{\alpha_2 x}}, & x^c < x < x^d \\ w^d, & x^d = x \\ 0 \ , & x^d < x \end{cases} \tag{2-51}$$

Now fuzzy Shannon entropy under s-curve function is maximized, with pdf as optimization constraint. The objective function is defined as:

$$\begin{aligned} &maximize = -\int_{-\infty}^{\infty} \{\mu(x) log\mu(x) + (1 - \mu(x))\log(1 - \mu(x))\} dx \\ &Subject\ to\ E(\mu(x)) = \int_{-\infty}^{\infty} \mu(x) p(x) dx \geq m, \quad 0 \leq \mu(x) \leq 1, \forall x \end{aligned} \tag{2-52}$$

Where m is the initial avg. membership value and *p(x)* is pdf.

Optimization examples include uniformly distributed pdfs and partially uniform distributed pdfs generating appropriate MFs. However a disadvantage for this technique is its large convergence time for obtaining optimization results, also for complicated pdfs the optimization result is tough to obtain. To overcome these computational challenges Hasuike, et al [69] proposed an algorithm based on piecewise linear functions utilizing Lagrange functions. Using least squares method equation (2-52) can be rewritten as:

$$\begin{aligned} &Minimize \quad \sum_{i=1}^{n+1}\{\mu_i log\mu_i + (1-\mu_i)\log(1-\mu_i)\} \\ &\qquad\qquad + W\sum_{i=2}^{n}(\mu_{i+1} - 2\mu_i + \mu_{i-1})^2 \qquad (2\text{-}53) \\ &subject\ to \quad \sum_{i=1}^{n+1}\mu_i p(x_i) \geq c, \quad 0 \leq \mu_i \leq 1, \ (i = 1,2 \ldots . n) \end{aligned}$$

here W is: weight defined for smoothing of the membership function.

The Lagrange for this problem is:

$$\begin{aligned} L = &\sum_{i=1}^{n+1}\{\mu_i log\mu_i + (1-\mu_i)\log(1-\mu_i)\} + W\sum_{i=2}^{n}(\mu_{i+1} - 2\mu_i + \mu_{i-1})^2 \\ &+ \lambda\left(c - \sum_{i=1}^{n+1}\mu_i p(x_i)\right) + \sum_{i=2}^{n-1}\mu_i(\mu_i - 1) - \sum_{i=2}^{n-1}\vartheta_i\mu_i \end{aligned} \qquad (2\text{-}54)$$

The smoothing function further minimize the objective function as

$$\begin{aligned} &Minimize \quad \sum_{i=1}^{n+1}\{\mu_i log\mu_i + (1-\mu_i)\log(1-\mu_i)\} + W\sum_{i=1}^{n}(\mu_i - (i-1)/n)^2 \\ &subject\ to \quad \sum_{i=1}^{n+1}\mu_i p(x_i) \geq c, \quad 0 \leq \mu_i \leq 1, \ (i = 1,2 \ldots . n+1) \end{aligned} \qquad (2\text{-}55)$$

The optimization for proposed objective function is computed using mathematical programming. The smoothing function however reduces the computational requirement of the optimization problem, but comes at the loss in generality of the defined function.

2. 3. 2 Research gap

The literature discussed above focuses on the developemnt of membeship function & various optimization processes. It can be concluded from the literature, that majority of exploration in fuzzy memebrship function optimization is focused towards obtaining optimal mathematical function (shape) of fuzzy set and not regarding the calculation of the optimal support for a individual fuzzy sets for a fuzzy variable. Numerous optimized mathematical functions have been propsoed by overlooking the support of FSs. The proposed research addresses this gap and formulates an optimizing algorithm which computes the optimized support for each FS in a given fuzzy variable. This is achieved by using fuzzy entropy as evaluation parameter to determine the optimized support. Scetion 2. 4 illustrates the detailed algorithm for proposed optimization technique.

2. 4 Objective function for optimization

The earliest and most efficient methods for MF generation were aligned towards statistical methods. Standard deviation (SD) is one of the most effective and common tool to analyze data statistically, it provides information about distribution in a data set around the mean. According to the literature survey it can be noted that variation of a MF in accordance to SD always yields a properly designed FLS. For instance Devi and Sarma [46] developed a statistical data based MF generation technique which utilizes histograms obtained from a data set.

The optimization technique developed in the proposed research integrates the use of statistical data along with fuzzy entropy to generate optimized MF from predefined MFs. The predefined sets used in proposed method are triangular FSs, as they are the most commonly employed sets for designing FLS [45]. A triangular MF is expressed as:

$$\mu_A(x) = \begin{cases} 0, & x \leq a \\ (x-a)/(b-a), & a < x \leq b \\ (x-c)/(b-c), & b < x < c \\ 0, & x \geq c \end{cases} \tag{2-56}$$

Here “a” and “c” are the support of the MF and “b” is the value with MF value of 1. The MF considered for the algorithm is a normal convex (having maximum membership value of 1) triangular membership function as shown in Figure 2-15

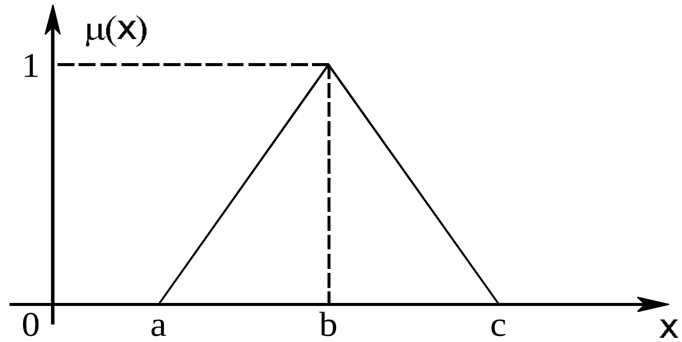


Figure 2-15 Normal triangular MF

Using De Luca and Termini [34] entropy function [equation (1-2)] the fuzzy entropy for a triangular FS [equation (2-4)] can be written as:

$$H(A) = \int_{-\infty}^{a} f(0)dx + \int_{a}^{b} f\left(\frac{x-a}{b-a}\right)dx + \int_{b}^{c} f\left(\frac{x-c}{b-c}\right)dx + \int_{c}^{\infty} f(0)dx \qquad (2\text{-}57)$$

Since the membership value for $(-\infty, a)$ and (c, ∞) is0 in a triangular MF, above equation can be rewritten as:

$$H(A) = \int_{a}^{b} f\left(\frac{x-a}{b-a}\right)dx + \int_{b}^{c} f\left(\frac{x-c}{b-c}\right)dx \qquad (2\text{-}58)$$

Substituting the MF in above equation we get:

$$H(A) = -\left[\int_{a}^{b}\left(\frac{x-a}{b-a}\right)ln\left(\frac{x-a}{b-a}\right)dx + \int_{a}^{b}\left(1-\frac{x-a}{b-a}\right)ln\left(1-\frac{x-a}{b-a}\right)dx\right]$$
$$-\left[\int_{b}^{c}\left(\frac{x-c}{b-c}\right)ln\left(\frac{x-c}{b-c}\right)dx + \int_{b}^{c}\left(1-\frac{x-c}{b-c}\right)ln\left(1-\frac{x-c}{b-c}\right)dx\right] \qquad (2\text{-}59)$$

The predefined FS used throughout this research is elucidated in Figure 2-16. These sets use a simple linguistic rule for its nomenclature and are established for their relevant position with error (ε) [42], [70]. For any system the desired error is always 0, hence FS associated with 0 error is named as “z” (zero). FS associated with positive error are named as: “sp” (small positive), “mp” (medium positive) and “lp” (large positive). FSs associated with negative error are named as: “sn” (small negative), “mn” (medium negative) and “ln” (large negative) as illustrated in Figure 2-16.

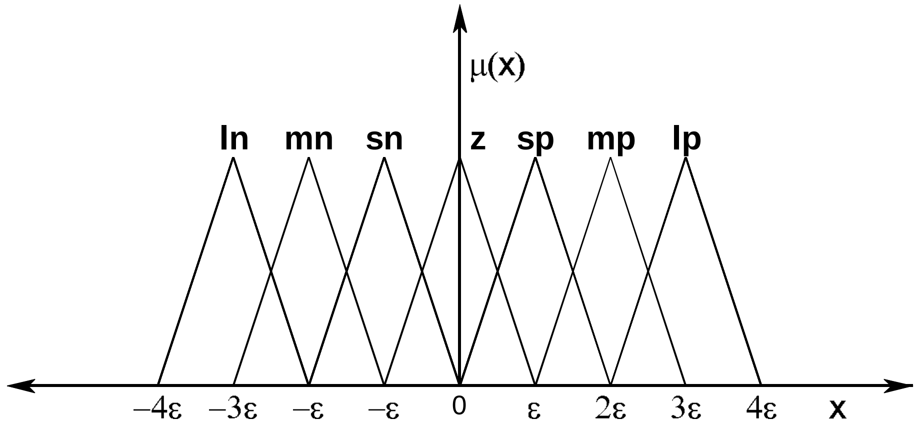


Figure 2-16 Predefined triangular MFs

The predefined sets are defined as:

1. Fuzzy set z: triangular fuzzy set having support [─ε, ε] and μ=1 at x=0
2. Fuzzy set sp: triangular fuzzy set having support [0, 2ε] and μ=1 at x=ε
3. Fuzzy set mp: triangular fuzzy set having support [ε, 3ε] and μ=1 at x=2ε
4. Fuzzy set lp: triangular fuzzy set having support [2ε, 4ε] and μ=1 at x=3ε
5. Fuzzy set sn: triangular fuzzy set having support [0,─2ε] and μ=1 at x=─ε
6. Fuzzy set mn: triangular fuzzy set having support [─ε,─3ε] and μ=1 at x=─2ε
7. Fuzzy set ln: triangular fuzzy set having support [─2ε,─4ε] and μ=1 at x=─3ε

Considering the FS "ze" its entropy function can be written as:

$$H(\mu_z) = -\left[\int_{-\varepsilon}^{0}\left(\frac{x+\varepsilon}{\varepsilon}\right)log\left(\frac{x+\varepsilon}{\varepsilon}\right)dx + \int_{-\varepsilon}^{0}\left(-\frac{x}{\varepsilon}\right)log\left(-\frac{x}{\varepsilon}\right)dx\right] - \left[\int_{0}^{\varepsilon}\left(\frac{\varepsilon-x}{\varepsilon}\right)log\left(\frac{\varepsilon-x}{\varepsilon}\right)dx + \int_{0}^{\varepsilon}\left(\frac{x}{\varepsilon}\right)log\left(\frac{x}{\varepsilon}\right)dx\right] \quad (2\text{-}60)$$

The key objective for proposed method is to obtain optimum support for a given FS; initial support for predefined sets is varied using SD. This is portrayed in Figure 2-17. Here FS "z" is displaced and is represented by dashed lines. Displaced FS is denoted as z^*.

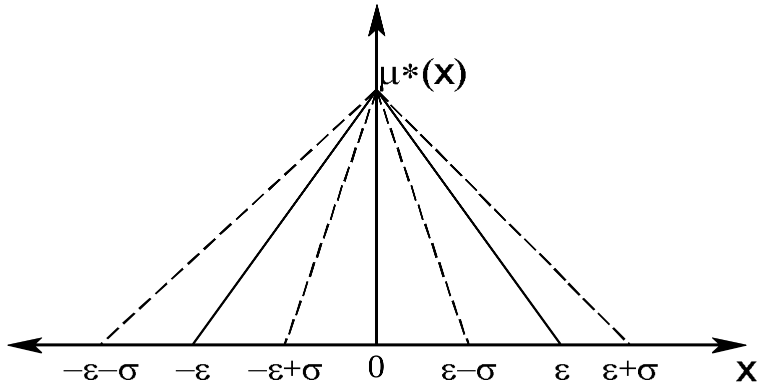


Figure 2-17 Displaced MF

The entropy for displaced set is:

$$H(\mu_{z^*}) = -\left[\int_{-\varepsilon\mp\sigma}^{0}\left(\frac{x+\varepsilon\mp\sigma}{\varepsilon\mp\sigma}\right)log\left(\frac{x+\varepsilon\mp\sigma}{\varepsilon\mp\sigma}\right)dx + \int_{-\varepsilon\mp\sigma}^{0}\left(-\frac{x}{\varepsilon\mp\varepsilon}\right)log\left(-\frac{x}{\varepsilon\mp\varepsilon}\right)dx\right] - \left[\int_{0}^{\varepsilon\pm\sigma}\left(\frac{\varepsilon\pm\sigma-x}{\varepsilon\pm\sigma}\right)log\left(\frac{\varepsilon\pm\sigma-x}{\varepsilon\pm\sigma}\right)dx + \int_{0}^{\varepsilon\pm\sigma}\left(\frac{x}{\varepsilon\pm\sigma}\right)log\left(\frac{x}{\varepsilon\pm\sigma}\right)dx\right] \quad (2\text{-}61)$$

Similarly entropy for other predefined sets can be computed. Next step is to evaluate total entropy, which is depicted as:

$$H(\mu^*) = H(\mu_{\ln^*}) + H(\mu_{mn^*}) + H(\mu_{sn^*}) + H(\mu_{z^*}) + H(\mu_{sp^*}) + H(\mu_{mp^*}) + H(\mu_{lp^*})$$
$$Or\, H(\mu^*) = \sum_{i=1}^{n} H(\mu_i^*) \quad (2\text{-}62)$$

The objective function is illustrated as:

Maximize the entropy function of each FS subjected to total maximum fuzzy entropy for all sets in the system.

Mathematically:

$$Maximize\, H(A^*) = \int_a^b f\left(\frac{x-a}{b-a}\right)dx + \int_b^c f\left(\frac{x-c}{b-c}\right)dx$$
$$Subject\ to\ maximum\ \ H(\mu^*) = \sum_{i=1}^{n} H(\mu_i^*) \quad (2\text{-}63)$$

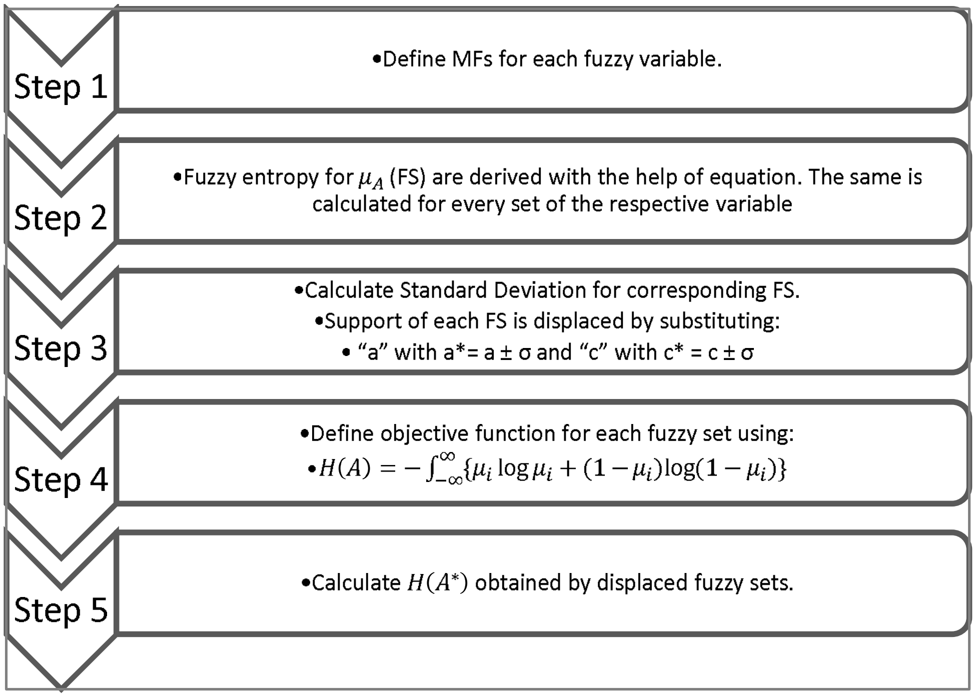
For instance the objective function for FS z^* can be written as:

Maximize

$$H(\mu_{z^*}) = -\left[\int_{-\varepsilon\mp\sigma}^{0}\left(\frac{x+\varepsilon\mp\sigma}{\varepsilon\mp\sigma}\right)log\left(\frac{x+\varepsilon\mp\sigma}{\varepsilon\mp\sigma}\right)dx + \int_{-\varepsilon\mp\sigma}^{0}\left(-\frac{x}{\varepsilon\mp\varepsilon}\right)log\left(-\frac{x}{\varepsilon\mp\varepsilon}\right)dx\right] - \left[\int_{0}^{\varepsilon\pm\sigma}\left(\frac{\varepsilon\pm\sigma-x}{\varepsilon\pm\sigma}\right)log\left(\frac{\varepsilon\pm\sigma-x}{\varepsilon\pm\sigma}\right)dx + \int_{0}^{\varepsilon\pm\sigma}\left(\frac{x}{\varepsilon\pm\sigma}\right)log\left(\frac{x}{\varepsilon\pm\sigma}\right)dx\right] \quad (2\text{-}64)$$

$$Subject\ to\ maximum\ \ H(\mu^*) = \sum_{i=1}^{n} H(\mu_i^*)$$

The objective function is formulated for all 7 predefined sets. The flowchart below summarizes the procedure for obtaining the objective function:



The next step is to solve the proposed objective function. This is done using genetic algorithms based optimization. GA based optimisation technique have been employed to obtain fuzzy entropy optimization solutions in numerous applications. Following table summarizes applicability & benefits of GA based optimization methods as compared to other methods.

Table 2-1 Comparison of GA with other optimisation methods

Reference	Description	Key findings
Andrew Starkey et. al. [71]	Comparison of Genetic Algorithms, and Particle Swarm Optimization for a real world problem: Geographical distribution of mobile field engineers for utilities like electricity, gas, water, or telecom. Entire country (UK) was divided into various regional areas which were further subdivided into work areas, where field engineers were to be distributed.	1. Objective functions were defined as: (a) maximum utilization of workforce, (b) area balancing w.r.t. the size of area. 2. Experimental results indicate GA clearly outperforms PSO 3. For multiobjective optimization, experimental results indicate MOGA clearly outperforms MOPSO.
S.Z. Yahaya et. al. [72]	A FLC was designed for a functional electrical simulation assisted elliptical stepping exercise. The controller is employed to smooth out the exercise movement. This results in a low knee joint torque.	1. GA outperforms PSO in convergence time by exhibiting a lesser convergence time. 2. GA however exhibits a slight higher RMSE as compared with PSO. 3. However GA provides better assistance hence reducing the knee joint torque. This is due to consecutive overshoots in PSO based optimization.
Ricardo Martinez-Soto et. al. [73]	Robust FLC was designed for trajectory tracking of autonomous mobile robots. A comparison is carried out between: Genetic Algorithms,	1. GA exhibited minimum convergence time. 2. GA+PSO exhibited least average error while PSO exhibited maximum

	Particle Swarm Optimization, GA+PSO techniques.	average error.
Yazmin Maldonado et. al. [74]	A type-2 FLC is implemented to control speed of DC motor by utilizing FLGA. The FLC is designed by optimizing the uncertainty parameters of fuzzy inference systems. A comparison is carried out between GA and PSO.	1. GA exhibited minimum convergence time. 2. GA exhibited minimum error for simulation results and experimental results as well.
Fevrier Valdez et. al. [75]	Comparison of dynamic optimization algorithms for fuzzy logic control, by evaluating the fuzzy control surface. Optimizatin techniques compared were: Particle Swarm Optimization, Genetic Algorithms, GA+PSO	1. Objective functions for optimized fuzzy control surface was defined and optimized using different optimization techniques. 2. GA outperforms other optimization functions for all objective functions 3. GA exhibited minimum convergence time for obtaining the optimization results.
Fevrier Valdez et. al. [76]	Comparison of various evolutionary method for optimization. Methods include: particle swarm optimization, genetic algorithms, and neural network optimization.	1. number of benchmark objective functions were optimized using different algorithms. 2. GA outperforms other optimization methods for most functions
Fevrier Valdez et. al. [77]	Comparison of various evolutionary method including Particle Swarm Optimization, Genetic Algorithms, hybrid optimization for various mathematical functions for Fuzzy Logic Decision Making	1. number of fuzzy based objective functions were optimized using different techniques. 2. GA outperforms other optimization functions for various objective functions (except Rosenbrock's function)
Malihe M. Farsangi et. al. [78]	Comparison of various multiobjective optimization algorithm for maximizing performance in a large scale power plant. Following parameters are computed to evaluate controller performance: voltage deviation, system losses and cost function.	1. Results obtained by GA exhibit best voltage deviation, and cost function 2. Convergence time for GA was found best for the techniques compared.

GAs are based on the concept of natural evolution i.e. survival of the fittest. GA does not mandate the derivative calculations to find maxima or minima for a given function. Instead it is based on the subsequent evaluations of objective function. The search direction is decided from heuristic guidelines [79]. In GA each element of parameter space is encoded into a binary string known as "chromosome", associated to the optimization objective function. A group of multiple chromosomes comprises the "population". The function values for the population are computed using objective function and new generation is evolved using various "genetic operations". Flowchart for objective function optimization is depicted in Figure 2-18 [80], [81], [82]. This flowchart illustrates the process of optimization for a given objective function. [83]

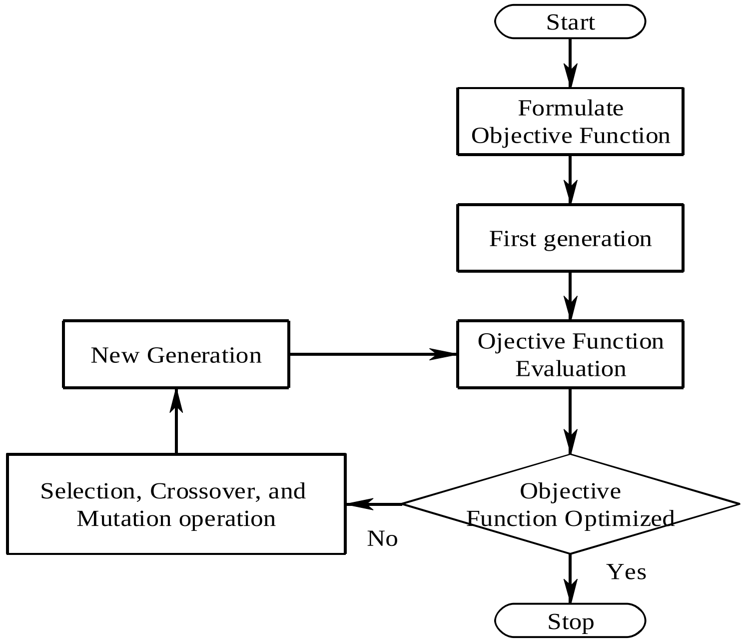


Figure 2-18 Flowchart for GA based optimization.

The parameters employed for GA optimization are summarized below in Table 2-2.

Table 2-2 Parameters used for GA

Name	Value (Type)
Generation size	300
Population	175
Selection method	Uniform
Crossover technique	Arithmetic
Mutation technique	Uniform
Termination method	Maximum generation

2. 5 Aim and scope of work

Non-linear & uncertain systems have continually challenged control system designers. The classic control theory based on linear system model and state-space control system design extensively relies on mathematical model of system. However intelligent algorithm based control systems adapt to system non-linearity and uncertainty efficiently as these controllers are inherently non-linear in nature [84]. The applicability of proposed controller is tested for some carefully chosen systems as discussed below:

2. 5. 1 Inverted pendulum system

The inverted pendulum system often the basic non-linear system encountered by control system designer. The dynamic models of robotic systems are largely based on pendulum system dynamics, the inverted pendulum system poses adequate non-linear dynamics that makes it a benchmark system to test new methods of non-linear control [85], [86]. This makes inverted pendulum system a impeccable choice to assess the proposed controller performance.

2. 5. 2 Twin rotor MIMO system

Cross-coupled MIMO system introduces a new paradigm of challenge for control system design. The conventional control system design approach for coupled systems is to design first design a non-interacting loop for individual input/output pair. Then a dedicated controller is formulated for each individual loop [87]. With recent advancements in aerospace sector unmanned aerial vehicles and autopilot systems are the recent talk of industry [88]. Twin rotor MIMO system is a principally a 2-degree-of-freedon helicopter model which offers significant cross coupling effects, making it an ideal system to assess the proposed controller applicability.

2. 5. 3 Magnetic levitation system

When non-linearity of a system increases the practicality of linearized PID controller degrades as it is based on the approximate linear model of the system. Magnetic levitation is a non-linear control problem which offers a higher non-linearity and mandates adaptive or advanced control strategies [89]. This makes it as a worthy system to evaluate adaptability for proposed controller.

2. 5. 4 Controller performance and evaluation parameters

The system performance is analyzed as follows:

1. Primarily the PID controller for system is designed and response is obtained.
2. Using system expert knowledge FLC is designed and response is obtained.
3. The initially designed FLC is than optimized, which yields optimized-FLC, response of which is compared with PID and FLC.
4. The response for optimized controller is now compared to standard works to assess the applicability of proposed controller performance.

The above steps are repeated for the systems discussed above, and relevant details are described in next few chapters. The controller's performance will be evaluated on the basis of following points:

1. Settling time
2. Rise time
3. Peak overshoot
4. Error indices

The settling time and peak overshoot illustrated the transient response nature and error-indices give a comprehensive detail for both the transient and steady-state response by computing with the accumulated error. Following error indices are used to evaluate the controller performance:

- Integral square error (ISE)

$$ISE = \int_0^{\infty} e^2(t)dt \tag{2-65}$$

- Integral time-square error (ITSE)

$$ITSE = \int_0^{\infty} t(e^2(t))dt \tag{2-66}$$

- Integral absolute error (IAE)

$$IAE = \int_0^{T} |e(t)dt| \tag{2-67}$$

- Integral time-absolute error (ITAE)

$$ITAE = \int_0^{T} t|e(t)|dt \tag{2-68}$$

The time multiplied error indices are particularly beneficial while evaluating the steady-state performance as these indices increase significantly if the error occurs during steady-state, and are therefore sometimes also called as time-penalized error indices [2].