NOVEL FUZZY TECHNIQUES FOR MODELLING HUMAN DECISION MAKING

SALANG MUSIKASUWAN, MSc.

Thesis submitted to the University of Nottingham
for the degree of Doctor of Philosophy

FEBRUARY 2013

Novel Fuzzy Techniques for Modelling Human Decision Making

Abstract

Standard (type-1) fuzzy sets were introduced to resemble human reasoning in its use of approximate information and uncertainty to generate decisions. Since knowledge can be expressed in a more natural by using fuzzy sets, many decision problems can be greatly simplified. However, standard type-1 fuzzy sets have limitations when it comes to modelling human decision making.

In many applications involving the modelling of human decision making (expert systems) the more traditional membership functions do not provide a wide enough choice for the system developer. They are therefore missing an opportunity to produce simpler or better systems. The use of complex non-convex membership functions in the context of human decision making systems were investigated. It was demonstrated that non-convex membership functions are plausible, reasonable membership functions in the sense originally intended by Zadeh.

All humans, including ‘experts’, exhibit variation in their decision making. To date, it has been an implicit assumption that expert systems, including fuzzy expert systems, should not exhibit such variation. Type-2 fuzzy sets feature membership functions that are themselves fuzzy sets. While type-2 fuzzy sets capture uncertainty by introducing a range of membership values associated with each value of the base variable, but they do not capture the notion of variability. To overcome this limitation of type-2 fuzzy sets, Garibaldi previously proposed the term ‘non-deterministic fuzzy reasoning’ in which variability is introduced into the membership functions of a fuzzy system through the use of random alterations to the parameters.

In this thesis, this notion is extended and formalised through the introduction of a notion termed a *non-stationary fuzzy set*. The concept of random perturbations that can be used for generating these non-stationary fuzzy sets is proposed. The *footprint of variation* (FOV) is introduced to describe the area covering the range from the minimum to the maximum fuzzy sets which comprise the non-stationary fuzzy sets (this is similar to the *footprint of uncertainty* of type-2 sets). Basic operators, i.e. union, intersection

and complement, for non-stationary fuzzy sets are also proposed. Proofs of properties of non-stationary fuzzy sets to satisfy the set theoretic laws are also given in this thesis.

It can be observed that, firstly, a non-stationary fuzzy set is a collection of type-1 fuzzy sets in which there is an explicit, defined, relationship between the fuzzy sets. Specifically, each of the instantiations (individual type-1 sets) is derived by a perturbation of (making a small change to) a single underlying membership function. Secondly, a non-stationary fuzzy set does not have secondary membership functions, and secondary membership grades. Hence, there is no 'direct' equivalent to the embedded type-2 sets of a type-2 fuzzy sets. Lastly, the non-stationary inference process is quite different from type-2 inference, in that non-stationary inference is just a repeated type-1 inference.

Several case studies have been carried out in this research. Experiments have been carried out to investigate the use of non-stationary fuzzy sets, and the relationship between non-stationary and interval type-2 fuzzy sets. The results from these experiments are compared with results produced by type-2 fuzzy systems. As an aside, experiments were carried out to investigate the effect of the number of tunable parameters on performance in type-1 and type-2 fuzzy systems. It was demonstrated that more tunable parameters can improve the performance of a non-singleton type-1 fuzzy system to be as good as or better than the equivalent type-2 fuzzy system.

Taken as a whole, the techniques presented in this thesis represent a valuable addition to the tools available to a model designer for constructing fuzzy models of human decision making.

Declaration

The work in this thesis is based on research carried out at the *Intelligent Modelling and Analysis* (and, originally, the *Automated Scheduling, Optimisation and Planning*) Research Group, the School of Computer Science, the University of Nottingham, England. No part of this thesis has been submitted elsewhere for any other degree or qualification and it is all my own work unless referenced to the contrary in the text.

Acknowledgements

Firstly, I would like to express my most sincere gratitude to my supervisor, Professor Jon Garibaldi, for providing support, guidance and opportunities throughout my PhD has made it possible to successfully complete this work. I am extremely grateful for his valuable comments and help he has afforded throughout this research programme.

In particular, I would like to thank Dr. Turhan Ozen for the technical help, encouragement and advice provided at the beginning of my research.

I would also like to thank the staff and all students of the IMA/ASAP Group at the University of Nottingham. Thank you for creating a warm and friendly environment for academic research.

Again, I would like to express my most sincere gratitude to my sponsor, The Royal Thai Government, for all the support that they provided throughout my study.

Finally, I would like to express my sincerest thanks and gratitude to my family in Thailand who have provided their deep love and unfailing encouragement throughout my recent and previous studies.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background and Motivation

Fuzzy sets were first introduced by Zadeh [2] in order to model the imprecision and uncertainty inherent in assigning memberships of elements to real-world sets, such as the set of *old* people or the set of *tall* people. These fuzzy sets were specifically designed to represent uncertainty and vagueness and provided formalised tools for dealing with the imprecision in real world problems. Knowledge can often be expressed more naturally by using (type-1) fuzzy sets and many complex decision making problems can be significantly simplified. However, type-1 fuzzy sets still have limitations, in that they are unable to model the effects of all uncertainties. Further, there is actually no *fuzziness* in the standard type-1 membership grade, as has been pointed out by many people including Klir and Folger [3]

> "... it may seem problematical, if not paradoxical, that a representation of fuzziness is made using membership grades that are themselves precise real numbers".

Zadeh addressed this problem originally in his seminal paper of 1975 [4] in which he introduced the concept of linguistic variables. Zadeh proposed "fuzzy sets with fuzzy membership functions" and went on to define fuzzy sets of type n, n = 2, 3, ..., for which the membership function ranges over fuzzy sets of type n - 1. The use of type-2 fuzzy sets was advocated many years ago by people including Dubois and Prade [5]

> "To make into account the imprecision of membership functions, we may think of using a type-2 fuzzy set",

and Yager [6]

> "The usefulness of fuzzy subsets of type II [type-2] is that it enables us to extend membership grades to linguistic values".

However, the use of type-2 sets in practice has been limited due to the significant increase in computational complexity involved in their implementation. Type-2 fuzzy sets can model uncertainties better and minimize their effects. The use of type-2 sets was advocated and extended by people: Dubois and Prade gave a formula for the composition of type-2 relations as an extension of the type-1 sup-star composition for the minimum t-norm [5], Mizumoto and Tanaka studied the set theoretic operations of type-2 sets and properties of membership degrees of such sets [7] and examined type-2 sets under the operations of algebraic product and algebraic sum [8], etc.

More recently, type-2 sets have received renewed interest mainly due to the effort of Mendel [1] but also, possibly, by the increases in computational power over recent years. Mendel has established a set of terms to be used when working with type-2 fuzzy sets and, in particular, introduced a concept known as the *footprint of uncertainty* which provides a useful verbal and graphical description of the uncertainty captured by any given type-2 set. The interested reader is particularly referred to [9] for a summary tutorial and/or [1] for a more detailed treatment. Mendel has particularly concentrated on a restricted class of general type-2 fuzzy sets know as *interval valued* type-2 fuzzy sets. Interval valued fuzzy sets are characterised by having secondary membership functions which only take the value in $\{0, 1\}$. This restriction greatly simplifies the computational requirements involved in performing inference with type-2 sets and Mendel has provided closed formula for intersection, union and complement, and computational algorithms for type reduction (necessary for type-2 defuzzification).

It is well accepted that all humans including 'experts', exhibit variation in their decision making. Variation may occur among the decisions of a panel of human experts (inter-expert variability), as well as in the decisions of an individual expert over time (intra-expert variability). Up to now it has been an implicit assumption that expert systems, including fuzzy expert systems, should not exhibit such variation. While type-2 fuzzy sets capture the concept of introducing uncertainty into membership functions by introducing a range of membership values associated with each value of the base variable, they do not capture the notion of variability — as a type-2 fuzzy inference system (FIS) will always produce the same output(s) (albeit a type-2 fuzzy set with an implicit representation of uncertainty) given the same input(s). Garibaldi et al. [10–14] have been investigating the incorporation of variability into decision making in the context of fuzzy expert systems in a medical domain. In this work, Garibaldi proposed the notion of *non-deterministic fuzzy reasoning* in which variability is introduced into the membership functions of a fuzzy system through the use of random alterations to the parameters of the generating functions. This notion was later extended and formalised through the introduction of a notion termed a *non-stationary fuzzy set*.

1.2 Aims of this Thesis

The ultimate goal of this research is to establish the techniques to model human decision making, with a particular focus on modelling the variation apparent in all human decision making. The project focuses on the development of a new type of fuzzy set and the associated systems that are able to capture the concept of introducing uncertainty into membership functions by introducing a range of membership values associated with each value of the base variable. In this way, they are specifically able to capture the notion of *variability* (current type-2 fuzzy sets are unable to do this).

In order to achieve this aim, the following objectives were identified:

- Investigate the use of a wider range of membership functions than often found in fuzzy inference systems. Specifically, non-convex membership functions together with traditional fuzzy sets were used in a case-study to build a fuzzy expert system to predict energy demand.

- Investigate the type-1 and type-2 fuzzy systems for time-series forecasting to examine the relationship between the number of model parameters and the performance of type-1 and type-2 fuzzy systems.

- Understand the notion of *non-deterministic fuzzy reasoning* in which variability is introduced into the membership functions of a fuzzy system through the use of random alterations to the parameters of these functions. The objective is to determine, extend, and formalise these through the introduction of notion term a *non-stationary fuzzy sets*.

- Implement the fuzzy systems to illustrate the use of non-stationary fuzzy sets and to investigate the relationships between interval type-2 and non-stationary fuzzy sets.
- Explore how the form of the primary membership function affects the inference process within a non-stationary fuzzy system.
- Examine the relationship between the primary membership functions and output uncertainties in non-stationary fuzzy sets compared with interval type-2 fuzzy sets.

1.3 Organisation of the Thesis

In this thesis, standard set theory, (type-1) fuzzy sets and systems, and then continues to present the concept of type-2 fuzzy sets and systems. Chapter 3 examines the use of non-convex membership functions for linguistic terms and presents a case-study. Chapter 4 presents the investigations into the performance of type-1 and type-2 fuzzy systems for time series forecasting. In Chapter 5, the new concept of non-stationary fuzzy sets is proposed. In Chapter 6, the relationships between interval type-2 and non-stationary fuzzy sets are presented. Chapter 7 presents the investigation of the primary membership functions in non-stationary fuzzy sets by (i) exploring the outputs of Gaussian and Triangular primary membership functions in non-stationary fuzzy sets, and (ii) investigating the relationship between primary membership functions and output uncertainties in interval type-2 and non-stationary fuzzy sets, and (iii) investigating and comparing the performance of type-1, type-2, and non-stationary fuzzy systems through MacKey-Glass Time-Series.. Finally, Chapter 8 draws conclusions, lists the contributions arising from this work, mentions the limitations and suggests some interesting potential directions for future research presented in this thesis.

Chapter 2

Theory of Fuzzy Sets and Systems

2.1 Sets Theory

A set is a collection of objects called elements of the set. The use of the word '*set*' means that there is also a method to determine whether or not a particular object belongs in the set. An object contained by a set is called a *member*, or *element*. We then say that the set is well-defined. For example, it is easy to decide that the number 6 is in the set consisting of the integers 1 through 9. After all, there are nine objects to consider and it is clear that 6 is one of them by simply checking all nine.

In this section, capital letters denote sets, while members of a set are written in lower-case. To indicate the *universe of discourse*, often referred to as the universal set, we use $\mathbb{U}$. All sets are members of the universal set. Additionally, a set with no elements is called a *null*, or *empty* set and is denoted $\emptyset$.

If we have an element x of set A, it can be represented as:

$$x \in A$$

while if x is not a member of A, it can be written as:

$$x \notin A.$$

There are two methods used to describe the contents of a set, the *list method* and the *rule method*. The list method defines the members of a set by listing each object in the set

$$A = \{a_1,\ a_2,\ \ldots,\ a_n\}.$$

The rule method defines the rules that each member must adhere to in order to be considered a member of the set

$$A = \{a \mid a\ has\ properties\ P_1,\ P_2,\ ...,\ P_n\}.$$

When every element in the set A is also a member of set B, then A is a *subset* of B

$$A \subseteq B.$$

If every element in A is also in B and every element in B is also in A, i.e. $A \subseteq B$ and $B \subseteq A$, then A and B are *equal*

$$A = B.$$

If at least one element in A is not in B or at least one element in B is not in A, then A and B are *not equal*

$$A \neq B.$$

Set A is a *proper subset* of B if A is a subset B but A and B are not equal, i.e. $A \subseteq B$ and $A \neq B$

$$A \subset B.$$

To present the notion that an object is a member of a set either fully or not at all, we introduce the function μ. For every $x \in \mathbb{U}, \mu_A(x)$ assigns a value that determines the strength of membership of each x in the set A,

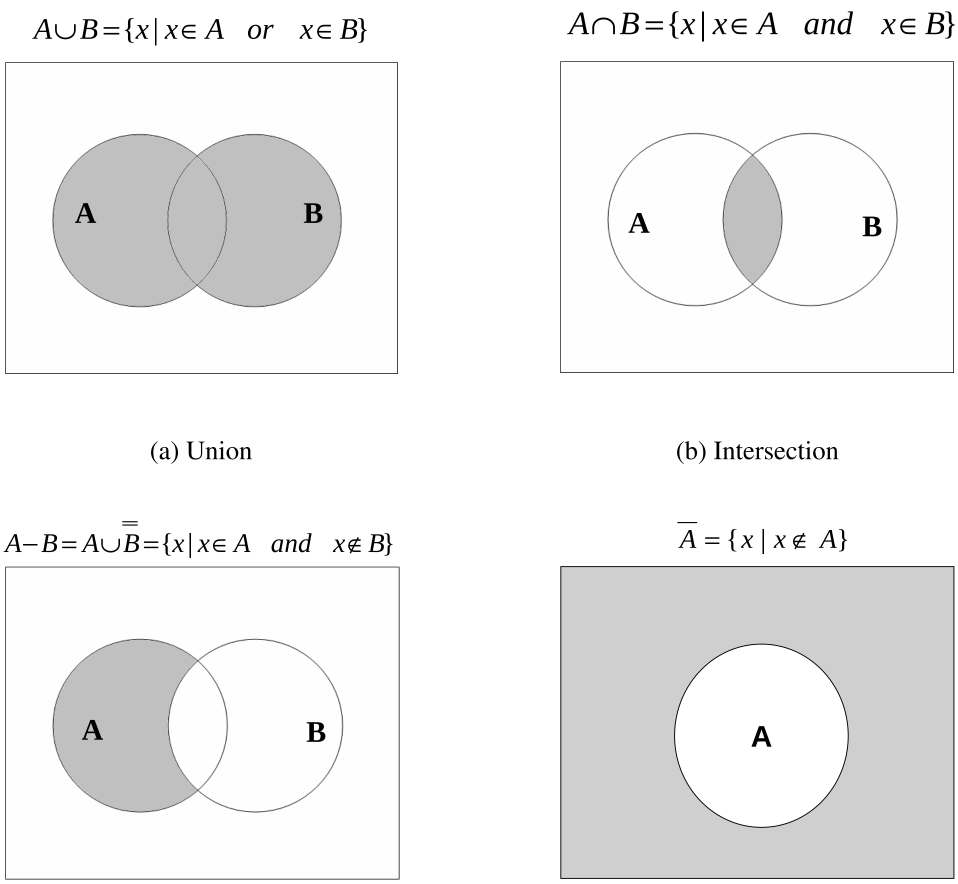
$$\mu_A(x) = \begin{cases} 1 & \text{if and only if} \quad x \in A, \\ 0 & \text{if and only if} \quad x \notin A. \end{cases}$$

Therefore, μ_A maps all elements of the universal set into the set A with values 0 and 1

$$\mu_A : \mathbb{U} \rightarrow \{0,1\}.$$

Using the given notation, four basic operations that can be used on sets are shown in Figure 2.1 using Venn diagrams and also written in set theoretic notation.

The operations shown in Figure 2.1 are routinely combined to produce more complex functions. Note that these examples use only two sets, but union and intersection can be defined for any number of sets. This is due to the properties of the basic operations shown in Table 2.1. Preserving these behaviours is important as fuzzy sets are a generalisation of classic sets and must able to reproduce exactly their behaviour [15].



(a) Union (b) Intersection

(c) Difference (d) Complement

Figure 2.1: Basic crisp set operations: The shaded region indicates the result of applying the given function

Table 2.1: Summary of some crisp set properties

Set Property	Description
Involution	$\bar{\bar{A}} = A$
Commutativity	$A \cup B = B \cup A$ $A \cap B = B \cap A$
Associativity	$(A \cup B) \cup C = A \cup (B \cup C)$ $(A \cap B) \cap C = A \cap (B \cap C)$
Idempotence	$A \cup A = A$ $A \cap A = A$
Distributivity	$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
Absorption	$A \cup (A \cap B) = A$ $A \cap (A \cup B) = A$

2.2 Boolean Logic

The well known Boolean logic system includes only two values, true or 1 and false or 0, which make up the *truth values* of this system. These values can be combined using *operators* to produce the *vocabulary* of Boolean logic [17]. The *truth tables* are used to show the response of the various operators to different combinations of truth values, as shown in Table 2.1. In the table, A and B are variables that can take on either of two possible truth values. For a logic system with n possible truth values, there would be 2^n possible combinations of these values using two variables.

Table 2.2: Boolean logic truth tables. The Symbols $\wedge$, $\vee$, $\leftrightarrow$, $\rightarrow$ *and*, *or*, *equivalence* and *implication* respectively

A	B	$\wedge$	$\vee$	$\leftrightarrow$	$\rightarrow$
0	0	0	0	1	1
0	0	0	0	1	1
1	0	0	1	0	0
1	1	1	1	1	1

In this example, there are 16 possible truth values, each defining one operator. In Table 2.2, four operators are shown, while there are 12 others left undefined. Each of these can be interpreted with a meaning attached; in the table we have *and*, *or*, *equivalence*, and *implication*, while those not listed may not be immediately obvious. There are other logic systems that can be defined using similar systems. For instance, there have been a number of three-valued logics defined that allow an *in-between* value such as that by Lukasiewicz [16]. These three-valued logics would have a maximum of $3^{3^2} = 729$ different truth tables. As more values are allowed, the number of truth tables grows and becomes extremely unwieldy.

Reasoning in a particular logic system is carried out using the operators allowed in the system. The reasoning procedure relies on *tautologies*, or rules that remain true in that logic system regardless of the values assigned to the variables involved. A tautology is true due to the logic of its construction. Four examples are shown for Boolean logic in Table 2.3. A *contradiction* is considered the opposite of a tautology, where the statement made is always false. For more information on classical logic see Klir& Folger [3].

Table 2.3: Boolean logic tautologies

Name	Definition
modus ponens	$(A \wedge (A \rightarrow B)) \rightarrow B$
modus tollens	$((A \Rightarrow B) \wedge \neg B) \Rightarrow \neg A$
hypothetical syllogism	$((A \Rightarrow B) \wedge (B \Rightarrow C)) \Rightarrow (A \Rightarrow C)$
contraposition	$(A \Rightarrow B) \Rightarrow (\neg B \Rightarrow \neg A)$

2.3 Fuzzy Sets and Systems

Although knowledge can be expressed more naturally by using (type-1) fuzzy sets and many complex decision making problems can be sigificantly simplified. A fuzzy set is a generalisation of a crisp set. It has been defined on a universe of discourse X and is charecterised by a *membership function*, $\mu(x)$, that takes on values in the interval [0,1] [1].

2.3.1 Membership Functions

The new concept was based on a simple modification of the most basic and fundamental of mathematical notions, that of *set*, by allowing elements to have a partial degree of membership, expressed by a number between 0 and 1. The *membership grade* of an element x in the set A is denoted by the function $\mu_A(x)$. μ_A maps all elements of the universal set into the set A with values in the continuous interval 0 to 1 [2],

$$\mu_A : \mathbb{U} \rightarrow [0,1].$$

As for crisp sets, a fuzzy set may be defined formally in two ways, each introduced by Zadeh [2]. The *list method* for a fuzzy set lists the strength of membership of each element of a discrete, countable *universe of discourse* ($\mathbb{U}$) to the set in equation [2],

$$A = \sum_{i=1}^{n} \mu_i / x_i \tag{2.1}$$

$$A = \{\mu_1/x_1 + \cdots + \mu_n/x_n\} \tag{2.2}$$

where x_i denotes the i^{th} member of $\mathbb{U}$ and μ_i/x_i is the strength of membership of element x_i. The use of the plus symbol to separate individual elements is a departure from standard set theory notation, which uses the comma. To describe a fuzzy set on a continuous universe we write [2],

$$A = \int_{\mathbb{U}} \mu_A(x)/x. \tag{2.3}$$

In either case, $\mu_A(x)$ is a function that assigns membership to A from every element in $\mathbb{U}$. For example, if we wish to represent temperature close to 25 °C using a fuzzy set with a continuous universal set we can define $\mu_A(x)$ to be

$$\mu_A(x) = 1/(1 + \frac{(x-25)^2}{25}). \tag{2.4}$$

The function shown in Figure 2.2, maps every real number into the set of temperatures close to 25 °C. If the temperature was 10 °C, that would be assigned a valued of 0.05, while 20 °C gets 0.33 and 25 °C, 1. Obviously, the closer number to 25 °C, the higher its membership in the set.

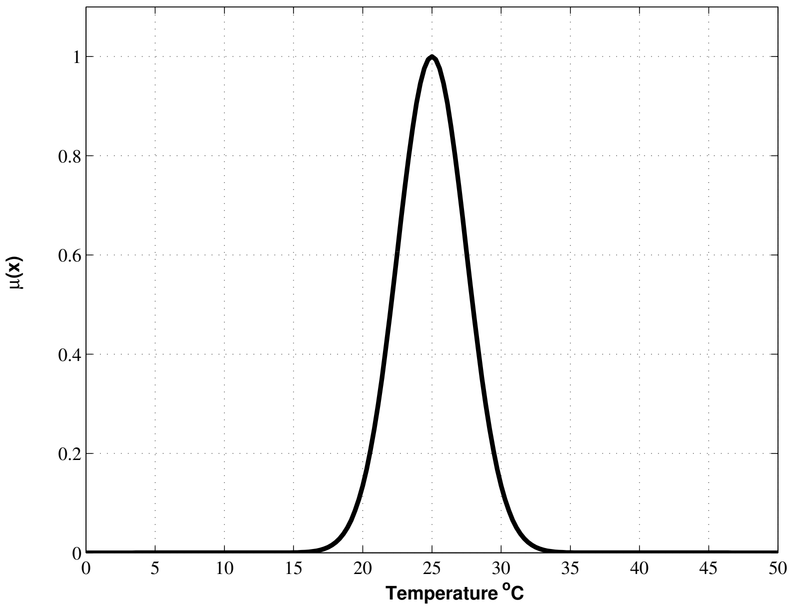


Figure 2.2: a fuzzy representation of the continuous definition for ‘Close to 25 °C’.

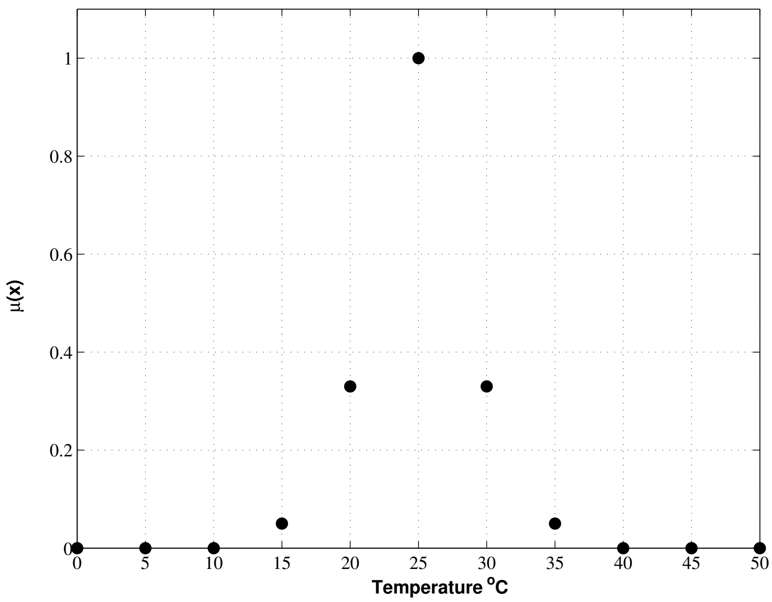


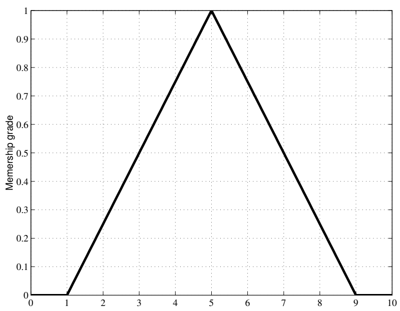
Figure 2.3: fuzzy representation of the discrete definition for ‘Close to 25 °C’.

Alternatively, if we are dealing with a discrete universe of discourse, a similar function can be defined in accordance with Equation 2.2 as:

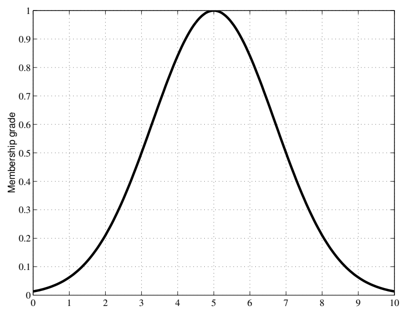
$$A = \{0.03/5 + 0.05/10 + 0.11/15 + 0.33/20 + 1.0/25 \quad (2.5)$$

$$+0.33/30 + 0.11/35 + 0.05/40 + 0.03/45\,\}. \quad (2.6)$$

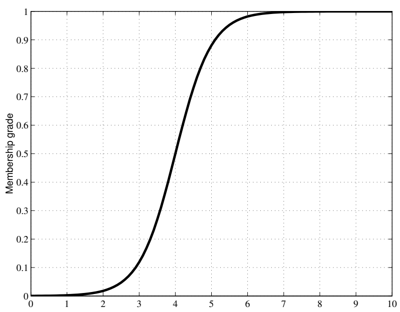
which consists of points when plotted, as seen in Figure 2.3.



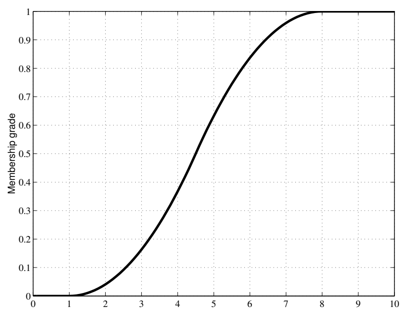
(a) Triangular



(b) Gaussian



(c) Sigmoid



(d) S-Shaped

Figure 2.4: Membership functions of shape (a) triangular, (b) Gaussian, (c) Sigmoid, and (d) S-shaped.

In general the shape of the membership function can be any thing more precise, but some of the more common shapes in practice are *triangular*, *Gaussian*, *Sigmoid*, or *S-shaped*. Each of this is shown in Figures 2.4.

2.4 Fuzzy Operators

The three main operators on any set, whether crisp of fuzzy are *complement*, *union*, and *intersection*. These three operations are capable of producing more complex ones when used in combination. In the classical set theory these operations can be uniquely defined as seen in [17]. In fuzzy set theory these operations are no longer uniquely defined, as membership values are no longer restricted to $\{0,1\}$ and can be in the range [0,1]. Any definition of these operations on fuzzy sets must include the limiting case of crisp sets. These operators are usually defined by the Zadeh [2] as follows:

- Complement:- A': $\mu_{A'}(x) = 1 - \mu_A(x)$
- Intersection:- $A \cap B$: $\mu_{A \cap B}(x) = min[\mu_A(x), \mu_B(x)]$
- Union:- $A \cup B$: $\mu_{A \cup B}(x) = max[\mu_A(x), \mu_B(x)]$

Then, the complement corresponds to the *not* of multi-valued logic, the union corresponds to *or* and the intersection corresponds to *and*.

2.4.1 Fuzzy intersection

The operation of fuzzy intersection takes two sets and returns a single set representing their difference. There are an infinite number of ways to define the fuzzy intersection. Intersection operators that adhere to specific requirements are considered to be a part of the general class of aggregation operators called *t-norms* denoted by the symbol T. These fuzzy intersection operators, which are usually referred to as T-norms operators, meet the following basic requirements. A T-norm operator is a two-place function $T(.,.)$ satisfying [2]

- boundary: $T(0,0) = 0, T(a,1) = T(1,a) = a$
- monotonicity: $T(a,b) \leq T(c,d)$ if $a \leq c$ and $b \leq d$
- commutativity: $T(a,b) = T(b,a)$
- associativity: $T(a,T(b,c)) = T(T(a,b),c)$

The first requirement imposes the correct generalisation to crisp sets. The second requirement implies that a decrease in the membership values in A or B cannot produce an increase in the membership value in A intersection B. The third requirement indicates that the operator is indifferent of the order of the fuzzy sets to be combined. Finally, the fourth requirement allows us to take the intersection of any number of sets in any order of pairwise groupings. Additionally, there are two further requirements that are useful [2].

- T is continuous function
- idempotent: $T(a,a) = a$

The fifth requirement deals with continuity while the sixth requirement ensures that the fuzzy intersection of a set with itself returns the original set.

The classical intersection ($\cap$) of ordinary subsets of X can be extended by the following formula which was proposed by Zadeh [2]:

$$\mu_{A\cap B}(x) = min(\mu_A(x), \mu_B(x)), \forall x \in X$$

where $\mu_{A\cap B}(x)$ is the membership functions of $A \cap B$. This formula gives the usual intersection when the valuation set is reduced to $\{0,1\}$. Obviously, there are other extensions of $\cap$ coinciding with the binary operators. Alternatively, this may be shown as: [2]

$$T(A,B) = min(a,b), \forall a \in A, \forall b \in B.$$

It is easily shown that this function satisfies requirements 1 to 6. A pictorial of fuzzy intersection is shown in Figure 2.5.

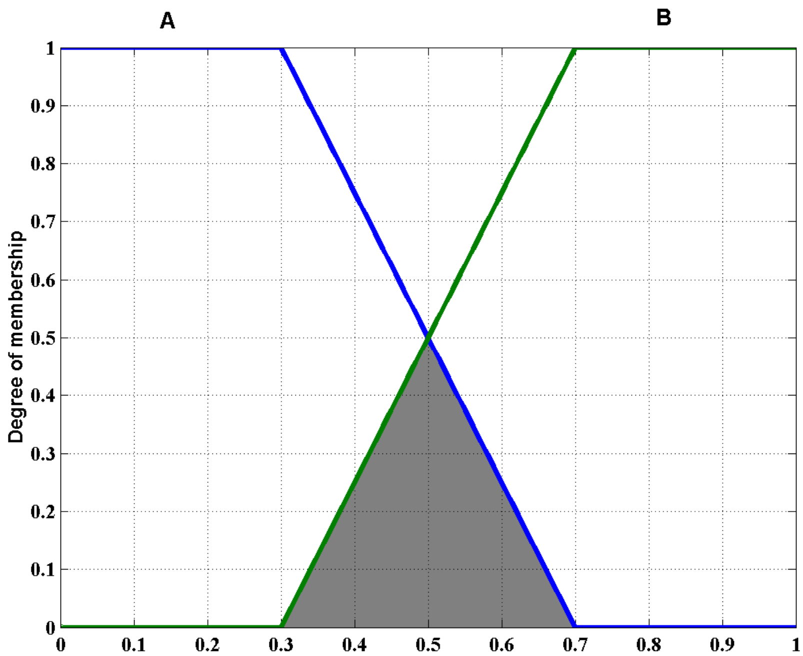


Figure 2.5: The classic intersection. The shaded region represents the intersection of sets A and B.

An alternate definition is a member of the Yager class [18],

$$T_w(a,b) = 1 - min(1, ((1-a)^w + (1-b)^w)^{1/w}), w \in (0, \infty).$$

As $w \to \infty$, this function behaves like the classic intersection.

2.4.2 Fuzzy union

The operation of fuzzy union takes two sets and returns a single set representing their union. For each element in $\mathbb{U}$, the fuzzy union operator S takes the element's membership grade in set A and B and returns the new membership in the set $A \cup B$. Like fuzzy intersection, the fuzzy union operator is specified in general by a function S: [2]

$$\mu_{A \cup B}(x) = S(\mu_A(x), \mu_B(x)).$$

Fuzzy union operators which are often referred to as *T-conorms* or *S-norms* operators, satisfy the following basic requirements. There are an infinite variety of definitions for fuzzy union with the four axioms shown below. Such operators are considered to be a part of the general class of aggregation operators called *T-conorms* or *S-norms*, denoted by the symbol S below [2].

- boundary: $S(1,1) = 1, S(a,0) = s(0,a) = a$
- monotonicity: $S(a,b) \leq S(c,d)$ if $a \leq c$ and $b \leq d$
- commutativity: $S(a,b) = S(b,a)$
- associativity: $S(a,S(b,c)) = S(S(a,b),c)$

The justification of these basic requirement is similar to that of the requirements for the T-norm operators. Additionally, there are two further requirements that can be useful:

- S is continuous function
- idempotent: $S(a,a) = a$

The fifth requirement deals with continuity while the sixth requirements ensures that the fuzzy union of a set with itself returns the original set.

The classic fuzzy union is shown in Figure 2.6 and is defined as: [2]

$$S(A,B) = max(a,b) \forall a \in A, \forall b \in B.$$

It is easily shown that this operator satisfies requirements 1-6. However, this operator makes intuitive sense. When two or more sets are joined together, the strongest membership of an element is used. The strength of membership is also equal to the element's

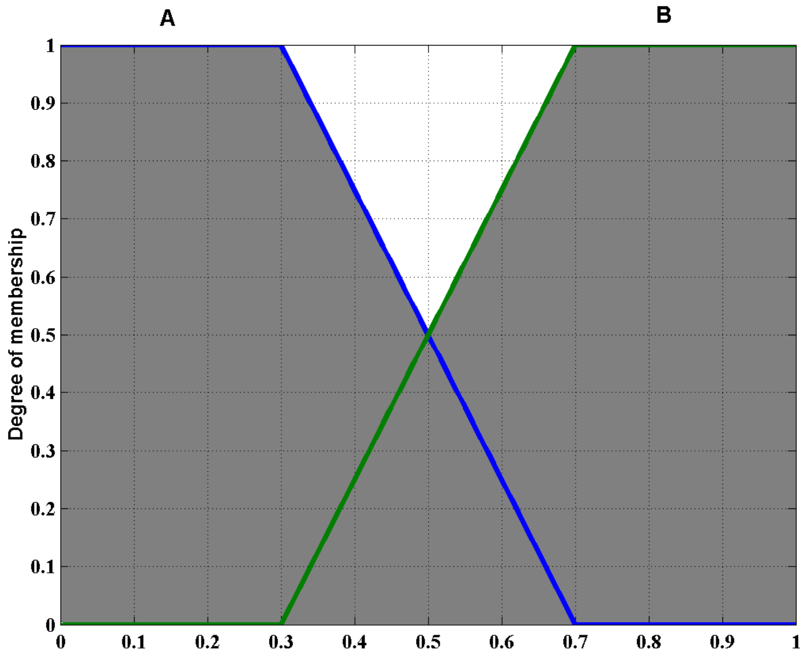


Figure 2.6: The classic fuzzy union. The Shaded region represents the union of set A and B.

largest strength in either of the two sets. An alternate definition is a member of the Yager class [18],

$$S_w(a,b) = min(1, (a^w + b^w)^{1/w}), w \in (0, \infty).$$

First introduced in a paper by Yager [18], this function behaves as the classic union as $w \to \infty$.

2.4.3 Complement

The least complex of the three operations, fuzzy complement, describes the difference between an object and its opposite. There are two rules that every fuzzy complement operator must follow to be compatible with crisp logic.

- $C(0) = 1, C(1) = 0$: This rule defines the boundary conditions, the bivalent case.
- If $a < b$ then $C(a) \geq C(b)$; $a, b \in (0, 1)$: This rule defines the fuzzy complement to be monotonic increasing.

In order to comply with the requirement that all fuzzy complements can mimic a crisp complement, the first rule is needed. It makes sense that as the degree of membership

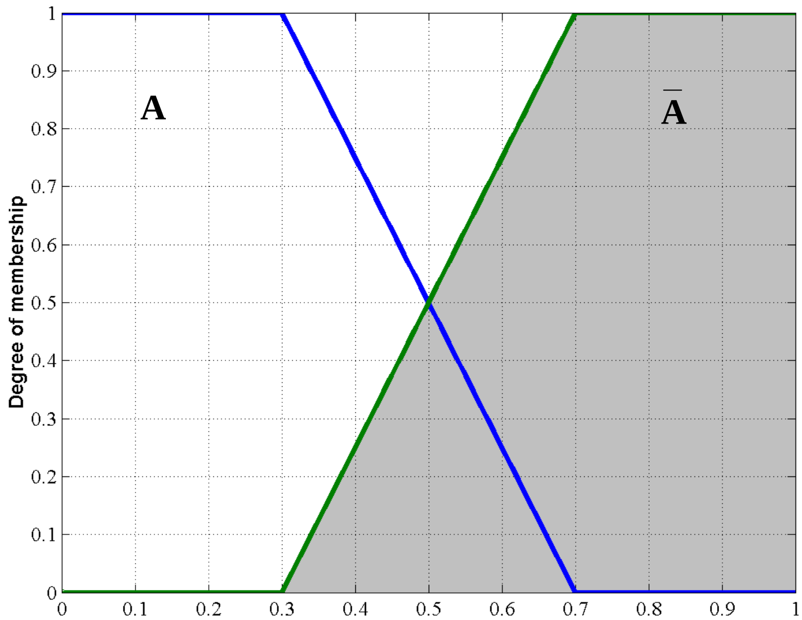


Figure 2.7: The classic fuzzy complement. The Shaded region represents the complement of set A.

of an element in set A increases, its membership in the complement set should decrease, and this is required by second rule. Operators that adhere to the first and second rules are members of the general class of fuzzy complements. There are two further rules that are not required for an operator to act as a fuzzy complement, but are nevertheless useful.

- C is a continuous function.
- $C(C(a)) = a$ for all $a \in [0,1]$: The fuzzy complement is involutive.

A set that has been defined using the rule method is continuous and requires continuous definitions for the complement. An additional restraint that can be imposed is involution. It may be desirable for the complement of a set to be reversible, by a further use of the complement function, as shown in the fourth rule. Although the first and second rules must be satisfied, it is possible to produce an infinite number of definitions. Some of the more popular are shown in Fig. 2.7.

The original fuzzy complement is shown in Figure 2.7 and was given by Zadeh [2] as $C(a) = 1 - a$. This definition follows rules 1-4 and is also shown in Figure 2.8 and 2.9 as the straight line. Later, Sugeno [19] introduced the λ-complement as $C_\lambda(a) = \frac{1-a}{1+\lambda a}$, $\lambda \in (-1, \infty)$. This is an alternative definition of fuzzy complement and follows rules 1-3 and is shown in Figure 2.8.

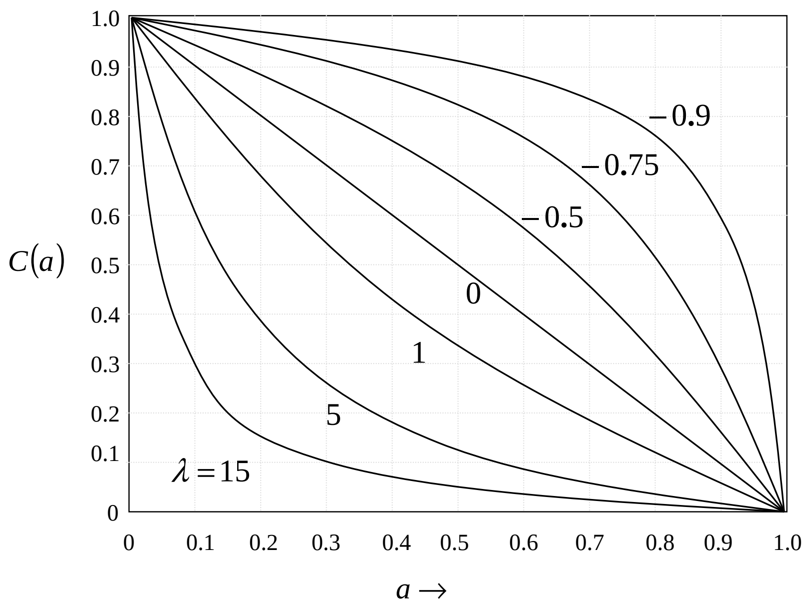


Figure 2.8: The Sugeno complement: Each graph corresponds to a given value of λ. Note that when $\lambda = 0$, this function behaves as the classic complement, $C(a) = 1 - a$

In 1980 Yager [18] defined the Yager Class as $C_w(a) = (1 - a^w)^{1/w}$, $w \in (0, \infty)$. This definition follows rules 1-3 and is shown in Figure 2.9.

2.5 Fuzzy Logic

That there are only two truth values, true and false, is a property of classical logic that has been challenged for some time [20]. Fuzzy logic is built upon fuzzy set theory which contains an infinite number of truth values. Fuzzy sets do not have crisply defined members and can contain elements with only a partial degree of membership. Similarly, in fuzzy logic the truth of statements is a matter of degree. The construction of traditional truth tables to express fuzzy logic is therefore impossible. A modified version of the truth table can be created using continuous operators. In Table 2.4, we have a truth table for a fuzzy logic used to compute the values, allowing the calculation of truth values using the full range of possible truth values.

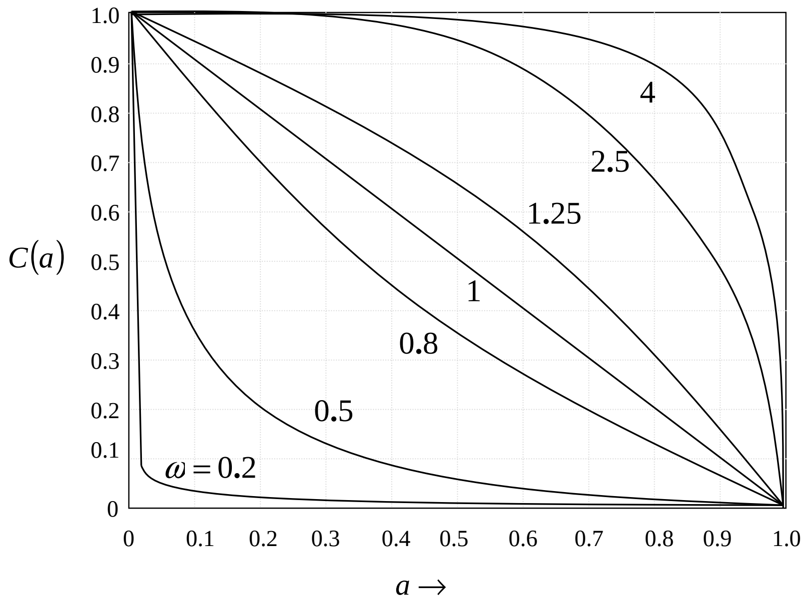


Figure 2.9: The Yager complement: Each graph corresponds to a given value of w. Note that when $w = 1$, this function behaves as the classic complement, $C(a) = 1 - a$

2.5.1 Linguistic variables

A variable in the classic sense is a placeholder that can take on any value defined in its universe of discourse. So for example, to describe the temperature of a room, the variable would be *temperature* and the value would be a number, such as 25°C. The equivalent in fuzzy logic is called a *linguistic variable*. A linguistic variable differs in that in addition to accepting a crisp number as input, it also has any number of fuzzy terms defined over its universe of discourse. To continue the temperature example, the linguistic variable would be *temperature* while the terms are fuzzy sets such as, *low*, *medium*, *high*, etc.

A *term*, or *linguistic label*, is a fuzzy set that describes a fuzzy concept and is defined on the universe of discourse of its parent linguistic variable, i.e.

$$\{term_1, term_2, \ldots, term_n\} \subset variable, \tag{2.7}$$

where $\subset$ indicates that terms are subsets of a variable.

If one wanted to capture the meaning of *medium*, a fuzzy set would be constructed which maps temperatures to membership values using a membership function. Each point in the graph would correspond to a specific temperature's membership in the term *medium*. The following is a possible definition, shown graphically in Figure 2.10.

Table 2.4: Fuzzy logic truth tables

A	B	$T(A,B)$	$S(A,B)$
0	0	0	0
0	1	0	1
1	0	0	1
1	1	1	1

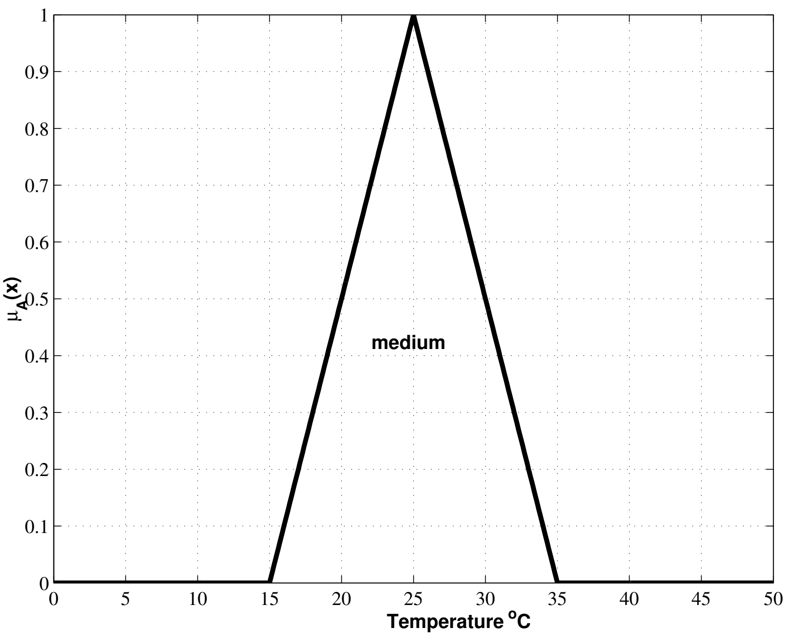


Figure 2.10: Definition for the *medium* term

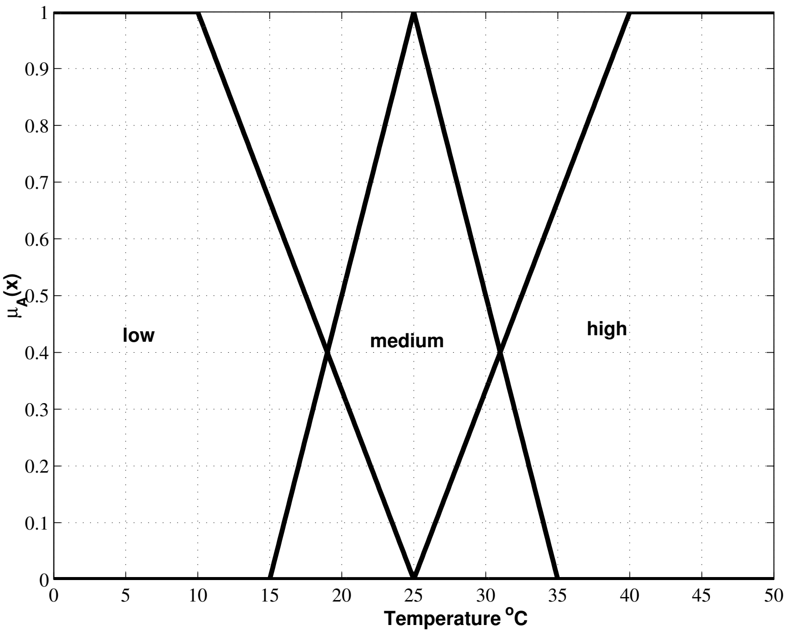


Figure 2.11: Linguistic variable (Temperature) defined with terms *low*, *medium*, and *high*.

If one were to define more terms for *temperature*, the graph shown in Figure 2.11 could be produced. It can be seen that three terms have been defined; *low*, *medium*, and *High*. Triangular and trapezoidal membership functions have been used in this example, but are not by no means required for the definition of terms. Many other membership functions can be used, including Gaussian and sigmoidal functions. For example, the Gaussian function is defined as

$$f(\sigma, c) = e^{\frac{-(x-c)^2}{2\sigma^2}}, \tag{2.8}$$

where c and σ are the mean and variance of the function.

Figure 2.11 shows three terms that create a fuzzy partition on the variable *Temperature*; the maximum value for a term is a zero value for all other terms and the sum of membership for each temperature in each term is one. These two properties are generally found in variables that use triangular and trapezoidal membership functions, but are not a requirement. For Gaussian membership functions these properties do not hold.

2.5.2 Fuzzy Hedges

There are two ways to create the terms of a linguistic variable. The first is to simply define them from the beginning. The second method is to modify an existing term and

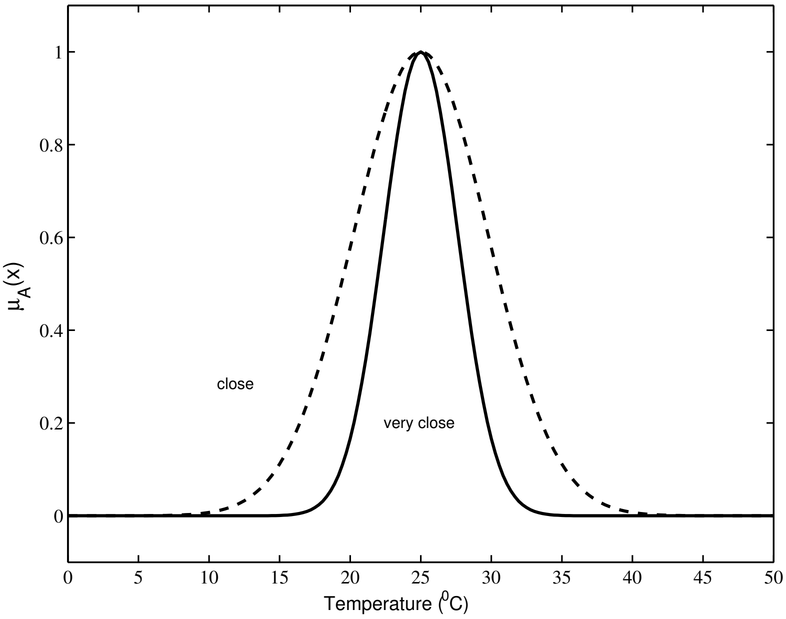


Figure 2.12: Hedge sample: very close to 25, the dash line is the definition for close while the solid line is defined for very close.

is generally referred to as a *hedge* or *linguistic modifier*. For instance, if the term *low* is already defined and the term *very low* is required, it is not necessary to create the definition from scratch, rather the term *low* can be mathematically altered by the hedge *very* to create the new term. This definition would depend on the situation; for example, it could be quite different for temperature than distance. A possible definition for *very* is to square the term being modified, i.e. $very(A) = A^2$.

If A is the fuzzy set *close to 25 °C* from the example in Equation 2.9,

$$\mu_A(x) = \frac{1}{1 + \frac{(x-25)^2}{25}}, \tag{2.9}$$

then the term *very close to 25* would be

$$\mu_{very(A)}(x) = \left(\frac{1}{1 + \frac{(x-25)^2}{25}} \right)^2. \tag{2.10}$$

This is shown graphically in Figure 2.12.

There are many different classes of hedges, such as powered hedges [21] and shifted hedges [22]. Each has its own benefits and should be chosen to suit the task at hand. Hedges are a useful tool in that a wide range of terms can be created in a standard way. This allows an expansion of available terms with little extra effort.

2.5.3 Fuzzy propositions

Once fuzzy variables and terms have been defined, a method which used to connect a variable with an associated term is accomplished using a *fuzzy proposition* representing a statement such as *temperature is cold* where *cold* is a term defined on the universe of discourse of variable *temperature*. The method used to evaluate this proposition is reviewed in Section 2.5.5. Fuzzy propositions are the structure that all fuzzy reasoning is built.

2.5.4 Logical connectives

Fuzzy propositions can be strung together to form more complex statements. When this occurs with propositions on different universes, a relation is formed. Consider the proposition p:

$$P:\ x_1\ \textit{is}\ A_1\ \textbf{and}\ x_2\ \textit{is}\ A_2, \tag{2.11}$$

where A_1 and A_2 have membership functions $\mu_{A_1}(x_1)$ and $\mu_{A_2}(x_2)$, respectively. A fuzzy relation P can represent this proposition with the membership function

$$\mu_P(x_1,x_2) = T(\mu_{A_1}(x_1),\mu_{A_2}(x_2)), \tag{2.12}$$

where T is any general T-norm that represents the *and connective*. A T-conorm is used to implement the *or connective* and must be compatible with the T-norm selected, i.e. they must be T-dual. The three most used operators are from Zadeh [2], Lukasiewicz [16] and the probabilistic functions (in Table 2.5):

Table 2.5: Three most used operators (*and* and *or* connectives) from Zadeh, Lukasiewicz, and Probabilistic

Source	*and*	*or*
Zadeh	$min(a,b)$	$max(a,b)$
Lukasiewicz	$max(a+b-1,0)$	$min(a+b,1)$
Probabilistic	ab	$a+b-ab$

When the propositions are related to the same universe of discourse, the relation reduces to a one dimensional fuzzy set.

2.5.5 Fuzzy rules

Fuzzy rules are used to create the conditional statements that form the backbone of fuzzy logic and approximate reasoning. A simple fuzzy rule has the form

if A **then** B

or using classical logic notation,

$A \rightarrow B$.

Either way, this denotes an implication operation from the *antecedent* or *premise* A to the *consequent* B. This is shorthand for the more specific,

if x **is** A **then** y **is** B,

which is a fuzzy rule made up of two propositions. It is possible to fashion a rule from multiple premises and consequents, resulting in a form similar to

if x_1 **is** A_1 **and** $\ldots x_n$ **is** A_n **then** y_1 **is** B_1 **and** $\ldots y_m$ **is** B_m,

where the connective used is *and* but could be any valid connective operator. More detail can be obtained from the comprehensive text by Zadeh [23]).

A fuzzy rule is by nature a relation between the premise and consequent. As such, a fuzzy rule

if x_1 **is** A_1 **and** x_2 **is** A_2 **then** y **is** B, can be written in another form as the relation R:

$$R = I(T(A_1, A_2), B), \tag{2.13}$$

where T is any T-norm and I is the implication function.

Classical logic allows only one implication function (Table 2.2), but for fuzzy sets and logic an infinite number of implications are possible. Zadeh [23] defined the implication as

$$A \to B = A \times B, \tag{2.14}$$

where $\times$ indicates the Cartesian product of two fuzzy sets,

$$A \times B = \int_{U \times V} T(\mu_A(u), \mu_B(v))/(u, v), \tag{2.15}$$

where U and V are crisp sets and $U \times V$ denotes their Cartesian product,

$$U \times V = (u, v) | u \in U, v \in V. \tag{2.16}$$

Therefore $A \to B$ results in a relation on the Cartesian product of the two universe A and B where each tuple has a membership grade given by the T-norm applied to the given inputs. There are a number of commonly used implication operators, such as S-, QL-, and R-implications. Each of these can be classified as belonging to either of two classes of implications: conjunction and disjunction. A thorough review of fuzzy implications is given by Dubois & Prade [24]. When more than one rule has been defined, an additional relation can be produced that defines an entire *rulebase*. *Aggregation* is the process by which the relation is constructed. If we have n rules and m premises per rule given by [24]

r_1 : if x_1 is $A_{1,1}$ and $\ldots x_m$ is $A_{m,1}$ then y is B_1

r_k : if x_1 is $A_{1,k}$ and $\ldots x_m$ is $A_{m,k}$ then y is B_k

r_n : if x_1 is $A_{1,n}$ and $\ldots x_m$ is $A_{m,n}$ then y is B_n

then the relation for the rulebase can be given as an aggregation,

$$R = \bigcup_k R_k. \tag{2.17}$$

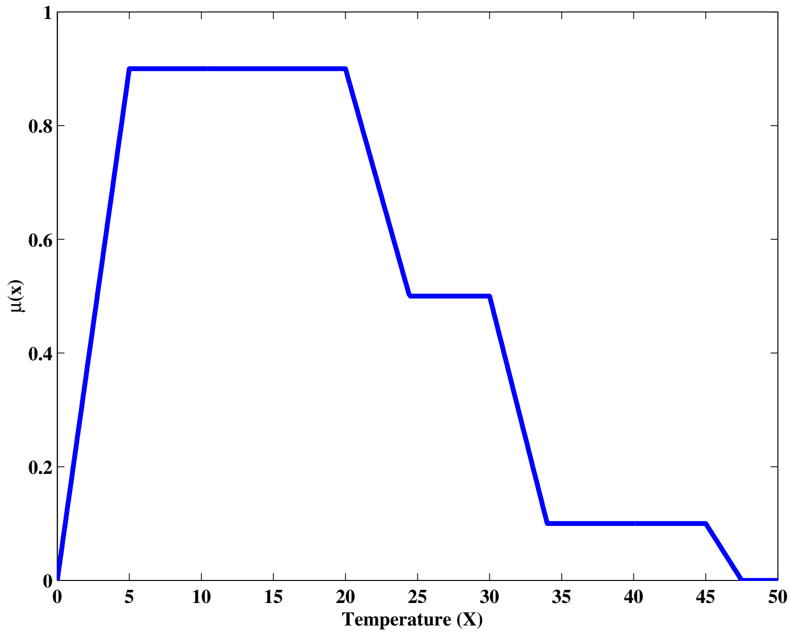


Figure 2.13: The region for fuzzy output which is to be defuzzified.

Here the implication used is a classical conjunction. If the implication function is a disjunction, the relation becomes a conjunction given by

$$R = \bigcap_k R_k. \tag{2.18}$$

Equation 2.17 and 2.18 show that the aggregation of fuzzy rules is dependent on the class of implication used to create the relation.

2.5.6 Defuzzification

The term *defuzzification* is a process to present the output of the reasoning in human understandable form of the fuzzy system. There are two principal classes of defuzzification, *arithmetic defuzzification* and *linguistic approximation*. Suppose we have a region to be defuzzified as shown in Figure 2.13.

The most popular defuzzification method used in applications is *centroid* (or *centre of gravity*), *height*, and *modified height* defuzzification. Centroid returns the centre of area under the curve. If we consider the area as a plate of equal density, the centroid is the point along the x axis about which this shape would balance. The Bisector is the vertical line that will divide the region into two sub-regions of equal area. It is sometimes, but

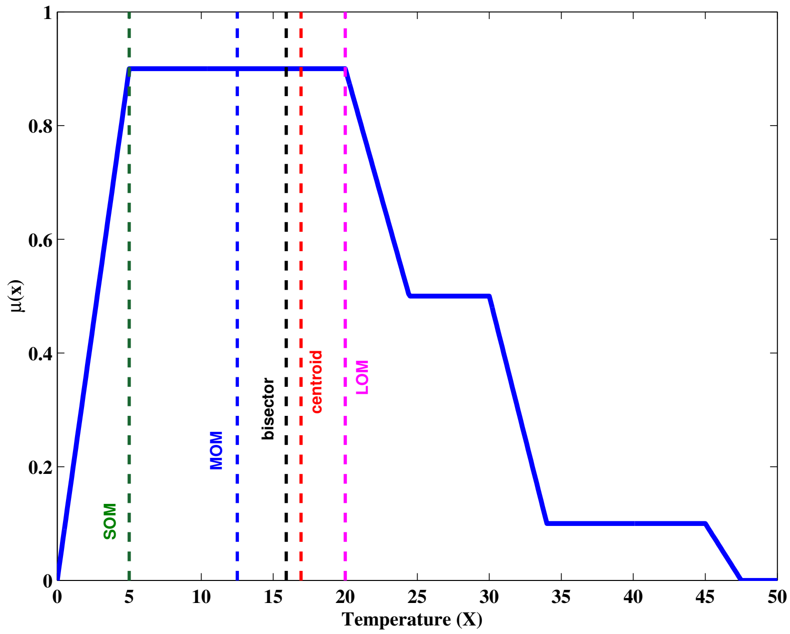


Figure 2.14: The output crisp values defuzzified by SOM, MOM, LOM, bisector, and centroid defuzzification methods.

not always same value as from the centroid. Middle of Maximum (MOM), Smallest of Maximum (SOM), and Largest of Maximum (LOM) key off the maximum value assumed by the aggregate membership function. In Figure 2.13 because there is a plateau at the maximum value, they are distinct. If the aggregate membership function has a unique maximum, then MOM, SOM, and LOM all take the same value. Examples of all of these methods mentioned above are shown in Figure 2.14

2.6 Fuzzy Inferencing Systems (FIS)

Fuzzy inference systems (FIS) are also known as fuzzy models, fuzzy rule-based systems and perhaps the more well known fuzzy controllers. There are two main types of FISs, *Mamdani* and *Takagi & Sugeno*. The Mamdani method [25] expects the output membership functions to be fuzzy sets that results in the possible need for a defuzzification stage in the inference process to convert the fuzzy output into a crisp output. In the Takagi & Sugeno method [26] the output membership functions are singletons, or spikes. This method uses a combination of linear systems to approximate a nonlinear system. The entire input space is broken down into a number of partial fuzzy spaces with the output space represented as a linear equation. The Takagi & Sugeno method greatly reduces the computation required in defuzzification, which increases the efficiency of a fuzzy system.

The fuzzy logic inference process consists of *fuzzifier*, *rules*, *inference engine*, and *defuzzifier* as shown in Figure 2.15. As shown in Figure 2.15, crisp inputs are first fuzzi-

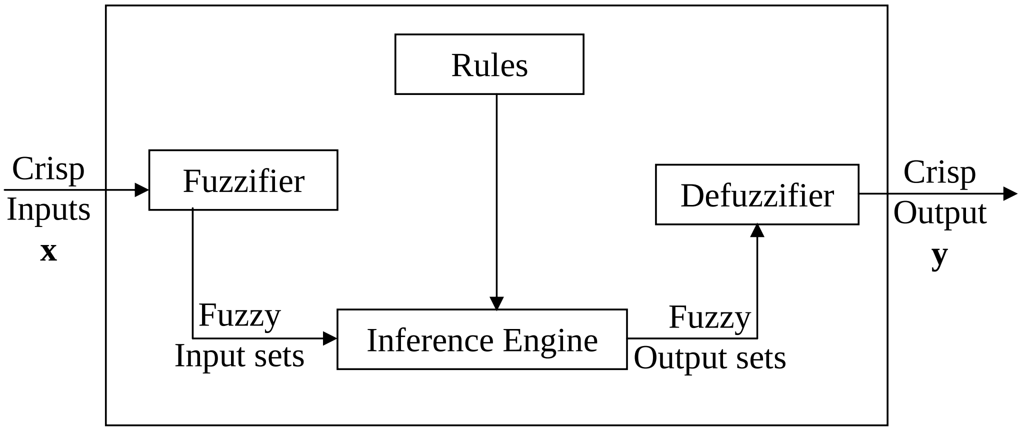


Figure 2.15: Machanism of type-1 fuzzy logic system

fied into fuzzy input sets. In inference engine, fuzzy logic principles are used to combine fuzzy IF-THEN rules from the fuzzy rule base into a mapping from fuzzy input sets to fuzzy output sets. Each rule is interpreted as a fuzzy implication. Finally, a defuzzifier produces a crisp output for fuzzy system from the fuzzy output set(s) from the inference engine.

2.7 Fuzzy Reasoning

Fuzzy reasoning is a body of procedures to represent various assertions about a system of interest, quantify or qualify their validity, and derive the estimation decisions based on imprecise data [23]. It has been used to perform well in situations where classical tools have not, particularly when complexity and imprecision are vital. Fuzzy reasoning is the process of deriving conclusions based on a set of fuzzy rules and given facts. There are at least six different type of fuzzy reasoning and Li [27] provides an analysis of the suitability of various methods.

In mathematical terms, the implication function defines a relation between a given premise and its consequent. What happens when a given premise is different from that which the relation is specifically created for? Consider the following two statements:

$$x \text{ is very fast} \tag{2.19}$$

and

$$\text{if } x \text{ is fast then } y \text{ is large.} \tag{2.20}$$

This leads naturally to the question of how to determine the value for y given the statement 2.19 and the relation in 2.20. Fuzzy and approximate reasoning are built upon such inferences.

An alternative way of conceptualising fuzzy reasoning is as the application of *expert knowledge* to *decision making*. Fuzzy reasoning is based on fuzzy **if-then** rules that contain all the knowledge that is used to make decisions in fuzzy reasoning. For further information, see the thorough review by Turksen [28] and additional material by Zadeh [29] and Castro *et al.* [30].

2.7.1 Compositional rule of inference

The characterization of implication relations between statements about the values of system variables are of particular importance is fuzzy reasoning application. In typical system analysis applications, the relation between two variables, A and B, is expressed by a function f mapping each value x of A into a value y of B. This situation is catered for by the compositional rule of inference, introduced by Zadeh in 1973 [23]. If R is a relation from A to B and x is a fuzzy subset of A, the fuzzy subset y of B that is induced by x is given by

$$y = x \circ R, \tag{2.21}$$

which is the composition of R and x. If the max-min composition is being used, the membership function for y in B is given by

$$\mu_B(y) = max_x(min(\mu_A(x), \mu_R(x,y))). \tag{2.22}$$

2.7.2 Generalised modus ponens

The result of application of the compositional rule of inference to a fuzzy dependence relation between two fuzzy propositions is called the *generalised modus ponens* [31]. The *generalised modus ponens* is a generalised version of the modus ponens from classical logic and is based on an **if-then** construct as follows,

if x is A then y is B

x is $\acute{A}$

---

y is $\acute{B}$

where $\acute{A}$ represents the input data and $\acute{B}$ the inferred result. This implies that given an **if-then** rule and a premise, the outcome can be determined. As a simple example, consider the saying *what goes up must come down* which can be written as *if it goes up then it will come down*. Using modus ponens it is possible to make the observation *it went up* and infer *it will come down*. This inference can be defined by the compositional rule of inference but is not limited to it. Other inference schemes are possible using modus ponens but not the compositional rule of inference as shown by Jager [32]. 'The generalised modus ponens is one of the many inferential procedures that may be employed to derive valid conclusions from valid premises in fuzzy logic. Depending on the particular notion of validity being employed or on the scheme chosen to measure degrees of truth, related expressions may be employed to proceed from premises to conclusions, Furthermore, other valid inferential procedures have been developed generalizing the classical inferential methods known as *generalized modus tollens* and the resolution principle.' stated by E.H. Ruspini *et al* [31]. Further details on fuzzy reasoning can be found in many texts, see for example in *Fuzzy Logic and Fuzzy Reasoning* by [33].

2.8 Type-2 Fuzzy Sets and Systems

As just discussed, fuzzy (type-1) logic was introduced by Lotfi Zadeh in 1965 [2] to resemble human reasoning in its use of approximate information and uncertainty to generate decisions. It was specifically designed to mathematically represent uncertainty and vagueness and provide formalised tools for dealing with the imprecision of many real problems. Since knowledge can be expressed in a more natural by using fuzzy sets, many decision problems can be greatly simplified.

Although many applications have been found for type-1 fuzzy logic, it is its application to rule-based systems that has most significantly shown its importance as a powerful design methodology, but yet it is unable to model and minimize the effects of all uncertainties. Thus, type-2 fuzzy logic should be introduced to handle uncertainties because it can model them and minimize their effects. Figure 2.16 shows a type-2 fuzzy logic system.

The concept of (general) type-2 fuzzy sets is also introduced by Lotfi Zadeh in 1975. Zadeh [34] proposed 'fuzzy sets with fuzzy membership functions' as an extension of the concept of an ordinary, i.e. type-1, fuzzy set and went on to define fuzzy sets of type n, n = 2, 3, ..., for which the membership function ranges over fuzzy sets of type n - 1 [4]. Type-2 fuzzy sets can model uncertainties better and minimize their effects. The use of type-2 sets was advocated and extended by people: Dubois and Prade gave a formula for the composition of type-2 relations as an extension of the type-1 sup-star composition for the minimum t-norm [5], Mizumoto and Tanaka studied the set theoretic operations of type-2 sets and properties of membership degrees of such sets [7] and examined type-2 sets under the operations of algebraic product and algebraic sum [8], etc. However, their use in practice has been limited due to the significant increase in computational complexity involved in their implementation.

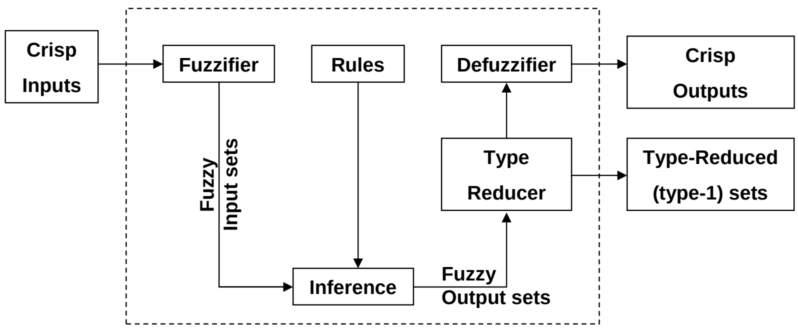


Figure 2.16: Diagram of type-2 fuzzy logic system

Recently, Mendel has established a set of terms to be used when working with type-2 fuzzy sets and, in particular, introduced a concept known as the *footprint of uncertainty* which provides a useful verbal and graphical description of the uncertainty captured by any given type-2 set. Mendel has particularly concentrated on a restricted class of general type-2 fuzzy sets known as *interval type-2 fuzzy sets* [1]. Interval type-2 sets are characterised by having secondary membership functions which only take the values in $\{0, 1\}$. This restriction greatly simplifies the computational requirements involved in performing inference with type-2 sets. Mendel and John developed a simple method to derive union, intersection, and complement, and computational algorithms for type reduction (necessary for type-2 defuzzification) [9].

2.9 General (non-interval) Type-2 Fuzzy Logic Systems

The aim of the remainder of this chapter is to present an overview of the important concepts in type-2 fuzzy logic systems. A type-2 FLS is constructed by the same structure of type-1 IF-THEN rules, which is still dependent on the knowledge of experts. Expert knowledge is always represented by linguistic terms and implied uncertainty, which leads to the rules of type-2 FLSs having uncertain antecedent part and/or consequent part:, which are then translated into uncertain antecedent or consequent MFs. The structure of rules in the type-2 FLS and its inference engine is similar to those in type-1 FLSs [1]. The inference engine combines rules and provides a mapping from input type-2 fuzzy sets to output type-2 fuzzy sets. To achieve this process, we must find unions and intersections of type-2 sets, as well as compositions of type-2 relations. The output of the type-2 inference engine is a type-2 fuzzy set. Using Zadeh's extension principle [34], type-1 defuzzification can derive a crisp output from type-1 fuzzy set; similarly, for a higher type set as type-2, this operation reduces the type-2 fuzzy sets to type-1 fuzzy sets. This process is usually called 'type reduction'. The complete type-2 fuzzy logic theory with the handling of uncertainties, such as the operations on type-2 fuzzy sets, centroid of a type-2 fuzzy sets, type-reduction, and etc., can be found in [9, 35–42].

2.9.1 General Type-2 Fuzzy Sets

Type-2 fuzzy sets were firstly defined by Zadeh [34]. A type-2 fuzzy set is characterized by a fuzzy membership function, i.e. membership value or membership grade for each element of this set is a fuzzy set in [0,1], whereas the membership grade of type-1 fuzzy set is crisp value in [0,1]. Mendel [1] defines the definitions of type-2 fuzzy sets as:

Definition 2.9.1 $\tilde{A}$ denotes a type-2 fuzzy set; $\mu_{\tilde{A}}(x,u)$ denotes the membership function in the type-2 fuzzy set $\tilde{A}$, where $x \in X$ and $u \in J_x \subseteq U = [0,1]$, i.e. [1]

$$\tilde{A} = \{((x,u),\mu_{\tilde{A}}(x,u))|\forall x \in X, \forall u \in J_x \subseteq U = [0,1]\} \tag{2.23}$$

where $0 \leq \mu_{\tilde{A}}(x,u) \leq 1$.

$\tilde{A}$ can also be expressed as

$$\tilde{A} = \int_{x\in X}\int_{u\in J_x} \mu_{\tilde{A}}(x,u)/(x,u) \quad J_x \subseteq U = [0,1] \tag{2.24}$$

where $\int$ denotes union over all admissible x and u. For discrete universes of discourse, use $\sum$ instead of $\int$.

Definition 2.9.2 At each value of x, say $x = x'$, the 2-D plane whose axes are u and $\mu_{\tilde{A}}(x,u)$ is called a vertical slice of $\mu_{\tilde{A}}(x,u)$. A secondary membership function is vertical slice of $\mu_{\tilde{A}}(x,u)$. It is $\mu_{\tilde{A}}(x=x',u)$ for $x \in X$ and $u \in J_x \subseteq U = [0,1]$, i.e. [1]

$$\mu_{\tilde{A}}(x=x',u) = \mu_{\tilde{A}}(x') = \int_{u\in J_{x'}} f_{x'}(u)/u \quad J_x \subseteq U = [0,1] \tag{2.25}$$

Using Equation 2.25, we can also re-express $\tilde{A}$ as a vertical slice manner, i.e.

$$\tilde{A} = \{(x,\mu_{\tilde{A}}(x,u)|\forall x \in X\} \tag{2.26}$$

or,

$$\tilde{A} = \int_{x\in X} \mu_{\tilde{A}}(x)/x = \int_{x\in X}\left[\int_{u\in J_{x'}} f_x(u)/u\right]/x \quad J_x \subseteq U = [0,1]$$

where $\int$ denotes union over all admissible x and u. For discrete universes of discourse, use $\sum$ instead of $\int$ as:

$$\tilde{A} = \sum_{x\in X} \mu_{\tilde{A}}(x)/x = \sum_{x\in X}\left[\sum_{u\in J_x} f_x(u)/u\right]/x \quad J_x \subseteq U = [0,1] \tag{2.27}$$

Definition 2.9.3 The domain of a secondary membership function is called the primary membership grade of x. In Equation 2.27, J_x is the primary membership function of x. where $J_x \subseteq [0,1]$ for $\forall x \in x$ [1].

Again in Equation 2.27, $f_x(u)$ is called secondary membership grade, which is the amplitude of the secondary membership function.

2.9.2 The Footprint of Uncertainty

The use of type-2 fuzzy sets in practice has been limited due to the significant increase in computational complexity involved in their implementation. Mendel introduced an important concept related to uncertainty, called the term 'footprint of uncertainty' (FOU), which consists of a bounded region with uncertainty in the primary membership of a type-2 fuzzy set [1]. The FOU is the union of all primary memberships. i.e.,

$$FOU(\tilde{A}) = \bigcup_{x \in X} J_x \tag{2.28}$$

The term FOU is very useful and was introduced to provide a very convenient verbal description of the entire domain of support for all the secondary grades of a type-2 membership function.

As examples of a FOU is the shaded region in Figure 2.17 and 2.18. The FOU is shaded uniformly to indicate that it is for an interval type-2 fuzzy set; thus, a uniformly shaded FOU also represents the entire interval type-2 fuzzy set. Figure 2.17 represents the FOU for Gaussian primary membership function with uncertain centre point, while Figure 2.18 represents the FOU for Gaussian primary membership function with uncertain standard deviation.

2.9.3 Embedded Type-1 and Embedded Type-2 Fuzzy Sets

Besides the FOU, there are two other important concepts to illustrate how to construct type-2 fuzzy sets with embedded type-1 and type-2 fuzzy sets. Mendel [1] uses these to help understand why it is so complicated to use type-2 fuzzy sets. A type-2 fuzzy set $\tilde{A}$ can be considered as a collection of type-2 fuzzy sets $\tilde{A}_e$, where $\tilde{A}_e$, is also called embedded type-2 fuzzy set in $\tilde{A}$. Furthermore, an embedded type-1 set A_e can be thought of as the union of all primary memberships of set $\tilde{A}_e$.

For continuous universe of discourse X and U,

- Embedded type-1 fuzzy set A_e, is [1]

$$A_e = \int_{x \in X} \theta / x \quad \theta \in J_x \subseteq U = [0,1] \tag{2.29}$$

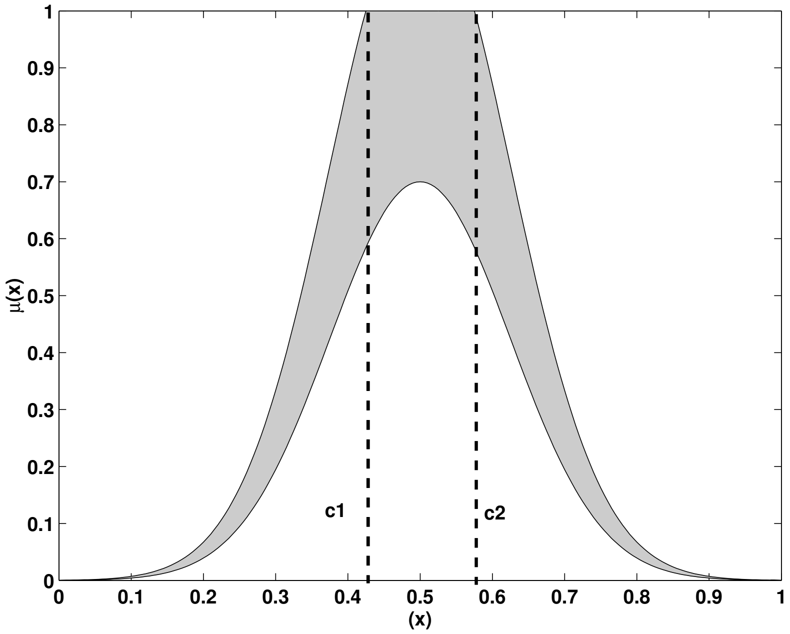


Figure 2.17: Footprint of Uncertainty for Gaussian primary membership function with uncertain centre point.

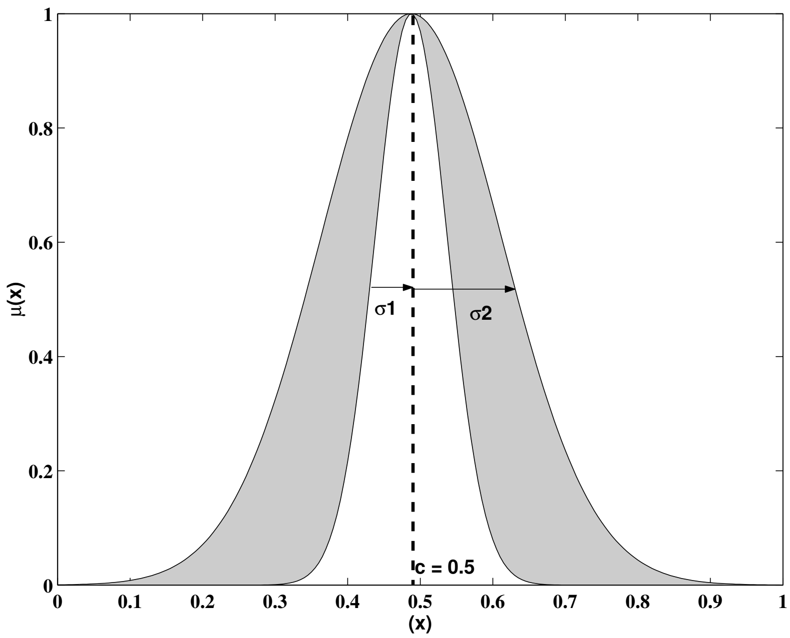


Figure 2.18: Footprint of Uncertainty for Gaussian primary membership function with uncertain standard deviation.

- Embedded type-2 fuzzy set $\tilde{A}_e$, is [1]

$$\tilde{A}_e = \int_{x\in X} [f_x(\theta)/\theta]/x \quad \theta \in J_x \subseteq U = [0,1] \tag{2.30}$$

where only one primary membership $\theta \in J_x$ of $\tilde{A}_e$ each x has an associated secondary membership function $f_x(\theta)$.

For continuous domains, both A_e, and $\tilde{A}_e$ are an uncountable number. For discrete universes of discourse X and U, use $\sum$ instead of $\int$. According to embedded type-1 and type-2 definitions, there exist $\sqcap_{i=1}^{N} M_i A_e$ and $\sqcap_{i=1}^{N} M_i \tilde{A}_e$, respectively.

2.9.4 Inference Process of General Type-2 Fuzzy Logic Systems

The structure of the l^{th} type-2 rule in a general type-2 system is: [1]

$$R^l : IF\ x_1\ is\ \tilde{F}_1^l\ and\ x_2\ is\ \tilde{F}_2^l,\ and,\ \cdots,\ and\ x_n\ is\ \tilde{F}_n^l\ THEN\ y\ is\ \tilde{G}^l \tag{2.31}$$

where $l = 1,2,\cdots,M$ and type-2 fuzzy relation R^l can be expressed by membership function as: [1]

$$\mu_{R^l}(\overset{x}{\rightarrow}, y) = \mu_{\tilde{F}_1^l \times \tilde{F}_2^l \times \ldots \tilde{F}_n^l \rightarrow \tilde{G}^l}(\overset{x}{\rightarrow}, y)$$

$$\mu_{R^l}(\overset{x}{\rightarrow}, y) = \mu_{\tilde{F}_1^l}(x_1) \sqcap \cdots \sqcap \mu_{\tilde{F}_n^l}(x_n) \sqcap \mu_{\tilde{G}^l}(y)$$

$$\mu_{R^l}(\overset{x}{\rightarrow}, y) = \left[\sqcap_{k=1}^{n} \mu_{\tilde{F}_k^l}(x_k\right] \sqcap \mu_{\tilde{G}^l}(y) \tag{2.32}$$

where $\sqcap$ denotes operation, whereas join and meet operations denoting by $\sqcup$ and $\sqcap$ will be used in equations 2.35 and 2.36, respectively. They are defined and explained in detail in [36, 43]. Type-2 union and intersection operations with their related join and meet operations are briefly explained as follows. Let two type-2 fuzzy sets $\tilde{A}$ and $\tilde{B}$ are: [1]

$$\tilde{A} = \int_{x\in X} \mu_{\tilde{A}}(x)/x = \int_{x\in X} \left[\int_{u\in J_x^u} f_x(u)/u\right]/x \quad J_x^u \subseteq [0,1] \tag{2.33}$$

and

$$\tilde{B} = \int_{x\in X} \mu_{\tilde{B}}(x)/x = \int_{x\in X} \left[\int_{u\in J_x^u} g_x(u)/u\right]/x \quad J_x^u \subseteq [0,1] \tag{2.34}$$

The union of secondary membership functions of $\tilde{A}$ and $\tilde{B}$ are given by [1]

$$\mu_{\tilde{A}\cup\tilde{B}}(x) = \int_{u\in J_x^u}\int_{w\in J_x^w} f_x(u)\bullet g_x(w)/(u\vee w) = \mu_{\tilde{A}}(x)\sqcup\mu_{\tilde{B}}(x), \quad x\in X \tag{2.35}$$

where $\vee$ means maximum, and $\bullet$ means minimum or product t-norm. The intersection of secondary membership function of $\tilde{A}$ and $\tilde{B}$, is: [1]

$$\mu_{\tilde{A}\cap\tilde{B}}(x) = \int_{u\in J_x^u}\int_{w\in J_x^w} f_x(u)\bullet g_x(w)/(u\wedge w) = \mu_{\tilde{A}}(x)\sqcap\mu_{\tilde{B}}(x), \quad x\in X \tag{2.36}$$

where $\wedge$ means any kind of t-norm.

According to the two operations stated above, *meet* and *join* should be used between two secondary membership function, i.e. $\mu_{\tilde{A}}(x)$ an $\mu_{\tilde{B}}(x)$; whereas $u\vee w$ or $u\wedge w$ must be computed between every possible pair of primary membership functions u and w, where $u\in J_x^u$ and $w\in J_x^w$.

Also the secondary membership of $\mu_{\tilde{A}\cup\tilde{B}}(x)$ or $\mu_{\tilde{A}\cap\tilde{B}}(x)$ must be computed as the t-norm operation between the corresponding secondary memberships of $\mu_{\tilde{A}}(x)$ and $\mu_{\tilde{B}}(x)$, $f_x(u)$ and $g_x(w)$, respectively.

The n-dimensional type-2 input fuzzy set $\mu_{\tilde{A}_x}$ whose membership function is [1]

$$\mu_{\tilde{A}_x} = \mu_{\tilde{X}_1}(x_1)\sqcap\cdots\sqcap\mu_{\tilde{X}_n}(x_n) = \sqcap_{k=1}^{n}\mu_{\tilde{X}_k}(x_k) \tag{2.37}$$

where $\tilde{X}_i \cdots i = 1,\cdots,n$ are the fuzzy inputs.

The output $\mu\tilde{B}^l$ of the type-2 fuzzy set can be derived from $\tilde{B}^l = \tilde{A}_x\circ R^l$, such that [1]

$$\mu_{\tilde{B}^l}(y) = \mu_{\tilde{A}_x\circ R^l}(y) = \sqcup_{x\in X}\left[\mu_{\tilde{A}_x}(\overset{x}{\rightarrow})\sqcap\mu_{R^l}(\overset{x}{\rightarrow},y)\right] \tag{2.38}$$

where $y\in Y,\ l = 1,\cdots,M.$

By substituting Equations 2.32 and 2.37 into 2.38, it can be shown that [1]

$$\mu_{\tilde{B}^l}(y) = \mu_{\tilde{A}_x\circ R^l}(y) = \sqcup_{x\in X}\left[\mu_{\tilde{A}_x}(\overset{x}{\rightarrow})\sqcap\mu_{R^l}(\overset{x}{\rightarrow},y)\right] \tag{2.39}$$

where $y\in Y$. Let

$$\mu_{\tilde{Q}_k^l}(x_k) = \mu_{\tilde{X}_k}\sqcap\mu_{\tilde{F}_k}(x_k). \tag{2.40}$$

Then,

$$\mu_{\tilde{B}^l}(y) = \mu_{\tilde{G}^l}(y)\sqcap\left\{\sqcup_{\overset{x}{\rightarrow}\in X}\left[\sqcap_{k=1}^{n}\mu_{\tilde{Q}_k^l}(x_k)\right]\right\}, \quad y\in Y. \tag{2.41}$$

Then let

$$F^l = \sqcup_{\underset{\rightarrow}{x} \in X} \left[\sqcap_{k=1}^{n} \mu_{\tilde{Q}_k^l}(x_k) \right] \tag{2.42}$$

so that

$$\mu_{\tilde{B}^l}(y) = \mu_{\tilde{G}^l}(y) \sqcap F^l, \quad y \in Y \tag{2.43}$$

Similar to a type-1 fuzzy logic system, F^l is also referred to as the firing strength. For singleton input, each $\mu_{\tilde{X}_i}(x_i)$ is non-zero only at one point, $x_i = \acute{x}_i$, so that equation 2.43 can be expressed as [1]

$$\mu_{\tilde{B}^l}(y) = \mu_{\tilde{G}^l}(y) \sqcap \left\{ \left[\sqcup \mu_{\tilde{X}_1}(\acute{x}_1) \sqcap \mu_{\tilde{F}_1}(\acute{x}_1) \right] \sqcap \cdots \sqcap \left[\sqcup \mu_{\tilde{X}_n}(\acute{x}_n) \sqcap \mu_{\tilde{F}_n}(\acute{x}_n) \right] \right\}$$

$$\mu_{\tilde{B}^l}(y) = \mu_{\tilde{G}^l}(y) \sqcap \left\{ \left[(1/1) \sqcap \mu_{\tilde{F}_1}(\acute{x}_1) \right] \sqcap \cdots \sqcap \left[(1/1) \sqcap \mu_{\tilde{F}_n}(\acute{x}_n) \right] \right\}$$

$$\mu_{\tilde{B}^l}(y) = \mu_{\tilde{G}^l}(y) \sqcap \left[\sqcap_{k=1}^{n} \mu_{\tilde{F}_1}(\acute{x}_1) \right], \quad y \in Y \tag{2.44}$$

In practice, a fuzzy logic system is a type-2 system as soon as at least one of its antecedent or consequent sets is a type-2 fuzzy set. Even a fuzzy logic system whose type-2 rules are activated by type-2 input is also called a type-2 fuzzy logic system.

2.9.5 Type Reduction and Defuzzification in General Type-2 Fuzzy Logic Systems

The defuzzifier of a type-1 FLS combines all fired output sets in some method to derive a crisp output result to represent the combined output set [6, 37]. For type-1 defuzzification methods, all the antecedent and consequent sets are type-1 sets; whereas for the type-reduction methods for type-2 FLS in Figure 2 some or all of the antecedent and consequent sets are type-2 fuzzy sets. The output set corresponding to each rule of the general type-2 FLS is a type-2 fuzzy set. Similar to the defuzzifier of type-1 FLS, type-2 FLS performs a centroid computation on all these output type-2 sets. The results of this process obtain a type-2 fuzzy set that is called the ‘type-reduced’ set. Consequently, each element of the type-reduced set can be taken as the centroid of some type-1 set embedded in the output set of the type-2 FLS. According to this concept, each of these type-1 embedded sets can be thought of many different type-1 FLSs. Each of such type-1 FLSs is embedded in the type-2 FLS, so the type-reduced set is a collection of the outputs of all type-1 FLSs embedded in the type-2 FLSs. As a result, using a fuzzy set to represent the output of the type-2 FLS is rather more complex than using a crisp number, i.e. the type-reduced set (type-1) many possesses more important information than a single crisp number.

The type-reduced output also can be interpreted as providing a measure of spread about the defuzzified output (i.e. crisp output) and can be thought of as a linguistic confidence interval, i.e. this confidence can be an interval supporting beliefs of different experts. Due to uncertainties in the type-2 membership function, the type-reduced set of the type-2 FLS can then be thought of as representing the uncertainty in the crisp output. Some measure of the spread of the type-reduced set may be taken to indicate the possible variation to the crisp output due to variations in the membership function parameters. Finally, the type-reduced set can be defuzzified to get a crisp output from the type-2 FLS that is called defuzzification in type-2 FLS.

To derive its centroid C_A from a type-1 set A in a discrete domain can be described as: [1]

$$C_A = \frac{\sum_{i=1}^{N} x_i u_A(x_i)}{\sum_{i=1}^{N} u_A(x_i)} \tag{2.45}$$

Similarly, to derive the centroid $C_{\tilde{A}}$ of type-2 fuzzy set $\tilde{A} = \{(x, \mu_{\tilde{A}}(x,u) | \forall x \in X\}$ whose x domain is also discretized into N points as: [1]

$$\tilde{A} = \sum_{i=1}^{N} \left[\int_{u \in J_{x_i}} f_{x_i}(u)/u \right] / x_i \tag{2.46}$$

Then $C_{\tilde{A}}$ can be a type-1 fuzzy set defined as follows: [1]

$$C_{\tilde{A}} = \int_{\theta_1 \in J_{x_1}} \cdots \int_{\theta_N \in J_{x_N}} [f_{x_1}(\theta_1) * \ldots * f_{x_N}(\theta_N)] \Big/ \frac{\sum_{i=1}^{N} x_i \theta_i}{\sum_{i=1}^{N} x_i} \tag{2.47}$$

From the definition of an embedded type-2 fuzzy set in Equation 2.30, every combination of $\theta_1, \cdots, \theta_n$ and its associated secondary grade $f_{x_1}(\theta_1) * \cdots * f_{x_1}(\theta_n)$ forms an embedded type-2 set $\tilde{A}_e$ in Equation 2.46. Each element of $C_{\tilde{A}}$ is determined by computing the centroid of the embedded type-1 fuzzy set A_e that is associated with $\tilde{A}_e$ and computing the t-norm of the secondary membership functions with $\theta_1, \cdots, \theta_n$, namely

$f_{x_1}(\theta_1) * \cdots * f_{x_1}(\theta_n)$. [1]

$$C_{\tilde{A}} = \frac{\sum_{i=1}^{N} x_i \theta_i}{\sum_{i}^{N} \theta_i} \tag{2.48}$$

Therefore by computing all this for all embedded type-2 sets in $\tilde{A}$, the complete centroid $C_{\tilde{A}}$ can be derived [1]. A practical sequence of computations to derive $C_{\tilde{A}}$ is as follows:

1. Discretise the x-domain into N points $x_1, \cdots, x_N$.
2. Discretise each J_{x_j} (the primary memberships of x_j) into a reasonable number of points, M_j, where $j = 1, \cdots, N$.
3. Enumerate all embedded type-1 sets; there will be $\sqcap_{j=1}^{N} M_j$ of them to compute $C_{\tilde{A}}$.

The computation in equation 2.45 of centroid of type-1 fuzzy set can be re-stated as a general form (as shown in Equation 2.49), where $v_l \in \Re$ (real numbers), and $w_i \in [0,1]$ for $l = 1, \cdots, N$ [1].

$$y(v_1, \cdots, v_N\, w_1, \cdots, w_N) = \frac{\sum_{l=1}^{N} v_l w_l}{\sum_{l}^{N} w_l} \tag{2.49}$$

For most of type-1 defuzzification, w_l becomes a type-2 set and v_l is crisp number, and Equation 2.49 will be no problem to fit. Nevertheless, the centre-of-sets defuzzifier of type-1 extends to centre-of-sets type reduction of type-2 set that requires both v_l and w_l to become type-1 sets. Then the general form for computing this centroid is called a generalised centroid [6, 38], and it is essential knowledge for type-2 to interpret type reduction. The general centroid (GC) is

$$GC = \int_{v_l \in V_l} \cdots \int_{v_N \in V_N} \int_{w_l \in W_l} \cdots \int_{w_N \in W_N} \left[T_{l=1}^{N} \mu_{v_l}(v_l) * T_{l=1}^{N} \mu_{w_l}(w_l) \right] / \frac{\sum_{l=1}^{N} v_l w_l}{\sum_{l}^{N} w_l} \tag{2.50}$$

where T is short from of t-norm and $*$ is t-norm operator.

Compared to the practical sequence of centroid computations in Equation 2.45, GC will be more complex in computing. It needs to discretise both v_l and w_l to the suitable number of points, M_l and N_l, respectively. In total, the number of computations will be $\sqcap_{j=1}^{N} M_j N_j$. There are four major type reduction methods which are described as follows:

- **Centroid type reduction** : Similar to the centroid defuzzifier to type-1, the union of type-2 fuzzy sets in equation 2.35 firstly requires computing the join of their secondary membership functions; i.e. to compute the secondary membership function $\mu_{\tilde{B}}(y)$ from $\tilde{B} = \bigcup_{l=1}^{M} \tilde{B}^l$, as [1]:

$$\mu_{\tilde{B}}(y) = \sqcup_{l=1}^{M} \mu_{\tilde{B}}(y), \quad \forall y \in Y \tag{2.51}$$

where $\mu_{\tilde{B}}(y)$ is the secondary membership function for the l^{th} rule, and it depends on many factors such as join, meet, and embedded sets in equations 2.29 and 2.30, respectively. The centroid type reduction calculates the centroid of $\tilde{B}$. Then extension from type-1 centroid defuzzifier to type-2 centroid type reduction can be shown as:

$$\mu_{\tilde{B}}(y) = \sqcup_{l=1}^{M} \mu_{\tilde{B}}(y), \quad \forall y \in Y \tag{2.52}$$

where $i = 1, \cdots, N$. For different fuzzy logic system inputs, different values of $y_c(\overset{x}{\rightarrow})$ will be derived. Similarly the sequence to compute this process, the y-domain is discretized into N points $y_1, \cdots, y_N$ and then J_{y_i}, is discretized into a suitable number of points $M_i (i = 1, \cdots, N)$. The total number of computations is $\sqcap_{i=1}^{N} M_i$. However, this process needs to compute $\mu_{\tilde{B}}(y)$ firstly (i.e. combined from all output sets to form one $\tilde{B}$) that is high computationally intensively. Mendel [1] notes that the centroid type-reduction here must use minimum t-norm to perform.

- **Height type reduction**: The extension from type-1 height defuzzifier to type-2 height type reduction can be described as [1]

$$\mu_{\tilde{B}}(y) = \sqcup_{l=1}^{M} \mu_{\tilde{B}}(y), \quad \forall y \in Y \tag{2.53}$$

where $l = 1, \cdots, M$. The $\bar{y}^l$ is the point having maximum membership in the l^{th} output set and θ_l, $J_{\bar{y}^l}$ and $f_{\bar{y}^l} (\forall l)$ are associated with $\mu_{\tilde{B}^l}(\bar{y}^l)$. The sequence to obtain $y_h(\overset{x}{\rightarrow})$ is firstly to choose $\bar{y}^l$ from each rule output, then discretise the primary membership of each $\mu_{\tilde{B}^l}(\bar{y}^l)$ into a suitable number of points M_l where $l = 1, \cdots, M$, i.e. rule number. In total there will be $\sqcap_{l=1}^{N} M_l$ computations. Compared to centroid type reduction, the difference is that the discretised number of points on the horizontal axis uses the number of rules M instead of N.

- **Modified height type reduction**: The extension from type-1 modified height defuzzifier to type-2 modified height type reduction can be shown as: [1]

$$\mu_{\tilde{B}}(y) = \sqcup_{l=1}^{M} \mu_{\tilde{B}}(y), \quad \forall y \in Y \tag{2.54}$$

where all symbols denote the same things as in Equation 2.53. The only difference between the modified height type-reduction and the height type-reduction is that each output set secondary membership function, $\mu_{\tilde{B}^l}(\bar{y}^l)$, in the modified height type-reduction is scaled by $1/(\delta^l)^2$.

- **Centre-of-Sets type reduction**: Similar to center-of-sets defuzzified of type-1 fuzzy logic system, the extension to type-2 center-of-sets type reduction needs to replace each type-2 consequent set, $\bar{G}^l$, by its centroid, $C_{\tilde{G}^l}$ (a type-1 set); and finds a weighted average of these centroids. The firing strength corresponding to the l^{th} rule is $\sqcap_{i=1}^{N} \mu_{\bar{F}_l}(x_i)$, indicated by W_l, i.e. using meet operation for type-2 to replace $T_{i=1}^{n} \mu_{F_i^l}(x_i)$ of type-1 center-of-sets defuzzified. W_l is also a type-1 fuzzy set. Then the center-of-sets centroid can be described by a generalized centroid expression as: [1]

$$y_{cos}(\overset{x}{\rightarrow}) = \int_{v_l \in C_{\tilde{G}1}} \cdots \int_{v_M \in C_{\tilde{G}M}} \int_{w_1 \in W_1} \cdots \int_{w_M \in W_M}$$

$$\left[T_{l=1}^{M} \mu_{vC_{\tilde{G}1}}(v_l) * T_{l=1}^{M} \mu_{w_l}(w_l) \right] / \frac{\sum_{l=1}^{M} v_l w_l}{\sum_{l}^{M} w_l} \tag{2.55}$$

To obtain $y_{cos}(\overset{x}{\rightarrow})$, a practical sequence is described as below: [1]

1. Discretise its output space Y and compute its centroid $C_{\tilde{G}^l}$ for each consequent using Equation 2.47.
2. Compute the firing strength W_l for each rule.
3. Discretise the domain of each type-1 fuzzy set $C_{\tilde{G}^l}$ and W_l into a suitable number of points as N_l and $M_l (l = 1, \cdots, M)$, respectively.
4. Enumerate all the possible combinations. The total number of combinations will be $\sqcap_{l=1}^{M} M_l N_l$.
5. Compute the centre-of-sets type reduction using Equation 2.55 with M, $C_{\tilde{G}^l}$, and W_l.

2.10 Interval Type-2 Fuzzy Logic Systems

According to the process of Equations 2.35 and 2.36, the operations of general type-2 fuzzy sets become computationally intensive due to the necessity to perform every pair of type-2 fuzzy sets. Especially, the number of its embedded type-1 fuzzy sets will be massive while it is a continuous universe of discourse. In [1], [39], and [44], it is suggested that there is good reason to implement type-2 fuzzy logic systems by using *interval* type-2 sets, for which it is relatively easy to compute meet and join operations and perform type-reduction. It also distributes the uncertainty evenly among all acceptable primary memberships.

2.10.1 Interval Type-2 Fuzzy Sets

When $f_x(u) = 1, \forall u \in J_x \subseteq [0,1]$ in equation 2.27, then the secondary membership functions are interval sets such that $\mu_{\tilde{A}}(x)$ can be called an interval type-2 membership function [1]. Therefore the type-2 fuzzy set $\tilde{A}$ can be shown as: [35]

$$\tilde{A} = \int_{x\in X} \mu_{\tilde{A}}(x)/x = \int_{x\in X} \left[\int_{u\in J_x} 1/u\right] /x \quad J_x \subseteq U = [0,1] \tag{2.56}$$

A Gaussian primary membership function with uncertain mean and fixed standard deviation having an interval type-2 secondary membership function can be called an interval type-2 Gaussian membership function as in equation 2.57. The interval type-2 Gaussian membership function with an uncertain *mean* in $[m_1, m_2]$ and fixed *standard deviation*, σ is: [35]

$$\mu_{\tilde{A}}(x) = exp\left[-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2\right], \quad m \in [m_1, m_2] \tag{2.57}$$

It is obvious that the type-2 fuzzy set is in a region, called the footprint of uncertainty (FOU), and bounded by an upper and lower membership function [44], which are denoted as $\overline{\mu}_{\tilde{A}}(x)$ and $\underline{\mu}_{\tilde{A}}(x)$, respectively. Both of these are type-1 membership functions. Hence equation 2.56 can be re-expressed as: [35]

$$\tilde{A} = \int_{x\in X} \left[\int_{u\in[\underline{\mu}_{\tilde{A}}(x),\overline{\mu}_{\tilde{A}}(x)]} 1/u\right] /x \tag{2.58}$$

2.10.2 Meet and Join for Interval Type-2 Fuzzy Sets

From general type-2 fuzzy sets in equations 2.33 and 2.34, let $\tilde{A}$ and $\tilde{B}$ be two interval sets $F = \int_{u\in F} 1/u$ and $F = \int_{w\in F} 1/w$, respectively, with domains $u \in [l_f, r_f] \subseteq [0,1]$, and $w \in [l_g, r_g] \subseteq [0,1]$. The meet between F and G is $Q = F \sqcap G = \int_{q\in Q} 1/q$. Equation 2.36 can be re-stated for an interval type-2 fuzzy set as: [35]

$$Q = F \sqcap G = \int_{q\in\left[l_f * l_g, rf * r_g\right]} 1/q \tag{2.59}$$

where $*$ denotes a t-norm. The *join* between F and G is $Q = F \sqcup G = \int_{q\in Q} 1/q$. Equation 2.35 can be re-expressed by interval type-2 fuzzy set as: [35]

$$Q = F \sqcup G = \int_{q\in\left[l_f * l_g, rf * r_g\right]} 1/q \tag{2.60}$$

From equations 2.59 and 2.60, the *meet* and *join* operation of interval sets are determined just by the two end-points of each interval set, i.e. $[l_f, r_f]$ and $[l_g, r_g]$. Also, the two end-points are associated with type-1 membership functions referred to as upper and lower membership functions.

2.10.3 Lower and Upper Membership Functions for Interval Type-2 Fuzzy Sets

The upper membership function is a subset that has the maximum membership grade of the FOU and the lower membership function is a subset that has the minimum membership grade of the FOU. For interval type-2 sets, the $\mu_{\tilde{Q}_k^l}(x_k)$ of equation 2.40 can be re-described as its upper and lower membership functions as: [35]

$$\mu_{\tilde{Q}_k^l}(x_k) = \int_{q^l\in\left[\underline{\mu}_{\tilde{Q}_k^l}(x_k), \overline{\mu}_{\tilde{Q}_k^l}(x_k)\right]} 1/q \tag{2.61}$$

where $\underline{\mu}_{\tilde{Q}_k^l}(x_k)$ denotes the lower membership function, and $\overline{\mu}_{\tilde{Q}_k^l}(x_k)$ denotes the upper membership function, as in equations 2.62 and 2.63, respectively [35].

$$\underline{\mu}_{\tilde{Q}_k^l}(x_k) = \int_{X_k} \left[\underline{\mu}_{\tilde{X}_k}(x_k) * \underline{\mu}_{\tilde{F}_k}(x_k)\right] / x_k \tag{2.62}$$

$$\overline{\mu}_{\tilde{Q}_k^l}(x_k) = \int_{X_k} \left[\overline{\mu}_{\tilde{X}_k}(x_k) * \overline{\mu}_{\tilde{F}_k}(x_k)\right] / x_k \tag{2.63}$$

In the same way, the $\mu_{\tilde{X}_k}(x_k)$ and $\mu_{\tilde{F}_k}(x_k)$ of equation 2.40 can also be re-stated as its upper and lower membership functions, respectively, such as: [35]

$$\mu_{\tilde{F}_k^l}(x_k) = \int_{u^l \in \left[\underline{\mu}_{\tilde{X}_k^l}(x_k), \overline{\mu}_{\tilde{F}_k^l}(x_k)\right]} 1/u^l \tag{2.64}$$

and

$$\mu_{\tilde{F}_k^l}(x_k) = \int_{w^l \in \left[\underline{\mu}_{\tilde{F}_k^l}(x_k), \overline{\mu}_{\tilde{F}_k^l}(x_k)\right]} 1/w^l \tag{2.65}$$

2.10.4 Inference Process of Interval Type-2 Fuzzy Logic Systems

The meet operation in Equation 2.42 just involves the t-norm operation between the points in two upper and lower membership functions, $\underline{\mu}_{\tilde{Q}_k^l}(x_k)$ and $\overline{\mu}_{\tilde{Q}_k^l}(x_k)$, i.e. equation 2.62 and 2.63. For all points $x \in X_k$, $k = 1, \cdots, n$, the result can be shown as $\underline{\mu}_{\tilde{Q}_k^l}(\overrightarrow{x})$ and $\overline{\mu}_{\tilde{Q}_k^l}(\overrightarrow{x})$, $\overrightarrow{x}$ is vector for all points. The join operation in Equation 2.42 leads to join the result from the meet operation above using the maximum value. The result F^l can be an interval type-1 set [3] as following:

$$F^l = \sqcup_{\overrightarrow{x} \in X} \left[\sqcap_{k=1}^{n} \mu_{\tilde{F}_k^l}(x_k)\right] = \left[\underline{f}^l, \overline{f}^l\right] \tag{2.66}$$

where

$$\underline{f}^l = \int_{X_1} \cdots \int_{X_n} \left[\underline{\mu}_{\tilde{X}_1}(x_1) * \underline{\mu}_{\tilde{F}_1^l}(x_1)\right] * \vdots * \left[\underline{\mu}_{\tilde{X}_n}(x_n) * \underline{\mu}_{\tilde{F}_n^l}(x_n)\right] / \overrightarrow{x} \tag{2.67}$$

and

$$\overline{f}^l = \int_{X_1} \cdots \int_{X_n} \left[\overline{\mu}_{\tilde{X}_1}(x_1) * \overline{\mu}_{\tilde{F}_1^l}(x_1)\right] * \vdots * \left[\overline{\mu}_{\tilde{X}_n}(x_n) * \overline{\mu}_{\tilde{F}_n^l}(x_n)\right] / \overrightarrow{x} \tag{2.68}$$

Consequently, for inference of interval type-2 fuzzy logic systems using the relation R^l to fire consequent fuzzy sets, the output result set $\mu_{\tilde{B}^l}(y)$ in equation 2.43 can be derived as: [35]

$$\mu_{\tilde{B}^l}(y) = \mu_{\tilde{G}^l}(y) \sqcap F^l = \mu_{\tilde{G}^l}(y) \sqcap \int_{f^l \in \left[\underline{f}^l, \overline{f}^l\right]} 1/f^l$$

$$\mu_{\tilde{B}^l}(y) = \int_{b^l \in \left[\underline{f}^l * \underline{\mu}_{\tilde{G}^l}(y), \overline{f}^l * \overline{\mu}_{\tilde{G}^l}(y)\right]} \tag{2.69}$$

where $\underline{\mu}_{\tilde{G}^l}(y)$ and $\overline{\mu}_{\tilde{G}^l}(y)$ are the lower and upper membership grades of $\mu_{\tilde{G}^l}(y)$. According to equation 2.60 the *join* of these n interval output sets can be obtained straightforwardly as: [35]

$$\mu_{\tilde{B}^l}(y) = \int_{b\in\left[\left[\underline{f}^1 * \underline{\mu}_{\tilde{G}^l}(y)\right]\vee\cdots\vee\left[\underline{f}^N * \underline{\mu}_{\tilde{G}^N}(y)\right],\left[\overline{f}^1 * \overline{\mu}_{\tilde{G}^l}(y)\right]\vee\cdots\vee\left[\overline{f}^N * \overline{\mu}_{\tilde{G}^N}(y)\right]\right]} 1/b. \tag{2.70}$$

For a singleton input, $\underline{f}^l$ and $\overline{f}^l$ in Equations 2.67 and 2.68 can be simplified as: [35]

$$\underline{f}^l = \underline{\mu}_{\tilde{F}_1^l}(x_1) * \cdots * \underline{\mu}_{\tilde{F}_n^l}(x_n) \tag{2.71}$$

and

$$\overline{f}^l = \overline{\mu}_{\tilde{F}_1^l}(x_1) * \cdots * \overline{\mu}_{\tilde{F}_n^l}(x_n) \tag{2.72}$$

2.10.5 Type Reduction for Interval Type-2 Fuzzy Logic Systems

- **Centre-of-Sets type reduction**: For Gaussian interval type-2 fuzzy sets, the upper membership function is a subset that has the maximum membership grade and the lower membership function is a subset that has the minimum membership grade. The join operation leads to join the result from meet operations using the supremum (i.e., maximum value), the result F^i can be interval type-1 set as follows: [35]

  $$F^i = \left[\underline{f}^i, \overline{f}^i\right] \tag{2.73}$$

  where $\underline{f}^i = \underline{\mu}_{\tilde{F}_1^l}(x_1) * \cdots * \underline{\mu}_{\tilde{F}_n^l}(x_n)$ and $\overline{f}^l = \overline{\mu}_{\tilde{F}_1^l}(x_1) * \cdots * \overline{\mu}_{\tilde{F}_n^l}(x_n)$

  In order to simplify the notation, we consider only a single output here. Then we have the centre-of-set type reduction method as follows: [35]

  $$y_{cos}(\overrightarrow{x}) = [y_l, y_r] = \int_{w^1\in\left[w_l^1, w_r^1\right]} \cdots \int_{w^M\in\left[w_l^M, w_r^M\right]} \int_{f^1\in\left[\underline{f}^1, \overline{f}^1\right]} \cdots \int_{f^M\in\left[\underline{f}^M, \overline{f}^{1M}\right]}$$

  $$y_{cos}(\overrightarrow{x}) = [y_l, y_r] = 1/\sum_{i=1}^{M} f^i w^i / \sum_{i=1}^{M} f^i \tag{2.74}$$

  where $y_{cos}(\overrightarrow{x})$ is an interval type-1 set determined by left and right end points (y_l and y_r), which can be derived from consequent centroid set $[w_l^i\ w_r^i]$ and firing strengths $f^i \in F^i = [\underline{f}^i, \overline{f}^i]$. The interval set $[w_l^i \; w_r^i]$ $(l = 1, \cdots, M)$ should be computed or set first before the computation of $y_{cos}(\overrightarrow{x})$. For any value $y \in y_{cos}$, y can be stated as: [35]

$$y = \frac{\sum_{i=1}^{M} f^i w^i}{\sum_{i=1}^{M} f^i} \tag{2.75}$$

where y is a monotonic increasing function with respect to w^i. Also, y_l in equation 2.74 is the minimum associated only with w_l^i and y_r is the maximum associated only with w_r^i. Note that y_l and y_r depend on a mixture of $\underline{f}^i$ or $\overline{f}^i$ values. Therefore, left-most point y_l and right-most point y_r can be shown as in Equations 2.76 and 2.77, respectively [35].

$$y_l = \frac{\sum_{i=1}^{M} f_l^i w_l^i}{\sum_{i=1}^{M} f_l^i} \tag{2.76}$$

$$y_r = \frac{\sum_{i=1}^{M} f_r^i w_r^i}{\sum_{i=1}^{M} f_r^i} \tag{2.77}$$

2.10.6 Type Reduction Algorithm for Interval Type-2 Fuzzy Logic Systems

Without loss of generality [44], assume the w_r^i will be arranged in ascending order, i.e. $w_l^1 \leq w_r^2 \leq \cdots \leq w_r^M$.

1. Compute y_r in equation 2.77 by initially using $f_r^i = (\underline{f}^i + \overline{f}^i)/2$ for $i = 1, \cdots, M$, where $\underline{f}^i$ and $\overline{f}^i$ are pre-computed by equations 2.71 and 2.72; and let $\acute{y}_r = y_r$.

2. Find $R(1 \leq R \leq M-1)$ such that $w_r^R \leq \acute{y}_r \leq w_r^{R+1}$.

3. Compute y_r in equation 2.77 with $f_r^i = \underline{f}^i$ for $i \leq R$ and $f_r^i = \overline{f}^i$ for $i > R$, then set $y_r^n = y_r$.

4. If $y_r^n \neq \acute{y}_r$, then go to step 5. If $y_r^n = \acute{y}_r$, then set $y_r = y_r^n$ and go to step 6.

5. Let $\acute{y}_r = y_r^n$ and return to step 2.

6. Stop.

This algorithm decides the point to separate two sides by the number R, one side using lower firing strengths $\underline{f}^i$'s and another side using upper firing strengths $\overline{f}^i$'s Therefore, the y_r in equation 2.77 can be re-stated as: [44]

$$y_r = y_r(\underline{f}^1, \cdots, \underline{f}^R, \overline{f}^{R+1}, \cdots, \overline{f}^M, w_r^1, \cdots, w_r^M)$$

$$y_r = \frac{\sum_{i=1}^R \underline{f}^i w_r^i + \sum_{i=R+1}^M \overline{f}^i w_r^i}{\sum_{i=1}^R \underline{f}^i + \sum_{i=R+1}^M \overline{f}^i}$$

$$y_r = \frac{\sum_{i=1}^R \underline{f}^i w_r^i}{\sum_{i=1}^R \underline{f}^i + \sum_{i=R+1}^M \overline{f}^i} + \frac{\sum_{i=R+1}^M \overline{f}^i w_r^i}{\sum_{i=1}^R \underline{f}^i + \sum_{i=R+1}^M \overline{f}^i}$$

$$y_r = \sum_{i=1}^R \frac{\underline{f}^i}{D_r} w_r^i + \sum_{i=R+1}^M \frac{\overline{f}^i}{D_r} w_r^i$$

$$y_r = \sum_{i=1}^R \underline{q}_b^i w_r^i + \sum_{i=R+1}^M \overline{q}_b^i w_r^i \tag{2.78}$$

where $D_r = (\sum_{i=1}^R \underline{f}^i + \sum_{i=R+1}^M \overline{f}^i)$, $\underline{q}_b^i = \underline{f}^i / D_r$ and $\overline{q}_b^i = \overline{f}^i / D_r$.

The procedure to compute y_l is similar to compute y_r. In step 2, it only needs to find $L(1 \leq L \leq M-1)$, such that $w_l^L \leq \acute{y}_l \leq w_l^{L+1}$ in step 3, let $f_l^i = \overline{f}^i$ for $i \leq L$ and $f_l^i = \overline{f}^i$ for $i > L$. y_l in equation 2.76 can be also re-expressed as: [44]

$$y_l = y_l(\overline{f}^1, \cdots, \overline{f}^L, \underline{f}^{L+1}, \cdots, \underline{f}^M, w_l^1, \cdots, w_l^M)$$

$$y_r = \frac{\sum_{i=1}^L \overline{f}^i w_l^i + \sum_{i=L+1}^M \underline{f}^i w_l^i}{\sum_{i=1}^L \overline{f}^i + \sum_{i=L+1}^M \underline{f}^i}$$

$$y_r = \frac{\sum_{i=1}^L \overline{f}^i w_l^i}{\sum_{i=1}^L \overline{f}^i + \sum_{i=L+1}^M \underline{f}^i} + \frac{\sum_{i=L+1}^M \underline{f}^i w_l^i}{\sum_{i=1}^L \overline{f}^i + \sum_{i=L+1}^M \underline{f}^i}$$

$$y_r = \sum_{i=1}^L \frac{\overline{f}^i}{D_l} w_l^i + \sum_{i=L+1}^M \frac{\underline{f}^i}{D_l} w_l^i$$

$$y_r = \sum_{i=1}^L \overline{q}_a^i w_l^i + \sum_{i=L+1}^M \underline{q}_a^i w_l^i \tag{2.79}$$

where $D_l = (\sum_{i=1}^L \overline{f}^i + \sum_{i=L+1}^M \underline{f}^i)$, $\overline{q}_a^i = \overline{f}^i / D_l$ and $\underline{q}_a^i = \underline{f}^i / D_l$.

The defuzzified crisp output from an interval type-2 fuzzy logic system is usually taken as the average of y_l and y_r, i.e., [44]

$$\frac{y_l + y_r}{2}.$$

2.11 Recent Work on Type-2 Fuzzy Sets and Systems

Recently, research interest in type-2 fuzzy logic has seen significant growth. Most of this research has only been concerned with interval type-2 fuzzy systems, in which the membership grade of a fuzzy set is given as an interval set, because it simplifies the complexity of generalised type-2 fuzzy sets. It is clear from the recent work that type-2 fuzzy logic should have a role to play in modeling uncertainty.

Type-2 fuzzy set has been first defined and discussed by Zadeh in [34, 45, 46]. Zadeh concentrated on the notion of a fuzzy set where the membership functions of a fuzzy set are measured with linguistic terms. Zadeh only explored the use of the minimum and maximum operators *t-norm* and *t-conorm*. Later, Mizumoto and Tanaka [7, 8] and Dubois and Prade [5] studied the logical connectives of what became known as a *secondary membership functions*. They also studied the *join* and *meet* under a variety of t-norm and t-conorm operators.

The use of type-2 fuzzy sets, called interval valued or IV fuzzy sets, has been promoted later by Turksen [47–49], Schwartz [50] and Klir and Folger [3] They discussed that type-2 interval fuzzy sets should be employed when the linguistic uncertainty of a term cannot be sufficiently modeled by the type-1 fuzzy sets. Zadeh [51, 52] claimed that fuzzy logic equates to computing with words (CWW) and provides examples using fuzzy granules (which is actually the FOU of an interval type-2 fuzzy set) to model words. Mendel [1, 53, 54] and Turksen [55] also agreed that CWW requires type-2 fuzzy sets by using the simpler interval type-2 representations. As "Words Mean Different Things to Different People", Mendel [53] demonstrated that human models of words as obtained through a survey require at least interval representations. These ideas led to more work on type-2 fuzzy sets and this has kept growing.

The technique used for defuzzifying type-2 fuzzy sets, called type-reduction, has been defined by Karnik and Mendel [37–39,56] by applying the extension principle to a variety of type-1 defuzzifiers. In those papers, Karnik and Mendel also provided a complete description of the fuzzy inferencing process, which is allow work on the application of type-2 fuzzy logic to proceed. John [42,57–59] also published a number of review papers on type-2 fuzzy systems. In 2001, Mendel published the first, and only textbook on the subject of type-2 fuzzy logic as *Uncertain Rule-Based Fuzzy Logic System: Introduction* *and New Directions* [1]. This greatly increased the interest in type-2 fuzzy logic and they have since been widely used in many areas of application.

The representation theorem of type-2 fuzzy sets has been given by Mendel and John [9]. They defined operations of type-2 fuzzy set without the use of the extension principle by representing a type-2 fuzzy set as a collection of simpler type-2 embedded sets. An example of their representation theorem for type-2 fuzzy sets is the definition of arithmetic operators for type-2 fuzzy numbers, as proposed by Coupland and John [60].

As previously noted, the complexity of join and meet operations and type-reduction of a type-2 fuzzy set limit the applicability of type-2 methods and, even though interval type-2 fuzzy sets are simpler, type-reduction is still a problem because of the inherent complexity and redundancies. The Karnik-Mendel [38] and Wu-Mendel [61,62] methods have been developed to make the type-reduction of interval type-2 fuzzy sets more efficient. The complexity of join and meet was resolved by those methods. This work has also been discussed by other authors. e.g. Coupland et al. [63–65] discussed this issue with some aspects of the geometric approach and Greenfield et al. [66] provided an efficient method for approximating the type-reduced set of a type-2 fuzzy set using a stochastic approach. Hisdal studied rules and interval sets for higher-than-type-1 FLS [67]. John and Coupland continued to discussed in this issue in [68–73]. Additional discussions on the use of interval sets in fuzzy logic can be found in [74–79].

2.11.1 Recent Applications of Type-2 Fuzzy Sets and Systems

2.11.1.1 Control Applications

To date, type-2 fuzzy logic has been widely used in control applications, and most applications are using type-2 interval fuzzy sets with the Karnik-Mendel iterative algorithms and the Wu-Mendel minimax uncertainty bounds, allowing fast execution of type-2 fuzzy systems.

Many researchers have begun to use type-2 fuzzy logic in control applications. For example, Melin and Castillo [80, 81] and Castillo et al. [82] have used type-2 interval systems in the context of plant control. Hagras [83] presented type-2 fuzzy logic control application to three challenging domains including industrial, mobile robots, and ambient intelligent environments control. Lynch et al [84, 85] are continuing to build a type-2 interval control system for large marine diesel engines. Hagras et al. [86, 87] and Doctor et al [88] used a type-2 interval system to model and adapt to the behaviour of people in an intelligent dormitory room. Wu and Tan [89] applied type-2 interval systems to the control of a complex multi-variable liquid level process and in [90] they simplified type-2 fuzzy logic control to real-time control applications. Melgarejo et al. [91] have developed a limited hardware implementation of a type-2 interval controller. Lin et al. [92, 93] designed type-2 fuzzy controller for buck DC-DC converters.

2.11.1.2 Time Series Forecasting Application

There are more researchers interested in using type-2 fuzzy sets to deal with forecasting applications. Uncu et al. [94] proposed a system modelling approach based on type-2 fuzzy sets to predict the price of a stock. Baguley et al. [95] found that a model with type-2 fuzzy sets can leverage design process knowledge and predict time to market from performance measures is a potentially valuable tool for decision making and continuous improvement. Kim and Park [96] used type-2 fuzzy logic system to forecast the Box-Jenkin's gas furnace time series and compare the results with type-1 fuzzy logic system. Huarng and Yu [97] proposed the use of a type-2 fuzzy time series model to improve the prediction performance by using the TAIEX, Taiwan stock index, as the forecasting target.

Medina and Mendez [98] presented an application of the interval singleton type-2 fuzzy logic system to one-step-ahead prediction of the daily exchange rate between the Mexican Peso and US dollar (MXNUSD). Mencattini et al. [99, 100] used type-2 fuzzy systems for meteorological forecasting. Li et al. [101] proposed a new method for short-term traffic forecasting using type-2 fuzzy logic.

Karnik and Mendel [102] used a type-2 interval system to predict the next value in a chaotic time series. Musikasuwan et al. [103] investigated the effect of number of model parameters on performance in type-1 and interval type-2 systems. Both systems were designed to predict a Mackey-Glass time series.

Liang and Wang [104] presented a new approach for sensed signal strength forecasting in wireless sensors using interval type-2 fuzzy system and compare with type-1 fuzzy system. Pareek and Kar [105] demonstrated an application of type-2 fuzzy system to predict a critical parameter of Gas Turbine in a power plant, that is the compressor discharge pressure. Mendez et al. [106] presented the experimental results of the application of type-2 fuzzy systems for scale breaker entry temperature prediction in a real hot strip mill.

2.11.1.3 Medical Applications

There are researchers using type-2 fuzzy logic to model in medical application. John et al. [107–109] used type-2 fuzzy sets to assist in the pre-processing of tibia radiographic images, while John and Lake investigated the use of type-2 fuzzy sets in modelling nursing intuition. Innocent et al. [110–112] represented the perceptions of lung scan images by experts in order to predict pulmonary emboli by using type-2 fuzzy relations. Garibaldi et al. [10–14] have done extensive work on assessing the health of a new born baby using knowledge of acid-base balance in the blood from the umbilical cord. Di Lascio et al. [113] presented a model of differential medical diagnosis for the pathologies based on type-2 fuzzy sets to indicate the elements needs to have more precise diagnosis and it can control its same accuracy. Finally, Herman et al. [114] examined the potential of the type-2 fuzzy system methodology in devising an EEG-based brain-computer interface to classify imaginary left and right hand movements.

2.11.1.4 Mobile Robot Applications

Type-2 fuzzy systems were successfully applied in mobile robot controllers. Phokharatkul and Phaiboon [115] implemented the type-2 fuzzy logic controller to process the data output to control the direction of the mobile robot movement. Hagras [116,117] implemented the type-2 fuzzy logic controller in different types of mobile robots navigating in indoor and outdoor unstructured and challenging environments. Coupland et al. [118] designed and compared three fuzzy logic control using type-1 , interval type-2 and general type-2 fuzzy logic to the robot control for completing the task of following the edge of a curved wall, and found that both type-2 fuzzy systems outperformed the type-1 system. Figueroa et al. [119] explored how the type-2 fuzzy logic controller, in the context of robot soccer games, overcomes uncertainty in the control loop without increasing the computational cost of the application. Finally, Wu [120] designed and implemented type-2 fuzzy logic control on Motorola 68HC11 8-bit micro-controllers to navigate a miniature robot in an unknown maze without touching the walls.

2.11.1.5 Others Applications

The others areas of application that type-2 fuzzy sets and systems have been successfully implemented. For examples, Gu and Zhang [121] created the web shopping expert based on the interval type-2 fuzzy inference system to provide a reasonable decision for online users. Tang et al. [122] constructed an online system 'hotstore.com' using type-2 fuzzy reasoning.

Lee and Lee introduced a ranking method for type-2 fuzzy values and used this result in solving the shortest path problem in a type-2 weighted graph [123, 124].

2.12 Summary

This chapter has described crisp set theory and the concepts of fuzzy sets. A fuzzy set can be defined using either the list or rule methods when the universe of discourse is countable and continuous, respectively. The classical union, intersection, and complement set theoretic operations of fuzzy sets have also presented in this chapter.

Since Zadeh first introduced the concept of a fuzzy set [2] and subsequently went on to extend the notion via the concept of linguistic variables [4] the popularity and use of fuzzy logic has been extraordinary. Fuzzy principles have been applied to a huge and diverse range of problems such as aircraft flight control, robot control, car speed control, power systems, nuclear reactor control, fuzzy memory devices and the fuzzy computer, control of a cement kiln, focusing of a camcorder, climate control for buildings, shower control and mobile robots [125, 126].

The use of fuzzy logic is not limited to control. Other successful applications include, for example, stock tracking on the Nikkei stock exchange [126], and information retrieval [127].

Then, the concept of type-2 fuzzy sets, both interval and non-interval (generalised), have been described. The footprint of uncertainty (FOU) was introduced to provide a very convenient verbal description for the entire domain of support of all the secondary grades of a type-2 membership function. The inferencing processes, type reduction and defuzzifcation, and operators for both interval and generalised fuzzy systems are also presented. Finally, recent work on type-2 fuzzy sets and systems and their applications are discussed.

In next chapter, the use of non-standard membership functions to better model reasoning in a variety of complex domains, including when modelling human reasoning, has been described.

Chapter 3

Non-Convex Fuzzy Sets

3.1 Introduction

This chapter presents the use of complex non-convex membership functions in the context of human decision making systems. In particular, this chapter attempts are made to address criticisms that were made as to whether the shapes being presented were really 'true', 'allowable' or in any way 'meaningful' membership functions.

It is suggested that in many applications involving the modelling of human decision making (expert systems) the more traditional membership functions do not provide a wide enough choice for the system developer. They are therefore missing an opportunity to, potentially, produce simpler or better systems.

This case study highlights a number of membership functions outside the paradigm of fuzzy control. In particular, the merits of non-convex fuzzy sets are discussed and a case study is presented which investigates whether it is possible to build an expert system featuring usual Mamdani style fuzzy inference in which a time-related non-convex fuzzy set is used together with traditional fuzzy sets. It is shown that this is indeed possible and an examination is made of the resultant output surface generated by four different sub-classes of non-convex membership functions.

The rest of the case study is structured as follows. Sections 3.2 (conventions of fuzzy terms) presents the discussion of membership functions and restates accepted definitions for completeness. Section 3.2.1 (non-convex membership functions) describes the concept and classes of non-convex membership functions. Notes that these two sections are abbreviated from the previous work [128] and are included for completeness. The new case study of using a time-related non-convex membership function is presented in Section 3.3. Finally, Section 3.4 and Section 3.5 present a discussion of the issues raised and summarised of this chapter, respectively.

3.2 Conventions of Fuzzy Terms

To enable a discussion of membership functions, we need to formally define the terminology used. In this Section, accepted definitions are restated for completeness.

Definition 3.2.1 (Linguistic variable) A linguistic variable is characterised by a quintuple $(X, T(X), U, G, M)$ in which X is the name of the variable, $T(X)$ is the term set, U is a universe of discourse, G is a syntactic rule for generating the elements of $T(X)$ and M is a semantic rule for associating meaning with the linguistic values of X [129].

Definition 3.2.2 (Normal) A fuzzy set, A, is *normal* if $\exists x'$ such that $\mu_A(x') = 1$ [129].

Definition 3.2.3 (Sub-normal) A fuzzy set, A, is *sub-normal* if it is not normal i.e. $\exists$ no x' such that $\mu_A(x') = 1$ [129].

Definition 3.2.4 (Convex) A fuzzy set, A, is said to be *convex* if and only if all of its α-cuts are convex in the classical sense. That is, for each α-cut, A_α, for any $r, s \in A_\alpha$ and any $\lambda \in [0,1]$ then $\lambda r + (1-\lambda)s \in A_\alpha$ [129].

Definition 3.2.5 (Non-convex) A fuzzy set, A, is said to be *non-convex* if it is not convex [129].

As well as being interested in sub-normal, non-convex fuzzy sets we also consider fuzzy sets that are *contained* in, or *included* in, another fuzzy set(s). For clarity, we use the term *subsumed* to describe a fuzzy set that is contained within another. We consider that such fuzzy sets can play an important role in human decision making. A subsumed fuzzy set is a special case of a non-distinct fuzzy set. In this chapter, we particularly investigate the use of non-convex membership functions for linguistic terms which is presented as a case study in section 3.3.

Definition 3.2.6 (Distinct) A fuzzy set, A, for a particular linguistic variable L, on the universe of discourse X is *distinct* from a fuzzy set, B (another term of L), on the universe of discourse X if and only if for all $x' \in X$ when $\mu_A(x') > 0$ then $\mu_B(x') = 0$ and when $\mu_B(x') > 0$ then $\mu_A(x') = 0$ [129].

Definition 3.2.7 (Non-distinct) A fuzzy set, A, for a particular linguistic variable L, is *non-distinct* if $\exists$ a fuzzy set B (another term of L) such that A is not distinct from B [129].

Non-distinct fuzzy sets are also referred to as *overlapping* fuzzy sets. There are many types of non-distinct fuzzy sets. For clarity, we further define *partially overlapping* and *subsumed* fuzzy sets.

Definition 3.2.8 (Partially overlapping) A fuzzy set, A, on the universe of discourse X is *partially overlapping* another fuzzy set, B, on the universe of discourse X if and only if $\exists x'$ where $\mu_A(x') = max(\mu_A)$ but $\mu_B(x') \neq max(\mu_B)$, and $\exists x''$ where $\mu_B(x'') = max(\mu_B)$ but $\mu_A(x'') \neq max(\mu_A)$ [129].

Definition 3.2.9 (Subsumed) A fuzzy set, A, on the universe of discourse X is *subsumed* within a fuzzy set, B, on the universe of discourse X if and only if for all $x' \in X$ $\mu_B(x') \geq \mu_A(x')$ [129].

Definition 3.2.10 (Regular) Fuzzy terms that are normal, convex and distinct using the above definitions will be referred to as *regular* terms [129].

It is often implicitly accepted, and occasionally explicitly stated (e.g. [130, 131]), that the terms of a linguistic variable should be *justifiable in number* (5 ± 2), *distinct*, *normalised* and covering the entire universe of discourse.

3.2.1 Non-Convex Membership Functions

It would appear that the class of fuzzy sets which might have non-convex membership functions can be naturally split into three sub-classes:

- Those where the universe of discourse is not time-related. Such sets will be termed *elementary* non-convex membership functions.

- Those where the universe of discourse is time-related. Such sets will be termed *time-related* non-convex membership functions.
- Those which result from the inferencing process in the Mamdani method. Such sets are termed *consequent* non-convex membership functions.

3.2.2 Elementary Non-Convex Sets

Plausible discrete domain non-convex fuzzy sets which are not defined over a time-related universe of discourse are quite easy to imagine. There are three 'well-known' principles that govern the ideal number of people for forming a mountain rescue team:

1. there should be an odd number of people so that in any decision-making vote a simple majority is possible (i.e. voting does not result in ties);
2. three is not a good number to have, because there is a tendency to end up with a 2-1 split which causes the single person to feel resentful; and
3. too many people cause too many arguments.

Hence a discrete fuzzy set expressing the compatibility of various numbers of people with a suitable mountain-rescue team might look as in Figure 3.1.

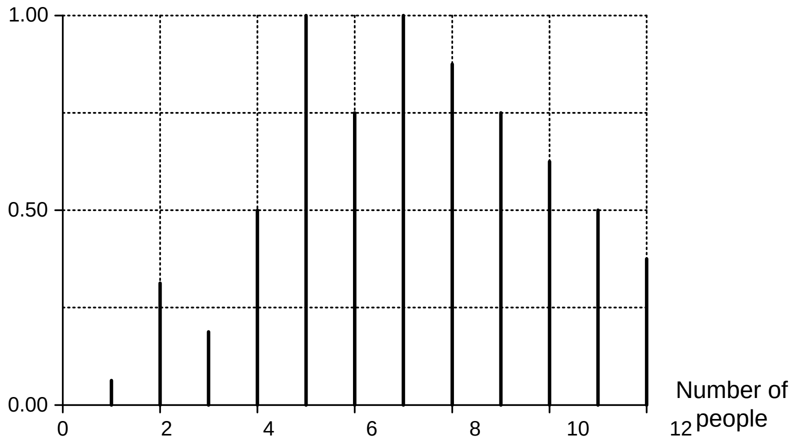


Figure 3.1: A discrete non-convex set: suitability of number of people to comprise a mountain rescue team

Continuous domain non-convex fuzzy sets may be less common. Consider though, as an example, the desirability (drinkability) of a glass (cup) of milk according to the temperature of the milk. Most people (who like drinking milk) prefer it 'ice-cold' out of the fridge as opposed to room temperature (although actually 'ice-cold' refers to several degrees above freezing). Many people also agree that hot milk is also quite pleasant to drink. Hence a fuzzy set expressing the drinkability of milk by temperature might look like Figure 3.2.

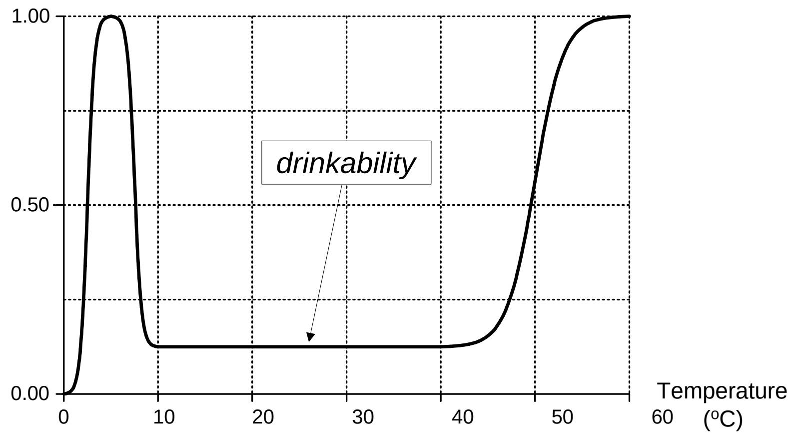


Figure 3.2: Drinkability of milk by temperature

Alternatively, the representation of this concept could be transformed by defining a variable *temperature* with perhaps four convex terms, *icy*, *cold*, *medium*, and *hot* (Figure 3.3), an output set *drinkability* with two convex terms, *low* and *high* (Figure 3.4), and an associated set of rules of the form:

IF *temp* is *cold* THEN *drinkability* is *high*
IF *temp* is *medium* THEN *drinkability* is *low*
IF *temp* is *high* THEN *drinkability* is *high*

The drinkability for a given temperature could then be found by inputting the temperature into the above set of fuzzy rules, executing the rules and then defuzzifying the consequent set by, for example, the centroid method. A plot of the resultant drinkability obtained for each temperature is shown in Figure 3.5. Note that the resultant set (Figure 3.5) is also now sub-normal. Of course, it can be normalised to obtain a closer match to Figure 3.2.

But how should the rules and membership functions to obtain the precise shape required be determined? And why incur the additional time and effort of eliciting the 5-6 membership functions required when it is simpler to elicit the set shown in Figure 3.2 directly?

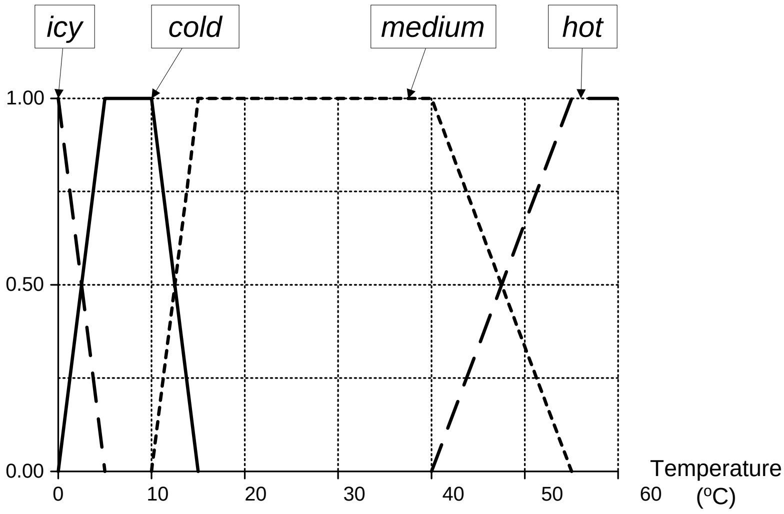


Figure 3.3: Temperature as an input variable

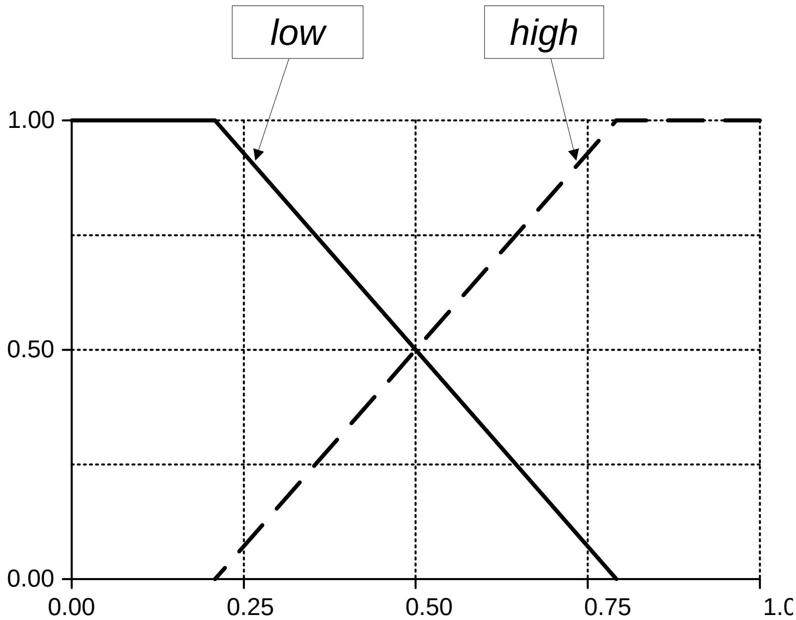


Figure 3.4: Drinkability as an output variable

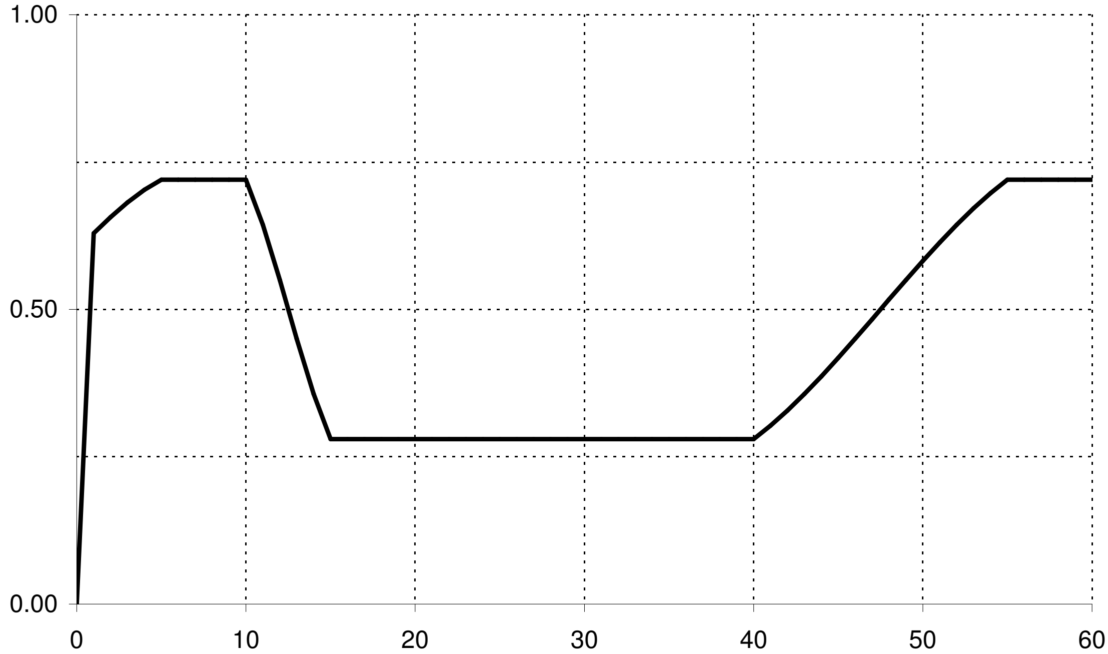


Figure 3.5: Drinkability of milk elicited from the rule base: c.f. Fig. 3.2

3.2.3 Time-Related Non-Convex Sets

As a plausible time-related fuzzy set, suppose that an energy-supply company is creating an expert system to predict demand load. Amongst other factors that are considered may be the time of day and the prevailing temperature outside. We want to capture the concept that energy demand increases at mealtimes. Of course, mealtime is a fuzzy concept as breakfasts, lunches and dinners occur at variable times and indeed *may* occur at any time. Hence a non-convex fuzzy set for *mealtime* defined on time-of-day may be defined as shown in Figure 3.6. Rules may then be created of the form:

IF *time-of-day* is *mealtime* AND *temp* is *low*
THEN *energy-demand* is *high*

Note that this fuzzy set is interesting as it is also sub-normal and never has a membership of zero.

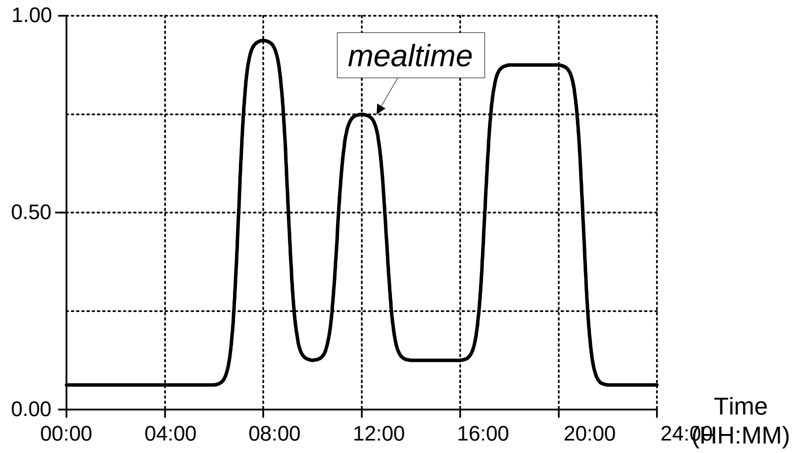


Figure 3.6: Mealtime by time of day

Another example of a time-related fuzzy set is that of disposable income. By this we mean the amount of money (as a percentage of salary) somebody has available after paying out all their commitments (e.g. loan(s), electricity bill, etc.). It is well known that if you are young you have more disposable income than if you are middle aged (typically with a mortgage and children) and also as you get past middle age your disposable income increases. Depending on the application we could look at high disposable income in two ways. In the first case we might have a fuzzy set *high* for the linguistic variable *disposable income* as in Figure 3.7 that has a domain which is the percentage of disposable income. However we may not know this information but have someone's age. In this case the domain would be age and we would have, for example, the non-convex fuzzy set in Figure 3.8, in which the domain is time but the fuzzy set relates to income.

3.2.4 Consequent Non-Convex Sets

In a rule-based fuzzy system the result of, for example, Mamdani fuzzy inferencing, is a fuzzy set. Figure 3.9 provides an example of a typical result of Mamdani inferencing (prior to defuzzification) where the antecedent and consequent fuzzy sets are triangular and/or trapezoidal. In the context of fuzzy control, this is usually defuzzified to produce the precise value required for the output variable. However, when modelling human decision making in a rule-based fuzzy system we might want to use the output directly as part of a chained inference process or we might like the output to be defuzzified somehow to a linguistic term. If it is argued that sub-normal, non-convex sets have no meaning, then what should be done with consequent sets? If, on the other hand, it is accepted that such consequent sets are meaningful, so that they can be interpreted or chained in further processing, then why should not the original inputs be similarly formed. Hence, we believe we need to improve our understanding of sub-normal, non-convex sets in order to lead us toward better 'Computing with Words' [51].

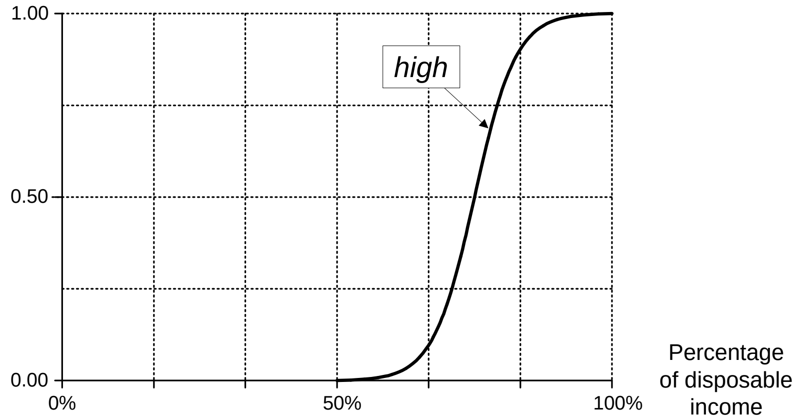


Figure 3.7: A convex set *high* defined on the percentage of disposable income

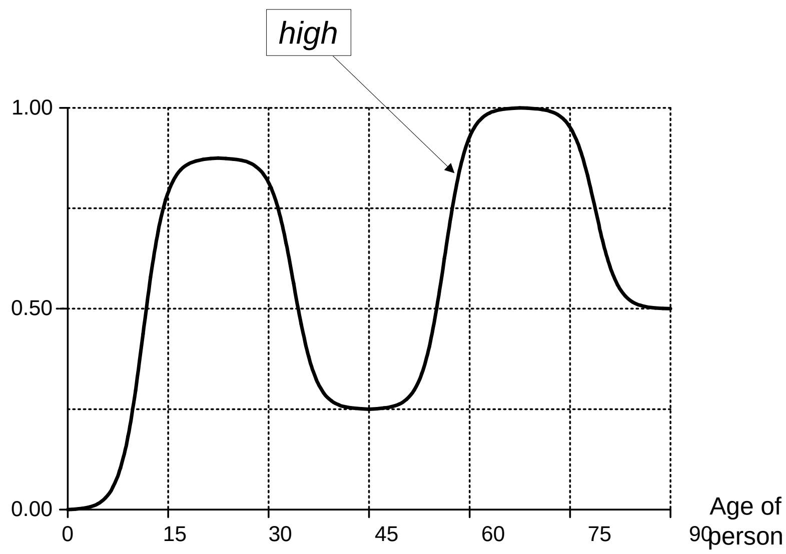


Figure 3.8: A non-convex set *high-disposable-income* defined on the age

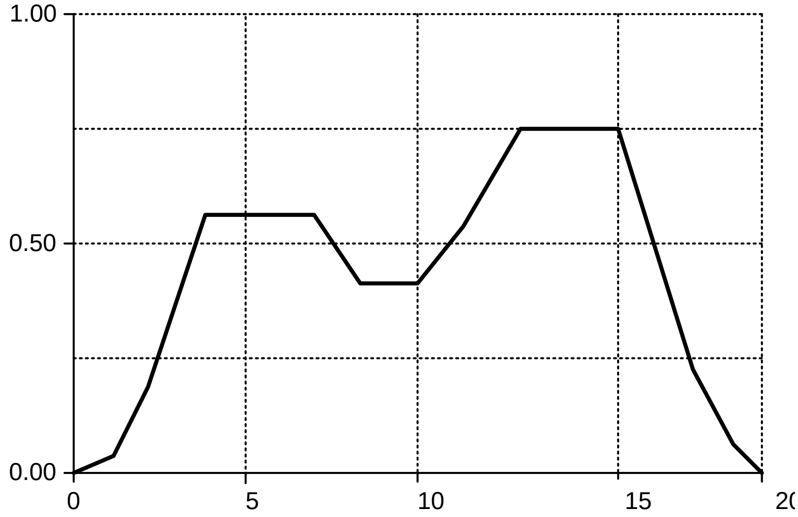


Figure 3.9: An example of a typical non-convex, sub-normal consequent set

3.3 A Case Study to Illustrate the Use of Non-Convex Membership Functions for Linguistic Terms

This section focuses on time related non-convex membership functions. It investigates whether the generated expert systems would work properly or not when time related non-convex membership functions are used together with normal membership functions. Suppose that an energy supply company is developing an expert system to predict demand load. Although there may be many factors that effect the demand load, *time* of the day and the prevailing *temperature* outside are chosen as the two system inputs. A simple model is more appropriate at this stage since it is only aimed to demonstrate the feasibility of the system.

3.3.1 Methodology

Firstly, 500 data sets of *time* and *temperature* are generated randomly where *time* is between 0 and 24 in the hh:mm format and temperature varies between 0°C and 40°C.

The illustrative system consists of two input variables, *Time* and *Temperature*, and one output variable, *Energy Demand* as shown in Figure 3.10 - 3.15. In addition to other usual membership functions, the *Time* variable is associated with the term *MealTime* which is a time-related non-convex membership function. In order to observe the influence of *MealTime* on the performance, four systems are created by only changing the term *MealTime* in each system. The four different shapes of *MealTime* membership functions that are used are as follows:

Case 1: mf of MealTime is in [0,1]

Case 2: mf of MealTime is in [0.2,1]

Case 3: mf of MealTime is in [0,0.9]

Case 4: mf of MealTime is in [0.2,0.9]

Note that membership functions which never reach zero are also unconventional but, once again, there appears no reason why this convention cannot be violated.

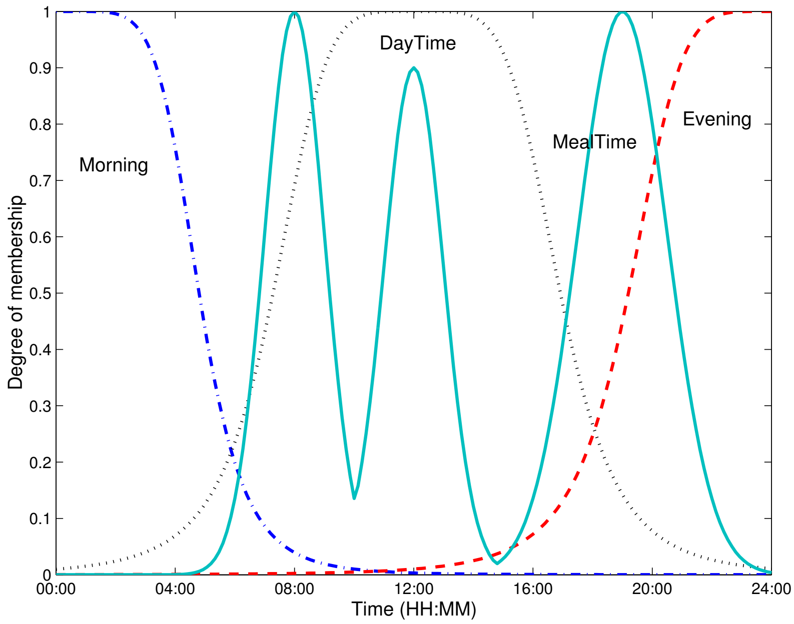


Figure 3.10: Case 1: Time with MealTime mf in [0,1]

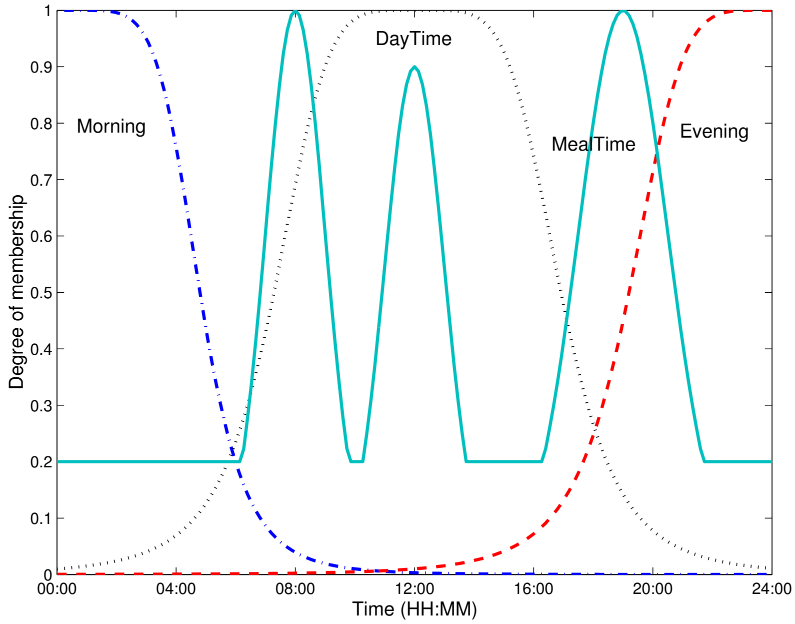


Figure 3.11: Case 2: Time with MealTime mf in [0.2,1]

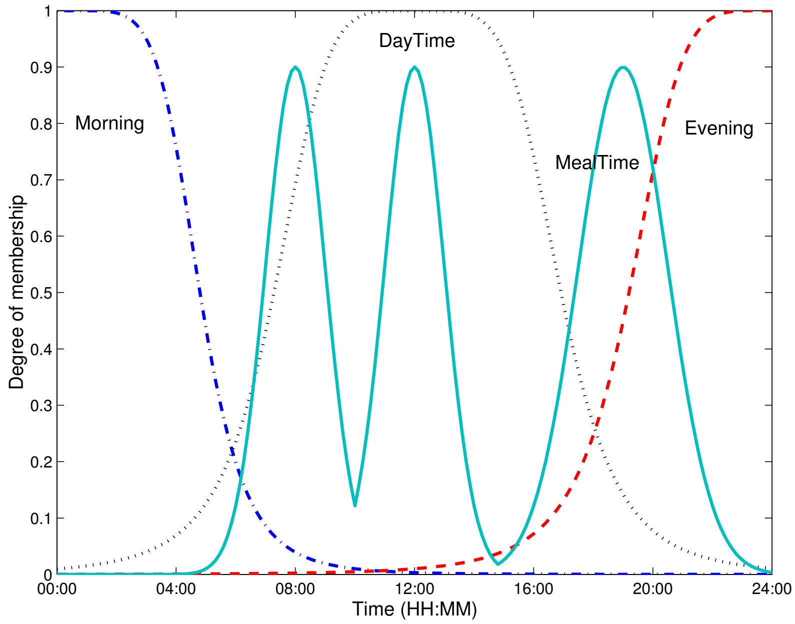


Figure 3.12: Case 3: Time with MealTime mf in [0,0.9]

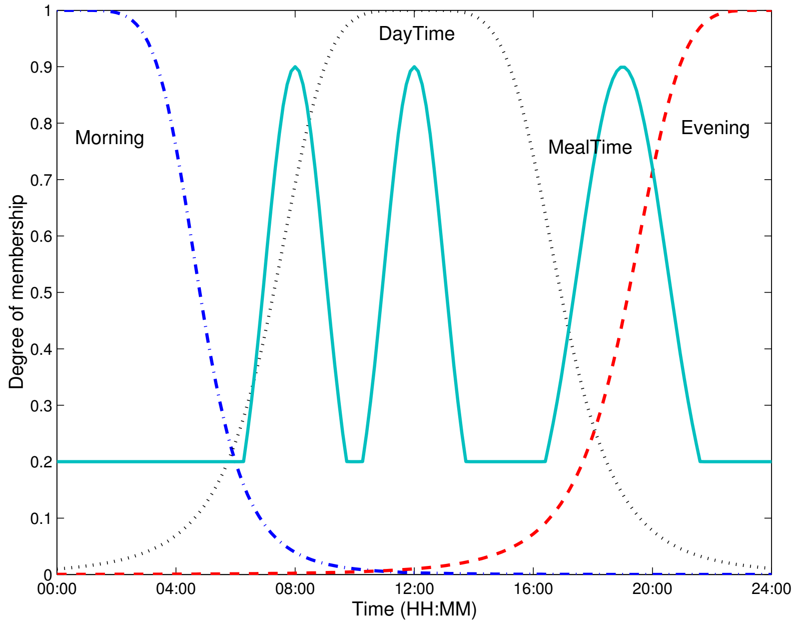


Figure 3.13: Case 4: Time with MealTime mf in [0.2,0.9]

As seen in Figures 3.10 - 3.13, *Time* is an input variable which consists of four membership functions; *Morning*, *DayTime*, *Evening*, and *MealTime* where the membership values of MealTime vary in the range of [0,1], [0.2,1], [0,0.9], and [0.2,0.9], for each generated system respectively. The values of *time* is between 0 and 24.

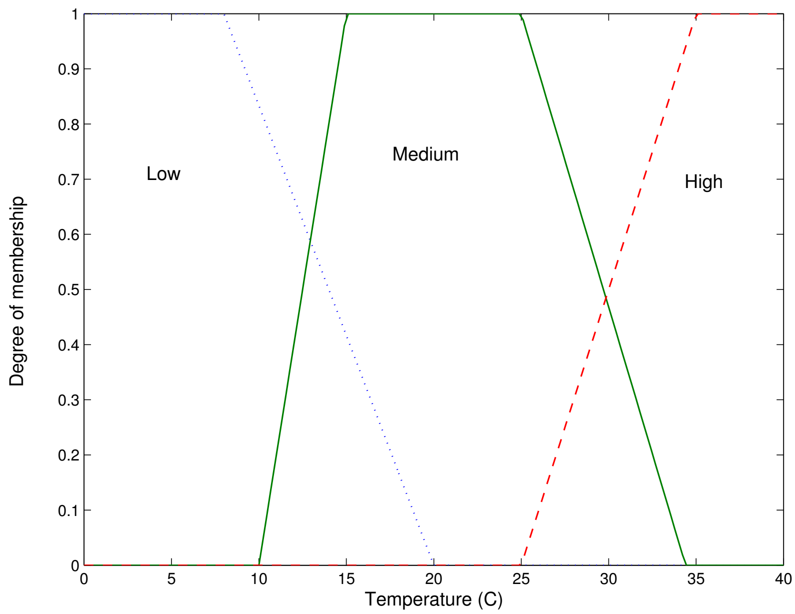


Figure 3.14: Temperature as an input variable with normal membership functions

Figure 3.14 shows the input variable *Temperature* which consists of three membership functions; *Low*, *Medium*, and *High*. The same linguistic variable, *Temperature*, is used in all four systems. The values of temperature is between 0°C and 40°C.

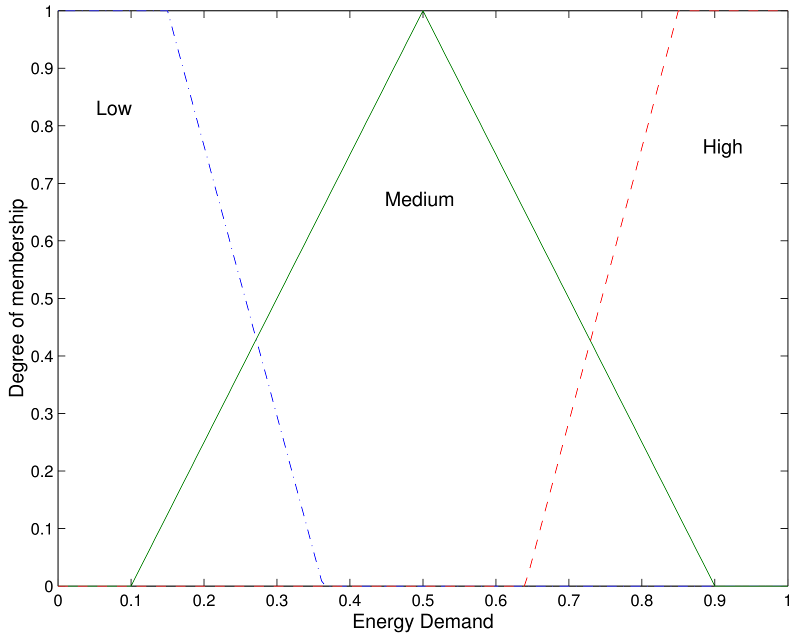


Figure 3.15: Energy Demand as an output variable with normal membership functions

Figure 3.15 shows the output variable *Energy Demand* which consists of three membership functions; *Low*, *Medium*, and *High*. The same linguistic variable, *Energy Demand*, is used in all four generated systems. The values of Energy Demand varies between 0 and 1.

The following 12 rules are used within the four expert systems. These rules are purely illustrative at the moment and do not correspond to a real application. The rules are:

1. IF *Time* is *Mealtime* AND *Temperature* is *Low*
   THEN *Energy-Demand* is *High*
2. IF *Time* is *Mealtime* AND *Temperature* is *Medium*
   THEN *Energy-Demand* is *High*
3. IF *Time* is *Mealtime* AND *Temperature* is *High*
   THEN *Energy-Demand* is *Medium*
4. IF *Time* is *Evening* AND *Temperature* is *Low*

THEN *Energy-Demand* is *High*

5. IF *Time* is *Evening* AND *Temperature* is *Medium*
   THEN *Energy-Demand* is *Medium*
6. IF *Time* is *Evening* AND *Temperature* is *High*
   THEN *Energy-Demand* is *Medium*
7. IF *Time* is *DayTime* AND *Temperature* is *Low*
   THEN *Energy-Demand* is *High*
8. IF *Time* is *DayTime* AND *Temperature* is *Medium*
   THEN *Energy-Demand* is *Medium*
9. IF *Time* is *DayTime* AND *Temperature* is *High*
   THEN *Energy-Demand* is *Low*
10. IF *Time* is *Morning* AND *Temperature* is *Low*
    THEN *Energy-Demand* is *Medium*
11. IF *Time* is *Morning* AND *Temperature* is *Medium*
    THEN *Energy-Demand* is *Medium*
12. IF *Time* is *Morning* AND *Temperature* is *High*
    THEN *Energy-Demand* is *Low*

3.3.2 Results

After time related non-convex membership function (*MealTime*) is applied into linguistic variable *Time*, it is observed that all four systems work perfectly well (as expected). The prediction results of *energy demand* are shown as three dimensional plots in Figures 3.16 - 3.19. Table 3.1 shows the summary of the results obtained from each system. The difference in the predicted results of each system is due to the different time-related non-convex *MealTime* term added to the variable *time*. It is clearly observed that the *energy demand* predictions have incorporated the information introduced by addition of the *MealTime* term. The surface plots, particularly in Figures 3.16 and 3.18 exhibit a marked increase in output at the peak meal times.

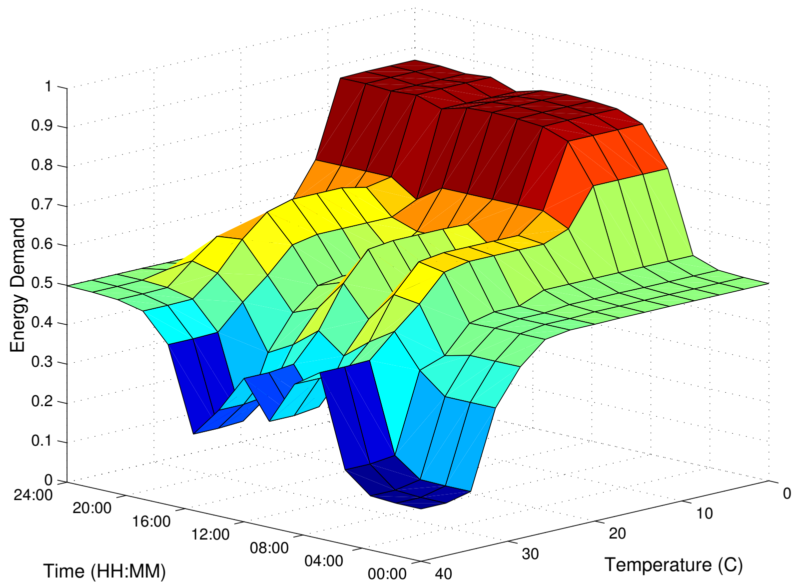


Figure 3.16: Three dimensional plot of the result from Case 1

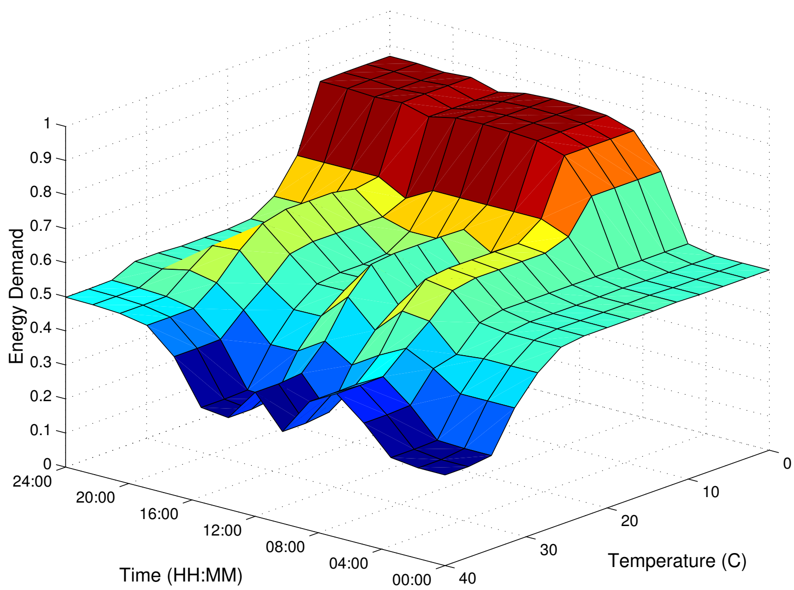


Figure 3.17: Three dimensional plot of the result from Case 2

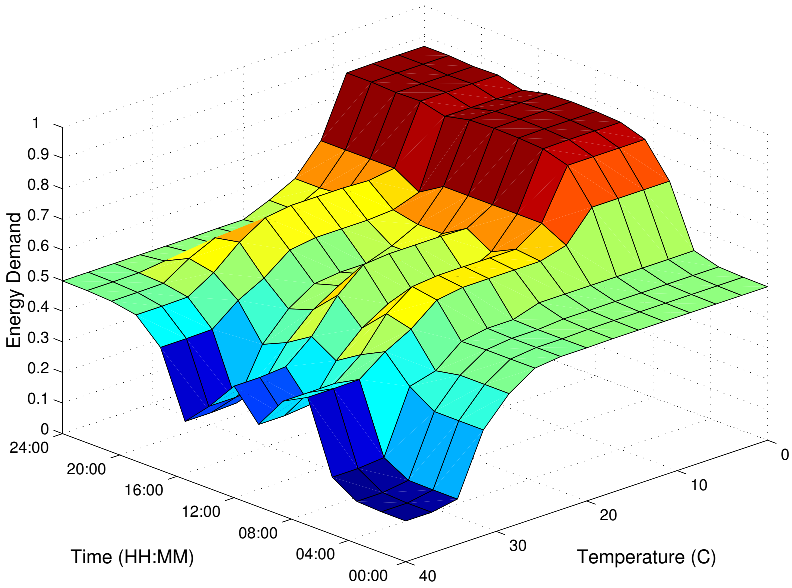


Figure 3.18: Three dimensional plot of the result from Case 3

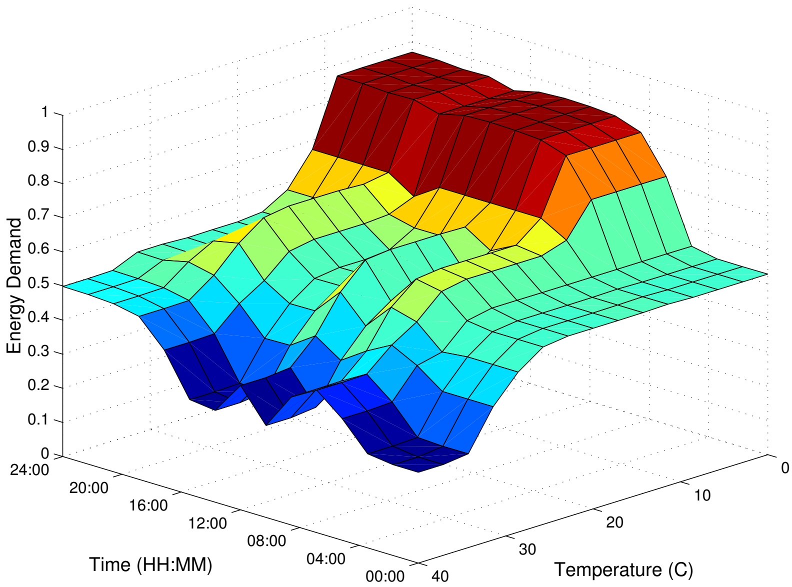


Figure 3.19: Three dimensional plot of the result from Case 4

The maximum predicted result is the same for each system as seen in Table 3.1. This is 0.8679 because it is produced when the consequents of the rules are *High* = 1, *Low* = 0, and *Medium* = 0 and the COG method is used for defuzzification results in 0.8679.

Table 3.1: Summary results from the outputs

System	Mean	STD	Max	Min
Case 1	0.5696	0.1690	0.8679	0.1334
Case 2	0.5822	0.1554	0.8679	0.2659
Case 3	0.5679	0.1688	0.8679	0.1334
Case 4	0.5808	0.1552	0.8679	0.2659

3.4 Discussion

This section has described the well-known properties *normal*, *convex* and *distinct* used in the vast majority of terms implemented in fuzzy systems in the literature. It has been argued that while these properties are undoubtedly useful in the context of fuzzy control, they restrict the more general shapes of terms that might be used within linguistic variables in fuzzy systems. Examples have been given in which potentially useful terms do not adhere to each of these three properties, and a case study is presented to demonstrate the use of such non-regular membership functions in a fuzzy expert system.

We repeat our assertion in our previous ideas [128] that we are not cognitive scientists and are *not* arguing that the unusual membership shapes described in this work are how such concepts are internally represented at a cognitive level. Whether concepts can be non-convex at a cognitive level has been discussed by, for example, Gärdenfors [132], in which he asserts that:

> "*most* properties expressed by simple words in natural language *can* be analysed as convex regions of a domain in a conceptual space" (our italics)

However, while he supports this (rather hedged) assertion with some examples, it remains far from proven. Whatever the reality at the cognitive level, we merely assert that non-regular fuzzy sets may be useful to consider when *modelling* human reasoning in a fuzzy system.

The results of the case study presented in this section has shown that non-regular terms can be used in a fuzzy logic system and they can perform together with regular membership functions. From these illustrations, we firmly believe that non-convex membership functions such as *MealTime* featured in the *Time* (of day) variable are plausible, reasonable membership functions in the sense originally intended by Zadeh.

We are particularly interested in the role of linguistic variables, and their associated terms as used in the fuzzy inferencing process. Within the general category of inferencing (rule-based) systems there are two broad aspects: *control systems* and *expert systems* (emulating human reasoning). Although human reasoning has been investigated since the inception of fuzzy logic (e.g. [4, 133]), by far the majority of published work has been concerned with fuzzy control. Indeed, both the two main methods of implementing fuzzy inferencing, namely the Mamdani method and the Takagi-Sugeno method, were introduced to solve control applications [25, 26].

This historical bias towards the control domain has, we believe, led to a relative neglect of aspects of inferencing in the context of human decision making. Thus, there has been a tendency to restrict membership functions to well-known forms. Triangular, left-shoulder, right-shoulder and trapezoidal, or more generally piecewise linear, functions are common. Also used are standard Gaussian or Sigmoid type curves.

In the case study to illustrate the use of non-convex fuzzy sets, the shapes of terms used in fuzzy systems have adopted several 'conventions'. Terms are almost invariably normalised (having a maximum membership value of 1), convex (having a single maximum or plateau maxima) and distinct (being restricted in their degree of overlap: often expressed as some variation on the concept that all membership values at any point in the universe of discourse sum to 1 across that universe). The shape of these terms are generated by certain accepted membership functions: piecewise linear functions (with restrictions), Gaussians or Sigmoids are almost exclusively used. As such these constitute only a small subset of the total set of possible shapes of terms. These conventions are largely empirical or are justified by arguments based on what might loosely be called 'fuzzy control principles'. However, in many applications involving the modelling of human decision making (expert systems), these traditional membership functions may not provide a wide enough choice for the system developer. They are therefore missing an opportunity to, potentially, produce simpler or better systems. This work extends previous work in which it was suggested that non-convex membership functions might be considered for use in the context of fuzzy expert systems. In particular, the merits of non-convex fuzzy sets are discussed and a case study is presented which investigates whether is is possible to build an expert system featuring usual Mamdani style fuzzy inference in which a time-related non-convex fuzzy set is used together with traditional fuzzy sets. It is shown that this is indeed possible and an examination is made of the resultant output surface generated by four different sub-classes of non-convex membership functions.

3.5 Summary

In this Chapter, the use of non-standard membership functions to better model reasoning in a variety of complex domains, including when modelling human reasoning, has been described. It has been shown that the use of such membership functions has been limited in practice, for no good reason. It is concluded that non-convex membership functions are useful and their further use is encouraged.

In next chapter, type-1 and type-2 fuzzy systems with a varying number of tunable parameters are investigated. Their performance, in their ability to predict the Mackey-Glass time series with various levels of added noise, was compared. The concept of non-deterministic fuzzy reasoning is also presented and how to implement non-deterministic fuzzy sets is described.

Chapter 4

Investigating the Performance of Type-1 and Type-2 Fuzzy Systems for Time-Series Forecasting

4.1 Introduction

As we mentioned in Chapter 2, many decision-making and problem solving tasks are too complex to be understood quantitatively, but by using knowledge that is imprecise rather than precise [1] and [9] it is possible to overcome this. Fuzzy logic resembles human reasoning in its use of approximate information and uncertainty to generate decisions. It was specifically designed to represent uncertainty and vagueness and provide formalized tools for dealing with the imprecision in many real problems. Since knowledge can be expressed more naturally by using fuzzy sets, many complex decision problems can be significantly simplified. Although many applications have been found for type-1 fuzzy logic systems, it is its application to rule-based systems that has most significantly shown its importance as a powerful design methodology, but yet it is unable to model and minimize the effect of all uncertainties. To overcome this limitation, type-2 fuzzy systems can be introduced as they can model uncertainties better and minimize their effects. Type-2 fuzzy systems are characterized by IF-THEN rules, but their antecedent or consequent sets are type-1 or type-2. A type-2 fuzzy set can represent and handle uncertain information effectively. More details about type-2 fuzzy sets and fuzzy systems can be found in

Chapter 2.

The rest of this chapter has been organized as follow. Section 4.2 presents general information about fuzzy logic systems and Mackey-Glass Chaotic Time-Series. Section 4.3 presented about an investigation into the effect of number of model parameters on performance in type-1 and type-2 fuzzy logic systems. SectionSection 4.4 stated an introduction concept of non-deterministic fuzzy reasoning. Finally, Section 4.5 summarised all contents presented in this chapter.

4.2 Time-Series Forecasting Using Type-1 Fuzzy Systems

Time-series forecasting is a forecasting method that uses a set of historical values to predict an outcome. These historic values, often referred to as a *time series*, are spaced equally over time and can represent anything from any period of data such as: yearly, monthly, daily, hourly, and so on call volumes.

Time-series forecasting assumes that a time series is a combination of a pattern and some random error. The goal is to separate the pattern from the error by understanding the pattern's *trend*, its long-term increase or decrease, and its seasonality, the change caused by *seasonal* factors such as fluctuations in use and demand.

4.2.1 Data Sets Preparation

Equation 4.1 has become known as the Mackey-Glass equation. It is a non-linear delay differential equation [1]:

$$\frac{dx(t)}{dt} = \frac{0.2x(t-\tau)}{1+x^{10}(t-\tau)} - 0.1x(t) \tag{4.1}$$

For $\tau > 17$ is known to exhibit chaos. This equation is converted into a discrete time-series equation by using Euler's method as shown in Equation 4.2:

$$f(x,t) = \frac{0.2x(t-\tau)}{1+x^{10}(t-\tau)} - 0.1x(t) \tag{4.2}$$

Then

$$x(t+1) = x(t) + hf(x,t) \tag{4.3}$$

Where $h = 1$ and $\tau = 30$. The initial values of $x(t)$ where $t \leq 30$ are set randomly. The value of each data is in [0,1].

0	1.200000	10	0.441455
1	1.085805	11	0.399445
2	0.982477	12	0.361433
3	0.888982	13	0.327038
4	0.804384	14	0.295916
5	0.727837	15	0.267756
6	0.658574	16	0.242276
7	0.595902	17	0.219220
8	0.539195	18	0.246463
9	0.487884	19	0.307757
			...

Figure 4.1: Sample data sets which generated by Mackey-Glass time-series

The sample of data set after generated is shown in Figure 4.1. The 1200 data points were generated and Figure 4.2 shows the simulation plot of time series.

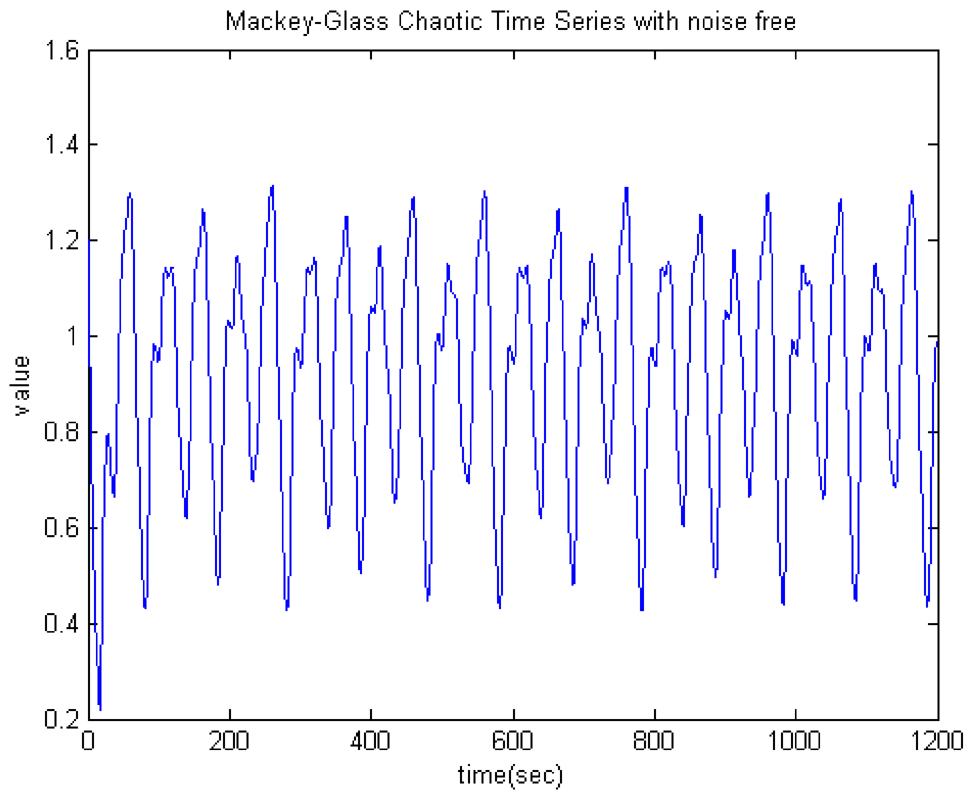


Figure 4.2: The time-series data plot with noise free

In this simulations, the sampled time series $x(k)$ is corrupted by uniformly distributed additive noise $n(k)$ and only noisy measured values of $s(k) = x(k) + n(k), k = 1, 2, \cdots, N$. For this simulation, the random noise was applied to the previous data by the following function.

$$S(k) = x(k) + randn()$$

The sample generated noise is shown in Figure 4.3

0	-0.0003	10	0.0007
1	0.0011	11	-0.0003
2	-0.0011	12	0.0006
3	0.0005	13	-0.0002
4	0.0012	14	-0.0005
5	-0.0010	15	-0.0010
6	0.0009	16	0.0020
7	0.0007	17	0.0003
8	-0.0004	18	0.0010
9	-0.0004	19	-0.0002
			...

Figure 4.3: The sample generated noise to be added into the data sets

The 1200 noisy data sets were generated and Figure 4.4 shows the simulation plot of time series with noise corrupted.

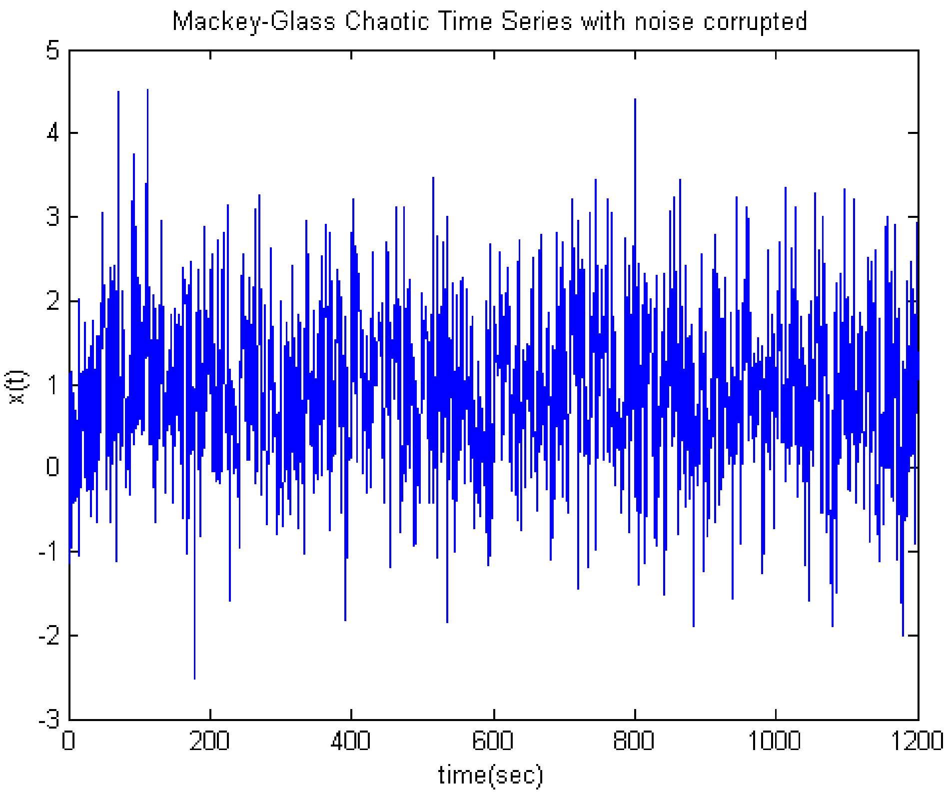


Figure 4.4: The time-series data plot with added noise

This experimental is based on $N = 1000$ points, $s(125), s(126), \cdots, s(1124)$. The first 500 data series $s(125), s(126), \cdots, s(624)$ are for training data set and the remaining 500 data series $s(625), s(626), \cdots, s(1124)$ are for testing data set.

4.2.2 Designing Type-1 Fuzzy Systems

In time series prediction we want to use known values of the time series up to the point in the time, t, to predict the value at some point in the future, $t+P$. The standard method for this type of the prediction is to create a mapping from D sample data points. Sampled every P units in time,

$$[s(t-(D-1)P), \cdots, s(t-P), s(t)]$$

To predict future value $s(t+P)$. we define $D=4$ and $P=1$. For each t, the input data is a four dimensional vector of the following form.

$$I(t) = [s(t-3)s(t-2)s(t-1)s(t)]$$

The output data is agreed to be the trajectory prediction.

$$O(t) = s(t+1)$$

In Figure 4.5 shows sample training data sets.

1.9370	1.7471	2.9251	0.2807	-0.6397
2.1747	1.0506	4.5190	1.0084	1.4378
2.1067	0.4479	2.1470	2.0623	1.1803
0.7460	3.3930	0.5041	-0.2267	1.5298
0.4123	1.3342	0.2845	0.0062	0.0902
0.4979	1.3185	2.1646	1.9048	0.5808
1.7471	2.9251	0.2807	-0.6397	0.9331
1.0506	4.5190	1.0084	1.4378	2.9545
0.4479	2.1470	2.0623	1.1803	1.0094
3.3930	0.5041	-0.2267	1.5298	1.3769
...	...	...	...	...

Figure 4.5: Sample of the training data sets

-0.1187	1.3142	2.5808	2.0849	1.8184
0.2340	1.2073	1.5794	1.0438	0.4313
1.9348	1.3365	0.0881	1.0777	-0.2672
0.8570	1.1163	0.6772	1.3750	0.0926
1.7157	1.2497	1.6080	2.3807	0.5720
1.4041	1.9959	1.5475	2.0937	1.1655
1.3142	2.5808	2.0849	1.8184	1.0890
1.2073	1.5794	1.0438	0.4313	0.6541
1.3365	0.0881	1.0777	-0.2672	1.4494
1.1163	0.6772	1.3750	0.0926	0.3657
…	…	…	…	…

Figure 4.6: Sample of the testing data sets

Where, The first 4 columns represent the inputs and the last column represents the output. And the training data set has 500 records. Similar to the testing data set also has 500 records. In Figure 4.6 shows the sample of testing data sets.

we assign 4 antecedents for forecasting, i.e $s(t-3), s(t-2), s(t-1)$, and $s(t)$ to predict $s(t+1)$, we use only two membership functions for each antecedent, so the number of rule is $2^4 = 16$ rules. The initial locations of antecedent membership functions are based on the mean and standard deviation of the first 500 points, i.e., training data. The initial membership functions are shown in Figure 4.7.

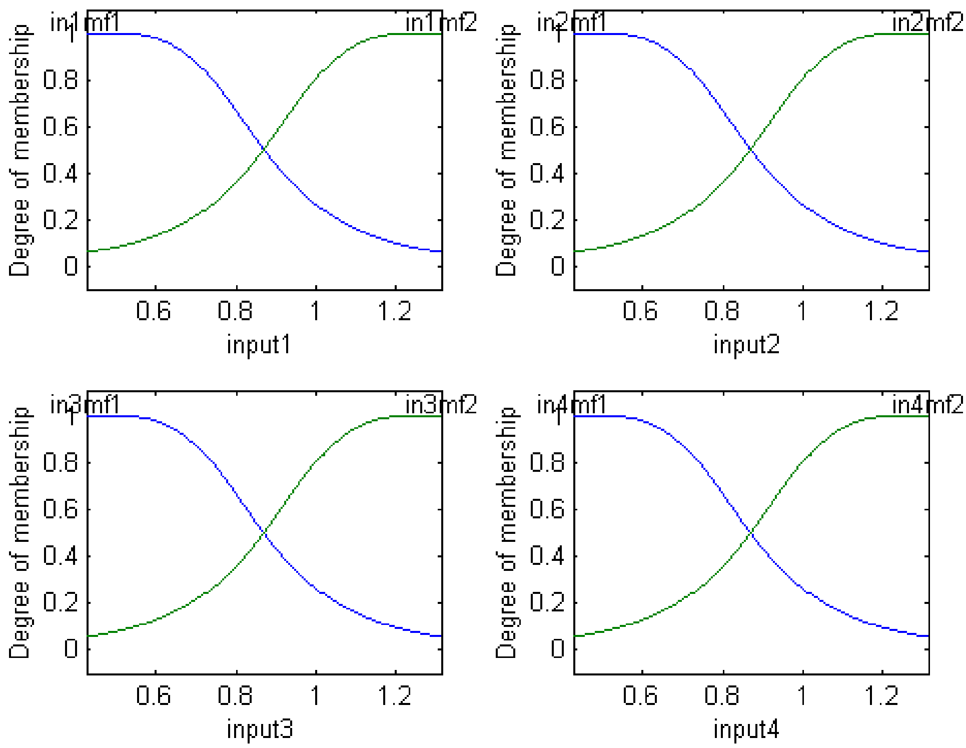


Figure 4.7: Initial membership functions for 4 inputs

4.2.3 Results

After the type-1 fuzzy system has been constructed, the parameters of type-1 system have been tuned to train the type-1 system by using the training data sets (500 records) and tested by using the testing data sets (500 records). The *root mean square error* (RMSE) has been calculated and recorded to compute the mean of RMSE. Figure 4.8 shows the plot between the mean of RMSE and epochs for the data sets with noise free. In Figure 4.9 shows the plot between Mackey-Glass time-series and the ouputs from type-1 fuzzy system obtained from the output of type-1 fuzzy system with noise free data sets and the plot of the prediction errors.

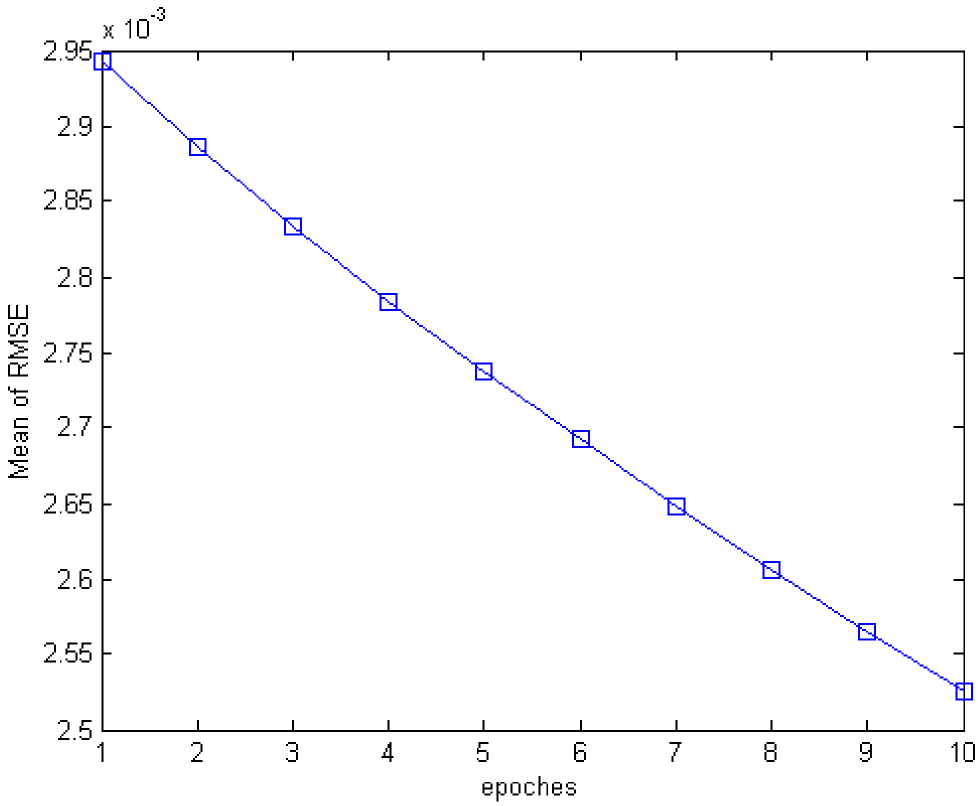


Figure 4.8: Plot of mean of RMSE vs epochs for the data sets with noise free

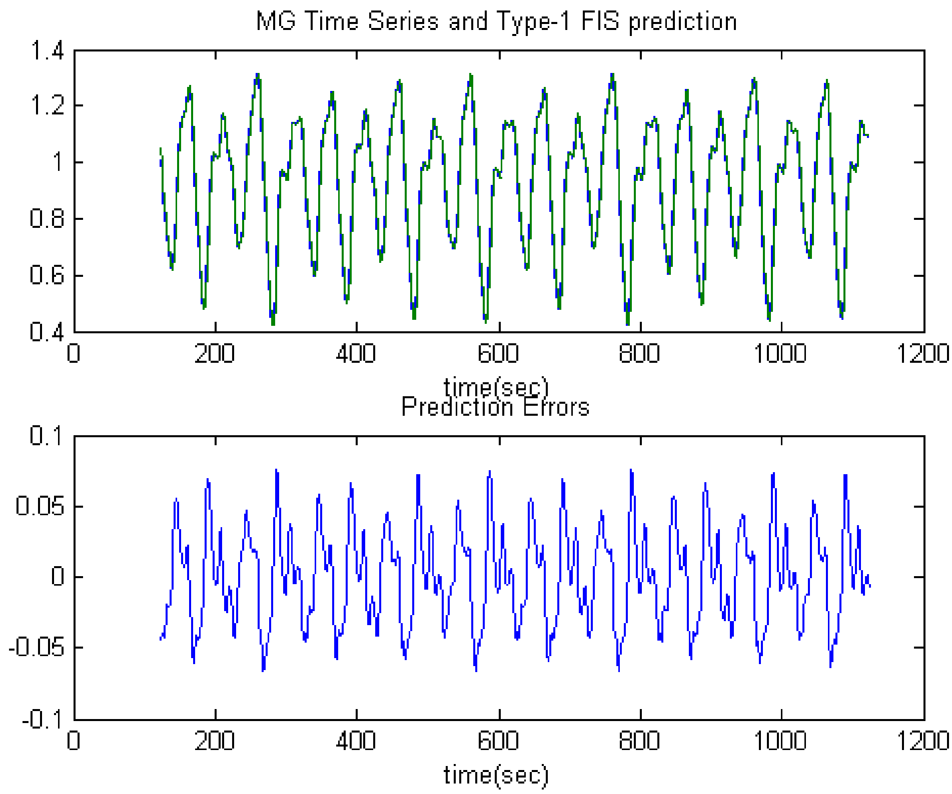


Figure 4.9: Mackey-Glass time-series vs type-1 fuzzy system prediction for the data set with noise free

Figure 4.10 shows the plot between the mean of RMSE and epochs for the data sets with added noise. In Figure 4.11 shows the plot between Mackey-Glass time-series and the ouputs from type-1 fuzzy system obtained from the output of type-1 fuzzy system with additive noise data sets and the plot of prediction errors.

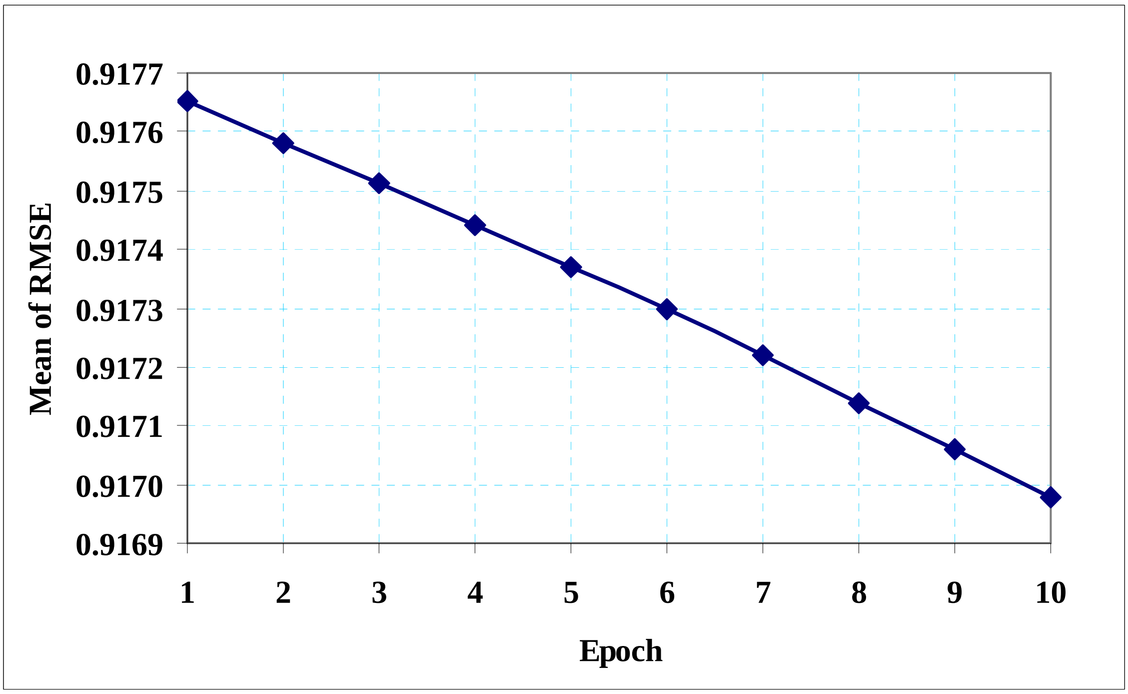


Figure 4.10: Plot of mean of RMSE vs epoch for type-1 fuzzy system with additive noises data sets

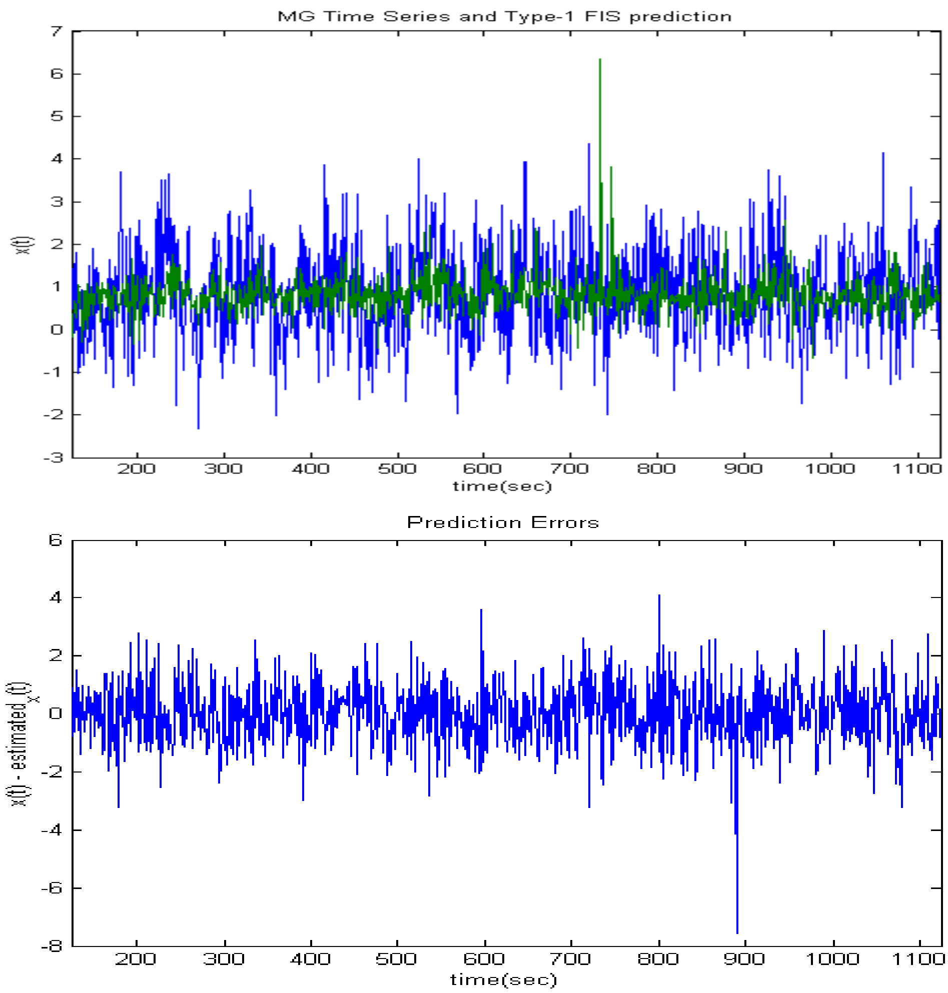


Figure 4.11: Mackey-Glass time-series vs type-1 fuzzy system prediction for the data set with noise added

As we seen from the results obtained from type-1 fuzzy systems, it is obvious that type-1 fuzzy system can predict the time-series forecasting with effectively performance for the data sets with noise free. In contrast, when the noise has been corrupted into the data sets (uncertainty in data-series) the performance of the type-1 fuzzy system is much larger errors. The type-2 fuzzy system has then been used to predict the time-series forecasting for the the data sets with additive noises. The results have been shown in Figure 4.12. It has been shown that the type-2 fuzzy system can perform much better results than type-1 fuzzy system (Figure 4.10).

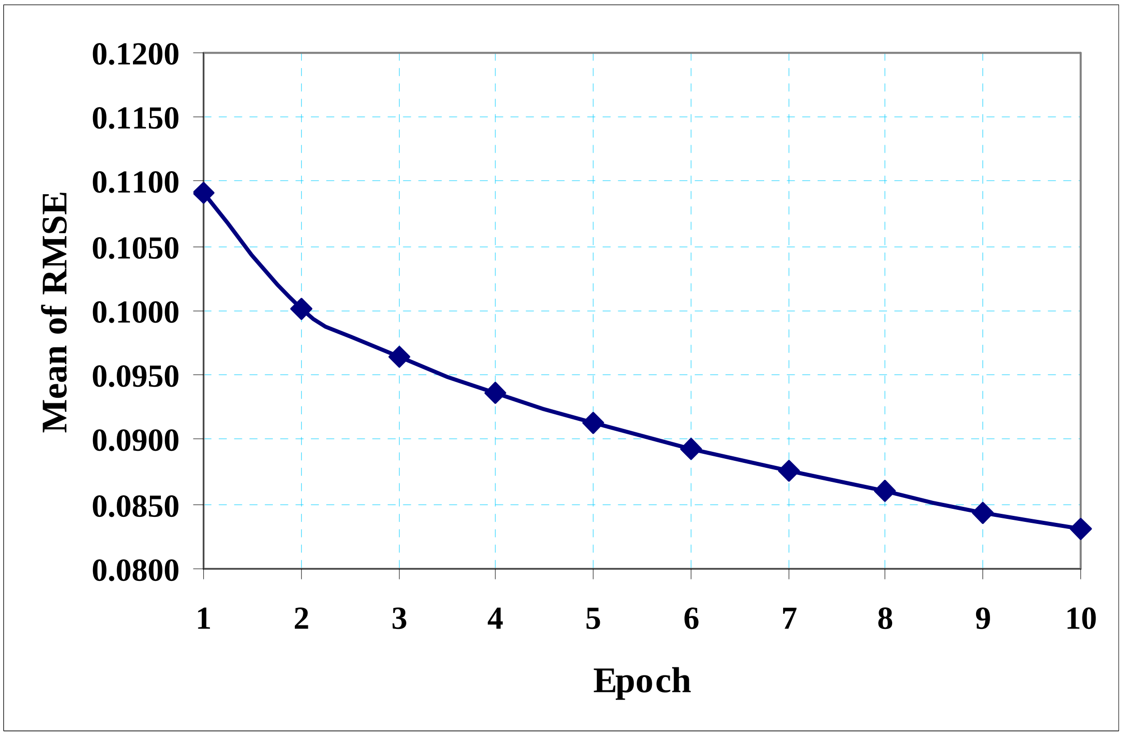


Figure 4.12: Plot of mean of RMSE vs epoch for type-2 fuzzy system with additive noises data sets

In Figure 4.13 shows the plot between Mackey-Glass time-series and the outputs from type-1 and type-2 (including predicted upper and lower bounds) fuzzy systems. It can be seen that the performance of type-2 fuzzy system is better than those in type-1 fuzzy system. But when we carefully considered into the number of model parameters type-2 fuzzy system has much larger number of parameters than type-1 fuzzy system. This curious finding led us to investigate the effect of number of model parameters on performance in type-1 and type-2 fuzzy systems which is shown in next section (section 4.3).

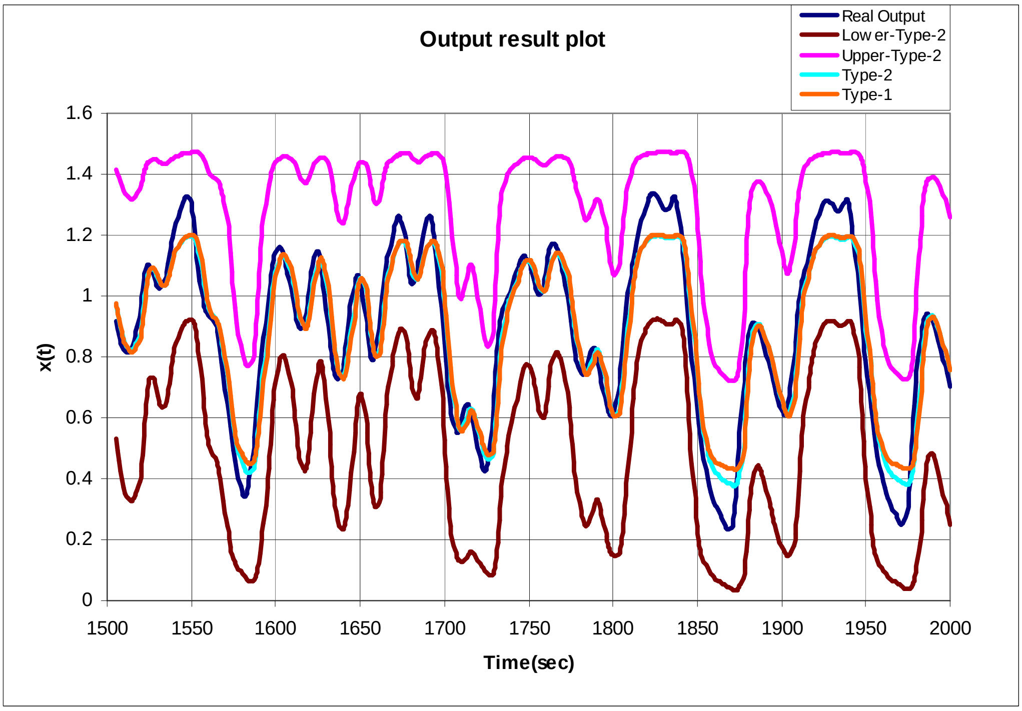


Figure 4.13: Mackey-Glass time-series vs type-1 and type-2 prediction for the data set with noise added

4.3 Investigation into the Effect of Number of Model Parameters on Performance in Type-1 and Type-2 Fuzzy Logic Systems

In 1977, Mackey and Glass published a paper in which they associate the onset of disease with bifurcations in the dynamics of first-order differential-delay equation, which model physiological systems. The Mackey-Glass time series has become one of the benchmark problems for time-series prediction in both the neural network and fuzzy logic areas. Mendel and Karnik [1, 44] have carried out experiments into forecasting Mackey-Glass Chaotic Time-series with noisy data by using type-1 and type-2 fuzzy systems. They have suggested that an interval non-singleton type-2 fuzzy system with type-1 fuzzy sets achieves the best performance and can be used in a real-time adaptive environment. Although the authors noted that the type-2 fuzzy system used in their experiment featured more internal tunable model parameters than the type-1 fuzzy system, they did not go on to investigate whether it was simply the number of model parameters that was responsible for the performance gains achieved.

The purpose of this work was firstly to attempt to reproduce the results of Mendel and Karnik and secondly to perform a careful analysis of whether the performance of the type-2 fuzzy system could be matched or surpassed by type-1 models with a similar or greater number of internal tunable model parameters. Four main classes of fuzzy systems are considered:

(i) T1-SFLS - 'conventional' fuzzy systems with singleton inputs and type-1 fuzzy sets throughout;

(ii) T1-NFLS - type-1 fuzzy systems with non-singleton (type-1) fuzzy inputs and type-1 fuzzy sets throughout;

(iii) T2-SLFS - fuzzy systems featuring interval type-2 sets with singleton inputs;

(iv) T2-NSLFS-T1 - fuzzy systems featuring interval type-2 sets with type-1 non-singleton inputs.

This work has evolved from recent studies on modelling of non-deterministic reasoning using type-2 fuzzy systems [10]. The software used in this experiment is that provided by Professor Mendel (at http://sipi.usc.edu/ mendel/software/).

4.3.1 Methodology

Five independent data sets with 5 different noise levels (in total 25 data sets 2200 series each) were generated. These data sets were generated by using Mackey-Glass time-series delay differential equation shown in Equation 4.1 above. After 5 data sets were generated, 5 different level of noise were generated as follows:

- Level 1: 0 noise (noise free)
- Level 2: 0.01 noise

- Level 3: 0.05 noise
- Level 4: 0.10 noise
- Level 5: 0.20 noise

where *noise* was a uniformly distributed random number in [-1,1]. Then these 5 different levels of noise were added into the data sets.

Type-1 singleton fuzzy logic system (SFLS) and type-1 non-singleton fuzzy logic system (NSFLS) have been designed with 4 antecedents, 2 and/or 3 membership functions for each antecedent, the number of rules are 16 and/or 81 rules (2^4 and/or 3^4) respectively, each rule is characterized by 8 antecedent MF parameters (means and standard deviations), and 1 consequent parameter ($\bar{y}$). The initial location of each Gaussian antecedent MF is based on the mean (m_x) and standard deviation (σ_x) and the mean of membership functions are:

- 2 MFs = $[m_x - 2\sigma_x, m_x + 2\sigma_x]$
- 3 MFs = $[m_x - 2\sigma_x, m_x, m_x + 2\sigma_x]$

Initially all standard deviation parameters are tuned to σ_x or $2\sigma_x$. Additionally the height defuzzifier and initial centre of each consequent's MF are random numbers in [0,1]. So, the total number of tunable parameter for Type-1 SFLS with 2 and 3 membership functions are 144 and 729, respectively. For type-1 NSFLS each of the 4 noisy input measurements are modelled using a Gaussian membership function, a different standard deviation is used for each of the 4 input measurement membership functions (σ_n). So, the total number of tunable parameters for Type-1 NSFLS with 2 and 3 membership functions are 145 and 730 respectively. Finally, 4 different models were created for both type-1 SFLS and type-1 NSFLS for each data set (25 data sets).

Interval type-2 singleton FLS (Type-2 SFLS) and type-1 non-singleton type-2 FLS (Type-2 NSFLS-T1) have been designed by using the partially dependent approach. First, the best possible singleton and non-singleton type-1 fuzzy systems were designed by tuning their parameters using back-propagation designs, and then some of those parameters were used to initialise the parameters of the interval type-2 SFLS and type-2 NSFLS-T1.

They consisted of 4 antecedents for forecasting, 2 membership functions for each antecedent and 16 rules. The Gaussian primary membership functions of uncertain means for the antecedents were chosen. The means of membership functions are:

- Mean of MF1 = $[m_x - 2\sigma_x - 0.25\sigma_n, m_x - 2\sigma_x + 0.25\sigma_n]$
- Mean of MF2 = $[m_x + 2\sigma_x - 0.25\sigma_n, m_x + 2\sigma_x + 0.25\sigma_n]$

where m_x is the mean of the data in the training parts, and σ_n is the standard deviation of the additive noise. Each rule of the type-2 SFLS and type-2 NSFLS-T1 were characterized by 12 antecedent MF parameters: left and right bounds on the mean, and the standard deviation for each of 4 Gaussian membership functions) and 2 consequent parameters (left and right hand end-points for the centroid of the consequent type-2 fuzzy set). So, in total the number of parameters tuned for type-2 SFLS is 224. Standard deviation for each of the 4 input measurement membership functions (σ_n) is used in type-2 NFLS-T1. So in total the number of parameters tuned for type-2 SFLS is 225.

Initially the final tuned results were used for the standard deviation of the input, σ_x or $2\sigma_x$, obtained from type-1 NSFLS design, and also $\bar{y}^i$ was obtained from type-1 SFLS and then initial $\bar{y}^i_r$ and $\bar{y}^i_l$ was chosen as:

$$\bar{y}^i_r = \bar{y}^i + \sigma_n$$

$$\bar{y}^i_l = \bar{y}^i - \sigma_n,$$

where i = 1, 2, ..., 16

Finally, two different models for both type-2 SFLS and type-2 NSFLS-T1 were created for each data set (totally 25 data sets). All designs mentioned above were tuned using steepest descent algorithm in which all of the learning parameters were set equal to the same value, 0.2. Training and testing were carried out for ten epochs. After each epoch the testing data was used to see how each fuzzy system performed, by computing root mean square error (RMSE).

All designs above were also based on 1,000 noisy data points: x(501), x(502), ..., x(1500). The First 500 noisy data, x(501), x(502), ..., x(1000) were used for training, and the remaining 500 , x(1001), x(1002), ..., x(1500), were used for testing the design. Four antecedents: x(k-3), x(k-2), x(k-1), and x(k) were used to predict x(k+1).

The performance of all the designs was evaluated using the RMSE as shown below:

$$RMSE = \sqrt{\frac{1}{500}\sum_{k=1000}^{1499}[x(k+1) - f(x^{(k)})]^2} \qquad (4.4)$$

where $x^{(k)} = [x(k-3), x(k-2), x(k-1), x(k)]^T$.

4.3.2 Results

After all models had been constructed and run, the performances of the type-1 and type-2 fuzzy systems were compared. The results of performance of 12 different models are shown as follows. Table 4.1 shows the number of parameters that were used in the experiment for each of the 12 models. Table 4.2 shows the result obtained from the mean of RMSE of the best model for 12 different fuzzy system models with five noise levels.

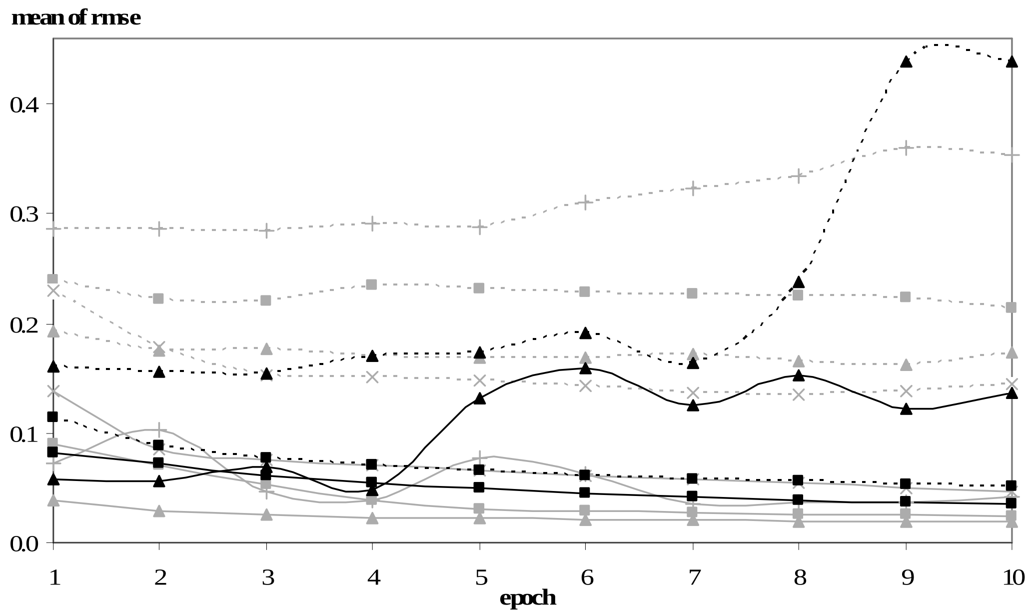


Figure 4.14: Graph of mean of RMSE of 12 models for noise level 1

Table 4.1: Number of parameters of each design

No.	**FLS**	**No. of Parameters**
M1	T1-SFLS-2mf-2σ	144
M2	T1-SFLS-2mf-σ	144
M3	T1-SFLS-3mf-2σ	729
M4	T1-SFLS-3mf-σ	729
M5	T1-NSFLS-2mf-2σ	145
M6	T1-NSFLS-2mf-σ	145
M7	T1-NSFLS-3mf-2σ	730
M8	T1-NSFLS-3mf-σ	730
M9	T2-SFLS-2mf-2σ	224
M10	T2-SFLS-2mf-σ	224
M11	T2-NSFLS-2mf-2σ	225
M12	T2-NSFLS-2mf-σ	225

Table 4.2: The mean of RMSE of the best model for 12 different fuzzy system models with 5 different noise levels

RMSE	**N1**	**N2**	**N3**	**N4**	**N5**
M1	0.2143	0.2351	0.2236	0.1253	0.1745
M2	0.2864	0.2447	0.2009	0.2071	0.2636
M3	0.1348	0.1326	0.1353	0.1343	0.1364
M4	0.1618	0.1723	0.1860	0.1800	0.1349
M5	**0.0243**	**0.0244**	**0.0274**	**0.0219**	**0.0226**
M6	0.0353	0.0555	0.0530	0.0467	0.0350
M7	0.0469	0.0469	0.0468	0.0466	0.0468
M8	**0.0189**	**0.0213**	**0.0239**	**0.0188**	**0.0211**
M9	0.0508	0.0285	0.0700	0.0529	0.0427
M10	0.1537	0.0450	0.0664	0.1291	0.1239
M11	**0.0264**	**0.0215**	**0.0243**	**0.0202**	**0.0216**
M12	0.0489	0.0263	0.0640	0.0434	0.0834

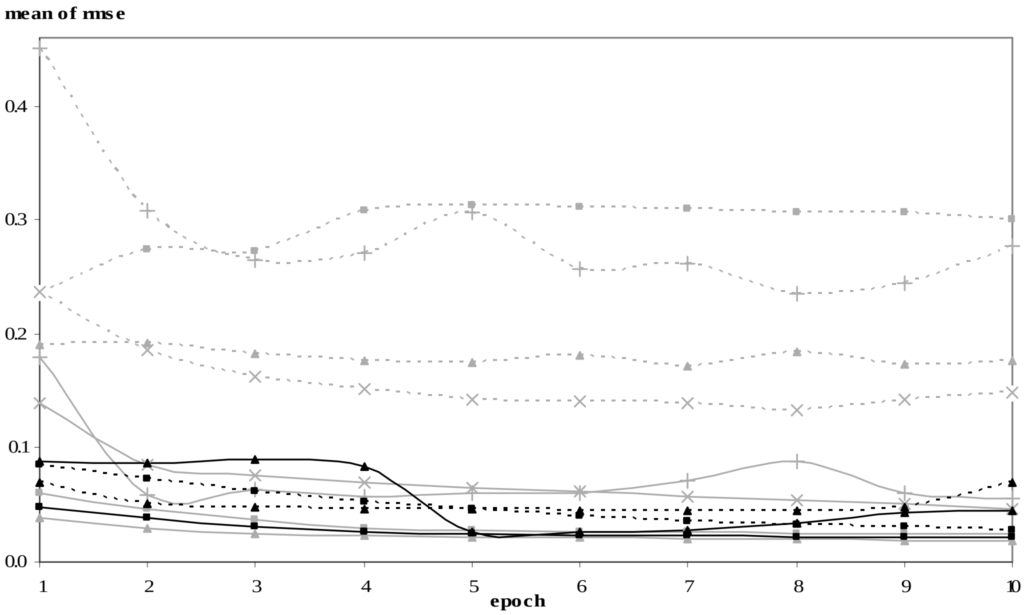


Figure 4.15: Graph of mean of RMSE of 12 models for noise level 2

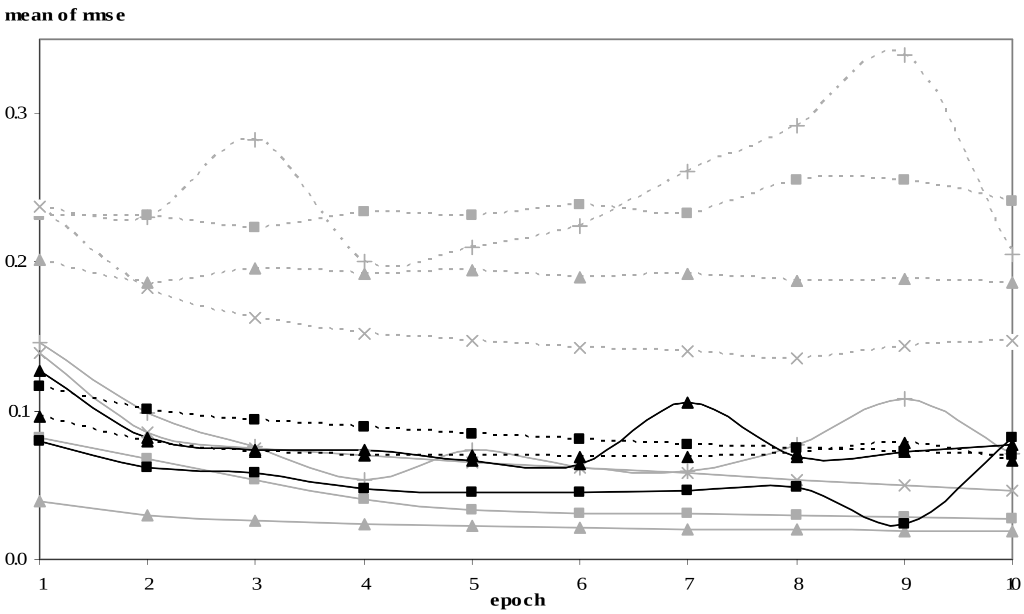


Figure 4.16: Graph of mean of RMSE of 12 models for noise level 3

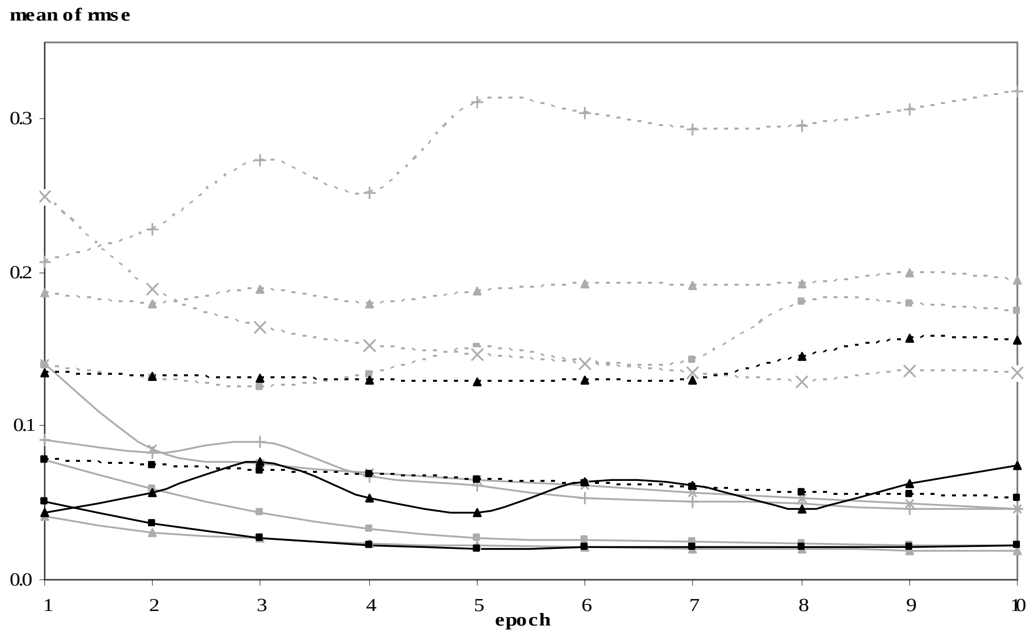


Figure 4.17: Graph of mean of RMSE of 12 models for noise level 4

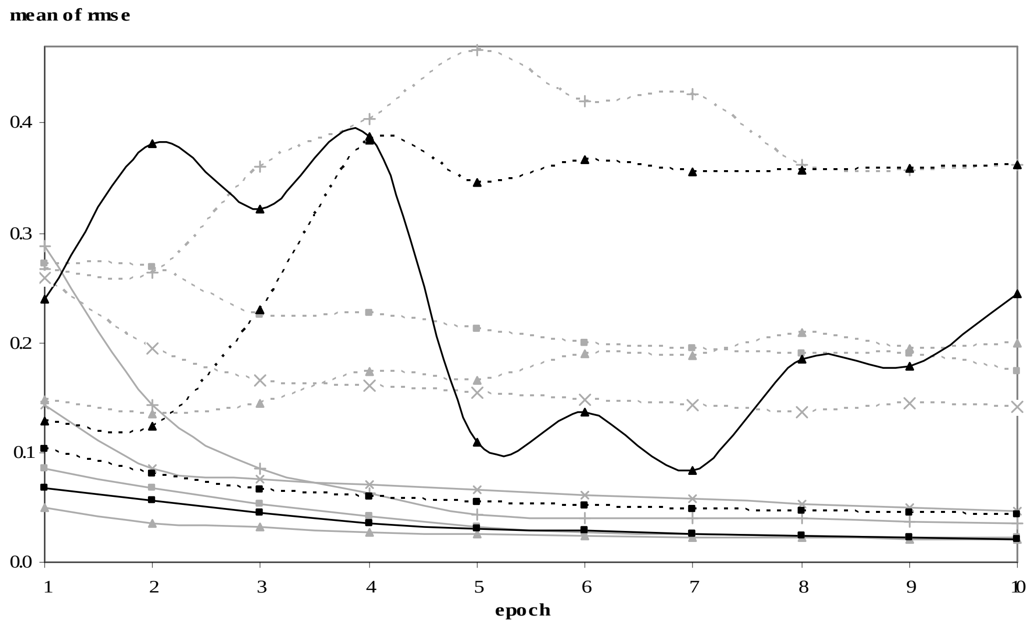


Figure 4.18: Graph of mean of RMSE of 12 models for noise level 5

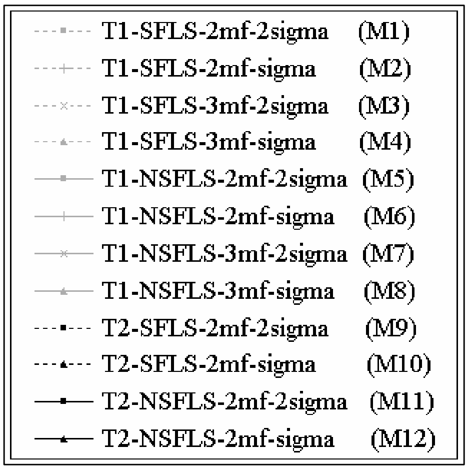


Figure 4.19: Key for figures 4.14- 4.18

Figures 4.14, 4.15, 4.16, 4.17, and 4.18 show the performance (RMSE) of 12 different models for the 5 different noise levels averaged over five separate runs, while Figure 4.19 shows the key that applies to Figures 4.14 - 4.18. Figures 4.20, 4.21, 4.22, 4.23, and 4.24 show the performance of just models M5, M8, and M11 with the y-axis expanded for more detail. From Figures 4.14 and 4.24, it can be seen that M5 and M8 show better results after epoch 3 than M9 (T2-SFLS) and M11 (T2-NSFLS), probably because these data sets are noise free and this finding agrees with Mendel's result.

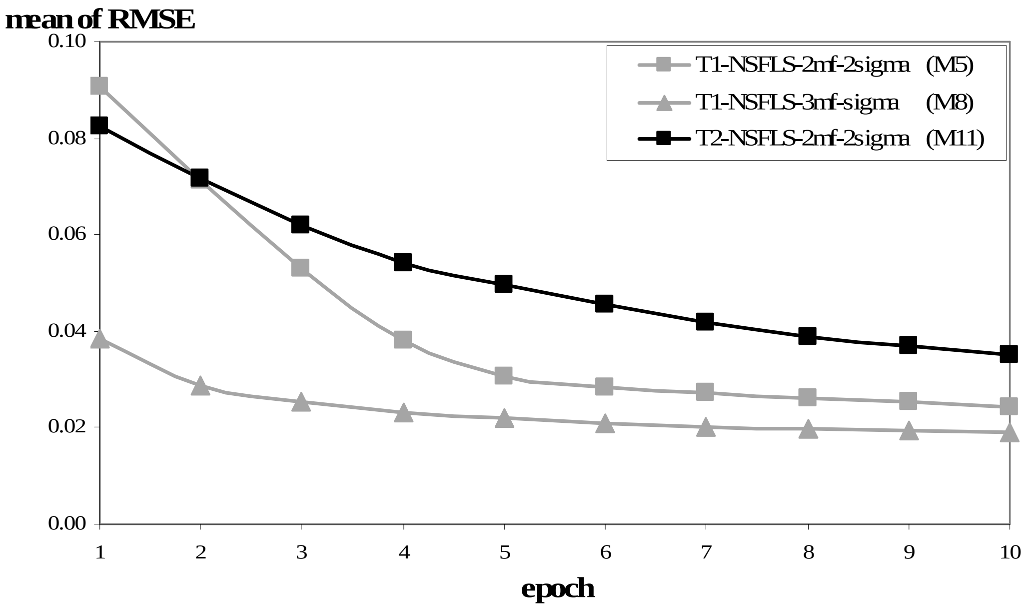


Figure 4.20: Graph of mean of RMSE of M5, M8 and M11 for noise level 1

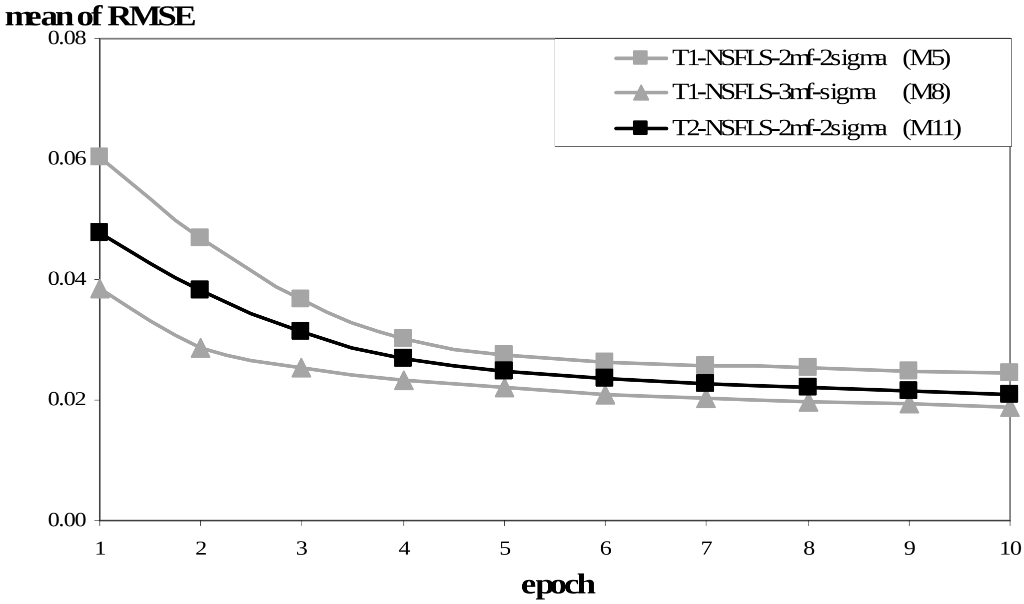


Figure 4.21: Graph of mean of RMSE of M5, M8 and M11 for noise level 2

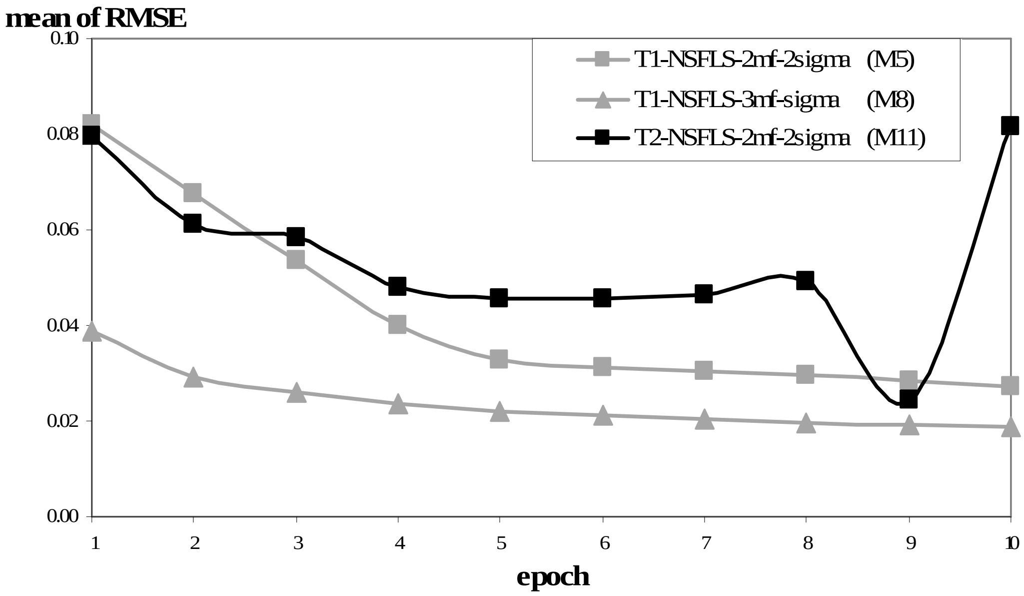


Figure 4.22: Graph of mean of RMSE of M5, M8 and M11 for noise level 3

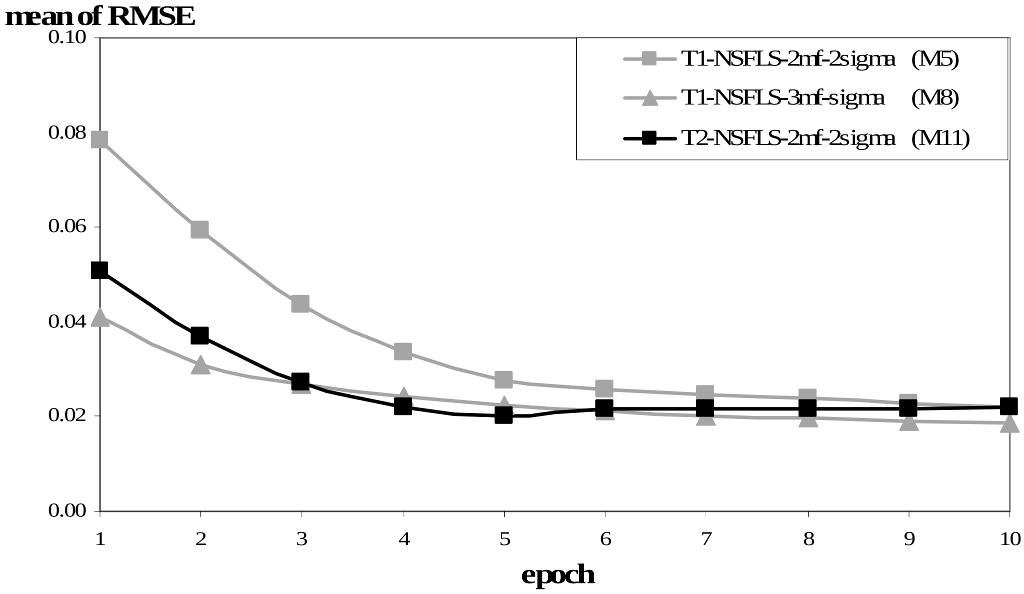


Figure 4.23: Graph of mean of RMSE of M5, M8 and M11 for noise level 4

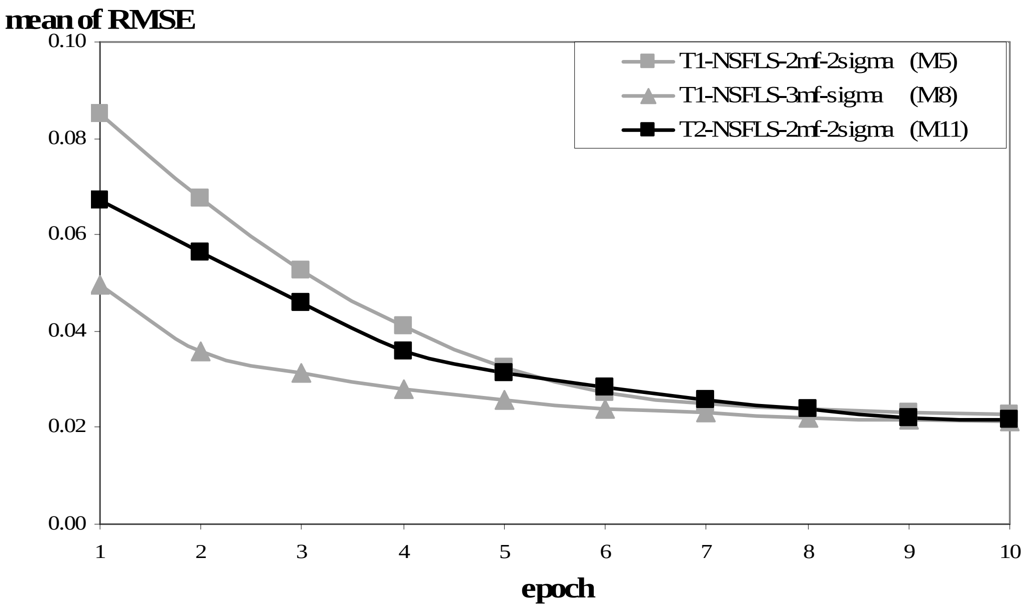


Figure 4.24: Graph of mean of RMSE of M5, M8 and M11 for noise level 5

4.3.3 Discussion

All cases the performance of type-1 fuzzy systems with singleton (crisp) inputs (M1 - M4), the most common found in practice, is worse than for the type-1 non-singleton fuzzy systems and the type-2 fuzzy systems. This is regardless of the number of parameters in the systems. Particularly, it should be noted that M3 and M4, each featuring 729 tunable parameters, whilst better than M1 and M2, achieve far worse performance than type-1 non-singleton or type-2 fuzzy systems with far fewer parameters (M5, M6, M9 and M10). This suggests that a high number of model parameters is not in itself sufficient to produce good performance.

With zero noise, M5 (with only 145 parameters) achieves better performance than M11 or M12. This agrees with Mendel's previous findings that in the absence of noise a type-1 fuzzy system with non-singleton inputs is an adequate model for capturing the uncertainty.

The best overall performance is achieved with M8. This is a type-1 fuzzy system with non-singleton inputs and with 3 membership functions for each variable, leading to a high number of tunable model parameters (730). From this, we may tentatively suggest that while type-2 fuzzy systems may not strictly be necessary in order to achieve 'optimal' performance, their benefit may lie more in achieving good performance in a more tractable model. Note also that M5 (T1-NSFLS with 'only' 145 tunable parameters) achieves very good performance, albeit slightly worse than the best models.

The best 3 models, M5 M8 and M11, are captured as in figures 4.20, 4.21, 4.22, 4.23, and 4.24. The comparison between the performances of M8 and M11 by using **Mann-Whitney U-test** has been found that M8 performs better than M11 with statistical significant at 95% in noise level 1, 3, 4, and 5. In noise level 2, M8 also performs better M11 but not statistical significant at 95%. So, the conclusion can be made that more number of parameters tuned can improve the T1-NSFLS's performance to be as good as or better than type-2 fuzzy systems.

Finally, we emphasise that these finding are for one particular data set (MG-Time series) only and hence, no general conclusions can be made from them alone. In order to reach more common conclusions it would be necessary to carry out similar experiments on a wide variety of data sets. There is no evidence at present to suggest that the similar results would necessarily be obtained for other kinds of data.

4.4 Non-Deterministic Fuzzy Reasoning

The purpose of developing an expert systems (based on fuzzy logic or not) is to encapsulate knowledge and expertise and use it like a human expert. Type-1 fuzzy logic systems, like a classical expert systems, are deterministic in the sense that for the same inputs the outputs are always the same. However, human expert exhibit a non-deterministic behaviour in decision making. Variation may occur among the decisions of a panel of human experts (inter-expert) as well as in the decision of an individual expert (intra-expert) for the same inputs. Understanding the dynamics of the variation in human decision making could allow the creation of *truly intelligent* systems that cannot be differentiated from their human counterparts. Moreover, in application areas where having an expert constantly available is not possible, such systems can produce a span of decisions that may be arrived at by a panel of experts.

Recently, Garibaldi *et al* have studied non-determinism by enhancing the fuzzy logic system developed in an earlier work [134]. The rule of the original fuzzy logic system were elicited in conjunction with several experts who took part in its development. However, when presented with the same data although the fuzzy system produced the same output each time, the same input was given. it was observed that the experts's conclusions varied both among themselves and from their previous conclusions.

For example, six expert clinicians who took part in the development of the earlier system (type-1 fuzzy logic system) were ask to rank 5 UAB assessments in terms of perceived likelihood of having suffered brain damage due to lack of oxygen. Figure 4.25 shows the rankings of 50 UAB assessments by six experts against the type-1 fuzzy logic system. A perfect agreement, which would be a straight line from (0,0) to (50,50), is the ideal desired result. However, as can be seen from Figure 4.25, there is neither perfect agreement with the fuzzy logic system nor among the experts. It can also be observed that at the extreme cases the experts tend to agree with each other and the fuzzy logic system but in the cases that fall in the middle of the range, there is less agreement. The distribution presents the characteristic of an elliptic envelope around the diagonal line from (0,0) to (50,50).

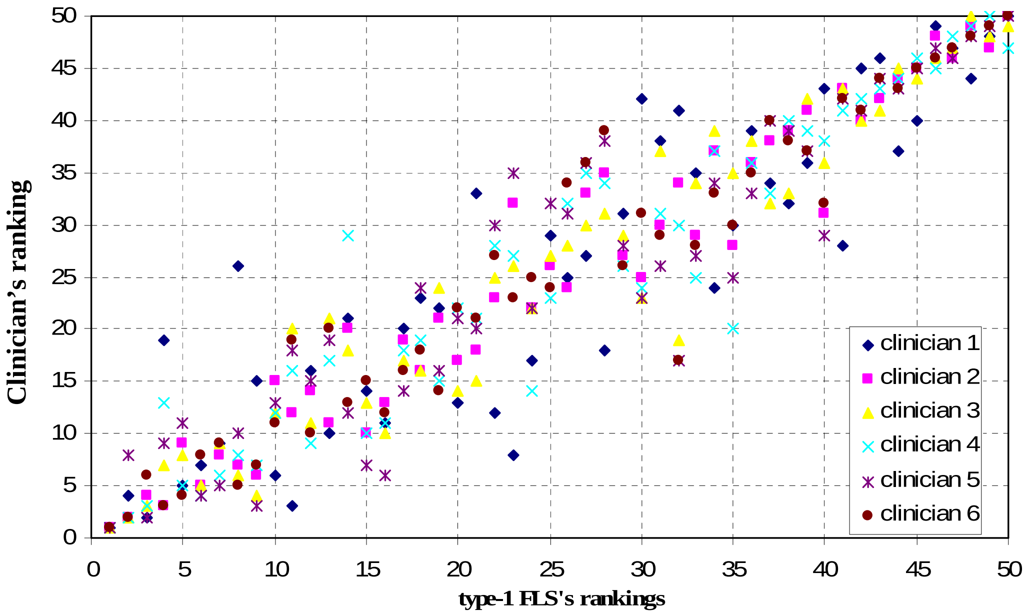


Figure 4.25: Variation in rankings of 50 assessments (original from [14])

The source of this non-determinism was suspected to be lying in the different interpretations of the linguistic terms used is the rules. The original fuzzy logic system used type-1 membership functions which are precise and cannot reflect the vagueness in the terms that they represent. However, these terms have different meaning for different experts and their interpretations may also vary depending on the environmental conditions or over time.

To explore the relationship between the vagueness of the terms used in an fuzzy logic system and the variation in its decision making, Garibaldi have carried out the experiments by introducing uncertainty to the membership functions associated with the linguistic terms.

Garibaldi *et al* [10–14] have been investigating the incorporation of variability into decision making in the context of fuzzy expert systems in the medical domain. In this work, Garibaldi proposed the notion of 'non-deterministic fuzzy reasoning' in which variability is introduced into the membership functions of a fuzzy system through the use of random alterations to the parameters of the generating function(s). For examples, in case of Gaussian membership functions — the primary membership functions are illustrated in figure 4.26. There are three alternative kinds of non-determinism have been proposed:

- Introducing the variability into the centre of membership functions (variation in location)

- Introducing the variability into the width of membership functions (variation in slope)
- Introducing the variability into the value of membership functions (noise variation)

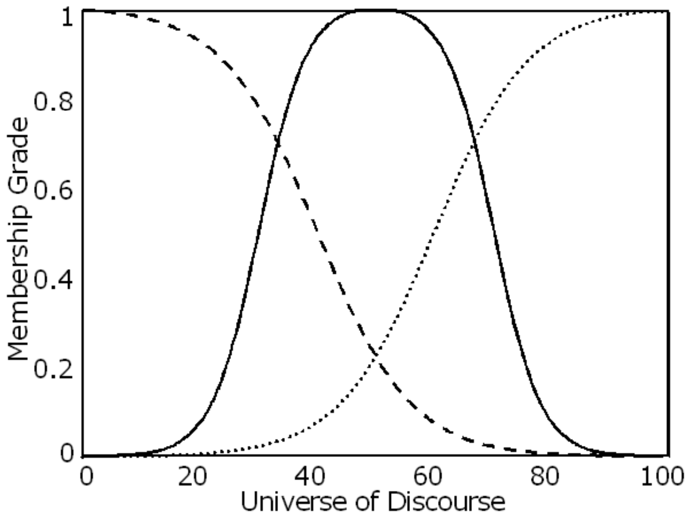


Figure 4.26: Illustration of the primary membership functions (original from [14])

Figure 4.27 illustrates the non-determinism membership functions with centre variation by shifting the centre of the primary membership functions (C_1, C_2, and C_3) by the amount of Δ ($\acute{C}_1$, $\acute{C}_2$, and $\acute{C}_3$).

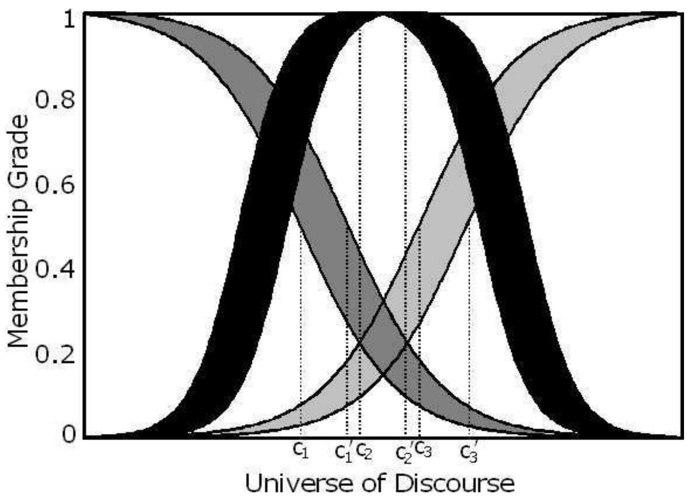


Figure 4.27: Illustration of non-deterministic membership functions with centre variation (original from [14])

Figure 4.28 illustrates the non-deterministic membership functions with width variation by shifting the standard deviation (width) of the primary membership functions (σ_1, σ_2, and σ_3) by the amount of Δ ($\acute{\sigma}_1$, $\acute{\sigma}_2$, and $\acute{\sigma}_3$).

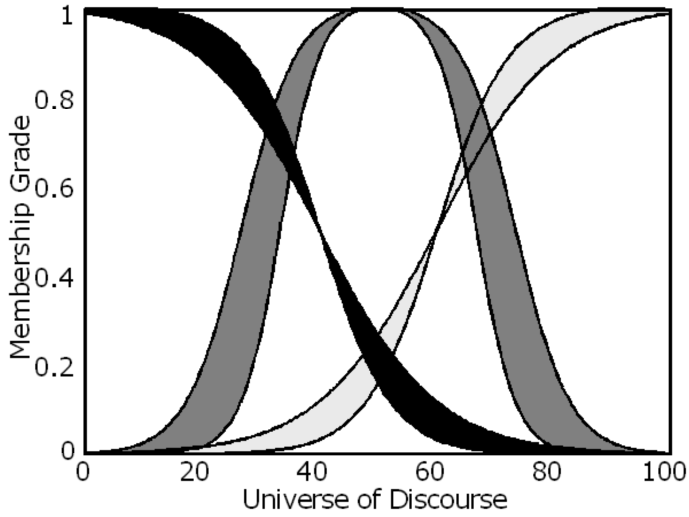


Figure 4.28: Illustration of non-deterministic membership functions with width variation (original from [14])

Figure 4.29 illustrates the non-deterministic membership functions by adding noise into the value of primary membership functions by the amount of Δ (ε_1, ε_2, and ε_3).

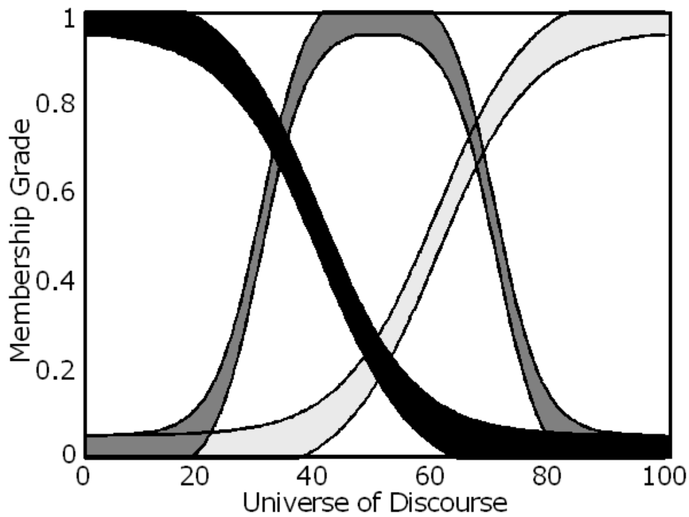


Figure 4.29: Illustration of non-deterministic membership functions with white noise (original from [14])

4.5 Summary

In this chapter, an investigation was carried out in which the performance of type-1 and type-2 fuzzy systems with varying number of tunable parameters were compared in their ability to predict the Mackey-Glass time series with various levels of added noise. Each of the fuzzy systems were tuned to achieve the best possible performance using a standardised gradient descent procedure. This experiments were repeated a number of times in order to establish the mean performance of each fuzzy system. The results show that the best performance was achieved with a type-1 fuzzy system, albeit featuring a high number of tunable parameters. A type-2 fuzzy system with far fewer parameters achieved performance very close to the best.

Finally, the concept of non-deterministic fuzzy reasoning has been presented and also described how to implement non-deterministic fuzzy sets. As mentioned in this chapter, Garibaldi et al have been investigating the incorporation of variability into decision making in the context of fuzzy expert systems in a medical domain. In those papers, Garibaldi proposed the notion of *non-deterministic fuzzy reasoning* in which variability is introduced into the membership functions of a fuzzy system through the use of random alterations to the parameters of these functions. In the next chapter, this notion is extended and formalised through the introduction of a notion that we will term a *non-stationary fuzzy sets*.

Chapter 5

Non-stationary Fuzzy Sets

5.1 Introduction

According to the use of type-1 fuzzy sets in practice has been limited due to the significant increase in computational complexity involved in their implementation. More recently, type-2 sets have received renewed interest mainly due to the effort of Mendel [1] but also, possibly, by the increases in computational power over recent years. Mendel has established a set of terms to be used when working with type-2 fuzzy sets and, in particular, introduced a concept known as the *footprint of uncertainty* which provides a useful verbal and graphical description of the uncertainty captured by any given type-2 set. The interested reader is particularly referred to [9] for a summary tutorial and/or [1] for a more detailed treatment. Mendel has particularly concentrated on restricted class of general type-2 fuzzy sets know as *interval valued* type-2 fuzzy sets. Interval valued fuzzy sets are characterised by having secondary membership functions which only take the values in $\{0, 1\}$. This restriction greatly simplifies the computational requirements involved in performing inference with type-2 sets and Mendel has provided close formulas for intersection, union and complement, and computational algorithms for type reduction (necessary for type-2 defuzzification).

As mentioned in Chapter 4, It is well accepted that all humans including 'experts', exhibit variation in their decision making. Variation may occur among the decisions of a panel of human experts (inter-expert variability), as well as in the decisions of an individual expert over time (intra-expert variability). Up to now it has been an implicit assumption that expert systems, including fuzzy expert systems, should not exhibit such variation. While type-2 fuzzy sets capture the concept of introducing uncertainty into membership functions by introducing a range of membership values associated with each value of the base variable, they do not capture the notion of variability — as a type-2 fuzzy inference system (FIS) will always produce the same output(s) given the same input(s), although any output(s) will be a type-2 fuzzy set with an implicit representation of uncertainty. Garibaldi et al. [10–14] have been investigating the incorporation of variability into decision making in the context of fuzzy expert systems in a medical domain. In those papers, Garibaldi proposed the notion of *non-deterministic fuzzy reasoning* in which variability is introduced into the membership functions of a fuzzy system through the use of random alterations to the parameters of these functions. In this chapter, this notion is extended and formalised through the introduction of a notion that termed a *non-stationary fuzzy sets*.

5.2 Non-stationary Fuzzy Sets and Systems

As mentioned in the section 5.1, Garibaldi previously proposed the notion of 'non-deterministic fuzzy reasoning' in which variability is introduced into the membership functions of a fuzzy system through the use of random alterations to the parameters of these membership functions. In this section, this notion is extended and formalised through the introduction of a concept termed a *non-stationary fuzzy set*. Informally, a non-stationary fuzzy set is a set (collection) of type-1 fuzzy sets in which there is a connection between (or restriction on) the membership functions of the fuzzy sets. This connection is expressed as a slight variation in the membership function over time. Figure 5.1 shows pictorial representation of repeated instantiations of a non-stationary fuzzy set in which the underlying Gaussian membership function has variation in its standard deviation. The sets were obtained by repeatedly generating (30 times) a Gaussian membership function with the centre of 0.5 and standard deviation that varies by ± 0.05 (i.e. between 0.45 and 0.55). That is, the parameters of the non-stationary set have been chosen in this example such that the extreme parameter values would match those used to generate the upper and lower bounds of type-2 fuzzy set as shown in Figure 2.18. It is apparent from Figure 2.18

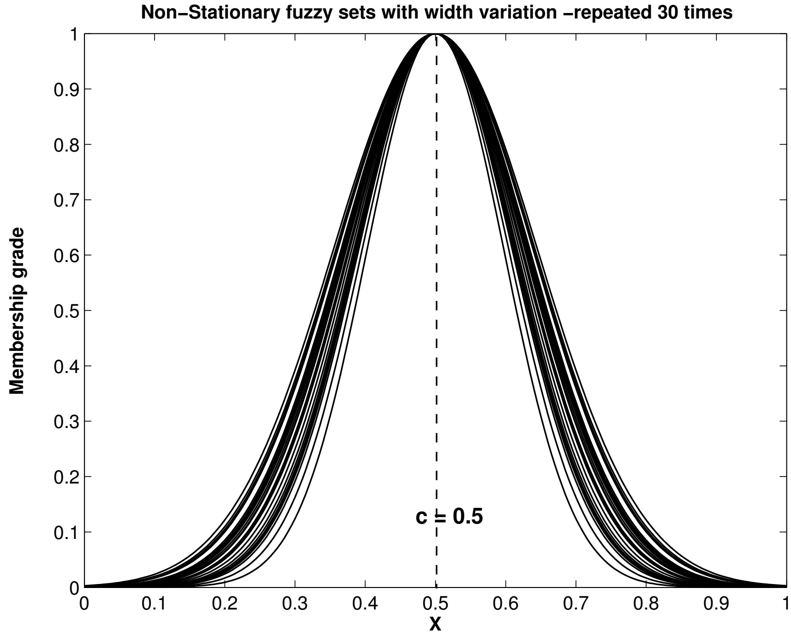


Figure 5.1: Illustration of a Gaussian non-stationary fuzzy set featuring variation in standard deviation and instantiated 30 times.

and 5.1 that the union of all possible instantiations of the non-stationary set is reminiscent of the FOU of the type-2 set. However, it is important to emphasize that non-stationary fuzzy sets are *not* type-2 fuzzy sets. Essentially, type-2 fuzzy sets are 'fuzzy sets with fuzzy membership functions' [34], while non-stationary fuzzy sets are collections of related fuzzy sets. From a formal point of view, non-stationary fuzzy sets are defined in a different way than type-2 fuzzy sets, and have distinct properties (as will be discussed). From a modelling point of view, they model different things: non-stationary fuzzy sets model *temporal variability* in (type-1) membership functions, while type-2 fuzzy sets model uncertain membership functions.

5.3 Non-stationary Fuzzy Sets

Let A denote a fuzzy set of a universe of discourse X characterised by a membership function μ_A. Let T be a set of time points t_i (possibly infinite) and $f : T \rightarrow \Re$ denote a *perturbation function*.

Definition 5.3.1 A non-stationary fuzzy set $\dot{A}$ of the universe of discourse X is characterised by a *non-stationary membership function* $\mu_{\dot{A}} : X \times T \to [0,1]$ that associates with each element x of X and t of T a time-specific variation of $\mu_A(x)$. The non-stationary fuzzy set $\dot{A}$ is denoted by:

$$\dot{A} = \int_{t \in T} \int_{x \in X} \mu_{\dot{A}}(x,t)/x/t. \tag{5.1}$$

However, an additional restriction is imposed on $\mu_{\dot{A}}$. To formulate it in a coherent and precise manner, let consider the first notice that $\mu_A(x)$ can be expressed as $\mu_A(x, p_1, ..., p_m)$, where $p_1, ..., p_m$ denote the parameters of $\mu_A(x)$. Now it is required that:

$$\mu_{\dot{A}}(x,t) = \mu_A(x, p_1(t), ..., p_m(t)). \tag{5.2}$$

where $p_i(t) = p_i + k_i f(t)$ and $i = 1, ..., m$. In this way, each parameter is varied in time by a perturbation function multiplied by a constant.

This definition establishes a relationship between standard and non-stationary fuzzy sets. Specifically, for a given standard fuzzy set A and a set of time points T, a non-stationary fuzzy set $\dot{A}$ is a set of duplicates of A varied over time. A time duplicate of A is termed an *instantiation* and denote it by $\dot{A}_t$, so that $\dot{A}_t(x) = \dot{A}(x,t)$. Thus, at any given moment of time $t \in T$, the non-stationary fuzzy set $\dot{A}$ instantiates the standard fuzzy set $\dot{A}_t$. The standard fuzzy set, A, termed the *underlying* fuzzy set and its associated membership function, $\mu_A(x)$, termed the *underlying* membership function.

Any membership function may be used for the underlying standard fuzzy set. In practice. of course, only a few alternative membership functions are found in standard fuzzy sets, namely: piecewise linear including Left-slope, Triangular, Right-slope, Left-shoulder, Trapeziodal, and Right-shoulder; Gaussian; and Sigmoidal.

Example 5.3.1 As an example, consider a linguistic variable for height. Let the universe of discourse, X, be the interval [1,2], with x interpreted as *height* (in metres). A non-stationary fuzzy set M of X, representing medium height, incorporating variability in the underlying membership function over time, might be represented by a Gaussian membership function where the standard deviation, σ, is a function of time:

$$\dot{M} = \int_{t\in T}\int_{1}^{2} e^{-\frac{(x-c)^2}{\sigma(t)^2}} /x/t. \quad (5.3)$$

Example 5.3.2 As an example, let us formalise these three forms in the context of Gaussian membership functions. A standard Gaussian membership function can be written in the parameterised notation as:

$$\mu(x,c,\sigma,\varepsilon) = e^{-\frac{(x-c)^2}{\sigma^2}} + \varepsilon. \quad (5.4)$$

(Of course, normally ε is zero.) Now, the three forms of non-stationarity described above can be expressed by , respectively:

$$\mu(x,c(t),\sigma,\varepsilon) = e^{-\frac{(x-c(t))^2}{\sigma^2}} + \varepsilon. \quad (5.5)$$

$$\mu(x,c,\sigma(t),\varepsilon) = e^{-\frac{(x-c)^2}{\sigma(t)^2}} + \varepsilon. \quad (5.6)$$

$$\mu(x,c,\sigma,\varepsilon(t)) = e^{-\frac{(x-c)^2}{\sigma^2}} + \varepsilon(t). \quad (5.7)$$

Note that for simplicity, (t) will be omitted from any parameter that does not vary over time. Naturally, there is no reason why these three different kinds of variation could not be combined together. In this case:

$$\mu(x,c(t),\sigma(t),\varepsilon(t)) = e^{-\frac{(x-c(t))^2}{\sigma(t)^2}} + \varepsilon(t). \quad (5.8)$$

but, for simplicity, at present such combined non-stationarity will not be considered.

5.4 Perturbation Functions

The original intention behind non-stationary fuzzy sets was to capture the notion of minor variations of a membership function corresponding to subtle differences in opinion over time. Additionally, the intention was that a non-stationary fuzzy set remains close to the underlying fuzzy set over time; that is , there is no permanent 'drift' or alteration of the membership function which is characteristic of learning processes. Thus, the term perturbation function has been deliberately chosen to imply that parameter changes induced by the function are 'small' or, more precisely, that parameter changes induce 'small' and temporary alterations in $\mu_A(x)$.

Note that there is an interesting relationship between small variations over time that are proposed here for non-stationary fuzzy sets and long-term changes in membership functions that are seen in adaptive (or 'learning') fuzzy sets. Such relationships are outside the scope of this thesis, and might require further research.

There are many ways in which an opinion may vary over time. However, three main alternative forms of non-stationarity which might be more useful in practice can be formalised as follows:

- Variation in location — the membership function is shifted, as a whole, left or right by small amounts along the universe of discourse, relative to the underlying membership function.

- Variation in width — the width of the membership function is increased or decreased by small amounts, relative to the underlying membership function.

- Noise variation — the membership function is shifted upward or downward by a small amount of 'white noise', relative to the underlying membership function.

The next issue to be addressed is the form of the perturbation function. In general, it would be appear that any function of time might be used as a perturbation function, within the formal restriction that the membership function remains in bounds (i.e. $\mu_A(x,t) \in [0,1]$). In theory, a perturbation function could be a true random function. Given that any measurement of time is arbitrary and relative, the actual set of functions that might be useful in practice is more restrictive. Any units might be used for time, t, but the most natural would be to express time in seconds, in the absence of any good reason not to. Again, given that any physical notion of time is relative, any arbitrary point in time might be chosen as zero. A few families of perturbation functions that might be used in practice are:

- periodic, e.g.:

$$f(t) = sin(\omega t) \tag{5.9}$$

- pseudo-random, e.g.:

$$f(t) = \frac{s(t+1) - 2^{47}}{2^{47}}, \tag{5.10}$$

  where $s(0)$ is the initial seed in $[0, 2^{48}]$ and

  $s(t+1) = (25,214,903,917 s(t) + 11) mod 2^{48}$.

- a differential time-series, such as the Mackey-Glass equation:

$$\frac{df(t)}{dt} = \frac{0.2 f * (t-\tau)}{1 + f^{10}(t-\tau)} - 0.1 f(t), \tag{5.11}$$

  where τ is a constant.

In last section, *c(t)*, σ*(t)*, and ε(t) can all be generated by using the following:

$$c(t) = c + kf(t) \tag{5.12}$$

$$\sigma(t) = \sigma + kf(t) \tag{5.13}$$

$$\varepsilon(t) = kf(t) \tag{5.14}$$

where c and σ are the centre and width of the initially type-1 fuzzy set, respectively, k is a constant value, and f(t) is the what will be termed the *perturbation function*. By perturbation function, a function (of time), that will generate small changes in the underlying membership function. In theory, this could be a true random function — i.e. the membership function parameter could be a true random variable: hence the terminology of *non-stationary* fuzzy sets.

5.5 Footprint of Variation (FOV)

As mentioned in section 2.8, the term 'footprint of uncertainty' (FOU) was introduced by Mendel to provide 'a very convenient verbal description of the entire domain of support for all the secondary grades of a type-2 membership function' [1], where uncertainty in primary memberships of a type-2 fuzzy set consists of a bounded region that called the footprint of uncertainty (FOU), e.g. pictorial in Figure 2.18. Each of the secondary membership functions of interval type-2 fuzzy sets has only one secondary grade that equal to 1.

A similar term, the '*footprint of variation*' (FOV), has been introduced as a similar verbal description of the area covering the range from the minimum to the maximum fuzzy sets which comprise the non-stationary fuzzy sets as shown in Figure 5.1. For non-stationary fuzzy sets which are generated by Uniformly distributed and Sinusoidal perturbation functions (producing random values within $[-1,1]$), the maximum area of FOV will be equivalent to the FOU of interval type-2 fuzzy sets with the same amount of variation. Normally distributed perturbation functions generate random values within $[-\infty,\infty]$, and so an FOV defined as the union of all primary memberships would fill the entire universe of discourse. This kind of FOV will need further investigation, and will be left for the future works.

5.6 Non-stationary Fuzzy Inference Systems

Although this thesis is not focused on a complete description of the fuzzy inference process, in order to clarify the difference between type-1, type-2, and non-stationary fuzzy sets. An FIS consists of four main inter-connected components: the *rules*, the *fuzzifier*, the *inference engine*, and an *output processor*. Type-1 FISs use only type-1 fuzzy sets, whereas an FIS which uses at least one type-2 fuzzy set is called a type-2 FIS. Figure 5.2 shows the mechanisms of a type-2 FIS (adapted from [1]).

Figure 5.3 shows the mechanism of the inferencing process in an FIS consisting of such non-stationary fuzzy sets. An FIS is naturally termed as a 'non-stationary FIS'. It should be emphasized that a non-stationary FIS is simply a repetition of a type-1 FIS with slightly different instantiations of the membership functions over time. Thus, im-

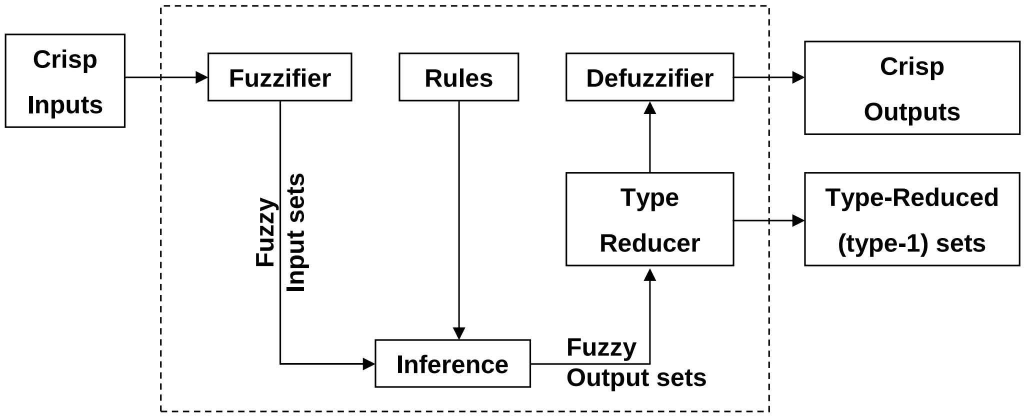


Figure 5.2: Mechanisms of a type-2 FLS (adapted from [1]).

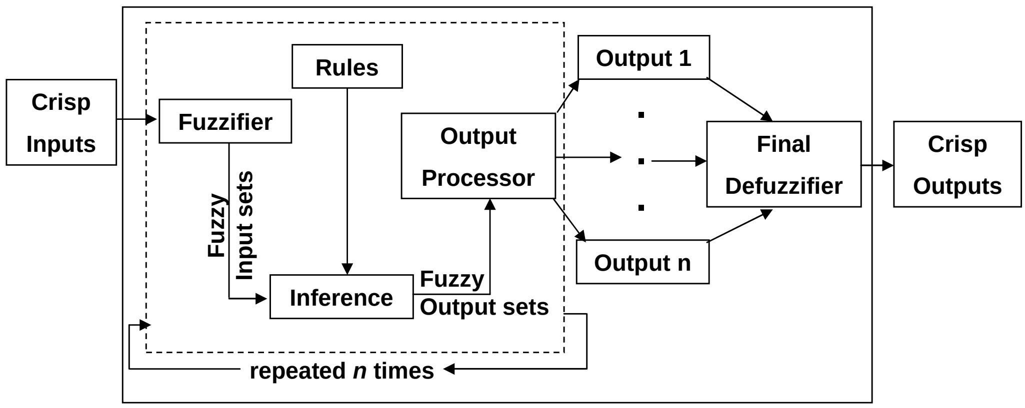


Figure 5.3: Proposed mechanisms of a non-stationary FLS.

plementing a non-stationary FIS is simply a matter of iterating over the required number of instantiations while perturbing the membership functions. Neither the form of non-stationarity (variation in location, variation in width, or noise variation) nor the form of perturbation function (periodic, random, chaotic, etc.) has any effect on the inference process. Hence, an inference with non-stationary fuzzy sets is clearly different from the type-2 inference, and does not suffer the difficulties of type-2 inference (particularly the inference using general type-2 fuzzy sets). A preliminary analysis of the relationship between non-stationary fuzzy sets and type-2 fuzzy sets will be explored in Chapter 6 and 7 in form of case studies.

5.7 Operations on and Properties of Non-stationary Fuzzy Sets

In this section, the operators of *union*, *intersection*, and *complement* of *non-stationary fuzzy sets* are introduced.

At first, let recall the familiar properties of type-1 fuzzy sets. Suppose, there are two fuzzy sets, A and B, that are characterised by membership functions $\mu_A(x)$ and $\mu_B(x)$:

$$A = \int_{x\in X} \mu_A(x)/x. \tag{5.15}$$

and

$$B = \int_{x\in X} \mu_B(x)/x. \tag{5.16}$$

Recall that:

$$\mu_{A\cup B}(x) = \int_{x\in X} \mu_A(x) \cup \mu_B(x)/x, \tag{5.17}$$

$$\mu_{A\cap B}(x) = \int_{x\in X} \mu_A(x) \cap \mu_B(x)/x, \tag{5.18}$$

$$\mu_{\bar{A}}(x) = \int_{x\in X} 1 - \mu_A(x)/x, \tag{5.19}$$

$$\mu_{\bar{B}}(x) = \int_{x\in X} 1 - \mu_B(x)/x, \tag{5.20}$$

The membership functions of the union and intersection of A and B, and the complement of A are, of course:

$$\mu_{A\cup B}(x) = \mu_A(x) \oplus \mu_B(x),\ \forall x \in X, \tag{5.21}$$

where $\oplus$ is a t-conorm,

Because A and B are type-1 fuzzy sets, their membership grades $\mu_A(x)$ and $\mu_B(x)$ are crisp number and, at each x, $\mu_{A\cup B}(x)$, $\mu_{A\cap B}(x)$, $\mu_{\bar{A}}(x)$ and $\mu_{\bar{B}}(x)$ are also crisp numbers.

$$\mu_{A\cap B}(x) = \mu_A(x) \otimes \mu_B(x),\ \forall x \in X, \tag{5.22}$$

where $\otimes$ is a t-norm, and

$$\mu_{\bar{A}}(x) = \overline{\mu_A(x)},\ \forall x \in X, \tag{5.23}$$

where $^{-}$ is a generic complement.

Using the maximum t-conorm, minimum t-norm and the standard complement, the previous equations become:

$$\mu_{A\cup B}(x) = max[\mu_A(x), \mu_B(x)],\ \forall x \in X, \tag{5.24}$$

$$\mu_{A\cap B}(x) = min[\mu_A(x), \mu_B(x)],\ \forall x \in X, \tag{5.25}$$

$$\mu_{\bar{A}}(x) = 1 - \mu_A(x),\ \forall x \in X. \tag{5.26}$$

Now, let $T = \{t_1, ..., t_n\}$ be a set of time points t_i, and let $\dot{A}$ and $\dot{B}$ be non-stationary fuzzy sets of a universe of discourse X. Thus

$$\dot{A} = \int_{t\in T}\int_{x\in X} \mu_{\dot{A}}(x,t)/x/t$$

and

$$\dot{B} = \int_{t\in T}\int_{x\in X} \mu_{\dot{B}}(x,t)/x/t.$$

Definition 5.7.1 (*Non-stationary Union Operator*): The union of $\dot{A}$ and $\dot{B}$, is a non-stationary fuzzy sets $\dot{A}\cup\dot{B}$ such that:

$$\dot{A}\cup\dot{B} = \int_{t\in T}\int_{x\in X} \mu_{\dot{A}\cup\dot{B}}(x,t)/x/t,$$

where

$$\mu_{\dot{A}\cup\dot{B}}(x,t) = \mu_{\dot{A}}(x,t) \oplus \mu_{\dot{B}}(x,t),\ \forall (x,t) \in X \times T.$$

Thus, using the maximum t-conorm, this becomes:

$$\mu_{\dot{A}\cup\dot{B}}(x,t) = max[\mu_{\dot{A}}(x,t), \mu_{\dot{B}}(x,t)],\ \forall (x,t) \in X \times T.$$

Definition 5.7.2 (*Non-stationary Intersection Operator*): The intersection of $\dot{A}$ and $\dot{B}$ is a non-stationary fuzzy set $\dot{A} \cap \dot{B}$ such that:

$$\dot{A} \cap \dot{B} = \int_{t \in T} \int_{x \in X} \mu_{\dot{A} \cap \dot{B}}(x,t)/x/t,$$

where

$$\mu_{\dot{A} \cap \dot{B}}(x,t) = \mu_{\dot{A}}(x,t) \otimes \mu_{\dot{B}}(x,t), \ \forall (x,t) \in X \times T.$$

where, using the minimum t-norm, this becomes:

$$\mu_{\dot{A} \cap \dot{B}}(x,t) = min[\mu_{\dot{A}}(x,t), \mu_{\dot{B}}(x,t)], \ \forall (x,t) \in X \times T.$$

Definition 5.7.3 (*Non-stationary Complement Operator*): The complement of $\dot{A}$ is a non-stationary fuzzy set, $\overline{\dot{A}}$, such that:

$$\overline{\dot{A}} = \int_{t \in T} \int_{x \in X} \mu_{\overline{\dot{A}}}(x,t)/x/t,$$

where

$$\mu_{\overline{\dot{A}}}(x,t) = \overline{\mu_{\dot{A}}(x,t)}, \ \forall (x,t) \in X \times T.$$

which, of course, using the standard complement, this becomes:

$$\mu_{\overline{\dot{A}}}(x,t) = 1 - \mu_{\dot{A}}(x,t), \ \forall (x,t) \in X \times T.$$

5.8 Proof of Properties of Non-stationary Fuzzy Sets

This section is dedicated to the proof of some of the fundamental properties of non-stationary fuzzy set operators defined earlier. This proofs are derived directly from Zadeh's proofs for standard type-1 fuzzy sets; these are included for completeness. Table 5.1 summarises the set theoretic laws that are satisfied by non-stationary fuzzy sets.

First, Let consider non-stationary fuzzy sets $\dot{A}$, $\dot{B}$, and $\dot{C}$:

Table 5.1: Summary of some set theoretic laws satisfied by non-stationary fuzzy sets.

Set theoretic laws	Maximum t-conorm	Minimum t-norm
Involution		
$\bar{\bar{\dot{A}}} = \dot{A}$	Yes	Yes
Commutativity		
$\dot{A} \cup \dot{B} = \dot{B} \cup \dot{A}$	Yes	yes
$\dot{A} \cap \dot{B} = \dot{B} \cap \dot{A}$	Yes	yes
Associativity		
$(\dot{A} \cup \dot{B}) \cup \dot{C} = \dot{A} \cup (\dot{B} \cup \dot{C})$	Yes	Yes
$(\dot{A} \cap \dot{B}) \cap \dot{C} = \dot{A} \cap (\dot{B} \cap \dot{C})$	Yes	Yes
Idempotence		
$\dot{A} \cup \dot{A} = \dot{A}$	Yes	Yes
$\dot{A} \cap \dot{A} = \dot{A}$	Yes	Yes
Distributivity		
$\dot{A} \cap (\dot{B} \cup \dot{C}) = (\dot{A} \cap \dot{B}) \cup (\dot{A} \cap \dot{C})$	Yes	Yes
$\dot{A} \cup (\dot{B} \cap \dot{C}) = (\dot{A} \cup \dot{B}) \cap (\dot{A} \cup \dot{C})$	Yes	Yes

$$\dot{A} = \int_{t \in T} \int_{x \in X} \mu_{\dot{A}}(x,t)/x/t,$$

$$\dot{B} = \int_{t \in T} \int_{x \in X} \mu_{\dot{B}}(x,t)/x/t,$$

and

$$\dot{C} = \int_{t \in T} \int_{x \in X} \mu_{\dot{C}}(x,t)/x/t.$$

Note that, for the sake of brevity in the formula below, whenever a non-stationary union, intersection or complement operator from the definitions given in Section 5.7 is used in this section, then $\forall(x,t) \in X \times T$ will be omitted.

5.8.1 Involution

Let us consider the complement of A, $\overline{\dot{A}}$:

$$\overline{\dot{A}} = \int_{t \in T} \int_{x \in X} \mu_{\overline{\dot{A}}}(x,t)/x/t.$$

By the definition of the complement operation for non-stationary fuzzy sets (Definition 5.7.3), we have:

$$\mu_{\overline{\dot{A}}}(x,t) = 1 - \mu_{\dot{A}}(x,t), \ \forall(x,t) \in X \times T. \tag{5.27}$$

Thus, the complement of $\overline{\dot{A}}$ can be expressed as:

$$\overline{\overline{\dot{A}}} = \int_{t \in T} \int_{x \in X} \mu_{\overline{\overline{\dot{A}}}}(x,t)/x/t,$$

where

$$\mu_{\overline{\overline{\dot{A}}}}(x,t) = 1 - \mu_{\overline{\dot{A}}}(x,t).$$

By replacing $\mu_{\overline{\dot{A}}}(x,t)$ with Equation 5.27, we obtain:

$$\mu_{\overline{\overline{\dot{A}}}}(x,t) = 1 - (1 - \mu_{\dot{A}}(x,t)).$$

It follows that:

$$\mu_{\overline{\overline{\dot{A}}}}(x,t) = 0 + \mu_{\dot{A}}(x,t).$$

Finally:

$$\mu_{\overline{\overline{\dot{A}}}}(x,t) = \mu_{\dot{A}}(x,t).$$

So we can claim that $\overline{\overline{\dot{A}}} = \dot{A}$. We then can conclude that non-stationary fuzzy set has an involution property.

5.8.2 Commutativity

Proof of commutativity for non-stationary fuzzy sets can be shown as following;
Let define a non-stationary fuzzy set, $\dot{A}$ and $\dot{B}$

$$\dot{A} = \int_{t \in T} \int_{x \in X} \mu_{\dot{A}}(x,t)/x/t,$$

and

$$\dot{B} = \int_{t \in T} \int_{x \in X} \mu_{\dot{B}}(x,t)/x/t,$$

Union

By the definition of the union operation for non-stationary fuzzy sets, we have:

$$\mu_{\dot{A} \cup \dot{B}}(x,t) = \mu_{\dot{A}}(x,t) \oplus \mu_{\dot{B}}(x,t).$$

As the t-conorm operator is commutative, we know that:

$$\mu_{\dot{A}}(x,t) \oplus \mu_{\dot{B}}(x,t) = \mu_{\dot{B}}(x,t) \oplus \mu_{\dot{A}}(x,t).$$

Again, by definition:

$$\mu_{\dot{B} \cup \dot{A}}(x,t) = \mu_{\dot{B}}(x,t) \oplus \mu_{\dot{A}}(x,t).$$

thus:

$$\mu_{\dot{A}\cup\dot{B}}(x,t) = \mu_{\dot{B}\cup\dot{A}}(x,t).$$

So we can claim that $\dot{A}\cup\dot{B} = \dot{B}\cup\dot{A}$.

We then can conclude that non-stationary fuzzy sets have an commutativity property for union operator.

Intersection

By the definition of the intersection operation for non-stationary fuzzy sets, we have:

$$\mu_{\dot{A}\cap\dot{B}}(x,t) = \mu_{\dot{A}}(x,t)\otimes\mu_{\dot{B}}(x,t).$$

As t-norm operator is commutative, we know that:

$$\mu_{\dot{A}}(x,t)\otimes\mu_{\dot{B}}(x,t) = \mu_{\dot{B}}(x,t)\otimes\mu_{\dot{A}}(x,t).$$

Again, by definition:

$$\mu_{\dot{B}\cap\dot{A}}(x,t) = \mu_{\dot{B}}(x,t)\otimes\mu_{\dot{A}}(x,t).$$

thus:

$$\mu_{\dot{A}\cap\dot{B}}(x,t) = \mu_{\dot{B}\cap\dot{A}}(x,t).$$

So we can claim that $\dot{A}\cap\dot{B} = \dot{B}\cap\dot{A}$.

We then can conclude that non-stationary fuzzy sets have an commutativity property for intersection operator.

5.8.3 Associativity

Proof of Associativity for non-stationary fuzzy sets can be shown as following;

Let define a non-stationary fuzzy set, $\dot{A}$, $\dot{B}$, and $\dot{C}$.

$$\dot{A} = \int_{t\in T}\int_{x\in X}\mu_{\dot{A}}(x,t)/x/t,$$

$$\dot{B} = \int_{t\in T}\int_{x\in X}\mu_{\dot{B}}(x,t)/x/t,$$

and

$$\dot{C} = \int_{t\in T}\int_{x\in X} \mu_{\dot{C}}(x,t)/x/t.$$

Union

By the definition of the union operation for non-stationary fuzzy sets, we have:

$$\mu_{\dot{A}\cup(\dot{B}\cup\dot{C})}(x,t) = \mu_{\dot{A}}(x,t) \oplus (\mu_{\dot{B}}(x,t) \oplus \mu_{\dot{C}}(x,t)).$$

As the t-conorm operator is associative, we know that:

$$\mu_{\dot{A}}(x,t) \oplus (\mu_{\dot{B}}(x,t) \oplus \mu_{\dot{C}}(x,t)) = (\mu_{\dot{A}}(x,t) \oplus \mu_{\dot{B}}(x,t)) \oplus \mu_{\dot{C}}(x,t).$$

Again, by definition:

$$\mu_{(\dot{A}\cup\dot{B})\cup\dot{C}}(x,t) = (\mu_{\dot{A}}(x,t) \oplus \mu_{\dot{B}}(x,t)) \oplus \mu_{\dot{C}}(x,t).$$

And so:

$$\mu_{\dot{A}\cup(\dot{B}\cup\dot{C}}(x,t)) = \mu_{(\dot{A}\cup\dot{B})\cup\dot{C}}(x,t).$$

So we can claim that $(\dot{A}\cup\dot{B})\cup\dot{C} = \dot{A}\cup(\dot{B}\cup\dot{C})$.
We then can conclude that non-stationary fuzzy sets have an associativity property for union operator.

Intersection

By the definition of the intersection operation for non-stationary fuzzy sets, we have:

$$\mu_{\dot{A}\cap(\dot{B}\cap\dot{C})}(x,t) = \mu_{\dot{A}}(x,t) \otimes (\mu_{\dot{B}}(x,t) \otimes \mu_{\dot{C}}(x,t)).$$

As the t-norm operator is associative, we know that:

$$\mu_{\dot{A}}(x,t) \otimes (\mu_{\dot{B}}(x,t) \otimes \mu_{\dot{C}}(x,t)) = (\mu_{\dot{A}}(x,t) \otimes \mu_{\dot{B}}(x,t)) \otimes \mu_{\dot{C}}(x,t).$$

Again, by definition:

$$\mu_{(\dot{A}\cap\dot{B})\cap\dot{C}}(x,t) = (\mu_{\dot{A}}(x,t) \otimes \mu_{\dot{B}}(x,t)) \otimes \mu_{\dot{C}}(x,t).$$

And so:

$$\mu_{\dot{A}\cap(\dot{B}\cap\dot{C}}(x,t)) = \mu_{(\dot{A}\cap\dot{B})\cap\dot{C}}(x,t).$$

So we can claim that $(\dot{A} \cap \dot{B}) \cap \dot{C} = \dot{A} \cap (\dot{B} \cap \dot{C})$.

We then can conclude that non-stationary fuzzy sets have an associativity property for intersection operator.

5.8.4 Distributivity

Proof of distributivity for non-stationary fuzzy sets can be shown as following;

Let define a non-stationary fuzzy set, $\dot{A}$, $\dot{B}$, and $\dot{C}$.

$$\dot{A} = \int_{t\in T}\int_{x\in X} \mu_{\dot{A}}(x,t)/x/t,$$

$$\dot{B} = \int_{t\in T}\int_{x\in X} \mu_{\dot{B}}(x,t)/x/t,$$

and

$$\dot{C} = \int_{t\in T}\int_{x\in X} \mu_{\dot{C}}(x,t)/x/t.$$

Union

By the definition of the union operation for non-stationary fuzzy sets, we have:

$$\mu_{\dot{A}\cap(\dot{B}\cup\dot{C})}(x,t) = \mu_{\dot{A}}(x,t) \otimes (\mu_{\dot{B}}(x,t) \oplus \mu_{\dot{C}}(x,t)).$$

As the t-conorm operator is distributive, we know that:

$$\mu_{\dot{A}}(x,t) \otimes (\mu_{\dot{B}}(x,t) \oplus \mu_{\dot{C}}(x,t)) = (\mu_{\dot{A}}(x,t) \otimes \mu_{\dot{B}}(x,t)) \oplus (\mu_{\dot{A}}(x,t) \otimes \mu_{\dot{C}}(x,t)).$$

Again, by definition:

$$\mu_{(\dot{A}\cap\dot{B})\cup(\dot{A}\cap\dot{C})}(x,t) = ((\mu_{\dot{A}}(x,t) \otimes \mu_{\dot{B}}(x,t)) \oplus (\mu_{\dot{A}}(x,t)\mu_{\dot{C}}(x,t)).$$

And so:

$$\mu_{\dot{A}\cap(\dot{B}\cup\dot{C})}(x,t) = \mu_{(\dot{A}\cap\dot{B})\cup(\dot{A}\cap\dot{C})}(x,t).$$

So we can claim that $\dot{A} \cap (\dot{B} \cup \dot{C}) = (\dot{A} \cap \dot{B}) \cup (\dot{A} \cap \dot{C})$.

We then can conclude that non-stationary fuzzy sets have an distributivity property for union operator.

Intersection

By the definition of the intersection operation for non-stationary fuzzy sets, we have:

$$\mu_{\dot{A}\cup(\dot{B}\cap\dot{C})}(x,t) = \mu_{\dot{A}}(x,t) \oplus (\mu_{\dot{B}}(x,t) \otimes \mu_{\dot{C}}(x,t)).$$

As the t-norm operator is distributive, we know that:

$$\mu_{\dot{A}}(x,t) \oplus (\mu_{\dot{B}}(x,t) \otimes \mu_{\dot{C}}(x,t)) = (\mu_{\dot{A}}(x,t) \oplus \mu_{\dot{B}}(x,t)) \otimes (\mu_{\dot{A}}(x,t) \oplus \mu_{\dot{C}}(x,t)).$$

Again, by definition:

$$\mu_{(\dot{A}\cup\dot{B})\cap(\dot{A}\cup\dot{C})}(x,t) = ((\mu_{\dot{A}}(x,t) \oplus \mu_{\dot{B}}(x,t)) \otimes (\mu_{\dot{A}}(x,t) \oplus \mu_{\dot{C}}(x,t)).$$

And so:

$$\mu_{\dot{A}\cup(\dot{B}\cap\dot{C})}(x,t) = \mu_{(\dot{A}\cup\dot{B})\cap(\dot{A}\cup\dot{C})}(x,t).$$

So we can claim that $\dot{A} \cup (\dot{B} \cap \dot{C}) = (\dot{A} \cup \dot{B}) \cap (\dot{A} \cup \dot{C})$.

We then can conclude that non-stationary fuzzy sets have an distributivity property for intersection operator.

5.8.5 Idempotence

It is well known that by restricting the t-conorm and t-norm operators to be idempotent, the only possible operators are *max* and *min* respectively.

Union

By the definition of the union operation for non-stationary fuzzy sets, we have:

$$\mu_{\dot{A}\cup\dot{A}}(x,t) = max(\mu_{\dot{A}}(x,t), \mu_{\dot{A}}(x,t)).$$

As the max operator is idempotent, we know that:

$$max(\mu_{\dot{A}}(x,t), \mu_{\dot{A}}(x,t)) = \mu_{\dot{A}}(x,t).$$

and so

$$\mu_{\dot{A}\cup\dot{A}}(x,t) = \mu_{\dot{A}}(x,t).$$

so we can claim that $\dot{A}\cup\dot{A} = \dot{A}$. We then can conclude that non-stationary fuzzy set has an an idempotence property for union operator.

Intersection

By the definition of the intersection operation of non-stationary fuzzy sets, we have:

$$\mu_{\dot{A}\cap\dot{A}}(x,t) = min(\mu_{\dot{A}}(x,t), \mu_{\dot{A}}(x,t)).$$

As the min operator is idempotent, we know that:

$$min(\mu_{\dot{A}}(x,t), \mu_{\dot{A}}(x,t)) = \mu_{\dot{A}}(x,t).$$

and so

$$\mu_{\dot{A}\cap\dot{A}}(x,t) = \mu_{\dot{A}}(x,t).$$

so we can claim that $\dot{A} \cap \dot{A} = \dot{A}$. We then can conclude that non-stationary fuzzy set has an an idempotance property for intersection operator.

5.9 Summary

In this chapter, a new concept termed *non-stationary fuzzy set* is defined. These have been created with the specific intention of modelling the variation (over time) of opinion, and then formalise the novel concept that previously proposed by Garibaldi [14] to model the variation in expert opinion. While apparently similar to type-2 fuzzy sets in some regards, non-stationary fuzzy sets possess some important distinguishing features. A non-stationary fuzzy set is, effectively, a collection of type-1 fuzzy sets in which there is an explicit, defined, relationship between the fuzzy set. Specifically, each of the instantiations (type-1 fuzzy set) is derived by a perturbation of (making a small change to) a single underlying membership function. While each instantiation is somewhat reminiscent of a embedded type-1 set of a type-2 fuzzy set, there is *not* a direct correspondence between these two concepts. It is also possible to view a standard type-1 fuzzy set, either as a single instantiation or as repeated instantiations of the underlying set with no perturbation. Again, a non-stationary fuzzy set does not have secondary membership function. Hence, there is no direct equivalent to the embedded type-2 sets of type-2 fuzzy set. Similarly, there are no secondary grades. While it is true that distributions of membership grades across 'vertical slices' are still not, formally, the same as secondary membership functions. The inference process is quite different. The crucial point is that, at any instant of time, a non-stationary fuzzy set is (instantiates) type-1 fuzzy set. Hence the non-stationary inference is just a repeated type-1 inference (albeit with slightly different type-1 sets at each time instant). In contrast, type-2 inference involves passing type-2 fuzzy sets through the process, resulting in type-2 output sets that require type reduction prior to defuzzification.

Some possible functions that can be used as a perturbation function are provided. The term , *footprint of variation* (FOV), is proposed to represent the area covering the range from the minimum to the maximum fuzzy sets which comprise the non-stationary fuzzy set. Operations on non-stationary fuzzy sets, i.e. *union*, *intersection*, *complement* are also introduced in this chapter. Finally, proof of some properties of non-stationary fuzzy sets, i.e. *involution*, *commutativity*, *associativity*, *idempotence*, and *distributivity* are presented.

In the next chapters, the investigation onto performances of non-stationary fuzzy sets comparing with interval type-2 fuzzy sets is proposed. The secondary membership functions, the output's interval, and the relationship between primary membership functions and output uncertainties are considered for these investigations.

Chapter 6

Relationships between Interval Type-2 and Non-stationary Fuzzy Sets

6.1 Introduction

In this chapter, The relationships between interval type-2 and non-stationary fuzzy sets is investigated, in terms of secondary membership functions. Two case studies were described which were carried out in order to illustrate the use of non-stationary fuzzy sets and to explore the relationship between the performance of non-stationary fuzzy inference systems (FISs) and interval type-2 FISs. All fuzzy inference systems were constructed to perform a fuzzy equivalent of the classical XOR operation, where Table 6.1 shows the classical XOR operation, where *Input1* and *Input2* are input variables and *xor(Input1,Input2)* is the output variable for those inputs.

In this study, FISs were constructed to predict the truth value of the XOR output where both input variables can take any value in the range of [0,1]. All FISs consist of two input variables which are *Input1* and *Input2*, one output variable which is *Output*, and four rules. Each variable consist of 2 membership functions, corresponding to meaning of the terms *Low* and *High*. The following four rules were used for all FISs. These rules are constructed based on the standard XOR problem.

1. IF *Input1* is *Low* AND *Input2* is *Low*
   THEN *Output* is *Low*

Table 6.1: XOR Truth Table.

	Input1	Input2	xor(Input1,Input2)
Case 1	0	0	0
Case 2	0	1	1
Case 3	1	0	1
Case 4	1	1	0

2. IF *Input1* is *Low* AND *Input2* is *High*
   THEN *Output* is *High*
3. IF *Input1* is *High* AND *Input2* is *Low*
   THEN *Output* is *High*
4. IF *Input1* is *High* AND *Input1* is *High*
   THEN *Output* is *Low*

The four cases of input values used throughout these studies are shown in Table 6.2.

Table 6.2: Input Values for Fuzzy Systems.

Case	Input1	Input2	Output
Case 1	0.25	0.25	?
Case 2	0.25	0.75	?
Case 3	0.75	0.25	?
Case 4	0.75	0.75	?

6.2 Case Study 1: Gaussian Membership Functions

6.2.1 The Non-stationary FISs:

In the first case study, non-stationary FISs utilising Gaussians of the form:

$$\mu_A(x,c,\sigma) = e^{-\frac{(x-c)^2}{\sigma^2}} \tag{6.1}$$

as the underlying membership functions were investigated.

Note that the ε parameter has now been dropped, as noise variation was not considered in this study. Two forms of non-stationary were implemented:

- Variation in location: only the centre , c, of the Gaussian (Equation 6.1) was varied over time, yielding non-stationary membership functions of the form:

$$\mu_{\dot{A}}(x,c+k_1 f(t),\sigma) = e^{-\frac{(x-(c+k_1 f(t)))^2}{\sigma^2}} \tag{6.2}$$

- Variation in width: only the standard deviation, σ, of the Gaussian (Equation 6.1) was varied over time, yielding non-stationary membership functions of the form:

$$\mu_{\dot{A}}(x,c,\sigma+k_2 f(t)) = e^{-\frac{(x-c)^2}{(\sigma+k_2 f(t))^2}} \tag{6.3}$$

The *Low* membership functions had centre 0.1, the *High* membership functions had centre 0.9, and all had a standard deviation of 0.25. The underlying Gaussian membership functions are shown in Figure 6.1.

Three different perturbation functions were used , as follows:

- a uniformly distributed pseudo-random function, e.g.:

$$f(t) \quad = \quad \frac{s(t+1)-2^{47}}{2^{47}}, \tag{6.4}$$

where $s(0)$ is the initial seed in $[0,2^{48}]$ and

$s(t+1) = (25,214,903,917s(t)+11)mod2^{48}$,

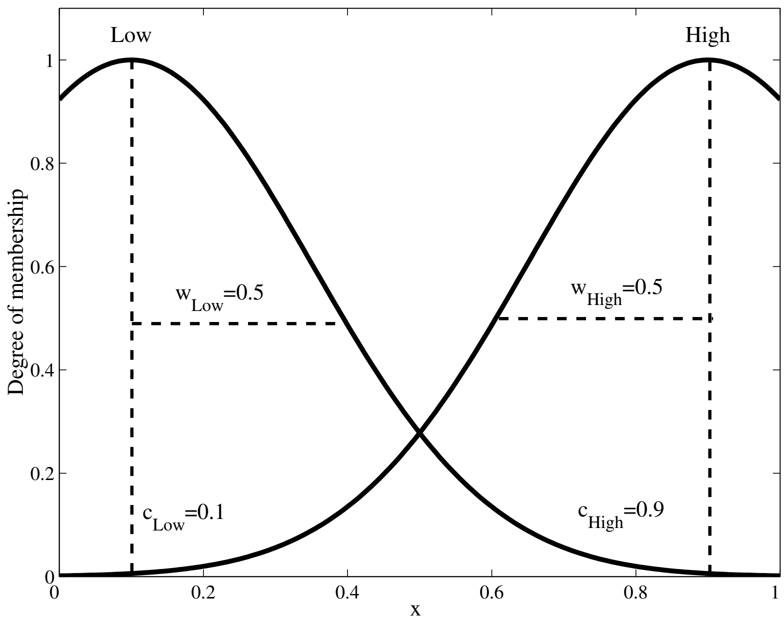


Figure 6.1: The underlying Gaussian membership functions for the terms *Low* and *High* as used in the case study 1.

- a random function with Gaussian distribution (Matlab *randn* function),
- Sine function (where $\omega = 127$), e.g.:

$$f(t) \quad = \quad sin(\omega t). \tag{6.5}$$

The first and third functions above returned values in the range $[-1, 1]$, while the second (the Matlab *randn* function) returned real values sampled from a normalised Gaussian distribution with mean zero and standard deviation one.

In all cases $k_1 = k_2 = 0.05$.

Four different non-stationary FISs for each of these three perturbation functions were designed (12 non-stationary FISs in total). These were distinguished by the number of instantiations (time points) used to construct the non-stationary fuzzy sets, as follows:

- 30 instantiations
- 100 instantiations
- 1,000 instantiations

- 10,000 instantiations

These non-stationary systems are denoted by NS1-#$-&*. '#' is either 'G' to denote Gaussian underlying membership functions or 'T' for Triangular ones (as used in the second case study, below). '$' is either 'L' to denote variation in location or 'W' for variation in width. '&' is either 'U' to denote uniform perturbation function, 'G' for Gaussian or 'S' for sinusoidal. Finally, '*' denotes the number of instantiations. For example, NS1-GL-G100 denotes a non-stationary system that utilises Gaussian underlying membership functions featuring variation in location, with a Gaussian perturbation function, instantiated 100 times.

6.2.2 The Type-2 FISs:

Two interval type-2 FISs were designed, featuring the same inputs and outputs, and the same four rules. The footprints of uncertainty of the membership functions were created by deviating the parameters of Equation 6.1 as follows. For the variation the location, the lower and upper bounds of the FOU were generated by:

$$\mu_{\tilde{A}}(x,c,\sigma) = e^{-\frac{(x-(c\pm k_1))^2}{\sigma^2}}. \tag{6.6}$$

For variation in width, the lower and upper bounds of the FOU were generated by:

$$\mu_{\tilde{A}}(x,c,\sigma) = e^{-\frac{(x-c)^2}{(\sigma\pm k_2)^2}}. \tag{6.7}$$

In both cases, again, $k_1 = k_2 = 0.05$.

Note that these formula were obtained by from Equation 6.2 and 6.3 by setting $f(t) = \pm 1$. This was purposefully chosen to establish a form of correspondence between the non-stationary and the interval type-2 FISs. The interval type-2 systems are denotes by IT2-#$. As before, '#' is either 'G' to denote Gaussian underlying membership functions or 'T' for Triangular ones (as used in the second case study, below), and '$' is either 'L' to denote variation in location or 'W' for variation width.

6.2.3 The Inference Process and Results:

For each FIS constructed, inference was performed and the results were obtained as follows. For the type-2 systems, the four input cases shown in Table 6.2 were presented to the systems. Inference was performed using the rules given to obtain the type-2 output sets. In each case, the usual Karnik-Mendel type reduction was used to obtain the *lower* and *upper* bound of the centre of gravity of the output. The *mean* of the output was taken as the average of the *lower* and *upper* bound.

For the non-stationary systems, each system was instantiated the specified number of times to obtain the set of non-stationary output fuzzy sets. In each case, defuzzification was applied to obtain the standard centre of gravity, g, As a result, a set of centres of gravity, denoted G, was obtained (it is obvious that $|G| = |T|$).

For the case of uniform or sinusoidal perturbation functions (for which the range was bounded to $[-1, +1]$), the minimum of G was taken as the lower bound, the maximum of G as the upper bound, and the arithmetic mean was taken as the *mean*.

For the case of Gaussian perturbation functions (for which the range was unbounded), the lower and upper bounds were derived by $m_G \pm \sigma_G$, where m_G is the mean of G and σ_G is the standard deviation of G.

The results obtained for variation in the centre of the underlying Gaussian membership functions are given in Table 6.3, while the results obtained for variation in the standard deviation of the underlying Gaussian membership functions are given in Table 6.4.

The non-stationary FISs featuring variation in centre were selected for further investigation. Specifically, for those featuring Gaussian perturbation functions, the distribution of membership values obtained over time was examined for specific values of x. As such a distribution is , in fact, obtained by taking a '*vertical slice*' through a non-stationary fuzzy set, it is in some way analogous to the secondary membership function of a type-2 fuzzy set. The distributions obtained in the y-axis of the membership values, at $x = 0.15$ for the *Low* term and $x = 0.85$ for the *High* term are shown in Figure 6.2. The non-stationary fuzzy sets were each generated with 10,000 instantiations. Similar distributions obtained at $x = 0.20$ and $x = 0.30$ for the *Low* term and those obtained at $x = 0.70$ and $x = 0.80$ for the *High* term are shown in Figure 6.3- 6.6 for centre variation and Figure 6.7- 6.10, respectively.

Table 6.3: Lower, Mean and Upper Bounds for Centre Variation.

	Case 1 (0.25,0.25)			*Case 2 (0.25,0.75)*			*Case 3 (0.75,0.25)*			*Case 4 (0.75,0.75)*		
FIS	*Lower*	*Mean*	*Upper*	*Lower*	*Mean*	*Upper*	*Lower*	*Mean*	*Upper*	*Lower*	*Mean*	*Upper*
it2-GL	0.3687	0.3956	0.4224	0.5780	0.6050	0.6320	0.5780	0.6050	0.6320	0.3687	0.3956	0.4224
NS1-GL-U30	0.3829	0.3904	0.3979	0.5940	0.6023	0.6106	0.5960	0.6054	0.6148	0.3849	0.3937	0.4025
NS1-GL-U100	0.3842	0.3928	0.4014	0.5962	0.6046	0.6130	0.5953	0.6054	0.6155	0.3840	0.3943	0.4046
NS1-GL-U1000	0.3849	0.3936	0.4023	0.5965	0.6053	0.6141	0.5969	0.6060	0.6151	0.3851	0.3945	0.4039
NS1-GL-U10000	0.3850	0.3941	0.4032	0.5968	0.6057	0.6146	0.5968	0.6057	0.6146	0.3849	0.3941	0.4033
NS1-GL-G30	0.3853	0.3937	0.4024	0.5972	0.6067	0.6142	0.5969	0.6068	0.6139	0.3851	0.3938	0.4019
NS1-GL-G100	0.3850	0.3940	0.4028	0.5968	0.6069	0.6144	0.5969	0.6073	0.6144	0.3851	0.3944	0.4027
NS1-GL-G1000	0.3850	0.3937	0.4031	0.5967	0.6066	0.6146	0.5967	0.6063	0.6146	0.3850	0.3934	0.4031
NS1-GL-G10000	0.3850	0.3935	0.4031	0.5967	0.6064	0.6146	0.5967	0.6064	0.6146	0.3850	0.3935	0.4031
NS1-GL-S30	0.3853	0.3937	0.4024	0.5972	0.6067	0.6142	0.5969	0.6068	0.6139	0.3851	0.3938	0.4019
NS1-GL-S100	0.3856	0.3931	0.4032	0.5972	0.6060	0.6150	0.5973	0.6056	0.6146	0.3856	0.3927	0.4031
NS1-GL-S1000	0.3854	0.3934	0.4033	0.5969	0.6063	0.6150	0.5969	0.6066	0.6150	0.3854	0.3937	0.4033
NS1-GL-S10000	0.3850	0.3935	0.4031	0.5967	0.6064	0.6146	0.5967	0.6064	0.6146	0.3850	0.3935	0.4031

Table 6.4: Lower, Mean and Upper Bounds for Width Variation.

	Input 1 (0.25,0.25)			*Input 2 (0.25,0.75)*			*Input 3 (0.75,0.25)*			*Input 4 (0.75,0.75)*		
FLS	*Lower*	*Mean*	*Upper*	*Lower*	*Mean*	*Upper*	*Lower*	*Mean*	*Upper*	*Lower*	*Mean*	*Upper*
it2-GW	0.3836	0.3921	0.4007	0.5993	0.6079	0.6164	0.5993	0.6079	0.6164	0.3836	0.3921	0.4007
NS1-GW-U30	0.3716	0.3909	0.4102	0.5898	0.6091	0.6284	0.5986	0.6155	0.6324	0.3676	0.3845	0.4014
NS1-GW-U100	0.3697	0.3911	0.4125	0.5875	0.6089	0.6303	0.5932	0.6106	0.6280	0.3720	0.3894	0.4068
NS1-GW-U1000	0.3731	0.3922	0.4113	0.5887	0.6078	0.6303	0.5911	0.6098	0.6275	0.3725	0.3907	0.4089
NS1-GW-U10000	0.3730	0.3917	0.4104	0.5896	0.6083	0.6270	0.5898	0.6085	0.6272	0.3728	0.3915	0.4102
NS1-GW-G30	0.3742	0.3932	0.4088	0.5912	0.6068	0.6258	0.5920	0.6066	0.6264	0.3736	0.3934	0.4080
NS1-GW-G100	0.3734	0.3937	0.4093	0.5907	0.6063	0.6266	0.5908	0.6056	0.6264	0.3736	0.3944	0.4092
NS1-GW-G1000	0.3733	0.3931	0.4097	0.5903	0.6069	0.6267	0.5903	0.6075	0.6267	0.3733	0.3924	0.4097
NS1-GW-G10000	0.3732	0.3926	0.4098	0.5902	0.6074	0.6268	0.5902	0.6075	0.6268	0.3732	0.3925	0.4098
NS1-GW-S30	0.3736	0.3934	0.4080	0.5920	0.6066	0.6264	0.5912	0.6068	0.6258	0.3742	0.3932	0.4088
NS1-GW-S100	0.3736	0.3944	0.4092	0.5908	0.6056	0.6264	0.5907	0.6063	0.6266	0.3734	0.3937	0.4093
NS1-GW-S1000	0.3733	0.3925	0.4097	0.5903	0.6075	0.6267	0.5903	0.6069	0.6267	0.3733	0.3931	0.4097
NS1-GW-S10000	0.3732	0.3925	0.4098	0.5902	0.6075	0.6268	0.5902	0.6074	0.6268	0.3732	0.3926	0.4098

Finally, the distribution of the centres of gravity, g, obtained for the NS1-GL-G10000 non-stationary FIS was also examined. The distributions obtained for the four input cases are shown in Figure 6.11- 6.14, respectively.

6.3 Case Study 2: Triangular Membership Functions

6.3.1 The Non-stationary FISs:

In the second case study, the same experiments as described above were repeated. However, this time Triangular functions were used as the underlying membership functions. That is, the underlying membership functions were of the form:

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

where a denotes the left-hand base-point of the triangle, b denotes the right-hand base-point and c denotes the centre of the triangle. Only one form of non-stationary was implemented in this case study, variation in location, yielding non-stationary membership functions of the form:

$$\mu_{\dot{A}}(x,a(t),b(t),c(t)) = \begin{cases} 0, & x - kf(t) \leq a \\ \frac{x-(a+kf(t))}{c-a}, & a < x - kf(t) \leq c \\ \frac{(b+kf(t))-x}{b-c}, & c < x - kf(t) < b \\ 0, & x - kf(t) \geq b \end{cases}$$

so that the whole triangle was shifted left or right over time by the amount $kf(t)$. *Low* membership functions all had $a = 0.10$, $b = 0.50$, and $c = 0.30$, and *High* membership functions all had $a = 0.50$, $b = 0.90$, and $c = 0.70$. The underlying type-1 membership function are shown in Figure 6.15.

The same three different perturbation functions were used:

- uniformly distributed,

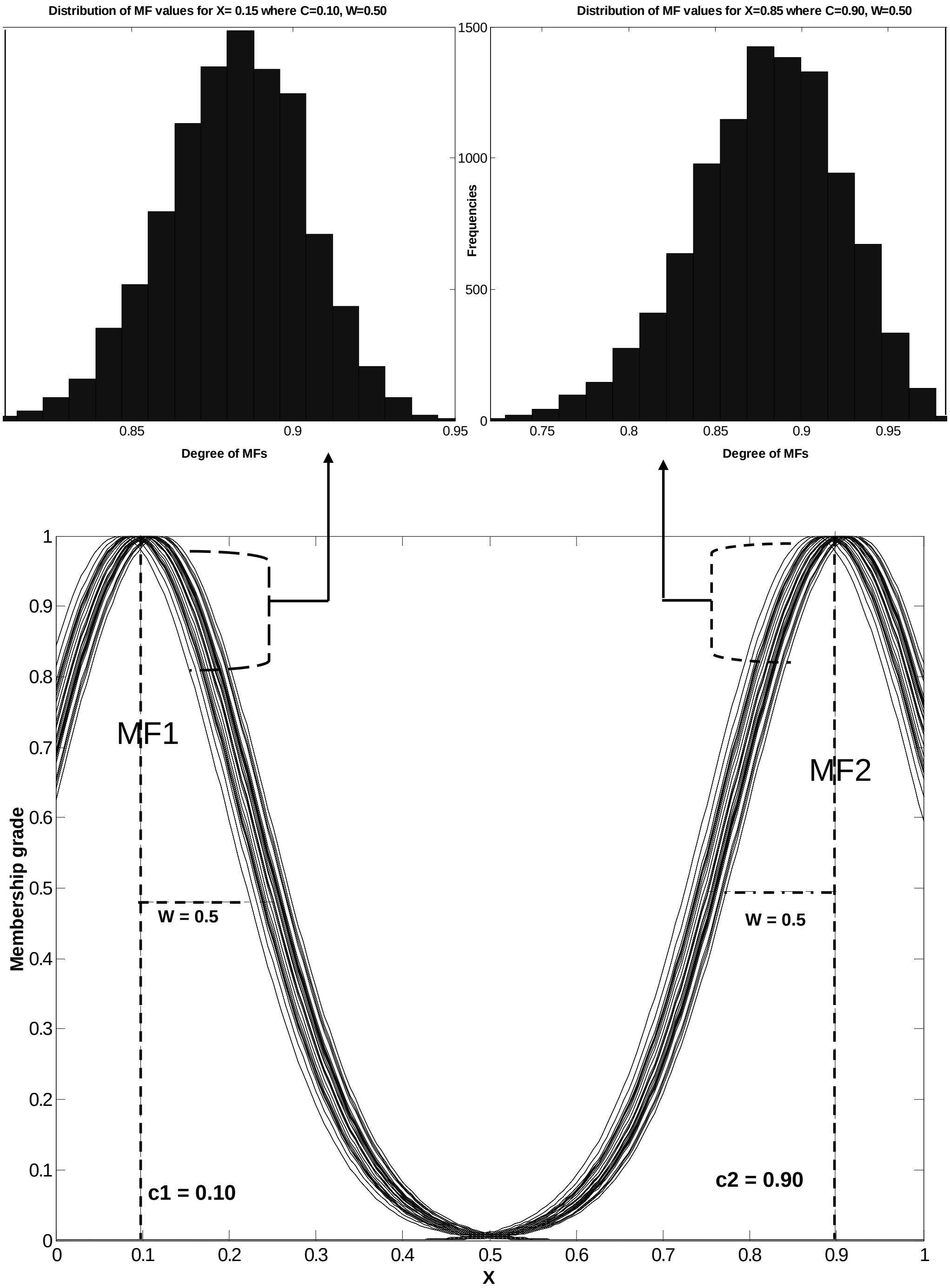


Figure 6.2: Distributions of membership grades over time at $x = 0.15$ and $x = 0.85$, for the non-stationary FIS NS1-GL-G10000.

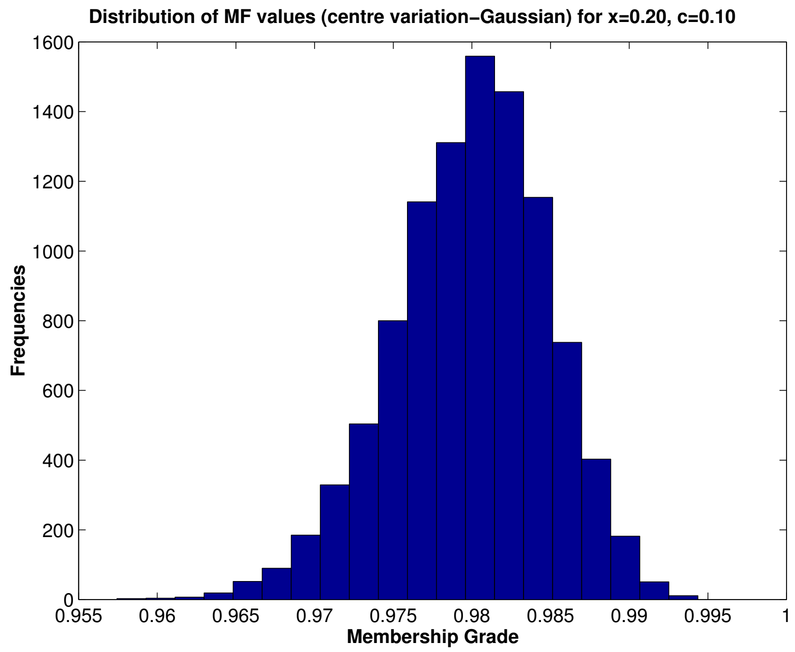


Figure 6.3: Distribution of MF values (centre variation-Gaussian MF) for x=0.20 and c=0.10.

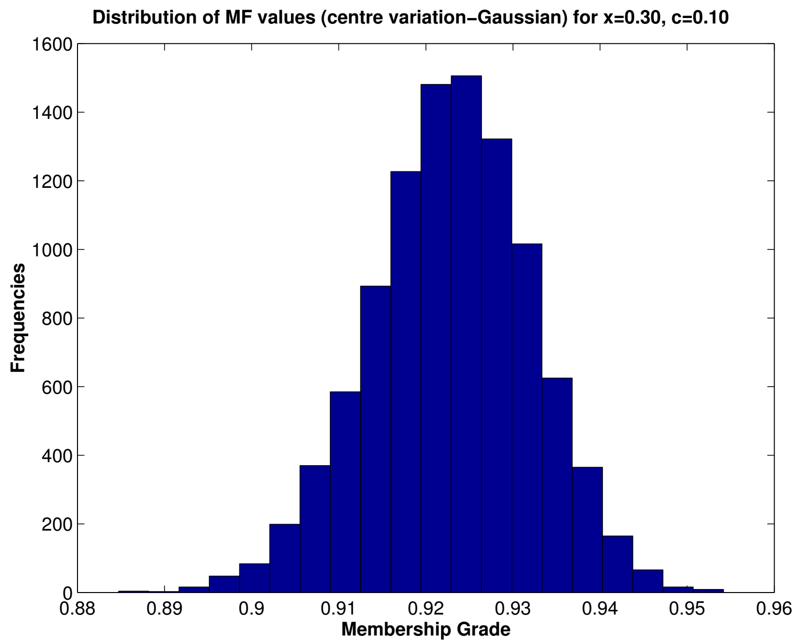


Figure 6.4: Distribution of MF values (centre variation-Gaussian MF) for x=0.30 and c=0.10.

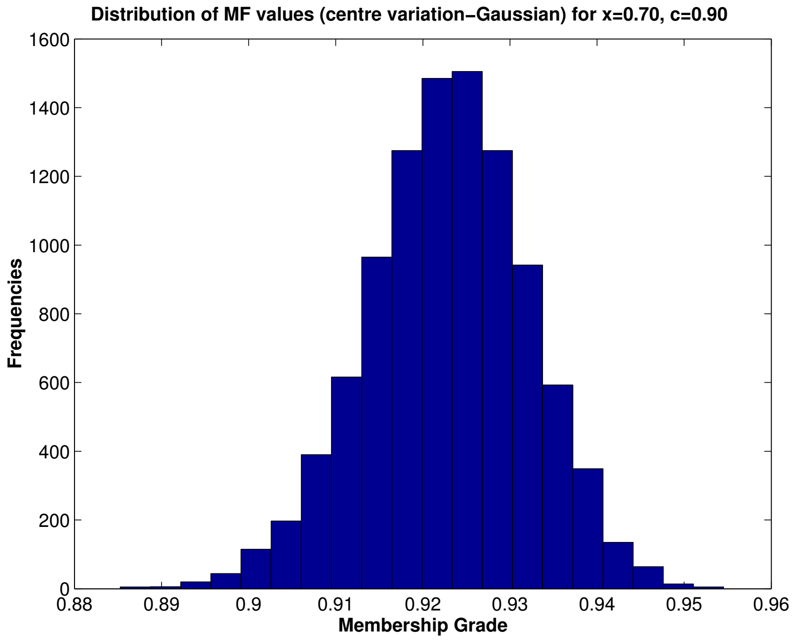


Figure 6.5: Distribution of MF values (centre variation-Gaussian MF) for x=0.70 and c=0.90.

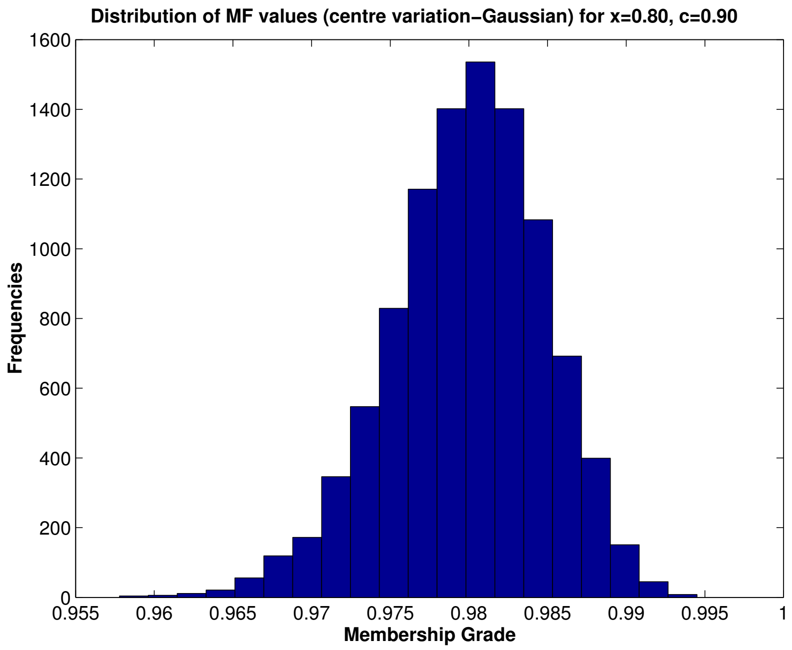


Figure 6.6: Distribution of MF values (centre variation-Gaussian MF) for x=0.80 and c=0.90.

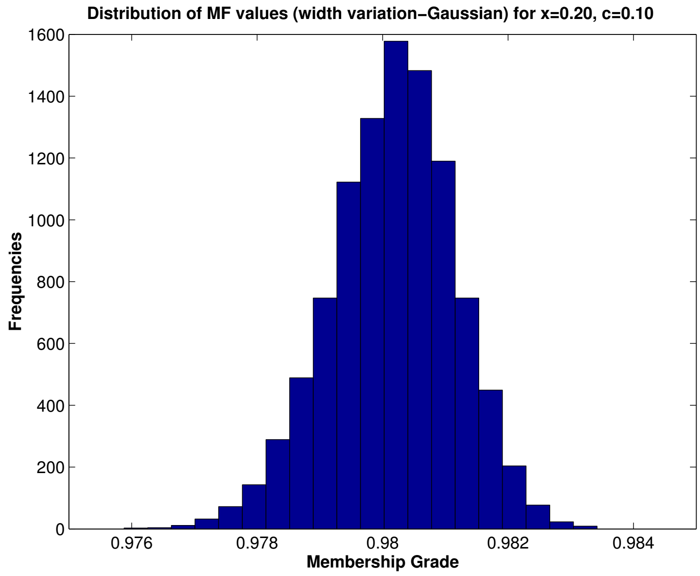


Figure 6.7: Distribution of MF values (width variation-Gaussian MF) for x=0.20 and c=0.10.

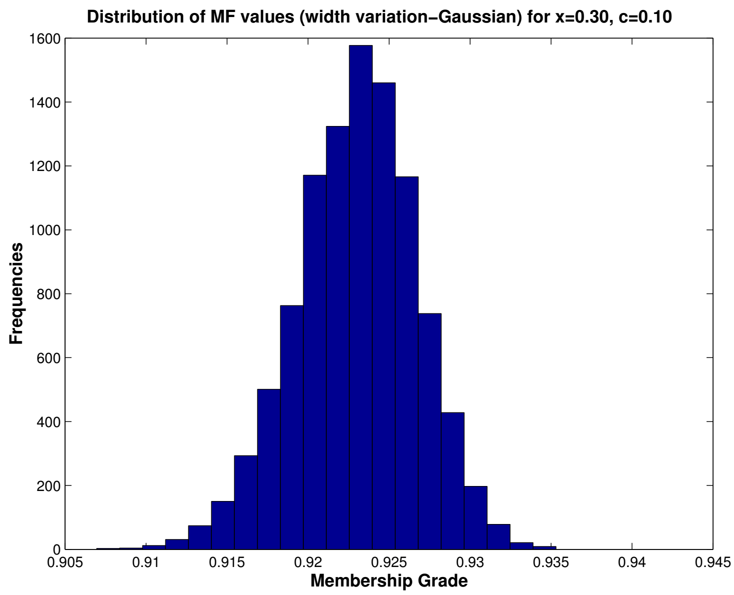


Figure 6.8: Distribution of MF values (width variation-Gaussian MF) for x=0.30 and c=0.10.

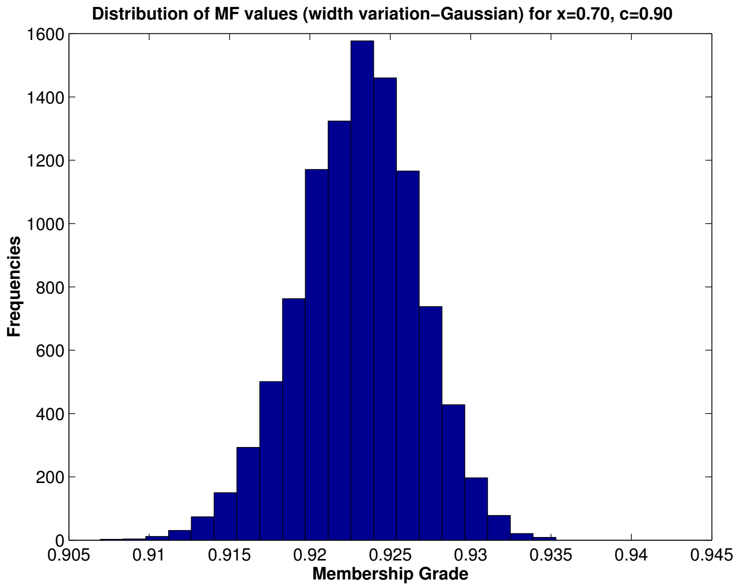


Figure 6.9: Distribution of MF values (width variation-Gaussian MF) for x=0.70 and c=0.90.

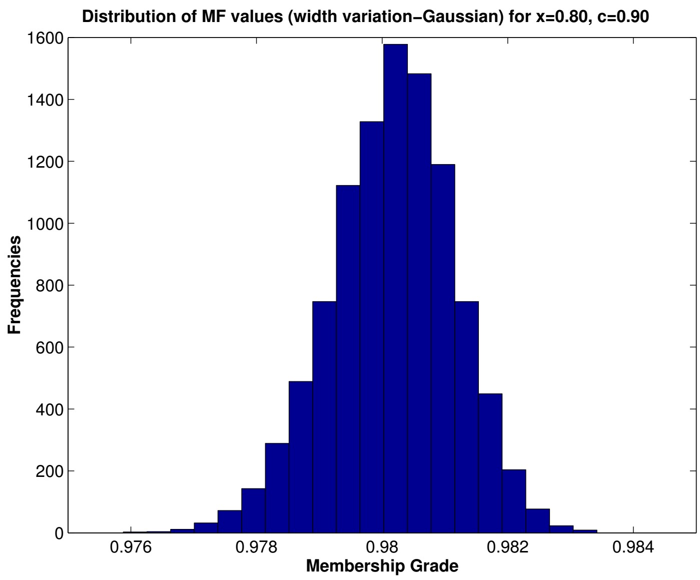


Figure 6.10: Distribution of MF values (width variation-Gaussian MF) for x=0.80 and c=0.90.

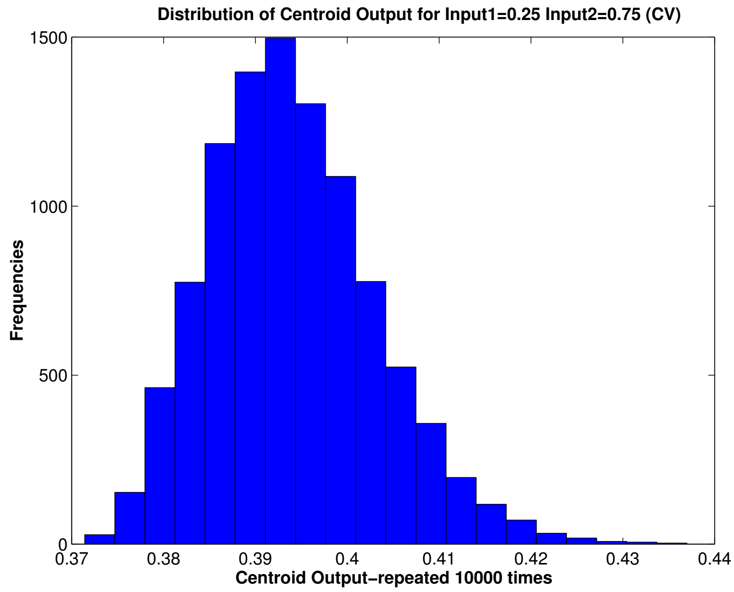


Figure 6.11: Distribution of centroid output (nsFLS-centre variation) for Input1= Input2 = 0.25.

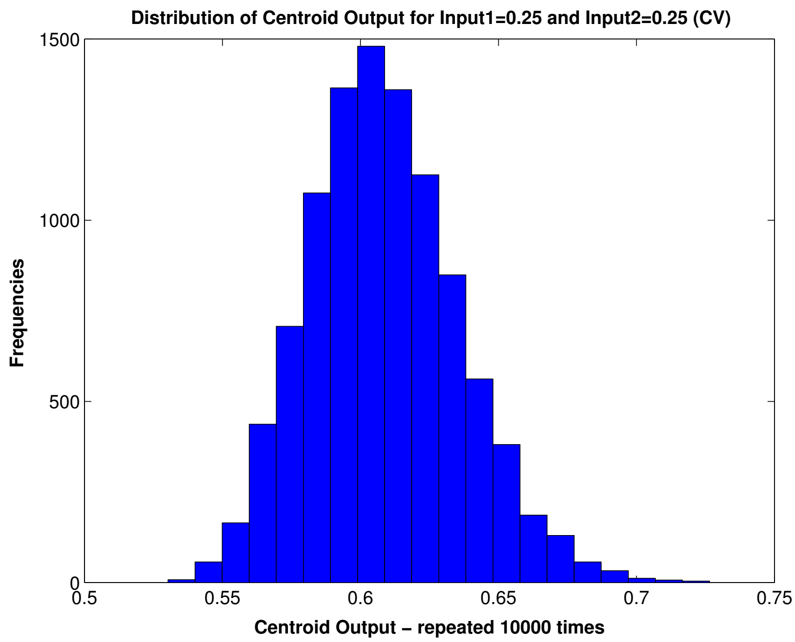


Figure 6.12: Distribution of centroid output (nsFLS-centre variation) for Input1= 0.25 Input2 = 0.75.

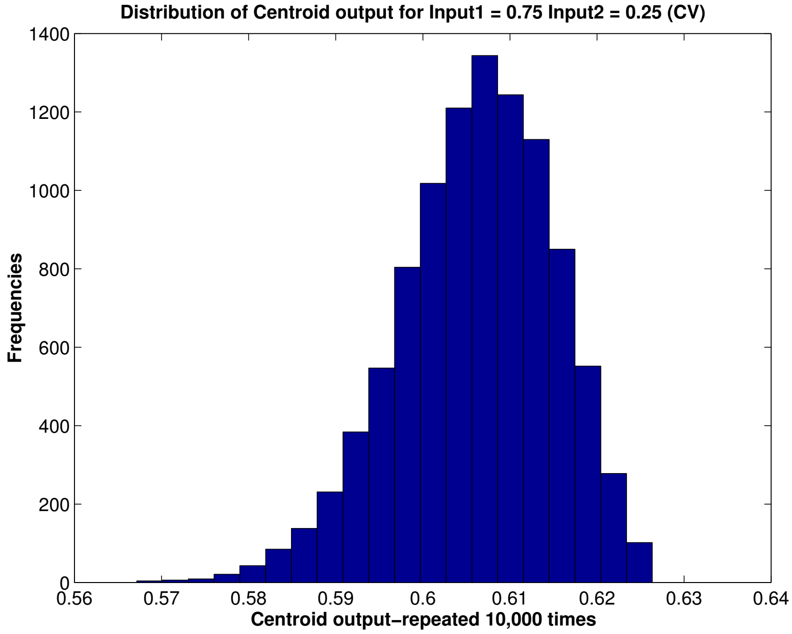


Figure 6.13: Distribution of centroid output (nsFLS-centre variation) for Input1= 0.75 Input2 = 0.25.

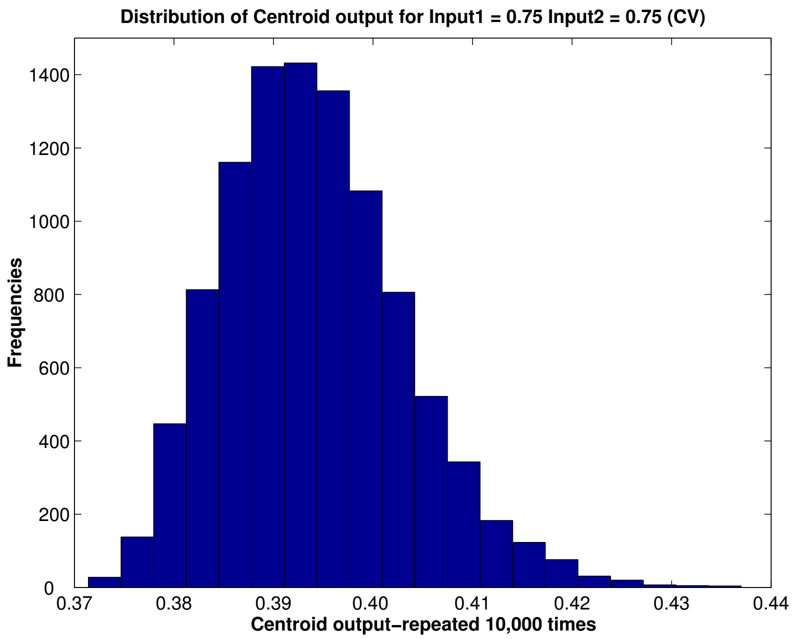


Figure 6.14: Distribution of centroid output (nsFLS-centre variation) for Input1= 0.75 Input2 = 0.75.

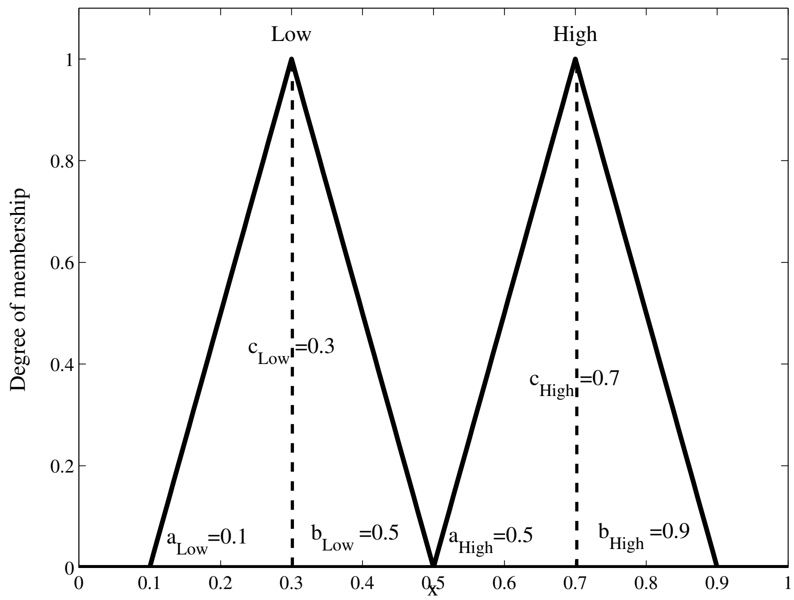


Figure 6.15: The underlying type-1 Triangular membership functions for the terms *Low* and *High* as used in the case study 2.

- Gaussian distributed,
- Sinusoidal.

In all case $k = 0.05$. Once again, four different non-stationary systems for each perturbation function were created, with:

- 30 instantiations
- 100 instantiations
- 1,000 instantiations
- 10,000 instantiations

6.3.2 The Type-2 FIS:

An interval type-2 system was generated using the same principle as described in 6.2.2. So, the FOU was bounded by:

$$\mu_{\tilde{A}}(x,a,b,c) = \begin{cases} 0, & x-k \leq a \\ \frac{x-(a+k)}{c-a}, & a < x-k \leq c \\ \frac{(b+k)-x}{b-c}, & c < x-k < b \\ 0, & x-k \geq b \end{cases}$$

where a denotes the left-hand base-point of the triangle, b denotes the right-hand base-point and c denotes the centre of the triangle.

$$\mu_{\tilde{A}}(x,a,b,c) = \begin{cases} 0, & x+k \leq a \\ \frac{x-(a-k)}{c-a}, & a < x+k \leq c \\ \frac{(b-k)-x}{b-c}, & c < x+k < b \\ 0, & x+k \geq b \end{cases}$$

with again $k = 0.05$.

6.3.3 The Inference Process and Results:

Inference was performed using the 13 systems (12 non-stationary FISs and the type-2 FIS), and the methodology as described above was using to derive the outputs of each FIS. The results obtained are shown in Table 6.5.

Table 6.5: Lower, Mean and Upper Bounds for Variation in Location of Underlying Triangular Membership Functions.

	Input 1 (0.25,0.25)			*Input 2 (0.25,0.75)*			*Input 3 (0.75,0.25)*			*Input 4 (0.75,0.75)*		
FIS	*Lower*	*Mean*	*Upper*	*Lower*	*Mean*	*Upper*	*Lower*	*Mean*	*Upper*	*Lower*	*Mean*	*Upper*
it2-TL	0.2500	0.3000	0.3500	0.6500	0.7000	0.7500	0.6500	0.7000	0.7500	0.2500	0.3000	0.3500
NS1-TL-G30	0.2506	0.2979	0.3453	0.6506	0.6980	0.7453	0.6506	0.6890	0.7453	0.2506	0.2980	0.3454
NS1-TL-G100	0.2505	0.2980	0.3481	0.6505	0.6979	0.7476	0.6505	0.6979	0.7476	0.2508	0.2980	0.3481
NS1-TL-G1000	0.2507	0.3001	0.3494	0.6501	0.6992	0.7483	0.6501	0.6992	0.7483	0.2511	0.3001	0.3490
NS1-TL-G10000	0.2505	0.3000	0.3494	0.6501	0.6996	0.7491	0.6501	0.6996	0.7491	0.2510	0.3004	0.3498
NS1-TL-S30	0.2500	0.2990	0.3500	0.6500	0.6990	0.7500	0.6501	0.6996	0.7496	0.2501	0.2996	0.3496
NS1-TL-S100	0.2500	0.3000	0.3500	0.6500	0.7000	0.7500	0.6500	0.6997	0.7500	0.2500	0.2997	0.3500
NS1-TL-S1000	0.2500	0.2999	0.3500	0.6500	0.6999	0.7500	0.6500	0.7000	0.7500	0.2500	0.3001	0.3500
NS1-TL-S10000	0.2500	0.3000	0.3500	0.6500	0.7000	0.7500	0.6500	0.7000	0.7500	0.2500	0.3000	0.3500
NS1-TL-U30	0.2509	0.3012	0.3466	0.6509	0.7012	0.7466	0.6505	0.7013	0.7497	0.2505	0.3013	0.3497
NS1-TL-U100	0.2503	0.3019	0.3485	0.6503	0.7019	0.7485	0.6504	0.7013	0.7497	0.2504	0.3013	0.3497
NS1-TL-U1000	0.2502	0.3011	0.3498	0.6502	0.7011	0.7498	0.6502	0.7008	0.7498	0.2502	0.3008	0.3498
NS1-TL-U10000	0.2500	0.2997	0.3500	0.6500	0.6998	0.7500	0.6500	0.6997	0.7500	0.2500	0.2997	0.3500

Again, for non-stationary fuzzy sets featuring Gaussian perturbation functions, the distribution of membership function of membership values obtained over time was examined for specific values of x. The distributions obtained at $x = 0.40$ and $x = 0.80$ for the *Low* and *High* terms, respectively, are shown in Figure 6.16.

Further distributions obtained for a variety of values of x for the two terms are shown in Figure 6.17- 6.22.

Finally, the distribution of the centres of gravity, g, obtained for the NS1-TL-G10000 non-stationary FIS was also examined. The distributions obtained for the four input cases are shown in Figure 6.23- 6.26, respectively.

6.4 Discussion

Examination of the results of both case studies in Tables 6.3, 6.4, and 6.5 highlights a number of interesting observations. Firstly, the results for input cases 1 and 2 are very similar to the results for input cases 4 and 3, respectively. This applies to both variation in location and variation in width, and is entirely as expected due to the symmetrical nature of the *Low* and *High* terms and rules of the XOR problem.

For variation in centre of Gaussian underlying membership functions (the first case study), the output interval of the type-2 FIS for input case 1 $(0.25, 0.25)$ is $[0.3687, 0.4224]$, with mean 0.3956. For the non-stationary FIS featuring, for example, Gaussian distribution of the perturbation function instantiated 10,000 times, the corresponding interval is $[0.3850, 0.4032]$, with mean 0.3941 — i.e. the mean is similar, but the interval is *narrower*. This finding is, in fact, repeated across all the results obtained for variation in centre. Similar results are observed for variation in standard deviation, except that in this case the intervals obtained are slightly *wider* than the corresponding type-2 intervals (the lower bounds are lower and the upper bounds are higher). Indeed, the same observation holds for the case of variation in location of Triangular membership functions. These preliminary finding need to be investigated further before any definitive conclusions can be reached as to whether this finding is independent of the underlying membership function.

These results emphasise the fact that non-stationary systems are different from type-2 fuzzy systems. However, it is also clear that, in some sense, a non-stationary system mim-

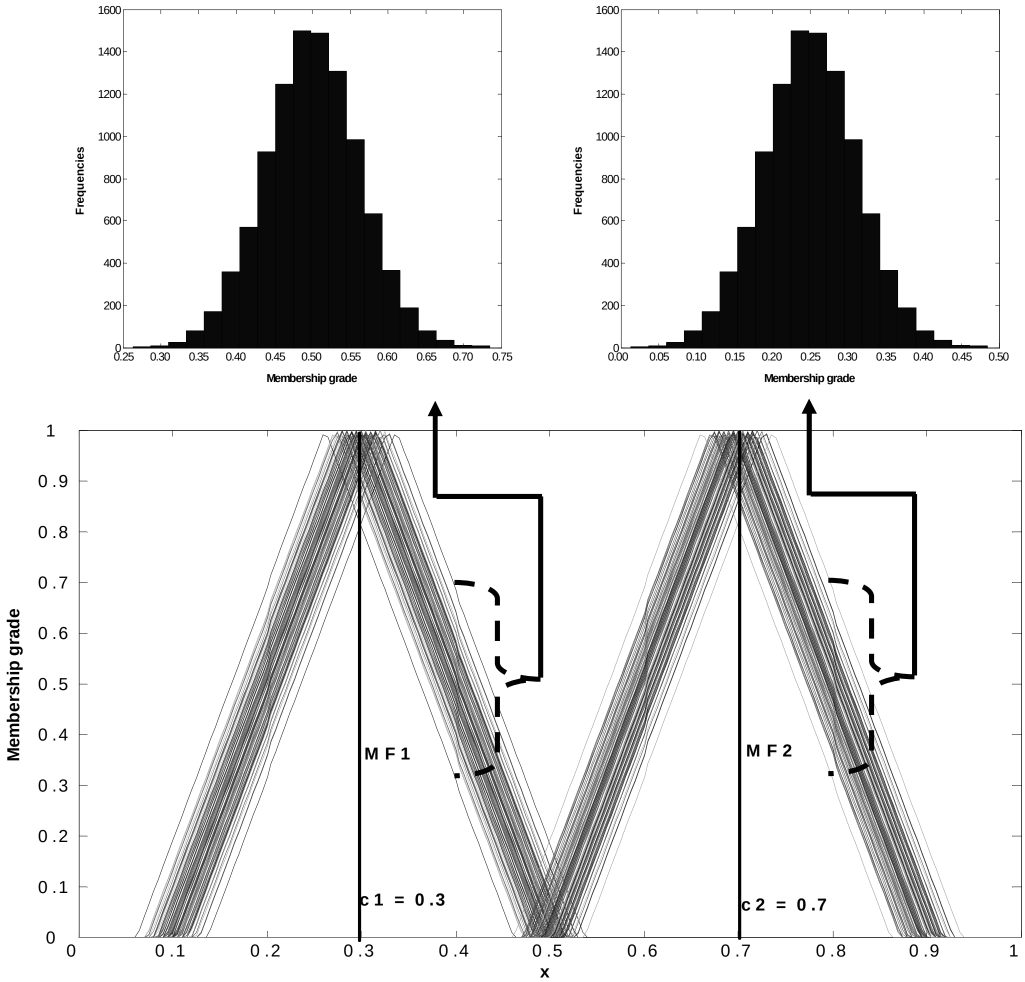


Figure 6.16: Membership grades (nsFLS) for X1=0.40, X2=0.80, where c=0.30 and 0.70.

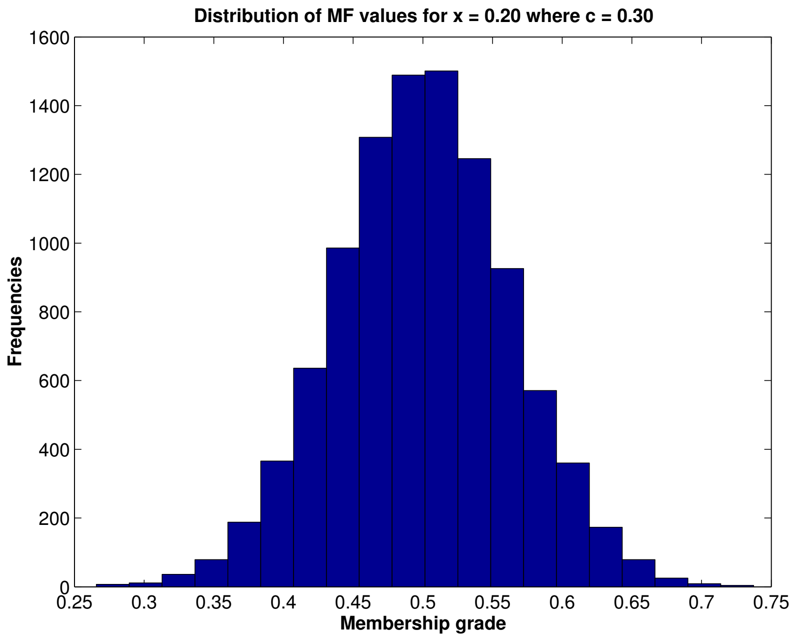


Figure 6.17: Distribution of MF values (centre variation-Triangular MF) for x=0.20 and c=0.30.

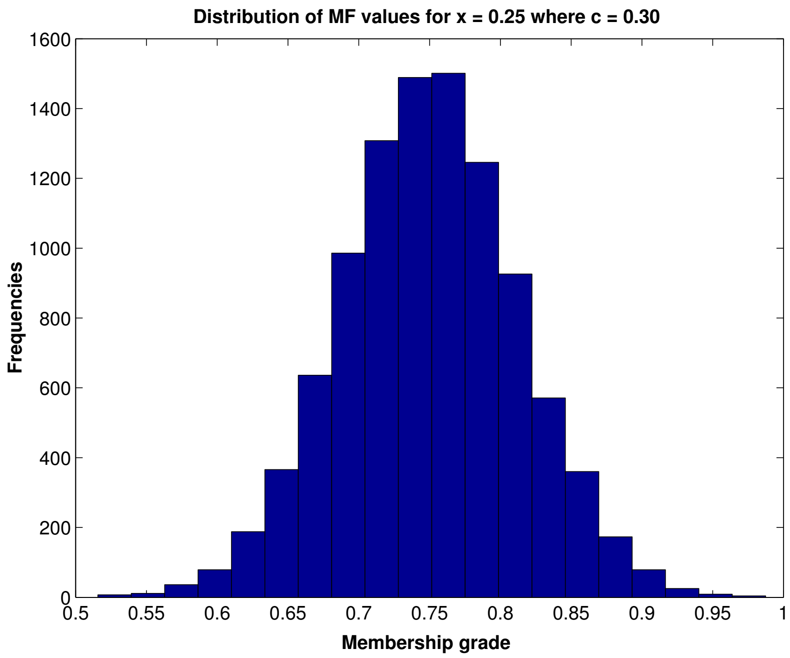


Figure 6.18: Distribution of MF values (centre variation-Triangular MF) for x=0.25 and c=0.30.

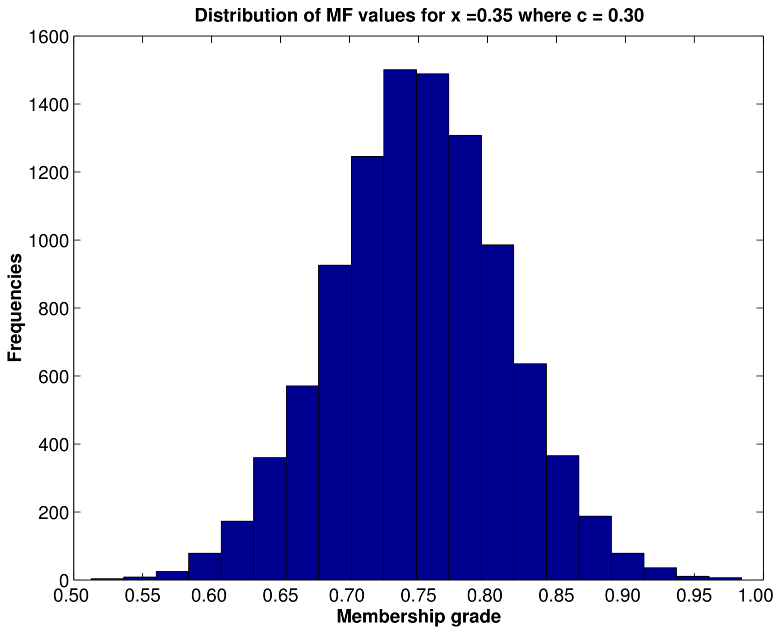


Figure 6.19: Distribution of MF values (centre variation-Triangular MF) for x=0.35 and c=0.30.

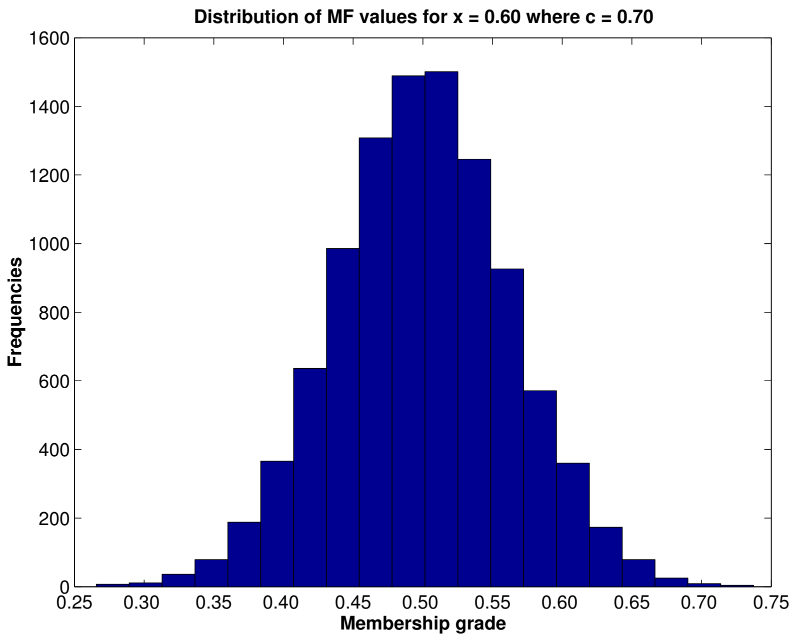


Figure 6.20: Distribution of MF values (centre variation-Triangular MF) for x=0.60 and c=0.70.

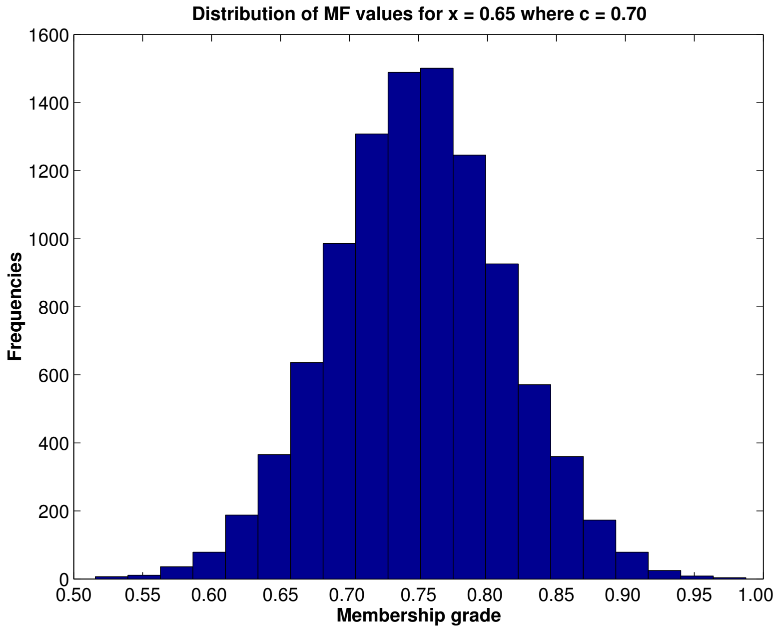


Figure 6.21: Distribution of MF values (centre variation-Triangular MF) for x=0.65 and c=0.70.

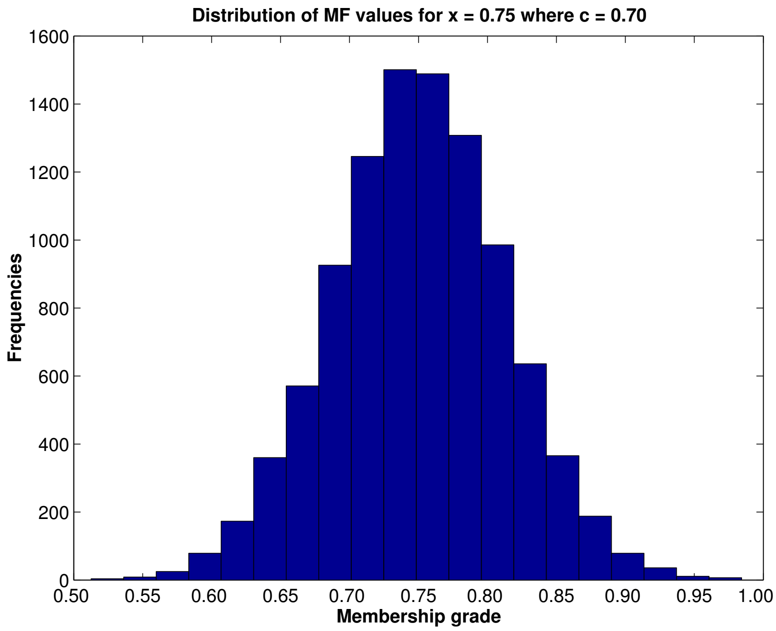


Figure 6.22: Distribution of MF values (centre variation-Triangular MF) for x=0.75 and c=0.70.

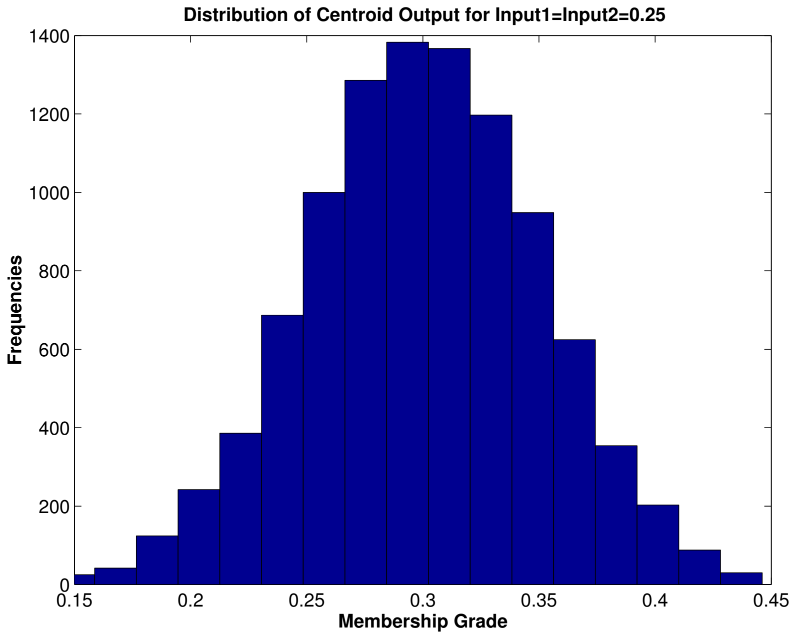


Figure 6.23: Distribution of Centroid Output (centre variation-Triangular MF) for Input1 = Input2 = 0.25.

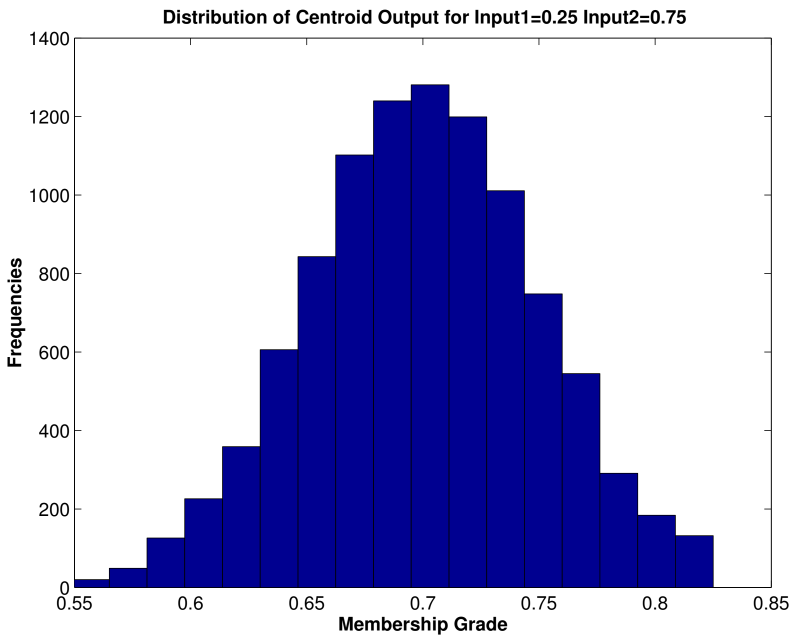


Figure 6.24: Distribution of Centroid Output (centre variation-Triangular MF) for Input1 = 0.25, Input2 = 0.75.

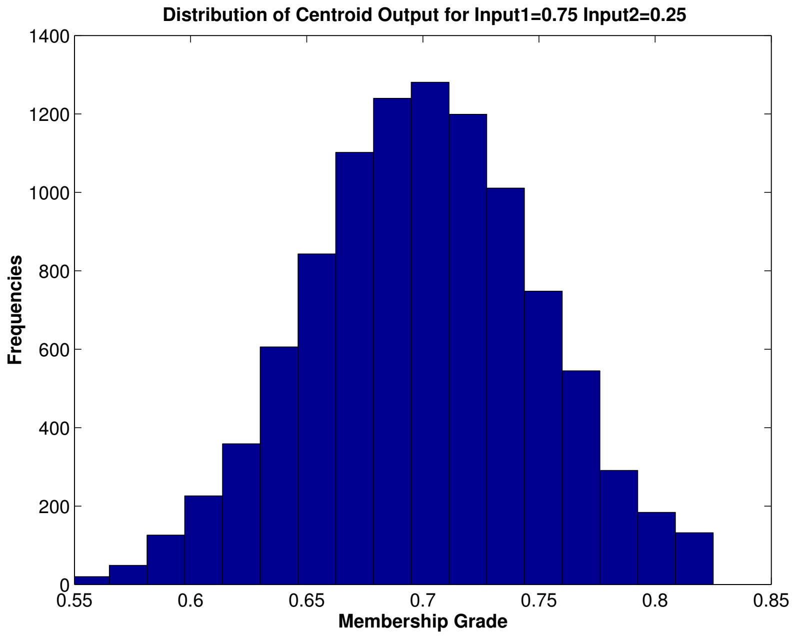


Figure 6.25: Distribution of Centroid Output (centre variation-Triangular MF) for Input1 = 0.75, Input2 = 0.25.

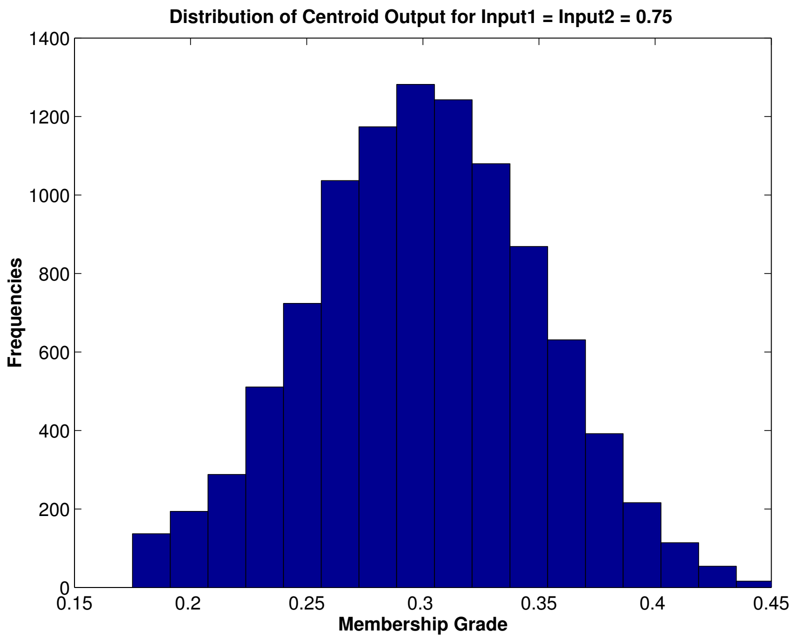


Figure 6.26: Distribution of Centroid Output (centre variation-Triangular MF) for Input1 = Input2 = 0.75.

ics a type-2 system in that the union of all possible instantiations of a non-stationary fuzzy set defines the region over which the underlying membership function varies, similar to the concept to the concept of the footprint of uncertainty of a type-2 set. The work on investigating and formalising this relationship is currently ongoing. At this stage, it can be stated that non-stationary fuzzy sets do allow the representation of a form of uncertainty in the membership function. In this way, they are moving away from the precise membership functions of type-1 fuzzy sets and moving toward satisfying the intention of type-2 sets.

Non-stationary systems do create an uncertainty in the output of inference. For non-stationary FISs featuring perturbation functions that are not uniformly distributed, the *vertical slices* (loosely analogous to the secondary membership functions of type-2 systems) of both the input sets and the output sets are not uniform. Furthermore, it is evident that changing the distribution of the perturbation function alters the distributions of the *vertical slices*. In essence, simply by switching the distribution of the perturbation function from a uniform distribution to a Gaussian one, the non-stationary FIS switches to exhibiting non-uniform distributions of *vertical slices*.

The results obtained suggest that there is some correspondence between the inferencing of non-stationary FIS and that of an interval type-2 FIS. That is, the non-stationary FIS is, in some sense, emulating the inferencing of a general type-2 FIS. Exactly established but, nevertheless, it would appear that it may be reasonable close. It is well known that, while interval type-2 fuzzy sets permit tractable inference and defuzzification, the use of general type-2 fuzzy sets render the inferencing process intractable (although some advances have been made recently in providing approximations [65]). On the other hand, altering the perturbation function within a non-stationary FIS has no effect on the difficulty of the inferencing process. This observation means that, non-stationary systems featuring non-uniformly distributed perturbation functions may allow approximations of general type-2 fuzzy inference to be carried.

6.5 Summary

In this chapter, two case studies has been carried out to illustrate the use of non-stationary fuzzy sets and to explore the relationship between the performance of non-stationary and interval type-2 fuzzy inferencing systems in terms of secondary membership functions. In first case study, Gaussian membership functions were used as the underlying membership function in all both non-stationary and interval type-2 fuzzy systems. There are two forms of variation were implemented such as variation in location or centre and variation in width or standard deviation. In second case study, Triangular functions were used as the underlying membership function. Only one form of variation, variation in centre or location, was implemented.

In summary, it is clearly that a non-stationary system mimics a type-2 system in that the union of all possible instantiations of non-stationary fuzzy set defines the region over which the underlying membership function varies, termed as *footprint of variation* (FOV), similar to the concept of the footprint of uncertainty (FOU) of a type-2 set. It can be observed that non-stationary fuzzy systems featuring non-uniformly distributed perturbation functions may allow approximations of general type-2 fuzzy inference to be carried. Of course, more work needs to be done on non-stationary fuzzy sets before any definitive claim can be made in regard to the correspondence between them and type-2 sets but it can be concluded that, even at this early stage, non-stationary fuzzy sets are a useful addition to the range of fuzzy methods.

Research on understanding and modelling the dynamics of variation in human decision making is ongoing and issues surrounding the use of non-stationary fuzzy sets. In next chapter, the relationships between the shape of the underlying membership functions and the uncertainties obtained in the output sets for both non-stationary and interval type-2 fuzzy systems will be explored.

Chapter 7

Investigate the Underlying Membership Functions of Non-stationary Fuzzy Sets

7.1 Introduction

The aim of this study was to explore relationships between the shape of the underlying membership functions (MFs) and the uncertainties obtained in the output sets for both non-stationary and interval type-2 fuzzy systems. The study was carried out on a fuzzy system implementing the standard XOR problem, in which either Gaussian or Triangular membership functions were employed, using a range of input values and recording the size of the output intervals obtained.

As mentioned in Chapter 6, Garibaldi et al. [10–14] have been investigating the incorporation of variability into decision making in the context of fuzzy expert systems (FESs) in the medical domain. In this work, Garibaldi proposed the notion of 'non-deterministic fuzzy reasoning' in which variability is introduced into the membership functions of a fuzzy system through the use of random alterations to the parameters of the generating function(s). Later, Garibaldi and Musikasuwan [135, 136] extended and formalised this notion through the introduction 'non-stationary fuzzy sets'. Full details about non-stationary fuzzy sets and systems are provided in Chapter 5.

The research presented here is continued from [135] and [137] (in Chapter 6) and divided into 2 main sections (2 experiments). The experiments were designed by constructing the interval type-2 and non-stationary fuzzy systems using Gaussian or Triangular

MFs as the underlying MFs in a system to predict the results of the standard XOR problem, over a wide range of input values ($21 \times 21 = 441$ pairs). The lower bound, mean, upper bound, and interval of the output for each system were computed and recorded.

7.2 Exploring Gaussian and Triangular Underlying Membership Functions

In the research presented here, in order to explore how the form of the underlying membership function affects the inference process withing a non-stationary fuzzy system, a study was carried out on a fuzzy system implementing the XOR problem, in which either Gaussian or Triangular membership functions were employed. Investigations were carried out onto different *perturbation functions* and different type of variation. This non-stationary fuzzy systems were also compared to conventional type-2 fuzzy systems featuring equivalent *Footprints of Uncertainty*. Non-stationary fuzzy systems using two difference shapes of underlying membership functions, i.e., Gaussian and Triangular membership functions are explored.

In order to investigate the effect of different underlying membership shapes in non-stationary fuzzy sets, Gaussian and Triangular membership functions were compared with interval type-2 sets. As stated earlier, this work is continued from [135] and this section focuses on constructing fuzzy systems to solve the standard XOR problem.

In this study, fuzzy systems were constructed to predict the output of truth value where both input variables can take any value in the range of [0,1]. All fuzzy systems consist of two input variables which are *Input1* and *Input2*, one output variable which is *Output*, and four rules. Each variable consist of 2 Gaussian or Triangular membership functions which are *Low* and *High*. The following 4 rules are used for all fuzzy systems. These rules are constructed based on the standard XOR problem.

1. IF *Input1* is *Low* AND *Input2* is *Low*
   THEN *Output* is *Low*
2. IF *Input1* is *Low* AND *Input2* is *High*
   THEN *Output* is *High*

3. IF *Input1* is *High* AND *Input2* is *Low*
   THEN *Output* is *High*
4. IF *Input1* is *High* AND *Input2* is *High*
   THEN *Output* is *Low*

There are three kinds of perturbation function that were used in this study, as follows:

- Sine based function (where $\omega = 127$)
- Uniformly distributed function
- Normally distributed random function

Sine based and Uniformly distributed functions return numbers in the range $[-1, 1]$, while the third (the Matlab *randn* function) returns real numbers sampled from a Normal distribution with mean zero and standard deviation one.

7.2.1 Gaussian Underlying Membership Functions

In this study the underlying Gaussian membership functions as shown in Figure 7.1 were used and two kinds of variation were investigated, i.e. *centre variation* and *width variation*.

7.2.1.1 The Non-stationary FISs:

In both case of centre and width variations, 3 different fuzzy systems (described by perturbation function used to generate membership functions, i.e.; Sine function, Uniformly distributed, and Normally distributed) were designed with two inputs (antecedents), one output (consequent), two Gaussian membership functions for each antecedents and consequent, and four rules. All terms (two inputs and one output) had two Gaussian membership functions, corresponding to meanings of *Low* and *High*. *Low* membership functions all had centre 0.1, *High* membership functions all had centre 0.9. Finally, the initial widths for all membership functions for all terms were 0.5. Note that the parameters of the underlying membership functions were chosen completely arbitrarily, since their precise values to be of any importance are not considered. The purpose of the study is purely

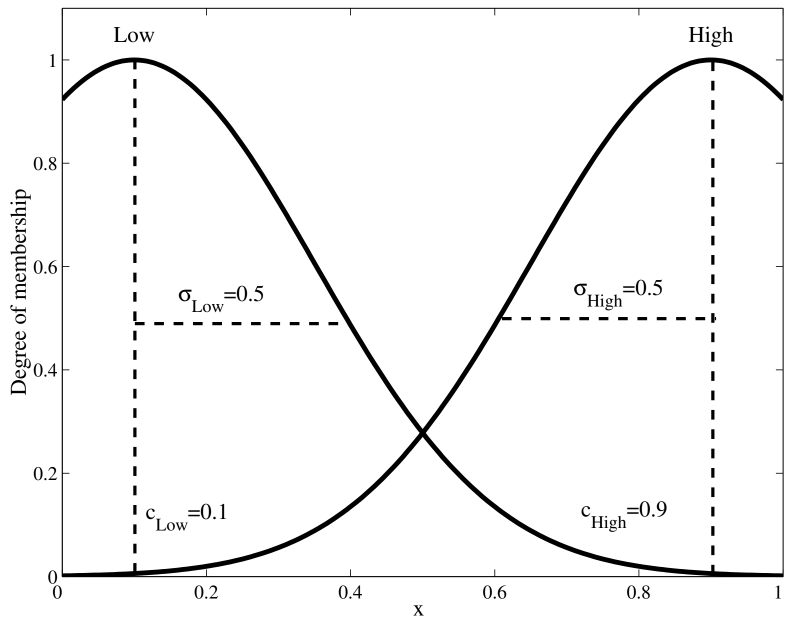


Figure 7.1: Gaussian underlying membership functions used in this case study.

to explore the similarities and differences between the non-stationary fuzzy systems and the equivalent interval type-2 systems in each case.

The four input vectors, (0.25,0.25) (0.25,0.75) (0.75,0.25) and (0.75,0.75), were presented to the system and each time the non-stationary fuzzy sets were generated by replacing centre (c) or width (σ) with $c = c + 0.05f(t)$ or $\sigma = \sigma + 0.05f(t)$ (where $f(t)$ represents the chosen *perturbation function*), respectively. To clarify, the non-stationary fuzzy sets were regenerated for *each* input vector. This process was repeated a fixed number of times (30 times for this study). As an aside, note that it would appear to be a perfectly acceptable design choice to generate the fuzzy sets of the non-stationary system once before presenting the four input vectors, and then to regenerate once again for the next set of four input vectors. The investigation of alternative design choice will be continued for the future work.

7.2.1.2 The Interval type-2 FISs:

Two interval type-2 systems were also designed with 2 inputs (antecedents), 1 output (consequent), 2 Gaussian membership functions for each antecedent and consequent, and four rules. The membership functions all had the same centre and width parameters as described above.

In the type-2 system, the footprint of uncertainty of the type-2 membership functions were created by deviating the parameters of the original type-1 membership functions by a percentage of the universe of discourse of the variables that they were associated to. Two different methods were used to create these type-2 membership functions: by varying the centre point, and varying the width around the original type-1 MF. In the case of varying the centre, the centre of lower and upper bounds membership functions were defined by shifting the initial centre point both left and right for 5% of universe of discourse of variable that MF belongs to, respectively, as follows:

- Centre of lower MF = $c \pm 0.05$

Similarly, in the case of varying the width, the width of lower and upper bounds membership functions were defined by shifting the initial width both left and right for 5% of universe of discourse of variable that MF belongs to, respectively, as follows:

- Width of lower MF = $\sigma \pm 0.05$

7.2.2 Triangular Underlying Membership Functions

For the case of Triangular membership functions, four kinds of variation were investigated, i.e. *centre variation*, *begin-point variation*, *end-point variation*, and *begin & end point variation*.

7.2.2.1 The Non-stationary FISs:

The Triangular shapes used throughout this case study to represent membership functions are shown in Figure 7.2.

The non-stationary fuzzy sets were then generated by replacing the begin-point a and/or end-point b, or centre-point c in Figure 7.2 with $a = a + 0.05f(t)$, $b = b + 0.05f(t)$, and $c = c + 0.05f(t)$, where $f(t)$ represents the chosen *perturbation function*. This process was again repeated 30 times. In all cases of variation, 3 different fuzzy systems (described by perturbation function used to generate membership functions, i.e.; Sine function, Uniformly distributed, and Normally distributed) were designed with two inputs (antecedents), one output (consequent), two Triangular membership functions for

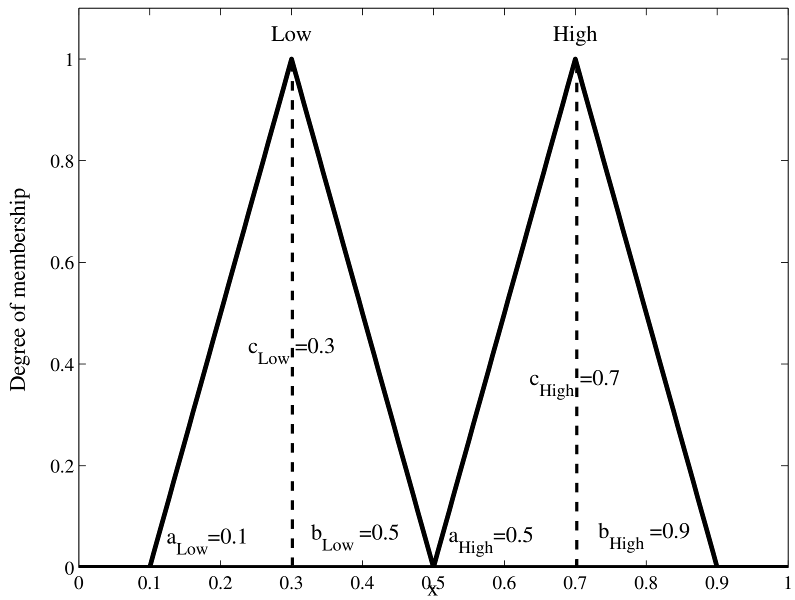


Figure 7.2: Underlying Triangular membership function used in this case study.

each antecedents and consequent, and four rules. All terms (two inputs and one output) had two Triangular membership functions, corresponding to meanings of *Low* and *High*. *Low* membership functions all have ordinary centre $c = 0.3$, $a = 0.1$, and $b = 0.5$; *High* membership functions all had ordinary centre $c = 0.7$, $a = 0.5$, and $b = 0.9$.

7.2.2.2 The Interval type-2 FISs:

Similarly, eight interval type-2 systems were also designed with 2 inputs (antecedents), 1 output (consequent), 2 Triangular membership functions for each antecedent and consequent, and four rules. The membership functions all had the same parameters as described above.

In the type-2 system, the footprints of uncertainty of the type-2 membership functions were created by deviating the parameters of the original type-1 membership functions by a percentage of the universe of discourse of the variables that they were associated with. Four methods were used to create these type-2 membership functions, to match those of the non-stationary systems.

(1) Varying the centre point of the original type-1 MF. The centre of lower and upper bounds membership functions were defined by shifting the initial centre c both left and right for 5% of the universe of discourse of the variable's MF, as follows:

- Centre of lower and upper MF = $c \pm 0.05$

(2) Varying the begin-point of the original type-1 MF. The begin-point of lower and upper bounds membership functions were defined by shifting the initial begin-point a both left and right for 5% of the universe of discourse of the variable's MF, as follows:

- begin-point of lower and upper MF = $a \pm 0.05$

(3) Varying the end-point of the original type-1 MF. The end-point of lower and upper bounds were defined by shifting the initial end-point b both left and right for 5% of the universe of discourse of the variable's MF, as follows:

- end-point of lower and upper MF = $b \pm 0.05$

(4) Varying both begin and end points around the original type-1 MF. The begin and end points of lower and upper bounds membership functions were defined by shifting the initial begin and end points a and b both left and right for 2.5% of the variable's MF, as follows:

- begin-point of lower and upper MF = $a \pm 0.025$
- end-point of lower and upper MF = $b \pm 0.025$

7.2.3 Methods

After all systems had been constructed, they were used to predict the output of each of the four input vectors ((0.25,0.25) (0.25,0.75) (0.75,0.25) and (0.75,0.75)). The lower, mean, upper, and interval of the results were computed and recorded.

In the case of interval type-2 systems, the lower and upper outputs were obtained directly [9], and the mean is simply the average of lower and upper bounds. In the case of non-stationary systems, for Sine and Uniform perturbation functions, the lower bound values were derived from minimum output value, the upper bound values were derived from maximum output value, and the mean were derived from average of the output value from 30 the repeated runs. Finally, the interval of the outputs were derived by computing the length between the lower and upper output values.

For the systems generated by Normally distributed random number (only), the lower and upper bounds are derived from $m \pm s$, where m is the mean of the outputs over time and s is the standard deviation. Finally, the outputs of four input sets ([(0.25,0.25) (0.25,0.75) (0.75,0.25) (0.75,0.75)]) were presented in Section 7.2.4.

7.2.4 Results

In the case of Gaussian membership functions, with centre variation, the lower and upper bounds of the obtained values and the final centroid output values for all 4 fuzzy systems are shown in Table 7.1. The same information is also presented for width variation.

Similarly, in case of Triangular membership functions, the lower and upper bounds predicted values and the final centroid output values for all systems are also shown in Table 7.2 and 7.3 — for centre variation; begin point variation; for end point variation; and for both begin and end points variation, respectively.

Table 7.1: Lower, Mean and Upper Bounds for Gaussian Membership Functions.

Variation	*Type*	*Perturbation*	*Input 1 (0.25,0.25)*			*Input 2 (0.25,0.75)*			*Input 3 (0.75,0.25)*			*Input 4 (0.75,0.75)*		
			Lower	*Mean*	*Upper*	*Lower*	*Mean*	*Upper*	*Lower*	*Mean*	*Upper*	*Lower*	*Mean*	*Upper*
Centre	*Type-2*	*Interval*	0.3687	0.3956	0.4224	0.5780	0.6050	0.6320	0.5780	0.6050	0.6320	0.3687	0.3956	0.4224
	Non-Stationary	*Normal*	0.3853	0.3937	0.4020	0.5970	0.6056	0.6141	0.5970	0.6056	0.6141	0.3853	0.3937	0.4020
		Uniform	0.3932	0.3933	0.4033	0.5791	0.6061	0.6331	0.5791	0.6061	0.6331	0.3932	0.3933	0.4033
		Sine	0.3854	0.3939	0.4033	0.5795	0.6065	0.6335	0.5795	0.6065	0.6335	0.3854	0.3939	0.4033
Width	*Type-2*	*Interval*	0.3836	0.3921	0.4007	0.5993	0.6079	0.6164	0.5993	0.6079	0.6164	0.3836	0.3921	0.4007
	Non-Stationary	*Normal*	0.3735	0.3911	0.4088	0.5912	0.6089	0.6265	0.5912	0.6089	0.6265	0.3735	0.3911	0.4088
		Uniform	0.3734	0.3933	0.4097	0.5903	0.6067	0.6267	0.5903	0.6067	0.6267	0.3734	0.3933	0.4097
		Sine	0.3732	0.3923	0.4098	0.5902	0.6078	0.6268	0.5902	0.6078	0.6268	0.3732	0.3923	0.4098

Table 7.2: Lower, Mean and Upper Bounds for Triangular Membership Functions with variation in centre point and begin point.

Variation	*Type*	*Perturbation*	*Input 1 (0.25,0.25)*			*Input 2 (0.25,0.75)*			*Input 3 (0.75,0.25)*			*Input 4 (0.75,0.75)*		
			Lower	*Mean*	*Upper*	*Lower*	*Mean*	*Upper*	*Lower*	*Mean*	*Upper*	*Lower*	*Mean*	*Upper*
Centre point	*Type-2*	*Interval*	0.2833	0.2981	0.3129	0.6871	0.7000	0.7129	0.6871	0.7000	0.7129	0.2871	0.3019	0.3167
	Non-Stationary	*Normal*	0.2843	0.2986	0.3130	0.6881	0.7002	0.7123	0.6881	0.7002	0.7123	0.2874	0.3017	0.3161
		Uniform	0.2835	0.2997	0.3124	0.6873	0.7004	0.7124	0.6873	0.7004	0.7124	0.2873	0.3009	0.3160
		Sine	0.2846	0.2989	0.3129	0.6871	0.7000	0.7129	0.6871	0.7000	0.7129	0.2871	0.3010	0.3166
Begin point	*Type-2*	*Interval*	0.2828	0.3004	0.3180	0.6825	0.7003	0.7180	0.6825	0.7003	0.7180	0.2825	0.3000	0.3175
	Non-Stationary	*Normal*	0.2812	0.3007	0.3203	0.6806	0.7003	0.7201	0.6806	0.7003	0.7201	0.2819	0.3001	0.3183
		Uniform	0.2829	0.3006	0.3173	0.6826	0.7006	0.7173	0.6826	0.7006	0.7173	0.2826	0.3005	0.3168
		Sine	0.2828	0.3001	0.3180	0.6825	0.7000	0.7180	0.6825	0.7000	0.7180	0.2825	0.2999	0.3175

Table 7.3: Lower, Mean and Upper Bounds for Triangular Membership Functions with variation in end point and begin & end point.

Vari-ation	*Type*	*Perturb-ation*	*Input 1 (0.25,0.25)*			*Input 2 (0.25,0.75)*			*Input 3 (0.75,0.25)*			*Input 4 (0.75,0.75)*		
			Lower	*Mean*	*Upper*	*Lower*	*Mean*	*Upper*	*Lower*	*Mean*	*Upper*	*Lower*	*Mean*	*Upper*
End point	*Type-2*	*Interval*	0.2825	0.3000	0.3175	0.6820	0.6998	0.7175	0.6820	0.6998	0.7175	0.2820	0.2996	0.3172
	Non-Stationary	*Normal*	0.2819	0.3001	0.3184	0.6775	0.6992	0.7210	0.6775	0.6992	0.7210	0.2788	0.3004	0.3221
		Uniform	0.2826	0.3006	0.3169	0.6822	0.7005	0.7168	0.6822	0.7005	0.7168	0.2822	0.3004	0.3166
		Sine	0.2825	0.2999	0.3175	0.6820	0.6998	0.7175	0.6820	0.6998	0.7175	0.2820	0.2998	0.3172
Begin & End point	*Type-2*	*Interval*	0.2827	0.3002	0.3177	0.6823	0.7000	0.7177	0.6823	0.7000	0.7177	0.2823	0.2998	0.3173
	Non-Stationary	*Normal*	0.2819	0.3003	0.3187	0.6814	0.7001	0.7188	0.6814	0.7001	0.7188	0.2815	0.2999	0.3183
		Uniform	0.2828	0.3006	0.3170	0.6825	0.7005	0.7170	0.6825	0.7005	0.7170	0.2824	0.3005	0.3167
		Sine	0.2827	0.3000	0.3177	0.6823	0.7000	0.7177	0.6823	0.7000	0.7177	0.2823	0.2999	0.3173

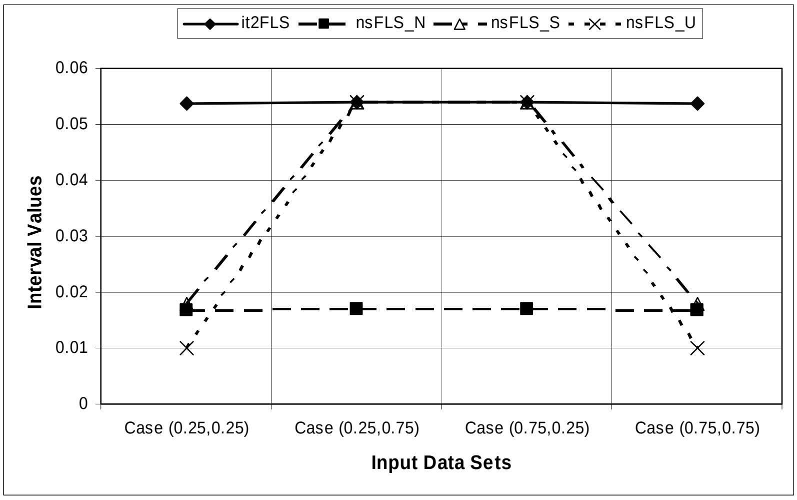


Figure 7.3: Means of the intervals of the outputs for Gaussian non-stationary and interval type-2 fuzzy systems with centre variation.

The length of each results interval was calculated and recorded. In case of Gaussian underlying MF, Figures 7.3 and 7.4 show the plots of mean of intervals for the non-stationary systems together with interval type-2 fuzzy systems with centre and width variation, respectively. Similarly, in case of Triangular underlying MF, the plots of mean of intervals for the non-stationary systems together with interval type-2 fuzzy systems with centre point, begin point, end point, and bengin & end points variation are shown in Figures 7.5, 7.6, 7.7, and 7.8, respectively.

7.2.5 Discussion

The class of a type-2 fuzzy set is determined by the secondary membership function. That is, if the secondary membership function simply takes the value zero outside the lower and upper bounds and 1 inside the bounds, then interval type-2 fuzzy sets are obtained. If (type-1) fuzzy sets are used for the secondary membership functions, then general type-2 fuzzy sets are obtained. In comparison, the class of a non-stationary fuzzy set is determined both by which kind of non-stationarity used (variation in location, variation in slope or noise variation) and by the form of perturbation function used to deviate the underlying membership function, in this study Normally distributed, Uniformly distributed, and Sine based perturbation functions are used to applied in both variation in location and variation in slope. It should be noted, therefore, that herein lies a subtle difference between

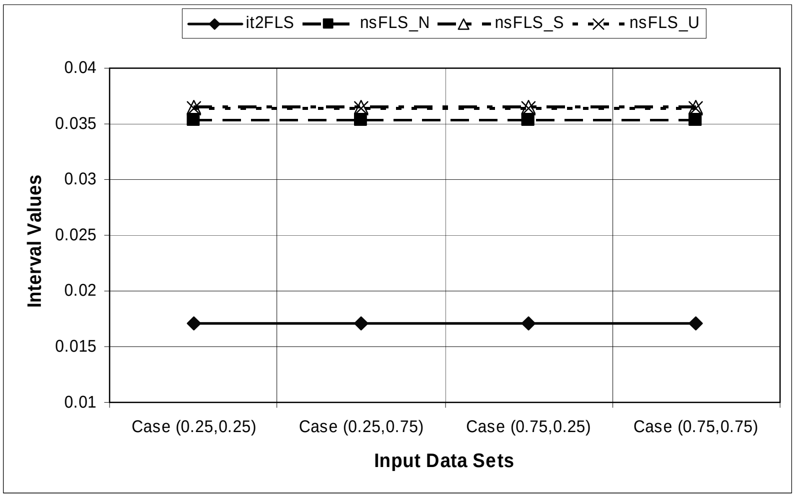


Figure 7.4: Means of the intervals of the outputs for Gaussian non-stationary and interval type-2 fuzzy systems with width variation.

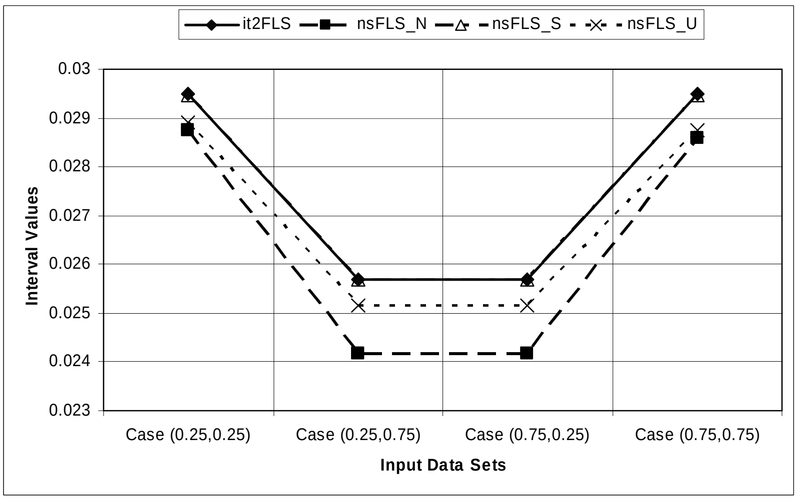


Figure 7.5: Means of the intervals of the outputs for Triangular non-stationary and interval type-2 systems with centre variation.

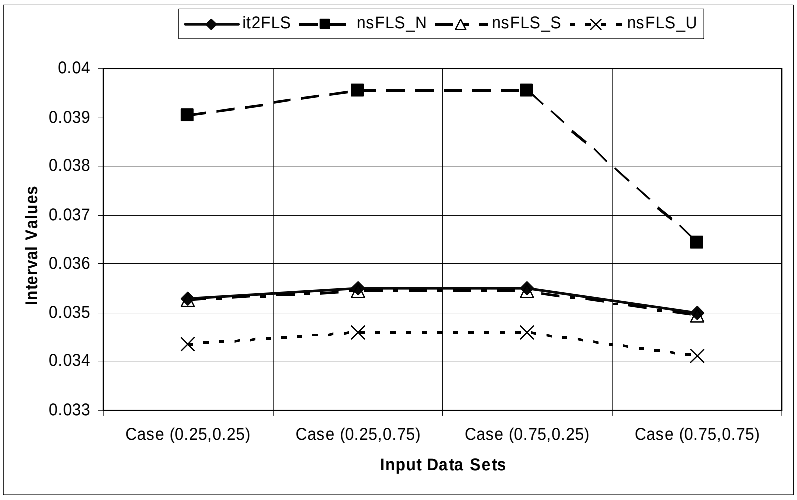


Figure 7.6: Means of the intervals of the outputs for Triangular non-stationary and interval type-2 systems with begin-point variation.

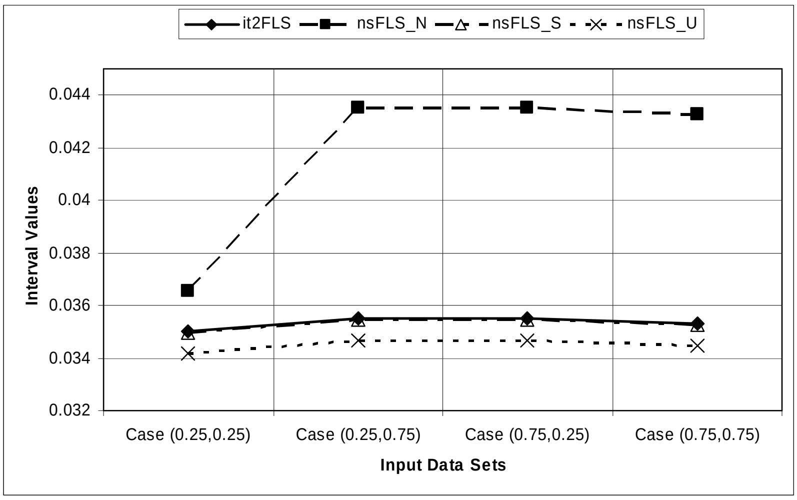


Figure 7.7: Means of the intervals of the outputs for Triangular non-stationary and interval type-2 systems with end-point variation.

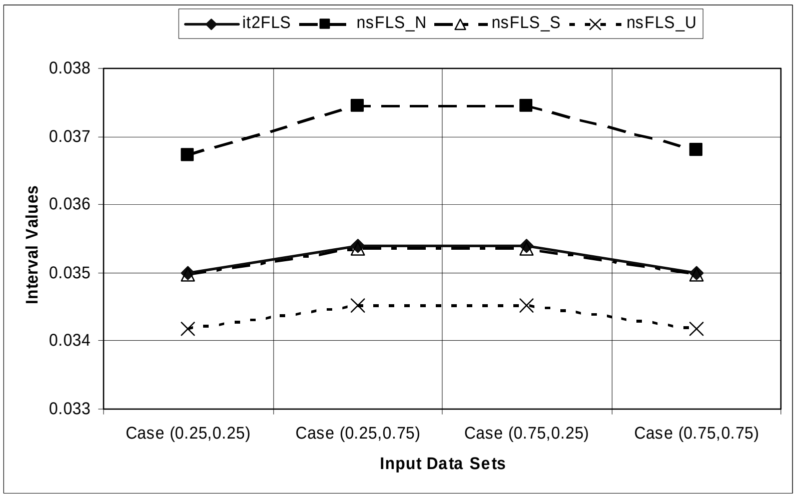


Figure 7.8: Means of the intervals of the outputs for Triangular non-stationary and interval type-2 systems with begin & end points variation.

non-stationary fuzzy sets used in this work and type-2 fuzzy sets. In the non-stationary fuzzy sets used here, the perturbation function acts *horizontally* across the universe of discourse; in type-2 fuzzy sets the secondary membership functions are defined *vertically* along the membership value μ.

For non-stationary fuzzy sets featuring 'noise variation', the perturbation function acts *vertically*. Of course, different perturbation functions can still be used and, thus, such non-stationary fuzzy sets might provide a more 'direct' comparison with type-2 fuzzy sets. Again, further exploration on these areas is continued in the ongoing work.

Turning to the results obtained for the interval of outputs obtained in the experiments carried out. In Figure 7.4 (Gaussian underlying membership functions with width variation), it can be seen that the output interval is constant for the type-2 system and for all the non-stationary systems. However, all the non-stationary systems exhibit (the same) larger output interval. This is a curious finding. In contrast, in Figure 7.3(Gaussian underlying membership functions with centre variation), the picture is very much more complex. The type-2 system has a constant output interval, as does the Normally distributed non-stationary system; however, the Normally distributed non-stationary system now has a *smaller* output interval. Furthermore, the output interval of the Uniform and Sine non-stationary system varies between that corresponding to the Normally distributed non-stationary system for 'symmetric' inputs (0.25,0.25) and (0.75,0.75), and corresponding to the type-2 system for the non-symmetric inputs (0.25,0.75) and (0.75,0.25). Again, these findings are curious.

In case of Triangular underlying membership functions, again the relationships are far from straight-forward. For begin and end-point variation (Figures 7.6 and 7.7), the output intervals appear to be non-symmetrical with the inputs. This is perhaps not surprising, as the membership functions are being altered in a non-symmetrical manner. However, the absolute value of output interval for the Normally distributed non-stationary systems is larger and the non-symmetry is more exaggerated. For the case of centre variation (Figure 7.5), all systems have approximately the same value of interval, which varies according to the input values. For begin and end points (i.e. width) variation (Figure 7.8), the interval of Normally distributed non-stationary systems are larger than all others. It is unable to draw any definitive conclusions from the results obtained here. For all cases except centre variation of Gaussian underlying membership functions, the Sine perturbation function produces results which are *very* close to the interval type-2 systems. Why if should be different for the one case, the answer for this question is needed to be find out in future work.

One might expect Normally distributed non-stationary systems to be different due to the fact that the variation is not hard-limited to the footprint of uncertainty of the corresponding type-2 system. Using a Normally distributed perturbation function, it is theoretically possible for the membership value to be any value for any given input (in effect the footprint of uncertainty is theoretically infinite). However, large deviations are both extremely unlikely theoretically and probably unachievable practically.

Non-stationary fuzzy sets provide a relatively straight-forward mechanism for carrying out inference with fuzzy sets that are *uncertain* in some way. Clearly, non-stationary systems are not direct equivalents of type-2 systems. However, non-stationary fuzzy systems may provide a mechanism whereby a form of fuzzy reasoning which *approximates* (in some meaning of the word) general type-2 fuzzy inference in a simple, fast and computationally efficient manner. An investigation into the relationship between the two frameworks (non-stationary systems and type-2 systems) in order to explore this approximation of interval and general type-2 inference is further explored.

7.3 Underlying Membership Functions vs Output Uncertainties in Type-2 and Non-stationary Fuzzy Sets

The aim of this study was to investigate the relationship between Gaussian and Triangular underlying membership functions used in both non-stationary fuzzy sets and interval type-2 fuzzy sets, and the uncertainties obtained in the outputs.

Fuzzy systems were constructed to predict the output of the XOR truth value where both input variables can take any value in the range of [0,1]. All fuzzy systems consist of two input variables which are *Input1* and *Input2*, one output variable which is *Output*, and four rules. Each variable consist of 2 Gaussian or Triangular membership functions which are *Low* and *High*. In previous work only a restricted range of input values were examined; specifically, the pairs (0.25,0.25), (0.25,0.75), (0.75,0.25) and (0.75,0.75). In this work, each input is varied over the range $[0,1]$ in increments of 0.05, giving a total of 441 pairs of Input1 and Input2 as [(0,0),(0,0.05), ..., (1,0.95), (1,1)]. The following 4 rules were used for all FISs:

1. IF *Input1* is *Low* AND *Input2* is *Low*
   THEN *Output* is *Low*
2. IF *Input1* is *Low* AND *Input2* is *High*
   THEN *Output* is *High*
3. IF *Input1* is *High* AND *Input2* is *Low*
   THEN *Output* is *High*
4. IF *Input1* is *High* AND *Input2* is *High*
   THEN *Output* is *Low*

There are three kinds of perturbation function that were used in this study, as follows:

- Sinusoidal function (where $\omega = 127$)
- Uniformly distributed function
- Normally distributed random function

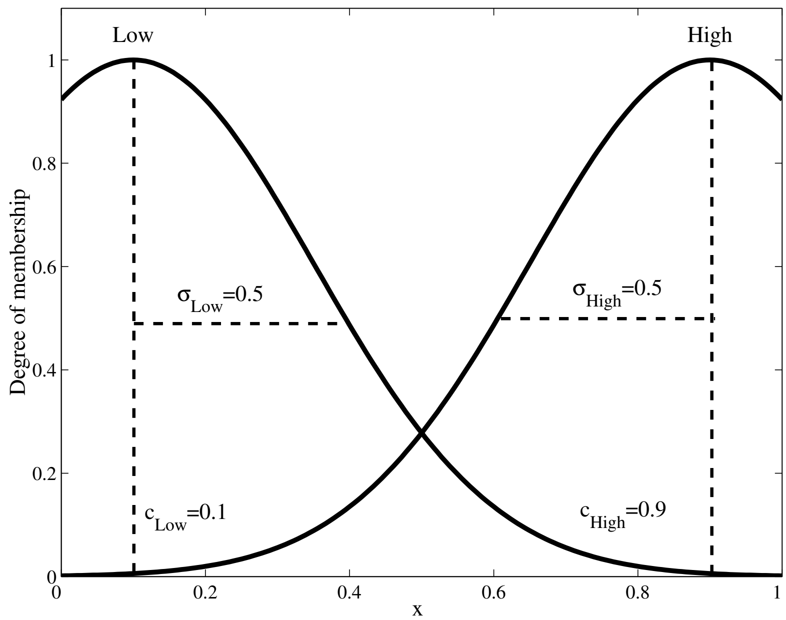


Figure 7.9: Underlying Gaussian membership function used in this case study.

The sinusoidal and uniformly distributed functions return numbers in the range $[-1, 1]$, while the third (the Matlab *randn* function) returns real numbers sampled from a Normal distribution with mean zero and standard deviation one.

7.3.1 Gaussian Membership Functions

The underlying Gaussian membership functions as shown in Figure 7.9 were used and two kinds of variation were investigated, i.e. *centre variation* and *width variation*.

7.3.1.1 The Non-stationary FISs:

The non-stationary fuzzy sets were generated by replacing centre (c) or width (σ) with $c = c + 0.05f(t)$ or $\sigma = \sigma + 0.05f(t)$, where $f(t)$ represents chosen *perturbation function*. The three different perturbation functions described above were used to generate the membership functions. All terms (two inputs and one output) have two Gaussian membership functions, corresponding to meanings of *Low* and *High*. *Low* membership functions all have centre 0.1, *High* membership functions all have centre 0.9. Finally, the initial widths for all membership functions for all terms were 0.5.

7.3.1.2 The Interval Type-2 FISs:

Two interval type-2 FESs have also been designed, where the membership functions all have the same centre and width parameters as described above. The footprints of uncertainty of the type-2 membership functions were created by deviating the parameters of the original type-1 membership functions by a percentage of the universe of discourse of the variables that they are associated with. In the case of centre variation, the centre of lower and upper bounds membership functions were defined by shifting the initial centre point both left and right for 5% of universe of discourse, as follows:

- *Centre of MF* = $c \pm 0.05$

Similarly, in the case of width variation, the width of lower and upper bounds membership functions were defined by shifting the initial width both left and right for 5% as follows:

- *Width of MF* = $\sigma \pm 0.05$

7.3.2 Triangular Membership Functions

In this study four kinds of variation were investigated, i.e. *centre variation*, *begin-point variation*, *end-point variation*, and *both begin and end points variation*.

7.3.2.1 The Non-stationary FISs:

The Triangular underlying membership functions used throughout this case study to represent membership function are shown in Figure 7.10. The non-stationary fuzzy sets were generated by replacing begin-point a and/or end-point b, or centre-point c in Figure 7.10 with $a = a + 0.05f(t)$, $b = b + 0.05f(t)$, and $c = c + 0.05f(t)$, where $f(t)$ represents the chosen *perturbation function*). Once again, the same three perturbation functions were used. Again, all terms (two inputs and one output) have two Triangular membership functions, corresponding to meanings of *Low* and *High*. *Low* membership functions all have ordinary centre (c) 0.3, a is 0.1, and b is 0.5; *High* membership functions all have ordinary centre 0.7, a is 0.5, and b is 0.9.

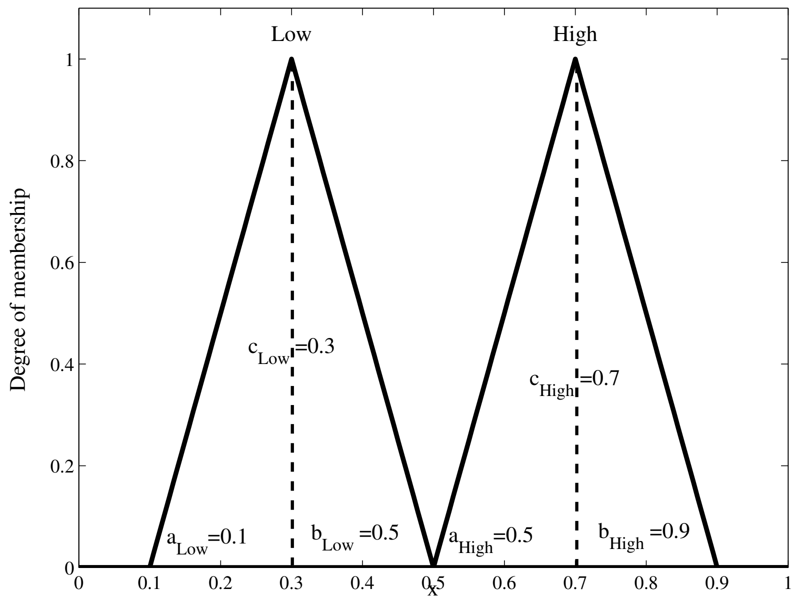


Figure 7.10: Underlying Triangular membership function used in this case study.

7.3.2.2 The Interval Type-2 FISs:

Similarly, four interval type-2 FISs were designed, where the membership functions all have the same parameters as described above. In the type-2 FES, the footprint of uncertainty of the membership functions are created by deviating the parameters of the original type-1 membership functions by a percentage of the universe of discourse of the variables that they are associated with. The four methods used to create these type-2 membership functions were: by (i) varying the centre point around the original type-1 MF both left and right for 5% of the universe of discourse of the variable, as follows:

- *Centre of lower and upper MF* = $c \pm 0.05$

(ii) varying the begin-point (a) both left and right for 5% of the universe of discourse, as follows:

- *begin-point of lower and upper MF* = $a \pm 0.05$

(iii) varying the end-point (b) both left and right for 5% of the universe of discourse, as follows:

- *end-point of lower and upper MF* = $b \pm 0.05$

and (iv) varying both begin and end points (a and b) left and right for 2.5% of the universe of discourse, as follows:

- *begin-point of lower and upper MF* = $a \pm 0.025$

- *end-point of lower and upper MF* = $b \pm 0.025$

7.3.3 Methods

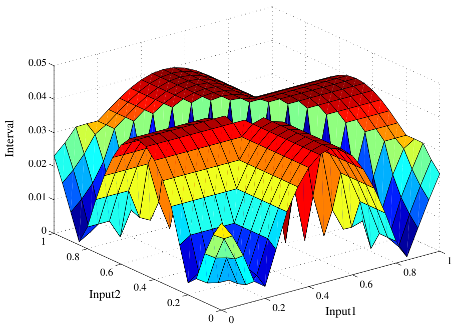
After all systems had been constructed, they were used to obtain the output of each pair of data sets (in total 441 pairs). The lower bound, mean, upper bound, and interval of the outputs were computed and recorded. In the case of interval type-2 systems, the lower and upper output bounds are those obtained directly from the systems; the mean was simply derived from the average of lower and upper outputs.

In the case of non-stationary systems, this process was repeated a fixed 30 times. For the Uniform and Sinusoidal perturbation functions, the lower bound, upper bound and the mean values were simply derived from minimum observed value, the maximum observed value and the mean of the observed values obtained in the 30 repeats, respectively. For the non-stationary systems utilising Normally distributed perturbation functions (only), the lower and upper bounds are derived from $m \pm s$, where m is the mean of the outputs over the 30 repeats and s is the standard deviation. Finally, the lower and upper bounds of the outputs from 441 input data pairs ([(0,0), (0,0.05), ..., (1,0.95), (1,1)]) were used to calculate the length of the interval of the outputs.

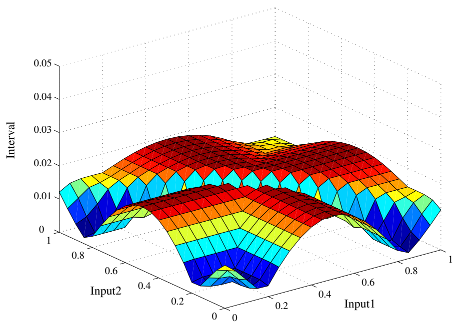
7.3.4 Results

Figures 7.11 and Figures 7.12 are surface plots showing the size of the interval obtained for the output, as a function of the two inputs, of the various systems utilising Gaussian membership functions. Figure 7.11 (a) shows the size of the output interval obtained for the interval type-2 system with centre variation, (b) shows that obtained for the non-stationary system utilising Normally distributed perturbation functions, (c) the non-stationary system using sinusoidal perturbation functions, and (d) using uniformly distributed perturbation functions. Figure 7.12 shows the similar surfaces obtained for systems having width-variation.

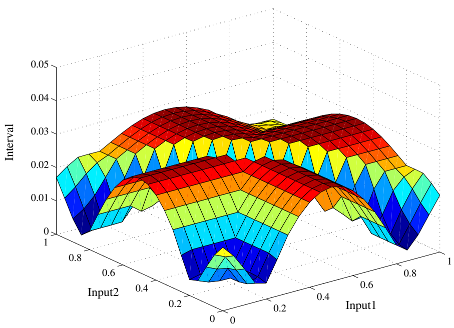
Figures 7.13 – 7.16 show similar plots obtained for systems featuring Triangular underlying membership functions exhibiting centre variation, begin-point variation, end-point variation and begin-end-point variation, respectively. Again, in each case the sur-



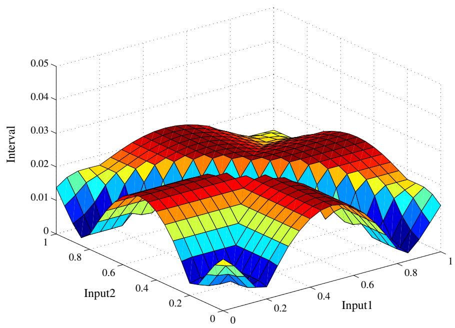
(a) interval type-2 system.



(b) NS system with Normally distributed function.

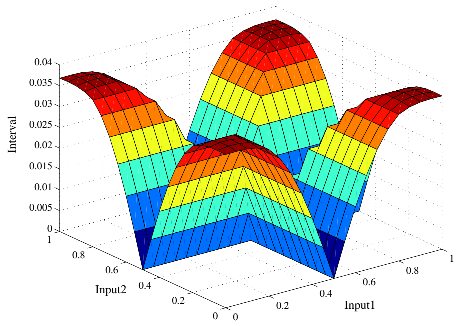


(c) NS system with Sinusoidal function.

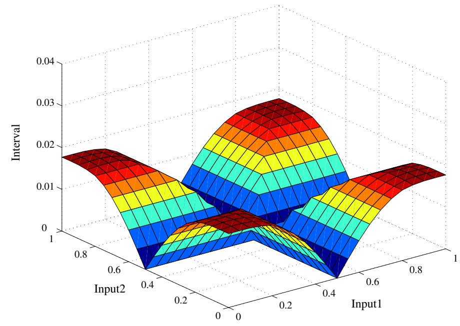


(d) NS system with Uniformly distributed function.

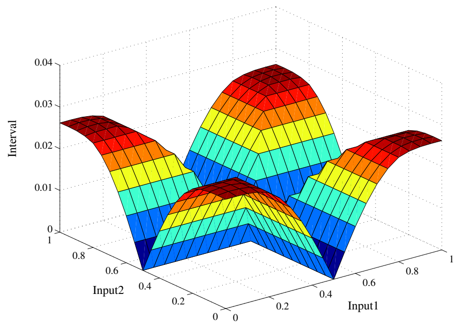
Figure 7.11: The output's intervals for centre variation with Gaussian underlying membership function in type-2 and non-stationary (NS) fuzzy systems.



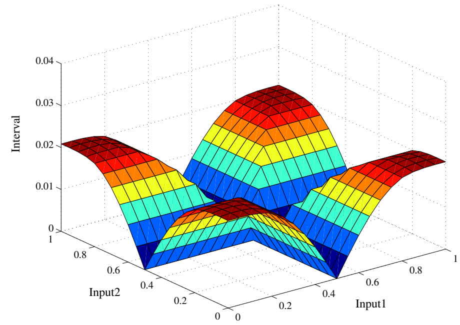
(a) interval type-2 system.



(b) NS system with Normally distributed function.

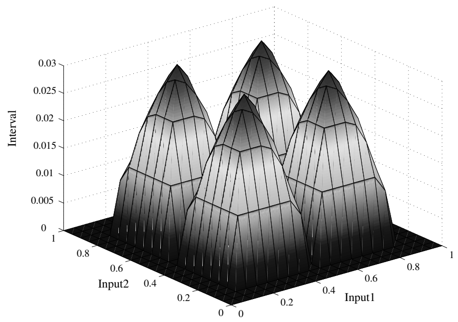


(c) NS system with Sinusoidal function.

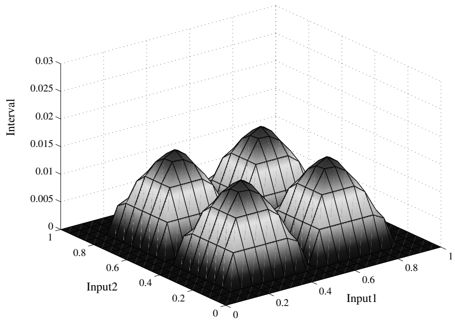


(d) NS system with Uniformly distributed function.

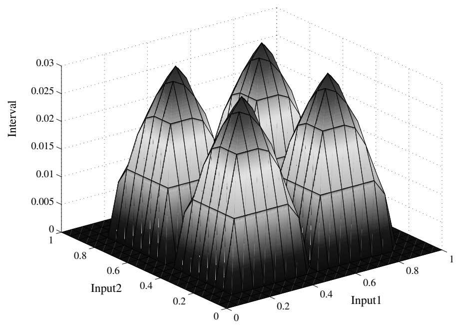
Figure 7.12: The output's intervals for width variation with Gaussian underlying membership function in type-2 and non-stationary (NS) fuzzy systems.



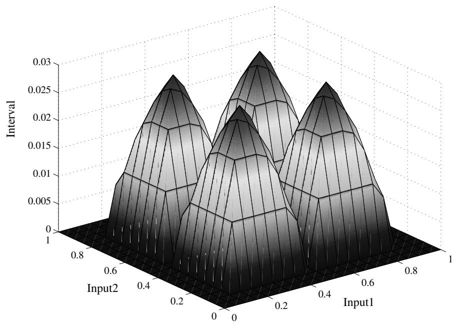
(a) interval type-2 system.



(b) NS system with Normally distributed function.



(c) NS system with Sinusoidal function.

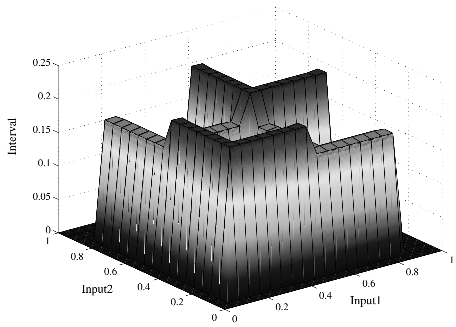


(d) NS system with Uniformly distributed function.

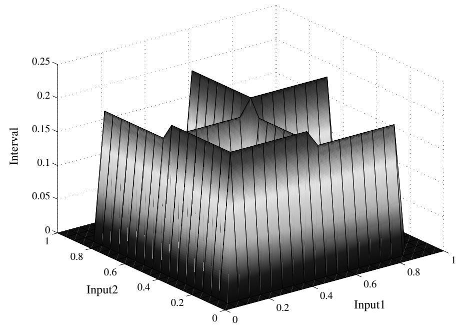
Figure 7.13: The output's intervals for variation in centre point with Triangular underlying membership function in type-2 and non-stationary (NS) fuzzy systems.

faces for (a) type-2, (b) Normally perturbed non-stationary, (c) sinusoidally perturbed non-stationary and (d) uniformly perturbed non-stationary systems are shown.

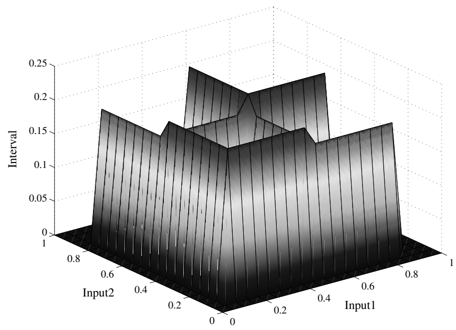
It can be seen that, in general, the shape of the surface obtained is similar for each of the different types of fuzzy system in each case (i.e. the shapes in Figures 7.11 (a) – (d) are similar), although the magnitude varies. In order to explore this further, the difference between the intervals of the outputs of the interval type-2 fuzzy systems and the non-stationary fuzzy system with uniformly distributed perturbation functions were plotted. Figure 7.17 shows the differences for the case of (a) Gaussian underlying membership functions with centre variation, (b) Gaussian underlying membership functions with width variation, (c) Triangular underlying membership functions with centre variation,



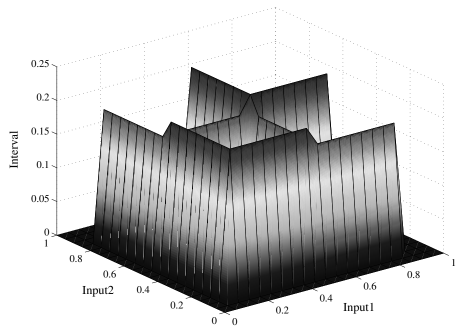
(a) interval type-2 system.



(b) NS system with Normally distributed function.

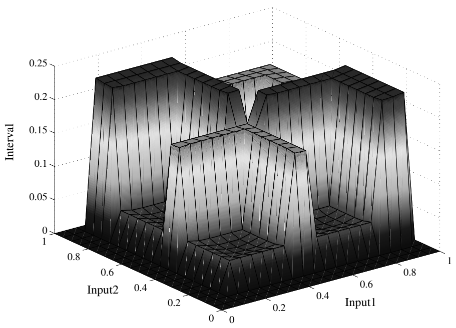


(c) NS system with Sinusoidal function.

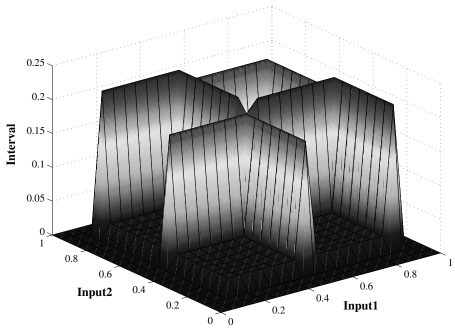


(d) NS system with Uniformly distributed function.

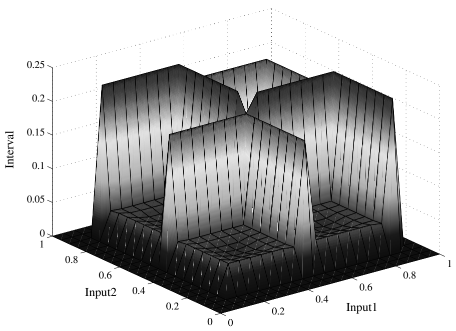
Figure 7.14: The output's intervals for variation in begin point with Triangular underlying membership function in type-2 and non-stationary (NS) fuzzy systems.



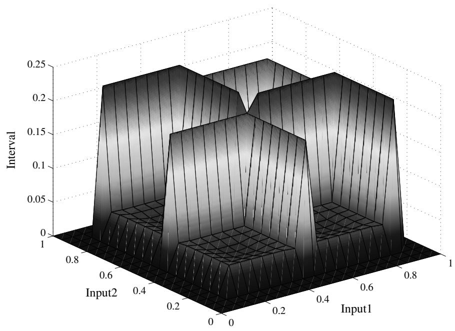
(a) interval type-2 system.



(b) NS system with Normally distributed function.

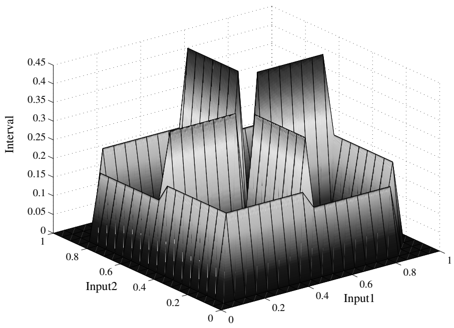


(c) NS system with Sinusoidal function.

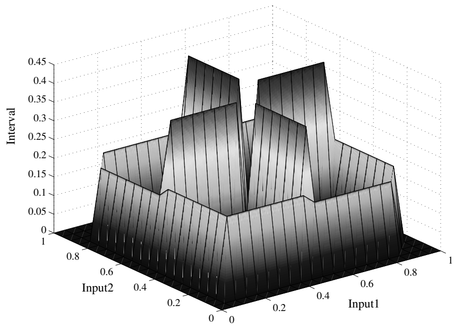


(d) NS system with Uniformly distributed function.

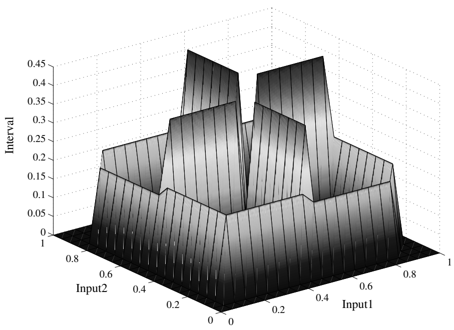
Figure 7.15: The output's intervals for variation in end point with Triangular underlying membership function in type-2 and non-stationary (NS) fuzzy systems.



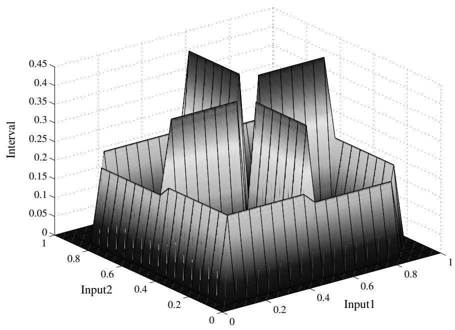
(a) interval type-2 system.



(b) NS system with Normally distributed function.

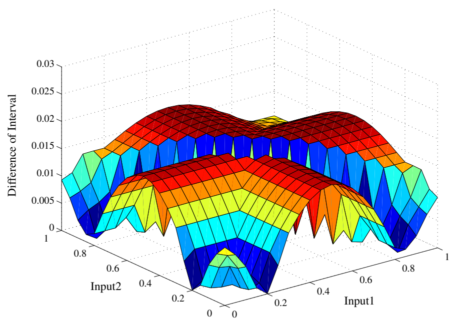


(c) NS system with Sinusoidal function.

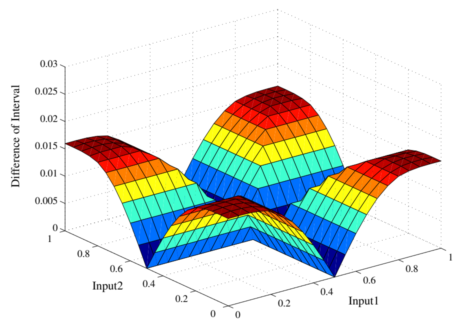


(d) NS system with Uniformly distributed function.

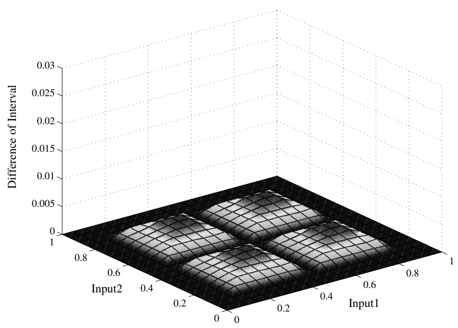
Figure 7.16: The output's intervals for variation in both begin-end point with Triangular underlying membership function in type-2 and non-stationary (NS) fuzzy systems.



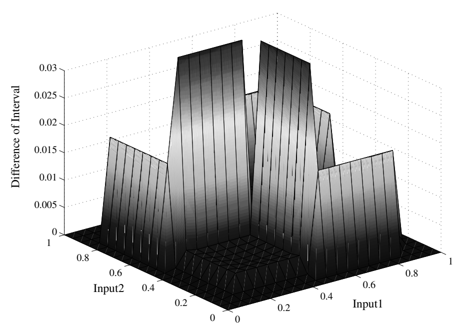
(a) Centre variation with Gaussian underlying MF.



(b) Width variation with Gaussian underlying MF.



(c) Centre variation with Triangular underlying MF.

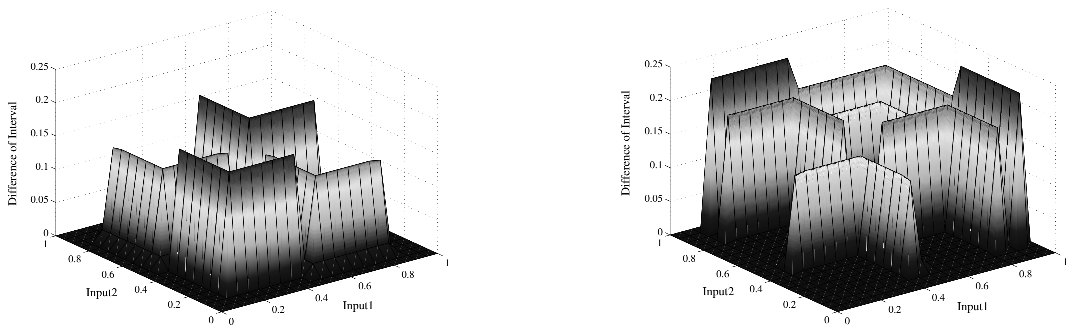


(d) Begin-end point variation with Triangular underlying MF.

Figure 7.17: The difference between the output's intervals of interval type-2 fuzzy system and non-stationary fuzzy system with Uniformly distributed function.

and (d) Triangular underlying membership functions with begin-end point (i.e. width) variation. The sub-Figures (a) – (d) are all plotted with the same z-axis. It can be seen that the differences for the systems with Gaussian underlying membership functions are quite small, the differences for Triangular underlying membership functions with centre-point variation are extremely small, whereas the differences for Triangular membership functions with begin-end point variation are larger and non-symmetrical.

Finally Figure 7.18 shows similar plots for Triangular underlying membership functions with begin-point variation and end-point variation. It should be noted that these have been plotted on a different scale on the z-axis. The differences are all much larger than those in Figure 7.17, and they are (as might be expected) non-symmetrical.



(a) Begin point variation with Triangular underlying MF.

(b) End point variation with Triangular underlying MF.

Figure 7.18: The difference between the output's intervals of interval type-2 fuzzy system and non-stationary fuzzy system with Uniformly distributed function.

7.3.5 Discussion

The term 'footprint of uncertainty' (FOU) was introduced by Mendel to provide 'a very convenient verbal description of the entire domain of support for all the secondary grades of a type-2 membership function' [1]. A similar term, the '*footprint of variation*' (FOV) is introduced, as a similar verbal description of the area covering the range from the minimum to the maximum fuzzy sets which comprise the non-stationary fuzzy sets as shown in Figure 7.19. For non-stationary fuzzy sets which are generated by Uniformly distributed and Sinusoidal perturbation functions (producing random values within $[-1,1]$), the maximum area of FOV will be equivalent to the FOU of interval type-2 fuzzy sets with the same amount of variation. Normally distributed perturbation functions generate random values within $[-\infty,\infty]$, and so an FOV defined as the union of all underlying memberships would fill the entire universe of discourse. This kind of FOV will need further investigation.

In Figure 7.11 and Figure 7.12 (both for systems with Gaussian underlying membership functions) it can be observed that the surfaces are (very approximately) comprised of four superimposed Gaussians. In the case of Figure 7.11 the Gaussian-like shapes are located centrally on the x and y axes, while in Figure 7.12 they are located on the corners. It would appear that there is some complex relationship between the FOV in non-stationary or FOU in type-2 systems and the size of the interval in the output, although this rela-

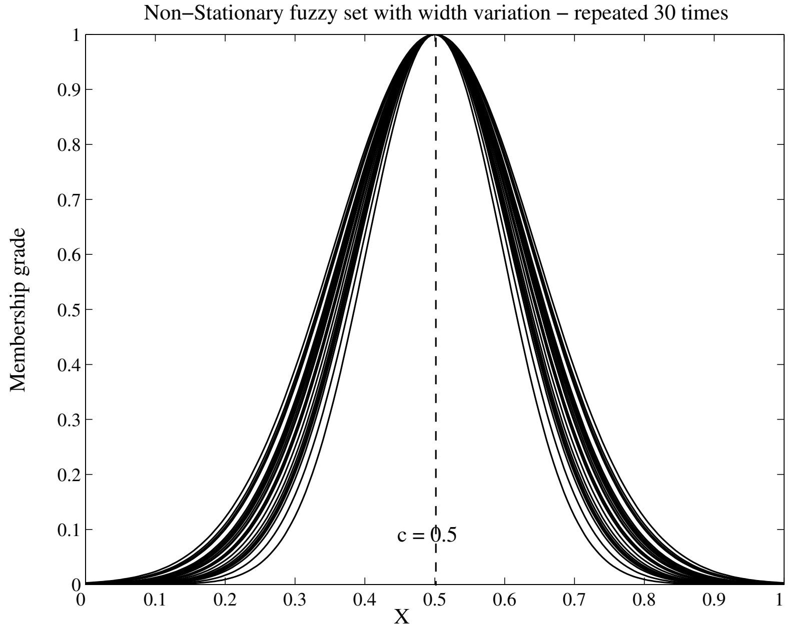


Figure 7.19: Non-stationary fuzzy set with width variation.

tionship cannot yet fully be explained. Again, it can be observed that the surfaces are divided into four equivalent (symmetrical) parts. This might be expected because of the symmetrical nature of the four rule XOR problem and this observation will be further explored. Similarly, Figures 7.13, 7.14, 7.15, 7.16 (Triangular underlying membership functions with centre variation) exhibits a similar four way symmetry, with vaguely Triangular shapes almost appearing as projections of the underlying membership functions. Figure 7.16 also appears to exhibit four way symmetry, but have a more complex form.

An interesting observation from Figures 7.14 and 7.15 is that Figure 7.14 (a) and Figure 7.15 (a) are *not* reflections of each other, as might be expected from the fact that varying the begin-point of the membership functions of all the variables is a reflection of varying the end-point of the membership functions. On examination, it is found that Figure 7.15 (a) can be obtained by rotating Figure 7.14 (a) through 90^o and then rotating each of the four quadrants through 90^o *separately*. It is believed that this is due to the lack of reflective symmetry in the rule set. That is, to obtain total reflection rule 1 would need to be changed to 'IF *Input1* is *Low* AND *Input2* is *Low* THEN *Output* is ***High***', and so on.

There is an interested finding for the systems with Triangular underlying membership functions with one-side slope variation (begin point and end point) as shown in

Figures 7.14 and 7.15. It can be observed that there are plateau in the surface of the intervals of the outputs for the interval type-2 systems only, but this observation does not occur in the equivalent systems with centre and width (begin-end point) variation. At this stage, It cannot be clearly explained the results obtained here, and more investigation and exploration is needed.

It can be seen from Figure 7.17 that the difference of output's interval between type-2 system and non-stationary systems with uniformly distributed perturbation functions is extremely small. Without performing a detailed mathematical analysis of the relationship between type-2 and non-stationary systems, it might be expected that uniformly distributed non-stationary systems with the same FOV as FOU of interval type-2 systems will match closely. It should be remembered that there is a stochastic element to the non-stationary systems, such that after 30 repeats an exact match would not be expected. As mentioned above, for the systems with one-side slope variation, Figures 7.18 (a) and (b), there is a big difference between the output's interval of type-2 system and non-stationary system with uniformly distributed perturbations. Once again, there is unsure of the significance of this finding at present. Non-stationary fuzzy systems may provide a mechanism for implementing a form of fuzzy reasoning which approximates general type-2 fuzzy inference in a simple, fast and computationally efficient manner. We are continuing investigations into the relationship between the two frameworks (non-stationary and type-2 systems).

7.4 Comparison on the Performances of Type-1, Type-2, and Non-stationary Fuzzy Systems for MacKey-Glass Time-Series Prediction

In this section, an experiment was carried out in which the performances of type-1, type-2, and non-stationary fuzzy logic systems (FLSs) were compared in their ability to predict the Mackey-Glass time series with 5 levels of additive noises. Each of the FLSs was tuned to achieve the best possible performance using a standardised gradient descent procedure. These experiments were repeated a number of times in order to establish the mean performance of each FLS. The results show that the best performance was achieved with a type-1 FLS, albeit featuring a high number of tunable parameters. A type-2 FLS with far fewer parameters achieved performance very close to the best.

The purpose of this work was firstly to compare the performance of the type-1, type-2, and non-stationary fuzzy systems. Seven main classes of fuzzy systems are considered:

- **T1-SFLS** - 'conventional' fuzzy systems with singleton inputs and type-1 fuzzy sets throughout;
- **T1-NFLS** - type-1 fuzzy systems with non-singleton (type-1) fuzzy inputs and type-1 fuzzy sets throughout;
- **T2-SLFS** - fuzzy systems featuring interval type-2 sets with singleton inputs;
- **T2-NSLFS-T1** - fuzzy systems featuring interval type-2 sets with type-1 non-singleton inputs;
- **NS-FLS-N** - fuzzy systems featuring non-stationary fuzzy input sets with Normal distribution pertubation;
- **NS-FLS-S** - fuzzy systems featuring non-stationary fuzzy input sets with Sine function distribution pertubation;
- **NS-FLS-U** - fuzzy systems featuring non-stationary fuzzy input sets with Uniform distribution pertubation;

This work has evolved from recent studies in section 4.3. The software for producing result of type-1 and type-2 fuzzy system those used in this experiment is that provided by Professor Mendel (at http://sipi.usc.edu/ mendel/software/).

7.4.1 Methodology

Five independent data sets with 5 different noise levels (totally 25 data sets 2200 series each) were generated. These data sets were generated by using Mackey-Glass time-series delay differential equation shown in Equation 4.1. After 5 data sets were generated, 5 different level of noise were generated as follows:

- Level 1: 0 noise (noise free)
- Level 2: 0.01 noise
- Level 3: 0.05 noise
- Level 4: 0.10 noise
- Level 5: 0.20 noise

where *noise* was a uniformly distributed random number in [-1,1]. Then these 5 different levels of noise were added into the data sets.

All designed models below were based on 1,000 noisy data points: x(501), x(502), ..., x(1500). The First 500 noisy data, x(501), x(502), ..., x(1000) were used for training, and the remaining 500 , x(1001), x(1002), ..., x(1500), were used for testing the design. Four antecedents: x(k-3), x(k-2), x(k-1), and x(k) were used to predict x(k+1).

The performance of all the designs was evaluated using the RMSE as shown below:

$$RMSE = \sqrt{\frac{1}{500}\sum_{k=1000}^{1499}[x(k+1)-f(x^{(k)})]^2} \tag{7.1}$$

where $x^{(k)} = [x(k-3), x(k-2), x(k-1), x(k)]^T$.

7.4.1.1 Type-1 FISs

Type-1 singleton fuzzy logic system (T1-SFLS) and type-1 non-singleton fuzzy logic system (T1-NSFLS) have been designed with 4 antecedents, 2 membership functions for each antecedent, the number of rules are 16 rules (2^4) respectively, each rule is characterized by 8 antecedent MF parameters (means and standard deviations), and 1 consequent parameter ($\bar{y}$). The initial location of each Gaussian antecedent MF is based on the mean (m_x) and standard deviation (σ_x) and the mean of membership functions are:

- Mean of MF1 = $m_x - 2\sigma_x$
- Mean of MF2 = $m_x + 2\sigma_x$

Initially all standard deviation parameters are tuned to σ_x or $2\sigma_x$. Additionally the height defuzzifier and initial centre of each consequent's MF are random numbers in [0,1]. For type-1 NSFLS each of the 4 noisy input measurements are modelled using a Gaussian membership function, a different standard deviation is used for each of the 4 input measurement membership functions (σ_n). Finally, two models were created for type-1 SFLS and type-1 NSFLS for each data set (25 data sets).

7.4.1.2 Type-2 FISs

Interval type-2 singleton FLS (Type-2 SFLS) and type-1 non-singleton type-2 FLS (Type-2 NSFLS-T1) have been designed by using the partially dependent approach. First, the best possible singleton and non-singleton type-1 fuzzy systems were designed by tuning their parameters using back-propagation designs, and then some of those parameters were used to initialise the parameters of the interval type-2 SFLS and type-2 NSFLS-T1. They consisted of 4 antecedents for forecasting, 2 membership functions for each antecedent and 16 rules. The Gaussian primary membership functions of uncertain means for the antecedents were chosen. The means of membership functions are:

- Mean of MF1 = $[m_x - 2\sigma_x - 0.25\sigma_n, m_x - 2\sigma_x + 0.25\sigma_n]$
- Mean of MF2 = $[m_x + 2\sigma_x - 0.25\sigma_n, m_x + 2\sigma_x + 0.25\sigma_n]$

where m_x is the mean of the data in the training parts, and σ_n is the standard deviation of the additive noise. Each rule of the type-2 SFLS and type-2 NSFLS-T1 were characterized by 12 antecedent MF parameters: left and right bounds on the mean, and the standard deviation for each of 4 Gaussian membership functions) and 2 consequent parameters (left and right hand end-points for the centroid of the consequent type-2 fuzzy set). So, in total the number of parameters tuned for type-2 SFLS is 224. Standard deviation for each of the 4 input measurement membership functions (σ_n) is used in type-2 NFLS-T1. So in total the number of parameters tuned for type-2 SFLS is 225.

Initially the final tuned results were used for the standard deviation of the input, σ_x or $2\sigma_x$, obtained from type-1 NSFLS design, and also $\bar{y}^i$ was obtained from type-1 SFLS and then initial $\bar{y}^i_r$ and $\bar{y}^i_l$ was chosen as:

$\bar{y}^i_r = \bar{y}^i + \sigma_n$

$\bar{y}^i_l = \bar{y}^i - \sigma_n,$

where i = 1, 2, ..., 16

Finally, two different models for both type-2 SFLS and type-2 NSFLS-T1 were created for each data set (totally 25 data sets).

7.4.1.3 Non-stationary FISs

Non-stationary fuzzy systems has been initially designed same as type-1 FLSs mentioned above. Afer that, the non-stationary fuzzy sets were generated by replacing centre $(m_x) mx = m_x + 0:05 f(t)$, where $f(t)$ represents chosen perturbation function.

There are three kinds of perturbation function that were used in this study, as follows:

- Sine based function (where $\omega = 127$) (for NS-FLS-S)
- Uniformly distributed function (for NS-FLS-U)
- Normally distributed random function (for NS-FLS-N)

In this experiments, there is only centre variation has been investigated. The three different perturbation functions (normal distribution, sine, and uniform functions) were used to generate the membership functions with 1000 iterations. All terms (four inputs and one output) have two Gaussian membership functions, corresponding to meanings of MF1 and MF2. The initial location of each Gaussian antecedent MF is based on the mean (m_x) and standard deviation (σ_x) and the mean of membership functions are:

- Mean of MF1 = $m_x - 2\sigma_x$
- Mean of MF2 = $m_x + 2\sigma_x$

Finally, three different models among NS-FLS-N, NS-FLS-S, and NS-FLS-U were created for each data set (totally 25 data sets).

7.4.2 Results

After all models had been constructed and run, the performances of the type-1, type-2, and non-stationary fuzzy systems were compared. The results of performance of 7 different models are shown as follows. Table 7.4 shows the result obtained from the mean of RMSE of the best model for 7 different fuzzy system models with five noise levels.

Table 7.4: The mean of RMSE of the best model for 7 different fuzzy system models with 5 different noise levels

FLS	DataSet-1	DataSet-2	DataSet-3	DataSet-4	DataSet-5
Type-1-SFLS	0.2764	0.2544	0.2874	0.2920	0.2832
Type-1-NSFLS	0.2227	0.2296	0.2351	0.2322	0.2466
Type-2-SFLS	0.2206	0.2285	0.2374	0.2322	0.2522
T1-NS-T2-FLS	0.2204	0.2276	0.2354	0.2302	0.2433
NS-FLS-N	0.1816	0.1995	0.1687	0.1836	0.2064
NS-FLS-S	0.2348	0.2563	0.2220	0.2338	0.2590
NS-FLS-U	0.2326	0.2544	0.2198	0.2317	0.2570

Figure 7.20 shows the performance (RMSE) of 7 different models for the 5 different noise levels averaged over five separate runs.

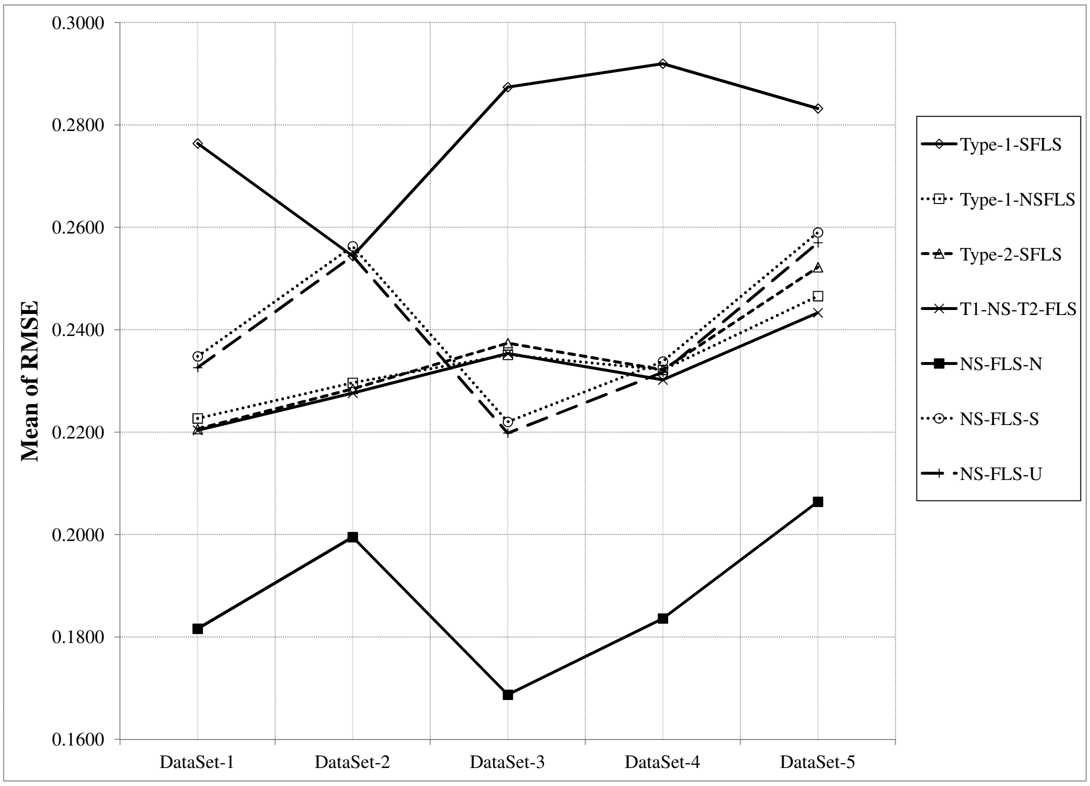


Figure 7.20: Graph of mean of RMSE of 7 models for 5 noise levels

7.4.3 Discussion

All cases the performance of type-1 fuzzy systems with singleton inputs (Type-1-SFLS) as shown in Table 7.4 and Figure 7.20, the most common found in practice, is worse than for the type-1 non-singleton fuzzy systems (Type-1-NSFLS), the type-2 fuzzy systems (Type-2-SFLS and T1-NS-T2-FLS), and non-stationary fuzzy systems (NS-FLS-N, NSFLS- S, and NS-FLS-U).

Again, we found that, type-1 non-singletion fuzzy system (Type1-NSFLS), type-2 fuzzy systems (both Type-2-SFLS and T1-NS-T2-FLS) are performed as good as each other for all data sets. Anyway, these systems were performed better than non-stationary fuzzy systems (NS-FLS-S and NS-FLS-U) except data set 3 that these 2 non-stationary fuzzy systems performed better.

The interesting finding from the experiments, we found that the best overall performance is achieved with NS-FLS-N. This is a non-stationary fuzzy system uisng normal distribution function as a pertubation function. From this, we may tentatively suggest that while type-2 fuzzy systems may not strictly be necessary in order to achieve optimal performance, their benefit may lie more in achieving good performance in a more tractable model.

Finally, we note that these finding are for one particular data set (Mackey-Glass Time-Series) only and hence, no general conclusions can be made from them alone. In order to reach general conclusions it would be necessary to carry out similar experiments on a wide variety of data sets. There is no evidence at present to suggest that the similar results would necessarily be obtained for other kinds of data.

With zero noise, M5 (with only 145 parameters) achieves better performance than M11 or M12. This agrees with Mendel's previous findings that in the absence of noise a type-1 fuzzy system with non-singleton inputs is an adequate model for capturing the uncertainty.

7.5 Summary

The class of a type-2 fuzzy set is determined by the secondary membership function. That is, if the secondary membership function simply takes the value zero outside the lower and upper bounds and 1 inside the bounds, then interval type-2 fuzzy sets are obtained. If (type-1) fuzzy sets are used for the secondary membership functions, then general type-2 fuzzy sets are obtained. In comparison, the class of a non-stationary fuzzy set is determined both by which kind of non-stationarity used (variation in location, variation in slope or noise variation) and by the form of perturbation function used to deviate the underlying membership function, in this study Normally distributed, Uniformly distributed, and Sine based perturbation functions are used to applied in both variation in location and variation in slope. It should be noted, therefore, that herein lies a subtle difference between non-stationary fuzzy sets used in this work and type-2 fuzzy sets. In the non-stationary fuzzy sets used here, the perturbation function acts *horizontally* across the universe of discourse; in type-2 fuzzy sets the secondary membership functions are defined *vertically* along the membership value μ. For non-stationary fuzzy sets featuring 'noise variation', the perturbation function acts *vertically*. Of course, different perturbation functions can still be used and, thus, such non-stationary fuzzy sets might provide a more ‘direct’ comparison with type-2 fuzzy sets.

In section 7.2 and 7.3, two case studies have been carried out (i) to explore output’s interval of *Gaussian* and *Triangular* underlying membership functions in non-stationary fuzzy sets and (ii) to investigate the relationship between underlying membership functions and output uncertainties in interval type-2 and non-stationary fuzzy sets. From these case studies, non-stationary fuzzy sets provide a relatively straight-forward mechanism for carrying out inference with fuzzy sets that are *uncertain* in some way. Clearly, non-stationary systems are not direct equivalents of type-2 systems. However, we believe that non-stationary fuzzy systems can provide a mechanism whereby a form of fuzzy reasoning which *approximates* (in some meaning of the word) general type-2 fuzzy inference in a simple, fast and computationally efficient manner. Of course, the further investigations into the relationship between the two frameworks (non-stationary systems and type-2 systems) in order to explore this approximation of interval and general type-2 inference need to be continued in the future work.

In section 7.4, we set up an experiments to compare the performance amongs type-1, type-2, and non-stationary fuzzy systems. It can be found that, the best overall model is non-stationary fuzzy system with normal distribution function used as perturbation function while the centre variation has been applied with 1000 iterations. Hence, we may tentatively suggest that while type-2 fuzzy systems may not strictly be necessary in order to achieve optimal performance, their benefit may lie more in achieving good performance in a more tractable model.

In the next chapter, the conclusions, contributions of this thesis are presented including limitations and direction of future works.

Chapter 8

Conclusions

It is well accepted that all humans including 'experts', exhibit variation in their decision making. Variation may occur among the decisions of a panel of human experts (inter-expert variability), as well as in the decisions of an individual expert over time (intra-expert variability). Up to now it has been an implicit assumption that expert systems, including fuzzy expert systems, should not exhibit such variation. While type-2 fuzzy sets capture the concept of introducing uncertainty into membership functions by introducing a range of membership values associated with each value of the base variable, but they do not capture the notion of variability — as a type-2 fuzzy inference system (FIS) will always produce the same output(s) (albeit a type-2 fuzzy set with an implicit representation of uncertainty) given the same input(s). Garibaldi et al. [10–14] have been investigating the incorporation of variability into decision making in the context of fuzzy expert systems in a medical domain. In this work, Garibaldi proposed the notion of *non-deterministic fuzzy reasoning* in which variability is introduced into the membership functions of a fuzzy system through the use of random alterations to the parameters of the generating functions.

In this thesis, a notion termed *non-stationary fuzzy set* is introduced and the concept of random perturbations that can be used for generating these non-stationary fuzzy sets is also presented. Basic operators, i.e. union, intersection and complement for non-stationary fuzzy are also proposed. Some of properties of non-stationary fuzzy sets has been proved. The uses of non-stationary fuzzy sets and relationship between non-stationary and type-2 fuzzy sets was investigated through several case studies. In this

Chapter, the contributions and conclusions of this thesis are summarised in the the next Section. The limitations and the direction of possible future works of this thesis are also discussed. Finally, the publications which have been produced from this research are listed.

8.1 Contributions

To reach the goal and fulfill the objectives stated in Chapter 1, this thesis has made the following contributions:

8.1.1 An Investigation of Non-convex Membership Functions

The objective of this research is to illustrate the use of non-convex membership functions for linguistic terms. This research focuses on time-related non-convex membership function. It investigates whether the generated expert systems would work properly or not when time-related non-convex membership functions are used together with normal membership functions. In Chapter 3, Fuzzy expert systems were developed to predict demand load for an energy supply company (simulation data sets). Although there may be many factors that affect demand load, in this research only two, *Time* and prevailing *Temperature* outside, were chosen as the the fuzzy input variables.

In this case study, 500 data sets of *time* and *temperature* were generated randomly, where *time* was between 0 and 24 in the hh:mm format and temperature varied between 0°C and 40°C. The system consisted of two input variables, *Time* and *Temperature*, and one output, *Energy Demand*. In addition to other usual membership functions, the *Time* variable was associated with the term *MealTime* which is a time-related non-convex membership function. In order to observe the influence of *MealTime* on the performance, four systems were created by only changing the term *MealTime* in each system. Four fuzzy expert systems were generated, using four different shapes of *MealTime* membership functions varying in the range of [0,1], [0.2,1], [0,0.9], and [0.2,0.9]. Twelve rules were used within the four generated fuzzy expert systems.

The results of the case study presented in Chapter 3 have shown that non-regular terms can be used in a fuzzy logic system and they can perform together with regular membership functions. The difference in the predicted results of each system is due to the different time-related non-convex *MealTime* term added to the variable *time*. It can be observed that the *energy demand* predictions have incorporated the information introduced by addition of the *MealTime* term. From these illustrations, it can be stated that non-convex membership functions such as *MealTime* featured in the *Time* (of day) variable are plausible, reasonable membership functions in the sense originally intended by Zadeh.

8.1.2 Investigation of Model Parameters in Type-2 Systems

The purpose of this work was to perform a careful analysis of whether the performance of type-2 fuzzy system could be matched or surpassed by type-1 models with a similar or greater number of internal tunable model parameters.

In Chapter 4, the data sets were generated by using Mackey-Glass time-series delay differential equation. Five independent data sets were generated and five different levels of noise were added to each data set. In total 25 data sets were generated with 1,000 time points each; the first 500 points were used for training the systems and the rest for testing the systems. Four antecedents, $x(k-3)$, $x(k-2)$, $x(k-1)$, and $x(k)$, were used to predict $x(k+1)$. Twelve fuzzy systems have been implemented for this experiment:

- Two type-1 singleton systems with 2 membership functions for input variables, featuring 144 tunable parameters.
- Two type-1 singleton systems with 3 membership functions for input variables, featuring 729 tunable parameters.
- Two type-1 non-singleton systems with 2 membership functions for input variables, featuring 145 tunable parameters.
- Two type-1 non-singleton systems with 3 membership functions for input variables, featuring 730 tunable parameters.
- Two type-2 singleton systems with 2 membership functions for input variables, featuring 224 tunable parameters.

- Two type-2 non-singleton systems with 3 membership functions for input variables, featuring 225 tunable parameters.

The performance of all systems was evaluated by using *root mean square error* (RMSE). In all cases, the performance of type-1 fuzzy systems with singleton inputs (the most common found in practice) was worse than for the type-1 non-singleton fuzzy systems and the type-2 fuzzy systems. This is regardless of the number of parameters in the systems. Particularly, it should be noted that type-1 fuzzy systems with singleton inputs where each input has 3 membership functions (featuring 729 tunable parameters), whilst better than type-1 fuzzy systems with singleton inputs where each input has 2 membership functions, achieve far worse performance than type-1 non-singleton or type-2 fuzzy systems with far fewer parameters. This suggests that a high number of model parameters is not in itself sufficient to produce good performance.

The best overall performance was achieved with a type-1 fuzzy system with non-singleton inputs and with 3 membership functions for each variable, leading to a high number of tunable model parameters (730). From this, it can be tentatively suggested that while type-2 fuzzy systems may not strictly be necessary in order to achieve 'optimal' performance, their benefit may lie more in achieving good performance in a more tractable model. So, the conclusion can be made that, by increasing the number of tunable parameters, a type-1 system's performance can be as good as or better than type-2 fuzzy systems.

8.1.3 The Introduction of Non-stationary Fuzzy Sets

In Chapter 5, a new concept termed non-stationary fuzzy set is defined. These have been created with the specific intention of modelling the variation (over time) of opinion, and then formalise the novel concept that previously proposed by Garibaldi [14] to model the variation in expert opinion. While apparently similar to type-2 fuzzy sets in some regards, non-stationary fuzzy sets possess some important distinguishing features. A non-stationary fuzzy set is, effectively, a collection of type-1 fuzzy sets in which there is an explicit, defined, relationship between the fuzzy set. Specifically, each of the instantiations (type-1 fuzzy set) is derived by a perturbation of (making a small change to) a single underlying membership function. While each instantiation is somewhat reminiscent of a embedded type-1 set of a type-2 fuzzy set, there is not a direct correspondence between these two concepts. It is also possible to view a standard type-1 fuzzy set, either as a single instantiation or as repeated instantiations of the underlying set with no perturbation.

In this work, some possible functions that can be used as a perturbation function are suggested. The term *footprint of variation* (FOV) is proposed to represent the area covering the range from the minimum to the maximum fuzzy sets which comprise the non-stationary fuzzy set. Operations on non-stationary fuzzy sets, i.e. *union*, *intersection*, *complement* are also introduced in this Chapter. Finally, proof of some properties of non-stationary fuzzy sets, i.e. *involution*, *commutativity*, *associativity*, *idempotence*, and *distributivity* are presented.

8.1.4 The Use of Non-stationary Fuzzy Sets

The relationships between interval type-2 and non-stationary fuzzy sets were investigated, in terms of secondary membership functions. In Chapter 6, Two case studies were carried out in order to illustrate the use of non-stationary fuzzy sets and to examine the relationship between the performance of non-stationary and interval type-2 fuzzy inference systems. In this study, all fuzzy inference systems were constructed to predict the truth value of the XOR output where both input variables can take any value in the range of [0,1].

All fuzzy inference systems consisted of two input variables, one output variable, and four rules. These rules are constructed based on the standard XOR problem. Each variable consists of two membership functions, corresponding to meaning of the terms *Low* and *High*. The four cases of input values used throughout these studies are: (0.25,0.25), (0.25,0.75), (0.75,0.25), and (0.75,0.75). Three different perturbation functions were used to generate non-stationary fuzzy sets, namely a *uniformly distributed function*, *Gaussian distributed function*, and a *sinusoidal function*.

In the first case study, Gaussian membership functions were used as the underlying membership function in both non-stationary and interval type-1 fuzzy systems. Two forms of variation were implemented such as variation in location and variation in width. In the second case study, Triangular functions were used as the underlying membership function.

Only one form of variation, variation in centre, was implemented.

For type-2 systems, inference was performed using the rules given to obtain the type-2 output sets. The usual Karnik-Mendel type reduction was used to obtain the *lower* and *upper* bound of the centre of gravity of the output. The *mean* of the output was taken as the average of the *lower* and *upper* bound. For non-stationary fuzzy systems, each system was instantiated the specified number of times to obtain the output fuzzy sets. In each case, defuzzification was applied to obtain the standard centre of gravity, g. As a result, a set of centres of gravity, G, was obtained. The minimum of G was taken as the lower bound, the maximum of G as the upper bound, and the arithmetic mean was taken as the *mean*.

From the results of experiments, it is clear that a non-stationary system mimics a type-2 system in that the union of all possible instantiations of non-stationary fuzzy set defines the region over which the underlying membership function varies, termed the *footprint of variation* (FOV). This is similar to the concept of the footprint of uncertainty of a type-2 set. It can be observed that non-stationary fuzzy systems featuring non-uniformly distributed perturbation functions may allow approximations of general type-2 fuzzy inference to be carried out.

8.1.5 Comparison of Outputs of Non-stationary and Interval Type-2 Fuzzy Inference

The aim of the study was to explore relationship between the shape of the underlying membership functions and the uncertainties obtained in the output sets for both non-stationary and interval type-2 fuzzy systems. In Chapter 7, this work was carried out on fuzzy systems implementing the standard XOR problem, in which either Gaussian or Triangular membership functions were employed as the underlying membership functions, using a range of input values and recording the size of the output intervals obtained.

The investigations were carried out onto three different *perturbation functions* (a *uniformly distributed function*, *Gaussian distributed function*, and *sinusoidal function*) and different type of variations: in the case of Gaussian underlying membership functions

- variation in centre or location, and

- variation in width or standard deviation;

in the case of Triangular underlying membership functions

- variation in centre point,
- variation in begin point,
- variation in end point, and
- variation in begin and end point.

Non-stationary fuzzy systems were compared to interval type-2 fuzzy system featuring equivalent *footprints of uncertainty*. All fuzzy systems consisted of two input variables (*Input1* and *Input2*), one output variable (*Output*), and four rules (constructed based on the standard XOR problem). Each variable consisted of 2 Gaussian or Triangular underlying membership functions (termed *Low* and *High*). Once the systems were implemented, they were used to predict the output for four input vectors. Then the *lower*, *mean*, *upper*, and *interval* of the results were computed and recorded, while these values were obtained by using the same method used in Section 8.1.4.

In the case of Gaussian underlying membership functions with width variation, it was found that the output interval was constant for both type-2 system and non-stationary systems, but all the non-stationary systems exhibit larger output interval. In contrast, in the case of Gaussian underlying membership function with centre variation, the type-2 system has the same constant output interval as in the normally distributed non-stationary system, but the latter system has a smaller output interval. In the case of Triangular underlying membership function with begin and end point variation, the output intervals appeared to be non-symmetrical with the inputs. For the case of centre variation, all systems had approximately the same value of interval, which varied according to the input values. Of all systems, the ones with sine perturbation function produced results which are *very close* to the type-2 systems.

From these results, it can be suggested that non-stationary fuzzy systems may produce roughly equivalent results to interval type-2 fuzzy systems. However, non-stationary fuzzy systems may provide a mechanism whereby a form of fuzzy reasoning which approximates general type-2 fuzzy inference in a simple, fast and computationally efficient manner.

8.1.6 Detailed Investigation into Non-stationary Fuzzy Inference

The aim of this study was to investigate the relationship between Gaussian and Triangular underlying membership function used in both non-stationary and interval type-2 fuzzy sets, and the uncertainties obtained in the outputs.

In Chapter 7, in a second case study, the non-stationary and interval type-2 fuzzy systems were implemented exactly the same as described in Section 8.1.5. The work in Section 8.1.5 restricted the range of input values examined; specifically, the pairs (0.25,0.25), (0.25,0.75), (0.75,0.25), and (0.75,0.75). In this work, each input is varied over the range [0,1] in increments of 0.05, giving a total of 441 pairs of *Input1* and *Input2* as [(0,0), (0,0.05),..., (1,0.95), (1,1)]. After all systems had been constructed, they were used to obtain the output of each pair of data sets (in total 441 pairs). The *lower*, *mean*, *upper*, and *interval* of output were computed and recorded. Again, while these values were obtained by using the same method used in Section 8.1.4.

From the experiments, it can be seen that the shape of the surface obtained is similar for each of the different types of fuzzy system in each case. In order to explore this further, the difference between the intervals of the outputs of the interval type-2 and non-stationary fuzzy systems with uniformly distributed perturbation functions were calculated and plotted. It can be observed that the differences for the systems with Gaussian underlying membership functions are quite small, the differences for Triangular underlying membership functions with centre point variation are extremely small, whereas the differences for Triangular underlying membership functions with begin-end point variation are lager and non-symmetrical.

In this work, the *footprints of variation* (FOV) of the non-stationary fuzzy sets was examined by comparing with the *footprints of uncertainty* (FOU) of type-2 fuzzy sets. For non-stationary fuzzy sets which are generated by uniformly distributed and sinusoidal perturbation functions, the maximum area of FOV will be the same as that of the FOU of an interval type-2 fuzzy sets with the same amount of variation. Normally distributed perturbation functions generate random values within $[-\infty, \infty]$, and so an FOV defined as the union of all underlying memberships would fill the entire universe of discourse.

From the experiments, it can be seen that the difference in the interval of output between the type-2 system and non-stationary systems with uniformly distributed perturbation functions is extremely small. Without performing a detailed mathematical analysis of the relationship between type-2 and non-stationary systems, it might be predicted that uniformly distributed non-stationary systems with the same FOV as FOU of interval type-2 systems will match closely.

8.1.7 Comparison on the Performances of Type-1, Type-2, and Non-stationary Fuzzy Systems for Time-Series Prediction

The aim of this study was to investigate and compare the performance of type-1, type-2, and non-stationary fuzzy systems for predicting MacKey-Glass Time-Series.

In Chapter 7, in a third case study, the type-1 and type-2 fuzzy systems were implemented exactly the same as described in Section 8.1.2, while non-stationary fuzzy sytems were initially constucted exactly the same as type-1 fuzzy systems and then using chosen perturbation functions to generate non-stationary fuzzy sets for 1000 iterations. After all systems had been constructed, they were used to obtain the prediction outputs and the performance of all systems was evaluated by using root mean square error (RMSE).

In all cases, the performance of type-1 fuzzy systems with singleton inputs (the most common found in practice) was worse than for the type-1 non-singleton, the type-2, and the non-stationary fuzzy systems. In the mean while, type-1 non-singletion fuzzy system (Type1-NSFLS), type-2 fuzzy systems (both Type-2-SFLS and T1-NS-T2-FLS) were performed as good as each other for all data sets. Anyway, these systems were performed better than non-stationary fuzzy systems (NS-FLS-S and NS-FLS-U).

The interesting finding from the experiments, we found that the best overall performance is achieved with NS-FLS-N. This is a non-stationary fuzzy system uisng normal distribution function as a pertubation function. From this, we may tentatively suggest that while type-2 fuzzy systems may not strictly be necessary in order to achieve optimal performance, their benefit may lie more in achieving good performance in a more tractable model.

8.1.8 Summary

In this thesis, the use of non-regular terms and non-convex membership function has been investigated. It is shown that non-convex membership functions can be used in a fuzzy system and they can perform together with regular membership functions. It is suggested that non-convex membership functions are plausible, reasonable membership functions in the sense originally intended by Zadeh. As the limitations of type-2 fuzzy system is complexity and high cost of computational time, an investigation was carried out to perform an analysis of whether the performance of type-2 fuzzy system could be matched or surpassed by type-1 models with a similar o greater number of internal tunable model parameters. From this investigation, the conclusion can be made that increasing the number of tunable parameters can improve a type-1 non-singleton fuzzy system to be as good as or better than type-2 fuzzy systems.

This thesis proposes the theoretical of the new type of fuzzy sets, termed *non-stationary fuzzy sets*, and clearly defines the mathematical formulae that represent non-stationary fuzzy sets. Perturbation functions used to altered the non-stationary fuzzy sets, operations on non-stationary fuzzy sets, and proof of some properties of non-stationary fuzzy sets have been presented. The term that *footprint of variation* (FOV) has been proposed to represent the region covering the range from the minimum to the maximum fuzzy sets which comprise the non-stationary fuzzy sets. Non-stationary fuzzy sets possess some important distinguishing features:

- A non-stationary fuzzy set is a collection of type-1 fuzzy sets in which there is an explicit, defined, relationship between the fuzzy sets. Specifically, each of the instantiations (individual type-1 sets) is derived by a perturbation of (making a small change to) a single underlying membership function.

- A non-stationary fuzzy set does not have secondary membership functions, or secondary membership grades. Hence, there is no 'direct' equivalent to the embedded type-2 sets of a type-2 fuzzy sets.

- The non-stationary inference process is quite different from type-2 FIS. That is at any instant of time, a non-stationary fuzzy set is (instantiates) a type-1 fuzzy set.

Therefore, the non-stationary inference is just a repeated type-1 inference (albeit with slightly different type-1 sets at each time instant).

Experiments have been carried out to investigate the use of non-stationary fuzzy sets through several case studies. The results from all of these experiments are compared with results produced by type-2 fuzzy systems. From the results of these experiments, conclusions can be drawn. Specifically:

- A non-stationary system mimics a type-2 system in that the union of all possible instantiations of non-stationary fuzzy set defines the region over which the underlying membership function varies, termed the *footprint of variation* (FOV). This is similar to the concept of the footprint of uncertainty of a type-2 set.

- Non-stationary fuzzy systems featuring non-uniformly distributed perturbation functions may allow approximations of general type-2 fuzzy inference to be performed.

- Non-stationary system with sinusoidal perturbation functions may produce equivalent results to interval type-2 fuzzy system.

- Uniformly distributed non-stationary systems with the same FOV as the FOU of an interval type-2 system will match closely.

- Non-stationary fuzzy systems may provide a mechanism whereby a form of fuzzy reasoning approximates general type-2 fuzzy inference in a simple, fast and computationally efficient manner.

8.2 Limitations and Future Directions

In this section, in order to point out the limitations of this thesis and the possibly direction for future work, four sub-sections are clearly catagorised as follows.

8.2.1 Operations on Non-stationary Fuzzy Sets

In Chapter 5, operations on non-stationary fuzzy sets, *union*, *intersection*, and *complement*, are only defined with *maximum t-conorm*, *minimum t-norm*, and *standard complement*, respectively. Algebraic product is another popular t-norm, known as *product* *t-norm*. This may be very useful for some application domains, especially in engineering applications of fuzzy sets and logic i.e., fuzzy controller context. Work needs to be done to formalise the concept of product t-norm for non-stationary fuzzy sets in terms of mathematical formulae as follows. From the following equation

$$\mu_{\dot{A}\cap\dot{B}}(x,t) = \mu_{\dot{A}}(x,t) \otimes \mu_{\dot{B}}(x,t),\ \forall(x,t) \in X \times T. \tag{8.1}$$

where $\otimes$ denotes the product t-norm, this becomes

$$\mu_{\dot{A}\cap\dot{B}}(x,t) = [\mu_{\dot{A}}(x,t) \times \mu_{\dot{B}}(x,t)],\ \forall(x,t) \in X \times T. \tag{8.2}$$

A *product t-norm operator* on non-stationary fuzzy sets (in Equation 8.2) must be proved to satisfy the set theoretic laws such as *involution*, *commutativity*, *associativity*, *idempotence*, and *distributivity*. This can be continued in future work.

Some others operators (*and* and *or* connectives from Lukasiewicz and Probabilistic logic) as shown in Chapter 2 (Table 2.5) may be needed to extend and formalise into a collection of non-stationary fuzzy sets operators that can be used in the non-stationary fuzzy systems. These ideas can also be continued in the future.

8.2.2 Relationship between Non-stationary and Type-2 Fuzzy Sets

The class of a type-2 fuzzy set is determined by the secondary membership function. That is, if the secondary membership function simply takes the value zero outside the lower and upper bounds and one inside the bounds, then interval type-2 fuzzy sets are obtained. If (type-1) fuzzy sets are used for the secondary membership functions, then general type-2 fuzzy sets are obtained. In comparison, the class of a non-stationary fuzzy set is determined both by which kind of non-stationarity used (variation in location, variation in slope or noise variation) and by the form of perturbation function used to deviate the underlying membership function.

In experiments carried out in this thesis, Normally distributed, uniformly distributed, and sinusoidal perturbation functions were applied to variation in location and variation in slope but not in noise variation. It should be noted, therefore, that herein lies a subtle difference between non-stationary fuzzy sets used in this work and type-2 fuzzy sets. In the non-stationary fuzzy sets used here, the perturbation function acts *horizontally* across the universe of discourse; in type-2 fuzzy sets the secondary membership functions are defined *vertically* along the membership value μ. For non-stationary fuzzy sets featuring 'noise variation', the perturbation function acts *vertically*. Of course, different perturbation functions can still be used and, thus, such non-stationary fuzzy sets might provide a more **direct** comparison with type-2 fuzzy sets. According to the limitation of time, further investigation in these contexts must be left for the future.

8.2.3 Perturbation Functions

For non-stationary fuzzy sets which are generated by uniformly distributed and sinusoidal perturbation functions (producing random values within $[-1,1]$), the maximum area of *footprint of variation* (FOV) will be the same size as the *footprint of uncertainty* (FOU) of interval type-2 fuzzy sets with the same amount of variation. In contrast, Normally distributed perturbation functions generate random values within $[-\infty,\infty]$, and so an FOV defined as the union of all underlying memberships would fill the entire universe of discourse. This kind of FOV will need further investigation. Also, other random functions that can be used as perturbation functions to generate non-stationary fuzzy sets must be further investigated. For example

- a differential time-series, such as the Mackey-Glass equation:

$$\frac{df(t)}{dt} = \frac{0.2f*(t-\tau)}{1+f^{10}(t-\tau)} - 0.1f(t), \tag{8.3}$$

  where τ is a constant.

8.2.4 Non-stationary Fuzzy Sets in Real World Applications

In this thesis, all experiments carried out only used simulation data sets and hence no general conclusions can be definitely claimed. In order to reach general conclusions it would be necessary to carry out similar experiments on a wide variety of real world data sets. Again, to reach general conclusions, the potential directions for future research of non-stationary fuzzy sets in real world applications could be suggested (but not limited to):

- **Medical applications**: As mentioned earlier, all humans including 'experts', exhibit variation in their decision making. Variation may occur among the decisions of a panel of human experts (inter-expert variability), as well as in the decisions of an individual expert over time (intra-expert variability). Up to now it has been an implicit assumption that expert systems, including fuzzy expert systems, should not exhibit such variation. Specifically in medical context, the experts (doctors) may ask to provide their diagnosis on any cases. Each expert sometimes might provide a difference decisions depending on situation and their experiences. It might be tentatively suggested to employ non-stationary fuzzy sets in expert system to model the variation in expert opinions (there may be a need to compare with type-1 and type-2 fuzzy systems). These ideas are actually the focus of ongoing work.

- **Time-series forecasting applications**: In Chapter 4, an investigation was carried out in which the performance of type-1 and type-2 fuzzy systems with varying number of tunable parameters were compared in their ability to predict the Mackey-Glass time series with various levels of added noise. Each of the fuzzy systems were tuned to achieve the best possible performance using a standardised gradient descent procedure. The results show that the best performance was achieved with a type-1 fuzzy system, albeit featuring a high number of tunable parameters. A type-2 fuzzy system with far fewer parameters achieved performance very close to the best. As mentioned in Chapter 5, a non-stationary fuzzy inference system (FIS) is simply a repetition of a type-1 FIS with slightly different instantiations of the membership function over time. Hence, non-stationary inference does not suffer the difficulties of type-2 inference (particularly the inference using general type-2 fuzzy sets). It may be expected that non-stationary FIS with fewer tunable parameters can perform as well as or better than type-1 and type-2 systems. To determine this, further investigation need to be done in the context of time-series forecasting applications, e.g. financial forecasting.

- **Control applications**: Non-stationary fuzzy sets have been investigated by Coupland and John [138] to compare the performance of type-1, interval type-2 using type-reduction, and non-stationary fuzzy systems in the context of control applications. An interesting finding in their experiment was the performance of the non-stationary fuzzy system. They observed that under the minimum t-norm, the non-stationary system gave a much smoother surface than any of the other systems, meaning that this system has a smoother control performance. Under the product t-norm, the non-stationary system was as good as any of the other systems. In this work, they conclude that the non-stationary system outperforms the other fuzzy technologies.

Although, standard type-1 fuzzy sets have been greatly employed in many control applications, it is not yet possible to model and minimize the effect of all uncertainties (i.e., uncertainty in membership functions, and noisy data sets). To overcome this limitation, type-2 fuzzy sets can be introduced into the systems. However, the use of type-2 fuzzy sets in practice has been limited due to the significant increase in computational complexity involved in their implementation. From this preliminary work, it is strongly believed that non-stationary fuzzy sets would be useful to apply into the control systems.

8.3 Publications Produced

The research described in this thesis has been published through a number of journal and in international conference papers. Most of the work presented in each of the main body Chapters of this thesis has also been published in this manner. A formal list of publications and presentations derived from this work follows.

8.3.1 Journal Publications

1. Garibaldi, J.M., Jaroszewski, M., Musikasuwan, S., Nonstationary Fuzzy Sets, IEEE Transaction on Fuzzy System, V. 16, n. 4, August, 2008. (Chapter 5 and 6)

8.3.2 Conference Publications

1. Ozen, T., Garibaldi, J.M., Musikasuwan, S., Modelling the Variation in Human Decision Making, Proceeding of Fuzzy Sets in the Heart of Canadian Rockies (NAFIPS 2004), Banff, Alberta, Canada, 27-30 June 2004.

2. Musikasuwan, S., Ozen, T., Garibaldi, J.M., An Investigation into the Effect of Number of Model Parameters on Performance in Type-1 and Type-2 Fuzzy Logic Systems, in Proceedings of Information Processing and Management of Uncertainty in Knowledge Based Systems (IPMU 2004), Perugia, Italy, July 4-9, 2004. (Chapter 4)

3. Ozen, T., Garibaldi, J.M., Musikasuwan, S., Preliminary Investigations into Modelling the Variation in Human Decision Making, in Proceedings of Information Processing and Management of Uncertainty in Knowledge Based Systems (IPMU 2004), Perugia, Italy, July 4-9, 2004.

4. Garibaldi, J.M., Musikasuwan, S., Ozen, T., John. R.I., A Case Study to Illustrate the Use of Non-Convex Membership Functions for Linguistic Terms, in Proceedings of IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2004), Budapest, Hungary, July 25-29, 2004. (Chapter 3)

5. Garibaldi, J.M., Musikasuwan, S., Ozen, T., The Association between Non-Stationary and Interval Type-2 Fuzzy Sets: A Case Study, in Proceeding of IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2005), Reno, Nevada, USA, May 22-25, 2005. (Chapter 6)

6. Musikasuwan, S., Garibaldi, J.M., Exploring Gaussian and Triangular Primary Membership Functions in Non-Stationary Fuzzy Sets, in Proceeding of Information Processing and Management of Uncertainty in Knowledge Based Systems (IPMU 2006), Paris, France, July 2-7, 2006. (Chapter 7)

7. Musikasuwan, S., Garibaldi, J.M., On Relationships between Primary Membership Functions and Output Uncertainties in Interval Type-2 and Non-Stationary Fuzzy Sets, in Proceeding of IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2006), Vancouver, Canada, July 16-21, 2006. (Chapter 7)

8.3.3 Presentations and conferences attended

- Fuzzy Expert Systems, Automated Scheduling, Optimisation, and Planning (ASAP) Research Group seminar, School of Computer Science and IT, The University of Nottingham, United Kingdom, April 9, 2003.

- Effect of Number of Model Parameters on Performance in Type-1 and Type-2 Fuzzy Logic Systems, International Conference of IPMU 2004, Perugia, Italy, July 8, 2004.

- Modelling the Variation in Human Decision Making, International Conference of NAFIPS 2004, Banff, Alberta, Canada, June 29, 2004.

- The Use of Non-Convex Membership Functions for Linguistic Terms, IEEE International Conference of FUZZ-IEEE 2004, Budapest, Hungary, July 28, 2004.

- The Association between Non-Stationary and Interval Type-2 Fuzzy Sets, IEEE International Conference of FUZZ-IEEE 2005, Reno, Nevada, USA, May 22-25, 2005.

- Exploring Gaussian and Triangular Primary Membership Functions in Non-Stationary Fuzzy Sets, Information Processing and Management of Uncertainty in Knowledge Based Systems (IPMU 2006), Paris, France, July 2-7, 2006.

- On Relationships between Primary Membership Functions and Output Uncertainties in Interval Type-2 and Non-Stationary Fuzzy Sets, IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2006), Vancouver, Canada, July 16-21, 2006.

Bibliography

[1] J. Mendel, *Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions*. Prentice Hall PTR, 2001.

[2] L. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, pp. 338–353, 1965.

[3] G. Klir and T. Folger, *Fuzzy Sets, Uncertaintly, and Information*. London: Prentice-Hall Internaltion Editions, 1988.

[4] L. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning - I,II,III," *Information Sciences*, vol. 8;8;9, pp. 199–249;301–357;43–80, 1975.

[5] D. Dubois and H. Prade, *Fuzzy Sets and Systems: Theory and Applications*. New York: Academic, 1980.

[6] R. Yager, "Fuzzy subsets of type ii in decisions," *Journal of Cybernetics*, vol. 10, pp. 137 – 159, 1980.

[7] M. Mizumoto and K. Tanaka, "Some properties of fuzzy sets of type-2," *Information and Control*, vol. 31, pp. 312 – 340, 1976.

[8] ——, "Fuzzy sets of type-2 under algebraic product and algebraic sum," *Fuzzy Sets and Systems*, vol. 5, pp. 277 – 290, 1981.

[9] J. Mendel and R. John, "Type-2 fuzzy sets made simple," *IEEE Transactions on Fuzzy Systems*, vol. 10, no. 2, pp. 117 – 127, 2002.

[10] T. Ozen and J. M. Garibaldi, "Investigating adaptation in type-2 fuzzy logic systems applied to umbilical acid-base assessment," in *European Symposium on Intel-*

ligent Technologies, Hybrid Systems and Their Implementation on Smart Adaptive Systems, Oulu, Finland, June 2003.

[11] T. Ozen, J. M. Garibaldi, and S. Musikasuwan, "Preliminary investigations into modelling the variation in human decision making," in *Information Processing and Management of Uncertainty in Knowledge Based Systems*, Perugia, Italy, July 2004.

[12] T. Ozen and J. M. Garibaldi, "Effect of type-2 fuzzy membership function shape on modelling variation in human decision making," in *IEEE International Conference on Fuzzy Systems*, Budapest, Hungary, July 2004.

[13] T. Ozen, J. M. Garibaldi, and S. Musikasuwan, "Modelling the variation in human decision making," in *Fuzzy Sets in the Heart of Canadian Rockies, NAFIPS-04*, Banff, Alberta, Canada, June 2004.

[14] J. Garibaldi and T. Ozen, "Uncertain fuzzy reasoning: A case study in modelling expert decision making," *IEEE Transactions on Fuzzy Systems*, vol. 15, no. 1, pp. 16–30, 2007.

[15] D. Goldrei, *Classic set Theory: a guided independent study*. London: Chapman & Hall, 1996.

[16] J. Lukasiewicz, *Philosophical remarks on many-valued systems of propositional logic*. Oxford: Oxford University Press: Reprinted (1967) in Sorrs McCall (Ed.), 1930.

[17] J. Bell, *Boolean-valued models and independence proofs in set theory*. Oxford: Clarendon Press, 1977.

[18] R. Yager, "On a general class of fuzzy connectives," *Fuzzy Sets and Systems*, vol. 4, pp. 235–242, 1980.

[19] M. Sugeno, *Fuzzy measures and fuzzy integrals: a survey*. Oxford: Clarendon Press, 1977.

[20] B. Kosko, *Fuzzy Thinking*. London: Harper Collins Publishers, 1994.

[21] H. Zimmermann, *Fuzzy Set Theory and Its Applications - 3/ed.* Kluwer Academic Publishers, 1996.

[22] G. Lakoff, "Hedges: a study in meaning criteria and the logic of fuzzy concepts," *Journal of Philosophical Logic*, vol. 2, pp. 458–508, 1973.

[23] L. Zadeh, "Outline of a new approach to the analysis of complex systems and decision processes," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 3, pp. 28–44, 1973.

[24] D. Dubois and H. Prade, "Fuzzy sets in approximate reasoning, part 1: Inference with possibility distributions," *Fuzzy Sets and Systems*, vol. 40, pp. 143–202, 1991.

[25] E. Mamdani and S. Assilian, "An experiment in linguistic synthesis with a fuzzy logic controller," *International Journal of Man-Machine Studies*, vol. 7, pp. 1–13, 1975.

[26] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 15, no. 1, pp. 116–132, 1985.

[27] Z. Li, "Suitability of fuzzy reasoning methods," *Fuzzy Sets and Systems*, vol. 108, pp. 299–311, 1999.

[28] B. Turksen, "Type i and type ii fuzzy system modeling," *Fuzzy Sets and Systems*, vol. 106, pp. 11–34, 1999.

[29] L. Zadeh, "Fuzzy logic, neural networks, and soft computing," *Communication of the ACM*, vol. 37, no. 3, pp. 77–84, 1994.

[30] J. Castro, E. Trillas, and J. Zurita, "Non-monotonic fuzzy reasoning," *Fuzzy Sets and Systems*, vol. 94, pp. 217–225, 1998.

[31] E. Ruspini, P. Bonissone, and W. Pedrycz, *Handbook of Fuzzy Computation.* Institute of Physics Publishing, Bristol and Philadelphia, 1998.

[32] R. Jager, "Fuzzy logic in control," Ph.D. dissertation, Unversiteit Delft, 1995.

[33] J. Baldwin, *Fuzzy logic and fuzzy reasoning, in Fuzzy Reasoning and Its Applications, E.H. Mamdani and B.R. Gaines (eds.).* London Academic Press, 1981.

[34] L. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning - I," *Information Sciences*, vol. 8, pp. 199–249, 1975.

[35] Q. Liang and J. Mendel, "Interval type-2 fuzzy logic systems: Theory and design," *IEEE Transactions Fuzzy Systems*, vol. 8, pp. 535 – 550, 2000.

[36] N. Karnik and J. Mendel, "Operations on type-2 fuzzy sets," *Fuzzy Sets and Systems*, vol. 122, pp. 327–348, 2001.

[37] ——, "Type-2 fuzzy logic systems: Type-reduction," in *Proceedings IEEE Systems, Man and Cybernetics*, 1998, pp. 2046–2051.

[38] N. K. J. Mendel, "Centroid of a type-2 fuzzy set," *Information Science*, vol. 132, pp. 195 – 220, 2001.

[39] N. Karnik and J. Mendel, "Introduction to type-2 fuzzy logic systems," in *Proceedings of 7th International Conference on Fuzzy Systems FUZZ-IEEE'98*, 1998, pp. 915–920.

[40] R. John, "Type 2 fuzzy sets for knowledge representation and inferencing," *Proceedings of 7th International Conference on Fuzzy Systems FUZZ-IEEE'98*, pp. 1003–1008, 1998.

[41] ——, "Embedded interval valued fuzzy sets," *Proceedings of FUZZ-IEEE 2002*, pp. 1316–1321, 2002.

[42] R. John and C. Czarnecki, "A type 2 adaptive fuzzy inferencing system," in *Proceedings IEEE Systems, Man and Cybernetics*, 1998, pp. 2068–2073.

[43] J. Mendel, "Uncertainty, fuzzy logic and signal processing," *Signal Processing*, vol. 80, pp. 913–933, 2000.

[44] N. Karnik, J. Mendel, and Q. Liang, "Type-2 fuzzy logic systems," *IEEE Transactions on Fuzzy Systems*, vol. 7, no. 6, pp. 643–658, 1999.

[45] L. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning - II," *Information Sciences*, vol. 8, pp. 301–357, 1975.

[46] ——, "The concept of a linguistic variable and its application to approximate reasoning - III," *Information Sciences*, vol. 9, pp. 43–80, 1975.

[47] I. Turksen, "Representation and inference: Review and challenges," in *Fifth IFSA World Congress*, 1993, pp. 175 – 178.

[48] ——, "Interval-valued fuzzy uncertainty," in *Fifth IFSA World Congress*, 1993, pp. 35 – 38.

[49] ——, "Fuzzy normal forms," *Fuzzy Sets and Systems*, vol. 69, pp. 319 – 346, 1995.

[50] D. Schwartz, "The case for an interval-based representation of linguistic truth," *Fuzzy Sets Systems*, vol. 17, pp. 153 – 165, 1985.

[51] L. Zadeh, "Fuzzy logic = computing with words," *IEEE Transactions on Fuzzy Systems*, vol. 4, no. 2, pp. 103–111, 1996.

[52] ——, "From computing with numbers to computing with words — from manipulation of measurements to manipulation of perceptions," *IEEE Transactions on Circuits and Systems*, vol. 45, no. 1, pp. 105–119, 1999.

[53] J. M. Mendel, "Computing with words, when words can mean different things to different people," in *Proceedings of Third International ICSC Congress on Computational Intelligence Methods and Applications*, 1999.

[54] J. Mendel, "The perceptual computer: an architecture for computing with words," in *FUZZ-IEEE*, St. Louis, MO, USA, 2003, pp. 37 – 42.

[55] I. Turksen, "Representation and reasoning for cww," *Fuzzy Sets and Systems*, vol. 127, pp. 17–36, 2002.

[56] N. Karnik and J. Mendel, "An introduction to type-2 fuzzy logic systems," University of Southern California, Tech. Rep., 1998.

[57] R. John, “Type 2 fuzzy sets for knowledge representation and inferencing,” School of Computing Sciences, De Montfort University, Research Monograph 10, 1998.

[58] ——, “Type 2 fuzzy sets: An appraisal of theory and applications,” *International Journal of Uncertainty, Fuzziness and Knowledge Based Systems*, vol. 6, no. 6, pp. 563–576, 1998.

[59] R. John and C. Czarnecki, “An adaptive type-2 fuzzy system for learning linguistic membership grades,” in *Proceedings of 8th International Conference on Fuzzy Systems FUZZ-IEEE’99*, 1999, pp. 1552–1556.

[60] S. Coupland and R. John, “An approach to type-2 fuzzy arithmetic,” in *UK Workshop on Computational Intelligence*, Bristol, UK, September 2003, pp. 107 – 114.

[61] H. Wu and J. Mendel, “Introduction to uncertainty bounds and their use in the design of interval type-2 fuzzy logic systems,” in *Proceedings of FUZZ-IEEE 2001*, Melbourne, Australia, 2001.

[62] ——, “Uncertainty bounds and their use in the design of interval type-2 fuzzy logic systems,” *IEEE Transactions on Fuzzy Systems*, pp. 622–639, 2002.

[63] S. Coupland and R. John, “Fuzzy logic and computational geometry,” in *RASC 2004*, Nottingham, UK, December 2004, pp. 3 – 8.

[64] ——, “Towards more efficient type-2 fuzzy logic system,” in *FUZZ-IEEE 2005*, Reno, NV, USA, May 2005, pp. 236 – 241.

[65] ——, “Geometric type-1 and type-2 fuzzy logic systems,” *IEEE Transactions on Fuzzy Systems*, 2006.

[66] S. Greenfield, R. John, and S. Coupland, “A novel sampling method for type-2 defuzzification,” in *Proc. UKCI 2005*, London, September 2005, pp. 120 – 127.

[67] E. Hisdal, “The IF THEN ELSE statement and interval-values fuzzy sets of higher type,” *International Journal of Man-Machine Studies*, vol. 15, pp. 385 – 455, 1981.

[68] S. Coupland and R. John., “A new and efficient method for the type-2 meet operation,” in *Proc. FUZZ-IEEE 2004*, Budapest, Hugary, July 2004, pp. 959 – 964.

[69] ——, “Geometric interval type-2 fuzzy systems,” in *Proceedings of the Joint EUSFLAT-LFA (IV EUSFLAT and XI LFA)*, Barcelona, Spain, September 2005, pp. 449 – 454.

[70] S. Coupland, “Type-2 fuzzy sets: Geometric defuzzification and type-reduction,” in *Proc. FOCI 2007*, Honolulu, Hawaii, USA., April 2007, pp. 622 – 629.

[71] S. Coupland and R. John, “On the accuracy of type-2 fuzzy sets,” in *Proc. FUZZ-IEEE 2007*, London, UK, July 2007, pp. 131 – 136.

[72] S. Coupland and R. I. John, “Geometric type-1 and type-2 fuzzy logic systems,” *IEEE Transactions on Fuzzy Systems*, vol. 15, no. 1, pp. 3 – 15, February 2007.

[73] S. Greenfield, F. Chiclana, R. John, and S. Coupland, “The collapsing method of defuzzification for discretised interval type-2 fuzzy sets,” in *Proc. UKCI 2007*, Imperial College, London, UK, August 2007.

[74] I. Turksen, “Interval valued fuzzy sets based on normal forms,” *Fuzzy Sets Systems*, vol. 20, pp. 191 – 210, 1986.

[75] M. Gorzalczany, “A method of inference in approximate reasoning based on interval-valued fuzzy sets,” *Fuzzy Sets Systems*, vol. 21, pp. 1 – 17, 1987.

[76] R. Lea, V. Kreinovich, and R. Trejo, “Optimal interval enclosures for fractionally-linear functions and their application to intelligent control,” *Reliable Computing*, vol. 2, pp. 265 – 285, 1996.

[77] H. Nguyen, V. Kreinovich, and Q. Zuo, “Interval-valued degrees of belief: Applications of interval computations to expert systems and intelligent control,” *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 5, pp. 317 – 358, 1997.

[78] S. Mabuchi, “An interpretation of membership functions and the properties of general probabilistic operators as fuzzy set operators, ii: Extension to three-valued and interval-valued fuzzy sets,” *Fuzzy Sets Systems*, vol. 92, pp. 31 – 50, 1997.

[79] V. Kreinovich, A. Lakeyev, J. Rohn, and P. Kahl, *Computational Complexity and Feasibility of Data Processing and Interval Computations*. MA, Boston: Kluwer, 1998, p. ch 10.

[80] P. Melin and O. Castillo, “Intelligent control of non-linear dymamic plants using type-2 fuzzy logic and neural networks,” Budapest,Hungary, July 2004.

[81] ——, “A new method for adaptive control of non-linear plants using type-2 fuzzy logic and neural networks,” *International Journal of General Systems*, vol. 33, no. 2-3, pp. 289 – 304, 2004.

[82] O. Castillo, G. Huesca, and F. Valdez, “Evolutionary computing for optimizing type-2 fuzzy systems in intelligent control of non-linear dynamic plants,” Detroit, MI, United States, June 2005, pp. 247 – 251.

[83] H. Hagras, “Type-2 flcs: A new generation of fuzzy controllers,” *IEEE Computational Intelligence Magazine*, vol. 2, no. 1, pp. 30 – 43, February 2007.

[84] C. Lynch, H. Hagras, and V. Callaghan, “Embedded type-2 flc for real-time speed control of marine and traction diesel engines,” Reno, NV, USA, May 2005, pp. 347 – 352.

[85] ——, “Using uncertainty bounds in the design of an embedded real-time type-2 neuro-fuzzy speed controller for marine diesel engines,” Vancouver, BC, Canada, July 2006, pp. 1446 – 1453.

[86] H. Hagras, F. Doctor, A. Lopez, and V. Callaghan, “Evolving type-2 fuzzy agents for ambient intelligent environments,” Lake District, United Kingdom, September 2006, pp. 208 – 213.

[87] H. Hagras, F. Doctor, and A. Lopez, “An incremental adaptive life long learning approach for type-2 fuzzy embedded agents in ambient intelligent environments,” *IEEE Transactions on Fuzzy Systems*, vol. 15, no. 1, pp. 45 – 55, February 2007.

[88] F. Doctor, H. Hagras, , and V. Callaghan, “A type-2 fuzzy embedded agent to realise ambient intelligence in ubiquitous computing environments,” *Journal of Information Sciences*, vol. 171, no. 4, pp. 309 – 334, May 2005.

[89] D. Wu and W. W. Tan, "A type-2 fuzzy logic controller for the liquid-level process," vol. 2, Budapest, Hungary, July 2004, pp. 953 – 958.

[90] ——, "A simplified type-2 fuzzy logic controller for real-time control," *ISA Transactions*, vol. 45, no. 4, pp. 503 – 516, 2006.

[91] M. A. Melgarejo R. and C. A. Pena-Reyes, "Hardware architecture and fpga implementation of a type-2 fuzzy system," Boston, MA, United States, April 2004, pp. 458 – 461.

[92] P.-Z. Lin, C.-F. Hsu, and T.-T. Lee, "Type-2 fuzzy logic controller design for buck dc-dc converters," Reno, NV, United States, May 2005, pp. 365 – 370.

[93] P.-Z. Lin, C.-M. Lin, C.-F. Hsu, and T.-T. Lee, "Type-2 fuzzy controller design using a sliding-mode approach for application to dc-dc converters," *IEE Proceedings: Electric Power Applications*, vol. 152, no. 6, pp. 1482 – 1488, November 2005.

[94] O. Uncu, K. Kilic, and I. Turksen, "A new fuzzy inference approach based on mamdani inference using discrete type 2 fuzzy sets," vol. 3, The Hague, Netherlands, 2004, pp. 2272 – 2277.

[95] P. Baguley, T. Page, V. Koliza, and P. Maropoulos, "Time to market prediction using type-2 fuzzy sets," *Manufacturing Technology Management*, vol. 17, no. 4, pp. 513 – 520, 2006.

[96] D.-W. Kim and G.-T. Park, "Using interval singleton type 2 fuzzy logic system in corrupted time series modelling," vol. 3684 NAI, Melbourne, Australia, 2005, pp. 566 – 572.

[97] K. Huarng and H.-K. Yu, "A type 2 fuzzy time series model for stock index forecasting," *Physica A: Statistical Mechanics and its Applications*, vol. 353, no. 1-4, pp. 445 – 462, 2005.

[98] M. D. L. A. H. Medina and G. M. Mendez, "Modelling and prediction of the mxnusd exchange rate using interval singleton type-2 fuzzy logic systems," Vancouver, BC, Canada, 2006, pp. 2305 – 2308.

[99] A. Mencattini, M. Salmeri, S. Bertazzoni, R. Lojacono, E. Pasero, and W. Moniaci, "Short term local meteorological forecasting using type-2 fuzzy systems," vol. 3931 NCS, Vietri sul Mare, Italy, 2006.

[100] ——, "Local meteorological forecasting by type-2 fuzzy systems time series prediction," Giardini, Naxos, Italy, 2005, pp. 75 – 80.

[101] L. Li, W.-H. Lin, and H. Liu, "Type-2 fuzzy logic approach for short-term traffic forecasting," *IEE Proceedings: Intelligent Transport Systems*, vol. 153, no. 1, pp. 33 – 40, 2006.

[102] N. Karnik and J. Mendel, "Application of type-2 fuzzy logic system to forecasting of time-series," *Information Science*, vol. 120, pp. 89–111, 1999.

[103] S. Musikasuwan, T. Ozen, and J. M. Garibaldi, "An investigation into the effect of number of model parameters on performance in type-1 and type-2 fuzzy logic systems," in *Information Processing and Management of Uncertainty in Knowledge Based Systems*, Perugia, Italy, July 2004.

[104] Q. Liang and L. Wang, "Sensed signal strength forecasting for wireless sensors using interval type-2 fuzzy logic system," Reno, NV, United States, 2005, pp. 25 – 30.

[105] U. Pareek and I. Kar, "Estimating compressor discharge pressure of gas turbine power plant using type-2 fuzzy logic systems," Vancouver, BC, Canada, 2006, pp. 649 – 654.

[106] G. M. Mendez, A. Cavazos, R. Soto, and L. Leduc, "Entry temperature prediction of a hot strip mill by a hybrid learning type-2 fls," *Journal of Intelligent and Fuzzy Systems*, vol. 17, no. 6, pp. 583 – 596, 2006.

[107] R. John, P. R. Innocent, and M. Barnes, "Type 2 fuzzy sets and neuro-fuzzy clustering of radiographic tibia images," *Proceedings of Third Joint Conference on Information Science*, vol. 1, pp. 58–61, 1997.

[108] R. John, P. Innocent, and M. Barnes, "Type 2 fuzzy sets and neuro-fuzzy clustering of radiographic tibia images," *Proceedings of 7th International Conference on Fuzzy Systems FUZZ-IEEE'98*, pp. 1373–1376, 1998.

[109] ——, "Neuro-fuzzy clustering of radiographic tibia image data using type-2 fuzzy sets," *Information Sciences*, vol. 125, no. 1-4, pp. 65–82, 2000.

[110] P. Innocent and R. John, "Type-2 fuzzy medical diagnosis," *Proceedings of FUZZ-IEEE 2002*, pp. 326–1331, 2002.

[111] ——, "Computer aided fuzzy medical diagnosis," *Information Sciences*, vol. 162, no. 2, pp. 81 – 104, 2004.

[112] P. Innocent, R. John, and J. Garibaldi, "Fuzzy methods in medical diagnosis," *Applied Artificial Intelligence*, vol. 19, pp. 69–98, 2005.

[113] L. Di Lascio, A. Gisolfi, and A. Nappi, "Medical differential diagnosis through type-2 fuzzy sets," Reno, NV, United States, 2005, pp. 371 – 376.

[114] P. Herman, G. Prasad, and T. McGinnity, "Investigation of the type-2 fuzzy logic approach to classification in an eeg-based brain-computer interface," vol. 7, Shanghai, China, 2005, pp. 5354 – 5357.

[115] P. Phokharatkul and S. Phaiboon, "Mobile robot control using type-2 fuzzy logic system," vol. 1, Singapore, Singapore, 2004, pp. 296 – 299.

[116] H. Hagras, "A hierarchical type-2 fuzzy logic control architecture for autonomous mobile robots," *IEEE Transactions on Fuzzy Systems*, vol. 12, no. 4, pp. 524 – 539, 2004.

[117] ——, "A type-2 fuzzy logic controller for autonomous mobile robots," vol. 2, Budapest, Hungary, 2004, pp. 965 – 970.

[118] S. Coupland, M. Gongora, R. I. John, and K. Wills, "A comparative study of fuzzy logic controllers for autonomous robots," in *Proc. IPMU 2006*, Paris, France, 2006, pp. 1332 – 1339.

[119] J. Figueroa, J. Posada, J. Soriano, M. Mclgarejo, and S. Rojas, “A type-2 fuzzy logic controller for tracking mobile objects in the context of robotic soccer games,” Reno, NV, United States, 2005, pp. 359 – 364.

[120] K. C. Wu, “Fuzzy interval control of mobile robots,” *Computers & Electrical Engineering*, vol. 22, no. 3, pp. 211 – 229, 1996.

[121] M. Tang, Y. Zhang, and G. Zhang, “Type-2 fuzzy web shopping agents,” Beijing, China, 2004, pp. 499 – 503.

[122] L. Gu and Y. Zhang, “Web shopping expert using new interval type-2 fuzzy reasoning,” *Soft Computing*, vol. 11, no. 8, pp. 741 – 751, 2007.

[123] S. Lee and K. Lee, “A ranking method for type-2 fuzzy values,” *Journal of Korea Fuzzy and Intelligent Systems Society*, vol. 12, no. 4, pp. 341 – 346, 2002.

[124] ——, “Shortest path problem in a type-2 weighted graph,” *Journal of Korea Fuzzy and Intelligent Systems Society*, vol. 11, no. 6, pp. 528 – 531, 2001.

[125] C. Lee, “Fuzzy logic in control systems: Fuzzy logic controller, part II,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 20, no. 2, pp. 419–435, 1990.

[126] T. Schwarz, “Fuzzy systems in the real world,” *AI Expert*, vol. August, 1990.

[127] K. Nakamura and S. Iwai, “Topological fuzzy sets as a quantative description of analogical inference and its application to question-answering systems for information retrieval,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-12, pp. 193–204, 1982.

[128] J. Garibaldi and R. John, “Choosing membership functions of linguistic terms,” in *Proceedings of the 2003 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2003)*. St. Louis, USA: IEEE, New York, 2003, pp. 578–583.

[129] L. Zadeh, “Fuzzy logic and its application to approximate reasoning,” *Information Processing*, vol. 74, pp. 591–594, 1974.

[130] J. Valente de Oliveira, “A design methodology for fuzzy system interfaces,” *IEEE Transactions on Fuzzy Systems*, vol. 3, no. 4, pp. 404–414, 1995.

[131] W. Pedrycz and J. Valente de Oliveira, “Optimization of fuzzy models,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 26, no. 4B, pp. 627–636, 1996.

[132] P. Gärdenfors, *Conceptual Spaces - The Geometry of Thought*. Cambridge, Massachusetts: MIT Press, 2000.

[133] F. Wenstop, “Quantitative analysis with linguistic values,” *Fuzzy Sets and Systems*, vol. 4, no. 2, pp. 95–115, 1980.

[134] J. Garibaldi, “Intelligent techniques for handling uncertainty in the assessment of neonatal outcome,” Ph.D. dissertation, Univeristy of Plymouth, 1997. [Online]. Available: http://www.cs.nott.ac.uk/ jmg/papers/thesis.pdf

[135] J. Garibaldi and S. Musikasuwan, “The association between non-stationary and interval type-2 fuzzy sets,” in *Proceedings of the 2003 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2005)*, Reno, NV, USA, May 2005, pp. 224–229.

[136] J. Garibaldi, M. Jaroszewski, and S. Musikasuwan, “Nonstationary fuzzy sets,” *IEEE Transactions on Fuzzy Systems*, vol. 16, no. 4, pp. 1072–1086, August 2008.

[137] S. Musikasuwan and J. Garibaldi, “Exploring gaussian and triangular primary membership functions in non-stationary fuzzy sets,” in *IPMU 2006*, Paris, France, July 2006.

[138] S. Coupland and R. John, “An investigation into alternative methods for the defuzzification of an interval type-2 fuzzy set,” in *Proc. FUZZ-IEEE 2006*, July 2006, pp. 7196 – 7203.