THE MONOTONICITY AND SUB-ADDITIVITY PROPERTIES OF FUZZY INFERENCE SYSTEMS AND THEIR APPLICATIONS

By

TAY KAI MENG

Thesis submitted in fulfilment of the requirements

for the degree of

Doctor of Philosophy

January 2011

SIFAT MONOTONISITI DAN SUB-TAMBAHAN BAGI SISTEM INFERENS KABUR DAN APLIKASINYA

Oleh

TAY KAI MENG

Tesis yang diserahkan untuk

memenuhi keperluan bagi

Ijazah Doktor Falsafah

Januari 2011

Acknowledgments

This thesis and the research work presented herein would not have been possible without the support of many people. I wish to express my gratitude to my supervisor, Prof. Dr. Lim Chee Peng, who has been abundantly helpful with his invaluable assistance, support, and guidance for this research.

Special thanks also to all my graduate friends, for sharing the literatures and offering invaluable assistance, as well as for their encouragement.

Besides, I wish to express my love and gratitude to my beloved family; for their understanding and endless love, throughout the duration of this study.

TABLE OF CONTENTS

CHAPTER 3 THE MONOTONICITY AND SUB-ADDITIVITY PROPERTIES OF FUZZY INFERENCE SYSTEMS

CHAPTER 4 A NEW FUZZY INFERENCE SYSTEM-BASED OCCURRENCE MODEL

CHAPTER 5 AN IMPROVED FUZZY FMEA METHODOLOGY WITH AN FIS-BASED RPN MODEL

CHAPTER 6 APPLICATION OF FIS-BASED MODELS TO EDUCATION ASSESSMENT AND CONTROL

CHAPTER 7 AN EXTENSION OF THE FUZZY RULE INTERPOLATION TECHNIQUE

CHAPTER 8 CONCLUSIONS AND SUGGESTIONS FOR FURTHER WORKS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AARS	Approximate analogical reasoning scheme
AR	Analogical Reasoning
CBA	Computer-Based Assessment
CRA	Criterion-Referenced assessment
FATI	First aggregate then inference
FERI	Fundamental Equation of Rule Interpolation
FIS	Fuzzy inference System
FITA	First inference then aggregate
FM	Failure Mode
FMEA	Failure Mode and Effect Analysis
FPR	Fuzzy Production Rule
FRI	Fuzzy Rule Interpolation
IC	Integrated circuit
JPEG	Joint photographic experts group
MACI	modified α-cut-based interpolation
MOI	Mean-of-Inversion
NLP	Non-Linear Programming
PCB	Printed circuit board
QP	Quadratic Programming
rep	representative value
RPN	Risk Priority Number
SCA	Sneak Circuit Analysis
SQP	Sequential quadratic programming
SR	Similarity reasoning
WFPR	Weighted fuzzy Production rule

SIFAT MONOTONISITI DAN SUB-TAMBAHAN BAGI SISTEM INFERENS KABUR DAN APLIKASINYA

ABSTRAK

Sistem inferens kabur ialah satu rangka pengkomputeran yang popular untuk masalah pemodelan, klasifikasi, kawalan, dan membuat keputusan. Dalam tesis ini, kajian ditumpukan kepada dua sifat sistem inferens kabur iaitu, sifat monotonik dan sub-tambahan. Sifat tersebut telah ditakrifkan, dan aplikasi mereka untuk masalah-masalah di dunia nyata dibincangkan. Melalui kajian ini, satu prosedur sistematik yang berdasarkan satu asas matematik (iaitu syarat keperluan) untuk membangunkan satu model sistem inferens kabur yang memenuhi sifat monotonik telah direka. Satu cara untuk memperbaik sifat sub-tambahan juga direka. Kebolehan cara-cara yang dicadangkan diuji menggunakan masalah dunia nyata, iaitu, Analisis Mod dan Kesan Kegagalan, penilaian pendidikan dan kawalan. Penggunaan teknik interpolasi peraturan kabur untuk sistem inferens kabur yang mengandungi peraturan yang tidak lengkap turut dikaji. Kajian menunjukkan apabila sifat monotonik diperlukan, teknik interpolasi peraturan kabur yang meramal kesimpulan peraturan secara berasing tidak sesuai untuk pemodelen sistem inferens kabur yang lebih daripada satu masukan. Oleh itu, teknik interpolasi peraturan kabur dirumuskan sebagai satu masalah pengoptimuman berkonstrain untuk sistem inferens kabur yang mempunyai lebih daripada satu masukan. Satu teknik interpolasi peraturan kabur baru yang berdasarkan cara program tidak linear dengan syarat keperluan dicadangkan dan diaplikasikan ke atas Analisis Mod dan Kesan Kegagalan. Keputusan menunjukkan

teknik interpolasi peraturan kabur baru tersebut dapat memenuhi sifat monotonik bagi masalah Analisis Mod dan Kesan Kegagalan.

THE MONOTONICITY AND SUB-ADDITIVITY PROPERTIES OF FUZZY INFERENCE SYSTEMS AND THEIR APPLICATIONS

ABSTRACT

The Fuzzy Inference System (FIS) is a popular computing paradigm for undertaking modelling, control, and decision-making problems. In this thesis, the focus of investigation is on two theoretical properties of an FIS model, i.e., the monotonicity and sub-additivity properties. These properties are defined, and their applicability to tackling real-world problems is discussed. This research contributes to formulating a systematic procedure that is based on a mathematical foundation (i.e., the *sufficient conditions*) to develop monotonicity-preserving FIS models. A method to improve the sub-additivity property is also proposed. The applicability of these proposed approaches are demonstrated using real-world problems, i.e., Failure Mode and Effect Analysis (FMEA) methodology, education assessment problem, and control problem. The use of Fuzzy Rule Interpolation (FRI) for handling the incomplete rule base issue in FIS modelling is studied. This research indicates that whenever the monotonicity property is needed, FRI that predicts each rule consequent separately is not a viable solution to handling the incomplete rule base problem in multi-input FIS-based models. As such, FRI is formulated as a constrained optimization problem for the case of multi-input FIS-based models. A new FRI technique incorporating a Non-linear Programming (NLP) method with the *sufficient conditions* is proposed, and its application to the FMEA methodology is demonstrated. The results confirm the effectiveness of the new FRI scheme in satisfying the monotonicity property in undertaking FMEA problems.

CHAPTER 1

INTRODUCTION

1.1 Background

Inference is a process of drawing a conclusion by applying heuristics (based on logic, statistics, etc.) to observations or hypotheses; or by interpolating the next logical step in an intuited pattern (Kneebone, 2001, Russell and Norvig, 2003). There are two main types of inference, i.e., deductive inference and inductive inference. On one hand, in deductive inference, if its premises are true, then its conclusions must also be true. It is impossible for the premises to be true and yet the conclusions to be false (Kahane, 1990). On the other hand, inductive inference is the process of reaching a general conclusion from specific examples (Russell and Norvig, 2003, Kahane, 1990).

An inference technique is a method that attempts to derive answers from a knowledge base. It can be viewed as the "brain" that reasons about the information in the knowledge base for the ultimate purpose of formulating new conclusions (Russell and Norvig, 2003). From the literature, various inference techniques have been reported, e.g. automatic logical inference (Harrison, 2009), Bayesian inference (Box and Tiao, 1992), probabilistic inference (Pearl, 1988), and fuzzy inference (Jang *et al.*, 1997).

The focus of this thesis is on the Fuzzy Inference System (FIS). A general FIS is a popular model used to tackle a wide variety of problems. Examples of successful application of FIS models include modelling (Du and Zhang, 2008, Jang *et al.*, 1997, Lin and Lee, 1995), classification (Sengur, 2008, Jang *et al.*, 1997, Lin and Lee, 1995), decision (Oluseyi Oderanti and De Wilde, 2010), and control (Kurnaz, *et al.*, 2010, Feng, 2006) problems. An FIS model can be viewed as a computing paradigm based on the concepts of fuzzy set theory, fuzzy production rule (If-Then rule), and fuzzy reasoning (Jang *et al.*, 1997). Examples of popular FIS models include the Mamdani FIS (Mamdani and Assilian, 1975), Sugeno/TSK FIS (Takagi and Sugeno, 1985, and Sugeno and Kang, 1988), and Tsukamoto FIS (Tsukamoto, 1979).

The success of FIS models is largely owing to the following key factors: (i) they are able to utilize linguistic information from human experts (Jang *et al.*, 1997, Lin and Lee, 1995 and Wang, 1992); (ii) they are able to simulate human thinking (Zadeh, 1973); (iii) they are able to capture approximate and inexact nature (i.e., uncertainty) of the real world (Jang *et al*., 1997, Lin and Lee, 1995 and Wang, 1992); (iv) they can be expressed with linguistic variables, which can easily be interpreted by humans (Jang *et al.,* 1997, Lin and Lee, 1995); (v) they are able to act as universal approximators to approximate any real continuous functions to any degree of accuracy (Wang, 1992 and Kosko 1994).

1.2 Problem Statements and Motivations

In view of the popularity and numerous successful applications of FIS models in various domains, researches on the monotonicity property of FIS models have received a lot of attention lately. Consider an FIS model, $y = f(x_1, x_2, \ldots, x_i, \ldots, x_n)$, that fulfils the condition of monotonicity between its output, y, with respect to each of its i^{th} input, x_i within the universe of discourse. The output either monotonically increases or decreases as x_i increases. Hence, $f(x_1, x_2, \dots, x_i^1, \dots, x_n) \le f(x_1, x_2, \dots, x_i^2, \dots, x_n)$ or $f(x_1, x_2, \dots, x_i^1, \dots, x_n) \ge f(x_1, x_2, \dots, x_i^2, \dots, x_n)$, respectively, for $x_i^1 < x_i^2$.

Even though the importance of FIS models and the monotonicity property of FIS models have been studied, the problem of designing and developing monotonicity-preserving FIS models has not been fully studied. There are relatively few investigations addressing the problem of designing monotonicity-preserving FIS models (Kouikoglou and Phillis, 2009). More importantly, there is a lack in the development of systematic methods to construct a monotonicity-preserving FIS model that can be easily applied to solve FIS modelling problems. Thus, in this thesis, a systematic, easy, and yet reliable approach to design and develop monotonicity-preserving FIS models is examined and investigated in details.

Besides, a search in the literature reveals that the use of Similarity Reasoning (SR) methods, such as Analogical Reasoning (AR) and Fuzzy Rule Interpolation (FRI) techniques, in monotonicity-preserving FIS models is not common. However, both AR and FRI are important techniques to provide solutions to FIS modelling problems when the fuzzy rule base is incomplete. Therefore, in this thesis, the applicability of FRI techniques to monotonicity-preserving FIS modelling is examined. An FRI formulation for monotonicity-preserving FIS models is also proposed. In addition to theoretical studies, the practicality of the proposed approach is further demonstrated with FIS-based modelling problems.

1.3 Research Methodology

The methodology adopted in this research is depicted in Figure 1.1. First, the background and literature review on related theory, dynamics, and operations of FIS models are described. The *sufficient conditions* (as explained Section 2.4.4) are extended and a monotonicity-preserving FIS modelling approach is developed. The proposed approach is applied to several practical FIS modelling problems, i.e., an FIS-based Risk Priority Number (RPN) model in Failure Mode and Effect Analysis (FMEA) methodology, an FIS-based education assessment problem, and an FIS-based control problem. A monotonicity-preserving FIS-based occurrence model for FMEA is first proposed and examined. An FMEA methodology procedure with a monotonicity-preserving FIS-based RPN model is suggested, and empirical experiments with information/data collected from a semiconductor manufacturing plant are presented. Then, a monotonicity-preserving FIS-based education assessment model is proposed, and examined with a case study. The use of the proposed procedure in FIS-based control problems is also investigated.

The effectiveness of FRI in developing monotonicity-preserving FIS models is studied. An FRI technique and the *sufficient conditions* are synthesized. An FRI formulation for developing monotonicity-preserving FIS models is proposed. In addition, a new FRI framework that incorporates the *sufficient conditions* is examined. Finally, conclusions from this research are drawn, and suggestions for further work are presented.

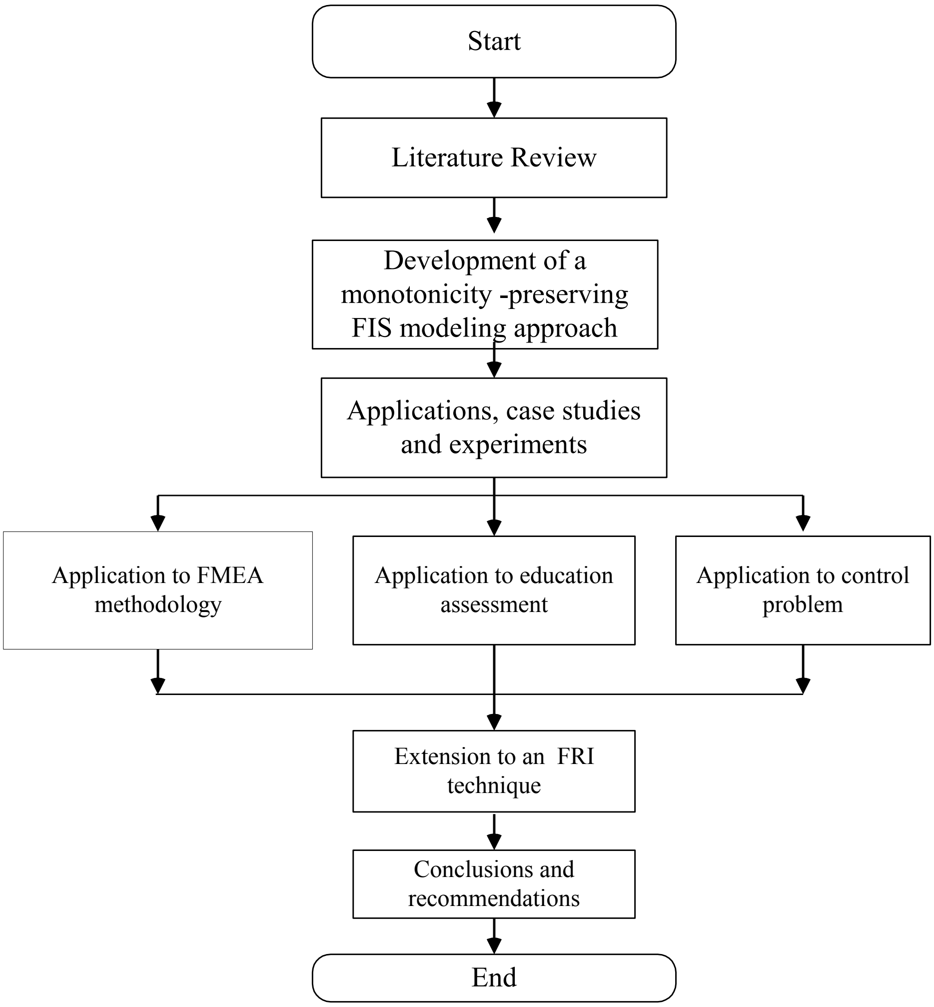

Figure 1.1 Research methodology

1.4 Objectives of the Research

The main aim of this research is to investigate the use of *sufficient conditions* in monotonicity-preserving FIS modelling and to examine the applicability of resulting monotonicity-preserving FIS models. The specific objectives are as follows.

- To examine the use of the *sufficient conditions* as a systematic method for designing and developing monotonicity-preserving FIS models.

- To extend the monotonicity property to another useful property, i.e., the sub-additivity property (a property inspired from the measure theory and the *length function*), and to embed these two properties into FIS-based assessment models.
- To propose an extension of the *sufficient conditions* to FRI techniques and to propose a new FRI framework for designing and developing monotonicity-preserving FIS models.
- To demonstrate the applicability of the resulting FIS models to various problems in the domains of FMEA, education assessment, and control.

1.5 Scope of the Research

In this thesis, the *sufficient conditions* are viewed as a solution to the monotonicity property and the sub-additivity property of FIS models. The scope of research is on the exploitation of the *sufficient conditions* for modelling of a zero-order Sugeno FIS model that preserves the monotonicity property and, at the same time, improves the sub-additivity property. The *sufficient conditions* are further extended to a systematic approach that is proposed in this research to construct monotonicity-preserving FIS models. The effectiveness of the proposed approach is demonstrated using three FIS-based applications, i.e., FMEA, education assessment and control problems. In addition, as a solution to FIS models with an incomplete rule base, the proposed approach is further extended to the use of FRI in modelling of FIS models with monotonic constraints.

1.6 Organization of the Thesis

This thesis is organized as follows. In this introductory chapter, the research background is first described. The problem statement and motivations are explained. The research methodology, objectives, and scope are also presented.

In Chapter 2, the background and literature review on fuzzy set theory, fuzzy ordering, fuzzy distance, FIS models, and FRI techniques are presented. The literature review covers mainly the monotonicity property of the FIS models. In Chapter 3, the monotonicity and sub-additivity properties are defined, and their importance is discussed with a practical example on FMEA. The *sufficient conditions* of an FIS model to be of monotonicity are derived. A novel method to design and construct a monotonicity-preserving FIS model, that is developed based on a sound mathematical foundation, is further proposed. Its applicability is demonstrated with simulated data. Another method to improve the sub-additivity property of an FIS model is also proposed. The derived *sufficient conditions* are further discussed.

In Chapter 4, an FIS-based occurrence model is studied, as an improvement for the conventional FMEA methodology. The FIS-based occurrence model is an example of a single-input monotonicity-preserving FIS model. The applicability of the *sufficient conditions* to this model is discussed and evaluated with benchmark and real-world problems.

In Chapter 5, an improved FMEA methodology, which is incorporated with the *sufficient conditions* and a rule refinement technique, is presented. To examine the effectiveness of the FMEA methodology, a series of experiments with real data sets collected from a semiconductor manufacturing plant is conducted. In Chapter 6, an FIS-based education assessment model, i.e. Criterion-Referenced assessment (CRA), is presented. The FIS-based CRA model incorporates the *sufficient conditions* and the rule refinement technique as a solution to fulfil the monotonicity and sub-additivity properties. A case study on laboratory evaluation is used to demonstrate the usefulness of the FIS-based CRA model. In addition, an FIS-based controller for water level problem is presented.

In Chapter 7, an extension of the *sufficient conditions* to the FRI technique is presented. A generalization of FRI is explained. FRI is further presented as an input-output mathematical model. Together with the *sufficient conditions*, the use of FRI in monotonicity-preserving FIS models is analysed. A simulated problem and a benchmark problem are used to support the analysis. A new formulation for FRI is further proposed, and a new FRI framework for monotonicity-preserving FIS models is developed and examined.

Finally, concluding remarks and contributions of this research are presented in Chapter 8. Suggestions for further works are also included.

CHAPTER 2

BACKGROUND AND LITERATURE REVIEW

2.1 Introduction

In this chapter, the background and literature review on fuzzy set theory, fuzzy ordering, and fuzzy distance, fuzzy set theoretical operations, Fuzzy Production Rule (FPR), fuzzy reasoning, Fuzzy Inference Systems (FISs), and Fuzzy Rule Interpolation (FRI) techniques are presented. A review on the monotonicity property of FIS is further described. Note that the literature review is mainly focused on theoretical aspect of FIS models. The background of sequential quadratic programming (SQP) technique is also presented. Other related literature reviews, especially those on the application of FIS models, are presented in the appropriate sections in subsequent chapters.

This chapter is organized as follows. In Section 2.2, fuzzy set theory, fuzzy ordering, fuzzy distance, and fuzzy set theoretical operations are presented. The FIS-based models and the monotonicity property are discussed in Sections 2.3 and 2.4, respectively. In Section 2.5, a review on Analogical Reasoning (AR) and FRI is presented. In Section 2.6, a review on the measure theory and length function is presented. In Section 2.7, SQP technique is presented. Finally, concluding remarks are presented in Section 2.8.

2.2 Background on Fuzzy Set Theory and Related Operations

In this section, fuzzy set theory, and several important concepts of fuzzy set theory, i.e., representative value, fuzzy ordering, and fuzzy distance are explained.

2.2.1 Fuzzy Set Theory

The theory of sets as a mathematical discipline was introduced by a German mathematician, G. Cantor (1845-1918) (Stoll, 1975). Cantor suggested that a set is made up of objects called members or elements, and one can determine whether or not an object is a member of a set (Stoll, 1975). Let X be a space of objects and x be a generic element of X. A set, A, is defined as a collection of elements or objects $x \in X$, as such that each x can either belong to or not belong to set A. Thus, the characteristic function or membership function of a set can be represented with Equation (2.1), either belong ($\mu_A(A) = 1$) or not belong ($\mu_A(A) \neq 1$) to A. In this thesis, Cardon's set is named classical set.

$$\mu_A(A) = \begin{cases} 1 & if\ and\ only\ if\ x \in A \\ 0 & if\ and\ only\ if\ x \notin A \end{cases} \quad (2.1)$$

A fuzzy set, on the other hand, introduces vagueness by eliminating the sharp boundary that divides members from non members in the group (Zadeh, 1965). The transition from members to non members is gradual, rather than abrupt. Thus, the characteristic function of a fuzzy set is allowed to have a value between 0 and 1, indicating the degree of membership in a given set. A fuzzy set, A, in X is defined as a set of ordered pairs:

$$A = \{(x, \mu_A(A)) | x \in X\}$$

where $\mu_A(A)$ is the membership function of x in A.

Several types of membership functions can be used to represent a fuzzy set, such as triangular, trapezoidal, Gaussian, and generalized bell functions (Jang *et al.*, 1997, Lin and Lee, 1995). A Gaussian membership function is fully specified by two parameters, i.e. centre c and standard deviation σ. Figure 2.1 shows a Gaussian membership function as defined in Equation (2.2). The derivative of a Gaussian membership function with respect to x, is shown in Equation (2.3).

$$\mu_G(x: c, \sigma) = e^{-[x-c]^2/2\sigma^2} \quad (2.2)$$

$$G'(x) = -((x - c)/\sigma^2)G(x) \quad (2.3)$$

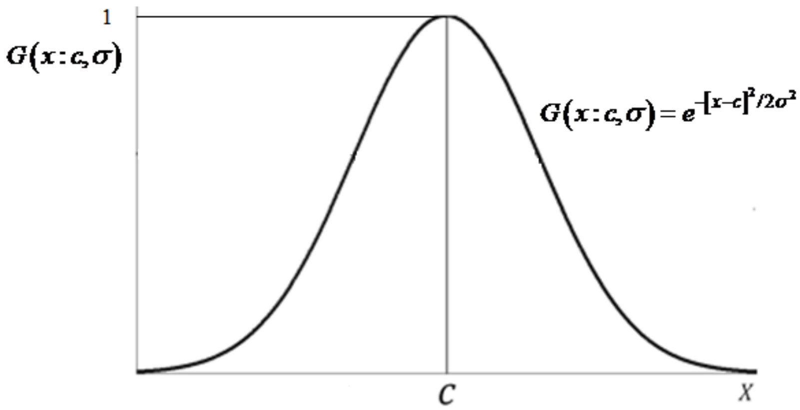

Figure 2.1 A Gaussian membership function

A α-cut of a fuzzy set A is a crisp set of A_α that contains all the elements of the universe set X that have a membership grade equals to or greater than α, where $1 \geq \alpha \geq 0$, as shown in Equation (2.4).

$$A_\alpha = \{x \in X | \mu_A(x) \geq \alpha\} \quad (2.4)$$

2.2.2 Representative Value of a Fuzzy Set

The representative value (rep) of a fuzzy set carries important information about the overall location, or the "most typical" location of a fuzzy set in its domain (Huang and Shen, 2006, 2008, Baranyi *et al.* 2004). For a fuzzy set A, in X, its representative, $rep(A)$ is a numerical value in the X domain. There are several ways how this value can be derived. Defuzzification is one of the most popular methods to obtain a crisp representative value of fuzzy membership functions within the universe of discourse (Jang *et al.*, 1997, Lin and Lee, 1995). Jang *et al.* (1997) listed five defuzzification operators, namely centroid of gravity, mean of maximum, bisector of area, the smallest of maximum, and the largest of maximum.

Assume that the lower and upper bounds of fuzzy set A in X are given by $\underline{A_\alpha} = \underline{support(A)}$ and $\overline{A_\alpha} = \overline{support(A)}$, respectively, the centre point of A, $cp(A)$ is defined in Equation (2.5). If A is convex and normal, with the α-cut method, $cp(A)$ is defined in Equation (2.6).

$$cp(A) = \frac{\overline{A_\alpha} - \underline{A_\alpha}}{2} \tag{2.5}$$

$$cp(A) = \frac{Sup_\alpha(A) - Inf_\alpha(A)}{2} \tag{2.6}$$

Note that $Inf_\alpha(A)$ and $Sup_\alpha(A)$ refer to infima and suprema of A in their α-cut. Alternatively, it can be determined by the point whereby the value of the fuzzy membership function equals to 1 (Huang and Shen, 2006, 2008).

2.2.3 Fuzzy Ordering and Distance of Fuzzy Sets

In 1990's, several important concepts of fuzzy sets were introduced, which included fuzzy ordering and fuzzy distance. For a bounded and gradual domain X, with a generic element x, a full ordering of x exists (Dubois and Prade, 1992). Kóczy and Hirota (1993a, 1993b, 1997) showed the possibility of introducing fuzzy ordering among all elements of x.

Consider two convex and normal fuzzy sets of universe X, namely A^1 and A^2 If $Inf_\alpha(A^1) \leq Inf_\alpha(A^2)$ and $Sup_\alpha(A^1) \leq Sup_\alpha(A^2)$, then $A^1 \preccurlyeq A^2$. Figure 2.2 illustrates the concepts of fuzzy ordering and fuzzy distance. From the fuzzy ordering principle, the basic concept of fuzzy distance for comparing fuzzy sets of the same universe, as well as for measuring the distance of each α-cut, separately, is introduced.

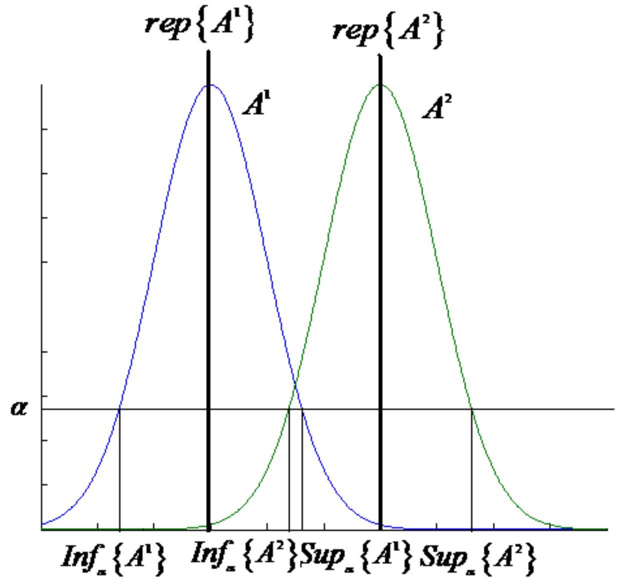

Figure 2.2 Fuzzy ordering and fuzzy distance

Based on Figure 2.2, the lower distance is defined as the distance of infima A^1 and A^2 at their α-cut, and the upper distance is calculated in a similar way with respect to their suprema, as in Equations (2.7) and (2.8), respectively.

$$Lower\ distance,\ d_l\left(A^1, A^2\right) = Inf_\alpha\left(A^2\right) - Inf_\alpha\left(A^1\right) \tag{2.7}$$

$$Upper\ distance,\ d_u\left(A^1, A^2\right) = Sup_\alpha\left(A^2\right) - Sup_\alpha\left(A^1\right) \tag{2.8}$$

The concept of fuzzy distance is important. It acts as the principle of various FRI techniques (Kóczy and Hirota, 1993a, 1993b, and 1997). Fuzzy distance and fuzzy ordering between two fuzzy sets of the same universe of discourse can also be defined by their representative values. For example, in Figure 2.2, the representative values are determined by the point whereby the fuzzy membership function value is 1. If $rep(A^1) \leq rep(A^2)$, then $A^1 \preccurlyeq A^2$. Equation (2.9) defines a simple fuzzy distance (known as general closeness) between two fuzzy sets, $d\left(A^1, A^2\right)$.

$$Fuzzy\ distance, d\left(A^1, A^2\right) = rep\left(A^2\right) - rep\left(A^1\right) \tag{2.9}$$

This definition is used in solid cut fuzzy set interpolation (Baranyi *et al.*, 2004) and in FRI techniques proposed by Huang and Shen (2006, 2008).

2.2.4 Fuzzy Set Theoretic Operations

Three of the most basic operations on classical sets are union, intersection, and complement. Corresponding to these three operations, fuzzy sets have similar operations, as defined by Zadeh (1965).

The union of two fuzzy sets, A and B, is a fuzzy set C, written as $C = A \cup B$ or $C = A\ OR\ B$. The membership function of C can be related to those of A and B, $\mu_C(x) = \mu_A(x) \vee \mu_B(x)$. Jang *et al.* (1997) listed several frequently used union operators, as follows.

Minimum: $\mu_C(x) = \min\big(\mu_A(x), \mu_B(x)\big)$

Algebraic product: $\mu_C(x) = \mu_A(x) \times \mu_B(x)$

Drastic product:

$$\mu_C(x) = \begin{cases} \mu_A(x) & if\ \mu_A(x) = 1 \\ \mu_B(x) & if\ \mu_B(x) = 1 \\ 0 \quad if\ \mu_A(x), \mu_B(x) < 1 \end{cases}$$

The intersection of two fuzzy sets, A and B, is a fuzzy set C, written as $C = A \cap B$ or $C = A\ AND\ B$. The membership function of C can be related to those of A and B, $\mu_C(x) = \mu_A(x) \wedge \mu_B(x)$. Again, Jang *et al.* (1997) listed several frequently used intersection operators, as follows.

Maximum: $$\mu_C(x) = \max\left(\mu_A(x), \mu_B(x)\right)$$

Algebraic sum: $$\mu_C(x) = \mu_A(x) + \mu_B(x) - \mu_A(x)\mu_B(x)$$

Drastic sum: $$\mu_C(x) = \begin{cases} \mu_A(x) & if\ \mu_A(x) = 0 \\ \mu_B(x) & if\ \mu_B(x) = 0 \\ 1 & if\ \mu_A(x), \mu_B(x) > 0 \end{cases}$$

The complement of fuzzy set A is denoted by $\bar{A}$ ($NOT\ A$). The membership function of $\bar{A}$, can be written as $\mu_{\bar{A}}(x) = 1 - \mu_A(x)$.

2.3 Background on Fuzzy Inference Systems and Related Operations

In this section, a review on FPR and fuzzy reasoning for FIS models is described. Besides, a popular FIS model, i.e., the zero-order Sugeno/TSK model, is explained.

2.3.1 Fuzzy Production Rules (Fuzzy IF-THEN Rules)

A major component of an FIS model is its FPRs (Mendel, 1995, Jang *et al.*, 1997). An FPR is expressed as a logical implication, i.e., in a form of an If-Then statement. Each FPR comprises two parts: an antecedent and a consequent. An example of an FPR is $IF\ A\ \ Then\ B$, where A is the antecedent and B is the consequent. It is a form of proposition, whereby a proposition is an ordinary statement involving terms which has been defined, e.g. “the damping ratio is low” (Mendel, 1995). From the proposition, the relevant rule can be obtained: “IF the damping ratio is low THEN the system’s impulse response oscillates a long time before it dies out”. Propositions can be combined or modified in many ways, via the set-theoretic operations, i.e., AND, OR, and NOT.

The main idea of FIS models resembles that of "divide and conquer", i.e., at the antecedent, an FPR defines a fuzzy region at the input space, while the consequent describes the behaviour of the region (Jang *et al.*, 1997). There is a number of strategies to partition the input space to form the antecedent. Among them are grid partition, tree partition, and scatter partition (Jang *et al.*, 1997, Lin and Lee, 1995).

The grid partition is popular, and it is often chosen for designing FIS models (Jang *et al.*, 1997). With the grid partition, an FPR with *n* antecedents has the form:

$$If\ (x_1 is\ A_1)AND\ (x_2 is\ A_2) \dots AND(x_n is\ A_n), THEN(y\ is\ B)$$

where x_i and y are the inputs and output of the FIS model, A_1, A_2,…A_n and B are linguistic variables/fuzzy sets for the inputs and output, respectively.

A Weighted Fuzzy Production Rule (WFPR) is an enhancement of an FPR. A WFPR allows knowledge imprecision to be taken into account by adding extra knowledge representation parameters, which include threshold value, certainty factor, local weight, and global weight (Yeung and Tsang, 1997, Lau and Chan, 1997). Generally, a WFPR with *n* antecedents can be represented by:

$$If\ (x_1 is\ A_1)AND\ (x_2 is\ A_2) \dots AND(x_n is\ A_n), THEN(y\ is\ B)$$

$$CF_R, LW_1, LW_2, \dots, LW_n, \lambda_{A_1}, \lambda_{A_2}, \dots, \lambda_{A_n}, GW$$

$$Fact_i: x_i\ is\ A_i^{'}\ with\ CF_i, 1 \le i \le n$$

$$Conclusion: Y\ is\ B'\ with\ CF_{B'}$$

The parameters are explained as follows. A threshold value, λ_{A_i} is assigned to a proposition. It ensures that the degree of similarity between the proposition (x_i is A_i) and its fact, i.e., greater than or equal to λ_{A_i}. The assignment of λ_{A_i} to "x_i is A_i " is not only to ensure the result of an approximate reasoning method is reasonable but also to prevent or reduce rule mis-firing (Yeung and Tsang, 1997). The certainty factor for a given fact ($Fact_i$), CF_i determines how certain the proposition is. It is used to express how accurate, truthful, or reliable the fact is. The certainty factor can also be applied to a rule (CF_R). It means how certain the relationship the antecedent and the consequent is (Yeung and Tsang, 1997).

For an FPR that comprises more than a proposition connected by "AND", the local weight for a proposition (x_i is A_i), LW_i is used to indicate the degree of importance of the proposition in relation to the antecedent (Yeung and Tsang, 1997). Global weight, GW , is used to indicate the degree of importance of each rule's contribution to the final goal. There are two different applications of the global weight (Yeung and Tsang, 1997): (i) to compare the relative degree of importance of a particular rule with those from other rules in a given inference path leading to a specific output membership function; (ii) to show the relative importance of a rule when it is used in different inference paths leading to different output membership functions.

2.3.2 Fuzzy Reasoning

Fuzzy reasoning (also known as approximate reasoning) is an inference procedure that derives a conclusion from a set of FPRs. It can be written as

FPR:	$If\ x\ is\ A\ \ Then\ y\ is\ B$
Fact:	$x\ is\ A'$
Consequent:	$y\ is\ B'$

where A' is close to A, and B' is close to B.

2.3.3 The Zero-Order Sugeno Fuzzy Inference System

An FIS model can be explained as a computing paradigm that is based on the concepts of fuzzy set theory, FPRs, and fuzzy reasoning. Consider an FIS model with n inputs. Let $\bar{x} = (x_1, x_{2,} \dots, x_{n-1}, x_n)$ be the input vector in a rectangular region, $U = U_1 \times U_2 \times \dots \times U_n$, where $U_i = [L_i, H_i]$ for $1 \leq i \leq n$. Consider M_i terms at the i^{th} input space, A_i^1, A_i^2, …,$A_i^{M_i}$, which are represented by fuzzy membership functions $\mu_i^1(x_i)$, $\mu_i^2(x_i)$, …, and $\mu_i^{M_i}(x_i)$, respectively. The output of the FIS model, $y = f(x)$, falls within the range of $[L_v, H_v]$. If a full grid partition is used, the number of fuzzy rule is $\prod_{i=1}^{n} M_i$.

The FPRs of a single-input ($n = 1$) FIS model, i.e., $R^k: A^K \rightarrow B^K$, are represented as follows.

$R^1: If\ x\ is\ A^1 Then\ y\ is\ B^1$

$R^2: If\ x\ is\ A^2 Then\ y\ is\ B^2$

.

.

.

$R^M: If\ x\ is\ A^M Then\ y\ is\ B^M$

Note that A^1, A^2, …, and A^M are linguistic terms at the rule antecedent part, and are represented by fuzzy membership functions $\mu^1(x)$, $\mu^2(x)$, …, and $\mu^M(x)$, respectively; B^1, B^2, …, and B^M are membership functions at the rule consequent part. The output of the FIS model is obtained using an inference technique, as in Equation (2.10),

$$y = f(x) = \frac{\sum_{j=1}^{j=M} \mu^j(x) \times b^j}{\sum_{j=1}^{j=M} \mu^j(x)} \tag{2.10}$$

where b^j, is the representative value (as explained in Section 2.2.2) of membership function B^j.

The FPRs for single-input FIS models can be extended and used in multi-input FIS models ($n > 1$), as follows:

$R^{j_1,j_2,\ldots,j_n}$ (Rule $\#M'$):

$$If\ (x_1 is\ A_1^{j_1}) AND\ (x_2 is\ A_2^{j_2}) \ldots AND (x_n is\ A_n^{j_n}), THEN (y\ is\ B^{j_1,j_2,\ldots,j_n})$$

To simplify the notation, each fuzzy rule ($R^{j_1,j_2,\ldots,j_n}$) is represented by an index, M', where $1 \leq M' \leq \prod_{i=1}^{n} M_i$. Consider the AND operator as the product function. The output is obtained by using the weighted average of a representative real value, $b^{j_1,j_2,\ldots,j_n}$, with respect to its compatibility grade, as in Equation (2.11).

$$y = f(\bar{x}) = \frac{\sum_{j_n=1}^{j_n=M_n} \ldots \sum_{j_2=1}^{j_2=M_2} \sum_{j_1=1}^{j_1=M_1} \left(\mu_1^{j_1}(x_1) \times \mu_2^{j_2}(x_2) \times .. \times \mu_n^{j_n}(x_n) \times b^{j_1,j_2,\ldots,j_n} \right)}{\sum_{j_n=1}^{j_n=M_n} \ldots \sum_{j_2=1}^{j_2=M_2} \sum_{j_1=1}^{j_1=M_1} \left(\mu_1^{j_1}(x_1) \times \mu_2^{j_2}(x_2) \times .. \times \mu_n^{j_n}(x_n) \right)} \tag{2.11}$$

where $b^{j_1,j_2,\ldots,j_n}$ is the representative value of membership function $B^{j_1,j_2,\ldots,j_n}$.

FIS models can be classified into two categories: First Inference Then Aggregate (FITA) and First Aggregate Then Inference (FATI) (Cordon *et al.*, 1997, Emami *et al.*, 1999, Hisao *et al.*, 2006). Equations (2.10) and (2.11) belong to FITA (Hisao *et al.*, 2006). For an FITA model, the representative value is first determined. Then, the output estimate is obtained by aggregating the crisp values of the compatibility fuzzy rules. The weighted average is one of the methods to obtain the output estimate. An FIS model that uses a fuzzy set or a crisp value at its rule consequent as in Equations (2.10) and (2.11) is categorized as an FIS model with a high degree of interpretability (Casillas, et al. , 2003), which allows direct translation of the rules. Besides, Equations (2.10) and (2.11) represent a singleton, zero-order Sugeno FIS. The associated fuzzy reasoning method has several advantages, e.g., its reasoning mechanism is simple and it is suitable for gradient-based learning algorithm (Hisao *et al.*, 2006).

2.3.4 Recent Advances on Fuzzy Inference System Modelling

Over the years, researches to enhance FIS models have been reported. Examples include fuzzy systems with neural network learning, e.g. ANFIS (Jang, 1993, Jang and Sun, 1995), and with evolutionary computation learning (Ishibuchi *et al.,* 1995). A type-two FIS model that incorporates type-two fuzzy sets (Zadeh, 1975) has been investigated in Karnik *et al*. (1999) and Liang and Mendel (2000). An FIS model with a rule reduction technique based on a similarity measure and with interpretability improvement has been suggested in Jin (2000). Other advances include the development of AR techniques (Turksen and Zhao, 1988) and various FRI techniques (Kóczy and Hirota, 1993a, 1993b, 1997) for FIS models. AR and/or FRI techniques are developed from the principles of similarity measure, fuzzy ordering, fuzzy partial ordering, and fuzzy distance. They are introduced as a solution to an incomplete rule base, which allows an unknown rule consequent of an observation to be predicted. Details on FRI techniques are further presented on Section 2.5.

2.4 Background and Review on the Monotonicity Property of a Fuzzy Inference System

Motivated by the popularity and numerous successful applications of FIS models in various domains, researches on the monotonicity property of FIS models have received a lot of attention lately. Consider an FIS, $y = f(x_1, x_2, \dots, x_i, \dots, x_n)$, that fulfils the condition of monotonicity between its output, y, with respect to its i^{th} input, x_i within the universe of discourse. The output of the model either monotonically increases or decreases as x_i increases, hence, $f(x_1, x_2, \dots, x_i^1, \dots, x_n) \leq f(x_1, x_2, \dots, x_i^2, \dots, x_n)$ or $f(x_1, x_2, \dots, x_i^1, \dots, x_n) \geq f(x_1, x_2, \dots, x_i^2, \dots, x_n)$, respectively, for $x_i^1 < x_i^2$.

The importance of the monotonicity property of FIS models has been explained in a number of publications. Among them include (i) many real-world systems obey the monotonicity property (Angeli and Sontag, 2003, Kouikoglou and Phillis, 2009, Won *et al.*, 2002, Lindskog and Ljung, 2000); (ii) this property is important for undertaking some FIS modelling problems, e.g., queuing (Kouikoglou and Phillis, 2009), decision making (Kouikoglou and Phillis, 2009), control (Won *et al.*, 2002, Zhao and Zhu, 2000), assessment models (Kouikoglou and Phillis, 2009), JPEG models (Wu and Sung (1994, 1996); (iii) in the case whereby the number of data samples is small, it is important to fully exploit the monotonicity property as an additional qualitative information (Broekhoven and Baets, 2008, 2009); (iv) exploitation of the monotonicity property as an additional qualitative knowledge allows the development of various improved system identification or modelling procedures that are susceptible to noise and inconsistencies in data samples as well as able to suppress overfitting (Broekhoven and Baets, 2008, 2009).

From the literature, studies related to the monotonicity property of FIS models have been reported. Generally, these studies focus on two domains: (i) mathematical conditions of an FIS model to satisfy (or not to satisfy) the monotonicity property, (ii) development of a method to construct a monotonicity-preserving FIS model. In the first domain, Zhao and Zhu (2000) examined the conditions for single-input and two-input Mamdani FIS models to be of monotonicity, with analysis of the FIS operations step-by-step. However, their analysis focused on the case that the membership functions at the input space are equally divided. Thus, Won *et al.* (2002) derived a set of *sufficient conditions* for the first-order Sugeno fuzzy models by differentiating the output of an FIS model with respect to its input(s). This is more reliable, as it has a sound mathematical foundation. However, this approach may not be applicable to FIS models with non-derived operators, e.g., minimum operators. Broekhoven and Baets (2009) further analyzed the use of three T-norm operators, i.e., minimum, product and, Łukasiewicz, in monotonicity-preserving Mamdani–Assilian FIS models.

For the second domain, Wu and Sung (1994, 1996) proposed a new defuzzification operator, i.e., Mean-of-Inversion (MOI) for monotonicity-preserving FIS models. Lindskog and Ljung (2000) proposed a monotonicity-preserving FIS design procedure by adding parametric constraints. As pointed out in Kouikoglou and Phillis (2009), the methods from Wu and Sung (1996) and Lindskog and Ljung (2000) focused on triangular membership functions only. Kouikoglou and Phillis (2009) suggested that exploitation of the *sufficient conditions* in FIS modelling might be a better idea. The derived *sufficient conditions* can be combined with a least-square and an evolutionary computation-based learning methods (Koo *et al.*, 2004, Won *et al.* 2001). Li *et al.* (2009) further extended the *sufficient conditions* to Sugeno FIS models with type-two fuzzy sets. Kouikoglou and Phillis (2009) also extended the *sufficient conditions* to hierarchical FIS models. Some important findings that are closely related to this research are further discussed in the subsequent sections.

2.4.1 Findings from Wu and Sung (1994, 1996) and Wu (1997)

Wu and Sung (1996) and Wu (1997) described another research related to the monotonicity property. They suggested that the monotonicity property is important for the stability analysis of FIS-based control problems (Wu and Sung, 1996). They also stressed the importance of the monotonicity property for JPEG models in image compression (Wu and Sung, 1994, 1996).

They focused on FIS with triangular membership functions. A new defuzzification operator, i.e., mean-of-inversion (MOI) for monotonicity-preserving FIS models, was proposed (Wu and Sung, 1996 and Wu, 1997). The MOI operator defuzzifies each fired rule separately, instead of superimposing all fired rules before defuzzification.

2.4.2 Findings from Zhao and Zhu (2000)

Zhao and Zhu (2000) suggested that for most process control problems, regardless of single-input single-output, or multi-input, multi-output problems, the relationship of the input and output obeys the monotonicity property, i.e., the output of the process can be expressed as a monotonic function of the input variables. They further suggested that the monotonicity property is important to ensure the stability and the steady state error of an FIS-based control problem. The conditions for single-input and two-input Mamdani FIS models to be of monotonicity are also presented, with the FIS operations analyzed step-by-step. Their study considers membership functions that are well partitioned. Their findings suggest that as long as the rule base is monotonically-ordered, a single-input Mamdani fuzzy model can be of monotonicity, and a two-input Mamdani fuzzy model can be roughly of monotonicity.

2.4.3 Findings from Lindskog and Ljung (2000)

Lindskog and Ljung (2000) again pointed out the importance of the monotonicity property in FIS-based control problems. They focused on FIS models with triangular membership functions. It was further assumed that the triangular membership functions are orthogonal, i.e, summation of the membership value at every point of the input space is 1.

A procedure to construct a monotonicity-preserving FIS model was suggested. An FIS structure that ensures input–output monotonicity is proposed, and is used to identify the dynamic system whose output is monotonic with respect to its input. They further parametrized the FIS structure. Constraints for each parameter is developed and imposed in the FIS designing process. A case study related to a water heating system is reported.

2.4.4. Findings from Won *et al.* (2001, 2002)

Won *et al.* (2002) reported that many real-world engineering systems satisfy the monotonicity property. Two examples, i.e., the cart-pole system and the magnetic crane controller system, are explained. They suggested that FIS models that preserve the monotonicity property are able to better approximate the actual control mechanism.

A set of *sufficient conditions* for the first-order Sugeno FIS by differentiating the output of the FIS with respect to its input(s) is derived. The derived condition was later combined with least-square learning (Koo *et al.*, 2004) and evolutionary computation-based learning (Won *et al.* 2001). Besides, Li *et al.* (2009) further extended the *sufficient conditions* to the Sugeno FIS with type-two fuzzy sets. Combination of the *sufficient conditions* with learning algorithms allow a monotonic FIS model to be constructed from data samples, based on a learning theory.

2.4.5 Findings from Broekhoven and Baets (2008, 2009)

Broekhoven and Baets (2008, 2009) pointed out that it is important to fully exploit the monotonicity property as additional qualitative information, especially in the case whereby the number of data samples is small.

Even though the *sufficient conditions* from Won *et al.* (2002) are useful, they do not explain the scenario if the min operator is used as the AND operator. Thus, Broekhoven and Baets (2008, 2009) further investigate the use of three basic AND