Acquisition of Fuzzy Measures in Multicriteria Decision Making Using Similarity-based Reasoning

PhD Thesis

Amol S. Wagholikar
B.E (Industrial Engg), MInfTech

Gold Coast, Australia
School of Information and Communication Technology

Submitted in fulfilment of the requirements for the degree of Doctor of Philosophy (PhD)

January 2007

ABSTRACT

Continuous development has been occurring in the area of decision support systems. Modern systems focus on applying decision models that can provide intelligent support to the decision maker. These systems focus on modelling the human reasoning process in situations requiring decision. This task may be achieved by using an appropriate decision model. Multicriteria decision making (MCDM) is a common decision making approach. This research investigates and seeks a way to resolve various issues associated with the application of this model.

MCDM is a formal and systematic decision making approach that evaluates a given set of alternatives against a given set of criteria. The global evaluation of alternatives is determined through the process of aggregation. It is well established that the aggregation process should consider the importance of criteria while determining the overall worth of an alternative. The importance of individual criteria and of sub-sets of the criteria affects the global evaluation. The aggregation also needs to consider the importance of the sub-set of criteria. Most decision problems involve dependent criteria and the interaction between the criteria needs to be modelled. Traditional aggregation approaches, such as weighted average, do not model the interaction between the criteria.

Non-additive measures such as fuzzy measures model the interaction between the criteria. However, determination of non-additive measures in a practical application is problematic. Various approaches have been proposed to resolve the difficulty in acquisition of fuzzy measures. These approaches mainly propose use of past precedents. This research extends this notion and proposes an approach based on similarity-based reasoning. Solutions to the past problems can be used to solve the new decision problems. This is the central idea behind the proposed methodology. The methodology itself applies the theory of reasoning by analogy for solving MCDM problems.

This methodology uses a repository of cases of past decision problems. This case base is used to determine the fuzzy measures for the new decision problem. This work also analyses various similarity measures. The illustration of the proposed methodology in a case-based decision support system shows that interactive models are suitable tools for determining fuzzy measures in a given decision problem. This research makes an important contribution by proposing a similarity-based approach for acquisition of fuzzy measures.

ACKNOWLEDGEMENTS

I thank my principal supervisor Assoc Prof. Peter Deer for his valuable guidance in this journey of discovery. He has been a great inspiration to me. I am very fortunate to have worked with him during this task. I sincerely thank my associate supervisor Dr Michael Blumenstein for his guidance and motivation. I am very grateful to School of Information and Communication Technology, Griffith University, for giving me this valuable opportunity and providing state-of-the-art research facilities.

I also thank all the researchers in my field whose work has been like a lighthouse on this path. My special thanks go to my colleagues Yuliang Fan, K. Moon, Andrew Petersen and Dr Michael Cowling for their valuable input and motivation. It has been a great pleasure to work with these talented people and learn from them. I also thank my students for their interest in this research.

Finally, yet importantly, I thank my mother, father, sister and all other family members for their support. I appreciate and thank all my friends for their encouragement during difficult phases. I am especially grateful to my wife Meghana, for her love and support.

Statement of Originality

This work has not been previously submitted for a degree or diploma in any university. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made in the thesis itself.

Amol S. Wagholikar

Date: 3 January 2007

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

List of Acronyms

AI: Artificial Intelligence

DSS: Decision Support Systems

IDSS: Intelligent Decision Support Systems

ES: Expert Systems

KBDSS: Knowledge-based Decision Support Systems

MCDM: Multicriteria Decision Making

AHP: Analytical Hierarchy Process

OWA: Ordered Weighted Average

SM: Similarity Measures

AR: Analogical Reasoning

CBR: Case-based Reasoning

KBS: Knowledge-based Systems

IR: Information Retrieval

CI: Case Indexing

Chapter 1 Introduction

Decision-making is one of the central functions of the human brain. Everybody makes decisions. However, human choice behaviour is complex. The decision making process has been a subject of scientific investigation. Different decision theories have been developed with different characteristics (Tsoukiàs 2003). These theories lead to different computational algorithms in various Decision Support Systems (DSS). These systems (DSS) are a class of systems whose main objective is to assist the decision maker in the decision-making process (Bonczek et al. 1981). Modern DSS apply artificial intelligence (AI) concepts and algorithms to incorporate "intelligent" behaviour in the systems. Such systems can be described as intelligent decision support systems (IDSS). These systems focus on providing the decision support by processing information in a human-like way (Anderson 2000).

With the advent of disciplines like AI and its subsequent advances, there is a continuous need of better mechanisms that can model human reasoning. This revolution has been in the direction of making "machines smarter" so that they can perform human tasks. Thus the conventional DSS are evolving into IDSS that provide intelligent assistance to the decision maker. IDSS plays an important role in situations where human expertise is involved. Unlike the conventional DSS, IDSS are useful in supporting decision makers in evaluating the qualitative tasks such as:

- Discovery of opportunities or problems (i.e. monitoring and interpretation) (King 1990);
- Suggesting alternative courses of action (i.e. repair) (King 1993); and
- Advising on the use of statistical, database, simulation and operations research techniques (Klein & Methlie 1995; Alter 1980).

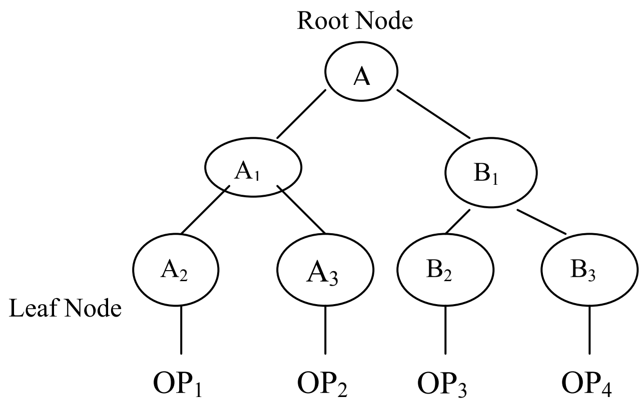
The importance of rich information is highly significant especially in decision-making situations. The information assists the decision-maker to decide their course of action. The more accurate is the information, the greater is the possibility of an optimum decision. With the continuous evolution in the discipline of computing, modern systems focus on modelling the human tasks (Mallach 2000). Indeed there appears to be a growing need to develop approaches for building intelligent systems that can specifically conduct human tasks. Such intelligent systems should have the capability to provide decision support based on a knowledge-base built using expert input. These systems include an Expert System (ES) as one of the main components (Klein & Methlie 1995). This component provides knowledge as per the decision making context using AI to the decision maker. Knowledge-based decision support systems (KBDSS) play the role of intelligent decision support systems (Zeleny 1987). These systems incorporate domain knowledge, modelling and analysis to provide decision makers with the capability of intelligent decision support. Such systems use AI tools to create DSS that can provide intelligent decision support to complex decision-making problems. The undertaken research continues the pursuit of developing approaches for automating the decision making process for providing intelligent decision support to the decision maker. This chapter focuses on introducing the research question and justifying its importance. The chapter outlines the research scope and sets a roadmap for this research towards the solution of the undertaken research question.

1.1 Decision Making Models

There are various decision making models. Many of them are rule-based models. The most commonly used models are decision trees and decision tables (Hoffer 1996). Decision trees are used for decision making under uncertainty. The method uses probabilities of certain events to choose the best course of action. Decision tables model the decision problem in the form of a matrix that uses certain production rules to specify possible conditions for the decision and the resulting actions. These models are explained below.

1.1.1 Decision Trees

The decision tree is a commonly used graphical tool for solving decision problems. It consists of graphical representation of decisions and their possible consequences. It is a predictive model that involves mapping observations about an item to conclusions about the item’s target value (Berikov & Litvinenko 2003). The decision tree comprises root nodes, leaf nodes and branches or arcs joining two nodes. In this decision model the outputs (i.e. alternatives) are evaluated based on the outcomes at each node. Each node corresponds to an event; an arc to a child represents a possible outcome of that event. A leaf node represents the predicted value of a final outcome variable given the values of the outcomes represented by the path from the root. The overall path from the root node to the final outcome can represent a rule for the given decision making process. The following figure shows a decision tree in a general form.



A → Root Node

A_i → Event A where i= 1, 2,…, n

B_1→ Child Node

Bj → Event B where j = 1,2,…,m

OP_k→ Outcome where k = 1, 2,…, l

Figure 1.1 A Generic Form of a Decision Tree

Decisions trees are easy to understand. They can easily be mapped to a set of production rules and applied to real problems. They can process both numerical and categorical data and are computationally faster than other approaches such as Neural Networks (Roiger 2003). However, decision trees also have certain disadvantages. The decision path leads to a single outcome. In reality, there might be multiple outcomes for a single path. Trees created from numeric datasets can be complex (Roiger 2003). The decision tree results in a categorical outcome. This analysis of decision trees shows that these models are not suitable for modelling complex decision problems that involve multiple output attributes and large numeric data sets.

1.1.2 Decision Tables

Decision tables are an engineering approach of modelling production rules (Vanthienen 2001). In this model, conditions are formed into a table, which also holds appropriate actions. Classical decision tables use binary logic to express states of conditions and actions to be performed. A generic decision table is shown below.

Table 1.1 Decision Table

$Condition_1$	$Rule_{11}$	$Rule_{12}$	…….
$Condition_2$	$Rule_{21}$	$Rule_{22}$	…….
…. ….	…. ….	…. ….	…. ….
$Action_1$	W_{11}	W_{12}	….
$Action_2$	W_{21}	W_{22}	….

Each condition in a decision table corresponds to a variable whose possible values are listed among the condition alternatives. Condition alternatives are represented by the rules. Each action is a procedure or operation to perform, and the entries specify whether (or in what order) the action is to be performed for the set of condition alternatives to which the entry corresponds. The rules model the condition alternatives. For example, in table 1.1, Rule_11, Rule_12 and subsequent rules are the condition alternatives for $Condition_1$. The $Action_1$ and $Action_2$ are specified by the values of these rules.

The main advantage of a decision table is its simple and intuitive interpretation. Decision tables are a compact but precise way to model decision problems. One of the main disadvantages is that such tables can be extremely large for complex decision problems making the process complicated. The approaches of decision trees and decision tables have certain limitations. These approaches are not the most suitable for complex decision problems, especially for automating the decision making process. They are also weak in quantifying the uncertainties involved in a decision making process. Hence it is necessary to investigate another decision model for this research.

1.2 Multicriteria Decision Making

MCDM is a highly structured, disciplined and formal approach to decision making (Deer et al. 2001). It consists of evaluating the alternatives in the given set A against the set C of criteria (Zeleny 1982). The global evaluation of the alternatives is then determined by aggregating the individual evaluations of the alternatives against the criteria. This global evaluation could be used for selection of the best possible alternative or for ranking the alternatives (Marichal 1999). The general MCDM process can be explained by a graphical representation based on Malczewski's flow process chart on MCDM (Malczewski 1999). The following notations are used.

- A set of 'm' alternatives is represented as $A = \{A_1, A_2, ..., A_m\}$. A set of 'n' criteria is represented as $C = \{C_1, C_2, ..., C_n\}$ where n > 1, m >1.
- The scores or individual evaluation of alternatives in terms of individual criteria are represented as x_{ij} where $i = \{1,2,......... .., m\}$ and $j = \{1,2,........, n\}$. The scores are commonly represented in the scale [0,1]. The scores model the degree of satisfaction of a particular criterion j for the given alternative i. The degree of satisfaction is represented using the fuzzy set theory (Zadeh 1965).

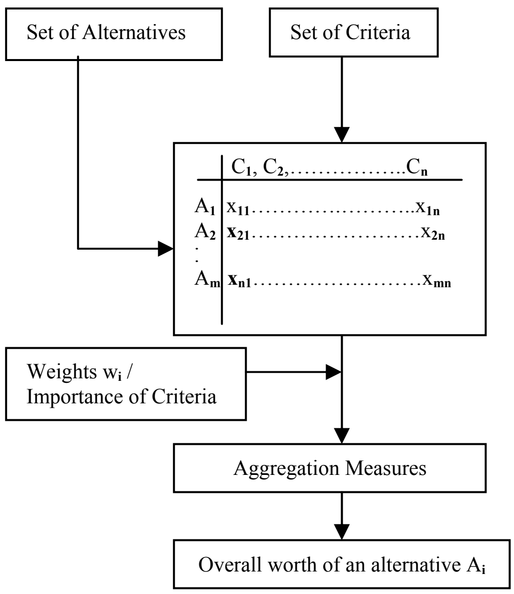


Figure 1.2 MCDM Process (adopted from (Malczewski 1999))

In the above process, the criteria can be of varying degree of importance. This importance can be modelled by weights w_i on the criteria i. For a set of n criteria, there will be 2^n-1 sub-sets of criteria. The weights of these sub-sets should also be modelled as well as these weights contribute in determining overall worth or global evaluation of an alternative.

1.3 Advantages of MCDM

It is suggested that it is very difficult to model the human decision making process in an objective manner (Olson 1996). Human preferences are difficult to measure objectively, so a model which has the ability to closely resemble the human decision making process should be chosen. The approaches such as decision trees and decision tables do not consistently resemble the human decision making process. The MCDM model described earlier certainly has the ability to represent the human preferences in a formal and systematic way. By using this approach, one can model all kinds of decision problems (Tsoukiàs 2003). Based on the study of the relevant literature (Hoffer 1996), the decision models can be compared. The literature compares the rule-based approaches such as decision tables and decision trees. This work has extended the comparison by adding MCDM model in the comparative analysis of decision models. The following comparison is based on the various criteria shown in the table below.

Table 1.2 Comparison of the rule-based decision models

Sr. No	Criteria	Decision Tables	Decision Trees	MCDM
1	Modelling complex problems	Worst	Worst	Best
2	Modelling simple problems	Worst	Best	Best
3	Making decisions	Worst	Best	Best
4	More compact	Best	Worst	Best
5	Easier to manipulate	Best	Worst	Best
6	Resemble human decision making	Worst	Worst	Best

From the above comparison, it can be observed that the MCDM model can certainly be applied for modelling all kinds of decision problems. This ability justifies the choice of MCDM for the undertaken research question.

1.4 Motivation

There are various decision-making models presented in the literature (Anderson et al. 2000, Klein & Calderwood 1989). As discussed earlier, MCDM is a systematic and formal model of decision-making. Multi-criteria decision-making (MCDM) is one of the well-known branches of decision-making. The area of MCDM has developed rapidly (Hwang & Yoon 1981). Some bibliographic surveys (Steurer et al. 1996) at the international level suggest that this can be seen by the significant increase in the importance of MCDM in the research community, in the form of increase in number of conferences, books, volumes, refereed journal articles (Martel 1999; Miettinen 1999; Roy 1996; Roy & Vanderpooten 1996). The undertaken research work also relates to one of the several issues associated with the MCDM model. This work is the extension of the previous work where the issue of dependence of criteria was raised (Wagholikar 2002).

1.4.1 Issue of Dependence between Criteria

In MCDM problems, the issue of trade-off between the alternatives is considered significantly important (Keeney and Raifa 1976). It was explained earlier that in the MCDM model, weights model the importance of criteria. One point is worth considering here: this is of the importance of sub-sets of criteria. The worth of an alternative is determined by individual criteria as well as sub-sets of criteria. The importance of sub-sets of criteria will depend on the nature of interaction between the criteria.

According to the literature (Grabisch 1995a), there can be positive or negative dependence between the criteria. Accordingly, criteria can be sub-additive or super-additive. If w_A & w_B are the individual importance of criteria A & B respectively and w_{AB} is their combined importance then the combined importance of criteria can be shown as below.

$w_{AB} < w_A + w_B$ ……………(i)

$w_{AB} > w_A + w_B$ ……………(ii)

$w_A * w_B \leq w_{AB} \leq w_A + w_B$ …..(iii)

where $w_A = 0\sim1$ and

$w_B = 0\sim1$

Case (i) is the case of weak dependence between the criteria. This means that criteria are considered less important when considered together than when they are considered separately; for example, fuel economy and power could have weak dependence among them especially in a case where decision maker prefers a fuel efficient vehicle. For a fuel efficient vehicle, high power is not expected.

Case (ii) is the case of strengthening dependence between the criteria. This means satisfaction of criteria A & B considered individually is not as important as their combined importance. The criteria support each other in combination with each other. In the literature, the above phenomenon is generally represented by the terms “dependence”, “interaction” and “relationship”. In this thesis, these terms are used interchangeably.

In this problem, acquisition of weights by an intelligent decision support system continues to be a necessary area to investigate (Klir et al. 1997). If the MCDM model is used for decision-making situations where the same decisions are made as in the past, then the same weights can be acquired for the new decision-making.

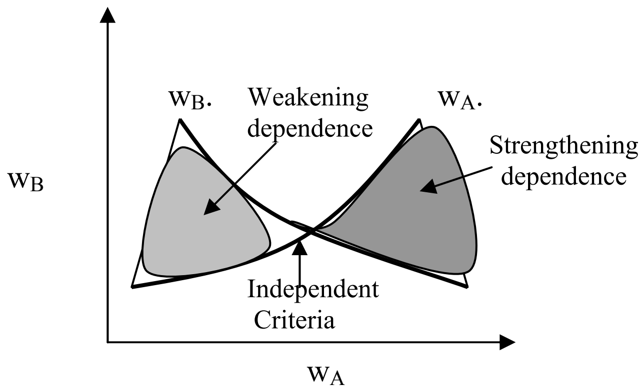


Figure 1.3 Interaction between criteria

The above figure represents the two curves modelling the weights w_A and w_B of the criteria A and B respectively. The combined importance of A and B increases with the increasing weights after the point of independent criteria, if the criteria interact with each other positively. The combined importance decreases if there is a negative interaction between the criteria.

1.4.2 Discussion

It is important to model the interaction between the criteria. The process of aggregation should take into account dependence between criteria. The non-additive integral, i.e. fuzzy integral, is an incremental summation of the product of the fuzzy measures and criteria scores. It has the ability to model the combined importance of criteria through fuzzy measures (Grabisch 1995a; Marichal 1999; Wagholikar 2002). Non-additive measures such as fuzzy measures and fuzzy integrals are effective in addressing the issue of dependence between criteria (Grabisch 1995a).

However, the application of the fuzzy integrals is an important issue as it involves determination of 2^n-1 coefficients of fuzzy measures. This determination makes the computation complex, as the number of criteria 'n' increases. It also reduces the use of fuzzy measures to reasonable values of 'n' where n >1 (Grabisch 1995a; Grabisch & Modave 1997). The central problem of determining fuzzy measures is important in practical applications (Klir 1997). There are a few methods available to deal with this problem, but their quality is hard to determine. Thus the question, "Which is the best method to solve the fuzzy measure determination problem?" has become one of the most important and challenging ones. The main problem here is neither computational cost in terms of memory and computational time, nor the computational complexity, but actually how to acquire the coefficients of the fuzzy measure in a practical situation, whatever their number is and to model this in an applicable algorithm. This is a really important problem and many researchers have pointed out this importance (Marichal & Roubens 2000). There are many practical issues associated with it. The approaches suggested by the various researchers involve the use of learning data about the same past decisions (Grabisch 1995a; Grabisch & Modave 1997; Marichal 1999; Miettinen 1999). Whenever an entirely new decision is to be made and the decision problem is to be solved under the MCDM setting in the presence of interactive criteria, then the form of learning data plays a significant role. In many instances, the decisions are based on the data about the same past decisions. In this case, determination of importance of a sub-set of criteria in the given set of criteria is an interesting issue. This is the main reason for this research and is the motivation behind it.

1.5 Research Objective

It has already been mentioned before that non-additive measures or fuzzy integrals can model the dependence of criteria. To effectively use the fuzzy integrals, the main challenge is how to determine the fuzzy measures. Various researchers have proposed various approaches to determine the fuzzy measures. These approaches use a training data about the same past decisions. This thesis proposes a new approach, which allows determination of fuzzy measures in the presence of similar data sets. Thus, the main research objective of this work is to investigate various methods of fuzzy measure acquisition and propose a method using similarity-based reasoning.

1.6 Original Contribution

- The research investigates the issue of interaction or dependence between the criteria in detail. Non-additive measures are important tools to handle the problem of dependence between criteria. This research addresses the practical issues associated with the determination of non-additive measures in the context of decision-making. It brings knowledge of the issue of determination of non-additive measures to a new stage of development.
- This research furthers the state-of-the art on the issue of determination of non-additive measures. Thinking at present suggests the use of past precedents for determining the non-additive measures. This work suggests use of reasoning by analogy for determination of fuzzy measures. Thus, it combines two important theories, the MCDM theory and reasoning by analogy to solve the problem of fuzzy measure determination. Such combination of two different theories justifies the original contribution of this work. It proposes a new way of thinking towards the problem of fuzzy measures determination.
- This work investigates various similarity measures. It illustrates the application of semantic similarity measures for solving the problem of fuzzy measure acquisition. It proposes a new semantic similarity measure and illustrates its application for solving the selected research problem.
- This work compares the state-of-the art approaches using a common data set. Such a unique comparison also illustrates the significance of the proposed approach and this research.

1.7 Thesis Structure

This chapter has introduced the research question. This work is divided into the following chapters.

Chapter 2 reviews previous literature and research findings related to the research question. It discusses in detail concepts that are important for the undertaken research. The knowledge obtained from this literature review not only contributes to the research process but also reveals differences between previous research and this study, thereby confirming the significance of this study. Some additive measures like OWA operators, which have gained significant recognition since their introduction by Yager (1988), will be explained. This chapter will discuss the concept of fuzzy measure and fuzzy integrals in detail as well as properties of these measures. The chapter subjects to critical analysis of various approaches suggested by various researchers. This will focus on some of algorithms proposed for determination of fuzzy measures using combination of learning data and semantic consideration about criteria. This chapter will also discuss the theoretical foundations of the proposed methodology. The area of analogical reasoning and Case-based reasoning and its suitability in the context of the research question are explained. The concept of similarity, various similarity measures and their suitability in the context of the undertaken research are discussed.

Chapter 3 presents the methods proposed to solve the research question and evaluates their aptitude for solving the research question. This chapter shows how the proposed interactive model of acquisition of fuzzy measures is an appropriate solution for the selected task.

Chapter 4 discusses the implementation of the proposed methodology in a case-based decision support system.

Chapter 5 compares the proposed approach with some existing approaches. It discusses the results of the comparative analysis. This chapter summarizes the results of this study and its significance to the chosen field of research.

Chapter 6 concludes the undertaken study. It will also throw light on the possible future research work.

1.8 Summary

This chapter has identified the undertaken research question. The primary focus of this research is to resolve the issue of dependence between the criteria in MCDM problems. Fuzzy measures and fuzzy integrals are potential tools to model the dependence or interaction between the criteria. The challenge lies in developing approaches for the practical application of these tools. This work addresses the issue of acquisition of fuzzy measures in practical applications. This will not only address the issue of dependence or interaction between the criteria but will also propose approaches to simplify the difficulties in the practical applications. The remainder of the thesis will discuss the important concepts relevant to the research question, critical analysis of the state-of-the-art literature, the proposed methodology and its outcomes.

Chapter 2 Literature Review

2.1 Introduction

The desired outcome of a MCDM-based decision support system is the overall worth of the given alternatives against a given set of criteria. The determination of the global evaluation of alternatives is primarily based on the data aggregation methods. These methods have different characteristics. This chapter presents an investigation of various aggregation methods or measures that will form the foundation of this work. It will also justify the choice of fuzzy measures for resolving the issue of dependence between the criteria. In the first chapter, the problems of dependence between criteria and the problem of acquisition of fuzzy measures were introduced. These problems will be further explored. The important and necessary concepts will be discussed. The main aim of this literature review is to investigate the relevant approaches to resolve the research question. This investigation will critically analyse the commonly used approaches. The concepts related with proposed methodology will be discussed in detail. The theory of similarity-based reasoning is central to the proposed approach. It will explain the use of case-based reasoning as a tool in developing approaches for the acquisition of fuzzy measures in practical applications. Thus, this chapter will establish a theoretical foundation of the proposed approach.

2.2 Established MCDM Approaches

MCDM is a specific perspective to deal with decision making problems. There are certain commonly used MCDM approaches (Miettinen 1999; Fishburn & Lavalle 1999) that propose a specific perspective towards solving MCDM problems. These methods include analytical hierarchy process, outranking methods, multiattribute utility and value theories, MACBETH (Measuring Attractiveness by a Categorical Based Evaluation Technique), UTA methods (Figueira et al. 2005). These approaches are applied for solving decision problems under the MCDM setting. The analytic hierarchy process (AHP) (Satty 1980) and outranking are the most common approaches. Hence they are assessed for solving the undertaken research problem outlined earlier. AHP is the original theory of prioritization. It derives relative scales of absolute numbers known as priorities from judgements expressed numerically on an absolute fundamental scale. It is a descriptive approach to decision making. The outranking methods compare all feasible alternatives or actions by pair building up some binary relations, crisp or fuzzy, and then to exploit in an appropriate way these relations in order to obtain final recommendations.

2.2.1 Analytical Hierarchy Process

The analytic hierarchy process is considered as a powerful decision making process. This approach is considered useful in decision problems involving qualitative as well as quantitative aspects. The Analytic Hierarchy Process (AHP) is a powerful and flexible decision making process to help people set priorities and make the best decision when both qualitative and quantitative aspects of a decision need to be considered. This approach simplifies complex decisions to a series of one-on-one comparisons and then synthesizes the results of those comparisons to provide decision aid. It is considered suitable for complex decision elements, which are difficult to quantify. It is based on the assumption that when faced with a complex decision the natural human reaction is to cluster the decision elements according to their common characteristics. There are various software applications based on AHP (Ehrgott et al. 2005). The main advantage of this approach is that it not only helps decision makers arrive at the best decision, but also provides a clear rationale that it is the best. A general representation of AHP is shown in the following figure. The overall objective i.e. the outcome of the decision problem is represented as the hierarchy of individual comparisons alternatives against a given set of criteria.

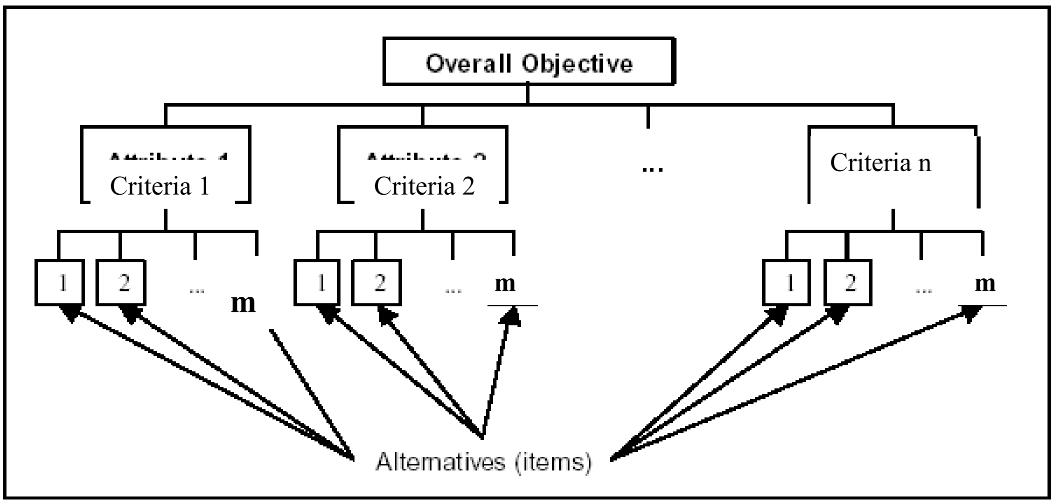


Figure 2.1 General Representation of AHP

AHP is mainly based on pair wise comparison. This approach helps the decision maker to structure the important components of a problem into a hierarchical structure similar to a family tree. The AHP process consists of the following steps.

1. Form the m x n decision matrix, involving a set 'm' of alternatives and set 'n' of criteria.
2. Form a comparison matrix of each alternative against each criterion using the scale shown below.

Table 2.1 AHP's pair wise comparison scale

Verbal Judgment of Preference	Numerical Rating
Extremely Preferred	9
Very Strong to extremely	8
Very Strongly preferred	7
Strongly to very strongly	6
Strongly preferred	5
Moderately to Strongly	4
Moderately Preferred	3
Equally to moderately	2
Equally Preferred	1

3. Synthesize the scores by using the above pair wise comparison
4. Calculate the overall priority ranking. The overall priority ranking can be calculated by the weighted average of the individual priorities for each criterion.

After studying the relevant literature (Satty 1980), the following disadvantages were observed.

1. This method considers comparison of alternatives against individual criteria. However, it does not consider comparison of alternatives against sub-set of criteria.
2. Overall priority ranking does not take into account the importance of sub-sets of criteria. It does not address the dependence issue.
3. The "pair wise comparison" technique gets complicated if there are more than a few alternatives. If there are only four alternatives, there are six pairs to consider (4*3/2). However, for 20 alternatives, there would be 190 pairs.

2.2.2 Outranking Methods

Roy (1968) proposed the idea of outranking in solving decision problems. The basic idea of an outranking method can be described as follows.

Let $C_1, C_2, \ldots\ldots\ldots\ldots, C_n$ and $A_1, A_2, \ldots\ldots\ldots\ldots, A_m$ be the set of criteria and set of alternatives respectively. Let $w_1, w_2, \ldots\ldots\ldots, w_n$ be the weights assigned to the criteria and let score a_{ij} represent the performance of the alternative A_i against a criterion C_j. It is assumed that higher score means better performance since any goal of minimization can be easily transformed into goal of maximization.

Consider an ordered pair of alternatives (A_i, A_j) for a given decision problem. Alternative A_i outranks A_j if on a great part of the criteria for the given decision problem, A_i performs at least as good as A_j (concordance condition), while its worse performance is still acceptable on the other criteria (non-discordance condition). After having determined for each pair of alternatives whether one alternative outranks another, these pair wise outranking assessments can be combined into a partial or complete ranking. There are different outranking methods (Vincke 1992; Zopounidis & Doumpos 2002). ELECTRE methods are commonly used outranking methods. These methods are based on the concordance and disconcordance indices. For an ordered pair of alternatives (A_j, A_k), the concordance index c_{jk} is the sum of all the weights for those criteria where the performance score of A_i is at least as high as that of A_j i.e.

$$c_{ij} = \sum_{i:a_{ij} > a_{jk}} w_i \quad j,k=1,\ldots,n, j \neq n \tag{2.1}$$

The concordance index lies between 0 and 1. The discordance index is computed as follows.

For each criterion where A_k outperforms A_j, the ratio is calculated between the difference in performance level between A_k and A_j and the maximum difference in score on the criterion concerned between any pair of alternatives. The maximum of these ratios is the discordance index. It is represented as follows.

$$d_{ij} = \max_{i=1,\ldots,n} \frac{a_{ik} - a_{jk}}{\max\limits_{j=1,\ldots,m} a_{ij} - \min\limits_{j=1,\ldots,m} a_{ij}}, \text{j,k=1,}\ldots\text{,n,j} \neq \text{k} \quad (2.2)$$

After computing the concordance and disconcordance index, a concordance threshold c* and discordance threshold d* are defined. Then, an alternative A_j outranks alternative A_k, if the $c_{jk} > c^*$ and $d_{jk} < d^*$, i.e. the concordance index is above and the discordance index is below its threshold, respectively.

This outranking defines a partial ranking on the set of alternatives. Consider the set of all alternatives that outrank at least one other alternative and are themselves not outranked. This set contains the promising alternatives for this decision problem.

The ELECTRE I method is used to construct a partial ranking and choose a set of promising alternatives. ELECTRE II is used for ranking the alternatives. In ELECTRE III an outranking degree is established, representing an outranking creditability between two alternatives, which makes this method more sophisticated. This method is more complicated and difficult to interpret (Figueira et al. 2005).

The ELECTRE methods are relevant when facing decision situations when the decision maker wants to include in the model at least three criteria. However, the outranking does not consider any interaction or dependence between criteria. It is purely based on the performance of an alternative against a given set of criteria. The concordance and dis-concordance index does not take into account the weights of sub-sets of criterion. Ultimately, the outranking methods do not address the problem of dependence between criteria. The above MCDM approaches i.e. AHP and Outranking methods deal with determining the global evaluation of alternatives in a given decision problem. This output is achieved through the process of aggregation. Aggregation is a central concept in the process of determining the global evaluation of an alternative (s).

2.3 Aggregation

Aggregation is a central activity in human reasoning in which the available facts are combined and synthesized. This aggregation process is applied in many disciplines such as decision-making, pattern recognition, data mining, information retrieval, neural networks and expert systems (Marichal 1999).

In everyday life, there are various decision making situations where one has to make a decision out of the various choices with each choice having multiple criteria. Under this situation of multi-criteria decision-making, a human reasoning process involves assessment of multiple criteria while making a decision.

The general idea of the aggregation process can be explained by the following example. Consider a computer-buying situation, where there are various alternatives such as IBM, Compaq, and Dell etc. There are multiple criteria of cost, system configuration, and reliability for this decision problem. The alternatives are assessed against these multiple criteria. During this reasoning process, these criteria are considered in terms of their importance, according to the decision-maker's requirements about the best possible alternative out of the given alternatives. A decision-maker looking for a computer with less cost, good system configuration and high reliability will consider the criteria of cost, system configuration and reliability according to the above requirement. The overall worth of each alternative will be evaluated against the criteria as per the decision-maker's requirement by combining the criteria evaluations. This process of combining the criteria evaluations to determine the overall worth of an alternative is termed as aggregation.

In a purely mathematical setting, the scores can be expressed in terms of numeric values. Thus, in this case the aggregation results in a single numeric value, which is the global evaluation of the given alternative against the given set of criteria. The aggregation procedure in a quantitative setting can be defined as below.

Consider a set of positive numeric values $x_1, x_2, \ldots x_n$ in the range of 0 and 1. Aggregation refers to the process of combining numerical values $x_1, x_2, \ldots x_n$ into a single value say X, so that the final result of aggregation takes into account all the individual values (Marichal 1999). In the MCDM context, the set of criteria scores represent the set of numeric values and X is the global worth of the alternative.

2.4 Aggregation Operators

The notion of multicriteria decision making was explained earlier. The importance of the aggregation process in global worth determination was emphasized. The input-process-output model of the aggregation process can be viewed as below.

Figure 2.2 Aggregation Measure Function

Thus the main processing module is aggregation measure. In most of the literature, this concept is termed as "aggregation operator" (Dubois & Prade 1985; Detyniecki 2000). It can be termed as aggregation measure as it measures the overall worth of an alternative. Both the terms will be used interchangeably as they convey the same meaning. Its general definition can be explained as below.

Consider n-tuples of objects belonging to a given set. Then the aggregation process results into a single object of the same set. This process can be defined in terms of a measure called as "Aggregation operator". In the case of a mathematical aggregation, consider a set of real numbers, which are to be aggregated. In this setting, an aggregation measure is simply a function f, which assigns a real number x to any *n*-tuple (x_1, x_2,..… x_n) of real numbers (Detyniecki 2000). Mathematically, it can be represented as follows-

$$X = f(x_1, x_2, \ldots\ldots\ldots\ldots x_n) \quad (2.3)$$

The properties of aggregation operators are explained in the literature (Dubois & Prade 1985; Marichal 1999).

2.4.1 Desirable Properties of Aggregation Operators

The aggregation operators possess various properties (Marichal 1999). It is important to consider the characteristics of the decision problem. An aggregation operator that has desirable properties can be applied for determining the global evaluation of alternative(s). The choice of an appropriate aggregation measure for modelling the interaction between the criteria will depend on its properties. The aggregation operator should possess the desired properties as needed for modelling the dependence between the criteria. It is necessary to understand various properties of aggregation operators so that they can be assessed for the applicability in resolving the undertaken research issue. This section discusses the various desirable properties of the aggregation operators.

2.4.1.1 Boundary Conditions

This property deals with upper and lower limits of aggregation operators. It characterizes the behaviour of the aggregation operator in the best and in the worst cases. An aggregation operator should satisfy the following conditions.

$f(0,0,..,0) = 0$ … (i)

$f(1,1,..,1) = 1$ … (ii)

The boundary condition (i) implies that the total aggregation is bad or not satisfactory if all the criteria in the MCDM setting are found to be bad or false or not satisfactory. This is the lower bound of aggregation. The boundary condition (ii) indicates that if all the criteria in the MCDM setting are found to be completely satisfactory or true, then the total aggregation is also true or satisfactory. This is the upper bound of aggregation. These boundary conditions are fundamental in the definition of aggregation operators (Mesiar & Komorníková 1997).

2.4.1.2 Monotonicity (non-decreasing)

The property of monotonicity means that if an argument or attribute value in the given set of attribute increases then the final value of the aggregation also increases or at least it remains equal. The aggregated value of a set of attribute values is greater than or equal to the aggregation value of the second set of values with the corresponding attribute value less than or equal to the attribute value in the former set. In a mathematical setting this can be explained. Consider a set of numeric values x_1, $x_2,\ldots\ldots,x_n$. If y_1 is another variable added in the set and y_2 is other variable in the set, then the aggregation with the variable y_1 is greater than the variable y_2

$$f(x_1, x_2, y_1, \ldots,x_n) \geq f(x_1,x_2,\ldots, y_2,\ldots,x_n) \text{ where } y_1 \geq y_2 \quad (2.4)$$

The above property shows that whenever there is an addition of criteria with its numeric value greater than or equal to its corresponding value in the other set of criteria, then there is increase in the aggregation.

2.4.1.3 Continuity

The aggregation function f is continuous with respect to each of its variables. There exists the aggregated value with every variable in the given set of attribute values.

2.4.1.4 Associativity

Aggregation operators are associative. This implies that if the arguments and small subsets in the given set are aggregated, then the order of the arguments and the subsets in the aggregation makes no effect on the aggregation.

$$f(x_1, x_2, x_3) = f(x_1, f(x_2, x_3)) = f(x_2, f(x_1, x_3)) \quad (2.5)$$

This property implies that even if the importance of criteria is interchanged the aggregation remains same.

2.4.1.5 Commutativity

This property is also known as symmetry or anonymity (Dubois and Prade 1985). This property implies that the aggregation remain same even if the order of the arguments change. This property is desirable for combining criteria of equal importance (Detyniecki 2000; Sharma 2001). For every permutation σ of $\{1,2,\ldots.\ n\}$ the operator satisfies:

$$f(x_{\sigma(1)}, x_{\sigma(1)}, \ldots\ldots\ldots, x_{\sigma(1)}) = f(x_1, x_2, \ldots\ldots, x_n) \quad (2.6)$$

2.4.1.6 Bisymmetry

This property of bisymmetry is applicable to the aggregation for n-ary operators where the inputs are equal to n^2. If these inputs are written in a square matrix (m x n) then this property implies that it does not matter whether the first column variables are aggregated and then the outputs of thereof, or vice versa, first row vectors are aggregated and then the relevant outputs. For a binary operator Y this means for all $x_{11}, x_{12}, x_{21}, x_{22}$, that:

$$Y\,(Y(x_{11}, x_{12}), Y(x_{21}, x_{22})) = Y\,(Y(x_{11}, x_{21}), Y(x_{12}, x_{22})) \quad (2.7)$$

It is to be noted that if an operator is commutative and associative then it is necessarily bisymmetric, but not necessarily commutative or associative.

2.4.1.7 Absorbent Element

If the aggregation operator has an absorbent element *a,* then it can be used like an eliminating score or like a veto. It can also be considered as a qualifying score:

$$Y\,(x_1,x_2,\ldots, a, \ldots,x_n) = a \quad (2.8)$$

This element is also called annihilator (Detyniecki 2000).

2.4.1.8 Neutral Element

If the operator of aggregation has a neutral element *e,* then it can be used to be associated with an argument that should not have any influence on the aggregation:

$$f^{[n]}\,(x_1,x_2,\ldots, e, \ldots,x_n) = f^{[n-1]}\,(x_1,x_2,\ldots, y_1, \ldots,x_n) \quad (2.9)$$

2.4.1.9 Idempotence

This property is also known as unanimity or agreement. This property implies that if a same value is aggregated *n* times, then the aggregation will result in the same original value.

2.4.1.10 Reinforcement

Some authors have pointed out a characteristic of many types of human information processing known as full reinforcement (Elkan 1994; Yager 1988). This property indicates the tendency, on one hand, of a collection of high scores to reinforce each other to give a resulting score more “affirmative” than any of the individual scores alone and on the other hand, the tendency of a collection of low scores to reinforce each other to give a resulting score more "disfirmative" than any of the individual scores. The first concept is called upward reinforcement and the second concept is called downward reinforcement. The property of reinforcement can be very interesting. For example, in medical diagnosis the appearance of a number of symptoms indicative of a disease will make us more confident in diagnosing a patient as having the disease than any symptoms alone while the lack of appearance of this symptoms will make us more confident diagnosing a patient as not having the disease.

2.4.1.11 Stability for a Linear Function

This property translates a stability of the operator for a change of measurement scale:

$$f\,(r.x_1 + t, r.x_2 + t, r.\,x_n + t) = r.\,(\,f\,(x_1 + x_2 + \ldots\ldots + x_n) + t \qquad (2.10)$$

2.4.1.12 Weights on the Arguments

It is crucial to have the possibility to express weights on the arguments. Weights on the arguments model the importance of criteria. The undertaken research questions focuses on modelling interaction between the criteria. Criteria weights are central in modelling the dependence or interaction between the criteria. Therefore aggregation measure must have the ability to incorporate criteria weights in computing the overall global evaluation of alternatives. This property is very important, the importance of criteria or attributes should be known for evaluating the alternatives in MCDM especially in procedures where criteria weights contribute in determining overall global evaluation of alternatives.

2.4.1.13 Decisional Behaviour

It is useful to have the possibility to express the behaviour of the decision-maker. For example: tolerant, optimistic, pessimistic or strict. These behaviours are usually named disjunctive and conjunctive behaviours (Dubois and Prade 1985). In order to model the behaviour of the decision-maker, the behavioural properties of the decision-maker need to be addressed in detail.

2.4.1.13.1 Conjunctive nature

Whenever an alternative has to be evaluated against multiple criteria, then many times the decision-maker is interested in satisfaction of all the criteria. As all the criteria may be important for evaluating the alternative. In a quantitative setting, it can be explained as follows.

Consider a set of numeric values $x_1,x_2,\ldots\ldots,x_n$ and $X = f(x_1,x_2,\ldots\ldots, x_n)$.

If $X \leq Min(x_1, x_2, \ldots\ldots x_n)$ then the aggregation is of conjunctive nature i.e. the global evaluation of the values requires that all the values in the given set must be equal to or greater than X. In multiple criteria setting, this means that all the criteria must be satisfied. Hence the aggregation has "And-like" behaviour (Dubois and Prade 1985).

2.4.1.13.2 Disjunctive nature

There are certain situations where the decision-maker wants at least one criterion out of the given criteria to be satisfied. This behaviour is exactly opposite to the conjunctive behaviour. Mathematically, if $X \geq Max(x_1, x_2, \ldots\ldots,x_n)$, then it implies that at least one criterion must be satisfied. Here the aggregation has "OR-like" nature.

2.4.1.13.3 Compensatory nature

This is an important property for aggregating interactive criteria. According to this property, the result of the aggregation is lower than the highest element aggregated (the maximum denoted as Max) and higher than the lowest one (the minimum denoted as Min) (Dubois and Prade 1985).

Thus,

$$Min(x_1,x_2,\ldots\ldots,x_n) \leq f(x_1,x_2,\ldots\ldots,x_n) \leq Max(x_1,x_2,\ldots\ldots,x_n) \quad (2.11)$$

This is the most important property of aggregation. In multicriteria setting, higher satisfaction of one criterion may influence the lower satisfaction of other criteria. The higher satisfaction of a particular criterion becomes dominant on the satisfaction of other criteria. For example, in the computer selection problem, the decision-maker is looking for a low cost computer for its business purpose and he finds that IBM computer is very cheap though its system configuration is minimal. However, since the decision maker is interested in buying a low cost computer, its system configuration will not bother their decision-making in the given situation and eventually they may prefer IBM PC. In this example, the criterion of cost is highly satisfied and the criterion of system configuration has lowest satisfaction. In other words the “cost criteria” has compensated the “system configuration” criteria. In general terms, a higher satisfaction of criteria can compensate for a lower satisfaction of another criteria (Fuller 1997). In a quantitative setting, the compensation means that the effect of higher scores cancels the effect of lower scores. Under such situation the aggregation of criteria aggregation will lie between the highest values in the given set i.e. the maximum value denoted by “max” and the lowest value in the set i.e. the minimum value denoted by “min”. In other words, the result of aggregation will be an intermediate value. Aggregation Operators with compensatory behaviour are called “Compensatory Operators” (Sharma 2001).

2.4.2 Notion of Compensation Vs Dependence between Criteria

It is important to mention the difference between the concepts of compensation and dependence between criteria. The concept of compensation deals with the result of aggregation lying between min and max. This does not necessarily indicate any interaction among the individual criteria in a given set of criteria. The concept of dependence of criteria however deals with modelling the importance of sub-sets of interactive criteria. The nature of interaction (i.e. positive or negative) between the individual criteria governs the importance of sub-sets of criteria (Keeny and Raiffa 1976). Compensation is the property of an aggregation operator whereas dependence of criteria is the property of sub-sets of criteria.

2.4.3 Types of Aggregation Operators

There are various aggregation operators (Dubois and Prade 1985). The aggregation operators can be broadly classified as additive measures and non-additive measures. Additive measures are the most commonly used aggregation measures. After describing the aggregation properties, some aggregation operators are discussed with their specific properties.

2.4.3.1 Minimum

The minimum operator commonly called as “min” gives the smallest value of a set. As it satisfies the axioms of the definition of aggregation operator, it can be considered as the aggregation operator. The main properties of the min operator are as below. These properties were explained earlier.

- Monotonicity
- Symmetry
- Associativity
- Idempotence.

If there is a restricted interval [*a, b*] the minimum has for absorbent element *a* and for neutral element *b*, the minimum has a conjunctive behaviour. Since the possibility of giving weights is important, the weighted minimum was introduced (Detyniecki 2000). If w_1, w_2………w_n are the weights given on the set of values $x_1, x_2, \ldots x_n$ then the weighted minimum is given as

$$\min_{w1,w2 \ldots\ldots wn} (x_1, x_2, \ldots\ldots x_n) = \min_{i}^{n} [\max(1-w_i, x_i)] \qquad (2.12)$$

2.4.3.2 Maximum

The maximum operator commonly called as "max" gives the largest value of a set. As it satisfies the axioms of the definition of aggregation operator, it can be considered as the aggregation operator.

Similar to min operator, the main properties of the max operator are as below.

- Monotonicity
- Symmetry
- Associativity
- Idempotence

If there is a restricted interval [*a, b*] the maximum has for absorbent element b and for neutral element a, the maximum has a disjunctive behaviour. Similar to the weighted minimum, the weighted maximum is given as

$$\max_{w1,w2 \ldots\ldots wn} (x_1, x_2, \ldots\ldots x_n) = \max_{i}^{n} [\min(w_i, x_i)] \qquad (2.13)$$

Mathematically speaking, both the minimum and maximum operators have compensation behaviour, but these are the limit cases. Using these operators one can never obtain an aggregated value "in the middle". For this reason, the compensation behaviour of these operators cannot be discussed in this case.

2.4.3.3 Arithmetic Mean

This is the most simplest and common aggregation operator. This is used very frequently in the aggregation. This is the common way to aggregate the numeric values in the given set (Dubois and Prade 1985).

Mathematically, if (x_1, x_2, x_3,…., x_n) is the given set of values then the arithmetic mean is given by

$$X = \frac{(x_1 + x_2 + \dots\dots + x_n)}{n} \tag{2.14}$$

The mean operator is compensatory. It gives an aggregated value that is smaller than the greatest argument and bigger than the smallest one. Therefore, the resulting aggregation is "a middle value". Mathematically this can be represented as following:

$$min(x_1,x_2,\dots\dots,x_n) \le f(x_1,x_2,\dots\dots,x_n) \le max(x_1,x_2,\dots\dots,x_n) \tag{2.15}$$

The mean operator is very simple to use. Apart from being compensatory, it satisfies the following properties of aggregation.

- Monotonicity
- Continuity
- Symmetry
- Associativity
- Idempotence

The classical extension of the mean operator is known as "Weighted Average" (WA) (Fuller 1997). Weighted average allows placing weights on the arguments. This implies that individual importance of criteria for a given set of alternatives is considered while evaluating the given alternative. If w_1, w_2, w_3,……,w_n are the weights assigned to the set of criteria whose attribute values are x_1, x_2, x_3………, x_n then the weighted average is given by

$$WA = \frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i} \qquad (2.16)$$

2.4.3.4 Quasi-arithmetic Means

The extensions of the arithmetic means are called Quasi-arithmetic means. The two important quasi-arithmetic means are Geometric means and Harmonic means.

Geometric mean:

A general mathematical definition of the geometric mean is:

For a set of attribute values of $x_1, x_2,\ldots\ldots,x_n$ the geometric mean is given by

$$\text{Geometric}_{x1,x2,\ldots\ldots,xn} = (\prod_{i}^{n} x_i)^{1/n} \qquad (2.17)$$

Harmonic Mean:

For a set of attribute values of $x_1,x_2,\ldots\ldots,x_n$ the harmonic mean is given by

$$\text{Harmonic}_{x1,x2,\ldots\ldots,xn} = 1/n \;\; (\sum_{i}^{n} 1/x_i) \qquad (2.18)$$

2.4.3.5 Median

Similar to the mean operator, the median operator is also compensatory. This operator also gives an intermediate value as the aggregated value. It consists in ordering the arguments from the smallest one to the biggest one and then taking the element in the middle.

If the cardinality of the set of arguments is not odd then there is not a middle argument but a pair, then the mean of the middle pair is considered as the median.

If $\{x_1, x_2, x_3,\ldots,x_n\}$ is the given set of values, then the median is given as

$$\text{Med}\{x_1,\ldots,x_n\} = x_{(n+1)/2} \qquad \text{if n is odd}$$

$$= \tfrac{1}{2}(x_{(n/2)}+x_{((n/2)+1)}) \qquad \text{if n is even}$$

Where parenthesis {} around the index show that elements have been arranged in ascending order, i.e. $x_1 \leq x_2 \leq_{\dots} \leq x_n$. This aggregation operator satisfies the boundary conditions, the monotonicity, the symmetry, the idempotence and evidently the compensation behaviour. There exists a generalization of this operator known as the k-order statistic, with which one can choose the element on the k^{th} position on the ordered list (from the smallest to the biggest element). The extension of the median operators is the "weighted median".

Let, $\{x_1, x_2, x_3, \dots, x_n\}$ be the given set of values, which are in ascending order.

Let $(w_1,w_2,\dots\dots,w_n)$ be the corresponding weights, then the collection of pairs of weights with its corresponding value is as follows.

$D = \{(x_1,w_1),(x_2,w_2)\dots\dots(x_n,w_n)\}$ & let the sum of the first *i* weights be denoted as

$$S = \sum_{i=1}^{n} w_i \tag{2.19}$$

From this the median (D) = x_k

Where k is such that $x_{k-1} \leq 0.5$ & $x_k \geq 0.5$

Thus the weighted median is the ordered value of arguments for which the sum of the weights first crosses the value of 0.5. The weighted median can be illustrated through the following example. Let there be a set of ordered arguments $X = \{x_1,x_2,x_3,x_4,x_5\}$ where-

$x_1 > x_2 > x_3 > x_4 > x_5$ and the corresponding weights be $W = \{w_1,w_2,w_3,w_4,w_5\} = \{0.3,0.2,0.1,0.2,0.2\}$.Thus the output is,

$$D = \{(x_1, 0.3), (x_2, 0.2), (x_3, 0.1), (x_4, 0.2), (x_5, 0.2)\}$$

After this the following result is obtained.

Table 2.2 Weighted Median

x_i	w_i	**S**
x_1	0.3	0.3
x_2	0.2	0.5
x_3	0.1	**0.6**
x_4	0.2	0.8
x_5	0.2	1.0

From the above table, it can be observed that the argument x_3 has the cumulative sum of the weights 0.6 and the arguments preceding and succeeding x_3 i.e. x_2 & x_4 are equal to and greater than 0.5 respectively. Thus according to the definition of weighted median, x_3 is the weighted median.

2.4.3.6 Symmetric Sum

Symmetric sum is a continuous self-dual aggregation operator *S*. The Self-duality is defined by:

$$S(x_1, x_2, x_3, \ldots, x_n) = 1 - S(1-x_1, 1-x_2, \ldots, 1-x_n) \qquad (2.20)$$

This operator was studied in detail by Silvert (1979). In particular it was shown that the symmetric sum of two arguments could be written under the form:

$$S(x, y) = G(x, y) / [G(x, y) + G(1-x, 1-y)] \qquad (2.21)$$

where *G* is a continuous, increasing, positive function satisfying *G*(0,0)=0. It is to notice there is not a unique function *G* characterizing each symmetric sum. It can be remarked that symmetric sums are in general not symmetric or commutative. A good example of symmetric sums is the weighted mean.

2.4.3.7 T-norm and T-conorm Operators

The concept of a triangular norm was introduced in order to generalize the triangular inequality of a metric (Menger 1942). The triangular norm was introduced as a dual operator viz. t-norm and t-conorm (Schweizer & Sklar 1960; Schweizer & Sklar 1983). Both of these operations can also be used as a generalisation of the Boolean logic connectives to multi-valued logic. The t-norms generalise the conjunctive 'AND' operator and the t-conorms generalise the disjunctive 'OR' operator. This situation allows them to be used to define the intersection and union operations in fuzzy logic. Some authors noted this possibility (Dubois & Prade 1981; Klir & Folger 1988) and appreciated the possibilities of this generalization (Alsina et al. 1983). The properties of these operators were investigated (Bonissone 1985). This investigation has the aim of using them in the development of intelligent systems. T-norm and t-conorms has been well-studied and very good overviews and classifications of these operators can be found in the literature.

Definitions:

t-norm: A t-norm is a function $T : [0,1]x[0,1] \rightarrow [0,1]$, having the following properties

$T(x,y) = T(y,x)$ (2.22) Commutativity

$T(x,y) \geq T(u,v)$, if $x \geq u$ and $y \geq v$ (2.23) Monotonicity (increasing)

$T(x, T(y,z)) = T(T(x,y),z)$ (2.24) Associativity

$T(x,1) = x$ (2.25) One as a neutral element

A well-known property of t-norms or conjunctive operator is:

$T(x, y) \leq \min(x, y)$ (Detyniecki 2000).

Proof: Using the monotonicity (2.23) and axiom (2.25),

$T(x, y)$ $T(x, 1) = x$ and using the commutativity, $T(x, y)=T(1,y) = y$.

So, $T(x, y)=\min(x, y)$.

This result implies that in MCDM the use of an "AND" ing allows for no compensation for one bad satisfaction. This implies that "All" the criteria must be satisfied.

t-conorm: Formally, a t-conorm is a function $S: [0,1] \times [0,1] \rightarrow [0,1]$, having the following properties:

$S(x, y) = S(y, x)$ (2.26) Commutativity

$S(x, y) \geq S(u, v)$, if $x \geq u$ and $y \geq v$ (2.27) Monotonicity (increasing)

$S(x, S(y, z)) = S(S(x, y), z)$ (2.28) Associativity

$S(x, 0) = x$ (2.29)

Zero as a neutral element

A well-known property of t-conorms or disjunctive operator is:

$S(x, y) \geq \mathrm{Max}(x, y)$ (Detyniecki 2000).

It is actually a consequence of axioms (2.26), (2.27), (2.29).

The above property implies that in MCDM the use of a pure "OR"ing requires at least one good satisfaction of the criteria. Thus it is clear that the t-norm and t-conorm operators satisfy the two extreme situations of aggregation problems *viz*. ALL of the criteria and AT LEAST ONE of the criteria respectively.

2.5 Additive Measures

In the context of MCDM, additive measures are those classes of operators where the overall result of the aggregation is the summation of individual worth of criteria. For example, consider a set C of criteria $\{c_1, c_2, \ldots\ldots\ldots.c_n\}$ with each having a degree of importance or weight attached to them as $\{w_1, w_2, \ldots\ldots\ldots w_n\}$ respectively. In case of additive measures, the combined importance of criteria is the algebraic sum of their individual criteria weight. This is the main property of additive measures. There are various additive measures found in the literature (Dubois & Prade 1985; Marichal 1999). Out of these, Ordered Weighted Average (OWA) operator suggested by Yager (Yager 1988) is investigated in detail. The OWA operator is explained next.

2.5.1 OWA Operator

Yager (1988) has introduced one special type of operator called the Ordered Weighted Average operator. According to Yager, the OWA are compensatory (between min & max). This property makes the investigation of OWA interesting as it may lead us towards a suitable solution and thus help us in achieving the objective mentioned in the beginning. OWA has the potential to model the sophisticated ways in which human beings process information. The OWA operator is defined as follows-

A Mapping f from $I^n \rightarrow I$ (where $I = [0,1]$) (2.30)

is called as an OWA measure of dimension n if associated with *f*, is a weighting vector W,

$$W = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}$$

Such that

$w_i \in (0,1)$

$\sum w_i = 1$

and where

$f(a_1,a_2,a_3,\ldots\ldots,a_n)= w_1b_1+w_2b_2+\ldots\ldots.w_nb_n$

Where b_i is the i^{th} *largest* element in the collection of criteria values $a_1, a_2, a_3,\ldots.., a_n$

2.5.2 OWA as Linguistic Quantifier

OWA operators are monotonic, symmetric, idempotent and compensatory (Marichal & Roubens 2000). Being a compensatory measure, the results given by an OWA fall between min and max. While dealing with the quantitative setting, the criteria satisfactions can be determined through the memberships functions. The criteria satisfactions are expressed numerically. For a pure quantitative setting this seems to be effective. But in the natural language of human reasoning, the quantification is always done in linguistic terms like “most of”, “a few”, “medium”, “more”, “excellent”, “good”, “average” etc. For example, in a car selection problem, criteria evaluations can be expressed as one of the linguistic values in the following set-

{good, very good, not good}

Thus the word “good” here stands for linguistic variable. The linguistic approach considers the variables, which participates the problem assessed by means of linguistic term instead of numerical values (Zadeh 1975). This approach is appropriate for many problems, since it allows a representation of the expert’s information in a more direct and adequate form where they are unable to express in numerical terms.

A linguistic variable Q differs from a numerical variable in that its values are not numbers; they are words or sentences in a natural or artificial language (Zadeh 1975). Since words, in general, are less precise than numbers, the concept of a linguistic variable serves the purpose of providing a means of approximated characterisation of phenomena which are too complex or too ill-defined to be amenable to conventional qualitative terms (Hohle 1978).

2.5.3 OWA as Quantifier-guided Aggregation:

Based upon the above concept of linguistic quantifiers proposed by Zadeh (1975), OWA has important applications in the area of quantifier-guided aggregation. Consider a set of criteria as $[c_1, c_2, c_3, \ldots\ldots, c_n]$ and let x be an object such that for any criterion c_i, $c_i(x) \in [0, 1]$ indicates the degree to which this criterion is satisfied by x. The degree to which x satisfies "all the criteria" is given by the following (Zadeh 1978).

$$A(x) = \min\{c_1(x), \ldots\ldots, c_n(x)\} \tag{2.31}$$

In this case x is required to satisfy all the criteria i.e $c_1, c_2, c_3, \ldots\ldots, c_n$. This is the boundary condition of aggregation. The degree to which x satisfies "at least one of the criteria" is given by $D(x) = \max\{c_1(x), \ldots\ldots\ldots\ldots, c_n(x)\}$ (2.32)

In this case x is required to satisfy c_1 or c_2 or……or c_n. This is the second boundary condition of the OWA aggregation. But in many applications rather than desiring that a solution that satisfies one of these extreme situations, "all" or "at least one", it may be required that x satisfies most or at least half of the criteria. Drawing upon Zadeh's concept of linguistic quantifiers these kinds of quantifier-guided aggregations can be accomplished.

2.5.4 OWA Calculator

Various properties of OWA operators were discussed earlier. Implementation of OWA algorithm will be presented here. A simple application is developed using Java. This application is termed as “OWA Calculator” and can be a part of the decision engine. The OWA calculator consists of reading data from CSV (Comma Separated Value) list file, processing it to calculate the OWA for different alternatives undertaken with scores for various attributes. The output of the application is the ranking of alternatives according to their OWA values. The OWA values are written in an output file. From this ranking one can determine the most preferred alternative. A car selection problem is considered. There are various attributes for each car. The following relevant attributes are selected for the car selection problem.

- Price
- Speed
- Mileage
- Comfort
- Safety

The car is to be chosen from the following alternatives.

- Mitsubishi
- Audi
- Mercedes
- Toyota

Each attribute is scored against the given alternatives. Thus a 4 x 5 decision matrix is formed. The decision matrix in the given situation can be represented as

Table 2.3 Car Decision Matrix for OWA calculator

	Price	Speed	Mileage	Comfort	Safety
Weights	W_1	W_2	W_3	W_4	W_5
Mitsubishi	Score $_{1,1}$	Score $_{1,2}$	Score $_{1,3}$	Score $_{1,4}$	Score $_{1,5}$
Audi	Score $_{2,1}$	Score $_{2,2}$	Score $_{2,3}$	Score $_{2,4}$	Score $_{2,5}$
Mercedes	Score $_{3,1}$	Score $_{3,2}$	Score $_{3,3}$	Score $_{3,4}$	Score $_{3,5}$
Toyota	Score $_{4,1}$	Score $_{4,2}$	Score $_{4,3}$	Score $_{4,4}$	Score $_{4,5}$

The weights and scores of the attributes for the cars are the numeric information stored in the CSV list file. Figure 2.3 shows the screen shot of the file with the numeric values used in the application. Note that the first row indicates the weights. Please note that the weights are chosen such that their algebraic sum equals to 1. The remaining four rows indicate the scores for the four cars respectively.

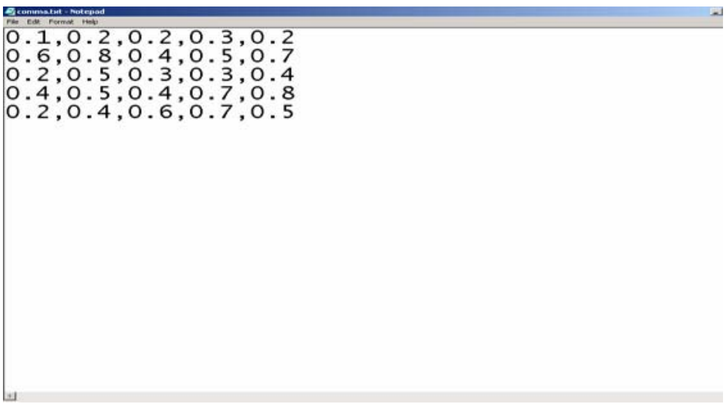

0.1,0.2,0.2,0.3,0.2
0.6,0.8,0.4,0.5,0.7
0.2,0.5,0.3,0.3,0.4
0.4,0.5,0.4,0.7,0.8
0.2,0.4,0.6,0.7,0.5

Figure 2.3 Input data for OWA calculation as a CSV List File

2.5.5 OWA Calculator Output

After running the application, the output on the screen is as below:

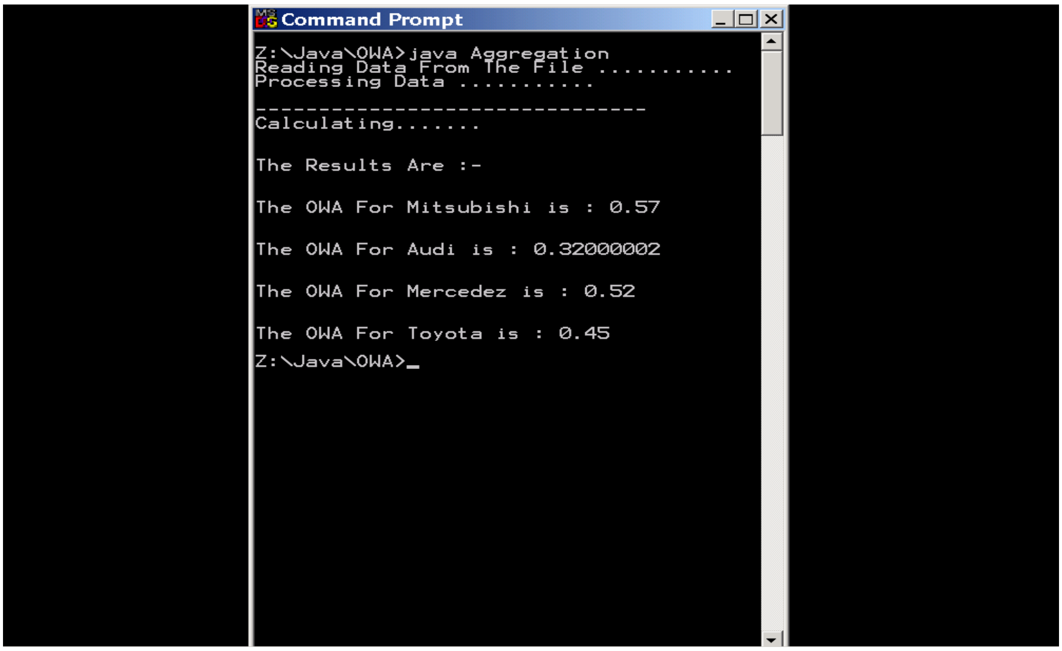


Figure 2.4 Output of OWA application

The calculated OWA are then sorted in descending order and are written to an output file. The output ranks the alternatives according to their OWA scores. From the output, it can be noted that Mitsubishi is the preferred car over the other alternatives.

2.6 OWA Analysis

- The above original implementation is done as a part of the undertaken research. From this implementation it can be observed that the ordering of argument values is an important step in finding the global worth of an alternative. The ordering step is perfectly valid when the argument values are distinct for each alternative. If the argument values are same, then the OWA proposes equal ranking of the alternatives. This situation may be a special situation where the decision maker happens to evaluate the criteria in an interlaced manner.
- Thus mathematically, if for a pair of alternatives the criteria scores are same but interlaced numeric values, then according to this OWA the alternatives will have the equal preference. However, such result is counter-intuitive.
- OWA does not consider the weights of sub-sets of criteria. In addition, there is no evident mechanism present to handle the issue of dependence of criteria.

Thus after analysing the above points and results from the previous work (Sharma 2001; Wagholikar 2002), there is a need for investigating other types of aggregations approaches for the undertaken research. In conclusion, the study of additive measures and traditional MCDM approaches shows that these tools are inadequate to model the interaction between the criteria.

2.7 Non-additive Measures

As discussed earlier the nature of interaction between the criteria i.e. positive or negative interaction, determines the positive or negative relationship between them. Hence the combined importance of criteria may not be equal to the summation of individual importance i.e. $w_{AB} \neq w_A+w_B$. This is the important property of the non-additive measures, which primarily distinguishes itself from the additive measures, which has the property of $w_{AB} = w_A + w_B$. Non-additive measures are termed as fuzzy measures and fuzzy integrals (Grabisch 1995a). The term 'non-additive' applies for any possible arithmetic operations between the individual criteria weights excluding the summation. The central idea is that the combined importance of criteria may not be equal to the summation of individual weights but it may be equal to other arithmetic operations. It is difficult or rather impossible to define an exact formulation for w_{AB} involving individual weights i.e. w_A and w_B.

2.7.1 Fuzzy Measure

Let $C = \{c_1, c_2, \ldots\ldots, c_n\}$ be the set of criteria and $\mathcal{P}(C)$ as the power set of C, i.e. the set of all subsets of C. A fuzzy measure μ is a set function on the set of C criteria such that

$\mu: \mathcal{P}(C) \rightarrow [0,1]$, which satisfies the following properties.

i. $\mu(\phi) = 0$ and $\mu(C) = 1$. (2.33)

This means that the combined importance of the set C of criteria is maximum when all the criteria are considered and its minimum when the number of criteria is zero.

ii. If $A \subset B \subset C$ then $\mu(A) \leq \mu(B)$, where $\mu(A)$ & $\mu(B)$ represent the importance of the set of criteria A & B respectively. This property implies that the importance of a subset of criteria is always less than the importance of criteria of its super-set. This is the property of monotonicity. It can be noted here that the additivity axiom for the additive measures $\mu (A \cup B) = \mu(A) + \mu(B)$, $A \cap B = \phi$ is been replaced by a weaker one i.e. monotonicity.

iii. Fuzzy measure is said to be super-additive if it satisfies the condition viz)

$$\mu (A \cup B) > \mu(A) + \mu(B) . \quad (2.34)$$

iv. Fuzzy measure is said to be sub-additive if it satisfies the condition viz.

$$\mu (A \cup B) < \mu(A)+ \mu(B) \quad (2.35)$$

In general, terms, fuzzy measure models the importance of sub-set of criteria by considering dependence of criteria. Due to this reason, fuzzy measure becomes a good candidate for solving the research question.

Please note that if a fuzzy measure is additive, then it suffices to define the n coefficients to the 2^n subsets of C. Otherwise, for a set of n criteria, there can be 2^n sub-sets of criteria. Grabisch (1996) and Grabisch & Modave (1997) showed that most of the real applications of fuzzy measure deal with multicriteria decision-making problems. They are able to model the relative importance of criteria as well as their interaction or dependence of criteria. Modelling of dependence among criteria is precisely the main use of fuzzy measures for MCDM problems (Grabisch 1996).

2.7.2 Fuzzy Integral

It is discussed earlier that fuzzy measure has the ability to model the importance of sub-set of criteria. Fuzzy integral is a tool to aggregate the fuzzy measure. There are two types of commonly known fuzzy integrals viz Choquet integral (Choquet 1953) and Sugeno integral (Sugeno 1974). Choquet integral is used for a quantitative setting whereas Sugeno integral is used for qualitative setting. Numerical setting is assumed for the undertaken research. Hence Choquet integral is investigated as the aggregation function.

2.7.2.1 Choquet Integral Definition

Let C = $\{c_1, c_2, \ldots, c_n\}$ is the set of criteria and μ be the fuzzy measure on X. The Choquet integral of the function $f : C \rightarrow [0,1]$ with respect to μ is given by the following expression.

$$C_\mu = \sum_{i=1}^{n} \{f(c_{(i)}) - f(c_{(i-1)})\}\ \mu(A(i)) \qquad (2.36)$$

Where $._{(i)}$ indicates that the indices have been permuted such that

$0 \leq f(c_{(1)}) \leq f(c_{(2)}) \leq \ldots\ldots. \leq f(c_{(n)}) \leq 1$ and $A(i) = \{ c_{(i)}, \ldots\ldots\ldots\ldots, c_{(n)} \}$

It can be observed that the fuzzy integral is the incremental summation of the product of the criteria evaluation and its fuzzy measure. Similar to the fuzzy measure, Choquet integral has the property of montonicity (Grabisch 1995a).

Non-additive measures can model the issue of dependence of criteria through fuzzy measure and fuzzy integrals. Though fuzzy measure can model the importance of sub-sets of criteria, this advantage comes with the computational complexity of determining $2^n - 1$ fuzzy measure coefficients. The recent research shows that one can use these tools for a same decision to be made over and over with the help of historical data and decision maker's preference (Grabisch 1995; Marichal & Roubens 2000). It is proposed to extend this notion. Development in this direction will be a contribution to the field of MCDM.

2.7.2.2 Set Relations between Aggregation Operators

The Choquet integral can range freely from the most tolerant behaviour (max) to the most intolerant (min). The next figure 2.5 gives the summary of all set relations between various aggregation operators and Choquet integral. From this figure, it can be observed that Choquet integral is the super set of other operators such as weighted sum, mean and OWA. These operators are considered as the special case of Choquet integral (Marichal & Roubens 2000). Choquet integral is more suitable for aggregation in quantitative setting. The undertaken research focuses on aggregation issues in a quantitative setting. Hence Choquet integral is a justifying choice for this research.

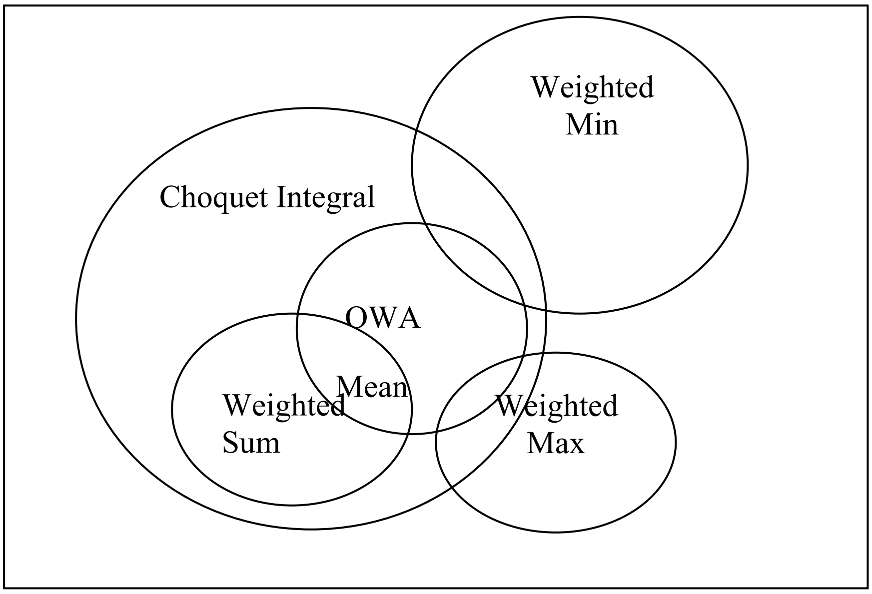


Figure 2.5 Set relations between aggregation operators

2.7.2.3 Sugeno Integral

As stated earlier, this is another type of fuzzy integral used for qualitative data set (Sugeno 1974). This is mainly used in the situation where the available information for the decision-making is of qualitative nature. Sugeno integral is defined as -

$$S_\mu (x) = V [x(i) \Lambda \mu (\{(i),........,(n)\})] \qquad (2.37)$$

Where (.) indicates a permutation on N such that $x_{(1)} \leq .... \leq x_{(n)}$, V stands for max and Λ stands for min. This integral can normally be applied for a qualitative setting especially using an ordinal scale (Marichal 1999). However, the undertaken research considers the quantitative setting for which Choquet integral is considered more suitable.

From the above discussion of non-additive measures, it is observed that fuzzy measures are flexible tools for handling the issue of dependence of criteria. However, they are not easy to apply for any practical problems. This difficulty is mainly due to the requirement of defining 2^n-1 positive real coefficients and difficulty in interpretation of fuzzy measures. These two problems are well described in the literature (Grabisch 1995a). It is agreed that due to these issues fuzzy measure remain difficult to interpret and understand. The application of fuzzy measure is easier if the practitioner chooses additive fuzzy measures with “n” real coefficients. However, in that case the issue of dependence between the criteria will not be handled due to limitations of additive measures described earlier.

2.8 Analysis Tools for Interaction between Criteria

Interpretation of fuzzy measures and its analysis is important. The fuzzy measures can be understood if the nature of interaction between the criteria is known. The tools that can assist in determining the nature of interaction are useful in the interpretation of fuzzy measures. The following concepts are useful for this purpose. These are the only concepts that can help in determining the nature of interaction.

2.8.1 Shapley Importance Index

Shapley importance index and interaction index are used for analysing the fuzzy measures (Shapley 1953). The Shapley importance index for every $i \in C$ i.e. for every criteria in a set of criteria, is defined by:

$$v_i = \sum_{k=0}^{n-1} \gamma_k \sum_{k \subset x \setminus i, |k|=k} (\mu_{ik} - \mu_k) \tag{2.38}$$

with $\gamma_k = \dfrac{(n-k-1)!k!}{n!}$

Where, n = Total number of criteria

k = Number of elements in a sub-set

The Shapley value of μ is the vector $v(\mu) = [v_1, ..., v_n]$.

The above Shapley index can be interpreted as a kind of average value of the contribution of element i, individual criteria, alone in all coalitions. Hence the summation of these Shapley values for a given set of elements would represent the importance of the complete set. The importance index can be used to understand the importance of an element in a coalition.

2.8.2 Interaction Index

This concept deals with the interaction between the elements (Grabisch 1995). If there are two elements i and j, their combined importance μ_{ij} would be equal to the summation of their individual importance provided there is no interaction between them. However, if there exists an interaction between the elements, their combined importance could be higher or lower than their summation depending upon positive or negative interaction between the elements. The interaction index between the sub-sets of elements is defined as below.

Let μ be a fuzzy measure on X . The interaction index I_{ij} is defined by

$$I_{ij} = \sum_{k=0}^{n-2} \xi_k \sum_{K \subset X \setminus \{i,j\}, |k|=k} (\mu_{ijk} - \mu_{ik} - \mu_{jk} + \mu_k) \tag{2.39}$$

With,

$$\xi_k = \frac{(n-k-2)!\,k!}{(n-1)!} = \frac{1}{\frac{(n-2)}{k}(n-1)}$$

where n = Total number of criteria

k = Number of elements in a sub-set

This interaction index can be interpreted as an average value of the added value given by putting i and j together. When the I_{ij} is positive (resp. negative), then elements are related in a positive (resp. negative) manner.

This index shows us the nature of interaction between the criteria. The index is positive if the criteria support each other and the index is negative if the criteria contradict each other. The above notion of interaction index can be extended to the sub-set of elements as follows:

Let μ be a fuzzy measure on X . The interaction index I_T for any sub set $T \subset X$ is defined by –

$$I_T = \sum_{k=0}^{n-|T|} \xi_k^{|T|} \sum_{K \subset X \setminus T, |K|=k} \sum_{L \subset T} (-1)^{|T|-|L|} \mu_{LK} \tag{2.40}$$

with $\xi_k^{\,p} = \dfrac{(n-k-p)!\,k!}{(n-p+1)!}$

Where T = Sub-set of criteria

n = Total number of criteria

2.8.3 Analysis of Importance and Interaction index

The above concepts of importance index and interaction index are useful in analysing the fuzzy measures. In a practical problem, these concepts can be used for understanding the behaviour of the fuzzy measures and to better understand the preference structure. However, these concepts do not address the issue for determining the fuzzy measures in a practical situation. They cannot be referred as tools for determination of actual fuzzy measures for the given decision problem. There are other measures suggested to simplify the issue of fuzzy measure determination.

2.8.4 K-Order Fuzzy Measures

The complexity of fuzzy measures puts limitations on its application. In order to simplify the complexity, Grabisch (1995a) has suggested K-order additive fuzzy measures. For defining k-order additive fuzzy measures, a pseudo-Boolean function must be defined. A pseudo-Boolean function is a real-valued function $f:\{0,1\}^n \rightarrow \Re$. In case of additive measures this pseudo-Boolean function would be a linear function i.e. $f(x) = \sum_{i=1}^{n} a_i x_i$.

The definition of k-order additive fuzzy measures depends on the nature of this pseuo-Boolean function. Let T be the set of elements in a given sub-set of criteria.

a_T be the set of fuzzy measure coefficients.

The k-additive fuzzy measure can be defined as the fuzzy measure whose corresponding pseudo-Boolean function is a multilinear polynomial of degree k such that the set of coefficients $a_T = 0$ for all T such that |T|> k and there exists at least one sub set T of k elements such that $a_T \neq 0$.For example 2-additive fuzzy measures are represented as

$$\mu(k) = \sum_{i=1}^{n} a_i x_i + \sum_{\{i,j\} \subset X} a_{ij} x_i x_j \text{, for any } K \subset X \text{ such that } |k| \geq 2 \qquad (2.41)$$

From the above formulation, a 2-order additive fuzzy measure would be determined by only 2 coefficients of the fuzzy measure.

The k-order additive measure certainly simplifies the representation of 2-order fuzzy measures. However, this representation of the fuzzy measures would not directly automate the decision making process as importance a_{ij} of the sub-set must be acquired.

This mechanism certainly simplifies representation of fuzzy measures especially in case of higher values of n. However, this approach does not suggest any mechanism for acquiring the importance of sub-set of criteria. Though k-additive fuzzy measures do not explicitly explain the acquisition of fuzzy measures, they are important tools for behavioural analysis of fuzzy measures and hence for understanding the nature of interaction between the criteria.

2.9 Fuzzy Measure Determination Methods

As mentioned earlier, fuzzy measure acquisition process can become very complex for large number of criteria 'n'. Also determining the fuzzy measure in a practically feasible manner is an important aspect. Researchers have suggested some methods. These methods are discussed in the following sections.

2.9.1 λ Measure

Sugeno (1974) has suggested a λ-Measure. This measure can be used as fuzzy measure in the fuzzy integral computation. λ is known as non-additivity constant in the range [-1, +∞]. Using this measure the Choquet integral can be computed through the following steps-

1. For a given set of weights (w_i), determine λ: [-1, +∞] measure from the equation,

$$1 + \lambda = \prod_{i=1}^{n} (1+ \lambda w_i) \tag{2.42}$$

   This formulation assumes that the n factors define the aggregate variable. Sugeno has shown that a unique root always exists which is non-zero and ≥-1. The λ value so determined is constant for each group of factors and is dependent on weights w_i.

2. Order the criteria evaluation values or scores x_{ij} in descending order.
3. For the ordered factors calculate the subset weight $\mu(A_i)$ where $A_i \subset n$, for increasing subsets of factors starting with the highest value using the following equation,

   $\mu(A_i) = \mu(A_i - 1) + w_i + \lambda\, w_i\, \mu(A_i - 1)$

   Where w_i = weight, and $\mu(A_0) = 0$
4. Use equation (1) above to form a global Choquet value for the group of values.

Several computations were conducted for determining λ measure for different set of values (Wagholikar & Deer 2004).

Based on the result the following points are observed.

a. The λ value is coming out to be invalid i.e. $\lambda < -1$, when the weights are normalized.
b. The λ gives all positive or all negative values. This indicates that there could be only positive or negative interaction among the criteria. It does not cater the case where there could be both positive as well as negative interactions among the criteria. This restricts the application of this measure for modelling only one type of interaction between the criteria.

Based on the above points, it can be concluded that λ measure cannot be used in fuzzy integral computation in the presence of positive as well as negative interaction between the criteria. Hence there is need to investigate some more methods of determining the combined importance of criteria.

2.9.2 Interactive Optimisation

Grabisch (1995a) has proposed an interesting approach for fuzzy measure identification through interactive optimisation. This approach involves use of historical data (i.e. data about the past decisions) and an interactive dialogue between the decision-maker and the fuzzy measure determination system. This approach involves active participation of the decision maker in the fuzzy measure determination process. This approach is based on the assumption of equal importance of criteria when the decision maker is unable to provide the preferences over the criteria. This implies that there is equal contribution of each criterion in the global evaluation of an alternative. Grabisch (1995a) describes this as an equilibrium situation as the weights on the criteria *"constitute an additive uniformly distributed measure"*.

This can be considered as an ideal situation as the criteria may not contribute equally in the evaluation process. Hence a solution, which is the proximity of the ideal solution, can be found. This process eventually focuses on minimizing the difference between the ideal measure and the learnt measure from the historical data.

Mathematically the additive uniformly distributed measure is represented as

$\mu = i / m$ where $i = 1,2,3...m$

By various values of "i" we get the respective i^{th} order measure. For e.g. a set of criteria where m=3, the first order measure $\mu=1/3 = 0.33$ which implies that the ideal contribution of criteria is 0.33.

The algorithm is as follows.

1. Based on the historical data about the same decisions made in the past, the constraints for the global evaluation for the current situation is expressed as follows –

$$Z_j - \delta_i \leq C_j^t X_j \leq Z_j + \delta_i \quad (2.43)$$

   Z_i = global evaluation in the past decisions

   δ_i = Acceptable allowance about global evaluation value in the learning data which is the data about the same decision(s) made in the past.

   $C_j^t X_j$ = Choquet evaluation

2. Determine the decision maker's preferences explicitly through the importance of degree of each criterion and the interaction between them. For this the preference table suggested for Analytical Hierarchy Process (AHP) (Satty 1980) is used.
3. The decision maker's preference between a pair of criteria i and j is expressed as follows.

   $x_i \leq \eta x_j$, where η determines degree of preference of criteria i over criteria j.

The type of relationship between the criteria is modelled by the following constraints.

$\mu(A \cup B) = \mu(A) + \lambda\, \mu(B)$ $0 \leq \lambda \leq 1$ (for positive interaction between the criteria)

$\mu(A \cup B) = \mu(A) + \mu(B) + \gamma(1 - \mu(A) - \mu(B))$ $0 \leq \gamma \leq 1$ & $\mu(A) + \mu(B) < 1$ (for negative interaction between the criteria).

4. After this input data set, it is shown that quadratic programming can solve the problem of minimizing the difference between the global evaluation and the Choquet evaluation of the past data (Grabisch 1995a).
5. The above process i.e. steps 2 and 3 are repeated until the decision maker's preferences are satisfied.

This method is explained in detail in Grabisch's work (Grabisch 1995a). In this approach the historical data is in the form of global evaluation Z_i and criteria evaluations or scores x_{ij} in the past decisions. However, it can be argued that the global evaluation may not be available from the decision problems. Therefore another approach is investigated, which does not require historical data unlike Grabisch's approach described above. It is mainly based on the decision maker's preference over the set of criteria.

2.9.3 Semantic Consideration about Criteria

This approach involves direct involvement of the decision maker and modelling his/her preferences in a linear constraint satisfaction problem (Marichal 1999). In this method, the input dataset consists of the following: -

-The set A of the alternatives and the set C of criteria

-A table of individual scores (utilities) x_{ij} given on a same scale

Thus, a general MCDM matrix is formed as shown in MCDM process in chapter 1. Then based on the decision maker's preference over the set of criteria a partial pre-order on the alternatives is formed. Along with this dataset, information such as signs of interaction between some pairs of criteria, degree of orness, veto and favour degrees etc. can be integrated, if it can be given by a linear expression. This data set forms a linear constraint satisfaction problem whose solution gives a 2-order fuzzy measure. It is assumed that the decision maker is able to provide information about the preferences between the two criteria. The formulation for this problem is as follows:

Maximize z = ε

Subject to

$$\sum_{j=1}^{m} a_{ij} x_j \leq b_j \quad \text{where } i = 1,......p \tag{2.44}$$

$$\sum_{j=1}^{m} c_{ij} x_j \leq d_i - \varepsilon \quad \text{where } i = 1,......,q \tag{2.45}$$

In the above formulation, the objective function is to maximize the inequality in the constraints expressed by the summation of importance of pair of criteria i and j and the criteria score. The coefficients b, d, p and q are constants. The decision maker's preferences are expressed based on criteria evaluation, which in turn form the linear constraints in the above formulation.

In this method, these constraints are shown to be bounded by fixed preference threshold δ. By solving the above linear program the 2-order fuzzy measure can be determined. The global evaluations of alternatives can be determined using Choquet integral. This method has been illustrated in detail by Marichal (1999). This method can determine only 2-order fuzzy measure. It can determine the importance of sub-sets involving two criteria. However, the applicability of this method for large number of criteria can be argued. This is because as the number of decision criteria increase, then linear programming may become complex. The value of threshold δ has a subjectivity associated with it. One has to try to compare solutions with different values of δ. The above method has been applied using the following dataset.

Table 2.4 Car Selection Decision Matrix for Marichal's approach

	Comfort	Fuel Economy	Power
Car A	0.9	0.75	0.95
Car B	0.75	0.9	0.95
Car C	0.75	0.9	0.55
Car D	0.9	0.75	0.55

The following results were obtained.

Table 2.5 Results using Marichal's approach

Car	Global Evaluation
Car A	0.925
Car B	0.85
Car C	0.725
Car D	0.65

Unlike previous method of interactive optimisation, this method does not require any historical data as the linear program is mainly formulated based on decision maker's preference over the set of criteria.

2.9.4 Determination of Fuzzy Measures by Transformations

The problem of determining the fuzzy measures can be considered as a problem of knowledge acquisition (Klir et al. 1996b). This difficulty arises mainly due to the definition of fuzzy measures over the sub-set of elements in a universal set. This implies an exponential increase in the computational complexity. In addition, fuzzy measures must satisfy the monotonicity requirement and when dealing with infinite sets the continuity (or semi-continuity) requirement. These requirements further increase the complexity in the process. This complexity can be simplified using suitable transformations (Dubois et al. 1993; Wang et al. 1999). Fuzzy measures can be constructed using appropriate transformations. A regular transformation is a continuous, strictly monotone function θ on X with $\theta(0) = 0$ and $\theta(1) = 1$.

Using these transformations, a given Sugeno measure can be transferred to another Sugeno measure with an assigned parameter. If the parameter is λ_0 and the assigned parameter is λ, then the transformation has the form

$$\theta\ (x) = \frac{(1+x)^{\log 1+(1+\lambda_0)} - 1}{\lambda} \qquad (2.46)$$

This approach requires determination of Sugeno measure (Leung et al. 2000). It requires a set of pre-defined Sugeno measures. As it has shown earlier, Sugeno measures can model only one type of interaction. Hence an approach based on the Sugeno measures is not suitable for decision problems where the criteria may be both positively as well as negatively related. These transformations can be applied when there exists only one type of interaction.

2.9.5 Determination based on Pre-defined Values of Choquet Integral

This approach suggests determination of fuzzy measures based on the values of pre-defined Choquet integral (Klir et al. 1996a). This approach assumes an existing monotone function μ. It defines the properties of the new monotone set function ν. Based on the existing set function and the required properties of the new function, new fuzzy measures can be determined using a Choquet integral. The approach is as follows. Let $\mu : \rho \to [0, \infty]$ be a given monotone set function with $\mu(\phi) = 0$. Choosing a non-negative measurable function f on X, a new set function ν can be obtained for all $A \in f$ via a Choquet integral. This can be expressed as below.

$$\nu(A) = c(A)\int_A f d\mu . \qquad (2.47)$$

The above expression shows that if a fuzzy measure is known and properties for the new fuzzy measure are defined, then new fuzzy measures can be determined. This approach is applicable if an acceptable fuzzy measure ν is constructed by determining an appropriate non-negative measurable function f. It is suggested that this approach is much simpler than the transformation methods discussed earlier as it requires determination of few real coefficients when $|X| < |\rho| - 2$. However, this approach can be complicated due the to problem of the construction of function f.

2.9.6 Data-driven Methods

This is a special class of methods where fuzzy measures can be determined using a data set about evaluations of the objects (Wang & Klir 1992; Yuan & Klir 1996). Consider a set of attributes X (criteria in a given decision problem), a function f on X. This function is the function assigning a value to each criteria. A fuzzy measure μ on the power set of x is also defined. This fuzzy measure specifies the importance of each sub-set of X. For this situation, the Choquet Integral $\int f d\mu$ is the tool summarising the information expressed by f according to the importance measure or the fuzzy measure μ. The Choquet Integral is the tool for obtaining the overall evaluation of the object against a set of criteria. Once the set of alternatives $X = \{x_k \mid k \in N_n\}$ and the fuzzy measure μ are given, and then either the Choquet or Sugeno integral can be used to determine the overall evaluation for any relevant object (Yuan & Klir 1995). When m objects are considered, m x n attribute valuations are involved, together with m overall summaries of the Choquet integral. This can be illustrated in the following table.

Table 2.6 Input data for Data driven methods

Object	Criteria scores	Choquet Evaluation
i=1	$f_1(x_1), f_1(x_2), \ldots\ldots\ldots, f_1(x_n)$	C_1
2	$f_2(x_1), f_2(x_2), \ldots\ldots\ldots, f_2(x_n)$	C_2
..		
..	$f_m(x_1), f_m(x_2), \ldots\ldots\ldots, f_m(x_n)$	
M		C_m

Where $f_i(x_k)$ denotes the valuation of attribute $x_k\,(k \in N_n)$ for object i and the Choquet Integral is given by

$$C_i = (c)\int f_i d\mu \qquad (2.48)$$

In the undertaken problem, the fuzzy measures μ are not known. Only the individual criteria scores and the overall assessment of the criteria are known. Then the fuzzy measure determination problem reduces to the inverse problem of determining the fuzzy measure μ from the linear system of m equations. The data of the criteria scores and the global evaluation can be used to solve the m equations of the following form.

$(c)\int f_i d\mu = a_i$, Where a_i is the overall assessment of an alternative. The above inverse problem is a complex problem especially for higher values of m.

This approach requires the knowledge of overall assessment of alternatives in a new decision problem. However, it can be argued that it is not possible to acquire this knowledge in a decision-making problem. Since this is a decision making process, the overall assessment of alternatives in a new decision problem would be absent. This approach however can be applicable in a situation where fuzzy measures obtained from the past data are applicable for the new decision problem.

2.9.7 Principal Components Analysis

From the studies done so far, it can be concluded that non-additive measures are important tools to solve the research question. However, the issue of fuzzy measure determination is complex especially for large values of n criteria. This issue may be simplified by using a transform so that less number of coefficients can be used for the computation. Principal Component Analysis or Karhunen-Loève transform (Karhunen 1946; Loève 1955) is a transform that is used to reduce the dimensionality of a dataset in which there are a large number of interrelated variables, while retaining as much as possible of the variation present in the data set (Dunteman 1989; O'Connell 1974). It is also known as Hotelling transform (Hotelling 1933). The principle component can be illustrated with the help of the following figure.

Figure 2.6 Principal components analysis

In the above figure the points indicate a two-dimensional data. The straight line is the direction of the first principal component that gives an optimal linear reduction of dimension from 2 to1.

The reduction in data dimensionality is achieved by transforming to a new set of variables, the principal components, which are uncorrelated, and which are ordered so that the first few retain most of the variations present in all of the original variables.

Computation of the principal components reduces to the solution of Eigen values. Principal components analysis can be useful in selecting a subset of variables to represent the total set of variables. This discussion does not take into account the use of outside criteria, such as their effectiveness in predicting a particular dependent variable, to select the subset of variables; only the internal structure of the data is considered. The rationale for selecting a subset of variables to represent the variation in a set of variables rather than the principal components themselves is based on two considerations. First, in order to compute the principal component scores, measures of all variables in the variable set are needed since each principal component is a linear combination of all of the variables. Some variables may be too difficult or too expensive to measure, and, therefore, collecting data on them in future studies may be impractical. Second, while the variables themselves are usually readily interpretable, the principal components are sometimes un-interpretable.

2.9.8 Literature Analysis

The above literature review so far explored various theoretical aspects of the research question. The undertaken problem is investigated. It was shown that aggregation and aggregation measures are the important and foundational aspects of this problem. Various aggregation measures were discussed. It was shown that non-additive measures could handle the issue of dependence between the criteria. However, there are issues in computing the non-additive measure coefficients. Especially determination of fuzzy measure coefficients is a major issue. Various fuzzy measure determination approaches were analysed. However, mainly it was discovered that these approaches are based on the learning data of the same decisions made in the past. They also involve considerable involvement of the decision maker. These approaches suggest the need of determining 2^n-1 fuzzy measure coefficients. None of them indicates the need and possibility of reducing the complex data set of 2^n-1 coefficients to some manageable form of data. It is difficult to derive a definite algorithmic solution for this problem. However, it is possible to have a different approach towards solving this fuzzy measures acquisition problem.

2.10 Theoretical Foundations of the proposed approach

After the critical analysis of the state-of-the-art approaches related with the research question, this work proposes an approach, which will address the issues identified in the literature review. The investigation of the issue of dependence of criteria has lead this research to develop an appropriate pragmatic mechanism for fuzzy measure determination such that non-additive measures can be used in the presence of interactive criteria. The literature analysis showed that the previous approaches mainly address the issue of interaction between the criteria using a set of historical data. The historical data can consist of data about the past similar decisions. Such data set can be used for fuzzy measures acquisition. This is the central idea behind the proposed approach.

2.10.1 Reasoning by Analogy

The main premise of this approach is that the use of similar precedents for fuzzy measure determination will resolve some pragmatic issues associated with fuzzy measures. Reasoning by analogy is one of the forms of human reasoning (Minsky 1963). It can be considered as core of human cognition (Gentner et al. 2001; Dubois & Prade 1994). People may choose to reason and solve problems using similar experiences (Lancaster & Kolodner 1987). It is a remarkable ability of people to easily understand new situations by resemblance to old ones, to comprehend metaphors, and to solve the problems based on previously solved, analogous problems (Read & Cesa 1991; Ross 1984). These characteristics of reasoning may be considered as abilities of analogical reasoning (Klein 1987; Vosniadou & Ortony 1989). Analogical reasoning has been considered as one of the key phenomenon of intelligence since ancient times (Helman 1988).

From a psychologist's perspective "With the invention of some simple version of the concept of a problem, it becomes possible to solve a problem by recognizing that it involves objects and relations similar to those involved in previous experiences." (Holyoak & Koh 1987; Faries & Schlossberg 1994). In other words, the capacities for analogy can become mental tools for problem solving. Analogy is a central concept in the theory of analogical reasoning.

2.10.1.1 Analogy

The English word "analogy" originates from the Greek *analogia*, meaning mathematical proportion (Such as 2:4::4:8). Basically, an analogy is a statement of a logical relationship between two similar things that are compared with each other (Vosniadou & Ortony1989). However, 'analogy' has taken on many meanings. The notion of analogy depends on the underlying context and domain.

This work considers analogy, as an inference that if two or more things agree with one another in one or more respects, then they will probably agree in yet other respects. The extension of this principle can be applied in the context of decision-making. Accordingly, if the new decision problem is similar to a past decision problem then the past solution can be applied for solving new problem. There are two kinds of analogies.

a. Metaphorical Analogies – These are the analogies in which the analogically related items are drawn from conceptually different or remote domains. They can also be termed as between-domain analogies (Kokinov 1996).

b. Literal Analogies – These are the analogies in which the items are drawn from the same domain, or at least from conceptually very close domains. They can also be termed as within-domain analogies. They are mainly used to solve the new problems using the solution of the old problems in the same domain (Holyoak & Koh 1987).

It is important not to lose sight of the distinction between the two kinds of analogies, because it is possible that somewhat different processes might be involved in the two cases (Spiro et al. 1989).

Analogical reasoning can be described as a reasoning process utilizing the relational information from a domain that already exists in memory (usually referred to as the source or base domain) to the domain to be explained (referred to as the target domain) (Keane et al. 1994). In this reasoning process, similarity is implicated because a successful, useful analogy depends upon there being some sort of similarity between the source domain and the target domain and because the perception of similarity is likely to play a major role in some of the key processes associated with analogical reasoning (Anderson & Thomson 1989).

Analogical reasoning mainly involves the following sequence of processes –

a. Gaining access to an appropriate analog;
b. Mapping some part of the information associated with that analog onto the target domain;
c. Mapping the source domain solution to the target domain.

2.10.2 Analogical Mapping

Analogical mapping involves setting up correspondences between properties in the two domains and transferring a relational structure that embodies some of the relations between these properties (Holyoak & Thagard 1995;Collins & Burstein 1989). In many instances of analogy comprehension, transfer of the relational structure is not needed because the relational structure already exists in the target domain (Thagard & Holyoak 1989). In the context of the undertaken research, this can be interpreted as the presence of same sub-set of criteria in the source and target domain. Hence the preferences of the sub-set of criteria in the source domain can be used directly for determining the fuzzy measures and hence eventually computing the global evaluation of the alternatives.

The above discussion directs the selection of one type of analogy for this work. As mentioned earlier, this research pursues the development of fuzzy measures acquisition methodology based on the notion of using similar data sets. These data sets represent the new decision problem and a set of the past decision problems. It is assumed that the data sets will belong to the same domain. Therefore literal analogies are applicable for the undertaken work. This work defines literal analogy-based reasoning as similarity-based reasoning. Thus, similarity based reasoning is considered as a sub-set of reasoning by analogy.

The application of similarity-based reasoning would require construction of a data set, retrieval of solution of a similar problem and adaptation of the solution for solving the new problem. Case-based reasoning (CBR) is considered as a good tool that implements the similarity based reasoning and can be useful for human decision-aiding (Pal et al. 2000; Kolodner 1991; Bergmann 2003; Mark & Kolodner 1992).

2.11 Case-based reasoning

There are various models suggested for CBR (Plaza & Aamodt 1994; Kolodner 1992). The CBR cycle mainly involves the following phases (Sun & Finnie 2003).

i. Retrieving similar previously experienced cases (e.g. problem-solution-outcome triples) whose problem is judged to be similar.
ii. Reusing the cases by copying or integrating the solutions from the cases retrieved.
iii. Revising or adapting the solution(s) retrieved in an attempt to solve the new problem.
iv. Retaining the new solution once it has been confirmed or validated.

The above phases are shown in the following CBR cycle proposed by Plaza and Aamodt (1994).

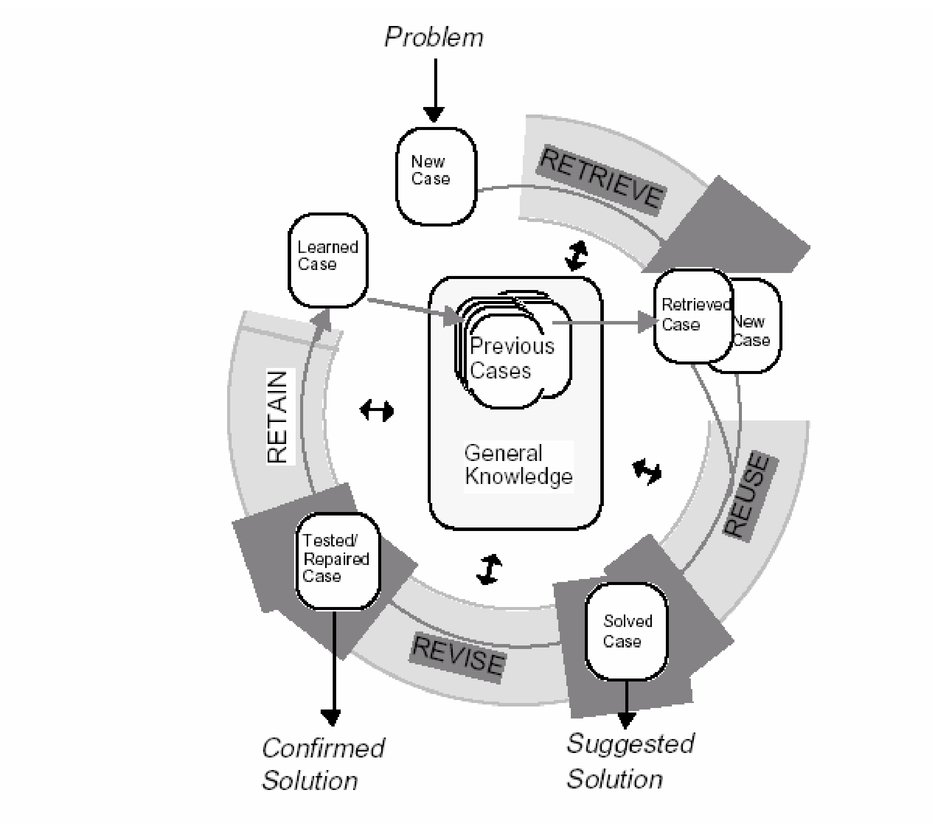


Figure 2.7 CBR Cycle (Adopted from Plaza and Aamodt (1994))

2.11.1 Case Representation

A case can be regarded as a contextualized piece of knowledge (Pal & Shiu 2004). Case representation is one of the problem areas associated with CBR (Yang et al. 2003). The representation problem is about deciding what to store in a case (Khator et al. 1995). In general terms, the case structure consists of a problem part and a solution part. These two parts can be structured in accordance with the real situation in the problem domain. The problem part consists of the desired goal or reasoner's aim, a set of constraints on this goal. The solution part may include the solution itself, the reasoning steps, the justification, the alternatives and the expectations. Different problems will have different case-structures (Koton 1988). Once the case structure is defined as per the underlying context, an appropriate case representation format can be selected. The case format can be regarded as the way in which the case is stored in the system. Case in a case base can be stored in different representational formats (Pal & Shiu 2004; Gebhardt et al. 1997). The following factors must be considered when choosing a representation format for a case:

- The form in which cases are available or obtained.
- Types and structures associated with the content or features that describe a case (Cortés et al. 2002). These types have to be available, or be capable of being created, in the case representation.
- The indexing and search mechanism planned. Cases have to be in a format that the case retrieval mechanism can deal with effectively.

The traditional case representations are relational case representation, Object-oriented case representation and predicate case representation (Pal & Simon 2000; Bergmann 2002). The following figure shows an example of a relational case representation.

Customer_Name	Xyz	Seats	4
Cusomer_ID	123	Colour	Gray
Age	30	Doors	4
Country	Aaa	Liter	2.5 L
Phone	679118	Horse Power	300
Profession	Researcher	Fuel_Capacity	20
Favourite_Colour	Silver		
Supplier_Phone			
Supplier_Name	My_supplier	Recommendation	My supplier type I
Price	$55,000	Outcome	Deal

Figure 2.8 Tuple-based case representation

After considering all the above factors, a relational case representation is chosen for the proposed approach. The criteria used for the decision problem form the attributes. The criteria values are chosen on an absolute scale. Each attribute would be weighted separately. The weights assigned to the case attributes allow them to have varying degrees of importance and may be selected by a domain expert or user (Watson et al. 2005).

2.11.2 Case Indexing

Case indexing refers to assigning indexes to cases for future retrieval and comparison. The choice of case index is important to enable retrieval of the right case at the right time (Kolodner 1993). This is because the case index will determine the context in which it will be retrieved in the future. Indexes must be predictive in a useful manner (Shin & Han 2001). Indexes should reflect the important features of a case and the attributes that influence the outcome of the case, and describe the circumstances in which a case can be retrieved in the future.

2.11.3 Case Retrieval

Case retrieval is the process of finding, within a case base, those cases that are the closest to the current case (Lenz 1999). To carry out effective case retrieval, there must be selection criteria that determine how a case is judged to be appropriate for retrieval and a mechanism to control how the case is judged to be appropriate for retrieval and a mechanism to control how the case base is searched. The selection criteria are necessary to determine which is the best case to retrieve, by determining how close the current case is to the case stored.

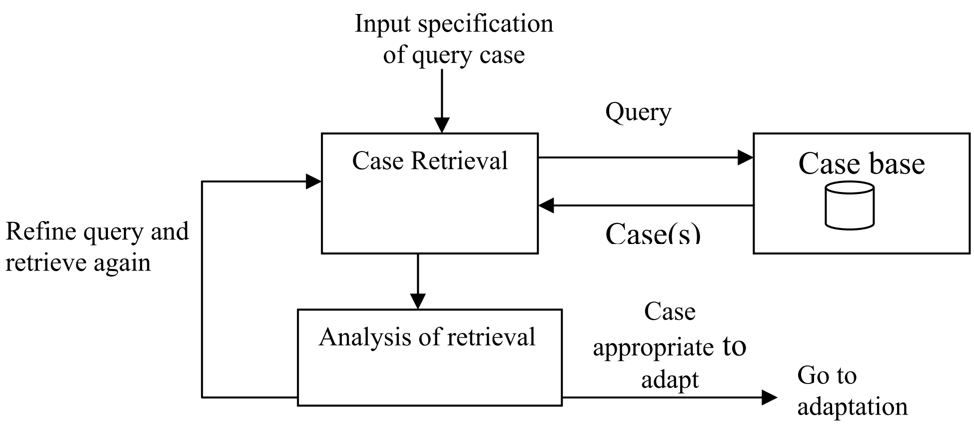


Figure 2.9 Case retrieval process

There are various retrieval techniques (Pal & Shiu 2004; Smyth 1998a; Cheetham 2000). The most commonly investigated retrieval techniques, by far, are the k-nearest neighbours (k-NN), decision trees and their derivatives (Perera & Watson 1998). These techniques involve developing a similarity metric that allows closeness (i.e. similarity) among cases to be measured. These techniques are described below.

2.11.3.1 K-Nearest-neighbour retrieval

In this approach, the retrieved case is chosen when the weighted sum of its features that match the current case is greater than other cases in the case base (Finnie & Sun 2002). In simple terms with all features weighted equally, a case that matches the present case on n features will be retrieved rather than a case that matches only on k features, where $k<n$. Attributes that are considered more important in a problem-solving situation may have their importance denoted by weighting them more heavily in the case-matching process.

2.11.3.2 Inductive approaches

Inductive approaches are used to determine the case-base structure, which determines the relative importance of features for discriminating among similar cases. The resulting hierarchical structure of the case base provides a reduced search space for the case retriever. This may, in turn, reduce the query search time (Bonzano et al. 1997; Falkman 2000).

2.11.3.3 Knowledge-guided approaches

Knowledge-guided approaches to retrieval use domain knowledge to determine the features of a case that are important for retrieving that case in the future (Barletta 1991). In some situations, different features of a case will have different levels of importance or contribution to the success levels associated with that case. As with inductive approaches to retrieval, knowledge-guided indexing may result in a hierarchical structure, which can be more effective for searching. This approach is important for case search.

2.11.3.4 Validated Retrieval

This is a 2-phase approach (Simoudis 1992). Phase 1 involves the retrieval of all cases that appear to be relevant to a problem, based on the main features of the present case. Phase 2 involves deriving more discriminating features from the initial group of retrieved cases to determine whether these cases are valid in the current situation. The advantage of validated retrieval is that inexpensive computational methods can be used to make the initial retrieval from the case base, while more expensive computational methods can be used in the second phase, where they are applied to only a subset of the case base. There are number of factors to consider when determining the method of retrieval:

- The number of cases to be searched
- The amount of domain knowledge available
- The ease of determining weightings for individual features
- Whether all cases should be indexed by the same features or whether each case may have features those vary in importance.

2.11.4 Case Adaptation

Case adaptation is the process of transforming a solution retrieved into a solution appropriate for the current problem (Leake 1995). It has been argued that adaptation may be the most important step of CBR since it adds intelligence to what would otherwise be simple pattern matchers (Pal & Shiu 2004). Case adaptation plays a central role in defining ability of a CBR system. The ability of a CBR system depends upon the adaptation of retrieved solution for solving new problems (Kinley et al. 1996). A number of approaches can be taken to carry out case adaptation.

- The solution returned (case retrieved) could be used as a solution to the current problem without modification, or with modifications where the solution is not entirely appropriate for the current situation.
- The steps or processes that were followed to obtain the earlier solution could be rerun without modification, or with modifications where the steps taken in the previous solution are not fully satisfactory in the current situation.
- Where more than one case has been retrieved, a solution could be derived from multiple cases or, alternatively, several alternative solutions could be presented.

The case adaptation process is shown below.

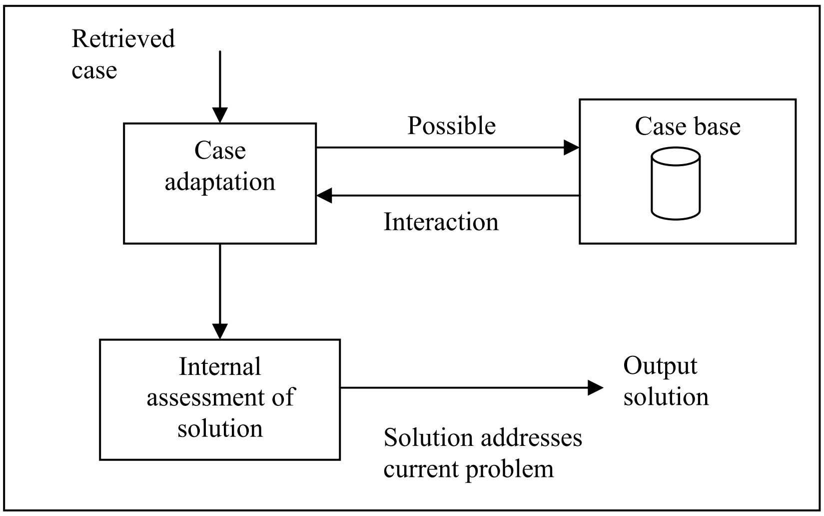


Figure 2.10 Case adaptation process

2.11.5 Case-base Maintenance

When applying CBR systems for problem solving, there is always a trade-off between the number of cases to be stored in the case library and retrieval efficiency (Smyth 1998c). The larger the case library, the greater the problem space covered. However, this would also downgrade system performance if the number of cases were to grow unacceptably high. Therefore, removing redundant or less useful cases to attain an acceptable error level is one of the most important aspects in maintaining CBR systems.

Case-base maintenance can be defined as the implementation of policies for revising the organisation or contents (representation, domain content, accounting information, or implementation) of a case base to facilitate future reasoning for a particular set of performance objectives (Wilson & Leake 1998).

The central idea of CBR maintenance is to develop some measures for case competence, which is the range of problems that a CBR system can solve. Various properties may be useful, such as the size, distribution, and density of cases in the case base, the coverage of individual cases and the similarity and adaptation knowledge of a given system (Smyth 1998b). Coverage refers to the set of cases that each case could solve and reachability refers to the set of cases that could provide solutions to the current problem. The higher the density of cases, the greater the chances of having redundant cases. By expressing the density of cases as a function of case similarity, a suitable case deletion policy could be formulated for removing cases that are highly reachable from others.

Another reason for CBR maintenance is the possible existence of conflicting cases in the case library due to changes in domain knowledge or specific environments for a given task. For example, more powerful cases may exist that can contain inconsistent information, either with other parts of the same case or with original cases that are more primitive. Furthermore, if two cases are considered equivalent (with identical feature values), or if one case subsumes another by having more feature criteria, a maintenance process may be required to remove the redundant cases.

After the adaptation has been completed, it is desirable to check that the solution adapted takes into account the differences between the case retrieved and the current problem (i.e. whether the adaptation specifically addresses these differences). At this point, there is also a need to consider what action is to be taken if this check determines that the solution developed is ready for testing and use in the applicable domain.

2.12 Learning in CBR Systems

Once an appropriate solution has been generated and output, there is some expectation that the solution will be tested in reality. To test a solution, one has to consider both the way it may be tested and how the outcome of the test will be classified as a success or a failure. This learning mechanism is shown in figure 2.11. In other words, some criteria need to be defined for the performance rating of the preferred solution. Using this real-world assessment, a CBR system can be updated to take into account any new information uncovered in the processing of the new solution. This information can be added to a system for two purposes: first, the more information that is stored in a case base, the closer the match found in the case base is likely to be; second, adding information to the case base generally improves the solution that the system is able to create (Althoff 2001).

Learning may occur in a number of ways (Segre 1989; Hammond 1989; Plaza et al. 1997). The addition of a new problem, its solution, and the outcome to the case base is a common method. The addition of cases to the case base will increase the range of situations covered by the stored cases and reduce the average distance between an input vector and the closest stored vector.

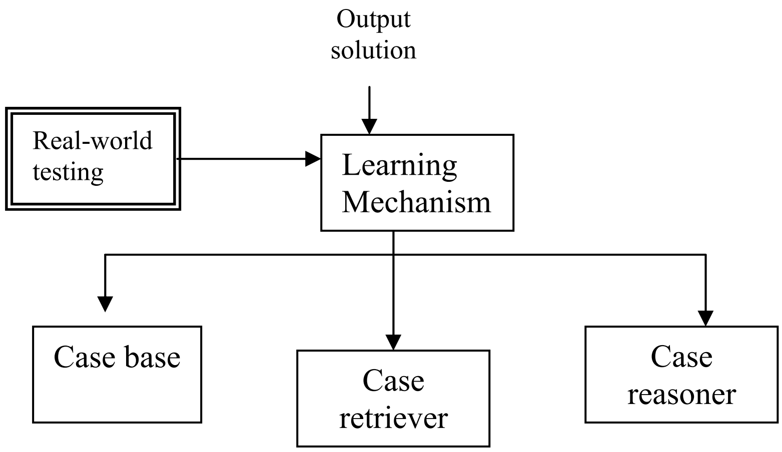


Figure 2.11 Learning Mechanism in CBR

A second method of learning in a CBR system is using the solution's assessment to modify the indexes of the stored cases or to modify the criteria for case retrieval. If a case has indexes that are not relevant to the specific contexts in which it should be retrieved, adjusting the indexes may increase the correlation between the occasions when a case is actually retrieved and the occasions when it ought to have been retrieved. Similarly, assessment of a solution's performance may lead to an improved understanding of the underlying casual model of the domain that can be used to improve adaptation processing. If better ways can be found to modify case with respect to the distance between the current and retrieved cases, the output solution will probably be improved.

2.13 Advantages of Case-based Reasoning

There are various advantages of using CBR for solving the undertaken research problem. These advantages are enunciated next.

2.13.1 Simplification of the Knowledge Acquisition Task

CBR consist primarily of the collection of relevant existing experiences/cases and their representation and storage. Unlike other model-based and rule-based systems, CBR does not require to depend upon a particular model or a set of rules for knowledge acquisition (Pal & Shiu 2004). This simplifies the task of knowledge acquisition. In the undertaken problem, the data about the past decisions will form the case library. It is assumed that such data would be easily available from the experts and since it is independent of any rules, it is easier to built.

2.13.2 Useful for Continuous Learning

CBR is useful for building systems that records success as well as failure. In systems that record failures as well as success, and perhaps the reason for those failures, information about what caused failures in the past can be used to predict potential failures in the future (Cunningham et al. 1997).

2.13.3 Flexibility in Knowledge Modelling

Due to their rigidity in problem formulation and modelling, model-based systems sometimes cannot solve a problem that is on the boundary of their knowledge or scope or when there is missing or incomplete data. In contrast, case-based systems use past experience as the domain knowledge and can often provide a reasonable solution, through appropriate adaptation, to these types of problems (Wilke & Bergmann 1998).

2.13.4 Reasoning in Ill-defined Domains

In a situation where insufficient knowledge exists to build a casual model of a domain or to derive a set of heuristics for it, a case-based reasoner can still be developed using only a small set of cases from the domain. The underlying theory of domain knowledge does not have to be quantified or understood entirely for a case-based reasoner to function (Pal & Shiu 2004).

2.13.5 Making Successful Predictions of a Preferred Solution

When information is stored regarding the level of success of past solutions, the case-based reasoner may be able to predict the success of the solution suggested for a current problem. This is done by referring to the stored solutions, the level of success of these solutions and the differences between the previous and current contexts of applying these solutions.

2.13.6 Learning Over Time

As CBR systems are used, they encounter more problem situations and create more solutions. If solution cases are tested subsequently in the real world and a level of success is determined for those solutions, these cases can be added to the case base and used to help in solving future problems. As cases are added, a CBR system should be able to reason in a wider variety of situations and with a higher degree of refinement and success (Pirolli & Anderson 1985; Leake & Wilson 1999).

2.13.7 Reasoning with Incomplete or Imprecise Data and Concepts

As cases are retrieved, they may not be identical to the current case. Nevertheless, when they are within some defined measure of similarity to the present case, any incompleteness and imprecision can be dealt with by a case-based reasoner. Although these factors may cause a slight degradation in performance, due to the increased disparity between the current and retrieved cases, reasoning can continue. In such situation a fuzzy logic based methodology works well (Hansen and Riordan 2002). This particular advantage supports the application of case-based reasoning for solving the undertaken research question.

2.13.8 Providing a Means of Explanation

Case-based reasoning systems can supply a previous case and its (successful) solution to help convince a user of, or to justify, a proposed solution to the current problem. In most domains there will be occasions when a user wishes to be reassured about the quality of the solution provided by a system. By explaining how a previous case was successful in a situation, using the similarities between the cases and the reasoning involved in adaptation, CBR system can explain its solution to the user. Even for a hybrid system, one that may be using multiple methods to find a solution, this proposed explanation mechanism could augment the casual (or other) explanation given to a user (Pal & Shiu 2004).

2.13.9 Applicability to Various Problems

The number of ways in which a CBR system can be implemented is almost unlimited. It can be used for many purposes, such as creating a plan, making a diagnosis, and arguing a point of view. Therefore, the data dealt with by a CBR system are able to take many forms, and the retrieval and adaptation methods will also vary. Whenever stored past cases are being retrieved and adapted, case-based reasoning is said to be taking place. CBR can be applied to extremely diverse application domains (Osborne & Bridge 1996; Schmidt et al. 1990; Watson & Gardingen 1999).

2.13.10 Reflecting Human Reasoning

As there are many situations where we, as humans, use a form of case-based reasoning, it is not difficult to convince implementers, users and managers of the validity of the paradigm. Similarly, humans can understand a CBR system's reasoning and explanations and are able to be convinced of the validity of an earlier solution, they are less likely to use this solution. The more critical domain, the lower the chance that a past solution will be used and the greater the required level of a user's understanding and credulity (Dillon & Shiu 2000). This is most important advantage of using CBR for addressing the undertaken research question.

2.14 Similarity Measures

The concept of similarity is central to the theory of case-based reasoning (Pal & Shiu 2004). The meaning of similarity always depends on the underlying context of a particular application, and it does not convey a fixed characteristic that applies to any comparative context. In CBR, the computation of similarity becomes a very important issue in the case retrieval process. In this process utility of a retrieved case determines the effectiveness of a similarity measurement. Therefore establishing an appropriate similarity function is important for handling the deeper or hidden relationships between the relevant objects associated with the cases. Broadly speaking, there are two major retrieval approaches (Liao et al. 1998).

a. The first approach is based on the computation of distance between cases, where the most similar case is determined by the evaluation of a similarity measure (i.e. a metric).
b. The second approach is related more to the representational/indexing structures of the cases. The indexing structure can be traversed to search for similar cases.

The undertaken work uses the first approach of case retrieval. In this approach several similarity measures are used to compute the most similar case. Basically, a similarity measure can be considered as the following ratio –

$$Similarity = \frac{Common}{Common + Different} \tag{2.49}$$

This definition can be expressed in following terms as well.

$$Similarity = 1 - \frac{Different}{Common + Different} \tag{2.50}$$

The above definition is represented in different forms by various similarity measures (Kontkanen et al. 2000). Some of the most commonly used similarity measures are explained below (Kardi 2004).

2.14.1 Euclidean Similarity Measure

This is one of the most commonly used similarity measure (Santini & Jain 1999). The similarity is computed based on the distance between the two objects in Euclidean space. The distance is calculated as the square root of the sum of squares of the arithmetical differences between the corresponding coordinates of two objects.

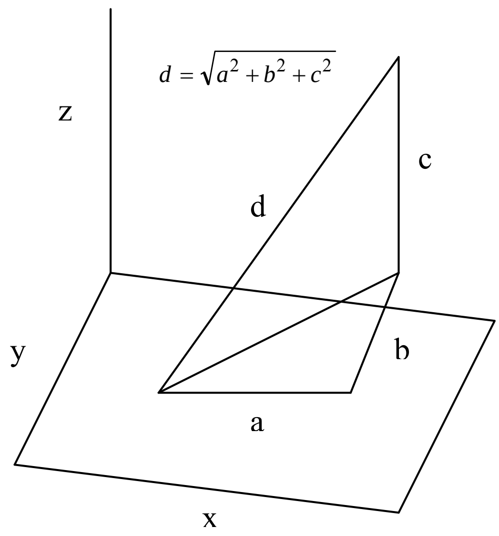


Figure 2.12 Euclidean distance representation

For n-dimensional space, this distance measure is an extension of the Pythagoras theorem.

This measure can be used to compute the similarity between the cases. In the undertaken work, the case attributes are weighted individually. Hence the weighted Euclidean distance can be used to compute the case similarity.

Let $CB = \{CB_1, CB_2, ..., CB_N\}$ be the case library with N cases. Each case in this library is composed of n attributes. The target case is also of the same form as the case from the case library. If CB_k is the target case and CB_j is the j^{th} case from the case library with x_{ki} and x_{ji} $(1 \le i \le n)$ as the values of the corresponding case attributes and $w_i (w_i \in [0,1])$ is the weight of the attributes, then the weighted Euclidean distance between them is given by

$$dist(CB_k, CB_j) = \left[\sum_{i=1}^{n} w_i^2 \left(x_{ki} - x_{ji} \right)^2 \right]^{1/2} \tag{2.51}$$

The case from the case library with a smallest distance can be considered as the most similar case and thus can be retrieved for solving new problem.

2.14.2 Simple Matching Coefficient

This similarity measure is applicable for binary data (Cheetham & Hazel 1969). This similarity measure is useful when both positive and negative values of an attribute carry equal information. For example, attribute gender (male and female) has symmetry attribute because number of male and female give equal information.

This measure is expressed as

$$Sim_{ij} = \frac{p+s}{t} \tag{2.52}$$

where

p = Number of variables that are positive for both objects

s = Number of variables that are negative for both objects.

q = Number of variables that are positive the i^{th} object and negative for the j^{th} object.

r = Number of variables that are negative for the i^{th} object and positive for the j^{th} object.

t = total number of variables = p + q + r + s

The distance between the two objects is given by –

$$d_{ij} = \frac{q+r}{t} \tag{2.53}$$

2.14.3 Jacquard's Coefficient

Jacquard's coefficient measure and Jacquard's distance measure are measurement of asymmetric information on binary variables (Gower 1971; Gregson 1975). These measures are derived from the simple matching coefficient discussed above. The simple matching coefficient measures the similarity between the attributes for both positive and negative values. However, in some applications, only one type of attribute value is required. For some applications the existence of 's' in simple matching makes no sense when number of variables that are negative for both objects do not contribute towards similarity computation. Therefore such variables need not be counted for computing the similarity or dissimilarity. This may happen when the value of positive and negative do not have equal information (asymmetry) (Bridge 1998).

The Jacquard's coefficient is given by

$$S_{ij} = \frac{p}{p+q+r} \tag{2.54}$$

where,

p = Number of variables that are positive for both objects

q = Number of variables that are positive for the i^{th} objects and negative for the j^{th} object.

r = Number of variables that are negative for the i^{th} objects and positive for the j^{th} object.

Jacquard's distance can be derived from the Jacquard's coefficient as

$d_{ij} = 1 - S_{ij}$

i.e.

$$d_{ij} = \frac{q+r}{p+q+r} \tag{2.55}$$

It is shown that these coefficients are ordering measures (Sneath & Sokal 1973). They are not equivalent similarity measures.

2.14.4 Hamming Distance

A finite binary 0 and 1 sequence is sometimes called a word in coding theory (Hamming 1986). If two words have the same length, we can count the number of digits in positions where they have different values. The total number of digits having different values from each other is called the Hamming distance. This measure can be explained as follows.

If q = number of variables with value 1 for the i^{th} object and 0 for the j^{th} object and r = number of variables with value 0 for the i^{th} object and 1 for the j^{th} object. The expression for the hamming distance is $d_{ij} = q + r$.

2.14.5 Manhattan (City-block) Distance

Similarity between two cases can be measured on an absolute scale as well. Manhattan or City-block distance measure is based on the absolute values of the attributes (Aksoy & Haralick 2000). It is similar to the walking distance between two points in a city like New York's Manhattan district where each component is the number of blocks in North-South and East-West directions. It can be represented graphically as shown in the following figure 2.13.

It is represented as follows.

$$d_{ij} = \sum_{k=1}^{n} \left| x_{ik} - x_{jk} \right| \tag{2.56}$$

This measure considers the absolute values of the attributes.

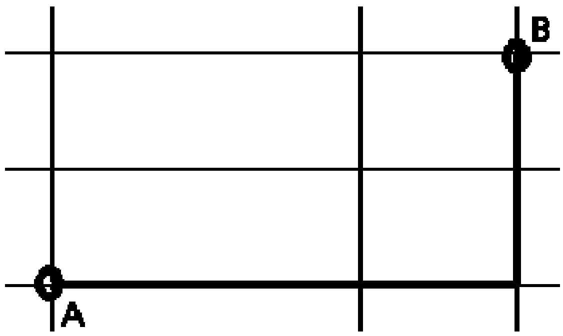


Figure 2.13 City block distance measure

2.14.6 Minkowski Distance

This is the generalized metric distance (Ichino 1994). This distance measure can be expressed as follows.

$$d_{ij} = \sqrt[p]{\sum_{k=1}^{n} \left| x_{ik} - x_{jk} \right|^{p}} \tag{2.57}$$

The value of p determines the type of distance.

When p=1 this distance measure becomes city block distance.

When p=2 this distance becomes Euclidean distance

2.14.7 Cosine Similarity

The cosine angle between any two vectors representing documents, queries, snippets or combination of these is also used as a similarity measure. The expressions cosine similarity, Sim (A, B), or COSIM are commonly used (Baeza-Yates & Ribeiro-Neto 1999).

$$Sim(A,B) = \cos(\theta) = \frac{A \bullet B}{|A||B|} = \frac{x_1 * x_2 + y_1 * y_2}{(x_1^2 + y_1^2)^{1/2}(x_2^2 + y_2^2)^{1/2}} \tag{2.58}$$

2.14.8 Mahalanobis Distance Measure

This is a distance measure based on the correlations between variables by which similarity of a new data set with a known data set can be determined. It is not dependent on the scale of measurements (Duda et al. 2000). It is defined as follows –

If $x=(x_1,x_2,x_3,......,x_p)$ is a multivariate vector with mean $\mu = (\mu_1, \mu_2, \mu_3, ... , \mu_p)$ and a covariance matrix Σ , then the Mahalanbois distance is given as

$$D_M(x) = \sqrt{(x-\mu)^T \Sigma^{-1} (x-\mu)} \tag{2.59}$$

If the co-variance matrix is the identity matrix, then the Mahalanbois distance measure is the same as the Euclidean distance measure.

This distance measure takes into account the correlation between the features and normalizes each feature to zero mean and unit variance.

$$d(X,Y) = (X - Y)^T W^{-1} (X - Y) \quad (2.60)$$

where W is the pooled covariance matrix and W^{-1} denotes its inverse. Assume that there are k groups (or clusters) of data, each having n_k points. These could correspond to data belonging to a node of the index tree of a multimedia database. The mean vector of the k^{th} group is

$$m^{(k)} = \left[m_1^{(k)}, m_2^{(k)}, \ldots, m_n^{(k)}\right]^T \quad (2.61)$$

where

$$m_i^{(k)} = \frac{1}{n_k} \sum_{j=1}^{n_k} x_i^{(k)}$$

The pooled mean m is the grand mean for all points:

$$m = \frac{1}{N} \sum_{j=1}^{k} n_k m^{(k)} \qquad \text{where} \quad n = \sum_{j=1}^{k} n_k \quad (2.62)$$

Now, this normalization can be done:

$$X^{(k)} = X^{(k)} - m \quad (2.63)$$

Then, the pooled sample covariance matrix is defined as:

$$W = \sum_{i=1}^{k} \sum_{j=1}^{n_k} \left(X_j^{(i)}\right)\left(X_j^{(i)}\right)^T \quad (2.64)$$

Mahalanobis distance takes into account correlation between features and normalizes each feature to zero mean and unit variance. It must be noted that if W is an identity matrix, i.e. the features are uncorrelated, and then Mahalanobis distance reduces to Euclidean distance.

2.14.9 Correlation Coefficient

Correlation coefficient is measure, which indicates the extent of correlation between the two variables in a given data set. It is a measure of interdependence between two random variables that ranges in value from −1 to +1, indicating perfect negative correlation at –1, absence of correlation at zero and perfect correlation at +1. It is a statistic representing how two variables co-vary. If there are two variables X and Y, with means $\bar{X}$ and $\bar{Y}$ respectively. The standard deviations of these variables S_X and S_Y respectively. The correlation is computed as

$$r = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{(n-1)S_X S_Y} \tag{2.65}$$

Another correlation coefficient used in the Pearson product-moment correlation coefficient (Ahlgren et al. 2003).

$$P = \frac{\sum_{i=1}^{n}(x_i - \bar{x}_i)\sum_{i=1}^{n}(y_i - \bar{y}_i)}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x}_i)^2 \sum_{i=1}^{n}(y_i - \bar{y}_i)^2}} \tag{2.66}$$

Note that $\bar{x}_i$ and $\bar{y}_i$ is the mean of the i^{th} feature over all the entries in the database. This measure is based on the moments around the mean and hence measures the mismatch in the features. This similarity measure is used mostly in cluster analysis.

2.15 Analysis of Similarity Measures

The above similarity measures can be applied in various problems. The application of a particular similarity measure would depend upon the nature of the problem (Geng & Hamilton 2006). While choosing a particular similarity measure, the computational complexity and its applicability for a given data set can be considered.

In a given decision problem, the values of the attributes can exist on different scales. The attributes can mostly be expressed in an ordinal scale or absolute scale. From a computational perspective, the similarity measure that can compute the similarity between the two data sets based on the default attribute scale can be effective.

Various important and commonly used similarity measures are discussed. Similarity measures such as simple matching coefficient, Jacquard's coefficient, Hamming distance measure are mainly used for binary data. This restricts their applicability for the undertaken work as the attribute values in a given decision problem may not be expressed in binary scale only. Mahalanbois distance measure considers the correlations between the variables and is applicable only if the variables are correlated. The criteria are scored individually. This does not imply a correlation between the criteria scores.

Euclidean (and squared Euclidean) distances are usually computed from raw data, and not from standardized data (Nguyen & Ho 2000). This method has certain advantages (e.g., the distance between any two objects is not affected by the addition of new objects to the analysis, which may be outliers). However, the distances can be greatly affected by differences in scale among the dimensions from which the distances are computed (Wilson & Martínez 1997). For example, if one of the dimensions denotes a measured length in centimetres, and then it is converted to millimetres (by multiplying the values by 10), the resulting Euclidean or squared Euclidean distances (computed from multiple dimensions) can be greatly affected (i.e., biased by those dimensions which have a larger scale), and consequently, the results may be very different. Generally, it is good practice to transform the dimensions so that they have similar scales.

This transformation may have a significant impact on the computation process for a large data set. Hence, it is clear that a similarity measure that can be applied directly to a given data set will perform better in terms of computational cost. It will be mainly easy to apply in practical problems such as the fuzzy measure determination problem.

Unlike the measures discussed above, City-block distance measure can be applied to a data set without any attribute value conversion. City block distance is computationally cheaper distance measure. It requires attributes in an absolute scale. In the undertaken problem of decision-making, the decision attributes can be best expressed in an absolute scale and ordinal scale. Unlike other similarity measures, it can be applied on continuous data. This property of the city-block distance measure makes it more suitable and easier to apply for the undertaken research. Hence city-block distance measure is chosen for developing the proposed approach.

2.16 Summary

This chapter explored the literature relevant to the issues of dependence between criteria. The process of data aggregation is a central aspect. The analysis of various aggregation operators showed that non-additive measures have the ability to model the dependence between the criteria. The critical analysis of the fuzzy measures determination methods identified the need of developing a new approach using similarity-based reasoning. This chapter has laid the theoretical foundation of the proposed methodology for the fuzzy measure acquisition. Case-based reasoning is proposed as the tool to implement the proposed methodology. The proposed approach and its investigation will be presented in the next chapter.

Chapter 3 Proposed Methodology

3.1 Introduction

Literature analysis suggests that the fuzzy measure determination issue can be addressed by using a dataset of past precedents. Various methods for determining fuzzy measures were discussed earlier. The issue of fuzzy measure determination is complex. It is not only because of the exponential (2^n-1) requirement of the fuzzy measure coefficients, but also because of the issues associated with their pragmatic determination. In addition to this complexity, the issue of fuzzy measure determination is subjective as well. This research suggests that similarity-based reasoning can solve these pragmatic issues associated with these monotonic set functions. This chapter discusses the proposed methods for fuzzy measure acquisition in a practical situation. These methods stand as the original contribution of this research. The proposed methodology can be stated as an approach that applies a data set consisting cases of past decision problems for determining fuzzy measures and global evaluation of alternatives in the new decision problem. It is stated earlier that Case-based reasoning is a suitable methodology for solving new problems using the solutions of the past problems (Watson 1998). In the context of this research, CBR can be used for solving a new decision problem based on the past decision problems and hence it would be useful for fuzzy measure determination. Due to the advantages of CBR in applying similarity-based reasoning, it is the central aspect in the development of the proposed approach. The application of similarity-based reasoning approach relies on past cases. In order to maintain the consistency between the new decision problem and the past cases, the contextual information about the decision making scenario is build in the case base (Adomavicius et al. 2005). The determination of values of fuzzy measures and the subsequent global evaluation of the alternatives will depend upon the degree to which the contextual information is incorporated in the case base. In the undertaken research, the contextual information is represented by the attributes and weights of the attributes. The decisions made in the past will act a reference for new decision problems. For a recurrent type of decisions, such case base can be built as the attributes will remain same. Thus, there is consistency between the data of the new decision problem and the case base in recurrent or repetitive type of decisions.

3.2 CBR-based Framework of the Proposed Methodology

CBR involves mapping of the solutions of the past problems to solve the new problems. There have been few models of CBR. Some models consider CBR as a five-step problem solving process (Allen 1994). Aamodt and Plaza (1994) introduced a R^4 model, which describes the process using the four Rs. In another model introduced by Leake (1996), when the CBR system gets a new problem from the user interface, it normalizes it into a problem description, P_0, and then searches the case base for a prior problem description which is most similar to the current problem description, p_1, that is, $P_1 \approx P_0$. The solution of the retrieved problem description, S_1, is used as the starting point for generating a solution to the new problem S_0. The transformation from a new problem to P_0 is case representation; the transformation process from P_0 to P_1 is case retrieval, transformation from P_1 to S_1 is case building and the transformation process from S_1 to S_0 is case adaptation. This model is used for the development of the proposed methodology.

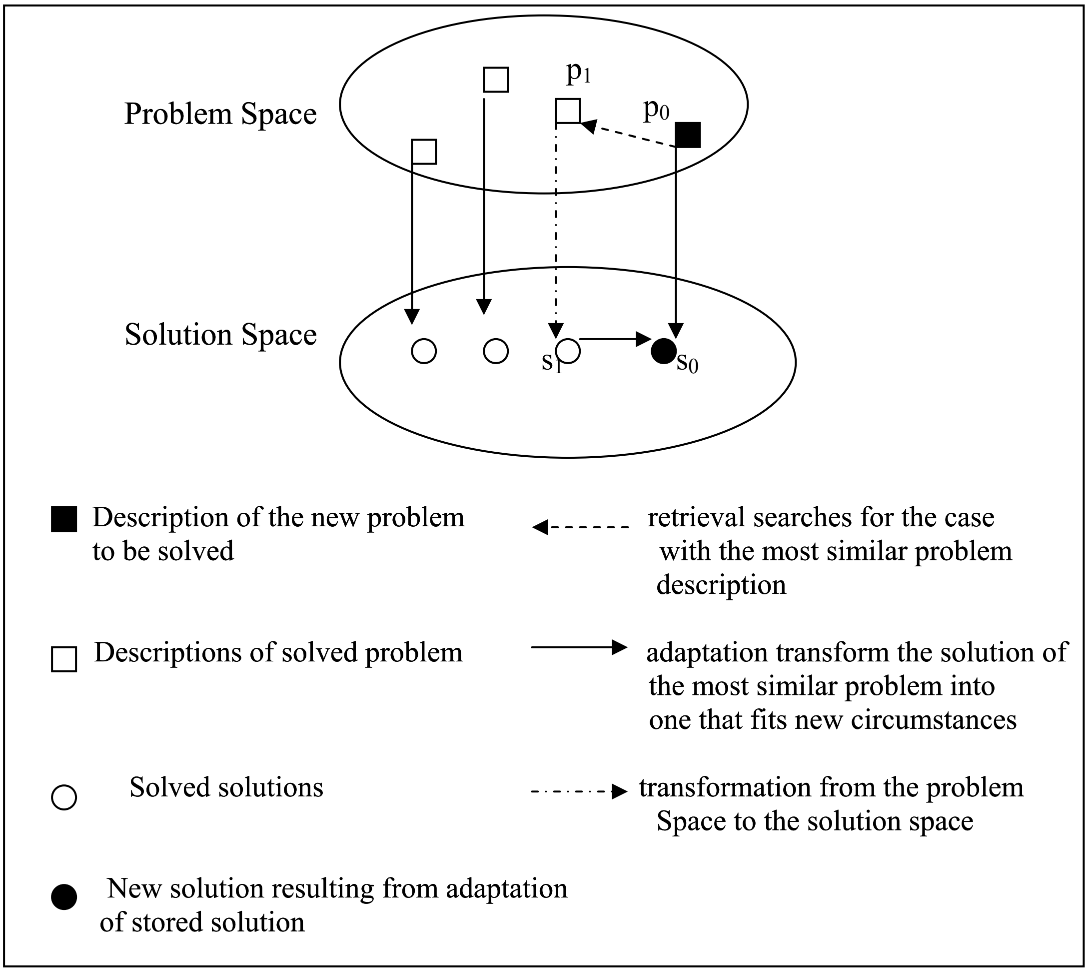


Figure 3.1 Mapping the problem space with case base

The above representation defines the conceptual foundation behind the development of the proposed approach. This representation can be analysed into two parts viz. the problem space and the solution space.

3.2.1 Problem Space

The new decision problem is in the form of MCDM "m x (n+1)" matrix, which models the evaluation process of given set A of alternatives. "n+1" stands for the number of criteria and global evaluation of the alternatives. This forms the problem space where each object is to be evaluated against a set C of criteria. The criteria satisfactions may be represented by the fuzzy membership functions. W_i represents the weight of the individual criteria i. The decision maker defines the weights. However, the importance of the sub-sets of criteria is unknown and hence the global evaluation of the alternatives is unknown. The problem statement is defined as follows:

How to acquire the weights of sub-sets of criteria or fuzzy measures for a given set of criteria in the new decision problem using similar data set? The problem is identified as the issue of the acquisition of fuzzy measures in practical application.

3.2.2 Solution Space

The solution space is a case-base of the past decision problems. The main interest here is to use the solution space and adapt it to solve new problem. This process would involve retrieving the similar case and adapting the case to solve the new problem (Pal & Shiu 2004). The expected outcome from this solution space will be the fuzzy measures of the most similar case from the case. These retrieved fuzzy measures will be adopted or directly applied for solving the new decision problem. The fuzzy measures will be used in determining the global evaluation of the alternatives.

3.3 Proposed System Model

The proposed system has a repository in the form of cases. This repository consists of the data about the attributes and their importance in the past similar situation. It also consists of the global evaluation of the objects in the past decision problems. This repository or the case-base is built using the contextual information about the decision problems.

The decision maker (DM) would provide input to the system in the form of a MCDM matrix for a new decision. Based on the suggested approach for fuzzy measure determination, the system would compute the worth of each alternative. These results are considered as the “first approximation”. DM then validates and confirms the system results in an interactive manner. The new results are added to the repository for future problem solving. Thus a learning mechanism is incorporated in the proposed model. The system learns from its results. The proposed model is shown below.

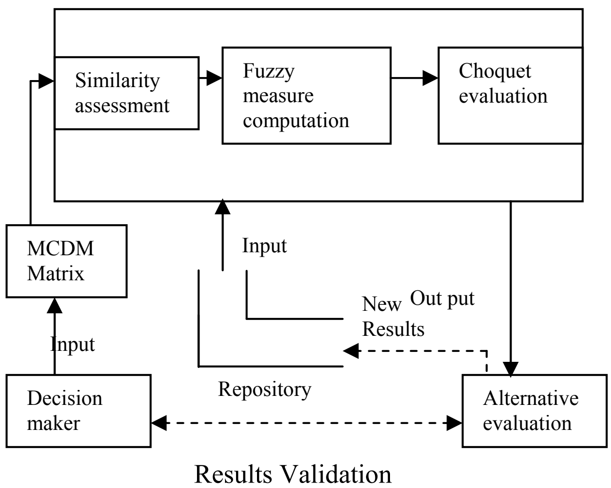


Figure 3.2 Proposed Schematic model

The proposed model is based on the assumption that the there is consistency between the data set in the case-base and the new decision problem. This characteristic of the proposed model makes the decision support system applicable for a class of decision problems where repetitive decisions are made. This task of building such consistent case is challenging. This research appreciates the difficulty involved in building a consistent case-base. It is proposed that this issue can be resolved by using expert input in case-base construction. This approach of using expert input is considered in developing the proposed methodology due to the following reasons:

- Expert input will provide consistent answers for repetitive decision making problems.

- Expert input will help in modeling the contextual information.
- The case base constructed using expert input will act as an authentic reference for solving new problems.
- The expert input-based decision support system can be modified, updated and expanded readily. There is a trade-off between acquisition of expert input and quality of the decision support systems as the process of expert input acquisition can be time consuming. However, the difficulty is acquiring the expert input is compensated by the simplification of the fuzzy measure determination.
- The expert input allows the system to provide an explanation of the reasoning behind its conclusion, which is necessary to provide the credibility and confidence that decision maker require before routinely accepting the decision support.

In figure 3.2, fuzzy measure computation is the important module in evaluating a given set of alternatives. The repository of the past decision problems can be used for this computation. Hence fuzzy measure computation based on the similarity assessment between the new decision problem and the case repository becomes the key aspect of proposed system. This work applies different computational strategies for development of this module. It has already been mentioned that case representation can take various forms. There can be various ways of determining the case similarity. The case similarity assessment involves assessing the similarity between the attributes (Leake et al. 1997). In the undertaken problem, the case similarity is determined based on the individual similarity between the attributes. In the literature review, it is observed that the widely used similarity measures mainly focus on determining the similarity between two single attributes (Wagholikar & Deer 2005a). For the undertaken problem of fuzzy measure determination, one approach is to construct the case-base where an individual attribute represents a sub-set of attributes in the new decision problem and hence it is equivalent to sub-set of attributes.

3.4 Similarity Assessment between Attributes

The similarity computation between the attributes can be explained in the following steps:

1. Let $S = \{s_1, s_2, \ldots\ldots, s_k\}$ be the set of similar attributes for a given case. The case represents the new decision problem. The case base will be build using such cases. $W_s = \{w_{s1}, w_{s2}, \ldots\ldots, w_{sk}\}$ be the importance of the corresponding attributes in the data set S for a given case, where k is the number of attributes >1.

Let $C = \{c_1, c_2, \ldots\ldots, c_n\}$ be the target data set, which are the criteria for a new decision. This data set forms the new case. From the above set of criteria, let c_l be the sub-set of criteria, where l<n.

2. The similarity between s_k and the sub set of attributes $c_{1\ldots n}$ is given as below

$$sim(s_k, c_l) = \sum_{k=1}^{m} \sum_{l=1}^{n-1} sim(s_k, c_l) \quad l = 1, 2, \ldots, n,\ m>1, n>1 \qquad (3.1)$$

That is, the similarity between the k-th single attribute from set S and sub-set of attributes from set C is equal to the algebraic summation of the similarities of individual elements of the c_l with s_k.

3. After this similarity assessment, the fuzzy measure can be modelled as the importance of the attribute s_k, which is the closest attribute from the data set S. Thus fuzzy measure for a sub-set of attributes $c_{1\ldots n}$ is –

$$\mu_C = w\{\max\,[sim\,(s_k, c_{1\ldots n})]\} \qquad (3.2)$$

4. Thus using the above result of fuzzy measures, the global evaluation of alternatives can be computed using fuzzy integral i.e. Choquet integral. Thus the Choquet evaluation as

$$Ch_\mu = \sum_{i=1}^{n} \{f(c_{(i)}) - f(c_{(i-1)})\}\ \mu_C \qquad (3.3)$$

where μ_C is the fuzzy measure of the sub-set of criteria obtained as a result of the similarity computation in eqn (3.2).

The evaluation of R.H.S in eqn (3.1) becomes the central aspect of determination of similarity and hence eventually the fuzzy measure coefficients. Thus the central problem in this algorithm is the computation of similarity between a single attribute and the sub-set of attributes in the past decision problems. Semantic similarity can be used to measure the similarity between these attributes.

3.5 Semantic Similarity Computation

Semantic similarity is used to determine the similarity between the terms based on their semantic content and a given similarity metric (Resnik 1995). There can be various similarity metrics for measuring the semantic similarity. The semantic association between the attributes may indicate the similarity between the attributes.

Latent Semantic Analysis (LSA) is a fully automatic mathematical/statistical technique for extracting and inferring relations of expected contextual usage of words in passages of discourse (Landauer et al. 1998). It is not a traditional natural language processing or artificial intelligence program; it uses no humanly constructed dictionaries, knowledge bases, semantic networks, grammars, syntactic parsers, or morphologies, or the like, and takes as its input only raw text parsed into words defined as unique character strings and separated into meaningful passages or samples such as sentences or paragraphs.

The first step is to represent the text as a matrix in which each row stands for a unique word and each column stands for a text passage or other context. Each cell contains the frequency with which the word of its row appears in the passage denoted by its column. Next, the cell entries are subjected to a preliminary transformation, whose details we will describe later, in which each cell frequency is weighted by a function that expresses both the word's importance in the particular passage and the degree to which the word type carries information in the domain of discourse in general. Next, LSA applies singular value decomposition (SVD) to the matrix. This is a form of factor analysis, or more properly the mathematical generalization of which factor analysis is a special case. In SVD, a rectangular matrix is decomposed into the product of three other matrices. One component matrix describes the original row entities as vectors of derived orthogonal factor values, another describes the original column entities in the same way, and the third is a diagonal matrix containing scaling values such that when the three components are matrix-multiplied, the original matrix is reconstructed. There is a mathematical proof that any matrix can be so decomposed perfectly, using no more factors than the smallest dimension of the original matrix. When fewer than the necessary number of factors are used, the reconstructed matrix is a least-squares best fit. One can reduce the dimensionality of the solution simply by deleting coefficients in the diagonal matrix, ordinarily starting with the smallest. (In practice, for computational reasons, for very large corpora only a limited number of dimensions—currently a few thousand— can be constructed). LSA is one approach for semantic similarity computation. This work introduces a new approach for semantic similarity computation.

In the undertaken research, similarity between the individual attributes of the new problem and the given case would determine the similarity of the entire case. The attribute selection would depend on the context of the problem (Pal & Shiu 2004). The attributes can have semantic relationship between them. The weight of the attribute acquired in the past decision problem would represent the fuzzy measure for the new decision, if this attribute is the most similar attribute from the case-base. It requires taxonomy of attributes for deriving the semantic relationship between the attributes.

3.5.1 Assumptions for the approach

The following assumptions are considered for this approach.

- The criteria in the new decision problem and the criteria in the case-base are part of a hierarchical taxonomy.
- The criteria in the case base of past decisions are considered as semantically superior attributes. This means that they exist at a higher node in a hierarchical taxonomy. These attributes are termed as "super-type" attributes.
- The criteria in the new decision problem are semantically related with criteria in the case-base. They exist at lower nodes in a hierarchical taxonomy than the super-type attributes. These attributes are termed as "sub-type" attributes.
- Weights of the super-type attributes would determine the weights of the sub-sets of criteria. These weights are acquired from an expert.
- In a given decision problem, the decision attributes can be modelled as sub-type attributes which are semantically related with attributes in the past decision problem. However, the importance of the attributes in the past decisions is known. Hence it can be assumed that the semantically similar attributes in the new

decision problem can have the same importance. In other words, the contextual information is assumed to be same. This assumption is applicable for repetitive type of decisions where the attributes do not change. This assumption makes this approach as the first approximation as discussed earlier.

- It is possible to have sub set of attributes in the new decision problem semantically similar to only one attribute in the case base.

3.6 Fuzzy measure determination using Taxonomical data

The above discussion leads to the investigation of means to determine semantic similarity. Resnik (1995) suggests various approaches to determine the similarity between the attributes, which are words or concepts. There are several potential types of relations. It can be hierarchical (e.g. IS-A or hypernym-hyponym, part-whole etc), associative (e.g. cause-effect), equivalence (synonymy), etc. Out of these types IS-A type of association is commonly used. In this association hypernym is a generic word that is broader than the other words. For example, vehicle represents all the objects denoted by the words train, automobile, airplane and is therefore hypernym of these words. In general, hypernym is a word that encompasses the meaning of another word. Hyponym is a word whose extension is included within that of another word. It has been suggested and employed to study a special case of semantic relations (Rada et al. 1989). WordNet (Miller 1990) is used for developing as well as illustrating this approach.

3.6.1 WordNet

WordNet is an electronic lexical reference system. This system is designed using current psycholinguistic theories of human lexical memory. English nouns, verbs, adjectives and adverbs are organised into synonym sets, each representing one underlying lexical concept. Different relations link the synonym sets. This data set is considered to be the most important resource available to the researchers in computational linguistics, text analysis and many areas (Brill 1995; Fellbaum 1998; Resnik 1998). This data set is used for computing the semantic similarity between two criteria in MCDM problem. A general structure of the taxonomy in the Wordnet is shown below.

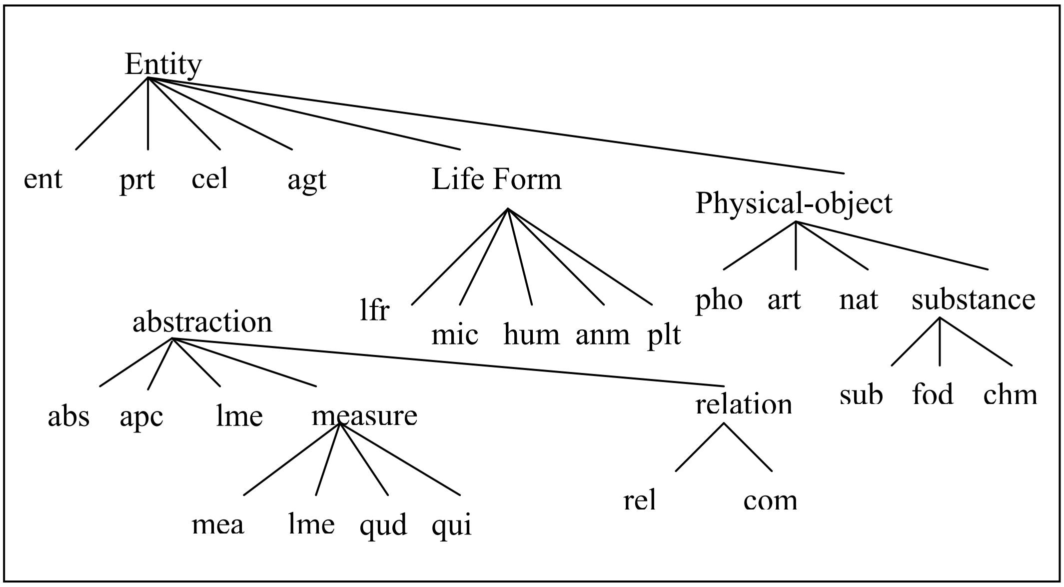


Figure 3.3 WordNet hierarchy

The WordNet hierarchy consists of application of polysemy. Polysemy is the linguistic phenomenon by which a term has more than one meaning. It is widespread both in oral and written communication, due to a human tendency to make a thrifty use of linguistic items. Actually, the adoption of an already defined term to convey a different meaning is a far more common practice than creating complex sentences for each of the meanings to be expressed. A word is judged to be polysemous if it has two senses of the word whose meanings are related. Polysemy allows the use of the same word in different contexts and thus endows language with indispensable flexibility" (Shanon 1993, 45). Since the vague concept of relatedness is the test for polysemy, judgements of polysemy can be very difficult to make. Wordnet is full of instances of polysemy. Therefore a word can have more than one meaning in different taxonomies for different concepts. Therefore the experimentation of semantic similarity based computation will consist of word or concept or an object appearing in more than one hierarchy of concepts. The similarity will not be transitive. However, due to the polysemy, the attribute in a given decision problem may appear in different concept hierarchies. This work has considered the polysemy in Wordnet in proposing semantic similarity based approach for fuzzy measure determination.

3.6.2 Node-based approach for similarity determination

There are two major approaches for semantic similarity determination. The important feature of this approach is that it computes the semantic similarity between the two concepts using the information content shared by them in IS-A type hierarchy (Resnik 1995).

In this approach, a node represents a unique concept consisting a certain amount of information and an edge represents a direct association between two concepts. The similarity between two concepts is the extent to which they share information in common. Each node subsumes one or more sub-nodes in a hierarchy. This hierarchical structure can be viewed as the association between super-class attributes and its corresponding sub-class attributes. Graphically it can be represented as below.

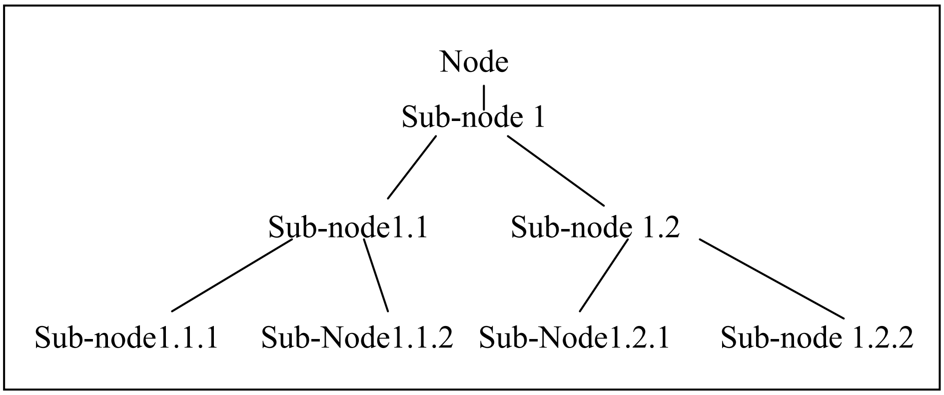


Figure 3.4 IS-A type hierarchy

According to the information theory (Shanon and Weaver 1963), the information content (IC) of a concept can be quantified as $IC(c) = -\log P(c)$, where $P(c)$ is the probability of encountering an instance of concept C. This probability is monotonic and for a top node in the hierarchy the information content is 0 (Lin 1998). Based on this information theory, the similarity between the concepts c_1 and c_2 is given by the following expression.

$$sim(c_1, c_2) = \max_{c \in Sup(c_1, c_2)} [IC(c)] = \max_{c \in Sup(c_1, c_2)} [-\log p(c)] \tag{3.4}$$

Where $\sup(c_1, c_2)$ is the set of concepts that subsume both c_1 and c_2.

This is the similarity between the concepts. The word similarity is given as below

$$sim(w_1, w_2) = \max_{c_1 \in sen(w_1) c_2 \in sen(w_2)} [sim(c_1, c_2)] \tag{3.5}$$

Where sen(w) denotes the set of possible cases for word w.

The results from the above expression can be substituted in eqn (3.1) and thus the overall similarity between the two attributes can be computed. From this proposed computation, it is clear that the probability of encountering an instance of a concept in a given taxonomy is the measure of determining the similarity between the two concepts. This probability can be computed as below.

$$p(c) = \frac{f(c)}{N} \tag{3.6}$$

Where $f(c)$ = frequency of instance and N = Total number of senses for a word.

The given taxonomy can be classified into classes and concepts for implementing the above approach. Words are the instances of the concepts in a given hierarchy. Classes can also be considered as the set of possible senses that the word w is contained. This approach can be illustrated by constructing a sample data set. A simple data set of 3 attributes for a car selection problem is considered.

3.6.2.1 Illustration

Consider the following Car selection problem.

Table 3.1 Data Set for Node-based similarity computation

	Power	Price	Comfort
Weights	0.4	0.3	0.3
Car A	0.4	0.5	0.2
Car B	0.6	0.7	0.9
Car C	0.4	0.3	0.9
Car D	0.8	0.7	0.2

In the above car decision problem, criteria are interactive. A case base of past decision problems is constructed. The attributes in the case base are semantically related with the attributes in the new decision problem. The attributes in this case base are the super-type attributes. The case base will consist of the set of criteria {Economy, Convenience, Value}. The decision maker assigns the corresponding weights of these criteria in the past decision problems. For this illustration, the following case base is constructed.

Table 3.2 Sample Case Base

	Economy	Convenience	Utility
Case 1	0.4	0.3	0.3
Case 2	0.5	0.2	0.3
Case 3	0.2	0.2	0.6
Case 4	0.3	0.5	0.2

A hierarchical taxonomy for the chosen criteria in the case base and new decision problem was constructed. Hypernyms in the word net are used for constructing the case base. Sample taxonomy between a super-type attribute in the case base and decision attributes in the new problem is shown below. The taxonomy for the semantically related attributes is considered. In the following taxonomy, the sub-type attributes are constructed in such a way that there exists a semantic relationship between the super-type attributes. The super-type attributes from the case base are at a higher level in the taxonomy. The possible senses of the super-type attributes were analysed. The sub-type attribute, which is one of the possible senses for the super-type attribute, is used for building this taxonomy.

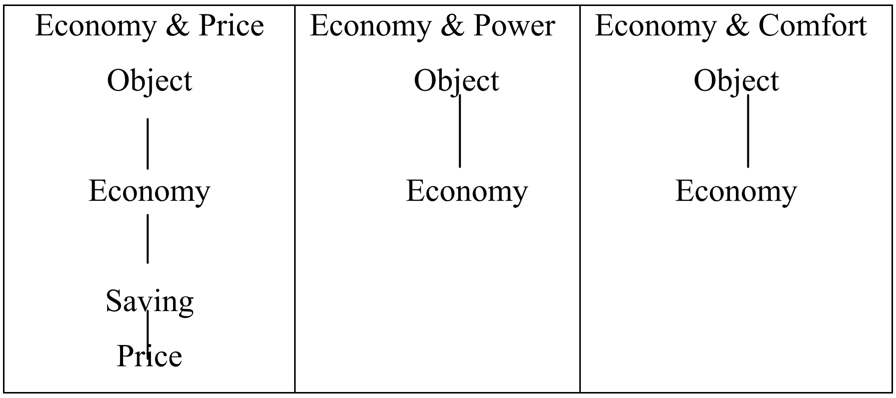


Figure 3.5 Taxonomy of the attributes

The similarity between Economy and power is zero. The similarity between Economy and comfort is zero as there are no instances of these attributes in the hierarchical taxonomy. Such taxonomies were constructed for other super-type attributes in the case base. The similarity computation can be summarised below.

Table 3.3 Information based similarity computation

Attribute	F(c)	N	P(c)	Log p(c)
Economy	1	4	0.25	-0.60
Convenience	1	4	0.25	-0.60
Utility	1	6	0.17	-0.77
Price	1	7	0.14	-0.84
Comfort	1	4	0.25	-0.60
Power	1	9	0.11	-0.95

As per the proposed similarity computation, the similarity between a super attribute and a sub-set of attributes from the new decision problem is the algebraic summation of the individual similarities. The similarity computation is shown below.

Table 3.4 Node-based similarity computations

	Utility	Convenience	Economy
{price, comfort}	0.75	0.6	0.8
{comfort, power}	0.6	0.8	0.7
{price, power}	0.5	0.7	0.65

From the above table, it can be observed that the sub-set of price and comfort is semantically similar to 'Economy'. Hence the weight given to this attribute in the past cases determines the importance of this sub-set. The computed fuzzy measure is the average value of the weight given to the semantically similar attribute in all the past cases. The computed fuzzy measures for each subset are as below.

$$\mu_{price\ ,comfort} = 0.35$$

$$\mu_{price\ ,power} = 0.30$$

$$\mu_{power,comfort} = 0.30$$

Based on the above fuzzy measures the global evaluation of the cars is as below.

Table 3.5 Evaluation of cars

Cars	Global Evaluation
Car A	0.305
Car B	0.705
Car C	0.51
Car D	0.835

From the above results it can be observed that the ranking of the cars is in accordance with the criteria scores. Cars having higher criteria scores are ranked higher than cars having medium and low criteria scores. This result illustrates the information content-based similarity approach for fuzzy measures determination. The ranking of Car D can be argued here. Car D does not outscore Car B on all criteria. Car B scores higher than Car D on the 'comfort' criterion. A decision maker with conjunctive behaviour may prefer an alternative with higher scores on all criteria. This result shows that there is still a scope for applying another approach for semantic similarity computation for this research.

3.6.3 Edge-based Distance Approach

This is another approach for determining the semantic similarity between the criteria (Resnik 2003). This approach estimates the distance between the nodes which correspond to concepts / classes being compared for a given taxonomy (Jiang 1997). The geometric distance between the nodes, representing the concepts, can conveniently measure the conceptual distance. The semantic similarity is given as

$$sim(w_1, w_2) = 2d_{\max} - \min_{c_1 \in sen(w_1) c_2 \in sen(w_2)} len(c_1, c_2) \tag{3.7}$$

where d_{max} = depth of the taxonomy

len = path length of the attribute in a hierarchical taxonomy.

c_1 = super type attribute

c_2 = sub type attribute

Path length is the distance of the current node from the topmost node in the hierarchical taxonomy. The definition of edge-based similarity is applied in the context of the undertaken research problem. Accordingly, the super attribute will have a shorter path length than the sub-attributes. The path length of super-type attributes determines the similarity between the two attributes.

3.6.3.1 Illustration

The following data set is used. The attribute weights are normalised. The weights and the scores or evaluation of each alternative against each attribute are represented in the scale {0,1}.

Table 3.6: Cars evaluation

	Price	Comfort	Power
Weights-	0.4	0.3	0.3
Car A	0.6	0.8	0.4
Car B	0.9	0.6	0.8
Car C	0.7	0.8	0.6
Car D	0.8	0.5	0.6

Based on that data set, the path lengths are calculated as follows.

Table 3.7 Attribute's path length

Attribute	d_{max}	*Len*
Economy	4	2
Convenience	4	2
Utility	6	2
Price	7	1
Comfort	4	3
Power	9	2

The similarity calculation as per the eqn. 3.7 is as shown below.

Table 3.8 Edge-based similarity computation

	Utility	Convenience	Economy
{price, comfort}	0.7	0.8	0.7
{comfort, power}	0.6	0.5	0.7
{price, power}	0.8	0.9	0.5

The computed fuzzy measure is the average value of the weight given to the semantically similar attribute in all the past cases as per the sample case in Table 3.2. The fuzzy measures are as below.

$$\mu_{price,comfort} = 0.3$$

$$\mu_{comfort,power} = 0.35$$

$$\mu_{price,power} = 0.3$$

By using the above similarity assessment and the data in Table 3.6, the following global evaluations were obtained.

Table 3.9 Edge-based similarity results

	Price	Comfort	Power	Global Evaluation
Weights-	0.4	0.3	0.3	
Car A	0.6	0.8	0.4	0.52
Car B	0.9	0.6	0.8	0.70
Car C	0.7	0.8	0.6	0.66
Car D	0.8	0.5	0.6	0.595

From the global evaluations of the cars, it can be observed that the results are largely acceptable as the worth of car B scoring high on the price and relatively high on comfort and space is maximum.

3.7 Discussion

- The semantic similarity based approaches are illustrated with a sample data set. The results given by these two approaches vary. The two approaches give a different preference order of the alternatives. The results mainly depend upon the definition of the hierarchical taxonomy. However, in a practical setting, it is difficult to build customized hierarchical taxonomy, as each decision problem would normally involve a different set of attributes. Wordnet provides a standard hierarchy of words. It does not provide similarity between the concepts.
- These approaches do not consider any agreement of the results with the decision maker. Such approaches might be implemented in an interactive manner. Hence it is worthwhile to develop another approach that can address these issues.

3.8 Determination of Fuzzy Measures by Solving Linear Systems

In the previous section, an approach based on the semantic similarity between the words or concepts was discussed. Another approach that is based on the same notion of using past cases is investigated. It is based on an assumption that the global evaluation of the alternatives in the past is equal to the global evaluation for the new decision problem. This approach is described in the following steps.

1. Consider the problem where a case base is built for a given decision problem. The case base consists of different past decision problems and its solution.

2. After the similarity computation, the most similar cases are retrieved and its corresponding global evaluation is used for computing the fuzzy measures in the new decision problem.
3. This forms the linear systems of equations as below.

$$a_{11}x_1+a_{12}x_2+a_{13}x_3+\dots\dots+a_{1n}x_n= b_1$$
$$a_{21}x_1+a_{22}x_2+a_{23}x_3+\dots\dots+a_{2n}x_n= b_2$$
$$a_{31}x_1+a_{32}x_2+a_{33}x_3+\dots\dots+a_{3n}x_n= b_3$$
$$a_{41}x_1+a_{42}x_2+a_{43}x_3+\dots\dots+a_{4n}x_n= b_4$$
$$\vdots$$
$$a_{n1}x_1+a_{n2}x_2+a_{n3}x_3+\dots\dots+a_{nn}x_n= b_n \qquad (3.8)$$

The above system can be represented in the matrix form as AX=B, where A is the square matrix of criteria evaluations, X is the column vector of fuzzy measure coefficients and B is the column matrix of global evaluations of alternatives.

3.8.1 Case Representation

In this approach it is assumed that the case representation is of the same form as the target problem. The case base consists of a set C of criteria or attributes $\{c_1, c_2,\dots\dots,c_n\}$ for a given set of cases. The case base also stores the global evaluation of the alternatives. This global evaluation is computed in the past decision problems. The attribute values are their absolute values. This case base can be represented as below.

Table 3.10. Absolute value based case representation

	c_1	c_2	c_3	c_4	…….c_n	Global Evaluation
Case 1	x_{11}	X_{12}	x_{13}	x_{14}	………x_{1n}	g_1
Case 2	x_{21}	X_{22}	x_{23}	x_{24}	………x_{2n}	g_2
Case 3	x_{31}	X_{32}	x_{33}	x_{34}	………x_{3n}	g_3
Case 4 … … Case n	.. .. … x_{m1}	.. .. … x_{m2}	.. .. … x_{m3}	.. .. … x_{m4}	.. .. … x_{mn}	.. .. … g_m

3.8.2 Similarity Measurement

In this approach, Manhattan or City-block distance is used for determining the case similarity. As defined earlier, Manhattan distance measure determines the similarity by computing the difference in absolute values of the attributes.

3.8.3 Determination of Fuzzy Measure Coefficients

After the most similar case is retrieved, its global evaluation can be applied to the problem of fuzzy measure determination. It is assumed that the global evaluation of the similar cases is equal to the global evaluation of the target case. This results into formation of a linear system of the form AX=B, where X is the column vector of fuzzy measures, A is the matrix of the coefficients and B is the column vector of global evaluations.

Solution of the above system would result in the set of fuzzy measure coefficients. However, obtaining the solution of the above system is not an easy task. In order to find $X= A^{-1}B$, A must be a square matrix. If the system is a non-square system then a matrix can be converted into a square matrix by introducing certain dummy variables. Iterative methods can also be used for a non-square system. Iterative methods tend to solve the problem by determining successive approximations to the solution starting from an initial guess. Preconditioning of a matrix is an iteration method that can be applied for solving a non-square system (Pan et al. 2005). Increasingly more accurate approximate solutions are obtained by iteratively correcting the available approximations. Effective corrections are usually obtained by solving an appropriate nearby system exactly. In contrast to the original system, this nearby system should be efficiently solvable. This approach is called preconditioning. The 'preconditioner' forms the nearby system (Saad 1996). According to this approach, if there is a system of equations as AX = B, then construct a matrix B^{-1} that approximates matrix A. A is symmetric and positive semi-definite matrix. This system can then be converted to a system of the following form using a pre-conditioning matrix.

$$B^{-1}AX = b\,B^{-1}$$

By using this pre-conditioner matrix, coefficients of X can be determined. However, there are certain conditions for the construction of pre-conditioner matrix. They are –

a. Product $B^{-1}A$ must be close to the identity.

b. B should be easily inverted (for instance, B is a diagonal or triangular matrix)
c. The pre-conditioned system is solved easily and faster than the original system.

The construction of the pre-conditioned matrix is a subjective issue associated with this approach. The pre-conditioned matrices are used in iterative methods. In this problem, the solution for the above system exists only when the system is a square system. This approach can be illustrated by the following example.

Table 3.11 Case base of the past decisions for linear systems

	Price	Power	Comfort	Global Evaluation
Weights	0.3	0.2	0.3	
Case 1	$25000	1000	Good	0.89
Case 2	$30000	1500	OK	0.7
Case 3	$25000	1000	OK	0.6
Case 4	$20000	1200	Good	0.7

The above data set is about the past decisions. The data set for the new decision problem is shown below.

Table 3.12 Data set for new decision problem

	Price	Power	Comfort
Weights	0.4	0.3	0.3
Car A	0.4	0.6	0.3
Car B	0.4	0.6	0.5
Car C	0.5	0.6	0.7
Car D	0.4	0.3	0.3

The above dataset was used for computing the similarity using SPSS for windows. It is assumed that the global evaluation from the past decisions will be approximately same for the alternatives in the new decision problem. From the similarity assessments using City-block distance measure, the following results were obtained.

Global evaluation of Car A = Global evaluation of Case 1
Global evaluation of Car B = Global evaluation of Case 2
Global evaluation of Car C = Global Evaluation of Case 3

Global evaluation of Car D = Global Evaluation of Case 4

The above equalities result into the linear system of equations of the form AX = B. After solving this linear system using matrix method, the following 2-order fuzzy measure coefficients were obtained.

$$\mu_{Price,Comfort} = 0.4$$

$$\mu_{price,power} = 0.4$$

$$\mu_{Power,Comfort} = 0.59$$

Thus the result of alternative evaluation for the car selection problem is as shown below.

Table 3.13 Global evaluation by the linear system

Alternative	Global Evaluation
Car A	0.4
Car B	0.489
Car C	0.589
Car D	0.34

3.8.4 Analysis

From the results of this approach, it is clear that this approach is applicable for fuzzy measure determination only under certain conditions. The applicability of this approach depends upon its dimensions. If it is a non-square system, then there is difficulty in solving the linear system. This approach for fuzzy measure determination is applicable in case of a square system only. Hence an approach, which allows determination of fuzzy measures empirically and interactively is suggested.

3.9 Experimental Elicitation of Fuzzy Measures

The purpose of any DSS is to help the decision maker choose the best course of action (Sefion et al. 2003). Sometimes "best" can be interpreted numerically such as the highest return on investment, the lowest cost for a required capacity, and the highest score on a weighted combination of factors. The approaches developed so far address the issue of fuzzy measure determination. As mentioned earlier, there exists subjectivity in determining the fuzzy measure coefficients. The issue of subjectivity can be handled if the fuzzy measures are directly elicited from the decision maker. Such approach directly validates the system results.

3.9.1 Interactive model

The interactive model is another approach applicable in case of fuzzy measures. For building a decision support system for practical applications, interactive models of decision support systems can be effective (Sharda et al. 1988). Interactive models require decision maker's input for evaluating the alternatives.

3.9.2 Proposed Approach

This is a two-phase approach. The first phase is to determine fuzzy measures using a given case-base. The second phase is to elicit the fuzzy measures from the decision maker and thus validate the results obtained in the first phase. The attributes or criteria in the case base have absolute values. The model is presented in the following text.

3.9.2.1 Phase I

1. Let $C = \{C_1, C_2, ..., C_n\}$ be the set of criteria against which a set A of alternatives $\{A_1, A_2, ..., A_n\}$ is to be evaluated. The individual weights w_i on the criteria are elicited from the decision maker such that $\sum_{i=1}^{n} w_i = 1$.

For the above decision problem, a case-base is constructed in the following form.

i) The set of attributes are identical to the new decision problem or target problem.
ii) Absolute values of each attribute.
iii) A set of individual weights defined by the decision maker at the previous instance of making a decision.
iv) A set S of fuzzy measure coefficients defined at the previous instance of decision-making. It is assumed that this knowledge acquisition is done through an expert or decision maker who is able to tell the importance of sub-set of criteria and nature of interaction between the criteria in the past decisions.

v) Global evaluation of the object obtained by using the Choquet integral and the fuzzy measure coefficients obtained from the experts.

2. After constructing the above data set, the similarity of the cases w.r.t to the new decision problem is determined using Manhattan distance measure. The similarity between a pair of attributes is computed as

$$d_{ij} = \sum_{k=1}^{n} | x_{ik} - x_{jk} | \tag{3.9}$$

Where x_{ik} and x_{jk} are the values of the attribute for attribute k, for i, j = 1,2,……,n. For computing the similarity, the distance values should be normalized in the range [0,1] by the maximum distance $d_{\max}$. For every value of k, the global similarity will be calculated by simple weighted average of the local similarities of the given attributes. Thus distance between the cases is given as

$$d(Case_t, Case_u) = \frac{\sum_{k=1}^{n} w_k d_{ij}}{\sum_{k=1}^{n} w_k} \tag{3.10}$$

w_k is the weight of the k^{th} criteria or attribute in the new decision problem.
Based on this distance, the similarity is calculated as

$$Sim\,(Case_t, Case_u) = 1 - d\,(Case_t, Case_u) \tag{3.11}$$

Where $Case^t$ = Target case i.e. the new decision problem & $Case^u$ = Case from the case library i.e. past decision problems.

3. After such similarity computation the case with maximum similarity would be retrieved and the values of the fuzzy measures would be considered for solving the new decision problem. However, the decision maker is likely to have different importance values. Hence using a direct rating method (Bilgic 1997), an Importance scale is designed to elicit the new fuzzy measure coefficients from the decision maker.

3.9.2.2 Phase II

The importance scale is designed to ask the decision maker about the importance of sub-set of criteria. The decision maker can directly rate the sub-set of criteria in a range between 0 and 1. This scale is derived from AHP and application of AHP (Satty 1980; Thirumalaivasan 2001).

The values in this scale represent all the possible levels of decision maker's agreement. These values would represent the values of the importance of the sub-sets in a given decision problem. These values will model the fuzzy measures in the given decision problem. The values are defined in such a way that the decision maker is able to express the agreement easily. This scale confirms the agreement of the decision maker. The designed scale is shown below.

Table 3.14: Importance Scale

Sub-set is Less important	Extremely	0
	Highly	0.1
	Very	0.2
	Strongly	0.3
	Moderately	0.4
	Equally	0.5
Sub-set is More important	Moderately	0.6
	Strongly	0.7
	Very	0.8
	Highly	0.9
	Extremely	1

Based on this scale, the new fuzzy measures can be elicited from the decision maker. The fuzzy measures elicited through the agreement of the decision maker with the values elicited from this experiment and the values given by the similarity computation can be considered as the measure of performance of the proposed system (Wagholikar & Deer 2006a). The above approach can be illustrated using the following sample data set.

3.9.3 Illustration

The following is a classic car selection problem where the following data is assumed.

Table 3.15. Sample Data set for car selection problem

	Price		Power		Comfort	
Weight→	0.4		0.3		0.3	
Car A	$30,000	0.6	2000	0.8	Good	0.8
Car B	$25,000	0.7	1000	0.6	OK	0.7
Car C	$20,000	0.8	1500	0.7	OK	0.7
Car D	$35,000	0.5	2500	0.5	Very good	0.9

For the above data set, the following case-base is constructed. The criteria evaluations are shown in the column next to the absolute value of each criterion. The criteria evaluations are obtained from the expert. Different experts assigned the criteria scores in these past cases. In this case construction, the scores given by the experts on the same criteria with different absolute values were same. This was due to the subjectivity in assessing criteria scores. It is assumed that the decision maker is interested in a budget car. A higher criteria weight models this preference. For the above decision problem, a sample case base is constructed. The sample case base is shown in the following table. As discussed earlier, the case base is built using expert input. The expert input is represented in the form of criteria scores. The illustration involves 3 most important criteria for the decision maker. The small dimension of the problem space do not change the way in which the fuzzy measures are acquired using the proposed approach. It is acknowledged that the larger dimension of problem space will require more expert input and thus forms a trade-off between acquiring criteria scores from the experts and computing fuzzy measures in a simple way. This basic idea is the key driver behind the development of this approach. The attribute values are the absolute values and the criteria scores are represented in a binary scale.

Table 3.16. Sample Case base of past decisions

	Price		Power		Comfort		Global Evaluation
Case1	$40,000	0.8	3000	0.9	Very good	0.7	0.82
Case2	$35,000	0.7	2000	0.6	Good	0.6	0.66
Case3	$45,000	0.7	3000	0.6	Very good	0.7	0.665
Case4	$20,000	0.6	1500	0.5	Good	0.5	0.575
Case5	$24,000	0.7	1300	0.6	Good	0.7	0.68
Case6	$29,990	0.6	1400	0.7	Good	0.7	0.68
Case7	$32,250	0.5	1600	0.6	Very Good	0.8	0.643
Case8	$14,990	0.8	1200	0.7	Ok	0.6	0.72
Case9	$16,990	0.8	1200	0.7	Good	0.8	0.77
Case10	$19,990	0.8	1300	0.6	Good	0.8	0.74

Case base of past fuzzy measure values will be in the form of a vector of dimension j x k where j is the number of cases and k is 2^n-1 fuzzy measure coefficients. The case base for this data set would be represented as a tuple C= [$Case_i$, {k_1, k_2,…., k_{2n-1}}]. This sample data set is built using Microsoft Access DBMS and SPSS. This developmental tool is chosen due to its simplicity and smaller scale of the application. After performing the similarity computations using SPSS for Windows, the following results were obtained. The four, most similar cases were retrieved out of the given case base.

Table 3.17: Proximity Matrix - SPSS output for Case dissimilarity

	Absolute City Block Distance			
	Case 1	Case 2	Case 3	Case 4
Car A	0.721	1.00	0.488	0.744
Car B	0.442	0.721	0.209	0.977
Car C	.0.442	0.721	0.209	0.977
Car D	.0.808	1.0	0.615	0.404

This is a similarity matrix in the range 0 to 1, ranging from most dissimilar cases to most similar cases respectively. Thus, Car A is most similar to Case 2. Hence for computing the worth of Car A, the default values of the fuzzy measures applied in case 2 could be used for evaluating worth of the car A. Similarly,

Car B is most similar to Case 4

Car C is most similar to Case 4

Car D is most similar to Case 2

As mentioned earlier, the fuzzy measure coefficients for the above similar cases are stored in the case repository. Thus the same fuzzy measure coefficients are used for computing the global evaluations of cars in the new decision problem. Based on the above similarities and corresponding fuzzy measures of the similar cases, the global evaluations of all the cars is given by the proposed case-based decision support system.

Table 3.18. Car Evaluations

Alternatives	Global Evaluations
Car A	0.76
Car C	0.74
Car B	0.68
Car D	0.62

By visual observation of the data in Table 3.15 and the results in table 3.18, this preference order is intuitive. Now, to test the above result and agreement of the decision makers, the fuzzy measures are elicited from the decision makers. For this a theoretical experimentation was performed with 10 decision makers using the proposed importance scale. The fuzzy measures were elicited from the decision makers. These fuzzy measures were used for computing the global evaluations using Choquet integral. Accordingly, the theoretical results of this experimentation are as follows.

Table 3.19 Evaluations based on decision maker's input

Alternatives	Global Evaluation using Importance scale
Car A	0.78
Car C	0.77
Car D	0.77
Car B	0.68

From the above results, it can be observed that the decision maker had ranked only car C and D equally. Other evaluations are same as the system's evaluations using the default case-base. Hence the system results are in an approximate agreement with the decision maker. This illustrates the efficacy of the proposed decision support system.

3.9.4 Evaluation of the proposed system

There can be various ways of evaluating a decision support system.

i. If the system is built using data, information and knowledge from one set of situations, then it can be evaluated using an independent set of data (Priya 2001; Rice & Cochr 1984). When a data-driven model is a significant part of the decision support system, sometimes the data can be randomly separated into two parts, one for model development and one for validation (Haberlandt et al. 2002).

ii. When the decision support system is not data-based but rather knowledge-based, then it can be empirically evaluated using outputs from the system against a historic data set (Sprague & Carlson 1982). This does assume that the logic underlying the system is constant over time.

The proposed system follows a data-driven model. Hence it can be evaluated using the 2^{nd} strategy in the above options where data in the first phase is used for model development and the data in the second phase is used for validation of the system. The difference in values of the fuzzy measures obtained in both the phases verifies the results of the proposed system. The results can be verified with an allowable variation between the similarity based fuzzy measure values and the values elicited from the decision maker.

3.9.4.1 Experimental Evaluation of the proposed interactive model

The proposed model can be evaluated empirically. Empirical evaluation of a decision support system can be done through experiments (Cohen 1995). For the undertaken research there is no standard data against which the results can be compared (Blake & Mertz 1998). The experimental evaluation methods can be used when there exists a high level of subjectivity in the problem domain and hence the solution.

3.9.4.2 Manipulation Experiments

The manipulation experiments test hypotheses about casual influences of factors by manipulating them and noting effects of the same on the output. The results of the proposed system can be evaluated by observing the preference structure of the alternatives given by the system and preference computed based on fuzzy measures elicited from the decision maker. Theoretical experiments were performed with a set of 10 respondents. Objective of the experiment was to determine the deviation between the preference order given by the system and the user's preference order.

Table 3.20 Evaluation of the approach

Alternative	System Ranking	No of Respondents Agreed	No of respondents disagreed
Car A	1	10	0
Car B	3	8	2
Car C	2	7	3
Car D	4	6	4

The deviation was computed based on the individual ranking of the alternatives.

3.10 Summary

This chapter presented the proposed methods to solve the problem of fuzzy measure acquisition. This investigation concludes that interactive model of fuzzy measure determination is a suitable approach for pragmatic acquisition of the fuzzy measures. This approach can be implemented in a case-based decision support system. The next chapter discusses the implementation of the proposed approach in a case-based decision support system. A more detailed evaluation of the proposed methodology will be presented in the following chapters. A comparative analysis of the proposed approach will be used for evaluating the results of the proposed approach.

Chapter 4 Experimental Results

4.1 Introduction

This work has proposed a two-phase approach for acquisition of fuzzy measures. The approach was illustrated with an example. This chapter discusses the implementation of the proposed methodology in an interactive case-based decision support system. This discussion will illustrate applicability of the proposed methodology in solving practical problems. This system illustrates the proposed model with a sample data set. The results obtained through an experimental implementation of the model will be discussed in this chapter.

4.2 Proposed System Development Cycle

The proposed approach can be applied in a practical application in the following cycle. This is a general representation of the proposed system for a given decision problem. A case base for any given decision problem can be constructed. The problems are represented in a MCDM matrix. The case base will be constructed in the same form as the new decision problem.

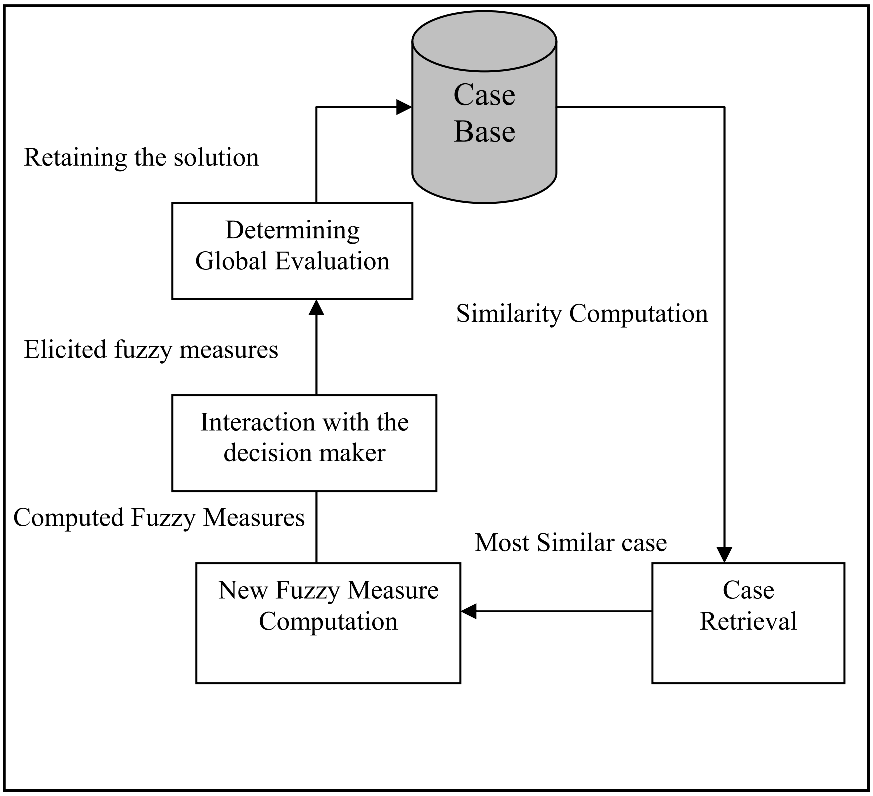


Figure 4.1 Proposed System Development Cycle

The above figure shows various phases of the proposed decision support system. A case base of past decision problems is constructed. The most similar case from the case base is retrieved using the city-block distance metric. This case is then used for the fuzzy measure determination. The system then interacts with the decision maker and elicits agreement of the decision maker on the fuzzy measure coefficients. The proposed importance scale is used for elicitation of the fuzzy measures from the decision maker. The global evaluation of the alternatives is then computed using the above fuzzy measures. The solutions are retained in the system for future problem solving. The system learns with every new decision problem. The case-base is continuously updated with the solutions of the decision problems. The initial case base is built using the expert's feedback about the past decision problems.

4.2.1 Case Construction

The case will consist of a problem description and a solution description. The problem description describes the state of the world when the case occurred while the solution description states the derived solution to that problem.

The case will consist description of the car selection problem in the form of attribute values. The solution description consists of global evaluation of each car and fuzzy measures values for each car. This can be represented as n-tuples of completely or incompletely described attribute values (Plaza et al. 1996). This set of attributes being divided in two non-empty dis-joint subsets: the sub-set of problem description attributes and the sub-set of solution description attributes denoted by P and S respectively (Banerjee et. al 1998; Dubois et al. 1999).

The case for a given car selection problem P is represented in the proposed system as $C_p = (p_1, p_2, ... p_k; s_1, s_2, ..., s_k)$, where $p = \{p_1, p_2, ..., p_k\}$ is a k-tuple standing for a car selection attributes. $s = \{s_1, s_2, ..., s_{n-k}\}$ is an n-k tuple, which represents the corresponding solution description. This consists fuzzy measure values acquired from the expert in the previous cases and the global evaluation of the cars.

Assume that a finite set C of known cases is given. Thus, the case base is denoted as C_p = (P, S). A new problem description denoted by P_o, for which the values of all attributes belonging to P are given. Then the proposed system estimates the values S_o of the attributes in S. This representation is the same as the relational data model (Rosenzweig & Silvestrova 2000; Ullman 1997). Hence this case base can be implemented using relational database technology. The following table shows the case representation for the proposed system using sample data. The case base is built using MS-Access RDBMS. It consists of attributes and their absolute values along with the attribute weights. The solution part consists of fuzzy measure values and global evaluation of the alternative. The global evaluation is stored as a part of historical data. It is not used for similarity computation. The absolute values of the criteria or attributes are used for similarity computation.

Table 4.1 Sample case representation

Car_Id	1
Features	
Price	$25,000
Power	1500
Comfort	Good
W_{price}	0.4
W_{power}	0.3
$W_{comfort}$	0.3
Solution	
Global Evaluation	0.7
$\mu_{price,power}$	0.4
$\mu_{price,comfort}$	0.8
$\mu_{power,comfort}$	0.6

4.2.2. Case Indexing

Case indexing (CI) is the method used to store the cases in the case-base. Like database systems, the proposed CBR system also uses indexes to speed up retrieval in the case base (Watson 1997). Information within a case in the case base is of two types:

- Indexed information that is used for retrieval

- Un-indexed information that may provide contextual information of value to the user but not to be used directly in retrieval. The global evaluation is considered as a part of the contextual information.

The process of case indexing is one of assigning labels to the case when entered in the case base to ensure its retrieval at the appropriate moment (Watson & Marir 1994). There are many different techniques available to index cases; such as choosing indexes using similarity and explanation-based methods and indexing vocabularies as well as using features and dimension of cases (Pal & Mitra 1992; Watson 1997). Case indexing can be done on a commonly used attribute. Price was chosen as the most commonly used attribute for this illustration. Using this attribute, the case base CB can be classified into 3 sub-classes i.e. CB $=\{CB_1, CB_2, CB_3\}$. The main reason for this classification is simply to reduce the search space. When a new case is added, it would belong to one of these sub-classes. This hierarchical indexing structure is shown below.

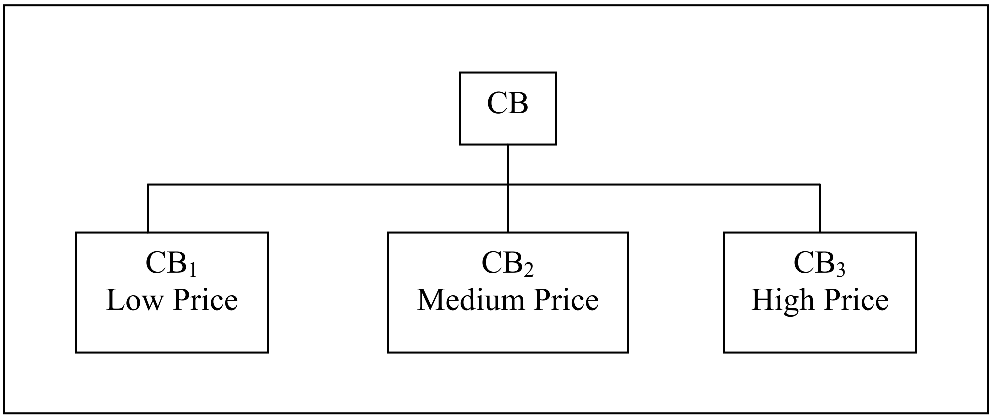


Figure 4.2 Case Indexing

In the proposed system, case indexing will certainly affect the system results with increasing number of cases as the case-base continues to grow. Case indexing is important for managing the time complexities for a larger data set. For a prototype of the proposed system the case indexing does not make any significant difference to the values of the fuzzy measures. It was considered only for faster case retrieval.

4.2.3 Similarity Computation

After constructing the case-base for the system, the most similar case was determined from the case base. This similarity computation also calculates the fuzzy measures for the new problem p_o.

In this module, case retrieval is of primary importance to the overall effectiveness of the proposed system. This is mainly due to -

- Retrieving the case ensures the best solution within the system's capability
- Retrieving the cases includes similarity computation and difference between the input problem and the retrieved cases. The subsequent modification uses this computation as a basis.

Case retrieval is the process of finding the closest case within the case-base. Case retrieval starts with a partial problem description and ends when a best matching previous case, in which the problem description is most similar to the current problem description has been found (Pal 2004).

When the problem description is given, the retrieval algorithm, using the indexes, retrieves the cases with the most similar problem(s) to the current problem description. The retrieval algorithm relies heavily on the indexes and the structure and organization of the case base to direct search to appropriate cases. In the undertaken research question, the new case is the alternative out of the given set of alternatives A= $\{a_1,a_2,\ldots..,a_n\}$. The retrieval algorithm uses the indexing structure defined in the earlier section where "price" was chosen as the index for the data set. The retrieval algorithm begins with locating the input case into the given sub-classes.

1. Starting from i =1, select one case e_j randomly from class C_i.
2. Compare the feature value of "price" for the new case e_{new} i.e. the alternative in set A and the case e_j from the case library.
3. If the two values are identical, $e_{new} \in C_i$, else i=i+1; go to step 1.
4. Repeat steps 1,2, and 3 until e_{new} belongs to some class C_i.

Figure 4.3 Algorithm for the case retrieval

Now, similar cases must be determined using a similarity measure. The similarity is computed using following steps.

1. For each case from the particular sub-class C_i, the decision alternative in the new problem is compared. If e_{new} belongs to c_i, for each case e_i calculate the degree of similarity between e_{new} and e_i. Manhattan distance metric is used for computing the similarity. In general terms, similarity is computed as below.

   $Sim(e_i,e_{new})$=common/ common + different (4.1)

   i. Rank the cases by similarity measure computed in step 1.

   ii. Choose cases such as $e_1,e_2,\ldots\ldots\ldots,e_k$ from C_i which are most similar to e_{new} i.e. $sim(e_{new},e_i)=\max\{Sim(e_{new},e_k), e_k \in C_i\}$ $i=1,2,\ldots\ldots..,k$.

2. The similar cases are ranked.
3. The case with maximum similarity is retrieved.

4.2.4 Fuzzy Measure Computation

After a matching case is retrieved, the system adapts the solution stored in the retrieved case to fulfil the needs of the current case. In the proposed system, the stored solution will consist of fuzzy measure values. The following algorithm used to carry out the task of case adaptation.

Step 1. Group the retrieved cases (which are most similar to the present case in the case library) based on their solution value into several categories. The solution value is the fuzzy measure value stored in the system.

Step 2: Count the number of cases retrieved in each category C_i $(i=1,2,\ldots\ldots,m)$ and record them as $n_1,n_2,\ldots..,n_m$.

Step 3: Choose category C_i for which $n_i=\max\{n_1,n_2,\ldots.,n_m\}$, take the average value b

Step 4: If there is only one category then the solution to the present case, the solution from the stored case can be applied directly. If there are several cases in the given category, then take the average value b defined by

$$\underline{b} = \frac{a_{i1}+ a_{i2}+ \ldots + a_{ik}}{k} \quad (4.2)$$

as the solution to the present case, where $a_{i1}+ a_{i2}+ \ldots + a_{ik}$ are the values of the global evaluations. Once an adapted solution is obtained, its validity should be tested in reality. Hence the decision maker will be asked if the solution is acceptable. If it is acceptable, the performance of this CBR system can then be improved by adding this new case record.

4.2.5 Agreement of the decision maker

The decision maker is asked about the validity of the solution. This process is carried out using the proposed importance scale discussed earlier. The agreement of the decision maker with the system computed fuzzy measure values is used for evaluation of the system.

4.3 Development of the System Prototype

Based on the above modules the proposed system is developed. This system is developed using relational database technology and SPSS. SPSS is used for similarity computation between the case-base and the new decision problem. The important aspects of the proposed system are discussed next.

4.3.1 Data Set

The data set comprises of the case-base of the past cases and a new decision problem. The data for the new decision problem is same as the data in table 3.15. The data about the decision problem mentioned in the earlier illustration is used for the development of the proposed system. The case base is built at a larger scale than mentioned in the earlier illustration. The case base comprises of 100 cases, which involves evaluations of 100 cars. The problem in validating this application is the lack of a standard data set (Wagholikar & Deer 2005b). Hence a customized data set is built. It is assumed that an expert in the past does the object evaluation. The global evaluation of the objects is stored in the case base. The global evaluation and the fuzzy measures applied in the past decision problems forms the solution part in this model.

4.3.2 Sample Data Set

Based on the above case representation the data set is as below. This data set is built using SPSS. The following data set shows 10 cases. A case-base with 100 such cases was built for the proposed case-based decision support system. The attribute values are represented in absolute scale. The values are numeric as well as linguistic. Such values are chosen so that the decision maker can easily assign the scores to the criteria for a given alternatives.

Table 4.2 Sample case for the proposed system

Case ID	Price	Power	Comfort	μ_1	μ_2	μ_3	Global Evaluation
1	$30,000	2000	OK	0.7	0.8	0.4	0.6
2	$25,000	1000	Good	0.6	0.7	0.3	0.55
3	$20,000	1500	OK	0.8	0.6	0.5	0.5
4	$35,000	2000	Very good	0.85	0.7	0.6	0.8
5	$25,000	1200	OK	0.7	0.6	0.7	0.8
6	$24,000	1300	Good	0.8	0.7	0.6	0.7
7	$30,000	1500	Ok	0.7	0.6	0.5	0.8
8	$35,000	1400	Good	0.75	0.7	0.7	0.8
9	$45,000	2500	Very Good	0.9	0.8	0.85	0.9
10	$50,000	2500	Very Good	0.85	0.9	0.8	0.9

The above case base is a sample of the actual case base constructed. The following factors were considered during the construction of the case base.

i. It was assumed that the global evaluation and the fuzzy measures are elicited from the experts in the previous decision problems.

ii. Attributes price and power are in numeric scale. The attribute comfort is on ordinal linguistic scale.

iii. The global evaluation and the combined importance of criteria are defined on a binary scale.

iv. The fuzzy measure coefficients μ_1, μ_2 and μ_3 represent the combined importance of the sub-sets of criteria {price, power}, {power, comfort} and {price, comfort} respectively.

The following table shows the descriptive statistics of the numerical attributes. These statistics are shown in order to understand the boundaries of the entire data set.

Table 4.3 Descriptive statistics of the data set

	Minimum	Maximum	Mean	Std. Deviation
Price	$13,200	$80,000	$35,339.99	$14,866.987
Power	900.00	3200.00	1633.3457	544.68835

The above data set is built into the system. The following figure shows the data model of the proposed system using relational database technology.

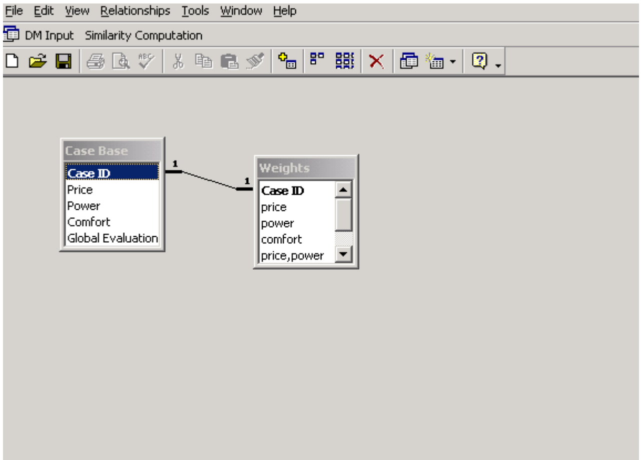


Figure 4.4 Data Model of the Proposed System

The case base is built in the form of tables. The next figure shows the snapshot of the case-construction module. This module allows the experts to construct the case base.

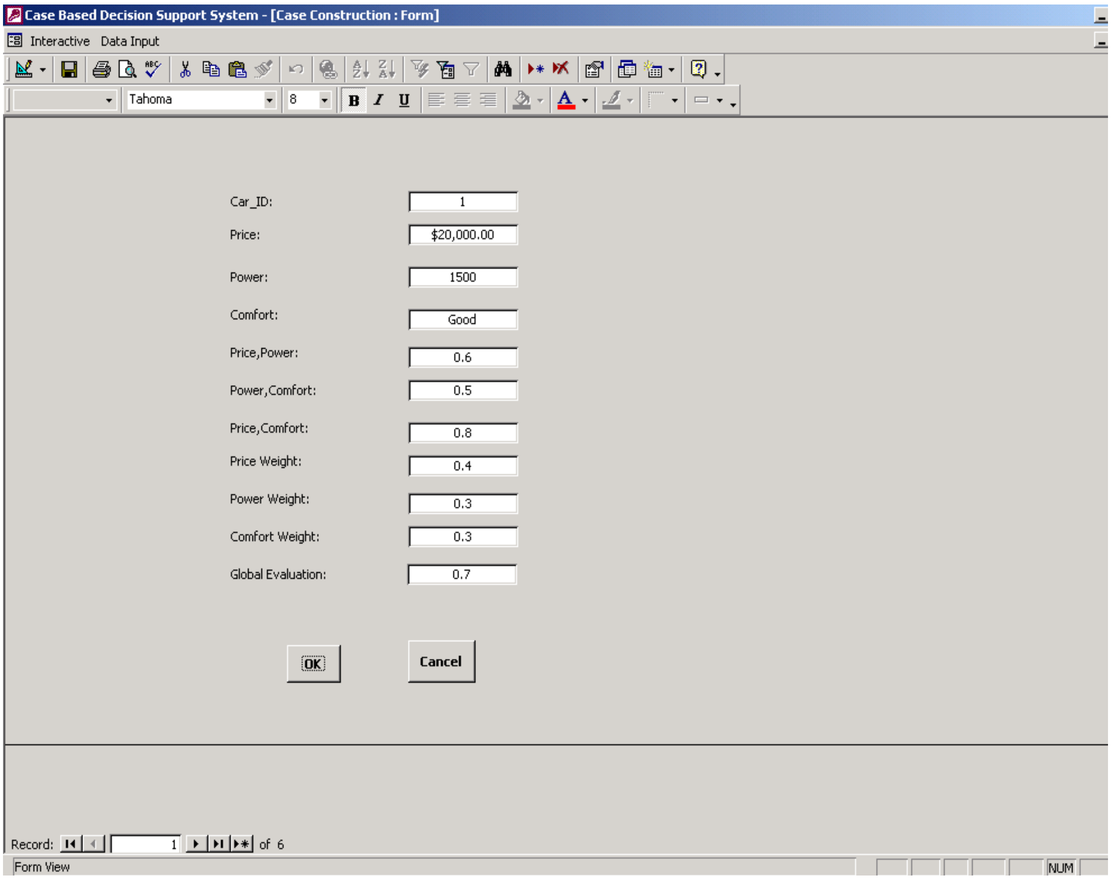


Figure 4.5 Case Construction Module

The expert defines the absolute values of the decision criteria, the individual criteria weights and weights of sub-sets of criteria. This case base is constructed in the form of tables in a relational database. The relational database allows easy construction of the case base. The constructed case base can easily be updated. The maintenance task is simpler with the proposed system. The proposed system is developed for any decision problem that can be expressed in MCDM form.

4.3.3 Similarity Computation

The decision maker enters the data for the new decision problem. Based on this data and the case base built earlier, the similarity computation is performed by the system using the Manhattan distance measure. After running the similarity algorithm, the results given by the system are as below.

Table 4.4 Computed Similarity Scores

Car	Similar Case ID	Similarity Score
Car A	Case 16	1
Car B	Case 89	0.998
Car C	Case 44	0.995
Car D	Case 16	1

4.3.4 Retrieved Fuzzy Measure Values

Using this similarity score, the system retrieves the fuzzy measures from the stored database. Thus the retrieved fuzzy measures are shown below.

Table 4.5 Retrieved Fuzzy measures

	Price, Power	Price, Comfort	Power, Comfort
Car A	0.6	0.57	0.7
Car B	0.4	0.6	0.75
Car C	0.8	0.7	0.6
Car D	0.6	0.5	0.45

4.4 Interactive Module

The interactive module of the system is shown below. The interactive module elicits the combined importance of the sub-set of criteria by displaying the sub-set to the decision maker. The decision maker enters the agreement on the importance of the sub-set of criteria. This agreement is modelled using the proposed importance scale.

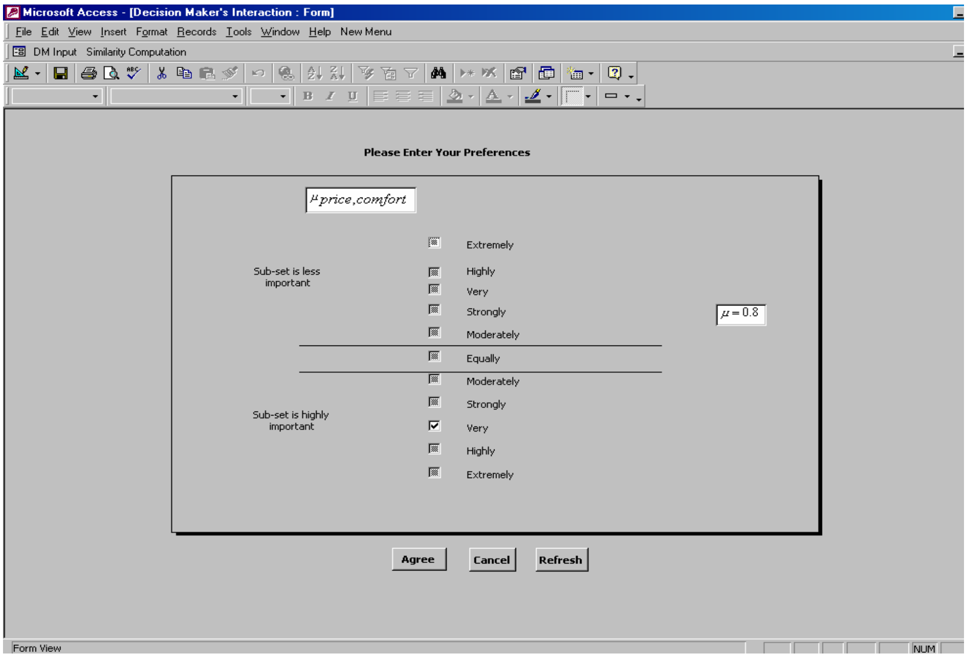


Figure 4.6 Interactive Module of the system

The results from the system are shown below. The system shows the global evaluation of the alternatives.

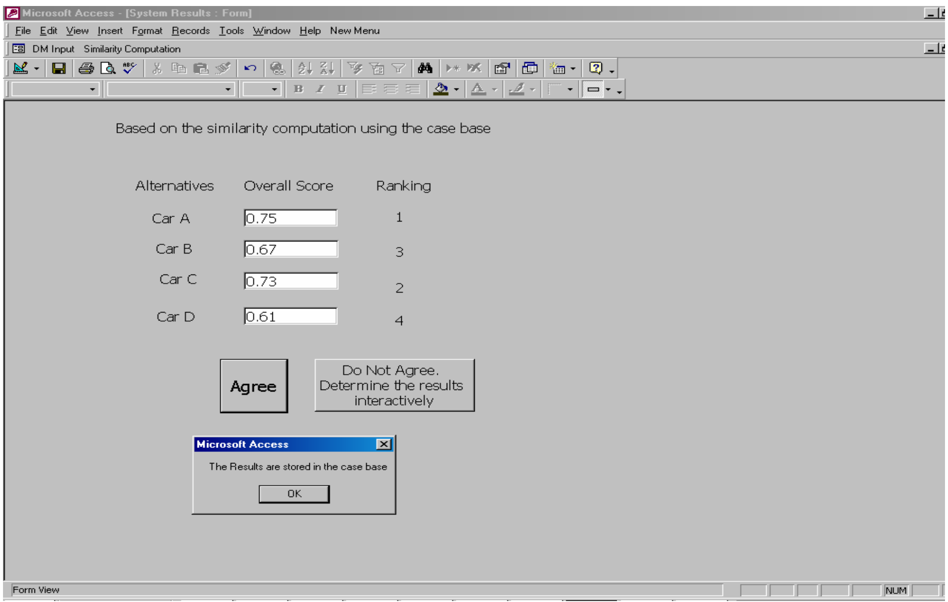


Figure 4.7 Application Results

The ranking of the alternatives is determined based on the global evaluation. If the decision maker agrees with the proposed ranking, the results are stored in the case-base for future problem solving. If the decision maker does not agree, then the system determines the global evaluations interactively using the proposed importance scale. The above results will be stored in the case base for future problem solving. Based on the fuzzy measure values derived in the earlier phases, the system computes the global evaluation using Choquet integral. The global evaluations are as below.

Table 4.6: Global evaluation of the objects obtained by the system

Car	Global Evaluation
Car A	0.75
Car B	0.67
Car C	0.73
Car D	0.61

The above global evaluation of the cars provides the intelligent decision support to the decision maker as the fuzzy measures model the interaction between the criteria. These results also incorporate agreement of the decision maker through the interactive module of the proposed system. This module validates and confirms the results obtained by the proposed algorithm. The proposed algorithm automates the task of acquisition of fuzzy measures using similarity-based reasoning in Phase I as discussed earlier in section 3.9.2.

4.5 Summary

The illustration of the proposed methodology showed the applicability of the proposed approach for solving the issue of fuzzy measure determination. The experimental results show that the proposed approach gives intuitive results. The proposed system allows the decision maker to validate the system results. Due to this, acceptance of the system results for making decisions in the given problem can be high. Thus, it can be said that the proposed approach is applicable especially for repetitive decisions and provides valid decision support to the decision maker in a given decision problem. The next chapter presents the comparison of the proposed approaches with other approaches and this comparison also becomes a basis for the evaluation of the claims made within this work.

Chapter 5 Analysis of the Results

5.1 Introduction

In the earlier discussion, the proposed methodology was illustrated with a sample data set. This chapter presents the analysis of the results. The proposed approach was compared with state-of-the-art approaches. The results were analysed using the resultant fuzzy measure values and the resultant preference order of the alternatives. The role of knowledge acquisition in the undertaken research will also be discussed. The comparison of the results with other state-of-the-approaches will evaluate the proposed methodology and thus will lead to conclusion of the undertaken research.

5.2 Comparative Results

The proposed methodology offers a new perspective towards solving the fuzzy measure acquisition problem. A common data set is used for comparative analysis of these approaches. The overall comparison and the individual comparison of the proposed methodology with the other approaches are presented. The proposed approach is compared using two different datasets involving interactive and non-interactive criteria respectively. The state-of-the-art approaches include Sugeno's λ Measure, quadratic programming based interactive optimisation approach (Grabisch 1995a) and a methodology based on semantic consideration about criteria (Marichal 1999). These approaches were discussed earlier in the literature review.

5.2.1 Results for Interactive Criteria

As mentioned earlier, there is no standard data set for comparison of all the approaches. The selection of data set for the comparative analysis is challenging. Since different state-of-the-art approaches differ in their problem space. It is difficult to have a common data set which becomes a test-bed for all MCDM approaches involving application of fuzzy measures. A data set that has the minimum number of attributes required for the comparative analysis is chosen. Therefore a dataset in table 3.15 is used for the comparative analysis. This research uses this data set for illustration, implementation as well as for the comparison of the approaches. It provides a single standard for illustration and comparison of the results.

The interaction index between the criteria models the nature of relationship between the criteria. The formulation was explained earlier in chapter 2, section 2.7.2. The positive value of interaction index conveys positive interaction between the criteria. The negative value of the interaction index conveys negative interaction between the criteria. The formulation for the interaction index for the given data set results into the following values.

Table 5.1 Interaction Indices between the criteria

Interaction Index	Value
$I_{price, power}$	0.5
$I_{price,comfort}$	0.3
$I_{power,comfort}$	-0.5
$I_{price,\{comfort, power\}}$	0.3
$I_{power,\{price,comfort\}}$	-0.3
$I_{comfort,\{price, power\}}$	0.1

The above table shows the interaction between individual as well all the possible sub-sets of criteria in the given problem. As mentioned earlier, the positive value of the interaction indices indicate that the criteria support each other. This implies that the values of the fuzzy measure coefficients will be greater than the summation of the individual importance of the criteria. The negative value of the interaction index indicates negative interaction between the criteria. This can be illustrated by the negative interaction index between power and comfort. It can be said that as the power of a car increases, the comfort decreases. This implies that the values of the fuzzy measure coefficients will be less than the summation of individual importance of criteria in a sub-set. This property will be used to analyse the values of the resultant fuzzy measure coefficients. The proposed methodology will be illustrated with data sets for interactive as well as non-interactive criteria. A data set with interactive criteria was constructed with both types of interactions between the criteria. Such a data set illustrated the usefulness of the proposed approach for solving decision problem under the setting of interactive criteria. The same data set was used to compare the other approaches. This comparison of the results can be analysed using the resultant fuzzy measure values and the resultant preference order. In case of a data set involving interactive criteria, the resultant fuzzy measure values should model the positive as well as negative interaction between the criteria. The sub-sets of criteria in the data set are positively as well as negatively related.

5.2.2 Comparative Results of all the approaches

The proposed approach was compared with other approaches based on resultant fuzzy measure values and global evaluation of given alternatives. The ranking of the cars must be intuitive. It should be in agreement with the individual criteria scores and importance of the criteria. Intuitive results mean that the results are in accordance with the human reasoning. The results will resemble the human expectations about the overall ranking of given alternatives based on their criteria evaluations. For comparative analysis, the following factors were taken into consideration.

i. An alternative having higher criteria scores with higher criteria weights must be ranked higher assuming that all the alternatives have equal criteria weights.

ii. The fuzzy measure values must be in accordance with the nature of interaction (either positive or negative) between the criteria.

It is stated earlier that the main research objective is to investigate and develop the pragmatic ways of determining the fuzzy measure coefficients. Computational complexity was not considered as a testing parameter in investigating the issue of dependence between the criteria. The time complexity in determining the fuzzy measures was not taken into consideration as well. Various approaches give different results on the same data set. The results about the fuzzy measure coefficients are presented below.

Table 5.2 Comparison of fuzzy measure coefficient values

	Fuzzy measure coefficients obtained by			
Fuzzy Measure coefficients	λ Measure	Interactive Optimisation	Semantic Consideration	Proposed methodology
Price, power	0.7	0.6	0.5	0.8
Price, comfort	0.6	0.7	0.6	0.7
Power, comfort	0.6	0.6	0.7	0.3

It can be observed from the interaction indices in table 5.1 that the criteria {price, power} and {price, comfort} are positively related. The criteria {power, comfort} are negatively related. The comparison of the fuzzy measure coefficients shows that the fuzzy measure values given by the proposed methodology certainly model the interaction between the criteria. For example, the criteria price and power are positively related with each other. Hence their combined importance or the fuzzy measure coefficient will be greater than the summation of their individual importance. The proposed methodology correctly represents this property as compared with other approaches.

The fuzzy measure values given by the proposed methodology follow the sub-additivity and super-additivity property. The values given by the other approaches do not model the interaction between the criteria adequately. This is evident from the resultant fuzzy measure values. For example, the values given by the λ measure shows that price and comfort are additive. It fails to model the super-additive relationship between price and comfort. Similarly, other approaches do not model the relationship between the criteria.

The global evaluation of the cars was computed using the resultant fuzzy measure values. The overall comparison of the results for global evaluation of the cars is shown below.

Table 5.3 Comparison of global evaluation

	Global Evaluation obtained by			
Alternatives	λ Measure	Interactive Optimization	Semantic Consideration	Proposed methodology
Car A	0.45	0.65	0.75	0.75
Car B	0.64	0.6	0.7	0.67
Car C	0.76	0.55	0.8	0.73
Car D	0.9	0.7	0.6	0.61

The λ measure approach ranks Car D higher than the other cars. The interactive optimization method ranks Car D higher than the other cars. The approach based on semantic consideration about criteria ranks Car C higher than the other cars. The proposed methodology ranks Car A higher than the other cars. Car A has higher criteria satisfaction. Therefore, intuitively Car A should be ranked higher than the other cars. The proposed methodology gives intuitive results.

The resultant preference order on the given set of alternatives is also studied in detail. The preference order must be in accordance with the absolute values and the decision maker's preferences. The weights model the preferences on the individual criteria. The preference order proposed by different approaches needs to be considered for analysis. The analysis of the resultant preference order for each methodology is presented next.

5.3 Comparison of the Results obtained using λ Measure

Using the formulation for the λ measure computation, the value of λ is computed using the following formulation. The formulation was explained earlier in chapter 2, section 2.8.1.

$$1 + \lambda = (1 + 0.4\lambda)(1 + 0.3\lambda)(1 + 0.3\lambda) \tag{5.1}$$

Solving the above equation, the value of λ is -9.167. This value of λ is invalid as it is less than -1. This implies that that the criteria are not interactive. The value of λ does not model either positive or negative interaction between the criteria. The interaction indices show that the criteria are interactive. However, this approach does not consider the interaction between the criteria.

This measure gives counter-intuitive values of the fuzzy measure coefficients. The fuzzy measure values for the sub-sets {price, power}, {price, comfort} and {power, comfort} shows that the criteria are additive. It does not model the super-additivity and sub-additivity between the criteria. This is in total contrast with the actual nature of criteria. The criteria are actually non-additive in nature. This was evident from the interaction indices determined earlier. Therefore fuzzy measure coefficients must model the non-additive nature of the criteria.

The approach of applying λ measure gives counter-intuitive results. It can model either positive or negative type of interaction. However, it does not model both types of interaction between the criteria. This measure can be useful if there exists only one type of interaction (positive or negative) among the criteria in a given decision problem. The application of λ measure requires the construction of decision problem involving only one type of interaction between the criteria. Such requirement restricts the applicability of this measure. In most of the decision problems the criteria can be positively as well as negatively related. The resultant preference order is also analysed. The λ measure proposes the following preference order.

$$\text{Car D} \succ \text{Car C} \succ \text{Car B} \succ \text{Car A} \quad (5.2)$$

This preference order can be argued using the criteria scores for each car. Car D does not have higher criteria scores on all criteria. For example it has the lowest score for the criteria price and power among all other alternatives. Car D can not have the highest preference. Car C has the highest score for the criteria price among all other alternatives. However, the scores on the other criteria are relatively lower than on the criteria Price. Therefore, Car C can not have the highest preference order. It can have second preference. The preference for car B is appropriate. The preference for car A is not appropriate. The resultant global evaluation for Car A is lowest among all alternatives. This contrasts the high criteria scores assigned to the Car A. Thus, it can be concluded that the preference order proposed by this approach is partially counter-intuitive.

The comparison of the global evaluation of the cars shows that the ranking given by the proposed methodology is more intuitive than given by λ measures. The criteria scores for Car A represent higher criteria satisfaction than values for Car C. Hence intuitively Car A should be ranked higher than Car C and other cars.

5.4 Sensitivity Analysis for λ Measure

In addition to the above comparison, a sensitivity analysis was conducted for observing the relationship between λ measure and the global evaluation of alternatives. The variation in the output was analysed with the help of a web-based software (Takahagi 2000). The sensitivity analysis for λ measure is shown in the figure below.

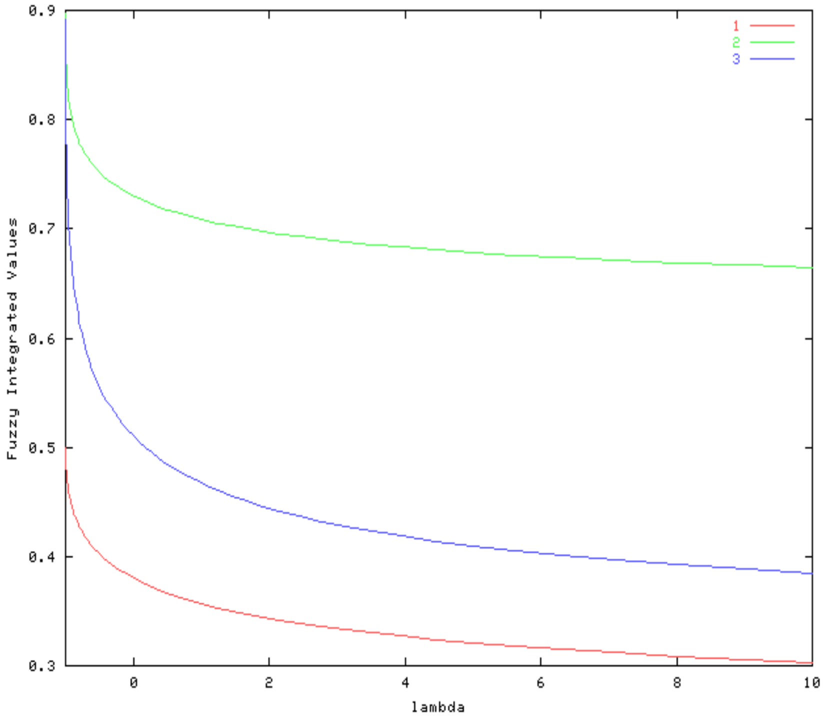


Figure 5.1 Sensitivity analysis for λ measure

The above sensitivity analysis shows that the global evaluation of objects is inversely proportional with λ measure. The sensitivity analysis also shows that for lower values of λ, the resulting fuzzy measures model only positive interaction between criteria. This sensitivity analysis supports the earlier conclusion about restricted ability of the λ measures.

5.5 Comparison of the results with the approach of Interactive optimisation

As discussed earlier, this approach is based on minimizing the distance between the global evaluation from the past data and the Choquet evaluation using the current data. The data in the table 3.16 and the case base in the table 4.2 from chapter 4 are used to compute the fuzzy measures. Using this data, the formulation is as follows.

$$Z_j - \delta_i \le C_j^{\,t} Xj \le Z_j + \delta_i \qquad (5.3)$$

The threshold $\delta = 0.5$. The above formulation shows that the Choquet evaluation for the new decision problem lies in between the lower and upper limits of the global evaluation from the past decision problem. Using the test data from table 3.16 and table 4.2, the decision maker's preferences are formulated as follows.

Power = η Price

Price = η Comfort

The decision maker's preferences are formulated using the criteria weights. It was shown that the above problem can be solved by the quadratic programming (Grabisch 1995a). Quadratic programming is a special type of optimization problem. For the undertaken problem of fuzzy measure determination, the quadratic problem can be formulated as:

$$f(\mathbf{x}) = \frac{1}{2}\mathbf{x}'Q\mathbf{x} + \mathbf{c}'\mathbf{x} \qquad (5.4)$$

Subject to one or more constraints of the form:

$A\mathbf{x} \leq b$ (inequality constraint)

$E\mathbf{x} = d$ (equality constraint)

Where, x is a 2^{n-1} dimensional vector containing all the fuzzy measure coefficients, Q is a 2^{n-1} dimensional square matrix and c is 2^{n-1} dimensional vector. Using a quadratic programming solver, the results for fuzzy measures and global evaluation in table 5.2 and 5.3 were obtained respectively. The results given by this approach were not in accordance with the criteria scores.

According to the results in table 5.2, the resultant fuzzy measure values contrast the nature of interaction between the criteria. In other words, the resultant fuzzy measure values model positive interaction between the criteria where the criteria are related negatively. For example, the sub-set {price, power} has positive interaction between the individual criteria. However, the value of this fuzzy measure coefficient, $\mu_{price,power}$ is less than the summation of the individual weights 0.4 and 0.3 of price and comfort respectively. The other fuzzy measure coefficients show similar result. Such results show a counter-intuitive preference order of the given alternatives. This approach proposes the following preference order

$$\text{Car D} \succ \text{Car A} \succ \text{Car B} \succ \text{Car C} \qquad (5.5)$$

As explained earlier, highest ranking of Car D is not acceptable. Car D should be the least preferred alternative due to its lower criteria scores. The proposed ranking for Car A is also not appropriate. It should be ranked higher than car D. Car B is ranked appropriately. The preference for Car C should be definitely higher than Car B and Car D as evidenced by its criteria scores. In addition to the counter-intuitive results, the quadratic programming has the following limitations.

i. The quadratic programming approach is expensive to implement in many cases.
ii. There is no commonly recognized implementation standard for Quadratic programming approach.
iii. It is uncertain how close a solution is to the true optimal

5.6 Comparison with method based on semantic considerations about criteria

This approach models the decision maker's preferences and determines the fuzzy measure by solving a linear programming problem. Using the same data set in table 4.1 and suggested formulations (Marichal 1999), the fuzzy measures are determined as follows.

i. Using the data in the table 4.1 the decision maker's preferences on the alternatives are formulated as follows

$$\text{Car B} \succ \text{Car A} \Leftrightarrow 0.2\ a_1 + 0.2\ a_2 + 0.9\ a_{13} + 0.9\ a_{23} > 0$$

$$\text{Car B} \succ \text{Car C} \Leftrightarrow 0.2\ a_1 + 0.4\ a_2 + 0.2\ a_{13} + 0.4\ a_{23} > 0$$

$$\text{Car D} \succ \text{Car C} \Leftrightarrow 0.4\ a_1 + 0.4\ a_2 - 0.7\ a_3 - 0.3\ a_{13} - 0.3\ a_{23} > 0$$

Where a_i and a_{ij} ($i, j \in N$) are the single order and 2-order fuzzy measure coefficients.

ii. The above formulations can be solved by the following linear program

Maximize $z = \varepsilon$

Subject to

$$0.2\ a_1 + 0.2\ a_2 + 0.9\ a_{13} + 0.9\ a_{23} > \delta + \varepsilon$$

$$0.2\ a_1 + 0.4\ a_2 + 0.2\ a_{13} + 0.4\ a_{23} > \delta + \varepsilon$$

$$0.4\ a_1 + 0.4\ a_2 - 0.7\ a_3 - 0.3\ a_{13} - 0.3\ a_{23} > \delta + \varepsilon$$

$$a_1 + a_2 + a_3 + a_{12} + a_{23} + a_{13} = 1$$

$$a_i \geq 0$$

$$a_i + a_{ij} \geq 0$$

Using a fixed value of the threshold δ=0.05, the above linear program resulted into the fuzzy measure values in table 5.2.

The central idea in the approach of determination of semantic consideration of criteria is to solve the problem using linear programming. This approach requires linearity in the expression for fuzzy measure determinations. The linearity assumptions lead to proportionality and additivity for determining the solution. Due to the property of proportionality a change in the fuzzy measure coefficient results in a proportionate change in that fuzzy measure coefficient's contribution to the overall global evaluation. In addition to these linearity assumptions, this approach assumes *certainty*; that is, that the coefficients are known and constant. The linearity cannot always be guaranteed. The fuzzy measure coefficients can be related non-linearly. It was shown earlier that the fuzzy measure coefficients could be determined by quadratic programming. The linearity requirement restricts the applicability of this approach. Problem formulation is another challenging task in implementing this approach. The challenge is in translating the decision maker's preference into a linear program where linearity cannot be granted. This approach results in fuzzy measures and the global evaluation of alternatives as shown in table 5.2 and 5.3 respectively. The resultant preference order is counter-intuitive as car A is not ranked higher than the other cars.

The resultant fuzzy measures in this approach do not model the interaction between the criteria correctly. This can be observed by the fuzzy measure values in table 5.2. The fuzzy measure coefficient for the sub-set {price,power} shows that the criteria are negatively related. This is in contrast with the interaction index of the criteria that shows that the criteria are positively related. The fuzzy measure coefficients for the sub-set {price, comfort} show that the criteria are sub-additive. The criteria price and comfort are actually super-additive. Similarly, the fuzzy measure coefficient for power and comfort models positive interaction instead of actual negative interaction present between the criteria.

According to this method, alternatives are ranked in the following order

$$\text{Car A} \succ \text{Car B} \succ \text{Car D} \succ \text{Car C} \tag{5.6}$$

In the above preference order, Car B is preferred over Car C and Car D is preferred over Car C. This is counter-intuitive. The criteria scores for Car B are lower than Car C for all criteria. Similarly, the criteria scores for Car D are lower than the criteria scores for Car C. Thus the proposed preference order is largely incorrect.

5.7 Comparison of the results given by proposed methodology

The proposed methodology gives the following preference order

$$\text{Car A} \succ \text{Car C} \succ \text{Car B} \succ \text{Car D} \tag{5.7}$$

The above order of alternatives is more intuitive as compared with the other approaches. The proposed ranking is in accordance with the criteria scores. Car A has the highest criteria scores for all criteria. Car D has lower scores for all criteria for the weights defined by the decision maker. The criteria scores for Car B are lower than that of the Car C for price and power. The criteria comfort has equal scores and weights for Car B and Car C respectively. After considering these criteria evaluations for Car B and Car C, the suggested preference order is acceptable. Thus it can be concluded that the resultant preference order suggested by the proposed methodology is intuitive. This illustrates the significance of the proposed methodology. It confirms that the applicability of the proposed methodology in the presence of interactive criteria.

5.8 Results for non-interactive criteria

The comparison of various approaches for the interactive criteria was explained earlier. A decision problem involving independent criteria is considered for the comparative analysis. The following data set was used for the illustration.

Table 5.4 Data set with non-interactive criteria

	Car Dealer	Fuel Economy	Comfort
Weights	0.4	0.3	0.3
Car A	0.2	0.3	0.4
Car B	0.8	0.7	0.9
Car C	0.3	0.5	0.7
Car D	0.8	0.4	0.2

It can be observed intuitively that the criteria in the above dataset are independent. It is highly unlikely to have an interaction between the criteria car dealer and fuel economy. The decision maker may still want to make a decision based on the above set of criteria. In this problem since the criteria are non-interactive, the interaction index is zero and therefore there is no interaction between all the sub-sets of the criteria.

5.9 Comparative Results of all the approaches for non-interactive criteria

In the case of non-interactive criteria, the interaction between the indices is not present. In real life decision problems, there can be some problems where the criteria can be independent of each other. The proposed methodology can be applied for non-interactive set of criteria as well. The main distinguishing characteristic of the proposed approach is that it is simple to apply in case of non-interactive criteria as well. Based on the data set in table 5.4, the obtained results are presented below.

Table 5.5 Comparison of fuzzy measure coefficient values

	Fuzzy Measure Coefficients Obtained by			
Fuzzy Measure coefficients	λ Measure	Interactive Optimization	Semantic Consideration	Proposed methodology
{FuelEconomy, Car Dealer}	0.62	0.6	0.6	0.7
{Fuel-Economy, comfort}	0.51	0.5	0.7	0.6
{Car-Dealer, Comfort}	0.62	0.5	0.5	0.6

It is stated earlier that when the criteria are non-interactive, the importance of sub-set of criteria should be the summation of individual importance of criteria in the given sub-set. This is the property of additive measures. The results given by the various approaches can be used to verify this property.

The global evaluation of the cars is computed by applying all approaches. The comparison of the results is shown below.

Table 5.6 Comparison of global evaluation

	Global Evaluation obtained by				
Alternatives	λ Measure	Interactive Optimization	Semantic Consideration	Proposed methodology	Weighted Average
Car A	0.27	0.3	0.4	0.35	0.29
Car B	0.78	0.5	0.7	0.75	0.8
Car C	0.44	0.6	0.75	0.5	0.48
Car D	0.45	0.55	0.6	0.53	0.5

The resultant fuzzy measure values are shown in table 5.5. The resultant fuzzy measure values do not follow the additivity property. This is important because of the non-interactive nature of the criteria. If there is no interaction between the criteria, then the fuzzy measures will be additive measures. It can be observed from the values of the fuzzy measure coefficients that this approach shows that the criteria are interactive. For example, the criteria fuel economy and car dealer are intuitively independent of each other. Therefore the value of the fuzzy measure coefficient will be the summation of the individual criteria weights. However, this approach shows that the criteria are negatively related. Similarly, other sub-sets model negative interaction between the criteria. Thus, this approach shows only one type of interaction between the criteria. This measure fails to model the independence between the criteria.

This approach of interactive optimisation also models negative interaction between the criteria. It shows that the criteria are sub-additive. The values of the fuzzy measure coefficients are less than the summation of the individual criteria weights.

In case of the approach based on the semantic consideration about criteria, the resultant fuzzy measure values show that the criteria in the sub-set {fuel economy, car dealer} and {Car dealer, comfort} are sub-additive. The criteria in the sub-set {Fuel-economy, comfort} are super-additive. This methodology shows that the criteria are positively as well as negatively related. This contradicts the fact that the criteria are independent.

It can be observed from the values of the fuzzy measure coefficients that the proposed methodology models the independence between the criteria. The resultant fuzzy measure coefficient values for all the sub-sets are summation of the individual criteria. Thus overall, it can be said that the proposed methodology is applicable in case of the non-interactive criteria.

The λ measure proposes the following preference order.

$$\text{Car B} \succ \text{Car D} \succ \text{Car C} \succ \text{Car A} \tag{5.8}$$

This preference order is acceptable as the criteria scores for Car A are lowest among all alternatives. The proposed preference order for interactive optimisation is

$$\text{Car C} \succ \text{Car D} \succ \text{Car B} \succ \text{Car A} \tag{5.9}$$

This approach also ranks Car A as the least preferred alternative. It is acceptable. However, the resultant preference for Car B is not acceptable. The approach based on semantic consideration about criteria proposes the following preference order

$$\text{Car C} \succ \text{Car B} \succ \text{Car A} \succ \text{Car D} \tag{5.10}$$

This order is clearly counter-intuitive as it ranks Car D as the least preferred alternative. The proposed methodology gives the following intuitive preference order

$$\text{Car B} \succ \text{Car D} \succ \text{Car C} \succ \text{Car A} \tag{5.11}$$

Since the criteria are non-interactive, weighted average is also applied for this data set. Car B has higher criteria satisfactions than the other cars. Therefore intuitively it should be ranked higher than the other cars. It can be observed that all the approaches with the exception of the approaches of determining fuzzy measure by interactive optimization and semantic consideration about criteria rank Car B higher than the other cars. The results given by the λ measure are acceptable for a data set with non-interactive criteria. It can be concluded that the proposed methodology gives intuitive results. Both the proposed methodology and the weighted average produce the same results in terms of car ranking. The results given by the proposed methodology are intuitive in the case of independent as well as dependent criteria. This illustrates the significance of the proposed methodology in the case of interactive as well as non-interactive criteria.

5.10 Role of knowledge acquisition

Knowledge acquisition includes elicitation, analysis and modelling of knowledge for solving real-life problems. The approaches discussed earlier can be thought of as having a knowledge acquisition component. This aspect is important for developing practical solutions for various decision problems. This is an important dimension of this research. It was stated earlier that the fuzzy measure determination problem is a knowledge acquisition problem. The important aspect in this is the way in which the knowledge is acquired in practical applications. The results can be analysed in this context as well. The approach, providing a better knowledge acquisition mechanism and resultant fuzzy measures and preference order, can be considered as a useful approach.

The approaches involve some degree of knowledge acquisition. It is important to understand the role of knowledge acquisition in each of the approaches. It is widely understood that the knowledge is acquired from the expert. The successful acquisition of the past data from the experts will depend upon a particular knowledge acquisition technique.

The approach of determining fuzzy measures using λ measure mainly rests on the weights of individual criteria. The decision maker assigns the individual weights. From a knowledge acquisition perspective, this approach needs to elicit individual weights. This is explicit knowledge acquisition and it is simple to acquire as it involves direct input from the decision maker. However, the simplicity in knowledge acquisition in this approach contradicts the overall results proposed by this approach. It can be observed and concluded from the previous discussion that the results given by this approach are counter intuitive and restrictive to one type of interaction between the criteria.

The approach based on semantic consideration of criteria requires knowledge in the form of criteria weights and the decision maker's preferences on the pair of criteria. It was assumed that this knowledge is acquired from an expert directly. This approach does not need any historical data. It can be said that the knowledge acquisition is direct and hence simple to acquire from the decision maker. The results given by this approach are partially intuitive.

The approach of fuzzy measure acquisition using interactive optimization uses the stored data about past decisions. The global evaluation of the alternatives in the past decisions is used for computing the fuzzy measures. It can be assumed that the data acquisition is done through an expert. This approach requires a larger amount of data as compared with other approaches. This difficulty is also complemented by the disadvantages of quadratic programming.

The proposed methodology primarily rests on acquiring knowledge about the past decisions. In the proposed methodology the knowledge acquisition takes place through expert inputs. The case-base is built using expert feedback about past decisions. In addition to this acquisition, the system also confirms the results using the importance scale. From the knowledge acquisition perspective, it can be said that this approach requires more efforts than the other approaches. However, the results given by the proposed methodology are intuitive. Hence the knowledge acquisition in this approach is justified. This also justifies the significance of the proposed methodology.

In summary, it can be said that knowledge acquisition plays an important role in developing fuzzy measure determination approaches. The overall results given by the decision support system and the underlying knowledge acquisition mechanism must be taken into consideration. These aspects help to choose an appropriate methodology for developing a decision support system using the MCDM model for various real-life decision problems.

5.11 Summary

This chapter discussed the comparison of the proposed approaches with state-of-the-art approaches for fuzzy measure determination. The approach was compared based on the resultant fuzzy measure values and the resultant preference order of the alternatives. The comparison was made in the presence of interactive as well as non-interactive criteria. The proposed approach delivered intuitive results in both cases. This comparative analysis also supports the evaluation of the proposed methodology.

Chapter 6 Conclusion and Future Research

6.1 Conclusion

This research focused on developing a new approach for the acquisition of fuzzy measures in MCDM problems. The application of approaches that resemble human reasoning processes was emphasized. The MCDM model evaluates the set of alternatives against a set of criteria. The most of the decision problems consists of dependent criteria. The interaction between the criteria was discussed in detail and it formed the foundation of the undertaken research question. This research identified the importance of modelling the dependence or interaction between the criteria.

Aggregation is the central aspect of this research. The intensive study of various aggregation operators showed that traditional additive measures such as weighted average do not model the interaction between the criteria. Non-additive measures such as fuzzy measures and fuzzy integrals are suitable tools for modelling the interaction between the criteria and subsequent aggregation process. The investigation of fuzzy measures showed that the acquisition of fuzzy measures is an important issue. This was due to the 2^n-1 fuzzy measure coefficients. The main problem is to determine the ways for acquiring these fuzzy measure coefficients in practical applications. Thus the main research objective was to investigate the ways of fuzzy measures acquisition in solving decision problems under the MCDM setting. In this pursuit, various approaches were analysed. Traditional approaches such as AHP and ELCTRE do not model the interaction between the criteria. The commonly used AHP approach uses the pair-wise comparison of criteria. The global evaluation of the alternatives does not consider the interaction between the criteria. The proposed methodology however, considers the interaction between the criteria as it uses fuzzy measures for this purpose. The other approaches mainly suggest use of past decisions for determining fuzzy measures in new decision problems.

The problem of fuzzy measure determination is considered as the problem of knowledge acquisition. The state-of-the-art methods were critically analysed. These methods suggest use of the same past precedents for fuzzy measure determination. However, these approaches do not handle the subjectivity associated with fuzzy measures. It is also difficult to determine generic values of fuzzy measures in decision problems due to subjective component in decision-making.

This research has extended this notion. Various methods have been proposed to resolve the issue of fuzzy measure acquisition. These methods mainly rest on computing the similarity between the current problem and past problems. The results of this similarity assessment can be used as fuzzy measure coefficients for determining the global worth of each alternative. The proposed approach suggests the application of case-based reasoning for acquisition of fuzzy measures. The proposed case-based decision support system delivered intuitive results.

The main contribution of this research was the development of the similarity-based reasoning model for determining the fuzzy measure coefficients. The proposed method allows eliciting the fuzzy measures from the decision maker. The experimental results indicate that the issue of subjectivity in fuzzy measure determination can be resolved by using expert input in constructing a case base for solving the given decision problem. The process of knowledge acquisition in the form of expert input can be time consuming for building the case base. There is a trade-off between the knowledge acquisition and quality of the case-base and hence eventually the resultant fuzzy measures. This work appreciates this trade-off. The proposed approach was illustrated with a simple but intuitive example to explain its applicability for solving repetitive type of decision making problems. The proposed approach offers an intuitive decision support for such class of decision problems. The decision problems with high number of criteria will require case base larger than the case base with three criteria. It is clear that more number of criteria will require larger case base. It was concluded that the proposed approach could be a suitable way for determining the fuzzy measures in a practical application. This research has made a contribution in advancing the MCDM theory by showing a simple approach that can address the complex issue of fuzzy measure acquisition in MCDM problems.

6.2 Future Research

This research offered a new perspective towards solving the problem of fuzzy measures determination. The proposed approach uses reasoning by analogy as its theoretical foundation. The proposed approach was illustrated using different data sets. This work has initiated interesting approaches. The further development of these approaches remains a future research. The following main issues needs to be considered and resolved in future research.

6.2.1 Semantic Similarity Approach

This work has initiated an interesting development for fuzzy measure determination using semantic similarity between criteria in a given decision problem. In case of semantic similarity-based approach of fuzzy measure determination, the main issue is to develop a mechanism for determining the semantic similarity between a single attribute and a set of individual attributes. This work has proposed and illustrated a basic approach for determining semantic similarity between criteria or attributes, where criteria in the new decision problem is semantically related with one or more attributes in the case-base. This work has shown that such similarity can be computed using a taxonomy of words where the criteria in the given decision problem exist in the hierarchy of concepts in the given taxonomy. However, building a practical application is yet to be considered. This work can be extended using approaches in natural language processing and information retrieval which are the sub-fields in AI and linguistics. This future work may also investigate the area of knowledge representation. Representing the knowledge using a given technique may enable the context for the underlying decision problem to be represented. The complex problems can be simplified by an appropriate choice of knowledge representation which will also address the practical problem of storing and manipulating the knowledge in the case-based decision support systems developed using fuzzy measures.

The extension of this work will involve the study of approaches that will enable automatic generation of fuzzy measure coefficients in a given decision problem using taxonomical data. It may involve study of approaches that search the case-base for the semantically similar attributes. Such study may lead to development of novel approaches that will make contribution to MCDM applications and may establish an important foundation for further research in disciplines such as text retrieval. This remains a future area of research.

6.2.2 Matrix Preconditioning Methods

This work has shown that fuzzy measures can be acquired in case of linear systems in the form of AX = B, where A is the square matrix of criteria evaluations, X is the column vector of fuzzy measure coefficients and B is the column matrix of global evaluations of alternatives. A matrix pre-inverse can be used in case of the system involving a non-square matrix. This work identified the applicability of matrix preconditioning methods for fuzzy measure acquisition. According to this approach, if there is a system of equations as AX = B, then construct a matrix B^{-1} that approximates matrix A. A is symmetric and positive semi-definite matrix. This system can then be converted to a system of the following form using a pre-conditioning matrix.

$$B^{-1}AX = B\,B^{-1}$$

By using this pre-conditioner matrix, coefficients of X can be determined. This theory can be extended and applied into real-life applications. Preconditioning techniques have emerged as an essential part of efficient iterative solutions of matrices. These techniques did not have much impact initially due to the simplicity of their heuristics and the relatively small size of the matrices to be solved. However, in case of iterative methods, a good pre-conditioner matrix is vital for solving the problem of fuzzy measure determination.

There is need to study matrix preconditioning methods in the context of MCDM problems. There is need to investigate the various matrix preconditioning methods for solving fuzzy measure acquisition problems. The theoretical basis of the pre-conditioning methods needs to be tested in real applications. Such future study may lead to development of methods or algorithms that can be applied for addressing the issue of fuzzy measure acquisition under the following conditions for a pre-conditioner matrix:

- Product $B^{-1}A$ must be close to the identity.
- B should be easily inverted (for instance, B is a diagonal or triangular matrix)

This future work will also investigate the applicability of different matrix preconditioning methods for solving complex problems.

Overall, this research has contributed in developing a different approach for solving the undertaken research issue. It has laid the theoretical foundation for development of important approaches and innovative decision support systems. On a concluding note, it can be said that the proposed approach can be useful in making rational decisions based on facts such as absolute values of the criteria in repetitive decisions. However, factors such as intuition can dominate the decision making process over the facts. There is still scope for development of decision support models that can handle human intuition completely or partially. This remains a future research work.

References

Adomavicius, G, Sankaranarayanan, R, Sen, S & Tuzhilin, A 2005, Incorporating contextual information in recommender systems using a multidimensional approach, ACM Transactions on Information Systems (TOIS), v.23 n.1, p.103-145.

Ahlgren, P, Jarneving, B & Rousseau, R 2003, Requirements for a cocitation similarity measure with special reference to Pearson's correlation coefficient. *Journal of the American Society for Information Science and Technology*, 54(6), 550-560.

Aksoy, S, & Haralick, RM 2000, Probabilistic vs. geometric similarity measures for image retrieval, In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, volume 2, pages 357–362, Hilton Head Island, South Carolina.

Alter, SL 1980, Decision Support Systems: Current Practice and Continuing Challenges, Reading, MA: Addison-Wesley.

Althoff, KD 2001, Case-based Reasoning. In S. K. Chang (editor), *Handbook of Software Engineering & Knowledge Engineering,* World Scientific Pub, Co.

Alsina, C, Trillas, E & Valverde, L 1983, On some logical connectives for fuzzy set theory, *Journal of Mathematics Analysis & Applications 93*, 15-26, 1983.

Anderson, DR, Sweeney, DJ & Williams, TA 2000, *An introduction to management science–Quantitative approaches to decision-making*, 9th Ed., Rochester Institute of technology.

Anderson, JR & Thomson, R 1989, Use of analogy in a production system architecture, *Similarity and analogical reasoning,* Vosniadou, S & Ortony, A, (eds), Cambridge University Press.

Baeza-Yates, R & Ribeiro-Neto, B 1999, *Modern Information Retrieval*, Addison– Wesley, ISBN 0-201-39829– X.

Banerjee, M, Mitra, S & Pal, SK 1998, "Rough fuzzy MLP: Knowledge encoding and classification," *IEEE Trans. Neural Networks*, vol. 9, pp. 1203-1216.

Barletta, R 1991, An introduction to case-based reasoning, *AI Expert*, *6*(8), pp. 42–49.

Brill, E 1995, Transformation-based error-driven learning and natural language processing: a case study in part of speech tagging, *Computational Linguistics*, *21*, pp. 543–565.

Bilgic, T & Turksen, I 1997, Measurement of Membership Functions: Theoretical and Empirical Work, *Handbook of Fuzzy Sets and Systems*, Vol.1.

Bergmann, R. 2002, Experience Management: Foundations, Development Methodology, and Internet-Based Applications, Berlin: Springer.

Bergmann R. 2003, Introduction to Case-based reasoning, Case-based reasoning on the web, http://www.cbr-web.org/, viewed 20 July 2003.

Berikov, V & Litvinenko, A 2003, "Methods for statistical data analysis with decision trees", Novosibirsk, Sobolev Institute of Mathematics, http://www.math.nsc.ru/AP/datamine/eng/decisiontree.htm, viewed July 2004.

Blake, C & Mertz, C 1998, UCI Repository of machine learning databases, http://www.ics.uci.edu/simmlearn/MLRepository.html, University of California-Irvine, Department of Information and Computer Sciences, viewed August 2004.

Bonzano, A, Cunningham, P & Smyth, B 1997, "Using Introspective Learning to Improve Retrieval in CBR: A Case Study in Air Traffic Control," in *Proceedings of second International Conference on Case-based Reasoning*, RI, USA, pp. 291-302.

Bonczek, RH, Holsapple, CW & Whinston, AB 1981, *Foundations of Decision Support Systems*, New York: Academic Press.

Bonissone, PP 1985, Selecting uncertainty calculi and granularity: An experiment in trading off precision and complexity, *Proceedings of the First Workshop on Uncertainty in Artificial Intelligence,* Los Angeles, pp. 57 - 66.

Bridge, D 1998, Defining and combining symmetric and asymmetric similarity measures, Procc. of *4th Eur. Workshop on Case-based Reasoning (EWCBR'98),* LNAI-1488, pp. 52-63.

Brown, AL, 1989, Analogical learning and transfer: What develops? *Similarity and analogical reasoning,* Vosniadou, S & Ortony, A, (eds), Cambridge University Press.

Cercone, N & McCalla, G, (eds), 1987, The knowledge Frontier: Essays in the Representation of Knowledge, New York, Springer-Verlag.

Cheetham, AH & Hazel, JE 1969, Binary (presence-absence) similarity coefficients. *J. Paleontol,* 43: 113-136.

Cheetham, W 2000, Case-based reasoning with confidence. In E. Blanzieri and L. Portinale, (eds), EWCBR 2000, volume 1898 of Lecture Notes in Artificial Intelligence, Springer-Verlag, pp.15-25.

Choquet, G 1953, Theory of capacities, *Ann.Inst.Fourier* 5, 131-295.

Cohen, P 1995, Empirical Methods for Artificial Intelligence, MIT Press, Boston

Collins, A & Burstein, M 1989, Afterword: A framework for a theory of comparison and mapping, *Similarity and analogical reasoning,* Vosniadou, S & Ortony, A, (eds), Cambridge University Press.

Cortés, U, Núñez, H, Sànchez-Marrè, M, Comas, J, Ignasi, R & Poch, M, 2002, Feature Weighting Techniques for Prediction Tasks in Environmental Processes, *ECAI Workshop on Binding Environmental Sciences and Artificial Intelligence* (BESAI'2002), Lyon, France.

Cunningham, P, Bonzano, A & Smyth, B, 1997, Using introspective learning to improve retrieval in CBR: a case study in air traffic control, case-based reasoning research and development. In: Leake, DB & Plaza, E, (eds), *Proceedings of the 1997 International Conference on Case-Based Reasoning, Springer Verlag Lecture Notes in Artificial Intelligence*, pp. 291–302.

Deer, P, Modave, F & Tilley, T 2001, Web-based collaborative decision making, *Proc IConIT'2001*, Thammasat University Press, Thailand, pp. 203-212.

Detyniecki, M 2000, Mathematical aggregation operators and their application to Video Querying, Ph.D. Thesis, University of Paris.

Dillon, TS. & Shiu, SCK, 2000, A tutorial on case-based reasoning. In Pal, SK, Dillon, TS & Yeung, DS (eds), *Soft computing in case based reasoning*. London: Springer.

Dubois, D & Prade, H 1981, Triangular norms for fuzzy sets, *Proceedings of the* 2nd International Symposium of Fuzzy Sets, Linz, pp. 39 – 68.

Dubois, D & Prade, H 1985, A review of fuzzy set aggregation connectives, *Information Sciences*, 36: 85 –121.

Dubois, D, Prade, H, & Sandri, S, 1993, "On possibility/probability transformations," *Fuzzy Logic\—State of the Art*, Lowen, R & Roubens, M, (eds), Dordrecht: Kluwer, pp. 103-112.

Dubois, D, Esteva, F, Garcia, P, Godo, L, de Màntaras, RL & Prade H 1999, Case-based Reasoning: A Fuzzy Approach, In: Ralescu, AL, Shanahan, JG, (eds), *Fuzzy Logic in Artificial Intelligence*, IJCAI' 97 Workshop, Berlin: Springer-Verlag, pp. 79-90.

Dubois, D, Prade, H 1994, Similarity-Based Approximate Reasoning, *Computational Intelligence: Imitating life*, New York: IEEE Press, pp. 69-80.

Dubois, D, Esteva, F, Garcia, P, Godo, L, de Màntaras RL, Prade H 1997, Fuzzy set-based models in case-based reasoning, IIIA Research Report, pp. 97-09.

Duda, RO, Hart, PE & Stork, DG 2000, *Pattern Classification,* John Wiley & Sons Inc., New York, 2nd edition.

Dunteman, G, H 1989, Principal Component Analysis, Series: Quantitative applications in the Social Sciences, Sage Publications, USA.

Elkan, C 1994, The paradoxical success of fuzzy logic, *IEEE Expert*, pp. 3-8.

Fellbaum, C, (eds), 1998, WordNet: An Electronic Lexical Database, MIT Press.

Figueira, J, Greco, S & Ehrgott, M, (eds), 2005, Multiple Criteria Decision Analysis: State of the Art Surveys. Springer Science Business Media, Inc., New York.

Finnie, G & Sun, Z 2002, Similarity and metrics in case-based reasoning, *International Journal of Intelligent System,* 17(3), pp. 273-287.

Fuller, R 1997, OWA operators in decision making, *Department of Computer Science*, Budapest, Hungary.

Gebhardt, F, Vo, A, Grather, W & Schmidt-Belz, B 1997, Reasoning with Complex Cases, International Series in Engineering and Computer Science, Vol. 393, Kluwer Academic Publishers, Boston.

Gentner, D, Holyoak, KJ, Kokinov, B, (eds.), 2001, The Analogical Mind: Perspectives from Cognitive Science, Cambridge, MA, MIT Press.

Geng L & Hamilton, HJ 2006, Interestingness Measures for Data mining: A Survey", ACM Computing Surveys, Vol.38, No.3, Article 9.

Gower, JC, 1971, A general coefficient of similarity and some of its properties, *Biometrics*, *27*, pp. 857–874.

Grabisch, M 1995a, Fundamentals of Uncertainty Calculi with Applications to Fuzzy Inference, Kluwer Academic Publishers.

Grabisch, M 1995b, Fuzzy integral in multi-criteria decision making, *Fuzzy Sets and Systems*, 69, pp. 279-298.

Grabisch, M 1996, The application of fuzzy integral in multi criteria decision making, *European Journal of Operational Research*, 89(3), pp. 445-456.

Grabisch, M & Modave, F 1997, Preferential independence and Choquet integral, In 8th International Conference On the Foundations and Applications of Decision under Risk and Uncertainty (FUR), Mons, Belgium.

Gregson, RAM 1975, *Psychometrics of Similarity,* Academic Press, New York.

Falkman, G 2000, Similarity measures for structured representations: a definitional approach. In: Blanzieri E, Portinale L, (eds), Advances in case-based reasoning. *Proceedings of the 5th European Workshop*, EWCBR 2000, Trento, Italy, September 6-9, Lecture notes in artificial intelligence 1898 (Springer-Verlag, Berlin Heidelberg, 380-392.

Faries, J & Schlossberg, K 1994, The effect of similarity on memory for prior problems, *In Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society*, Hillsdale, NJ: Lawrence Erlbaum, pp. 278-282.

Fishburn, PC & Lavalle, IH 1999, "MCDA: Theory, Practice and the Future." *Journal of Multi-criteria Decision Analysis* 8, no. 1, pp.1–12.

Hayes C & Cunnigham P 1999, Shaping a CBR view with XML, In Proceedings of International Conference on CBR (ICCBR'99) Munich, Germany, pp 468-81.

Hammond, K 1989, Case-Based Planning: Viewing Planning as a Memory Task, Chapter 1, Academic Press.

Hamming, RW 1986, Coding and Information Theory, Prentice-Hall, Englewood Cliffs.

Hansen, BK, Riordan, D 2002, A fuzzy case-based system for weather prediction, Engineering Intelligent Systems 3, pp.139–145.

Hohle, U 1978, Probabilistic uniformization of fuzzy topologies, *Fuzzy Sets and Systems Vol. 1*.

Holyoak, KJ & Koh, K 1987, Surface and structural similarity in analogical transfer, *Memory & Cognition*, 15(4), pp. 332-340.

Holyoak, KJ & Thagard, PR 1989, Analogical Mapping by Constraint Satisfaction, *Cognitive Science*, 13, pp. 295-355.

Holyoak, KJ & Thagard, P 1995, *Mental Leaps: Analogy in Creative Thought*, Cambridge, MA: MIT Press.

Helman, DH 1988, *Analogical Reasoning*, Kluwer Academic Publishers, pp. 65-103.

Hwang, CL & Yoon K 1981, Multiple Attribute Decision Making Methods and Applications, Springer-Verlag.

Hoffer, J 1996, *Modern Systems Analysis and Design*, The Benjamin/Cummings Publishing Company, USA.

Hotelling, H 1933, Analysis of a Complex Variables into Principle Components, *Journal of Educational Psychology*, vol. 27, pp. 417-441.

Ichino, M & Yahuchi, H 1994, Generalized Minkowski metrics for mixed feature-type data analysis, *IEEE Trans. Syst. Man Cybern*. 24, pp. 698-708.

Kardi,T 2004, http://people.revoledu.com/kardi/tutorial/Similarity/index.html viewed March 2005.

Karhunen, K 1946, 'Zur Spektraltheorie Stochastischer Prozesse', *Ann. Acad. Sci. Fennicae*, 37.

Khator, SK., Weber-Lee, R & Barcia, RM 1995, Case-Based Reasoning for Cash-Flow Forecasting using Fuzzy Retrieval, *Proceedings of the First International Conference on Case-Based Reasoning Research and Development*, October 23-26, Springer Verlag, London, UK, p.510-519.

King, D 1990, Intelligent Decision Support: Strategies for Integrating Decision Support, Database Management and Expert System Technologies, Expert Systems with Applications, Vol. 1, pp. 23-38.

King, D 1993, 'Intelligent support systems: art, augmentation, and agents', in Sprague, RH, Jr & Watson HJ, (eds), *Decision Support Systems: Putting Theory into Practice*, 3rd edn, Englewood Cliffs, NJ: Prentice Hall. (An account of integrated intelligent system development trends and the future of DSS and expert systems.)

Kinley, A, Leake, D & Wilson, D 1996, Acquiring case adaptation knowledge: A hybrid approach, *In Proceedings of the Thirteenth National Conference on Artificial Intelligence*, Menlo Park, CA: AAAI Press, pp. 684-689.

Klein, G & Calderwood, R 1988, How do people use analogues to make decisions? *In Kolodner, J, (eds), Proceedings of the DARPA Case-Based Reasoning Workshop*, Palo Alto: DARPA, pp. 209-223.

Klein, GA 1987, Applications of analogical reasoning, *Metaphor and Symbolic Activity*, 2, pp. 201-218.

Klein, G & Calderwood, R 1989, Decision models: Some lessons from the field, *IEEE Transactions on Systems, Man, and Cybernetics* 21(5), pp. 1018-1026.

Klein, M & Methlie, L 1995, *Knowledge-Based Decision Support Systems with Applications to Business*, M. Klein and L. Methlie, Wiley, New York.

Keane, MT, Duff, S & Ledgeway, T 1994, Constraints on analogical mapping: A comparison of three models, *Cognitive Science*, 18, pp. 387-438.

Keen, PGW & Scott-Morton, MS 1978, *Decision Support Systems: An Organizational Perspective*, Reading, MA: Addison-Wesley.

Keeney, RK & Raiffa H, 1976 Decisions with multiple objectives, preferences and value trade-offs, John Wiley and Sons, New York.

Klir, GJ & Folger, TA 1988, Fuzzy Sets, Uncertainty and Information, *Prentice-Hall:* Englewood Cliffs, NJ.

Klir, GJ, Wang, Z & Harmanec, D 1997, Constructing fuzzy measures in expert systems, *Fuzzy Sets and Systems* 92, pp. 251–264.

Klir, GJ, Wang, Z & Wang, W 1996a, Monotone set function defined by Choquet integral, *Fuzzy Sets and Systems*, 81(2), pp. 241–250.

Klir, GJ, Wang, Z & Wang, W 1996b, Constructing fuzzy measures by rational transformations, *Journal of Fuzzy Math.,* 4(3), pp. 665–675

Kokinov, B 1996, Analogy-Making: Psychological Data and Computational Models, In Kokinov, B, (eds), *Perspectives on Cognitive Science*, Vol. 2, NBU Press.

Kolodner, J 1991, "Improving human decision making through case-based decision aiding," Artificial Intelligence Magazine 12(2): pp. 52–68.

Kolodner, JL 1992, An introduction to case-based reasoning. Artificial Intelligence Review 6, pp. 3–34.

Kolodner, JL 1993, Case-Based Reasoning, Morgan Kaufmann, San Francisco, pp. 27-28.

Kontkanen, P, Lathinen, J, Myllymäki, P & Tirri, H 2000, An unsupervised Bayesiandistance measure, Proceedings of *5th E. W. on Case-based Reasoning (EWCBR'2000),* LNAI-1898, pp. 148-160.

Koton, P 1988, Reasoning about evidence in causal explanations, In *Proceedings of the Seventh National Conference on Artificial Intelligence*, San Mateo, CA: Morgan Kaufmann, pp. 256-261.

Lancaster, J & Kolodner, J 1987, Problem solving in a natural task as a function of experience, *In Proceedings of the Ninth Annual Conference of the Cognitive Science Society,* Hillsdale, NJ: Lawrence Erlbaum, pp. 727-736.

Leake, DB 1995, Becoming An Expert Case-Based Reasoner: Learning To Adapt Prior Cases, Proceedings of the Eigth Annual Florida Artificial Intelligence Research Symposium, pp. 112-116.

Leake, DB 1996, Case-Based Reasoning: Expriences, Lessons, & Future Directions. Menlo Park, California: AAAI Press.

Leake, D, Kinley, A & Wilson, D 1997, Case-based similarity assessment: Estimating adaptability from experience. In *Proceedings of the Fourteenth National Conference on Artificial Intelligence*, AAAI Press.

Leake, DB & Wilson DC 1999, When experience is wrong: examining CBR for changing tasks and environments, *In Proceedings of 3rd International Conf. on Case-Based Reasoning (ICCBR)*, Springer-Verlag, Lecture Notes in Computer Science 1650, pp. 218-232.

Lin, D 1998, An Information-theoretic definition of similarity, *Proceedings of the Fifteenth International Conference on Machine Learning*, ISBN: 1-55860-556-8, pp. 296-304.

Lenz, M 1999, Case Retrieval Nets as a Model for Building Flexible Information Systems, Dissertation Humboldt Universität, Berlin.

Leung, KS, Wang, Z & Wang, J 2000, Determining nonnegative monotone set functions based on Sugeno's integral: an application of genetic algorithms, *Fuzzy Set Syst.* 112, pp. 155–164.

Liao, TW, Zhang, Z & Mount, CR 1998, Similarity measures for retrieval in case-based reasoning systems, *Applied Artificial Intelligence,* 12, pp. 267-288.

Loève, MM 1955, *'Probability Theory'*, Princeton, N.J.: VanNostrand, Main.

Malczewski, J 1999, *GIS and Multicriteria Decision Analysis*, John Wiley and Sons, New York, NY, 392 pp.

Mallach, E 2000, Decision Support and Data Warehouse Systems, Irwin McGraw-Hill.

Marichal, J 1999, Aggregation Operators for Multicriteria decision aid, PhD Thesis University of Li`ege, Belgium.

Marichal, J & Roubens, M 2000, Determination of weights of interacting criteria from a reference set, *European Journal of Operational Research* 124 (3) 641-650.

Mark, W & Kolodner, JL 1992, Case-based reasoning, IEEE Expert, vol. 7, no.5, pp. 5-6.

Martel, J 1999, *Multicriterion Decision Aid: Methods and Applications*, CORS-SCRO, Annual Conference, Windsor, Ontario.

Mesiar, R & Komorníková M 1997, Aggregation Operators, *Proceeding of the XI Conference on applied Mathematics,* Herceg D., Surla K (eds), Institute of Mathematics, Novi Sad, pp. 193-211.

Menger, K 1942, Statistical metrics, *Proc. Nat. Acad. Sci. U.S.A. 8*, pp. 535-537.

Miettinen, K 1999, *Nonlinear Multiobjective Optimization,* Kluwer Academic Publishers.

Miller, G 1990, WordNet: An on-line lexical database, *International Journal of Lexicography*, 3(4).

Minsky, M 1963, *Computers and Thought*, 'Introduction', in Feigenbaum, EA & Feldman, J (eds.), McGraw-Hill, New York.

Nguyen NB & Ho TB 2000, A mixed similarity measure in near-linear computational complexity for distance-based methods, In: Zighed DA, Komorowski J, Zytkow J (eds), *Principles of Data Mining and Knowledge Discovery (PKDD* 2000), LNAI 1910, Berlin: Springer Verlag; pp 211-220.

Nguyen, BV 2002, AAA: Another Academic Advisor using Case-Based Reasoning, *CS690 (Case-Based Reasoning) report.*

O'Connell, MJ 1974, Search program for significant variables. *Comput. Phys. Commun.,* 8**,** pp. 49–55.

Olson, DL 1996, Decision Aids for Selection Problems, New York: Springer.

Osborne, HR & Bridge, D 1996, A case-based similarity framework. Procc. of *3rd Eur.Work.. on Case-based Reasoning (EWCBR'96)*. LNAI-1168, pp. 309-323.

Pal, SK, Dillon, TS & Yeung, DS (eds) 2000, *Soft Computing in Case Based Reasoning* London, U.K.: Springer-Verlag.

Pal, S & Shiu, S 2004, *Foundations of Soft Case-Based Reasoning*. Wiley-Interscience: Hoboken, NJ.

Pal, SK & Mitra, S 1992, Multi-layer perceptron, fuzzy sets and classification, IEEE Trans. Neural Networks, 3, pp. 683-697.

Pan, VY, Murphy, B, Rosholt, RE, Tang, Y, Yan, X 2005, Additive Preconditioning in Matrix Computations, Technical Report TR 2005009, *Ph.D. Program in Computer Science, Graduate Center, City University of New York*, New York.

Perera, RS & Watson, I 1998, "A hierarchical case representation using context guided retrieval" *Knowledge Based Systems Journal* 11(5-6), pp. 285-292.

Pirolli, P & Anderson, J 1985, The role of learning from examples in the acquisition of recursive programming skills. Canadian Journal of Psychology 39, pp. 240-272.

Plaza E & Aamodt, A 1994, Case-based reasoning, Foundational issues, methodological Variations, and system approaches, Artificial Intelligence Communications, IOS Press, 7(1), pp. 39-59.

Plaza E, Esteva F, Garcia P, Godo, L, & López de Màntaras R 1996, A logical approach to case-based reasoning using fuzzy similarity relations. *Info Scie,* 106, 105-122.

Plaza, E, Arcos, JL & Martin, F 1997, Cooperative case-based reasoning. In: *Distributed Artificial Intelligence meets Machine Learning,* LNAI 1221, Springer Verlag, pp. 180-201.

Priya, S & Shibasaki, R 2001, National spatial crop yield simulation using GIS-based crop production model, *Ecological Modelling*, 135, pp. 113-129.

Polya, G 1945, *How to Solve It*, Princeton University Press, Princeton, NJ.

Rice, JA & Cochr, PA 1984, Independent evaluation of a bioenergetics model for largemouth bass, *Ecology*, 65(3), pp. 732-739.

Read, S & Cesa, I 1991, This reminds me of the time when . . .: Expectation failures in reminding and explanation. *Journal of Experimental Social Psychology*, 27, pp. 1-25.

Resnik, P 1995, Using information Content to Evaluate Semantic Similarity in a Taxonomy, *Proceedings of the 14th International Joint Conference on Artificial Intelligence.*

Resnik, P 1998, Semantic Similarity in a Taxonomy: An Information-Based Measure and its Application to Problems of Ambiguity in Natural Language. Journal of Artificial Intelligence Research, 11, pp. 95-130.

Ross, B 1984, Remindings and their effects in learning a cognitive skill, *Cognitive Psychology* 16, pp. 371-416.

Roiger, R., & Geatz, M 2003, Data mining: a tutorial-based primer, Boston, Addison-Wesley.

Rosenzweig, B & Silvestrova, E 2000, Oracle PL/SQL, Upper Saddle River, NJ: Prentice-Hall PTR, pp 447.

Roy, B 1968, Classement et choix en présence de points de vue multiples (la méthode ELECTRE), *Revue d'Informatique et de Recherche Opérationnelle (RIRO)* 8, 57-75.

Roy, B 1996, Multicriteria Methodology for Decision Analysis, Kluwer Academic Publishers, Dordrecht.

Roy, B & Vanderpooten, D 1996, "The European school of MCDA: Emergence, basic features and current works", Journal of Multi-Criteria Decision Analysis 5, pp. 22-38.

Satty TL 1980, "The Analytic Hierarchy Process", New York: McGraw-Hill.

Saad, Y 1996, *Iterative methods for sparse linear systems*. PWS, ISBN 0-534-94776-X.

Santini, S & Jain, R 1999, Similarity Measures, IEEE Trans. Pattern Analysis and Machine Intelligence 21 (9), pp. 871–883.

Schmidt, H, Norman, G & Boshuizen, H 1990, A cognitive perspective on medical expertise: Theory and implications, *Academic Medicine* 65(10), pp. 611-621.

Schweizer, B & Sklar A 1983, Probabilistic metric spaces, *North Holland*, New

York.

Schweizer, B & Sklar, A 1960, Statistical Metric spaces, *Pacific J. Math. 10, pp.* 313-334.

Sefion, I, Ennaji, A & Gailhardou, M 2003, "ADEMA: A Decision Support System for Asthma Health Care", In *Proceedings of Eighteenth International Congress of the European Federation for Medical Informatics*, pp 623-628.

Shapley, LS 1953, A value for n-person games. In Kuhn, HW & Tucker, AW (eds), *Contributions to the Theory of Games, Vol. II*, number 28 in Annals of Mathematics Studies, Princeton University Press, pp. 307–317.

Shin, KS & Han, I 2001, A case-based approach using inductive indexing for corporate bond rating, Decision Support Systems, 32, pp. 41–52.

Silvert, W 1979, Symmetric summation: a class of operations on fuzzy sets, IEEE Transactions on Systems, Man and Cybernetics *9,* 659-667.

Simoudis, E 1992, Using case-based retrieval for customer technical support, IEEE Expert, pp. 7-12.

Spiro, RJ, Feltovich, PJ, Coulson, RL & Anderson, DK 1989, Multiple analogies for complex concepts: antidotes for analogy-induced misconception in advanced knowledge acquisition, *Similarity and analogical reasoning,* Vosniadou, S & Ortony, A (eds), Cambridge University Press.

Sprague, RH, Jr. & Carlson ED 1982, *Building Effective Decision Support Systems*, Englewood Cliffs, NJ.

Sharda, R, Bar, SH & McDonnell, JC 1988, Decision support system effectiveness: A review and an empirical test, Management Science 34(2), pp. 139-159.

Shannon, CE & Weaver W, 1963, A Mathematical Theory of Communication, Bell System Technical Journal, Vol. 27, pp. 379-423,623-656.

Sharma, P 2001, Comparative analysis of aggregation operators in fuzzy multicriteria decision making, Post-graduate thesis, Griffith University, Gold Coast, Australia.

Steurer RE, Gardiner LR & Gray J 1996, A Bibliographic Survey of the Activities and International Nature of Multiple Criteria Decision Making, *Journal of Multi-Criteria Decision Analysis, Vol.* 5, No.3, pp. 1255-1260.

Smyth, B & Keane, MT 1998a, Adaptation-Guided Retrieval: Questioning the Similarity Assumption in Reasoning, Artificial Intelligence, 102, pp. 249-293.

Smyth, B & McKenna, E 1998b, Modelling the competence of case bases, in *Proceedings of the Fourth European Workshop on Case-based Reasoning (EWCBR-98)*, Dublin, Ireland, Springer-Verlag, Berlin, pp. 208-220.

Smyth, B 1998c, Case-Base Maintenance, *Proceedings of IEA-98-AIE, Lecture Notes in Computer Science,* Springer-Verlag, Berlin Heidelberg, New York.

Sneath, PHA & Sokal, RR 1973, *Numerical Taxonomy: the Principles and Practices of Numerical Classification*. W. H. Freeman, San Francisco.

Segre, AM 1989, Machine Learning of Robot Assembly Plans, Kluwer Academic Publishers.

Sugeno, M 1974, Theory of fuzzy integrals and its applications, PhD Dissertation, Tokyo Institute of Technology, Japan.

Sun, Z & Finnie, G 2003, Intelligent techniques in E-Commerce: A Case Based reasoning Perspective, Physica-Verlag, Berlin.

Takahagi, E 2000, Fuzzy Integral Calculation Site, http://www.isc.senshu-u.ac.jp/~thc0456/Efuzzyweb/, viewed June 2006.

Thagard, PR & Holyoak, KJ 1989, A computational model of analogical problem solving, *Similarity and analogical reasoning,* Vosniadou, S & Ortony, A (eds), Cambridge University Press.

Thirumalaivasan, D 2001, Aquifer vulnerability assessment using analytic hierarchy process, *Proceedings of the 22nd Asian Conference on Remote Sensing*.

Tsoukiàs, A 2003, "From decision theory to decision aiding methodology", DIMACS Technical Report 2003-21.

Ullman, JD 1997, Information integration using logical views. In Proc. of the *6th Int. Conf. on Database Theory* ICDT-97, Lecture Notes in Computer Science, pp. 1940. Springer-Verlag.

Vanthienen, J 2001, Ruling the Business: About Business Rules and Decision tables, Vandenbulcke, J & Snoeck, M (eds.): *New Directions in Software Engineering*, Leuven University Press, Leuven, pp. 103–120.

Vincke, P 1992, *Multicriteria Decision-Aid,* John Wiley & Sons, New York, 1992.

Vosniadou, S & Ortony, A 1989, *Similarity and Analogical Reasoning*, Cambride University Press.

Wagholikar, A 2002, An investigation on dependence of criteria in multicriteria decision making, Post-graduate Thesis, Griffith University, Gold Coast, Australia.

Wagholikar, A & Deer, P 2004, Non-additive Measures In Multicriteria Decision Making, In *Proceedings of International Conference on Computational Intelligence for Modelling*, Control & Automation, Gold Coast, Australia.

Wagholikar A & Deer, P 2005a, Methods for Fuzzy Measure Determination: A Theoretical Study, *Proc. for International Conference on Artificial Intelligence* (ICAI'05), Las Vegas, USA.

Wagholikar, A & Deer, P 2005b, Empirical Determination of Fuzzy Measures Using Case-based Reasoning, *Proc. for 2nd Indian International Conference on Artificial Intelligence*, Pune, India.

Wagholikar A. and Deer, P 2006a, An Interactive Model for Fuzzy Measure Determination, In *Proc.* 2nd International Conference on Information Management and Business, Sydney, Australia, February.

Wagholikar A & Deer, P 2006b, A Similarity-based Approach for Determination of Monotonic Set Functions, *Proc. of The Third International Conference on Neural, Parallel & Scientific Computations*, Atlanta, USA, August

Wang, Z, Leung, KS & Wang, J 1999, "A genetic algorithm for determining non additive set functions information fusion," *Fuzzy Sets Systems,* vol. 102, pp. 463-469.

Wang, Z & Klir, GJ 1992, Fuzzy Measure Theory, Plenum Press, New York.

Watson, I 1997, Applying Case-based Reasoning: Techniques for Enterprise Systems. San Francisco, California: Morgan Kaufmann Publishers.

Watson, I & Gardingen, D 1999, A distributed case-based reasoning application for engineering sales support. In: *Proc 16th Int Joint Conf on Artificial Intelligence (IJCAI-99),* Vol. 1, Morgen Kaufmann Publishers, pp. 600-605.

Watson, I & Marir, F 1994, Case-based Reasoning: A Categorised Bibliography, *The Knowledge Engineering Review*, 9 (3).

Watson, I 1998, CBR is a methodology not a technology. *In, Research & Development in Ex-pert Systems XV, Miles*, R, Moulton, M, Bramer, M (Eds.), Springer, London, pp. 213-223.

Wilke, W & Bergmann, R 1998, Techniques and Knowledge Used for Adaptation During Case-Based Problem Solving, *Proceedings of IEA-98-AIE, Lecture Notes in Computer Science*, Springer-Verlag, Berlin Heidelberg New York.

Wilson, DR & Martínez, TR 1997, Improved Heterogeneous Distance Functions, *Journal of Artificial Intelligence Research*, 6, pp. 1-34.

Wilson, DC & Leake, DB 1998, Categorizing case-base maintenance: dimensions and directions, in Proceedings of the fourth European Workshop on Case-based reasoning (EWCBR-98), Dublin, Ireland, Springer-Verlag, Berlin, pp. 197-207.

Yager, RR, 1988, On ordered weighted averaging aggregation operators in multi-criteria decision making, IEEE Transactions on Systems, Man and Cybernetics, 18, pp. 183-190.

Yang, C, Orchard, R, Farley, B & Zaluski, M 2003, Automated case base creation and management, Proceedings of the Sixteenth International Conference on Industrial & Engineering Applications of Artificial Intelligence and Expert Systems, Loughborough, Springer, UK.

Yuan, B & Klir, GJ 1995, Fuzzy Sets and Fuzzy Logic: Theory and Applications, Prentice-Hall PTR, Upper Saddle River, NJ.

Yuan, B & Klir, GJ 1996, Constructing fuzzy measures: a new method and its application to cluster analysis. In: *Proc. NAFIPS'96*, pp. 567–571.

Zadeh, LA 1965, "Information and Control" *Fuzzy Sets,* 8(3), pp. 338-353.

Zadeh, L 1975, The concept of linguistic variable and its applications to approximate reasoning, Part I, *Information Sciences*, pp. 199-249, Part II, *Information sciences* 8, pp. 301-357, Part III, *Information Sciences* 9, pp. 43-80.

Zadeh, L 1978, Fuzzy sets as a basis for a theory of possibility, *Fuzzy sets and systems 1*, pp. 3-28.

Zeleny, M 1982, *Multiple Criteria Decision Making*, New York: McGraw-Hill.

Zeleny, M 1987, Management support systems: towards integrated knowledge management, *Human Systems Management* 7 (1), pp. 59–70.

Zopounidis, C & Doumpos, M 2002, *Multicriteria Decision Aid Classification Methods.* New York: Springer.