CZECH UNIVERSITY OF AGRICULTURE IN PRAGUE
FACULTY OF ECONOMICS AND MANAGEMENT
DEPARTMENT OF INFORMATION ENGINEERING

INTEGRATING MULTIPLE FUZZY EXPERT SYSTEMS UNDER VARYING REQUIREMENTS

DISSERTATION THESIS

By Eng.

SHADY REFAT SAAD EL-DIN ALY

M. Sc. In Production Engineering
Helwan University, Cairo

Supervised by

Prof. Eng. IVAN VRANA, DrSc.

Information Engineering

Prague, 2006

DISSERTATION THESIS

INTEGRATING MULTIPLE FUZZY EXPERT SYSTEMS UNDER VARYING REQUIREMENTS

By Eng.

Shady Refat Saad El Din Aly

Submitted to
Faculty of Economics and Management, Czech University of Agriculture in Prague,
In fulfillment of the equirements for the degree of Doctor.

June 2006,
Prague

Thesis Supervisor:

Prof. Ing. Ivan Vrana, DrSc.
Department of Information Engineering
Faculty of Economics and Management
Czech University of Agriculture
Prague, the Czech Republic

Contents

Acknowledgement

First and foremost, I am deeply thankful to the Great God, by the grace of whom the progress and accomplishment of this work became a fact.

I would like to express my sincerest gratitude to Professor **Ivan Vrana, DrSc**, my advisor, for his supervision, valuable suggestions, and continuous encouragement. Actually, I have learned from him a lot, and any success I have got during my study in Czech Republic is attributed to him after the Great God's help.

I would like to deeply thank all staff members of my department for love and help; they have been my family in Czech Republic.

Also, I wish to express my gratitude and love to the great Czech Republic.

This is dedicated to my lovely Egypt.

Abstract

Usually the binary or Yes-or-No decision making problems constitute critical and decisive problems in many disciplines, such as in business, economics, agriculture, military field, engineering, medicine…, etc. Most often these types of decision problems are ill-structured, and consequently requires multiple, different and specific expertise's. This is due to the nature of such problems in which not all input variables are known, and decisions may be influenced by multiple different, relevant aspects, and accordingly multiple corresponding expertise's are required. Fuzzy expert systems (FESs) are widely used to model expertise's due to their capability to model real world values, which are not always exact, but frequently vague, subjective, and uncertain. In this research thesis, different synergetic expertise's, relevant to the decision context, are to be integrated and modeled using corresponding FESs. The integration is done through combining or aggregating the decision outputs of the integrated FESs. In addition, there are several imposed restricting requirements on the integration problem. Some of these requirements are general, and are related to the way and type of information used to combine or aggregate the FESs, the format of combined or aggregated outputs, and the similarity or uniqueness of the participating FESs. Other requirements are specific, and are related to the roles and relationships among the FESs and special influences of some of them. In order to realize effective and objective integration of these FESs, while satisfying the imposed requirements, first a unified psychometric numerical scale is standardized for the outputs of all FESs. This scale ranges from 0 to 10, where the value 0 represents complete bias toward "No" decision and the value 10 represents complete bias toward "Yes" decision. Then, Every FES should produce a crisp numerical output expressing the degree of bias toward "Yes" or "No" decision. Intermediate values produced by FESs reflects the degree of bias either to "Yes" or "No" decision. Then, the integration problem is structured in order to organize the efforts toward its solution. The problem is basically structured into two cases of integration; integrating multiple FESs sharing same or common domain knowledge, and integrating multiple FESs each of which has different or unique domain knowledge. In the first case, the integration is done through combining the crisp outputs of the knowledge-equal FESs, whereas, in the second case, the integration is done through aggregating the outputs of the knowledge-unique FESs. For the first case, the classical combining criteria are adopted, and new promising criterion has been developed for dealing with the first case. Also, novel consensus-based heuristics are developed and used to compare the newly developed combining criterion to the prominent of classical one. The experimental results have showed the superiority of the newly developed criterion to the classical well-known Arithmetic Mean criterion and other classical ones, especially with regard to decisiveness. Then, a hierarchical fuzzy model is developed to select the most adequate combining criterion from among several considered ones. For the second case, an aggregation heuristic is developed to accumulate the decision outputs of the multiple knowledge-unique FESs. In both cases of the integration problem, the Analytical Hierarchy Process (AHP) was used to compute the weights of the participating FESs. Then, for the requirement of relying on the past expertise's' performance data and knowledge, a Multi-layer Feed-forward Back-propagation neural net was proposed to learn past numerical data patterns of expertise's performance, where the correct decision answer is known and recorded. Also, a hierarchical fuzzy model was developed to combine/aggregate the FESs' output in case of availability of past If-Then knowledge. Finally, heuristic algorithms and practical suggestions are presented for the satisfaction of specific requirements, which have clearly showed the flexibility of the established numerical outputs scale in converting logical notions into practical objective solutions.

Introductory summary

This dissertation work presents a novel design solution to the ill-structured, multi-aspect binary decision making problems. This design solution is in form of multiple, independent, separated FESs integrated to comprehend all decision aspects and expertise's that are relevant to a particular decision making context. The integration is proposed through combining/aggregating the decision outputs of the multiple FESs into a one consolidated group decision. In addition, there several imposed requirements that should be taken into account. Some of these requirements are general and are related to the method and information used in integration. Other requirements are specific and related to the roles and special influences of the participating FESs.

This thesis is devoted to deal with the proposed integration problem in order to obtain objective and adequate solution approaches, while satisfying the imposed requirements. The thesis consists of the following chapters:-

Chapter 1: contains an introduction, which reviews the basic component of a FES, the unit of integration, with especial emphasis on the crisp nature of its decision output. Then, the reasons behind constructing multiple independent FESs are explained in details. After, the basic types and main objectives of systems integration, in general, are briefly stated. Then, the topic of group decision making (GDM), which represents an underpinning foundation of this research, is highlighted. Next, the possible configurations of integrating expert systems are explained, with a focus on the proposed configuration of the integrated FESs.

Chapter 2: makes a review of literature of all topics those have relations to, or touch the integration problem, or those are strong candidate to solve the problem. Reviewed topics are systems integration (in general), GDM, AHP and its role in GDM, patterns classifiers in the field of pattern recognition, hierarchical fuzzy systems (HFS), and neural classifiers. Then, a special highlight is focused on the reported evidences of the importance and strength of group problem solving (GPS) and systems integration.

Chapter 3: structures and formulates the integration problem and the candidate solution approaches, according the imposed requirements. The general and specific requirements are stated and described more in details in this chapter.

Chapter 4: presents different approaches to weight the relative importance's of the participating FESs under different decision making circumstances. The evaluation is either based on the AHP, utilizing present information, or misclassification rate performance computed utilizing past expertise's' performance data. Illustrative examples are provided.

Chapter 5: considers two cases of the integration problem, defined in chapter 3, the combination and aggregation, and presents solution approaches adequate for both cases. The adequate well-known classical combining criteria are adopted and configured. New promising combining criterion is introduced, and verified according to the well-known social choice theorems. Finally, an aggregation heuristic is presented.

Chapter 6: presents a new approach to combine the outputs of FESs through the consensus analysis employed in GDM. The previously developed consensus indicators are adopted and configured to suit the integration problem structured in chapter 3. Then, some improvement is made for the previously developed set of indicators, and new consensus measures are also defined. Based on the information extracted from these measures, some consensus-based heuristics are introduced. An illustrative example is finally presented.

Chapter 7: conducts some experimentation to compare the performance of the newly developed combining criterion to that of the existing ones, all presented in chapter 5. One of consensus-based heuristics presented in chapter 6 is used to determine a datum level for this comparison. Finally, a comment on the obtained results is made.

Chapter 8: describes a HFS-based model for selecting among the combining criteria adopted and developed in chapter 5.

Chapter 9: presents two approaches to handle and exploit available past expertise's' performance data and knowledge in combining/aggregating the outputs of FESs. The first approach involves the adoption of the Multi-layer Feed-forward Back-propagation neural net for learning and mapping the relationship between the past expertise's' data patterns and the recorded correct decision answer. The second approach involves the use of a HFS-based model to combine/aggregate the outputs of FESs, based on an existing If-then past knowledge.

Chapter 10: specifies how to provide for satisfying the specific requirements stated in chapter 3, and based on handling present outputs' information. A heuristic, mathematical formulas and practical suggestions are presented.

Chapter 11: describes a potential, practical application for this research work. Then, it states the added values expected form applying the research results in such real application. Other possible applications are also suggested.

Chapter 12: terminates the thesis by stating the overall results and achievements realized in relation to the pre-established objectives, and making conclusions about the features, characteristics, and capabilities of the proposed solution approaches, and the justification of this research. Some suggestions regarding possible extensions and future research attempts are finally made.

Motivation

Yes-or-No type decision making problem (also binary decision making or two-group classification) is a common problem in today's business, economic and industrial world. Usually such type of decision making constitutes a crucial decision. For instance, medical diagnosis when assigning patients to one of two groups (at risk, not at risk) based upon some medical observations and tests. Banks are using classification rules to distinguish solvent firms from those companies that may soon end up in bankruptcy. Credit card companies employ classification methods to detect small percentage of credit cards that are being used fraudulently. Examples of other applications of two-group classification analysis include fault detection and machine failure, and decision to launch new investment in a new product. All of these two-group classification decisions have significant economic and social consequences, as an error of assigning an object to the wrong group may lead to catastrophic results. A failing savings and loan may be misclassified as solvent and allowed to operate to the detriment of its depositors. Because of the importance of making the correct decision, researchers and practitioners are constantly looking for better decision making procedure. Due to the continuous advancement and the followed specialization in the industrial and commercial fields, the increasing amount, speed, and diverse information exchange in our world and its associated uncertainty, and the ever changing economic conditions, the complexity of the binary decision making problem increases, in form of ill-structuredness, uncertainty, subjectivity and vagueness. FES is widely utilized in dealing with complex, ill-structured decision making problems, due to its capability to deal with complex ill-structured problems, and which has the ability to treat vagueness and uncertainty, and provide for handling subjective factors as well. Human expertise, knowledge and intuition modeled and represented within FES, have been proven to perform very well, under the condition of ill-defined boundaries, vagueness, and uncertainty. However, because of the critical nature of the binary decision making problems and their inherent complexity, and because of the need to obtain a reliable, good quality decision solution, usually several expertise's are required. The need for multiple expertise's ensues from that usually the decision problem has multiple aspects and touches different knowledge's and expertise's. In addition, dependence on experts is likely to continue as the industrial and commercial world keep specializing while professionals and experts increase their level of expertise in smaller and smaller areas. Consequently, the systems needed for decision support are required to be open-ended, flexible, adaptive, and cover a wide range of expertise's. Unfortunately, huge FES is not the correct choice, as it might be thought. This because that huge expert systems, are cumbersome to maintain, modify, control. In addition, large-scale expert systems containing thousands of rules quickly can overload the memory and make the application difficult to operate. Based on all the above, the integration of multiple FESs is proposed as a one way to deal with the aforementioned complexity inherent with the binary decision making problems, which requires multiple expertise's to cope with such complexity.

Actually, the idea of integrating multiple FESs was practically motivated by a currently held project at the Custom Administration of the Czech Republic, concerning the design of intelligent system to support decision making process in detecting suspicious custom declaration transactions made by exporters and importers. Determining whether or not a current declaration transaction is suspicious is a type of binary decision making, or two-group classification problem. There are many types of commodities about which custom declaration transactions are made; bicycles and motorcycles,…, etc., alcohol drinks, foods, cloths, and other commodities. Accordingly different types of commodity-related expertise's are needed. In addition, expertise's of different types like legal, economical, technical, historical, and territorial...etc, may be simultaneously required in judging a decision problem. Consequently, it is expected that many different relationships corresponding to different expertise's are needed to be modeled. Some of these relationships may involve different variables, some of them are quantitative, and some are subjective. In addition, due the dynamic economic and business environment, some of these variables may exhibit uncertainty and vagueness. Over and above such type of decision problem, which requires heterogeneous expertise's, often exhibits nonlinearity and ill-structuredness, which are difficult to manipulate with the use of conventional decision support system and analytical approaches. FESs are widely recognized for their capability to handle vague, inexact, and subjective inputs. FES is the only way to provide robust realistic solution to the decision problem, because without the quantification of vagueness, uncertainty, and subjectivity, the obtained decision solution will be inferior or unrealistic. Other analytical quantitative approaches, either stochastic or certain, do not have the capability to include qualitative or subjective input variables, the inclusion of them is necessary to provide a realistic solution. This is beside that FES provides a natural way to incorporate human expertise in form of If-then decision rules, based on and very close to the linguistic description of the human expert. Therefore, a FES was initially proposed as a solution to perform the custom declaration detection decisions. However, as it has been described previously that multiple different expertise's are needed to give more realistic and comprehensive decision answer, it is difficult to construct a large-scale expert system that includes all these heterogeneous expertise's, their relationships, and decision rules. One way to solve this dilemma is to construct multiple rule-based FESs, each of which corresponds to an expertise. These expertise's can be unique like: technical, legal, territorial,…, etc, or similar, but the views, skills, and way of thinking within each expertise may differ. The practical reasons behind the creation of independently separated FESs are: cohesion of knowledge units, control and final decision responsibility, avoidance of knowledge interaction or mutual influence, modularity in analyzing and explaining the final decision, sensitivity of aggregate knowledge, flexibility with the existence of context-based reasoning, improving maintainability, and consistency in handling relationships and reasoning. These reasons of separation will be described more in details within the introduction made in this thesis. Thus, the proposed solution of the practical problem described above is to integrate multiple FESs to suit the contextual requirements associated with judging custom declaration decision transactions. This problem is generalized into integration of multiple FESs for judging binary decision making problems, and which has become the subject of this thesis. The proposed way of integration is through combining or aggregating the decision outputs of these systems into a final, consolidated group decision. Another important issue that should be taken into account is that the integration should provide for satisfaction of some imposed restricting requirements. These requirements are divided into general and specific requirements. The general requirements are related to the way and information used to combine or aggregate the FESs, the format of combined or aggregated outputs, and the similarity or uniqueness of the participating FESs. The specific requirements are related to the roles and relationships among the FESs and special influences of the participating FESs. Some of these requirements are real requirements imposed by the aforementioned practical project, and some are elicited subjectively as possible requirements. These two sets of requirements will be described more in details in chapter 3.

Based on all above, this research is devoted to the problem of integrating multiple FESs under varying requirements for judging binary group decision making (GDM) problems. Hence, the key success factor is how effectively combine or aggregate the final judgments provided by individual FESs into a finally consolidated representative decision answer, "Yes" or "No".

The research objectives

The main aim of this research is to realize objective integration of multiple, relevant FESs, through adoption and development of combining or aggregating heuristics, criteria, methods, or models to arrive at consensus given their individual output decisions, and at the same time satisfying the imposed restricting requirements. Therefore, the research specific objectives are:-

A. Structuring the problem

1. Establishing a unified scale standardized for the output decisions produced by the individual FESs.

2. Structuring the integration problem according to the general requirements, and for possible decision making contexts.

3. Structuring the combination or aggregation approaches, methods, or models, used according to the general and specific requirements and for possible decision making contexts.

4. Formally stating the combination/aggregation problem.

B. Adopting and developing adequate combining/aggregating criteria, heuristics, methods, or models

1. Preparing necessary candidate decision aiding tools, and configuring them to problem requirements and specifying their roles.

2. Adopting and configuring the existing classical combining rules or algorithms to solve the combination and aggregation problems, given the crisp outputs based on the established unified numerical scale, and identifying their scope and characteristics.

3. Investigating the capabilities of these classical combining/aggregating criteria or methods to solve the problem, while satisfying various imposed requirements.

4. Developing more efficient combining/aggregating methods.

5. Testing and comparing criteria when applicable and identifying the superiors in terms of some performance measures.

C. Satisfying specific and general requirements

1. Developing new combining/aggregating criteria, heuristics, models, or methods to satisfy both specific and general requirements.

2. Identifying which criteria, heuristics, models, or methods satisfy or suitable to handle which requirement, and concluding on their relative strengths and contexts of their applicability.

Method of work

As the integration problem has been inspired by and emerged from a real need of a currently held practical project, mentioned earlier, the project requirements and context represents a basic guide in formulating the problem. The research problem is generalized for solving binary decision making problems through the integration of several context-relevant FESs. First, a decision regarding the adequate format of the output decisions of the individual FESs is investigated. Possible choices were whether these decision outputs should be in form of subjective decision answers, either “Yes” or “No”, or it should be in form of crisp numerical values. The final decision made is that every FES should provide its answer in form of crisp numerical values representing the degree of bias toward “Yes” or “No” decision options. This will permit the use of more sophisticated combination/aggregation methods working at a measurement level which uses more information about the degree of bias to every decision class rather than having abstract information about only which class is selected. The decision of output format will affect all the succeeding research effort, since the adoption and development of new combining/aggregating methods should be based on the output format they will manipulate. In order to realize the main aim of this research and the previously stated objectives, the research efforts are organized as follows:-

A. Generalizing and stating the research problem to a GDM problem evaluating a binary decision making or classification problem.

B. Gathering all general and specific requirements, mainly guided by and emerging from the need of the aforementioned current practical application, and subjectively eliciting other logically possible requirements.

C. Conducting an extensive computer search for related literature, taking into account the research context or surroundings of the proposed research point. This computer search is made on three levels or directions:-

- <u>First direction</u>: investigating the past research attempts in integrating expert systems, intelligent systems, and other decision support systems, pattern classifiers….etc., in an attempt to find analogy and to get benefits from this past experience, specially when they share similarity like in combining systems’ numerical outputs…, etc.

- <u>Second direction</u>: studying and surveying the GDM topic and adopting aggregation or combination methods utilized in this research field to be possibly exploited and configured for the proposed research.

- <u>Third direction</u>: surveying the artificial intelligence modeling tools and algorithm, which have wide applicability nowadays due to their capability to handle complex ill-structured decision making problems, and because they have high degree of adaptability to wide range of decision making contexts. This aim is to adopt some candidate tools that might match the requirements of the integration problem. Some of the tools that are candidate to be investigated are the Artificial Neural Networks (ANN), and the Hierarchical Fuzzy Systems (HFS). The reason behind this is that the ANN has the learning capability in handling past expertise’s historical data, which is one general requirement of the problem. In addition, there are many topologies and training algorithms of neural nets available, which may be helpful; this is beside the wide applicability of neural nets as classifiers. The HFS is investigated because it has a potential to be used to logically structure the combination/aggregation problem and to provide for separate combination/aggregation of mutually related systems. Other tools candidates to be investigated are k-nearest neighbors (k-NN) classifiers and regression analysis.

- Fourth direction: surveying the research attempts conducted in the field of pattern classification, specifically, the decision combination of multiple classifiers, which has close analogy to the proposed research. This aims to exploit the combination methods developed in that field, and to adopt and configure them for use in the integration problem.

- Fifth direction: studying and surveying the past literature concerning the decision aiding tools or methods used for weighting systems or experts, which is necessary notion to reflect the relative importance's of the participating FESs. The prominent Analytical Hierarchy Process (AHP) is mainly aimed, since it is logical, practical, simple and efficient method. Other different weighting methods are to be also investigated.

D. Adopting and matching the candidate techniques and approaches for combining/aggregating the FESs' outputs, and at the same time providing for the satisfaction of various requirements. Based on the comprehensive literature survey done within the research fields: GDM, pattern classification, artificial intelligence,…, etc., and guided by the requirements that need to be satisfied, the following actions are carried out:-

1. The integration problem is structured according to the basic requirements of the research problem. This will permit organizing the efforts toward solving the problem.
2. A research for match between the requirements of the integration problem and the already existing techniques is conducted, based on the research effort performed during the survey of literature. This should end with a list of candidate viable approaches
3. Structuring the candidate and necessary combining/aggregating approaches according to their roles in satisfying the requirements.
4. Selecting the most efficient in terms of performance quality, and computation cost, and effectiveness regarding requirements satisfaction.
5. Configuring and improving the existing combining/aggregating approaches.
6. Attempting to solve the bottleneck problems that confront the efficient application of existing combining/aggregating approaches.
7. Attempting to adapt and configure the artificial intelligence tools and decision aiding tools to manipulate the problem and satisfy the requirements.
8. Attempting to develop new combining/aggregating criteria and heuristic algorithms.
9. Attempting to integrate the adopted and developed solution approaches and methods to provide more reliable integration solutions.

Chapter 1

Introduction

This chapter is mainly intended to define the integration problem, position it relative to other forms and types of systems integration, and to identify the research topics that have relation to this work. Special focus will be devoted to the reasons for constructing multiple independent FESs, problem definition and the proposed configuration of FESs integration.

There are some topics that have close relation to this research work, and it is necessary to review and discuss their main features in relation to the integration problem. One important topic is the idea of integrating knowledge-based systems (KBSs) and decision support systems (DSSs). Another relevant topic is the fields of group decision making (GDM), and patterns recognition. GDM constitutes the basic foundation and major underpinning field of this research problem. This is because the idea of integrating multiple FESs through combining/aggregating their individual outputs is closely related and analogous to the notion of GDM, in which the opinions, preferences, or ranks of several decision makers are combined/aggregated to obtain a final group decision. The field of patterns recognition is a relevant topic because of the analogy that exists between the research efforts made in this field regarding combining multiple patterns classifiers, and the idea of combining the outputs of multiple FESs. This topic will be postponed to the next chapter, the literature review, and discussed only within the review. Another mother topic is the field of artificial intelligence (AI), especially with regard to the type of the integrated unit, which is a FES. It is important to grasp the components, nature, and information flows through a FES. Understanding the relationships between these topics and the proposed integration problem constitute a necessary step toward developing a successful, and objective integration solutions, and to avoid unrepresentative or inadequate ones.

In this chapter, the basic components of a FES, which is the unit of integration, will be briefly reviewed with special emphasis on the crisp nature of the system's output. Then, the main reasons of having separated FESs will be explained. The basic types of KBSs and DSSs integrations will be discussed. The position of integrating FESs relative to these types will be highlighted. After, the main characteristics and features of GDM, will be reviewed and discussed in relation to the problem of integrating multiple FESs. Finally, the possible configurations of integrating expert systems (ESs) will be explained, with special focus on the proposed configuration of the FESs integration.

1.1 Fuzzy expert system

The system consists of four components: a fuzzification subsystem, a knowledge-base, an inference mechanism, and a defuzzification subsystem which converts the implied fuzzy sets into crisp output value (see figure 1.1). The special concern around FESs is attributed to their wide applicability and use due to their capability to treat vagueness, uncertainty, and subjectivity. Especially important is the crisp nature of the output of the system, which is the format that will be manipulated when attempting to combine or aggregate the outputs of the multiple FESs. Through out this thesis, this crisp output should express in some way, which will be established later in a subsequent chapter, the degree of bias toward either "Yes" or "No" decisions. More detailed description of the fuzzy set theory and FES can be found in (Zadeh, 1965; Kilagiz et al., 2004).

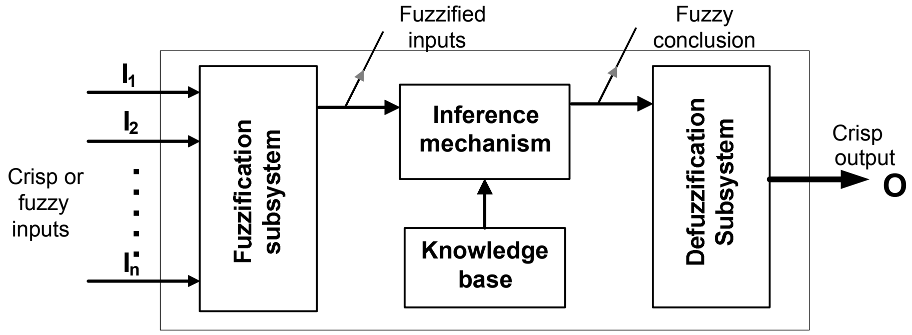

Fig. 1.1 The structure of a fuzzy expert system.

Next section, the practical reasons of constructing several independent FESs will be explained in details.

1.2 Why multiple fuzzy expert systems?

The need for multiple FESs can occur frequently when a complex problem in hand has multiple relevant aspects for which the existence of multiple independent and separated expertise's is necessary, and there is no available expertise that covers the whole aspects of the problem. Barret and Edwards in 1995 (Barret & Edwards, 1995) stated that when experts have different areas of expertise, each one is captured in a separate knowledge base. In addition, the crucial nature of the type of made decisions such as the case in binary decision making (Yes-or-No) usually requires several expertise's in the same field, each of which has its own skills and way of thinking necessary to cope with the decision making complexity and to improve reliability. There are still some other reasons behind independence amongst ESs, in general, related to the practical implementation and efficiency of operation. These reasons can be:-

- **Cohesion of knowledge units:**

Specific type of knowledge and skills associated with the modeled expertise's can be considered more homogenous and cohesive if represented separately in its single, own ES. This knowledge includes the view of the modeled expertise about how to approach the problem and solve it. Cohesive sets of variables utilized and linguistic scales used contribute to the clarity and controllability of the constructed ES. Maintaining cohesive chunks of knowledge also promote relevancy, responsibility, modularity, and facilitate maintainability of the modeled system. The cohesive ES is easier to build, which contribute to improving its performance and reliability.

- **Control and final decision responsibility:**

When every ES clearly corresponds to specific expertise, it becomes easier to analyze, understand, and control its performance and operation than if we have one huge ES that involves several expertise's the interaction among them can obscure the analyzability and the control of operation. In addition, it is easier to attribute the decision responsibility to every ES, and if past historical performance data are available, then the relatedness and responsibility of each individual decision can be more deeply investigated. Also, independent ESs, each of which provides its explanation how and why a particular decision is made can improve decision analyzability and responsibility. Decision responsibility provides important feedback information, which in case of independence among ESs can be easier to understand

and contribute to continuous adjustment and improvement of performance. When we have a single huge ES, it becomes cumbersome to test its performance, attribute the decision answer to specific reason or knowledge, and difficult to analyze. The responsibility of decision answer can be obscured in case of huge ES including several interacting expertise's.

- **Avoidance of knowledge interaction or mutual influence:**

When multiple expertise's are modeled in one huge ES, it becomes more likely that many different knowledge chunks contained in the large rule base affects the implication of each other, resulting in either compromising solution or distortion of some decisive conclusions. Particularly important is the decisiveness of the decision output, especially in case of Yes-or-No type decision problems. Compromising solution resulting from rules interaction is not desired, since they are not decisive. Whereas, in case of independent and cohesive expert systems, relatively fewer related rules are combined and fired together, and this eliminates the tendency of interaction between different knowledge chunks and production of compromising or distorted decisions.

- **Modularity in analyzing and explaining the final decision:**

Huge ES is difficult to analyze, due to vast amount of relationships and variables associated with multiple expertise's modeled. The explanation information provided by a large-scale ES is less valuable because it is difficult to follow. Many rules are involved and related to each other, the interpretation path may be cumbersome or complex, and it may exceed human capability to grasp. It is clear that the problem is lessened with more compact independent ESs.

- **Sensitivity of the aggregate knowledge:**

In some circumstances, the knowledge modeled, in form of decision rules and relationships among input and output variables, may be considered sensitive which necessitates keeping it hidden from acquaintance of other ESs or operators. This sensitive knowledge can found in particular business fields where the expertise and way of thinking should be kept away from customers or analysts of other systems, because the aggregate knowledge and way of thinking can be misused, for instance, by customers or irresponsible employees or analysts. Example of this case is the problem of currently held project of designing intelligent system of state Custom Administration in Czech Republic. The purpose of the system is to detect suspicious customer declaration transactions provided by commercial exporters and importers. The intelligent system should handle multiple specialized expertise's (eg., legal, technical, commercial...,etc.). Due to the sensitivity of the aggregate knowledge and also of the individual expertise's, it is necessary to keep such expertise's independently separated and kept hidden inside individual ESs from other systems. Similarly, in the military applications, the importance and nature of the knowledge and expertise is typically very sensitive, and the need to build separate, knowledge-hidden ESs can frequently occur.

- **Flexibility with the existence of context-based reasoning:**

Under certain circumstances, when there are a variety of decision making transactions, the set of available ESs may not be all relevant to all decision problems. Also, for each decision transaction it may happen that different relevant expertise's are needed. Then, the existence of independent ESs that cover all possible expertise's needed in all decision making contexts, provides a great flexibility to the performance of the intelligent system as a whole. It is clear that this proposed independence of the multiple systems is highly flexible to handle different possible decision making transactions with minimum or zero disturbance in individual systems, and in the whole system. Of course, this is not the case with a huge system with large knowledge base, where flexibility as well as performance efficiency are degraded, because much disturbance is required in order to search the relevant relationships and rules for every transaction with accumulated time loss. This could be severe problem in the case of dynamic business environments, where there are always high volume of processed transactions per unit time, frequently changing features of decision transactions, and a frequently needed system modifications.

- **Improving maintainability:**

Ease of maintenance is an important characteristic of a dynamically efficient and adaptive ES. This is in order to keep following the updates frequently occur in the business world. Maintaining and controlling huge ESs containing thousands of decision rules is a cumbersome and complex task, where many interrelated inference paths should be adjusted. In contrast in the case of more compact and separated ES, maintainability is improved, due to fewer number of related rules needs to be adjusted, and inference paths are more smooth and easy to follow.

- **Consistency in handling relationships and reasoning:**

When the variables and relationships associated with the modeled expertise are homogenous, in that they belongs to similar pieces of knowledge, reasoning process becomes more straightforward and consistent due to the increased relatedness among knowledge chunks. Consequently, decision of every individual ES is more likely to be more consistent, than it can be when obscured relationships are not eliminated. This reasoning consistency improves traceability of the rule bases and facilitates its construction. Also, consistency contributes to the clarity and reliability of the constructed systems, and to the understandability and analyzability of their individual decisions as well.

- **Improving performance of individual FESs:**

In a huge rule-based expert system, the increased number of rules can quickly overload the memory and makes the application difficult to implement, whereas in a more compact and separated FES, the performance of individual systems is improved.

Next section, the idea of integrating FESs, the main reasons behind integration, and the general possible ways of systems integration will be discussed.

1.3 What is FESs integration? How and why integrate?

The proposed integration is a novel idea toward objectively integrating multiple FESs through combining or aggregating their crisp outputs. These integrated FESs are arranged synergistically to suit a particular need of a given decision making context. The problem can be also defined as complementing together the most relevant FESs to match the problem contextual requirements, and to obtain a finally consolidated decision representative of the individual decision outputs, as a result of combining or aggregating these individual decisions. Generally, there are two types of DSSs integration defined by George Marakas (Marakas, 2003):

(1) Functional integration: in which various decision support functions are integrated and linked to those of the existing infrastructure.
(2) Physical integration: involves the architectural bundling of the hardware, software, and data communication characteristics associated with the modern DSSs into the existing set of physical systems.

Also, the integration can be both functional and physical at the same time. For instance, in multi-knowledge sources integration, the data bases for all sources can be pooled together in a common data base. At the same time the main purpose of the integration is to provide for different decision support functions provided by each knowledge source.

The integration proposed in this thesis belongs to the functional type of integration, in which several decision support functions are provided by corresponding independent, separated FESs. This is accomplished through combining the crisp outputs of the multiple FESs into a finally consolidated decision.

Further, functional integration can be in form of two levels: Across different management support systems (MSSs) or within a MSS ((Turban, 2001) used the term Management Support System (MSS) to refer to the application of any technology (e.g., DSS, ANN, ES,..., etc.), either as an independent tool or in combination with other information technologies, to support a management task in general and decision making in particular). The integration of different MSSs is called the first level of integration, whereas the integration within a MSS is called the second level or global integration. Combining several MSSs, each addressing a specific class of a decision problem is an example of the first level of integration. The second level refers to the integration of appropriate MSS technologies in building a specific integration especially to take the advantage of the strengths of specific MSS technologies. For instance, the ANN can be used for pattern recognition as part of the intelligence phase of the decision-making process, and an ES can be used to provide a solution to the problem. The proposed integration of multiple FESs belongs to the first level defined for integration, since FES is a one type of MSS.

Other classification of integration types is that integration can be of two forms:
- Integration of different types of systems, such as integrating ESs and DSSs together.
- Integration of systems of the same type, such as integrating multiple FESs.

Two general major objectives for system integration (Turban, 2001):
- **Enhancement**: one system enhances the function of another different one or ones.
- **Increasing the capabilities of the whole applications**: in order to increase the capability of the whole system, each system performs the subtasks at which it is the best.

In the next section, the GDM topic, which constitutes the basic foundation of the proposed integration problem, will be discussed in relation to the common and applicable features to this research work.

1.4 GDM: the foundation

As the integration of multiple FESs involves the cooperation between multiple decision making systems in solving a given decision problem, it is useful to make reference to the field of GDM. This is in order to position the integration problem correctly within such field, which constitutes the important foundation, and in order to succeed in dealing with the analogous integration problem. It is also important in order to be able to exploit and tailor the group decision techniques of the GDM for combining/aggregating the decision outputs of the multiple FESs.

DeSanctis and Gallupe in1987 (DeSanctis & Gallupe, 1987) highlighted a reason for the need of GDM that may be the problem is too significant for any single individual. GDM is among the most important and frequently encountered processes within companies and organizations in both public and private sectors. Bui and Jarke in 1986 (Bui & Jarke, 1986), defined co-operative GDM as a problem solving process in which (i) there are two or more persons, each characterized by his or her own perceptions, attitudes, and personalities, (ii) who have recognized the existence of a common problem and (iii) attempt to use a group decision support system (GDSS) to reach at a collective decision.

The majority of real world decision making problems involve multiple decision makers (Turban, 1988). Choi et al. in 1994 (Choi et al., 1994) stated that most group problems are complex and unstructured and are difficult to solve. They pointed out four properties of group decision problems which render them hard to attack:

(1) They are social problems not mathematical or scientific ones,
(2) They are difficult to satisfy all constraints and requirements,
(3) They are more difficult to model than single problems,
(4) There are few methodologies to verify fairness, a concept that is closely related to the aggregation of preferences.

Many of the decisions in today's workplace are made by groups of individuals. Groups bring several advantages to the choice process: the addition of multiple sources of knowledge and experience, a wider variety of perspectives, and potential synergy associated with collaborative activity. They also bring with them several limitations that, when ignored, can result in decision outcomes ranging from problematic to catastrophic. GDM term was generalized and replaced by Holsapple in 1991 (Holsapple, 1991) with the term multi-participant decision making (MDM). MDM is defined to be an activity, a decision making process, conducted by a collective entity composed of two or more individuals and characterized in terms of both the properties of the collective entity and of its individual members (Marakas, 2003). Figure 1.2 illustrates MDM's possible structures.

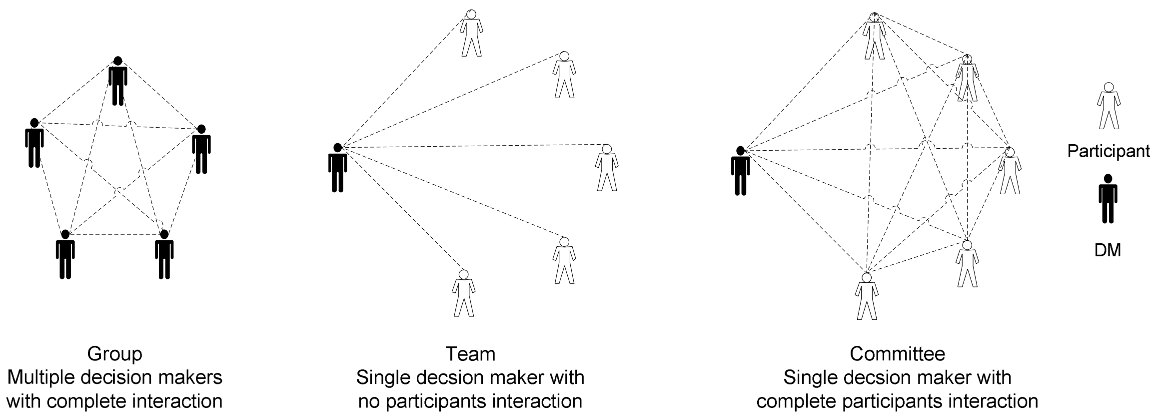

Fig. 1.2 Basic MDM structures.

The proposed type of integrating multiple FESs has a little different decision making aspect from the above described MDM structures in that, it is a group, a collaborative group of decision makers, but with no interaction. Every decision maker, which is an individual FES, provides its decision output; then all decision outputs are collected, and the problem is to combine or aggregate them into a final collective group decision.

In the next section, the possible configurations of multiple ESs integration will be reviewed and described, and the proposed configuration of the integrated FESs will be highlighted.

1.5 The configurations of the multiple expert systems integration

There are two main possible configurations that can connect together several ESs: series and parallel configurations (Beeri & Spiegler, 1996). Hybrid configurations of series and parallel could be then constructed. The two main configurations are described as follows:-

(a) Multiple series expert systems:
ESs are arranged in such a way that the input to the problem is presented to the first ES, which passes its output decision to its successor, and so on as in figure 1.3.

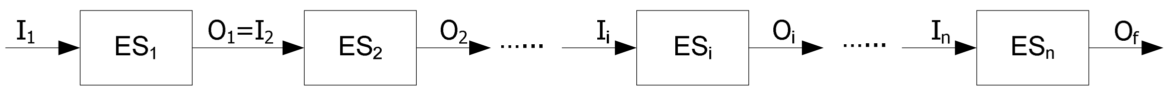

Fig. 1.3 Multiple series expert systems.

(b) Multiple parallel expert systems:
ESs are arranged in such a way that a same problem is presented to every ES concurrently in order to reach at one consolidated output decision (see figure 1.4). Integrating several ESs in parallel can be viewed as a form of GDSS. Further, parallel ESs can be arranged in hierarchical layers, where the outputs of sub-groups within each layer can be successively combined in parallel until obtaining the final decision (see figure 1.5). The proposed integration of multiple FESs follows this parallel configuration of integration (see figure 1.6). The combiner/aggregator could be a simple mathematical formula, a heuristic algorithm, a model…, etc.

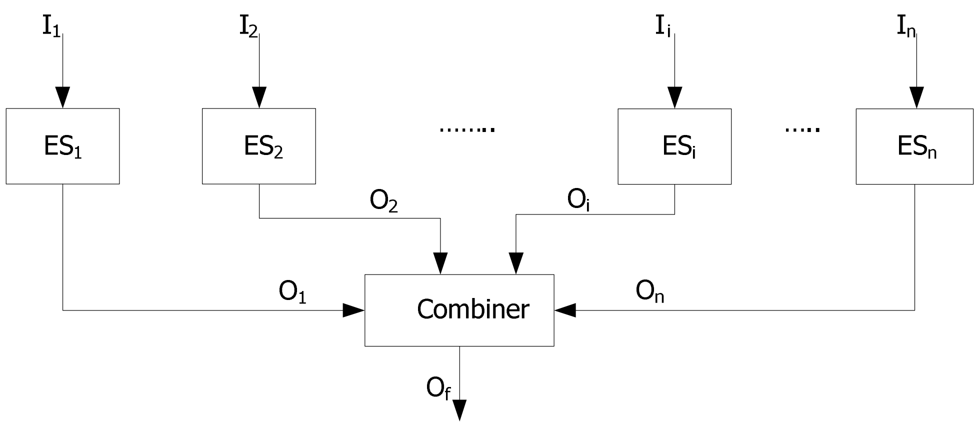

Fig. 1.4 Multiple parallel expert systems.

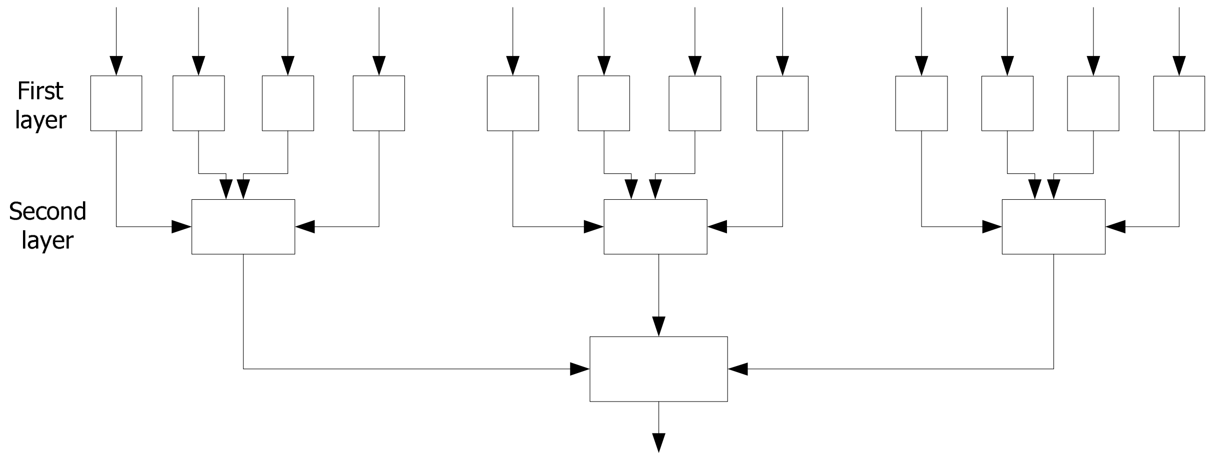

Fig. 1.5 multi-layer structure of parallel expert systems.

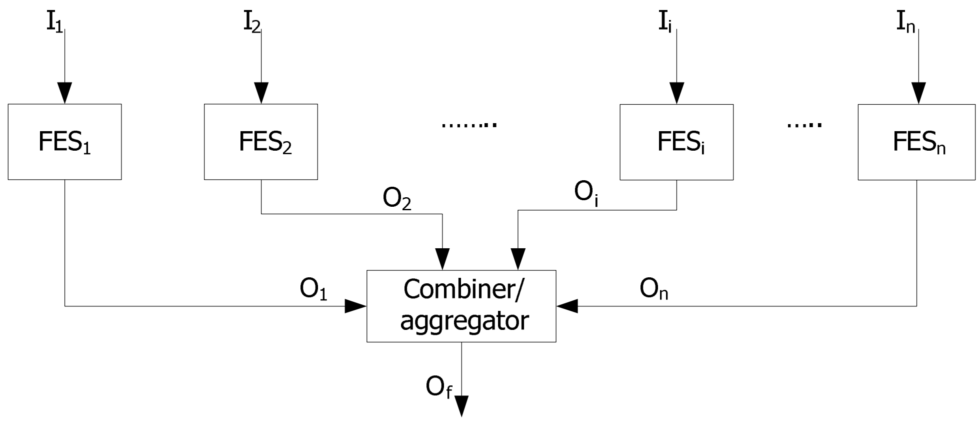

Fig. 1.6 The proposed configuration of the multiple
fuzzy expert systems integration.

At this point the definition, justification, research boundaries, and configuration have been specified. In summary, the proposed integration of multiple parallel FESs belongs to the GDM problems, and involves combining or aggregating the final crisp output of the individual FESs to reach at a final conclusion in judging a binary decision problem.

Next chapter will review and investigate the past research attempts made within all relevant research topics identified in this chapter in relation to the defined integration problem and its solution. This aims to understand the foundation of this research work necessary for success, and to get benefits from the solution approaches developed in these relevant fields, which could be adopted, configured, or improved to suit the particular needs of the proposed FESs integration.

Chapter 2

Literature review

In chapter 1, the basic research boundaries of the FESs integration problem have been identified. Four basic relevant research topics were highlighted: systems integration, GDM, pattern recognition, and AI. The field of systems integration constitutes the mother discipline of this research work, and involves various ideas regarding types and roles of the integrated systems and the existing techniques used to realize integration. The second relevant field is the GDM that represents the foundation of combining and aggregating the judgments of multiple decision makers. This is because the proposed idea of integrating FESs is to be realized through combining or aggregating the judgments of multiple decision making systems, which are the FESs. This field involves the preference aggregation and combination techniques for the group decisions, which could be adopted and configured to solve the proposed integration problem. Other relevant research topics are the fields of pattern recognition, and AI. Pattern recognition field contains the research works regarding combining multiple pattern classifiers, and the use of neural classifiers especially for two-groups or binary classification problems. This is analogous to the idea of combining the outputs of multiple FESs evaluating a binary decision problem. The field of AI provides the basic background knowledge about the components and the information flow through the units of integration, and contains a variety of tools like ANNs, and fuzzy systems that could be exploited to develop a solution for the integration problem. Some of these candidate tools are the neural classifiers, and the hierarchical fuzzy systems (HFSs). These AI tools will find their roles in the subsequent chapters of this thesis.

Because of the novelty of the proposed integration problem, and that the solution of the problem is touching many disciplines, especially those mentioned above, an extensive survey of literature has been conducted. Over one thousand research articles have been investigated concerning the relevant past research attempts conducted in the described research boundaries, and in relation to the objectives of this integration problem. This was in order to grasp all the research aspects of the problem, and to get acquainted with the candidate solution ideas and approaches.

Next section, the recorded results of the surveys concerning systems integration will be presented. A critical assessment of the conducted researches will follow the recorded literature in every section of the review.

2.1 Survey of past research attempts in integrating systems

It is important to study other forms of integration that previously have been proposed. This is first in order to be able to differentiate between them and the proposed integration of multiple FESs, and second for the purpose of getting benefits and experience from past efforts in form of applicable integration techniques and solutions, or new integration mechanisms and ideas. The results of this survey is presented and discussed below.

- In 1980, Erman et al. (Erman et al., 1980) proposed blackboard architecture for integrating knowledge of multiple knowledge sources. They stated that knowledge integration aims to remove uncertainty.

- In 1990, Engel et al. (Engel et al., 1990) used a blackboard-based tool named MKSMART, an acronym for Multiple Knowledge Source Management And Reasoning Tool, to integrate ESs with conventional problem-solving techniques.

- In 1990, Sullivian and Fordyce (Sullivian & Fordyce, 1990) analyzed a system called logistics management system that was developed by IBM for operations management. The system combines ESs, simulation, and DSSs. In addition, it includes computer aided manufacturing and distributed data processing subsystems. It provides plant manufacturing management with a tool to help in resolving crises and in planning.

- In 1990, Ritchie (Ritchie, 1990) suggested a novel AI-based solution approach to the problem of providing operator decision support in integrated freeway and arterial traffic management systems. A conceptual design is presented that is based on multiple real-time KBS integrated by a distributed blackboard problem solving architecture. He presented the conceptual design for the proposed distributed blackboard architecture.

- In 1992 and1993, several models have been proposed for integrating ESs and DSSs by (Watkins et al., 1992; Goldbogen and Howe, 1993; Van Weelderen and Sol, 1993).

- In 1994, Mitra and Dutta (Mitra & Dutta, 1994) presented an architecture which integrates several optimization models and human expertise to arrive at a complete solution for a complex problem. A blackboard and an associated truth maintenance system form the basis of computer-based support for the proposed architecture and the associated interactions. The architecture of the design tool is independent of the application domain.

- In 1996, Azzam and Nour (Azzam & Nour, 1996) proposed a decentralized ES consisting of multiple local ESs. Each local ES is concerned with a part of the global system. Such ES is useful in dealing with modern large scale interconnected power systems, which are characterized by large computational burdens.

- In 1996, Belz and Merten (Belz & Merten, 1996) described a combination of several ESs implemented as intelligent agents, with a scheduling systems and a simulation based DSS for rescheduling production lines when problems occur.

- In 1996, Beeri and Spiegler (Beeri & Spiegler, 1996) presented a model for integrating multiple ESs. The model- Synergetic Expert Systems (SES) - contains several ESs, which can be arranged synergistically to suit the particular needs of a problem. An object-oriented approach is used to design the model and to handle its various components. A formal definition and delineation of efficient and economic ES is made. Authors stated that the model may be applied when different experts or ESs are needed to tackle a complex problem.

- In 1997, Atanackovic et al. (Atanackovic et al., 1997) presented an integrated knowledge-based model to support various activities associated with the planning and design of electric power systems integrating several major ESs, and simulation tools.

- In 1997, Gams et al. (Gams et al., 1997) presented a general schema of integrating multiple reasoning systems, along with its implementation for emulsion quality control in a certain type of rolling mill. Integration is performed on the basis of training examples or hypothetical examples. Authors stated that multiple reasoning is often performed on the basis of predefined functions that combine proposals from single systems. The experiments conducted showed that the use of multiple reasoning systems is in general beneficial whenever an appropriate combination is found. Authors concluded that the proposed schema enables integration of an arbitrary number of systems as black boxes into a transparent knowledge base.

- In 1998, Ohtsuki et al. (Ohtsuki et al., 1998) developed an intelligent control system for wastewater treatment processes. In this control system, multiple software agents that model the target system using there own modeling method collaborate by using data stored in an abstracted database named “blackboard”. The software agents, which are called expert modules, include a FES, a fuzzy controller, a theoretical activated sludge model, and evaluator for raw data. The difficulty of controlling an activated sludge system by single conventional strategy was briefly reviewed, and then an approach to overcome these difficulties by using multiple modeling methods in the framework of an “intelligent control system” was proposed.

- In 1999, Li and Love (Li & Love, 1999) presented an integration of ES and ANN for estimating a contractor’s mark-up percentage in the construction industry.

- In 2001, Chi et al. (Chi et al., 2001) proposed a blackboard-based architecture that integrates several ESs applied to the blanking technology. It consists of a set of independent domain-specific modules, blanking technology knowledge sources or ESs, which interact via a shared global-data structure – the blackboard that organizes and stores the intermediate problem solving data. Knowledge sources produce changes to the blackboard that lead incrementally to a solution of a blanking technology problem.

Critical assessment

The previously conducted researches have addressed the integration problem from different views. The most extensively used mean of realizing integration is the blackboard architecture, attributed to its flexible control structure and modularity. However, one limitation of such architecture is that it does not specify how a specific piece of knowledge will have to be handled by other knowledge sources. Other form of integration considered is the use of multiple ESs or in general KBSs to provide multiple different decision support functions. Some forms of integration concentrated on the modeling aspect in terms of development technology or software. A little of past researches in the literature have considered the cooperation of multiple knowledge sources in providing for one decision support function. Also, few numbers of researches have concentrated about the form of data structure to be handled by different knowledge sources. Also, no or very few numbers of them have considered the idea of integration through combination. All the previously mentioned types of integration are different from the proposed integration in this research study; the proposed integration incorporates multiple FESs arranged in parallel, which are to produce a unified form of output decisions, and that the integration is done through combining their final output decisions, not through sharing a common knowledge base or cooperating in changing a common data base as in blackboard systems.

It should be noted that the conducted extensive literature review up to date has not revealed any previous research attempts that are directed specifically toward integrating multiple FESs, through combining their decision outputs; this research is a novel.

Next section, the recorded results of the surveys concerning GDM, the tools used in GDM, and group decision techniques will be presented.

2.2 GDM literature

GDM involves aggregation or combination of different individual preferences or judgments into a single collective one. This subject has received a great deal of attention from researchers in many disciplines (Ramanathan & Ganesh, 1994). Much of the researches have focused upon the area of social choice, and the early work of Arrow (Arrow, 1951) has been a major influence in this area. Much of the research that followed Arrow's work mostly concentrated on Utility Theory as a tool for GDM (Sen, 1970; Wendell, 1980). A comprehensive list of existing GDM methodologies and synthesizing mechanisms can be found in (Hwang & Lin, 1987). Two important subjects within the topic of GDM, with which the proposed integration problem has close relation and interest. They are: preference aggregation/combination, and the analytical hierarchy process (AHP) (Saaty, 1980). Some of the aggregation/combination techniques utilized in GDM could be adopted and configured to combine/aggregate the FESs' outputs. The AHP is one of the most prominent decision aiding tools that will have great role in weighting the importance's of the FESs in this study. This is beside its extensive and successful use in GDM. The results of the survey regarding the two subjects is presented and discussed below.

2.2.1 Preferences aggregation/combination in GDM

In group preference aggregation, a synthesizing mechanism is used to derive a collective decision, representative in some way to the individual opinions. Because the aggregation or combination of individual judgments or preference constitutes a major success factor in integrating the multiple FESs, special focus has been devoted to this subject. This is in order to get benefit from the previously developed techniques in this subject, which could be adopted, configured, or improved to provide an adequate solution for the proposed integration problem.

The aggregation of preferences has been widely studied by researchers, specially the aggregation of preference rankings and preference scores. The aggregation of preference rankings has wide applications in GDM, social choice, committee election and voting systems, and a large amount of researches has already been conducted in this area (Wang et al., 2005). How to aggregate individual preferences or a set of ordinal rankings into a group preference or consensus ranking is a typical GDM problem. A chronologically sorted review of aggregation literature is provided below.

- In 1784, Borda (Borda, 1784) was the first to examine the ordinal ranking problem for choosing candidates in an election and proposed a method of marks to rank candidates according to the sum of ranks assigned by voters to each candidate. In 1962, Kendall (Kendall, 1962) was the first to study the problem in a statistical framework. He proposed a ranking solution which was equivalent to Borda's method of marks, so the method is frequently referred to as Borda-Kendall (BK) method, which is probably the most widely used technique in determining a consensus ranking due to its computational simplicity.

- In 1962, Kemeny and Snell (Kemeny & Snell, 1962) spread headed a new trust in the area of group consensus formation by proposing “distance” measure between individual rank orderings. They have proposed a set of axioms to be satisfied by such a measure, and then proved its existence and uniqueness.

- In 1982, Cook and Seiford (Cook & Seiford, 1982) showed that in the presence of ties, Borda-Kendall method could not perform well as claimed.

- In 1983, Beck and Lin (Beck & Lin, 1983) developed a procedure called the maximize agreement heuristic (MAH) to arrive at consensus ranking that maximizes agreement among decision makers (DMs) or voters.

- In 1986, Korhonen et al. (Korhonen et al., 1986) described a computer interactive mathematical programming technique for solving group decision problems, once the utility functions of individual members have been specified.

- In 1987, Bui (Bui, 1987) presented Co-oP, a cooperative multiple criteria GDM system, one of the most well known and best-documented implementations within the multiple criteria decision makers’ context.

- In 1989, Lewandowski (Lewandowski, 1989) described the decision theoretic frame work of Selection Committee Decision Analysis and Support (SCDAS), a Group Decision Support System (GDSS) for selecting the best alternative from a given, finite set of alternatives.

- In 1992, Tapia and Murtagh (Tapia & Murtagh, 1992) presented an algorithm for solving a decision situation involving many decision makers who are all concerned with determining a compromise solution to a given multi-objective problem. Fuzzy programming enables the decision makers to vary, at any given iteration in the execution of a computer programme, their fuzzy aspiration levels in terms of input information known as preference criteria and underachievement tolerance values. A binary search technique is applied to the set of aspiration levels until a feasible efficient and acceptable compromise solution is obtained.

- In 1995, Salo (Salo, 1995) developed an interactive approach for the aggregation of group members’ preference judgments in the context of an evolving value representation.

- In 1996, Myung et al. (Myung et al., 1996) described an approach for aggregating the opinions of a group of experts that accounts for differences in competence among the experts and dependencies between their opinions. The authors derived aggregation rules for combining two or more expert predictions into a single aggregated prediction using Shannon’s information measure.

- In 1996, Cesa-Bianchi et al. (Cesa-Bianchi et al., 1996) examined the problem of deterministically predicting Boolean values by combining the Boolean predictions of several experts.

- In 1996, Bryson (Bryson, 1996) considered the GDM problem in which every decision maker provides his opinion about a given set of decision alternatives or objects utilizing the AHP to obtain a preference vector or weight vector containing the weights of AHP ranking. Given such preference or weight vector of each decision maker, Bryson proposed a framework for assessing the current level of group consensus, and described a decision procedure for consensus building.

- In 1996, Bryson together with Ngwenyama et al. (Ngwenyama et al., 1996) proposed three consensus-analysis indicators to evaluate the level of agreement in the group of decisions, and another three individual indicators related to the measure of the position of each individual relative to other group decisions.

- In 1998, Csaki et al.(Csaki et al.,1998) described a method, based on criterion trees, for decomposing a group decision model with decision tables. They proposed the aggregation of weights and scores into a group weight system and group score matrix, expressing this way the whole group's judgments on any single score or weight.

- In 1999, González-Pachón and Romero (González-Pachón & Romero,1999) presented a goal programming (GP) to aggregate ordinal rankings.

- In 2000, Major and Ragsdale (Major & Ragsdale, 2000) considered a general problem of combining the classifications or prediction of a number of local information systems into a single system. They introduced a new approach for solving an aggregation problem where a decision maker needs to classify an observation based on a group membership predictions coming from multiple experts. They empirically tested the proposed approach along with two alternative approaches found in the literature using real world data and four popular classification techniques. The results showed that the proposed approach produced the best performance according to certain performance measures. They concluded that researchers may use this method to create a knowledge base of information for a distributed ES or one that acquires prediction information from the Internet.

- In 2001, Matsatsinis et al. (Matsatsinis et al., 2001) reviewed some of the past approaches in the multi-criterion GDM. They pointed out that an examination of the literature, which is neither exhaustive nor complete, reveals that group decision-making and negotiation problems constitute a challenging area for multi-criteria decision scientists.

- In 2001, Mostaghimi (Mostaghimi, 2001) used a Bayesian estimation methodology for combining experts' information with the decision maker's prior. An information collection process is designed by setting constraints on this model. It was shown that the information produced in the process of producing a decision can also give insights into the impacts of the decision maker.

- In 2002, Hurley and Lior (Hurley and Lior, 2002) studied methods for aggregating expert rank-orders into an overall group rank order. They pointed out that there is an immense literature on the way expert opinions can be aggregated. A part of this literature considers aggregation based on a simple average (Hogarth, 1978; Armstrong, 1985; Ashton, 1986). They compared the performance of trimmed mean rank-order aggregation procedures in the case where a subset of individuals in the group charged with the decision vote strategically. Finally they employed a Mont Carlo simulation experiment on a specific decision instance and found that trimmed mean aggregation compares favorably with other procedures.

- In 2004, Cheng (Cheng, 2004) proposed a novel method to derive the collective opinion of a group of members as expressed in a grading process in which individual members evaluate objects or events by assigning numerical scores. The collective opinions are represented using triangular fuzzy numbers whose construction is based on the possibility distribution of the grading process. The usefulness of the proposed approach is demonstrated in a GDM problem involving multiple evaluation criteria. The results demonstrate that the fuzzy number construction method provides a better representation of the group preference than traditional methods.

- In 2004 Shih et al. (Shih et al. 2004) proposed a geometric-distance measures based consensus indicators. He used these indicators as guides to reach at consensus.

- In 2004, Hwang (Hwang, 2004) utilized the fuzzy set priority method to aggregate the ranking of multiple evaluators of multiple decision alternatives.

- In 2005, Ölcer and Odabai (Ölcer & Odabai, 2005) proposed a new fuzzy multiple attribute decision making (FMADM) method, which is suitable for multiple attributive GDM problem in a fuzzy environment. This method was proposed to deal with the problem of ranking and selection of alternatives. In the proposed approach, an attribute based aggregation technique for heterogeneous group of experts is employed and used for dealing with fuzzy opinion aggregation for the subjective attributes of the decision problem. The propulsion/maneuvering system selection as a real case study is used to demonstrate the versatility and potential of the proposed method for solving fuzzy multiple attributive GDM problems.

- In 2005, Wang et al. (Wang et al., 2005) developed a preference aggregation method for ranking alternative courses of actions by combining preference rankings of alternatives given on individual criteria or by decision makers. In the method, preference rankings were viewed as constraints on alternative utilities, which were normalized, and linear programming models are constructed to estimate utility intervals, which were weighted and averaged to generate an aggregated utility interval.

<u>Critical assessment</u>

The survey on aggregation has indicated that most of the researches have been devoted to the aggregation at the rank level, where a group of decision makers provide their rank preferences for several decision alternatives. Borda count was the most widely utilized in rank aggregation. At the abstract-level combination of preferences, where the preference is made as a selection of one preferred alternative, the majority voting and weighted majority voting, were the most widely used. At the measurement level, where the scores or weights are used to indicate priorities of alternatives, the arithmetic mean or simple average aggregation had the wide applicability. However, at any level there has not been any method to select or to character different combination or aggregation criteria. For instance, at the measurement level, no mean has been developed to assist in selecting between the arithmetic mean and the geometric mean criteria. On the other hand, there were no mean to decide whether or not to include weights of decisions of different experts or decision makers. Other forms of aggregation were based on Bayesian probabilistic approaches, and require estimation of prior probabilities, which considered one limitation; this is beside the uncertainty and risk implied with the use of such approaches. Mathematical programming based on a utility function had a considerable portion of past researches in GDM. However, it suffers from complexity in problem formulation and expression of qualitative requirements such as the dependencies between decision makers' judgments. This is besides other associated assumptions that can make the final decision solution either unreliable or unrealistic. Literature considering analyzing consensus exhibits few numbers of researches, although it constitutes a strong step toward getting more insight and information about individual decisions or judgments. Consensus analysis was based mostly on two measures: distance and angle measures of similarity among individual preference vectors. The main purpose of consensus analysis was to facilitate reaching consensus among various individual decisions, as another form of aggregation. In overall, all the above discussed applicability characteristics, limitations, and advantages of previously conducted research efforts will be considered when developing solution approaches for the proposed FESs combination/aggregation problem.

Next section, the literature regarding the use of AHP in GDM and as a decision aiding tool will be reviewed.

2.2.2 AHP

Several successful researches have been conducted regarding the utilization of AHP in GDM, and as a stand-alone decision aiding tool. Some of these prominent researches are reviewed hereinafter.

- In 1983, Hannan (Hannan, 1983) applied the AHP to evaluate contestants in a group decision problem.

- In 1984, Arrington et al. (Arrington et al., 1984) applied AHP in GDM problem, through using the simple arithmetic mean with equal weightings for all the group members to arrive at the group consensus.

- In 1985, Saaty and Kearns (Saaty & Kearns, 1985) formally proposed AHP for the GDM.

- In 1986, Lockett et al. (Lockett et al., 1986) used AHP to model research portfolio within the research and development (R&D) management discipline.

- In 1988, Khorramshahgol and Moustakis (Khorramshahgol & Moustakis, 1988) developed a hybrid methodology, based on Delphi method and AHP for setting priorities.

- In 1992, Dyer and Forman (Dyer & Forman, 1992) explained why AHP is so well suited to GDM and showed how AHP can be applied in a variety of group decision contexts. They argued that the AHP offers numerous benefits as a synthesizing mechanism in group decisions. They described four ways that AHP can be applied to

GDM: (1) Consensus, (2) Voting or compromising, (3) Forming the geometric mean of individual judgments, and (4) Combining results from individual models or parts of a model. Finally, they discussed four applications of AHP in group decision situations.

- In 1992, Carlsson et al. (Carlsson et al., 1992) described a system for formalizing consensus reaching within a set of decision makers trying to find and agree upon a mutual decision. The system uses the AHP in order to model the preferences of each decision maker.

- In 1994, Choi et al. (Choi et al., 1994) discussed the applicability and practicality of the AHP in GDM for a new provincial seat selection in South Korea. Their experience indicated that AHP forms a strong group support tool because of its fair and rational hierarchy contributes to understanding the group problem and reducing the gap between conflicting groups. Authors pointed out that the case study provided in their research confirmed the applicability and practicability of the AHP in GDM.

- In 1994, Ramanathan and Ganesh (Ramanathan & Ganesh, 1994) conducted an evaluation using the well established social axioms. They compared two AHP's group preference aggregation methods: the Geometric Mean Method (GMM) and the Weighted Arithmetic Mean Method (WAMM). They used counter-examples that the GMM does not always satisfy the Pareto Optimality axiom, which is one of the prominent and widely accepted social axioms. The authors finding was significant as the GMM has been the most commonly used method in AHP for combining individual opinions to form a group opinion. The other method WAMM has satisfied all the axioms, except "the independence of irrelevant alternatives" axiom. In order to use the WAMM, the weightings (importance's) to be assigned to the group's members must be stated. They proposed a simple eigenvector based method to intrinsically determine the weightings for group members using their own subjective opinions. Finally, they brought out the superiority of the proposed method over the previous.

- In 1997, Barzilai and lootsma (Barzilai & lootsma, 1997) applied the Multiplicative AHP, a variant of the original AHP, to arrive at a joint decision, by incorporating the relative power of group members.

- In 2005, Beynon, (Beynon et al. 2000) developed nascent DS/AHP method of multi-criteria decision making aggregation for GDM with non-equivalent importance of individuals in the group. A discount rate value was defined for each member of the group depending on their perceived individual levels of importance. This discount rate attenuates the evidence from an individual by re-assigning more value to their concomitant level of ignorance. The adjusted evidence from each group member was then combined to derive the group's collective decision.

Critical assessment

The survey of past research efforts regarding the use of AHP in GDM has showed that successful research applications have been accomplished. The applicability and practicability of AHP have been confirmed by several researchers. Several advantages of use of AHP in GDM have been identified by researchers. Some of these AHP is mathematically simple to use, and provide for fair evaluation. The hierarchy of AHP can be easily built to encompass or accommodate several decision makers or experts in a one of its levels. In addition, AHP has an associated procedure to assess consistency of judgments. This provides way to test and revise judgments. Also, several AHP models together can be combined using simple average or geometric mean..., etc. This is in addition to its previously proven capability to handle quantitative and qualitative factors as well, within individual decisions. All this makes AHP applicable to various decision making contexts, like voting, ranking, combining...etc. However, some researches argued upon the disadvantages or limitations of AHP. One of these limitations is that as the number of alternative or criteria increases, the hierarchy can become too complex and the evaluation results can be unreliable. Another limitation is that mutually exclusiveness in evaluation can not be guaranteed, and this exclusiveness if not kept can affect also the reliability of results. However, in spite of its limitations, its advantages weigh more, and still AHP is most widely recognized decision aiding tools especially in the field of GDM. In overall, one successful future application of AHP should make advantage of benefits of utilizing AHP, and try to relax the inherent limitations. This can be provided through avoiding having so many evaluating criteria or grouping criteria and construct hierarchical models rather than flat ones.

Next section, the results of the survey in the field of pattern recognition and binary classification will be reviewed and discussed.

2.3 Combining multiple classifiers

Closely related to the idea of integrating multiple FESs are the researches conducted in combining multiple classifiers in the field of pattern recognition. The idea of combining various expert classifiers with the aim of compensation for the weakness of each single expert while preserving its own strength has recently been investigated widely. The rationale lies in the assumption that by suitably combining the results of a set of experts according to a rule, a combining rule, the performance obtained can be better than that of a single expert. The successful implementation of multiple-expert classifiers requires the use of the most complementary experts possible, and the definition of a combining rule for determining the most likely class a sample should be attributed to, given the class to which it is attributed by each single expert. The result of literature survey in this field is summarized below.

- Xu et al. in 1992 (Xu et al., 1992) proposed four approaches based on different methodologies for combining multiple classifiers. One is suitable for combining any kind of individual classifiers such as Bayesian, K-nearest neighbor, and various distance classifiers. The other three could be used for combining any kind of individual classifiers. On applying these methods to combine several classifiers for recognizing totally unconstrained handwritten numerals, the experimental results showed that the performance of individual classifiers can be improved significantly.

- In 1992, Suen et al. (Suen et al., 1992) considered the idea of combining multiple classifiers applied to the recognition of unconstrained handwritten numerals.

- In 1994, Ho et al. (Ho et al., 1994) proposed five methods for combining multiple classifiers. These methods have been tested in applications and resulted in substantial improvements.

- In 1995, Cho and Kim (Cho & Kim, 1995) proposed fuzzy integral based method for combining multiple neural classifiers.

- In 1996, Ackermann and Bunke (Ackermann & Bunke, 1996) proposed a method for combining decision of multiple classifiers for face recognition.

- In 1998, Kittler et al. (Kittler et al., 1998) developed a common theoretical framework for combining classifiers, which uses distinct pattern representations and showed that many existing schemes can be considered as special cases of compound classification where all the pattern representations are used jointly to make a decision. An experimental comparison of various classifier combination schemes demonstrated that the combination rule developed under the most restrictive assumptions-the sum rule-outperformed other classifiers combination schemes.

- In 1999, Nordmann and Pham (Nordmann & Pham, 1999) dealt with reliability and cost evaluation of weighted dynamic-threshold voting systems. The particular voting system studied consists of n units that each provide a binary decision (0 or 1) or abstain from voting (x). The system output is 1, if the cumulative weight of all 1-opting units is at least a pre-specified fraction of the cumulative weight of all non-abstaining units. Otherwise the system output is 0. Systems of this type are encountered in many areas ranging from pattern recognition, safety monitoring, human organization systems. A general mathematical model of this system was presented.

- In 1999, Cordella et al. (Cordella et al., 1999) proposed a method for estimating the reliability of a single recognition act of an expert on the basis of information directly derived from its output. In this way, the reliability value of decisions is more properly estimated, thus allowing a more precise weighting during the combination phase. The approach was tested using four handwritten character recognition systems, combined in different ways to form 11 multi-expert systems, on the digits of a large standard data base. They made a comparison with classical rules.

- In 2001, Constantinidis et al. (Constantinidis et al., 2001), introduced a new multiple expert fusion algorithm, designated the augmented behavior-knowledge space method. The proposed method overcomes the problem of relying on large data sets in order for the classification methods to be properly utilized. The proposed method solves this problem as it exploits the confidence level of the decision of each classifier. It was shown that this approach is advantageous when small data sets are available.

Critical assessment

Most of the researches have been devoted to defining different combining rules able to solve the conflicts among classifiers, i.e., to determine the most likely class on the basis of the responses of the experts in the case of disagree. Several decisions combination methods like majority voting, arithmetic mean, and Borda count are also applicable to the proposed problem of combining multiple FESs. Still, the possibility of imposing different specific aggregation requirements, like preserving extreme values, has not been addressed. Several weighting methods have been proposed to weight multiple classifiers, but still AHP is superior to them, and has not been exploited much in this field. The research experience in this field will play an important role in dealing with the proposed problem of this research study. The existing combining methods can be configured and utilized analogously in the proposed research.

Next section, some of most recent and prominent research concerning the hierarchical fuzzy systems will be presented. HFS-based models will be utilized in subsequent chapters to develop a model for selection among combining criteria, and to combine/aggregate the FESs' outputs in case of availability of past historical If-then knowledge.

2.4 Review of HFS literature

The HFS was first proposed by Raju et al. in 1991, in order to reduce the number of rules required to design the rule base, which is a necessary performance factor of fuzzy systems. The idea of HFS was also reported in (Brown et al., 1995). HFS can be very useful in dealing with the combination problem of multiple FESs, in that it can provide a logical way to hierarchically handle specific combination relationships among FESs. A review of literature concerning the researches on HFSs is summarized below.

- In 2000, Wei and Wang (Wei & Wang, 2000) proved that the general n-dimensional HFSs are universal approximators, which is considered an extension of the results in (Wang, 1992).

- In 2001, Frantti and Mähönen (Frantti & Mähönen, 2001) presented a fuzzy logic based demand forecasting model, in which the fuzzy results are inferred in three sequential phases. In each phase the number of variables is split due to hierarchical structure of the inference module. A data base and a rule base are divided accordingly into three hierarchical levels.

- In 2003, Lee et al. (Lee et al., 2003) pointed out that one limitation of the HFS is that the intermediate outputs are artificial in nature in many cases and do not possess physical meaning; consequently it becomes so hard to design the intermediate layers, unless the exact relationships between inputs and intermediate outputs is understood and that the intermediate outputs physically can be interpreted. They proposed a new kind of mapping rule base scheme to get the fuzzy rules of HFSs. The algorithm of this scheme is developed such that one can easily design the involved fuzzy rules in the middle layers of the hierarchical structure. In contrast with the conventional single layer fuzzy controller, the presented method has approximate performance using the same scaling factors. The authors concluded that the simulated results showed that the algorithm is effective and feasible.

Critical assessment

The idea of HFS has been proposed recently in the beginning of the last decade. It involves the hierarchical structuring of several fuzzy systems for the purpose of reducing the total number of rules required to completely build the rule base, which is an inherent problem in any huge standard fuzzy system. The main aim of conducting the survey is to get deeper on how HFS can be practically used to structure the proposed problem of combining/aggregating multiple FESs, and at the same time to get acquainted with the limitations of such systems which could be overcome or avoided in this proposed research. Several successful researches have been studied. The idea of HFS can help constructing any fuzzy model or system that is logically and hierarchically structured, has a relatively small number of required rules, and has smooth reasoning paths. One potential contribution of the subject of HFSs in this study is to develop a HFS-based model for combining/aggregating FES's output, which handles past if-then knowledge. Another contribution is to develop another HFS-based fuzzy model for selection among combining criteria. One limitation of HFSs reported in the literature is the difficulty to assign physical meanings for the intermediate outputs of the hierarchical structure. In subsequent chapters, I will explain how this difficulty simply could be overcome if the relationships in the model are well understood and could be logically interpreted.

The next section the results of literature concerning neural classifiers will be presented.

2.5 Review of literature on neural classifiers

The neural classifiers will be used to combine/aggregate FESs' judgments through learning the past available expertise's' performance data and relationships. Some of the relevant researches, in which neural classifiers have been reported successful, are reviewed below.

- In 1987, Lippmann (Lippmann, 1987) reviewed six important neural net models that are used for patterns classification. He described the topology, training algorithm, and capability of each classifier, and discussed their potential advantages and limitations.

- In 1990, Denton et al. (Denton et al., 1990) evaluated the performance of a neural network as a classifier. It was found that the performance of the neural network is comparable to the best of other methods (e.g., linear discriminant function and non-parametric approaches like mathematical programming models) under a wide variety of modeling assumptions. Authors pointed out that the use of neural networks as classifiers thus relieves the modeler of testing assumptions which would otherwise be critical to the performance of the usual classification techniques.

- In 1991, Jacobs et al. (Jacobs et al., 1991) presented a supervised learning procedure for a system of many separate networks, each of which learn to handle a subset of the complete set of training cases, and are called local experts. The new procedure can be viewed either as a modular version of a multi-layer supervised network or associative version of competitive learning. The authors demonstrated that the learning procedure divides up a vowel discrimination task into appropriate subtasks, each of which can be solved by a very simple expert network.

- In 1993, Anhcer and Wang (Anhcer & Wang, 1993) applied back-propagation neural networks algorithm with monotonicity constraints for two-group classification problems.

- In 1995, Lee and Srihari, (Lee & Srihari, 1995) proposed a neural network approach for combining multiple classifiers.

- In 1996, Mak et al. (Mak et al., 1996) presented and compared five techniques for aggregating expertise's in the field of knowledge acquisition. These techniques were: logit regression, disciminant analysis (DA), ID3 pattern classification, k-NN (k-Nearest Neighbor), and neural networks. They conducted an experiment to extract experts' judgments in form of binary decisions on new product entry. The elicited knowledge was aggregated by the five techniques. The neural nets were shown to outperform the other methods in robustness and predictive accuracy.

- In 2004, victor et al. (Victor et al., 2004) proposed a principled approach to building and evaluating neural network classification models for DSSs implementations. The theory concerning model accuracy and generalization was presented.

Critical assessment

Neural classifiers have proved efficient in the field of pattern classification (Lippmann, 1987; Mak et al., 1996; Cordella et al. 1999). In handling past historical data, neural nets have been proven superior to the regression techniques, when there are few number of training patterns or small sample size is available, and when the decision outputs are subjective or non-metric. In most the classification researches, neural classifiers have outperformed other classifiers like K-NN and logistic regression, discriminant analysis. Neural classifiers also require fewer assumptions and offer less complexity than other mathematical and statistical classification techniques, and have the capability and flexibility to incorporate various relationships within the training patterns. In overall, the neural classifiers are strong candidate to be used to combine/aggregate FESs' outputs, through learning the past historical expertise's' performance data.

Next section, the reported evidences extracted out of the investigated literature, and which are in favor of the philosophy of group problem solving (GPS), GDM, and KBSs integration will be presented.

2.6 The reported evidences of the importance and decision support strength offered in integrating systems and expertise's

In chapter 1, the practical reasons of why we should have several independently separated FESs have been explained. In addition to those previously described reasons for independence, some reported evidences from the literature about the importance of integrating several systems and expertise's are presented below:

- In 1996, Sikora and Shaw (Sikora & Shaw, 1996) demonstrated that compared to a single-agent approach, a group problem-solving (GPS) approach, where the examples are distributed to different learning programs, can be very beneficial in terms of producing more accurate rules.

- Studies conducted in (Markham & Ragsdale, 1995; Wang, 1996) suggested that it may be preferable to use the information provided by several experts as a more robust classification schema for use within the two-group problem. An appealing reason for using the opinions of several experts when solving a problem is that a group approach may produce better solutions to complex problems.

- Empirical evidence of benefits of integrated systems in the health care industry has been reported by Forgionne and Kohl in 1995 (Forgionne & Kohl, 1995). Improvements have been found in both process and outcome.

- The dynamic and changing environment of an organization impacts decision making. This means that systems for decision support need to be open-ended, flexible, adaptive, and cover a wide range of expertise's (Mitra & Dutta, 1994).

- In 1996, Mak et al. (Mak et al., 1996) pointed out that based on Arrow's (Arrow, 1951) impossibility theorem of social choice; a consensual solution derived from all experts' judgments collected is a socially acceptable solution. This assumption is particularly useful when the experts posses different views of the problem.

- In the recent years, multiple reasoning has achieved widespread attention. There have been many reports of improvements when reasonably combining multiple systems (Gams et al., 1997).

- In 1999, Cordella et al. (Cordella et al., 1999) pointed out that practical experience suggests that in applications in pattern classification, single recognition system (expert), albeit very refined fails to achieve an acceptable performance level.

- The literature showed that a GPS approach where a problem is decomposed into smaller sub-problems, improves a decision makers performance in certain problem domains (Major & Ragsdale, 2000).

- In 2001, Constantinidis et al. (Constantinidis et al., 2001) stated that: "Practical applications in pattern recognition have been shown to benefit from the combination of the decisions of different algorithms and techniques ("experts"), since classifiers with different internal structures can complement each other. Such multiple expert strategies are likely to deliver more robust decisions than individual classifiers working alone. As a result, it is now common to adopt a variety of decision fusion algorithms to combine the individual expert decisions. The decision combination process has to merge the individual decisions in such a way that the final classification improves the classification profile of any of the individual experts".

All the provided evidences are in favor with the use of cooperative schemes, in form of GPS or GDM, multi-expert systems, multi intelligent or knowledge sources, multi-classifier, multiple ESs…, etc, and this is due to the inherent benefits expected from realizing integration. However, the key success factor of this proposed research is to objectively realize integration through developing effective combination/aggregation methods, which will be the main focus through out the incoming chapters of this thesis.

In the next chapter, I will get deeper into the first phase of problem solving; that is the formulation and structuring of the integration problem.

Chapter 3

Formulating and structuring the integration problem

Before attempting to solve any problem, it is important to define and understand such problem. This is necessary in order to correctly focus on solving the real problem. Otherwise solution could be far away from the expected. After defining the problem, the thinking of how to solve such problem should begin. One of the common strategy or approach generally employed in problem solving is to divide the main problem into smaller, easily manageable problems or components. This contributes to reduction of the complexity associated with the attempt to solve the main problem. In this chapter, the integration problem of a multiple FESs will be defined, stated and structured. This is in order to be able to organize the solutions to the problem or its components in such a way that enables realizing the expected contribution of these component solutions to the solution of main integration problem.

The problem of integrating multiple parallel FESs evaluating a Yes-or-No decision making problem is defined as combining or aggregating the crisp outputs provided by these individual systems into a finally consolidated binary decision. In order to consistently be able to formulate and structure the integration problem, it is important to agree on the format of the crisp outputs provided by the FESs, which will be the inputs to the combination/aggregation problem. Also, it is important to specify how such unified format is to express the subjective "Yes" and "No" decision options. It is necessary to have a unified and objective format for such crisp outputs in order to be able to develop objective and appropriate combining/aggregating methods. The next section will address this issue.

3.1 Unification of the output format of the FESs

In order to be able to consistently combine or aggregate the individual decision outputs of the integrated FESs, it is necessary for these outputs to follow a standardized or unified output scale. The subjective investigation of the combination/aggregation literature has revealed that combining/aggregating the outputs at the measurement level, in which the numerical score is used to determine the degree of bias or belonging to a class, permits the use of more sophisticated combination/aggregation criteria or algorithms, whereas combination at only abstract level, in which preferences are expressed by only identifying the preferred class, allows only low level or simple criteria like the majority voting to be used. The combination/aggregation at the measurement level also permits to use all criteria at all other lower levels of combination/aggregation. Therefore, the adopted unified output scale should be objective or numerical, and should allow combination/aggregation at the measurement level. Given this fact, every participating FES should produce a crisp numerical output within a unified numerical psychometric scale from 0 to a maximum arbitrary value; where 0 means complete bias toward "No" decision answer and the maximum value means complete bias toward "Yes" decision answer. Intermediate values give the degree of bias toward either decision option. The maximum value can be arbitrarily set to 100, 10, or 1…etc., without affecting the final combined decision solution. The basic notion behind this unified scale is to provide for consistency and homogeneity in assessing the FESs' decision outputs. Through out this research, a unified output scale ranging from 0 to 10 will be utilized to assess the binary judgments of the individual FESs. Figure 3.1 simply depicts the proposed unified output scale.

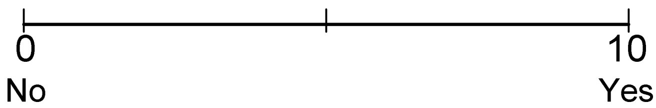

Fig.3.1 The objective numerical judgmental scale used for assessing FESs' outputs.

In the next section, the integration problem will be formally stated, based on the above established output scale.

3.2 Problem statement

Given a whole set of FESs, which complement each other for a finite set of decision making transactions or problem contexts, then, based on a current transaction or context, an appropriate relevant set of FESs, particularly suited for this current transaction or context is selected. Every FES should provide its decision output in form of crisp numerical value within 0 to 10, following the previously established unified scale. Then, the problem of combining or aggregating the decision outputs of the matched FESs can be formulated as a GDM problem with only two alternatives (i.e., "Yes" and "No" decisions). Therefore, the problem is then formally stated as follows:-

Let $A = \{a_1, a_2\}$ be a finite set of the two decision alternatives (i.e., a_1 = "Yes", a_2 = "No").
Let $F = \{f_1, f_2, \ldots, f_k, \ldots, f_m\}$ (m ≥ 2) be a finite set of FESs, $O = \{o_1, o_2, \ldots, o_k, \ldots, o_m\}$ ($0 \leq o_k \leq 10$) be the crisp numerical outputs of FESs representing the degree of bias toward either a_1 or a_2 (the value 0 exhibits complete bias toward a_2, and value 10 exhibits complete bias toward a_1), $W = \{w_1, w_2, \ldots, w_k, \ldots, w_m\}$ be the associated weights set of the FESs, where each k^{th} FES, f_k, gives an output o_k, and has a weight value w_k ($w_k \geq 0, \sum w_k = 1$). Then:

The problem now is to find a combining/aggregating criterion, or algorithm, C, with an interpretation function, I, which maps the set of individual decision outputs of FESs, O, into a one collective, consolidated group decision, O_f, $0 \leq O_f \leq 10$. The interpretation function, I, associated with the combining/aggregating criterion maps the combined/aggregated output value into either class, "Yes" or "No", of the set of two classes or alternatives. Formally stated:-

$$C: O^m \rightarrow [0, 10] \tag{3.1}$$

Where, O^m is the group preferences or outputs vector, and

$$I: [0,10] \rightarrow \{a_1, a_2\} \tag{3.2}$$

The combination/aggregation criterion should take into account and satisfy a set of pre-established general and specific requirements as needed. These requirements are identified, and most of them are emerging from a real need of the aforementioned currently held practical project. Other requirements are elicited subjectively as possible requirements.

<u>The general requirements:-</u>
They are related to three different issues. The first issue is concerned with the way in which a combination/aggregation method or criterion is used; whether it computes the combined/aggregated value based on the present or current information about the individual output values, or it handles past historical data and knowledge. The second issue is related to

the form of decision outputs to be manipulated; whether they are continuous-valued outputs, or binary-valued outputs. The third issue is related to whether we have several FESs that share common knowledge (i.e., same specialization area, but different skills or tools) or we have several FESs, each of which has a unique knowledge or different expertise that adds to the complete knowledge of the problem. These three general requirements can be summarized as follows:-

- Handling present outputs information versus handling past historical expertise's' performance data.

- Manipulating continuous-valued outputs combination versus manipulating binary valued outputs combination.

- Combination of knowledge-equal FESs versus aggregation of knowledge-unique FESs.

The specific requirements:-
They are more closely viewed as restrictions on combining/aggregating the individual outputs of FESs, and they are concerning the roles of individual FESs and relatedness among their individual decisions. These requirements are:-

- Preserving extreme values of FESs' output decisions; these extreme values should have extra importance, since they more strongly refer to one of the two decision alternatives, and consequently they should be dealt with some special care.

- Provision for null participation of certain FESs, based on some reasons like incomplete or insufficient input information, or recommendations of the corresponding expertise's to not to participate for any other reason.

- Provision for mutually related FESs' decisions. When some subsets of the FESs are related, then some special actions should be made to exploit this relatedness in getting more information about the reliability of the group decisions.

- Provision for veto-type privilege given to some critical decision outputs of one or more FESs. This veto-type privilege actually arises under certain circumstances, and in such cases all the group decisions are forced to be in favor of the decision direction of such critical FESs.

In the next section, the first step toward organizing the research efforts in dealing with the integration problem is to be accomplished, which is the structuring of the problem and the candidate solution approaches.

3.3 Structuring the problem

Based on the statement of the problem, and after conducting an extensive review of relevant literature, the integration problem and the combining/aggregating approaches can now be structured. This will help organizing the research efforts toward solving the problem in a more transparent and controlled way.

3.3.1 Structuring the integration problem

First, the integration problem can be structured and categorized into two main problem cases:

- Integrating multiple parallel FESs each of which holds the same domain knowledge (shortly called: “knowledge-equal” FESs).
- Integrating multiple parallel FESs each of which holds different or unique domain knowledge (shortly called: “knowledge-unique” FESs).

The main difference between the two problem cases is that, in the first problem case, the multiple participating FESs share the same domain knowledge or have same specialization, but at the same time each of which acquires different reasoning skills, tools and may be different views of the problem. The main reason of integration is to increase the reliability and quality of the obtained group decision, and hence the expertise’s are duplicated, and need to be combined in some way to reach at a final consolidated decision. In the second problem case, the participating FESs have different domain knowledge’s and expertise’s, different areas of specialization that are all necessary to comprehend all the corresponding relevant decision aspects of the problem, and hence these expertise’s should be accumulated or aggregated to obtain a final group decision. In addition, in the first case problem, every FESs is able solely to make the decision, but the synergism and multiplicity of FESs are necessary in order to increase reliability and quality of the group decision, especially when this decision type is critical. On the other hand, in the second case, every FES cannot solely make the decision, because it grasps only one decision aspect out of all relevant decision aspects to the problem, and the solution of the problem is obtained through aggregating or accumulating the knowledge’s about all relevant aspects. As a simple example for the need of multiple unique expertise’s, let us consider a modern car to be diagnosed as being either “safe” or “insecure”. Then, in order to judge this binary-type decision, several expertise’s are required to grasp all decision aspects that tell either the car is “safe” or “insecure”. Suppose that the needed expertise’s are related to the following decision aspects: vibration, electric power, electronics, structural mechanics, chemical or environmental…, etc. Here, only one expertise can not solely make the decision. In order for the final decision to be complete and representative, the opinions of all such expertise’s should be considered and accumulated.

Therefore, based on a current judgment problem, an adequate, relevant, and predefined group of FESs is selected. The problem is then either to combine the outputs of the multiple systems acquiring or specializing in the same knowledge domain or to aggregate the outputs of the multiple ones specializing in different, relevant domains of knowledge. The conceptual structure is depicted in Figure 3.2.

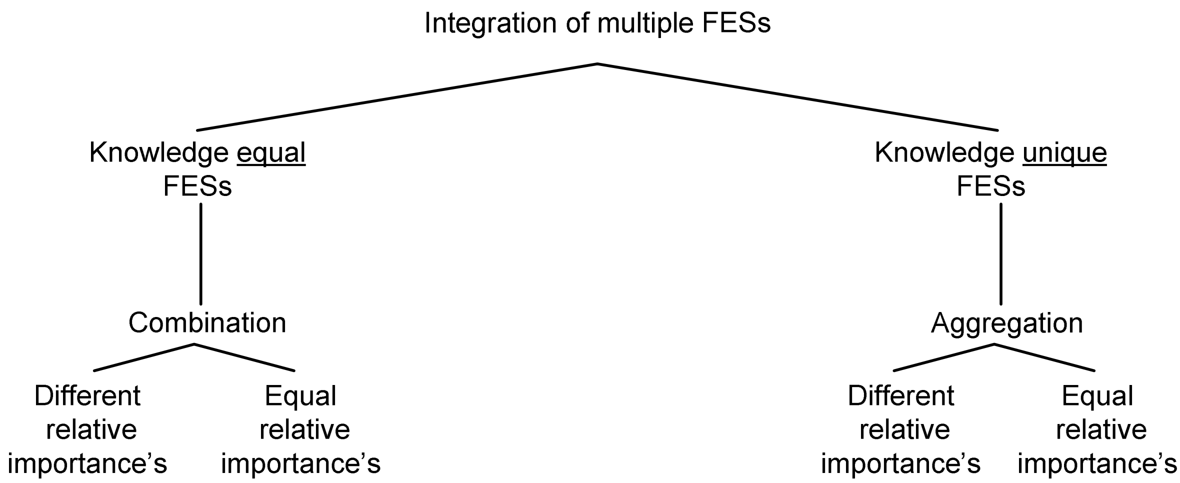

Fig. 3.2 The conceptual structure of the FESs' integration problem.

*I have published the conceptual structure of the integration problem in ((Aly & Vrana, 2005c) & (Aly & Vrana, 2006a)).

Next section, the candidate solution approaches to the integration problem will be structured.

3.3.2 Structuring the candidate solution approaches

Second, the candidate combination/aggregation approaches are to be also structured according to the established conceptual structure and the general requirements of the integration problem. The candidate solution approaches are structured as shown in figure 3.3. As it is shown in the figure, there are two possibilities to deal with the integration problem; the first possibility is concerned with handling the available past historical expertise's' performance data and knowledge. The candidate approaches are: the multi-layer feed-forward back-propagation network (BPN), and a HFS-based model. Both models are to combine/aggregate FES's outputs. The justification of the adoption of these two models will be stated in subsequent chapters. The BPN will be used to manipulate the past expertise's performance data patterns prior to combining/aggregating FESs' outputs. The HFS-based model will handle the past existing If-then knowledge to combine/aggregate FESs' outputs. The second possibility is concerned with handling present or current information about the FESs' output values. Some of the candidate approaches will be used to combine expertise's and some other will be used to aggregate them. Further, some approaches will work on continuous-valued outputs like the simple Arithmetic Mean combining criterion and an aggregation heuristic, and other methods will convert the FESs'outputs into binary form, like the Majority Voting combining criteria. All of these approaches will be presented in the subsequent chapters.

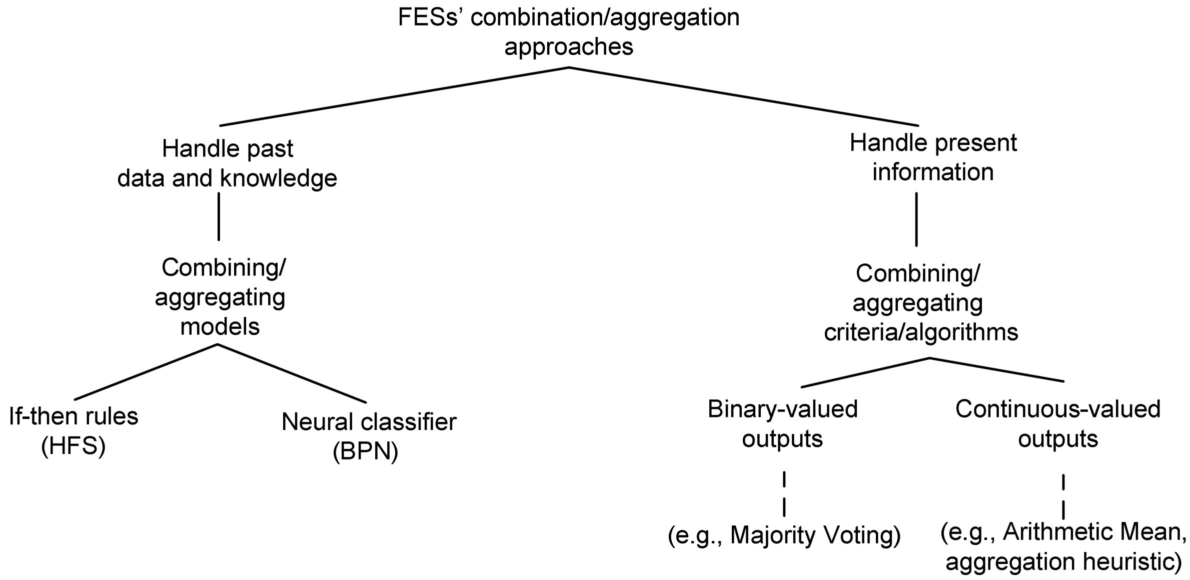

Fig. 3.3 FESs' combination/aggregation approaches structured according to the basic requirements of the problem.

Having structured the integration problem, and the possible or candidate solution approaches, all the subsequent steps will follow these two conceptual structures. Throughout the thesis, all approaches are to follow the established unified output numerical scale.

In the subsequent chapters, the solution approaches to the integration problem will be presented with concentration on the justification of the utilization, adoption, and development of these approaches, and explanation of how these approaches add value to the overall solution of the integration problem, and how these individual approaches contribute to the provision for the satisfaction of the stated general and specific requirements.

Next chapter will consider how to include and assess the relative importance's of the participating FESs, under various decision making circumstances.

Chapter 5

Combining/aggregating the crisp outputs of the fuzzy expert systems

In chapter 3, the problem of integrating multiple parallel FESs was structured into two basic problem types:

- Integrating multiple FESs each of which holds the same domain knowledge ("knowledge-equal" FESs),

- Integrating multiple FESs each of which holds different or unique domain knowledge ("knowledge-unique" FESs).

The main difference between the two problem types was summarized as being in the first case problem, we have several FESs, each of which grasps all decision aspects of the problem, each of which has different skills, and can make the decision solely, but the inclusion of several synergetic FESs aims to increase the reliability and quality of the decision made. In the other hand, in the second case problem, the incorporation of several FESs is necessary to encompass and grasp all relevant decision aspects of the problem, where every individual FES can not solely make the decision. In the first case problem, the integration of the FESs is to be accomplished through combination of all expertise's, whereas in the second case problem, the integration is to be accomplished through aggregation of all such expertise's. This chapter is concerned with presenting the solution approaches or the integration mechanisms for both problem cases.

Next section, the combination problem will be addressed.

5.1 Combining the crisp outputs of knowledge-equal FESs

In this section, possible combining rules, methods and criteria for the first problem of combining knowledge-equal FESs will be adopted and developed. The problem can be viewed as how to combine several decision opinions of several decision makers, the FESs, where their judgments may be conflicting. These multiple judgments are in the form of numerical real values within the range [0,10] expressing the degree of bias toward "Yes" or "No" decision options.

Based on the extensive survey of literature, there have been no previous attempt to integrate ESs through combining their final decision outputs; except for the pattern classifiers in the field of pattern recognition (reader may refer to the literature review in the thesis (chapter 2)). However, the literature has revealed that the arithmetic mean (AM), and geometric mean (GM), and their weighted versions, the weighted arithmetic mean (WAM), and weighted geometric mean (WGM) have been extensively used in combining numerical values in the GDM field. This is beside their general usage in statistics, specially the AM, which is generally used to measure the central tendency of group of numerical values. In the pattern classification, the majority voting rule (MV), and its weighted version, the weighted majority voting (WMV) has been widely utilized. Also, Borda count has been extensively used in aggregation of group preference ranking of several decision alternatives. In this section, the applicable classical combining criteria will be reviewed and categorized according to their context of usage and format of their inputs data. Then, a comment on their characteristics, limitations, and advantages will be made. After, these criteria are to be configured to the binary GDM involved in combining the crisp numerical outputs of the multiple FESs. Then, a promising new combining criterion will be introduced, which offers some more desirable features than the existing classical ones. The distinct features of this new criterion will be explained, and finally its weighted version will be presented.

5.1.1 Classical combining rules

Cordella et al. in 1999 (Cordella et al., 1999), pointed out that many methods used to combine experts' decisions have been proposed in the recent past. Many methods have been proposed in the area of combining multiple classifiers. Some of them are based on heuristic approaches, such as voting or ranking strategies, while others are based on probability theory, e.g., the Bayesian method (Ackermann & Bunke, 1999). Theoretically, the performance of the combining criteria of a given set of experts' decisions should augment the amount of information provided by each single expert's decision. In the literature, the various combining rules or criteria are divided into three types; depending on the output information experts provide (Xu et al., 1992). Type 1 classifiers output unique label, i.e., the label of the presumed class; they are also known as classifiers that works at an abstract level. Type 2 classifiers, which work at a rank level, rank all classes in a queue where the class at the top is the first choice. Type 3 classifiers, which work at a measurement level; they attribute each class a measurement value to represent the degree that input sample belongs to that class. Almost all the combining rules are defined with reference to Type 1 classifiers. However, (Cordella et al., 1999) stated that combining schemes that exploit information from the classifiers at the measurement level allow us to define combining rules that are more sophisticated and potentially more effective. Since combination methods span a wide variety of research areas, the term module is used to refer to the individual units to be combined. A module can be an expert, an ES, a forecaster, an estimator, or a classifier. Depending on what level of information received from the module, there are three types of combination (Al-Ghoneim & Kumar, 1998):-

a) Combination at an abstract level:

At this level, combination criteria or algorithms use only the abstract level information, the identity of the top class, provided by the modules. This methods are based on voting procedures that are adopted from GDM theory such as unanimity, majority, plurality,…, etc. The majority and plurality voting rules are the most widely used.

b) Combination at a rank level:

At this level, the classifier modules provide rank information; that is the preference ordering of classes from top to bottom rank. Every module provides a sorted list of classes arranged in order of preference. A well-known combination method at the rank level is the Borda count.

c) Combination at a measurement level:

At the measurement level, the combination criteria or algorithms have access to a set of numerical scores provided by the classifier modules. Cordella et al. in 1999 stated that combining schemes that exploit information from the classifiers at the measurement level allow us to define combining rules that are more sophisticated and potentially more effective.

The established unified scale for FESs' outputs provides for combination at a measurement level. Combination or aggregation at the measurement level is the main focus in this research.

Let us review the most widely recognized combining criteria or rules under each combination level.

I. Combination at the abstract level:
The majority and plurality voting rule are the most widely used. They are further divided based on whether or not they take account of relative importance's of modules. They are as follows:-

A. <u>Equal-weights voting rules:</u>

Definition: Majority Voting (MV): This rule chooses the classification made by more than half the voters; when no such class is found, the result is considered as an error or rejected (Hansen et al, 1990; Cordella et al., 1999).

In case of two or binary decision options, this rule becomes the same as the plurality voting rule, which chooses the class that gets more votes than any other class. However, I will call this rule the majority voting because it is most commonly used term in the binary classification literature. Now, the rule is defined mathematically as follows:

Let:
V_i^j : the vote of the ith module for the jth decision option, j = 1 for "Yes" decision option, and j = 0 for "No" decision option, i = 1,2,…n.
n : the total number of modules.

Then:

$$V_i^j = \begin{cases} 1 & \text{if the decision is made in favor of class j.} \\ 0 & \text{if otherwise.} \end{cases}$$

Hence, the total vote of the jth decision option is given by:

$$V^j = \sum_i V_i^j \tag{5.1}$$

Then, the final group decision of the multiple modules is:

$$O_f = \begin{cases} \text{"Yes"} & \text{if } \arg\max_j V^j = 1. \\ \text{"No"} & \text{if } \arg\max_j V^j = 0. \end{cases}$$

Where, $\arg\max_j V^j$ gives the class j, which has the greater number of voting's.

Note: If no such class exists; that is the total number of voting's received by each class were equal, then this final group or modules' decision is titled "Non-biased", or "unclassified".

B. Weighted voting rule

Given the same definition of the majority voting rule, the weighted voting utilizes additional information in combining the modules' outputs, which are the weights of the participating modules, w_i. This reflects the relative importance of every module. The rule is mathematically defined as follows:-

$$WV^j = \sum_i (w_i * V_i^j) \tag{5.2}$$

Then, the final group decision of the multiple modules is:

$$O_f = \begin{cases} \text{"Yes"} & \text{if } \arg\max_j WV^j = 1. \\ \text{"No"} & \text{if } \arg\max_j WV^j = 0. \end{cases}$$

Where,

$\arg\max_j WV^j$: gives the class j, which has the maximum weighted voting.

Note: If $WV^1 = WV^0$, that is the weighted voting's were equal, then this final group or modules' decision is titled "Non-biased", or "unclassified".

II. Combination at a rank level:

The most well-known combination method at the rank level is the Borda count. This rule is a generalization of majority voting rule. It is defined as follows:-

Definition 3. Borda count rule: for any class C_j, let B_j^i be the number of classes ranked below C_j by module i (i = 1, 2, …., n). The Borda count for class C_j is:

$$B_j = \sum_{i=1}^{n} B_j^i \tag{5.3}$$

n : the total number of modules.

The Borda count decision rule works by picking the class with the highest B_j.

Actually, the use of Borda count is confined to combining the preference ranking of a group of decision makers or systems, usually when there are more than two decision options or alternatives. In this thesis, the focus only on the binary or two-option decision making. This rule has been only described for the purpose of demonstrating the difference among the three ways or combination. The main focus will be on the combination at other types of combination.

III. Combination at a measurement level:

There are numerous approaches to the combination of the scores of various modules into a combined score for group of modules. A popular approach uses quasi arithmetic means as a family of algebraic combination methods (Smolíková & Wachowiak, 2002; Ben-Arieh 2005). These means defined as follows:-

A. Generalized means:

$$F_\alpha(x) = h^{-1}\left[\frac{1}{n}\sum_{i=1}^{n} h(x_i)\right], x \in I^n \tag{5.4}$$

Where h is continuous strictly monotonic function (h^{-1} is the inverse function of h.). Four types of means can be formed from this general function:

1. Root-power or generalized mean:

Let $h(x) = x^\alpha$ then $h^{-1}(x) = x^{1/\alpha}$.

$$F_\alpha(x) = \left(\frac{1}{n}\sum_{i=1}^{n} x_i^{\alpha}\right)^{1/\alpha}, x \in I^n \tag{5.5}$$

2. Geometric mean (GM):

For $\alpha \to 0$,

$$GM = F_0(x) = \lim_{\alpha \to 0} F_\alpha(x) = \sqrt[n]{\prod_{i=1}^{n} x_i} \tag{5.6}$$

3. Harmonic mean (HM):

For $\alpha \to -1$,

$$HM = F_{-1}(x) = \frac{n}{\sum_{i=1}^{n} 1/x_i} \tag{5.7}$$

4. Arithmetic mean (AM):

For $\alpha \to 1$,

$$AM = F_1(x) = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{5.8}$$

And for every $x \in I^n, F_{-1}(x) \le F_0(x) \le F_1(x)$.

B. Weighted means:

If we denote the weight of the i^{th} participating module as w_i, then:-

1. Weighted geometric mean (WGM):

$$WGM = \prod_{i=1}^{n} (x_i)^{w_i} \tag{5.9}$$

2. Weighted harmonic mean (WHM):

$$WHM = \frac{\sum_{i=1}^{n} w_i}{\sum_{i=1}^{n} w_i / x_i} \qquad (5.10)$$

3. Weighted arithmetic mean (WAM):

$$WAM = \frac{\sum_{i=1}^{n} w_i * x_i}{\sum_{i=1}^{n} w_i} \qquad (5.11)$$

It should be noted that according to the inequality that directly follows equation 5.8, it is always a condition that the combined value using the AM will be always greater than or equal to that produced by the GM, which is always greater than value produced by HM. Also, the HM tends always to give low combined values, and its use is not so common. From the decision making viewpoint, these distinct computational characteristics mean that, in most cases, the combined result of only one of the three formulas will be optimal. This is because logically the correct decision answer is unique. It also means that if we know the value of one formula, we can expect where the values of the others lie. Another notion is that the performance of these criteria will not only be affected by the mathematical form, but also, by this distinct dependency. This may make some of the three suitable in some situation and not appropriate in another. For instance, if there is agreement in high magnitudes of values, then this will be more reflected by the arithmetic mean, since it gives the highest magnitude of the three formulas, whereas, if there is agreement in low magnitudes of numerical values, then the HM gives the best result. Also, it should be noted that, the above three combining criteria, basically have statistical root, functionality and meaningfulness, and this does not mean that they are always suitable for decision making in combining multiple numerical judgments. For instance, the AM, from the statistical viewpoint measures the central tendency based on the magnitudes of a group of numerical values, but in such binary decision making problem, this measure is not always adequate, where in decisiveness and extreme values we might have special interest.

Based on the survey of the literature, the most widely utilized combining and preference aggregation criteria are the AM and GM criteria. Here, the term preference aggregation is different from the second case of the integration problem, which aggregation of expertise's. I have differentiated in this thesis between the terms aggregation and combination, but because the preference aggregation is a subject in the GDM field, I have mentioned it as it is known. The HM criterion is seldom used, if not never. So, I will not much concentrate on this criterion, but however all criteria, and even the HM will be compared in a subsequent chapter. The main concentration will be on the well-known and widely utilized AM.

One advantage of decision combination at the measurement level is that it allows more sophisticated combining criteria to be used, and which adds extra information about not only which class is selected or which rank it received, but also the degree at which class is evaluated by every module. This can be reflected on some decision aspects associated with the reaction to the finally obtained decision.

Next section, the classical combining criteria will be configured to combine the FESs' outputs, according to the established output unified scale.

5.1.2 Configuring the combining criteria for binary decision making

In order to be able to numerically combine the individual judgments of the group members, these judgments should be numeric; otherwise no way to combine them that takes into account the decisive degree of decisions made. The MV combining rule can be used but it does not accurately reflect the magnitudes of numerical judgments. It works only based on subjective values of judgments "Yes" or "No". In this study, I am focusing on the aggregation or combination at measurement level, that takes into account the decisive degree of "Yes" or "No" decision.

Given the numerical scale [0,10] used to distinct between "No" or "Yes" decisions, the previously described measurement-level combining criteria AM, GM, HM, can be utilized to combine multiple numerical judgments as follows:-

- If the resulting combination value is above the middle (> 5), then the group final decision is "Yes".
- If the resulting combination value is below the middle (< 5), then the group final decision is "No".
- If the resulting combination value is equal to the middle (= 5), then the group final decision is "Non-biased".

Also, a threshold value ± 0.5 could be used to more decisively attribute the resulting combination value into "Yes" or "No" decision answers. This means that, if the resulting value is greater than 5.5, then it is attributed to "Yes", and if the resulting value is less than 4.5, then it is attributed to "No"; otherwise it is attributed to "Non-biased" class. This threshold gives more distinction for decision attribution to either class. It should be noted that any other threshold value could be utilized, depending on the analysts' view of the required degree of decisiveness.

Combining criteria at abstract level, like majority voting, can be also utilized based on attribution of every individual FES's output to either decision class. This can be done as follows:-

- If the FES's crisp output value is above the middle (> 5), then it is classified as "Yes" decision.
- If the FES's crisp output value is below the middle (< 5), then it is classified as "No" decision.
- If the FES's crisp output value is equal to the middle (= 5), then it is classified as "Non-biased" decision.

A threshold value ± 0.5 could be used also to enhance the decisiveness of class's attribution. Then, majority voting rule can be applied.

Combining criteria at a rank level, like Borda count, are utilized frequently when there exist more than two decision options, in order to be significant and meaningful. Their utilization in the literature was not recorded for binary decision making, if not at all. So, I shall not concentrate on them any more, out of this section. However, the Borda count could be also utilized, after adequately converting the outputs of FESs as follows:

- If the FES's crisp output value is above the middle (> 5), then its first rank class is "Yes" and its second rank is "No" decision.
- If the FES's crisp output value is below the middle (< 5), then its first rank class is "No" and its second rank is "Yes" decision.
- If the FES's crisp output value is equal to the middle (= 5), then the two decision classes have equal rank, or in other meaning neither classes are classified first.

Then, the Borda count can be applied.

It should be noted that, the established output numerical and unified scale offers flexibility in allowing for the use of all types of combination criteria, as it has been described above. However, up to now, two important problems arise concerning the utilization and selection of the combining criteria. The first is that although it is agreed that the combination at the measurement level is more advantageous because it allows more sophisticated combining methods to be used, and also because that it enables processing of more detailed information about numerical judgments, there is no way to decide which combining criteria whether at abstract or measurement gives the optimal group output. Therefore, a method should be developed to decide, which criterion to apply in a given situation depending on the only available information; that is the individual outputs of FESs. This issue will be addressed in chapter 8.

The second problem is that both the widely used arithmetic mean and geometric mean combining criteria are compromising more than decisive, especially in this two-class or binary decision making, where the decisiveness is so important in order to classify the decision answers into one of the two decision options or classes. In addition, they result in most situations in information loss due to the smoothing effect they impose on the combined values. Both formulas never give the extreme value of a group of numerical judgments, unless all values are equal. Also, they work always by pointing to the center of the numerical data based on their values not on the direction of decision answer, and this is considered not adequate in many situations, especially when we need a considerable degree of decisiveness to reach at either "Yes" or "No" final decision. The AM, which most commonly utilized in combination, statistically measures the central tendency of group of numerical values, but not the degree of agreement of these values, and always results in missing the extreme values. Therefore, there is a need for more logical and decisive criterion, that more sensitively reflects degrees of agreement in group decisions, and that tends to point to either extremes, 0 or 10; correspondingly "No" or "Yes".

In the following section, a new promising logical combining criterion will be presented, which offers some more desirable features than the commonly utilized formulas AM and GM.

5.1.3 The multiplicative proportional deviative influence (MPDI) combining criterion

The multiplicative proportional deviative influence (MPDI) is a Black-board (Chi et al. 2001) inspired new combining criterion. It imitates some processing aspect of the blackboard concepts in integrating multiple knowledge sources. In the black-board, knowledge sources, which can be experts or any intelligent systems, interact via a shared global data structure – the blackboard that organizes and stores the intermediate problem solving data. Knowledge sources produce changes to the blackboard that lead incrementally to a solution of the problem. Communication between knowledge sources is conducted solely through changing the blackboard. Similarly, the MPDI combination is based on this idea in that there are multiple knowledge sources each of which changes a numerical value initially existing in the black-board. Every knowledge source bears numerical value within the range [0,10] expressing the degree of bias to either “Yes” or “No” decision answer. Initially, the blackboard contains the middle value 5 of the psychometric scale, which expresses that initially there is no bias. Then, all knowledge sources fairly participate in changing this initial value based on their <u>deviation</u> from that middle value. That is the <u>influence</u> of every numerical value is <u>proportional</u> to its deviation from the middle. Then, all deviations are accumulated on the middle by <u>multiplicatively</u> augmenting it, if the deviation is positive, and decrementing it, if the deviation is negative. This is why, I call it Multiplicative Proportional Deviative Influence, or MPDI, and mathematically defined and expressed as follows:-

Suppose that there are n experts currently participating in judging the binary decision problem. Then, let us define the followings:-

O_m : the initially non - biased middle value, 5.

O_i : the i^{th} expert's judgement, $i = 1,2,.....n$.

ΔO_i^+ : the absolute deviation of the i^{th} expert judgment O_i, above from the middle.

$$\Delta O_i^+ = \begin{cases} O_i - O_m & \text{if } O_i \geq O_m \\ 0 & \text{otherwise} \end{cases}$$

ΔO_i^- : the absolute deviation of the i^{th} expert judgment O_i, below from the middle.

$$\Delta O_i^- = \begin{cases} O_m - O_i & \text{if } O_i \leq O_m \\ 0 & \text{otherwise} \end{cases}$$

Then :

$$MPDI = \left[\frac{\prod_{i=1}^{n} \left[1 + (\Delta O_i^+ / O_m) \right]}{\prod_{i=1}^{n} \left[1 + (\Delta O_i^- / O_m) \right]} \right] * O_m \qquad (5.12)$$

The limiting range of values the MPDI criterion can take is within $[\,5/2^n\,,\,5*2^n\,]$, which is a function of the number of experts or decision makers involved in judgment.

<u>An illustrative example on applying MPDI:</u>

Suppose that there are 6 experts’ judgments, then the new combining criterion is used as follows:-

$O_1 = 1$, $O_2 = 4$, $O_3 = 2$, $O_4 = 3$, $O_5 = 6$, $O_6 = 10$.
$O_m = 5$.
Then, the deviations from the middle are computed as follows:

$\Delta O_1^+ = 0, \qquad \Delta O_1^- = 4.$

$\Delta O_2^+ = 0, \qquad \Delta O_2^- = 1.$

$\Delta O_3^+ = 0, \qquad \Delta O_3^- = 3.$

$\Delta O_4^+ = 0, \qquad \Delta O_4^- = 2.$

$\Delta O_5^+ = 1, \qquad \Delta O_5^- = 0.$

$\Delta O_6^+ = 5, \qquad \Delta O_6^- = 0.$

Then, according to equation 5.12:

$$\text{MPDI} = \left[\frac{1*1*1*1*1.2*2}{1.8*1.2*1.6*1.4*1*1}\right]*5 = 2.48 \quad (\text{“No” decision})$$

The AM and GM give:

AM = (1 + 4 + 2 + 3 + 6 + 10) / 6 = 4.33 (“No” decision)
GM = $(1 * 4 * 2 * 3 * 6 * 10)^{1/6}$ = 3.36 (“No” decision)

The MPDI clearly reflects the visually appearing consensus in the experts’ numerical judgments, where 4 out of 6 answers were biased to “No” direction answer. In addition, the MPDI outperforms the AM and GM formulas, in that it was more decisive in pointing to the “No” decision answer, since it has provided value 2.48, which is closer to 0 than 4.33 and 3.36 (that is more close to non-biased decision!). However, the experimentation to be conducted in chapter 7 will give more non-biased insight about the performance of the proposed MPDI relative to that of the described measurement-level criteria.

Next section will discuss the distinct characteristics of the proposed MPDI.

5.1.4 The distinct features and characteristics of the MPDI

First, the MPDI notion is logically understood as a group of numerical values imposing fair influences on an initially existing non-biased value (i.e., the middle value, 5), based on their deviations from that middle value. It can be considered as a more logical formula than the widely used AM, which only measures the central tendency of a group of numerical values. The AM and GM have important limitations in that they imposes smoothing effect on a group of numerical values, and that results in some type of information loss. In addition, they hide the extreme values of a group of numbers, which in our case of binary decision problem, are specially desired values. In contrast, the proposed MPDI formula is considered more decisive than the AM in the sense that the multiplication process done in using MPDI always gives more magnified values than a sum process. Multiplication magnifies the agreement of more than one numerical value, multiplied at the numerator or dominator of the MPDI, giving larger value than the result of addition as done with use of the AM or GM. This is especially important in GDM, where the agreement of several decisions should be reflected in the final group decision. For instance, consider the following three numerical values to be combined, 3, 7, 7:
AM = 5.67 (“Yes” decision)
GM = 5.27 (“Yes” decision)
MPDI = 7 (“Yes” decision)

It is clear that the result of MPDI is more decisive than that of AM, and this is logical since two experts' judgments out of three are biased to same "Yes" direction, and this is reflected more by MPDI, which has magnified the agreement of the two values by multiplication in the numerator. In contrast AM has provided a compromising solution; it is biased to the "Yes" side but not as decisive as result of MPDI. Also, GM has given a result that is even less decisive than that of AM. If someone sets a threshold ± 0.5 as a degree of confidence about "Yes" or "No", then in this case, the resulted GM value will be classified as "Non-biased" value. Another example, consider the following three numerical values to be combined, 3, 3, 7:

AM = 4.33 ("No" decision)
GM = 3.98 ("No" decision)
MPDI = 3.57 ("No" decision)

Still again in the other direction MPDI has provided more decisive result than AM. Consider the following three numerical values to be combined, 6, 8, 9:

AM = 7.67 ("Yes" decision)
GM = 7.56 ("Yes" decision)
MPDI = 5 * (1.2 * 1.6 * 1.8) = 17.28 (10) ("Yes" decision)

As mentioned earlier, upper limit of MPDI can exceed value 10, which should be interpreted as a greater consensus toward "Yes" decision answer. This can be also logically interpreted as with the increase in number of experts who agree on a particular option, this should be reflected into reinforcement of their concordant answers. The over flow of the combined values exceeding 10, does not cause problem, because it reflects the high degree of agreement that has led to saturation over 10, the maximum of the used scale. This is also due to the accumulation of knowledge.

Next sections, the developed MDI criterion will be verified against the most widely accepted social axioms.

5.1.5 The social choice axioms

Many social axioms governing the process of group preference aggregation have been proposed since the early work of (Arrow, 1951). (Richelson, 1981) evaluated many social choice functions (or voting systems) such as the MV and the Borda count using several social choice axioms. The four most commonly employed axioms are (Sen, 1969; Mirkin, 1979; French, 1986):

<u>Axiom 1</u> (**Universal domain**): the group preference aggregation method should define a group preference pattern for all logically possible individual preferences.

<u>Axiom 2</u> (**Pareto optimality**): Let A and B are two alternatives. If all the group members prefer A to B then the group preference should be in favor of A.

<u>Axiom 3</u> (**Independence of irrelevant alternatives**): if an alternative is eliminated from consideration, then the new group ordering for the remaining alternatives should be equivalent (i.e., same ordering) to the original group ordering for the same alternatives.
Note: this axiom is not relevant here, in binary decision problem, where we have only two fixed decision options. However, it was mentioned as for the generality.

Axiom 4 (**Non-dictatorship**): there is no individual whose preference automatically becomes the preference of the group, independently, of the preferences of other group members.

The Pareto Optimality axiom is almost universally accepted. The Independence of irrelevant alternatives axiom has received a number of criticisms from several researchers. There is a little argument with the Non-dictatorship axiom, and it is also universally accepted (Ramanathan & Ganesh, 1994). These social axioms usually form an important guide when developing new combining criteria.

Next section, the MPDI criterion will be evaluated against these most widely accepted axioms.

5.1.6 Satisfying the social axioms

Here, in order verify that the MPDI can be used consistently to combine group preference and that it is comparable to other two formulas AM and GM, it is necessary to check it against the well-known social axioms, which constitutes the basis for evaluating group preferences combining criteria. Checking the MPDI against the three social axioms has revealed the following:

The first axiom: (**Universal domain**):

The MPDI satisfies this axiom, since for all possible individual experts' or ESs' numerical judgments within the pre-established range [0,10], the group preference pattern, which is the resulting combined value of MPDI, is defined, and there is a defined method to attribute this value into a decision class.

The second axiom: (**Pareto optimality**):

Pareto optimality axiom is the most important and widely recognized one. The MPDI satisfies this axiom, in that by similarity we have two decision alternatives: "Yes" and "No". If all the individual experts or ESs give all numerical judgments over 5; that is they all in favor of "Yes" alternatives, then MPDI will give a resulting combined value in favor of "Yes". Similarly, If all the individual experts or ESs gives all numerical judgments below 5; that is they all in favor of "No" alternatives, then MPDI will give a resulting combined value in favor of "No".

For instance, suppose that the following four ESs' judgments have been made: 3, 4, 2, 1. This means that all ESs have preferred the "No" decision option over the "Yes" decision. Then, the combined value utilizing the MPDI criterion is:
MPDI = 1.03 ("No")

Similarly, in the other direction, suppose the four expert systems judgments were: 8,6,9,10. Then, the combined value utilizing the MPDI criterion is:
MPDI = 34.56 (10) ("Yes")

In this way, the MPDI criterion always gives combined value that in favor of one alternative as long as all individual preferences were in favor of the same decision alternative.

<u>The fourth axiom:</u> (**Non-dictatorship**):

The MPDI also satisfies this axiom, since the resulted combined value takes into accounts fairly all individual preferences made, based on multiplication, and there is no individual expert or ES whose judgments dominates all others or has the right to dominates others; the group preference must be made taking into account the effect of all judgments. The group decision is based on mathematical computations defined by the MPDI.

Based on all the above results, the developed MPDI criterion is viable and can be consistently utilized in combining the crisp outputs of FESs, and can be used in general in the field of GDM to combine binary preferences, when adhering to the established meaningful scale.

*I have proposed the newly developed MPDI combining criterion in (Aly & Vrana, 2006c).

In order to more fairly prove the superiority of the MPDI criterion over other well-known classical ones like the AM, and GM, chapter 7 will be devoted for this purpose, where the performance of all these combining criteria and that of the MPDI will be compared against some non-biased datum level.

Next section, the weighted version of the developed MPDI will be presented.

5.1.7 The weighted multiplicative proportional deviative influence (WMPDI) criterion

The weighted version of MPDI is based on simple notion that as the relative importance's of the experts' or ESs' differ, then in this case their computed deviative influences are weighted to reflect the varying importance's in imposing influences; that is every FES influences the combined value based or proportional to its weight. Formally stated as follows:-

Let

O_m : the initially non - biased middle value, 5.

O_i : the i^{th} expert's crisp output, $i = 1,2,....n$.

w_i : the i^{th} expert's weight.

ΔO_i^+ : the absolute deviation of the i^{th} expert judgment O_i, above from the middle.

$$\Delta O_i^+ = \begin{cases} O_i - O_m & \text{if } O_i \geq O_m \\ 0 & \text{otherwise} \end{cases}$$

ΔO_i^- : the absolute deviation of the i^{th} expert judgment O_i, below from the middle.

$$\Delta O_i^- = \begin{cases} O_m - O_i & \text{if } O_i \leq O_m \\ 0 & \text{otherwise} \end{cases}$$

Then :

$$WMPDI = \left[\frac{\prod_{i=1}^{n} \left[1 + w_i (\Delta O_i^+ / O_m)\right]}{\prod_{i=1}^{n} \left[1 + w_i (\Delta O_i^- / O_m)\right]} \right] * O_m \tag{5.13}$$

Next section, a solution approach to the second case of the integration problem FESs will be presented.

5.2 Aggregating the crisp outputs of knowledge-unique FESs

In chapter 3, the integration problem was structured into two main problem cases, and the difference between the two cases was explained in details. The first case involves combining knowledge-equal FESs, whereas the second case involves aggregating knowledge-unique FESs. The first case problem was dealt with in the previous sections through adopting and developing combination criteria. In this section and up to the end of this chapter, the second problem case will be considered. Before presenting the proposed solution to this problem, let us summarize the basic idea of how to deal with such case problem. Knowledge-unique FESs involves multiple FESs each of which holds different specialized knowledge's and expertise's, and all these knowledge's and expertise's are necessary together in order to obtain a complete and comprehensive solution to the decision problem. Therefore, these different knowledge's and expertise's should be accumulated or aggregated in order to obtain such complete and comprehensive solution. Every FES represents a unique distinctive knowledge and expertise of a specialization area. Every FES produces a partial solution to the decision, and the complete and inclusive solution of such decision problem (i.e., Yes-or-No problem) is reached by aggregating the conclusions of all the FESs. Accordingly, the problem of integrating multiple knowledge-unique FESs should be referred to as aggregation more accurately than combination. This is the main concern of this section, how to accumulate or aggregate the outputs of knowledge-unique FESs.

Next section, two heuristics to aggregate the outputs of FESs will be presented. The first heuristic aggregates the outputs of FESs having different relative importance's to the decision problem. The second heuristic aggregates the outputs of FESs having equal relative importance's.

5.2.1 Aggregating the outputs of knowledge-unique FESs of different relative importance's

Given the numerical outputs of the knowledge-unique FESs, each of which has different weight, the problem is how to accumulate such crisp outputs. This requires understanding the role of each output value. Every crisp output gives a partial confidence about the final group decision. This is because every FES can not solely make the decision. Then, in order to have a full confidence about the final decision all these outputs should be added in such away that accumulate their decision directions. Using the concept of partial and complete, let us assume that the final complete output is evaluated by a total arbitrary scale similar and parallel to the individual scales of FESs' outputs. Thus, the minimum value of this total scale is zero, and means "No", and its maximum is selected to any positive number, and means "Yes". Then, each individual output among all FES's outputs takes one portion along this total scale, and the complete group output is to be reached through adding or accumulating these individual outputs. Therefore, if the resulted accumulated value was above the middle of this total scale, then the final decision is "Yes"; were it below the middle of the total scale, then the group decision is "No"; otherwise it is "Non-biased". The aggregation heuristic is described formally below.

<u>Step 1</u>: Apply AHP to obtain the weights or priorities of FESs:
$w_1, w_2, \ldots, w_i, \ldots, w_n$.

Where,
w_i: the weight of the i^{th} FES.
n : total number of the knowledge-unique FESs.

Step 2: Establish a total numerical scale from 0 to an arbitrary maximum chosen value, S_{max}, to represent the decisive degree between "Yes" and "No" decisions. The value 0 corresponds to "No", and the value S_{max} corresponds to "Yes". The middle value of such total scale is denoted S_{mid}.

Step 3: Apportion the total numerical scale established in the previous step into smaller numerical scales allocated to every FES in proportion to its computed priority, as follows:-

$$S_i = w_i * S_{\max} \tag{5.14}$$

Where,
S_i: is the maximum value of the output scale of the i^{th} FES.

Then, the crisp output of each i^{th} FES should be produced within the allocated numerical scale, from 0 to S_i.

Step 4: Given the crisp output of each FES, accumulate expertise's by summing all crisp outputs to given the finally aggregated group output, O_f (eq. 5.15):

$$O_f = \sum_{i=1}^{n} O_i \tag{5.15}$$

Where,
O_i: is the crisp output of the i^{th} FES.
Then, the final output is judged as either "Yes" or "No" as follows:-
If $O_f > S_{mid}$, then the final group decision is "Yes".
If $O_f < S_{mid}$, then the final group decision is "No".
If $O_f = S_{mid}$, then the final group decision is "Non-biased".

It should be noted that any attempt to use the AM or any averaging method for aggregating the outputs of knowledge unique FESs should be deemed as erroneous, because of the difference in logical notion behind knowledge-equal and knowledge-unique FESs. AM is not suitable to be used here, because it is basically a combining criterion which makes averaging for multiple numerical values playing the same role for the averaged variable, whereas the crisp outputs of the knowledge-unique FESs are playing different role to the final group decision. Consequently, the conversion of the notion of partial and complete solutions into summing or aggregation rather than averaging is the correct and accurate representation of this notion.

Next section, the above described aggregation heuristic will be adjusted for the sub-case of having equal weights for individual FESs.

5.2.2 Aggregating the outputs of knowledge-unique FESs of equal relative importance's

In this case, the previously described heuristic used in case of different weights is to be adjusted. The aggregation heuristic for equal weight case will directly begin from step 2, without computing the priorities, up to the end of the described heuristic, except that equation 5.14 is adjusted as follows:-

$$S_i = \frac{S_{\max}}{n}, \ \mathrm{i} = 1,2,....,\mathrm{n}. \qquad (5.16)$$

Where,
n: the total number of knowledge-unique FESs.

Equation 5.15 was adjusted to allow for the division of the total judgment scale equally among the individual output scales of the FESs.

*I have published the aggregation heuristic described above in ((Aly & Vrana, 2005c) & (Aly & Vrana, 2006a)).

Now, for the two case integration problems, adequate combination and aggregation methods have been adopted and developed. The next three chapters will concentrate on addressing the first type of the integration problem, the combination of knowledge-equal FESs. This is because of the complexity attributed to the potential conflict among the opinions of equal-knowledge FESs, which is a typical GDM problem and that needs to be more deeply investigated. Also, further investigation for the combination case will be conducted because up to now no exact rule exists that tells which combining criterion is the optimal to adopt and apply, and under which condition.

Next chapter, a new combining approach based on consensus evaluation will be presented, which aims to enhance the reliability and explanation capability of the group decision.

Chapter 6

Group consensus-analysis and consensus-based heuristics

In this chapter, a new approach for combining the crisp outputs of FESs will be presented. This approach is based on analysis of consensus relevant information existing within the set of crisp outputs produced by FESs. Group consensus-analysis through consensus indicators provides important insight and information about how to combine a group of numerical experts' judgments. It applies also to measure consensus level within the judgments of multiple FESs. This chapter is concerned with the development of a set of indicators to be applied in analyzing group consensus within the outputs of FESs, and to further develop consensus-based heuristics to combine such crisp outputs through reaching at consensus. These indicators should be adapted to handle the judgments of the FESs, according to the established meaningful numerical scale. Two different sets of similarity-based indicators for analyzing consensus will be presented and explained in this study. The first is based on configuring the set of indicators previously developed by (Ngwenyama et al., 1996). The second set is considered an improved one that does not rely on existence of known or desired similarity significance levels or thresholds. New measures of consensus, the standard deviation, percentage of class voting's, and sum of weights of class voting's will be introduced. All these different kinds of indicators are intended to capture more deep information from the set of numerical judgments in order to assist in making more reliable decision based on consensus analysis. The meaningfulness and differences among these indicators will be explained. The final aimed result of this chapter will be explanatory consensus-based heuristics for combining the crisp outputs of FESs. This will be achieved through pooling the information obtained by all such developed measures of consensus.

I shall first review the approach developed by Ngwenyama et al., to measure group consensus within experts' or ESs' numerical judgments. Then, I will configure such set of indicators to the situation of binary decision making problem, and according to the established scale of FESs' outputs. Then, an improved version of such indicators will be presented. The advantages offered by the new developed set will be described. After, a new set of consensus measures will be introduced. Then, all these developed indicators will be pooled into developing consensus-based heuristics exploiting all such relevant information. Finally, an example will be provided to demonstrate how the developed heuristics can applied.

6.1 Consensus indicators: an introduction

Consensus relevant information should constitute an important guide in combining or aggregating the decisions of multiple experts or ESs. This information provide a more clear picture about the differences and similarities among the decisions made by the multiple FESs, that can help on either developing a decision making procedure based on the evaluation of such information. The main intention of such analysis is to improve the decision quality and reliability of the finally reached consolidated decision.

Consensus is one major topic in GDM (Shih et al., 2004). Analyzing consensus has drawn considerable attention in the past; for instance see (Fedrizzi & Kacprzyk, 1988; Cook & Seiford, 1978). Bryson in 1996 (Bryson, 1996) considered the GDM problem in which every decision maker provides his opinion about a given set of decision alternatives or objects utilizing the AHP (Saaty, 1980) to obtain a preference vector or weight vector containing the weights of AHP ranking. Given such preference or weight vector of each decision maker, Bryson proposed a framework for assessing the current level of group consensus, and described a decision procedure for consensus building. In 1996, he and together with Ngwenyama et al. (Ngwenyama et al., 1996) proposed three indicators related to the level of agreement, and another three individual indicators related to the measure of the position of each individual relative the other group decisions.

Next section, the previously developed consensus analysis approach of Ngwenyama will be reviewed.

6.1.1 Previously developed consensus indicators

Ngwenyama et al. in 1996 described an approach to assess group consensus given a set of preference vectors of each decision maker belonging to the group. This preference vector can be in form of scores, ranks or weights of multiple decision options or alternatives. The preference vector: $V^i = (w_1, w_2, \ldots . w_n)$, denotes the vector of the i^{th} group member or decision makers, out of g members belonging the decision making group G, and assessing the ranks or weights of n alternatives. For instance, such weights can be obtained via the application of the AHP. Ngwenyama's have proposed six indicators:

(1) Group Strong Agreement Quotient (GSAQ): measures the level of agreement in within the decisions made by the group members.
(2) Group Strong Disagreement Quotient (GSDQ): measures the level of disagreement within the decisions made by the group members.
(3) Group Strongest Disagreement Indicator (GSDI): measures the breadth of decision opinions in the group.
(4) Individual Strong Agreement Quotient ($ISAQ^i$): measures for each i^{th} individual decision how much it has concordance with other members' made decisions.
(5) Individual Strong Disagreement Quotient ($ISDQ^i$): measures for each individual's decision how much it has dis-concordance with other members' made decisions.
(6) Individual Strongest Disagreement Indicator ($ISDI^i$): gives the ultimate disagreement of the i^{th} individual's decision with any one of other members' decisions. It helps identify which individual has greatest disagreement with any of the group members.

Next section, the similarity measure as adequate to the established numerical judgments scale for the binary decision problem will be explained, and the Ngwenyama's indicators will be accordingly configured.

6.2 Similarity measures for binary decision making and the adequate consensus indicators

In order to configure the Ngwenyama's consensus indicators for the binary decision making problem, first we need to define the adequate form of preference vector. All FESs should provide their crisp output numerical values within the range 0 to 10, which expresses the degree of bias either to toward either "Yes" or "No" decision answer. A binary preference vectors will consist of two components express the degree of bias toward either decision alternative. In order to be able to utilize and configure the Ngwenyama's indicators, the crisp judgments of FESs should be first converted in the form of binary vector. Then, if the i^{th} FES' output is 8; this means that it assigns 8 for "Yes" option and 2 for "No" option. Then, the i^{th} FES's preference vector will be $V^i = (8,2)$. Let us, then, utilize the terms of GDM and interchangeably use sometimes the word member or expert instead of FES. This is because the same approach can be used for evaluating consensus within a group of FESs, experts, or generally decision makers, providing that the utilized numerical scale is same. Mathematically, the decision made by every i^{th} member, belonging to the decision making group G, is represented by the preference vector $V^i = (x^i_1,x^i_2)$, $x^i_2 = 1- x^i_1$. Where:

V^i : preference vector of the i^{th} group's member.
x^i_1: score, rank, or priority of "Yes" decision option.
x^i_2: score, rank, or priority of "No" decision option.

Remark: In order to further preserve the possibility of using the similarity measures between vectors, I keep using the vector notation in spite that the binary preference vectors have only one independent component.

Ngwenyama et al. (1996) used the cosine angle between two vectors to express similarity between any two preference vectors. As the cosine of angle increases, the similarity or agreement increases. Then, the similarity between two preference vectors k and l, $Sim^{k,l}$, is mathematically formulated as:-

$$\mathrm{Sim}^{k,l} = \frac{V^k \bullet V^l}{\|V^k\| * \|V^l\|} \tag{6.1}$$

where :

$V^k \bullet V^l$: the scalar product of vectors V^k and V^l.

$\|V^k\|$: the magnitude of the i^{th} vector.

If $Sim^{k,l} = 1$, then the two vectors have the same direction, which corresponds to $\theta = 0^o$. If $Sim^{k,l} = 0$, then the two vectors are said to be completely dissimilar, which corresponds to $\theta = 90^o$, the largest possible angle between two vectors. The cosine of angle between two vectors that are dispersed in two different direction will be half-discounted (divided by 2) to reflect the difference in the two decision directions. Then, based on this measure of similarity, six indicators were developed by Ngwenyama et al. The six indicators were mathematically formulated as follows:-

(1) $GSAQ_\alpha$:

$$GSAQ_\alpha = \sum_{i=1}^{g-1} \sum_{j=i+1}^{g} 2 * \Gamma(i,j) \Big/ g(g-1) \tag{6.2}$$

$$\Gamma(i,j) = \begin{cases} 1 & \text{if } Sim^{i,j} \geq \alpha \\ 0 & \text{otherwise} \end{cases}$$

Where,
g: the total number of members in the group G of decisions.

(2) GSDQ$_\delta$:

$$GSDQ_\delta = \sum_{i=1}^{g-1}\sum_{j=i+1}^{g} 2 * \Phi(i,j) \Big/ g(g-1) \tag{6.3}$$

$$\Phi(i,j) = \begin{cases} 1 & \text{if } Sim^{i,j} \le \delta \\ 0 & \text{otherwise} \end{cases}$$

(3) GSDI:

$$GSDI = \min_{\forall i,j \in G, i \neq j} \left\{ Sim^{i,j} \right\} \tag{6.4}$$

(4) ISAQ$^i_\alpha$:

$$ISAQ^i{}_\alpha = \sum_{j=1, j\neq i}^{g} \Gamma(i,j) \Big/ (g-1) \tag{6.5}$$

$$\Gamma(i,j) = \begin{cases} 1 & \text{if } Sim^{i,j} \ge \alpha \\ 0 & \text{otherwise} \end{cases}$$

(5) ISDQ$^i_\delta$:

$$ISDQ^i{}_\delta = \sum_{j=1, j\neq i}^{g} \Phi(i,j) \Big/ (g-1) \tag{6.6}$$

$$\Phi(i,j) = \begin{cases} 1 & \text{if } Sim^{i,j} \le \delta \\ 0 & \text{otherwise} \end{cases}$$

(6) ISDI^i:

$$ISDI^i = \min_{\forall j \in G, j \neq i} \left\{ Sim^{i,j} \right\} \tag{6.7}$$

Based on the above described consensus indicators, an important composed indicator of for individual decision in the group is the individual consensus vector, which is a whole characteristic of individual decisions. It is denoted as ICV^i, ICV^i = ($ISAQ^i$, - $ISDQ^i$, $ISDI^i$), which is the i^{th} individual vector. This vector contains important information about the status of each decision relative to other group decisions, and was first defined by Ngwenyama et al.. It tells whether this decision is agreeing or disagreeing with other group decisions. This leads to that there are two distinct decisions in the group. The individual with the best ICV is that one, who has ever maximum agreement and minimum disagreement with other individual decisions in the group. I denote this best ICV individual decision as ICV^b. In contrast, the individual which has ever minimum agreement and maximum disagreement with other group decisions is the decision with worst ICV, and is denoted as ICV^w. These decisions are of special importance because they identify which decision is the most agreeing with others, and which one is the most disagreeing. This provides additional necessary information in attempting to reach at consensus. From their definition, they can be stated as follows:-

$ICV^b = ({}^{max}ISAQ, {}^{min}ISDQ, {}^{max}ISDI)$
$ICV^w = ({}^{min}ISAQ, {}^{max}ISDQ, {}^{min}ISDI)$

Where:
${}^{max}ISAQ$: is the maximum value of $ISAQ^i_\alpha$, $\forall$ i $\in$ G.
${}^{min}ISAQ$: is the minimum value of $ISAQ^i_\alpha$, $\forall$ i $\in$ G.
${}^{min}ISDQ$: is the minimum value of $ISDQ^i_\delta$, $\forall$ i $\in$ G.
${}^{max}ISDQ$: is the maximum value of $ISDQ^i_\delta$, $\forall$ i $\in$ G.
${}^{max}ISDI$: is the maximum value of $ISDI^i$, $\forall$ i$\in$ G.
${}^{min}ISDI$: is the minimum value of $ISDI^i$, $\forall$ i$\in$ G.

Based on the definition of the configured indicators, it holds that the value of ${}^{max}ISAQ$ is always greater than or equal to the value of $GSAQ_\alpha$.

It should be noted that the chosen values for significance levels α, δ, influentially determine the values of 4 indicators: $GSAQ_\alpha$, $GSDQ_\delta$, $ISAQ^i_\alpha$ and $ISDQ^i_\delta$. Ngwenyama's have suggested two possible values for α and δ; 0.985 for α, which corresponds to cosine of 10 ^{o}, and surprisingly 0.966 for δ, which corresponds to cosine of 15^o. Unfortunately, there is no clear rule to help set the values for α, δ, or describe a relation between these values and indicators' values. Moreover, the threshold δ is set un-logically very high, without any justification. Actually, setting the values of these two thresholds exhibits some form of vagueness associated with optimizing the values of α and δ. These are considered apparent limitations of Ngwenyama's approach. Reader may refer to such article (Ngwenyama et al., 1996) for more detailed explanation. However, generally, setting the values of α and δ depends on the analyst's vision about which value of similarity can be considered a threshold of strong agreement or disagreement. Another limitation is that because there are two threshold levels for agreement and disagreement, then some similarities can be found that neither satisfy the α's threshold nor satisfy the δ's threshold. Thus some similarities can be classified neither as agreement nor as disagreement, and this introduces another source of ambiguity and information loss to the final decision solutions. Also, it is not logical that some similarity value is not classified as agreement or disagreement. One solution to this dilemma is to redefine the above indicators using only one level or threshold of agreement and disagreement, α. Therefore, if the similarity value exceeds α, then it is considered agreement; if it is below α, then it is considered disagreement. The above six indicators can be redefined only in terms of α, and in this case, the following formulas hold:-

$GSDQ_\alpha = 1- GSAQ_\alpha$, and $ISDQ_\alpha = 1- ISAQ_\alpha$.

However, still some source of vagueness associated with setting an adequate value for α. therefore, in order to get ride of all these limitations and vagueness, one solution is to make the set indicators independent of such significance threshold values. For this purpose, a new set of similarity-based indicators is developed. This set of indicators considered a modified version of the previously developed one. These new indicators offer more advantages and fewer limitations than the previous one, in that the values of these indicators do not depend on threshold values, and they contain more detailed information about the pair-wise similarities.

In next section, the modified set of indicators will be introduced.

6.3 The modified indicators

I have developed new six consensus indicators that are independent of significance threshold values of agreement and disagreement. This modified set of indicators is extension of the Ngwenyama's set of indicators, and is based on the average similarity rather than its counts. This is to avoid the difficulty and vagueness associated with the determination of the threshold values. Similarly, the first three are related to group level of agreement and breadth of opinions. The last three indicators measure the position of each individual decision relative to the others. The modified set of indicators is introduced as follows:-

(1) GASI (Group Average Similarity Indicator):
This is the average similarity of all pairs of group members, which measures the level of agreement within the group. As this indicator value reaches 1, this means complete agreement, and as it reaches 0, it means complete disagreement. As the value of GASI increases over 0.5, this means that agreement level within the group of decisions are more than disagreement level in them. Mathematically expressed as follows:-

$$GASI = \sum_{i=1}^{g-1} \sum_{j=i+1}^{g} 2 * Sim(i,j) \Big/ g(g-1) \qquad (6.8)$$

Where,
g: the total number of members in the group G.

(2) GADI (Group Average Dissimilarity Indicator):
This is the average dissimilarity of all pairs of group members, which measures the level of disagreement in the group.

$$GADI = \sum_{i=1}^{g-1} \sum_{j=i+1}^{g} 2 * [1\text{-}Sim(i,j)] \Big/ g(g-1) \qquad (6.9).$$

Or simply, *GADI= 1-GASI*

(3) GMDI (Group Maximum Dissimilarity Indicator):
It measures the breadth of opinions in the group. It is the minimum magnitude of similarity throughout all pairs of similarities.

$$GMDI = \underset{\substack{\forall i,j \in G, \\ i \neq j}}{Min} \{Sim(i,j)\} \qquad (6.10)$$

(4) IASI^i (Individual Average Similarity Indicator):
It measures the average amount of agreement every i^{th} individual decision has with other group members.

$$IASI^{i} = \sum_{j=1, j \neq i}^{g} Sim(i,j) \Big/ (g-1) \qquad (6.11)$$

The individual who has the highest value of IASI^i is said to have maximum agreement with other group members, its value is denoted IASI*. Its associated output value is denoted (arg

$IASI^*$). It has been found that this group member always is the median in case of odd group members' number, g. The individual who has lowest value of $IASI^i$, denoted as $IASI^-$, is said to have minimum agreement with other group members, and its associated output value is denoted as (arg $IASI^-$).

(5) $IADI^i$ (Individual Average Dissimilarity Indicator):
It measures the average amount of disagreement, every i^{th} individual decision has with other group decisions.

$$IADI^i = \sum_{j=1, j\neq i}^{g} \left[1\text{-}Sim(i,j)\right] \Big/ (g-1) \tag{6.12}$$

Or simply, $IADI^i = 1 - IASI^i$

(6) $IMDI^i$ (Individual Maximum Dissimilarity Indicator):
This indicator gives the minimum level of agreement each i^{th} individual has with other group's decisions.

$$IMDI^i = \underset{\substack{\forall j \in G, \\ j \neq i}}{Min} \{Sim(i,j)\} \tag{6.13}$$

Based on the definition of the proposed modified indicators, it holds that the value of $IASI^*$ is always greater than or equal the value of GASI.

Similarly as it was defined for the previous set of consensus indicators, for this set of average-based indicators, let us define the individual consensus vector. Conveniently, we can denote the individual consensus vector based on average-based indicators as AICV, in order to distinguish it from original ICV used for other set. The individual consensus vector $AICV^i$ is reduced into only two distinct components; $AICV^i = (IASI^i, IMDI^i)$, since $IADI^i = 1 - IASI^i$. The individual, who has a non-dominated AICV is the decision who has ever maximum agreement and minimum disagreement with other individual decisions in the group. We denoted this individual decision which has best AICV as $AICV^b$. In contrast, the individual which has ever minimum agreement and maximum disagreement with other group decisions is the decision with worst AICV, and denoted as $AICV^w$. These decisions are of special importance because they identify which decision is the most agreeing with others, and which one is the most disagreeing. This provides necessary information in attempting to reach at consensus. From their definition, they can be stated as follows:-

$AICV^b = ({}^{max}IASI, {}^{max}IMDI)$
$AICV^w = ({}^{min}IASI, {}^{min}IMDI)$

Where:
${}^{max}IASI$: is the maximum value of $IASI^i$, $\forall i \in G$, ${}^{max}IASI = IASI^*$.
${}^{min}IASI$: is the minimum value of $IASI^i$, $\forall i \in G$, ${}^{min}IASI = IASI^-$.
${}^{min}IMDI$: is the minimum value of $IMDI^i$, $\forall i \in G$.
${}^{max}IMDI$: is the maximum value of $IMDI^i$, $\forall i \in G$.

These indicators' values will be used as guides in leading consensus-based decision making procedure or heuristics toward combining the FESs' outputs.

It should be noted that the above introduced new set of indicators, which does not require specification of thresholds for agreement or disagreement, provides a relief from previously discussed vagueness of Ngwenyama's indicators. In addition, this new set takes account of all values of similarity not only the count of it, as it was previously.

*I have published this modified set of indicators along with the configured one of Ngwenyama in (Aly & Vrana, 2006d).

Next section, other consensus measures of consensus will be introduced.

6.4 New measures of group consensus

In this section, three measures of consensus will be presented. Two measures are related to the consensus level within the set of FESs' outputs, and a third measure is related to the set of weights associated with these outputs. The first measure is the standard deviation of judgments, and the second measure is the percentage of voting's received by a decision class or option.

Next section, I will explain how the standard deviation could assist in measuring consensus.

6.4.1 The standard deviation as a measure of consensus

Another measure of consensus based on statistical characteristics of incoming experts' numerical judgments is the standard deviation, which measures the amount of dispersion from the center of group of numerical values. The formula of standard deviation, defined for a group of experts' or FESs' numerical judgments, is as follows:-

$$\sigma = \sqrt{\frac{\sum_{i=1}^{g}(\mathrm{x}_{\mathrm{i}} - \overline{\mathrm{x}})^2}{\mathrm{g}-1}} \tag{6.14}$$

where :

x_{i} : the i^{th} expert's numerical judgment, i = 1,2,.......g.

$\overline{\mathrm{x}}$: the arithmetic mean of the experts' judgments.

g : the total number of experts in the judging group.

The standard deviation (σ), as a measure of dispersion, provides new information about the consensus level of the group. It differs from GASI, in that GASI quantify the average of total pair-wise similarity of the group, whereas σ measures the dispersion from the center or mean of the group.

It should be noted that the both measures of consensus, GASI and σ, give two different but related information about the consensus level, but do not give information about the consensus related to one of the two binary decision directions; that is both GASI and standard deviation do not tell which direction of decision answer, either "Yes", or "No" has which degree of consensus. So, for this purpose another third measure of consensus is to be introduced next section; that is the percentage of voting's of a decision class.

6.4.2 The percentage of voting's as a measure of consensus

In order to get information about how much degree of agreement each decision class or option has received, a third type of consensus measure will be used. This measure is the percentage of voting's for each decision class. This measure is simple and has been utilized frequently in the social choice and voting theories. It is a derived value from the Majority Voting rule. In order to compute the value of this indicator, first the given FESs' crisp outputs should be attributed to either decision class; that is if the value of this output was above the middle (> 5.5), then it is attributed to "Yes" decision direction; if it is below the middle (< 4.5), then it is attributed to "No" decision direction, otherwise it is attributed to non-biased class. The utilization of limits or threshold ± 0.5 around 5 aims to insure that the decision answer is well classified away from the non-biased class. After attribution of outputs to the decision classes, three indicators of consensus based on percentage of voting's can now be defined. They are three values:-

- Percentage "Yes" voting (%YV).
- Percentage "No" voting (%NV).
- Percentage "Non-biased" voting (%NBV).

High value of (%YV) means that there is a considerably high degree of consensus or voting's dominance assigned to "Yes" decision option. Similarly, high value of % NV and %NBV indicates a high degree of consensus in the corresponding directions.

Next section, another measure of consensus related to the weights of the FESs will be defined.

6.4.3 The sum of weights of voting's as a measure of consensus

Given the crisp outputs of FESs attributed to each decision class, the weight of voting's received by each class is then computed and based on this measure, three indicators are:-

- Sum of weights of "Yes" voting's (SWYV).
- Sum of weights of "No" voting's (SWNV).
- Sum of weights of "Non-biased" voting's (SWNBV).

High value of SWYV means that there is a considerably high degree of weight dominance level that imposes undertaking the "Yes" direction, which has been voted by the most important sub-group of the FESs, and so on for SWNV, and SWNBV. These groups of indicators are useful in detecting whether there is high degree of weight dominance level in the set of weights; that is, in other meaning, these indicators can tell whether there exists some decision class that has been favored by some influential subset of the FESs. The weights expressing relative importance's of individual FESs are to be computed utilizing the approaches introduced in chapter 4.

*I have proposed the utilization of these three measures of indicators in ((Aly & Vrana, 2006d) & (Aly & Vrana, 2006e)).

Next section, I shall use some of these measures along with the other modified similarity-based indicators to guide two consensus-based heuristics toward combining the crisp numerical outputs of the FESs into a finally consolidated decision.

6.5 Consensus-based heuristics

The development of consensus indicators was aimed to produce helpful information that can assist in making more reliable and explanatory decision than that obtained with only using blind combining criteria. Now, the information extracted from the set of FESs' outputs via several consensus indicators should be pooled into developing a decision making heuristic that exploit all these information in reaching at consensus, or in another meaning combine the outputs of FESs. Up to now, I have introduced three different sets of indicators measuring consensus. The first set consists of the similarity-based indicators. The modified version of these indicators which are based on the average are more advantageous than the old ones developed by Ngwenyama for the reasons mentioned before in sections 6.2 and 6.3. Other different measures of consensus degree are the percentage of voting's and the sum of voting's weight indicators that can be used to decide whether there exists some high level of majority or importance in favor of one alternative or class. They can conveniently be named voting's and weights' dominance level indicators, in order to distinguish them from other developed similarity-based consensus indicators. In the incoming sections, I will propose two consensus-based heuristics having slightly different decision making strategies, and guided by these three different measures of consensus. It will be shown in each heuristic, how the information embedded in the values of indicators could be used to reach at consensus. Two consensus-based heuristics will be introduced. The two heuristics utilize the modified version of the similarity-based consensus indicators, and the other two measures, percentage of voting's, and sum of voting's weights. In addition, the Arithmetic Mean (AM) combining criteria will be also used as another guide within the two heuristics, especially whenever there is no considerable degree of consensus to rely upon in making a reliable decision. It is also used when the agreement is found to be in favor of the "Non-biased" option; in this case this option is avoided as much as possible by applying the AM. If the application of AM still gives this option, then it is the final group decision. A direction threshold $\pm$ 0.5 is utilized to increase the reliability of identifying the decision obtained by the AM. This means that if the value of AM is greater than or equal to 5.5, then the direction is 10 or the "Yes" decision answer. If the value of AM is less than or equal to 4.5, then the direction is 0 or the "No" decision answer. Otherwise it is "Non-biased". The first heuristic uses the direct information obtained from the indicators values to tell what the direction of correct decision answer (DCA). The second heuristic reaches at consensus in slightly different strategy through employing consensus facilitation.

Next section, the first heuristic will be presented.

6.5.1 Consensus-based heuristic for eliciting the direction of the correct decision answer

In this section, a consensus-based heuristic is described that is guided by a subset of the consensus indicators introduced before. The inputs to the heuristic are the crisp numerical outputs representing the judgments of individual FESs. Given these inputs, the heuristic attempts to elicit the direction of correct answer (DCA), "Yes", or "No", based on the information extracted from the consensus analysis. The proposed heuristic relies mainly on the values of the two defined consensus indicators: GASI, the measure of group agreement level, and $IASI^*$, the maximum value of individual agreement indicator, $IASI^i$. These two indicators involves two other dependent indicators, GADI and $IADI^*$, respectively, since GADI = 1 – GASI, and $IADI^i$ =1- $IASI^i$. The other two remaining indicators, GMDI and $IMDI^i$ measure neither the agreement level, nor the position of individual decisions relative to others. Consequently, the proposed heuristic will not rely on the information obtained from these two indicators but rather on the information telling what the level of agreement is, and which individual decision has maximum agreement with other group decisions. This information is found in the values of the two indicators, GASI and IASI*. In addition to these two similarity-based indicators, the heuristic exploit also the information of two other consensus measures, the percentage of voting's (%V), and the sum of weights of voting's (SWV). The use of all these indicators is aimed to detect different forms of consensus in order to adhere to such agreement in making the decision. The values of these indicators constitute an explanatory guide toward making more reliable and understood group decision.

The heuristic begins with specifying the acceptable levels of consensus indicators values, and threshold for percentage of voting's and sum of voting's weights. Default acceptation levels for GASI, IASI*, are: 0.75, 0.75. These indicators' default acceptable values were suggested in (Bryson, 1996) as the three fourth majority is widely acceptable ratio. Nevertheless, I have another opinion regarding specifying acceptable levels for these guiding indicators. According to the nature of developed indicators, I suggest a change in the value of GASI, from 0.75, to 0.5. The acceptable value set for IASI* remains at 0.75 as was suggested by (Bryson, 1996). The rationale for selecting these values is that it is enough that half of the decisions in the group are agreeing. Also, the value 0.75 of IASI* guarantees that the decision which has IASI*, arg IASI*, is agreeing at least with three fourth of the group decisions. However, the specification of indicators acceptable levels should depends on the vision of the decision analyst about which levels of agreement and disagreement are sufficient to make a reliable decision based on consensus. This also depends on the characteristics of the problem and the participating ESs. Any other values of the acceptable levels can be specified as convenient. The value 0.75 of GASI can be also used, and its effect is to make the heuristic stricter in assessing the level of agreement. After specifying the acceptable levels for the similarity-based indicators, the the FESs' outputs or voting's are attributed to the possible decision classes, "Yes", "No", "Non-biased", and a preliminary check is made to determine whether these exist dominance level in the voting's or in the weights. If so, the heuristic is to be terminated and the decision is made, in which the dominant decision class is undertaken. If not, the heuristic proceeds relying after on the values of two similarity-based indicators, GASI and IASI*.

The rationale and decision making philosophy for the proposed heuristic is that as long as the acceptable levels of indicators GASI and IASI* are satisfied (GASI $\geq$ 0.5, and IASI$^*\geq$ 0.75) , then we can make reliable decision based on consensus, in that the decision which has IASI* will be selected to point to the group decision, since this individual decision exhibits maximum agreement with other decisions especially in case of acceptable consensus degree (GASI $\geq$ 0.5). Then, the DCA should be that to which decision of IASI* points (i.e., arg IASI*); that is if the arg IASI* > 5.5 then the DCA is 10; if the arg IASI* < 4.5 then the DCA is 5; otherwise DCA is 5 ("Non-biased"). If the degree of consensus GASI is lower than the acceptable level, then the DCA is judged by arithmetic mean (AM) by using ± 0.5 direction threshold. It should be noted that if the value of IASI* satisfy the acceptable level value, and arg IASI* happens to be equal 5, then again DCA is judged by the AM. The heuristic algorithm is formally stated as follows:-

Step 0: Initializing: specify the acceptance level values for GASI and IASI*. Default values are: 0.5 and 0.75 respectively.

Step 1: Attributing outputs: every FES's judgment, O_i, is attributed to one of three classes or consensus sub-group depending on whether or not this value is above, below or at the middle value of the used scale, 5 ± 0.5:

Condition 1.1: if $O_i > 5.5$, then O_i is attributed to "Yes" class.
Condition 1.2: if $O_i < 4.5$, then O_i is attributed to "No" class.
Condition 1.3: if $4.5 \leq O_i \leq 5.5$, then O_i is attributed to "Non-biased" class.

Step 2: Preliminary checking: the heuristic is to be terminated because of either high voting's or weights dominance levels under the following two conditions:-

Condition 2.1: if max{%YV, %NV, %NBV} ≥ 75 %, then there is a high degree of voting's dominance level and the DCA is given by the class argument of max{%YV, %NV, %NBV}(arg max{%YV, %NV, %NBV}). Stop.

Condition 2.2: if max {SWYV, SWNV, SWNBV} ≥ 0.75, then a high degree of weighting dominance level and the DCA is given by the class argument of max{SWYV, SWNV, SWNBV}(arg max {SWYV, SWNV, SWNBV}). Stop.

Note: arg max {SWYV, SWNV, SWNBV}or arg max{%YV, %NV, %NBV}gives the decision class which has either high voting's or weights dominance levels respectively.

Step 3: Computing indicators: computing the values of consensus indicators: GASI and $IASI^*$. Find the decision with $IASI^*$, arg $IASI^*$.

Step 4: Testing and decision making: comparing the value of GASI to 0.5 acceptance level. There are three cases:-

Condition 4.1: If GASI ≥ 0.5, and $IASI^* \geq 0.75$ (It is usually guaranteed that $IASI^* >$ GASI) then the DCA is that is pointed to by arg $IASI^*$; except for if arg $IASI^*$ happens to be equal 5, then the DCA is that is pointed to by AM using ± 0.5 direction threshold.

Condition 4.2: If GASI ≥ 0.5, but $IASI^* < 0.75$, then the DCA is that pointed to by AM using ± 0.5 direction thresholds.

Condition 4.3: If GASI ≤ 0.5, then the DCA is that pointed to by AM using ± 0.5 direction thresholds.

*I have published the heuristic described above in (Aly & Vrana, 2006c).

Next section, a more strict heuristic will be introduced which employ consensus facilitation procedure in order to reach at consensus.

6.5.2 A consensus-based heuristic employing a consensus facilitation procedure

This section presents another consensus-based heuristic, which differs from the previously explained one in that it employs a consensus facilitation procedure similar to that proposed by (Ngwenyama et al, 1996). This consensus facilitation procedure is conducted, whenever the acceptable values of indicators are not satisfied, but there is a hope to realize consensus. In this consensus facilitation, a problematic decision is identified, which has the worst individual consensus vector, $AICV^w$. Then, the indicators values are recomputed without considering this problematic decision. If the removal of this element results in an increase in the value of GASI, then in this case this individual is actually removed from the group of decisions. This process continues until the occurrence of two conditions; either the acceptable indicators' values are reached, or the next decision to remove belongs to the last remaining subgroup, in which all decisions belong to the same decision class. This because the agreement within this last subset of decision is already existing then no need to remove additional decision from this remaining subgroup. The details of this procedure will be described in the formal steps of the heuristic. The heuristic is same like the first one up to step 3, and utilize same acceptable values for all used consensus indicators and measures. The difference between the two heuristics is in step 4. In the second heuristic, step 4 includes consensus facilitation procedure, which attempts to reach at consensus through reaching the indicators' acceptable levels, whereas in the first heuristic, no such procedure exists and the consensus information is assessed only to find whether it is possible to make a reliable collective decision based on agreement or not. The heuristic is described below.

Step 0: Initializing: specify the acceptance level values for GASI and $IASI^*$. These values are set: 0.5 and 0.75 respectively.

Step 1: Attributing outputs: every numerical judgment of each FES, O_i, is attributed to one of three classes or consensus sub-group depending on whether or not this value is above, below or at the middle value of used scale, 5 ± 0.5:

Condition 1.1: if $O_i > 5.5$, then O_i is attributed to "Yes" class.
Condition 1.2: if $O_i < 4.5$, then O_i is attributed to "No" class.
Condition 1.3: if $4.5 \le O_i \le 5.5$, then O_i is attributed to "Non-biased" class.

Step 2: Preliminary checking: the heuristic is to be terminated because of either high voting's or weights dominance levels under the following two conditions:-

Condition 2.1: if max{%YV, %NV, %NBV} ≥ 75%, then a high degree of voting dominance level and the DCA is given by the class argument of max{%YV, %NV, %NBV}(arg max{%YV, %NV, %NBV}). Stop.

Condition 2.2: if max {SWYV, SWNV, SWNBV} ≥ 0.75, then a high degree of weighting dominance level and the DCA is given by the class argument of max{SWYV, SWNV, SWNBV}}(arg max {SWYV, SWNV, SWNBV}). Stop.

Step 3: Computing indicators: computing the values of consensus indicators: GASI and $IASI^*$. Find the decision with $IASI^*$, arg $IASI^*$.

Step 4: Testing and decision making: comparing the value of GASI to 0.5 acceptance level. There are three cases:-

Condition 4.1: if GASI ≥ 0.5 and $IASI^* \ge 0.75$. In this case, the DCA is that is pointed to by arg $IASI^*$; except for if the arg $IASI^*$ happens to be equal 5; in this case the DCA is pointed to by the AM using ± 0.5 direction thresholds.

Condition 4.2: If GASI $\geq$ 0.5, but IASI* < 0.75, then consider consensus facilitation procedure:-

Step 4.2.1: Based on the information that has been gotten from step 1, divide the individual decisions into three consensus subgroups when applicable ("Yes", "No", and "Non-biased").

Step 4.2.2: Consider removal of the first decision that has the worst AICV, AICV^w = (^{min}IASI, ^{min}ISDI). If such decision does not exist, then consider the decision with ^{min}IASI. This decision is called "problematic decision". Re-compute the consensus indicators. Remove the decision only if the following condition occurs:-

Condition 4.2.2.1: there is an increase in GASI(t) (i.e. GASI(t) $\geq$ GASI(t-1)) (where GASI(t) is the agreement at facilitation trial t).

Stop the consensus facilitation procedure until occurrence of one of the following two conditions:

Condition 4.2.1: The next individual decision to remove is the first element in the last consensus sub-group. Here a check is to be made; if all indicators values were satisfying the acceptable values; that is GASI(t) $\geq$ 0.5, and IASI*(t) $\geq$ 0.75, then go to condition 4.1; otherwise the DCA is pointed to by the AM using $\pm$ 0.5 direction thresholds.

Condition 4.2.2: GASI(t) $\geq$ 0.5, and IASI*(t) $\geq$ 0.75. In this case, the DCA is that is pointed to by the current arg IASI*(t), except for if arg IASI*(t) is happens to be equal 5, in this case AM is used with $\pm$ 0.5 direction thresholds.

Condition 4.3: If GASI < 0.5, then in this case the DCA is pointed to by the AM) using $\pm$ 0.5 direction threshold.

The rationale of the above proposed heuristic is similar to that of the first heuristic. The first three steps remained unchanged. Then, the heuristic proceeds as follows: condition 4.1 is executed if the values of GASI and IASI* were satisfying the acceptable levels. Condition 4.2 tells that if the level of agreement, GASI, at least is greater than or equal to the disagreement level, GADI; that is when GASI $\geq$ 0.5, but there is no such decision who has 0.75 agreements with other decisions, then a consensus facilitation procedure is conducted in order to reach the acceptable levels of these indicators. This facilitation procedure is terminated when the acceptable values of indicators are reached or the last homogenous decisions subgroup is remaining as a result of decisions removal. For low level of agreement identified; that is when GASI is less than 0.5, which means that the disagreement in the decisions group is more than the agreement, in this case little hope is beard to realize consensus and the decision based on the value of AM is to be made.

Up to now, two consensus-based heuristics have been presented. The two heuristics are similar but have slightly different strategies. The first heuristic exploits the consensus information to determine whether it is possible make a final group decision based on agreement; if it is possible the AM combining criteria is used to determine the DCA. The second heuristic make more attempts to realize consensus through consensus facilitation. The two heuristics utilize acceptable values for the consensus indicators to detect agreement, and consequently make more reliable and explanatory decision accompanied with the associated degree of consensus.

It should be noted that higher values of the indicators GASI and IASI* could be utilized, and its affect is to make the two heuristics stricter in judging degree of agreement within the group and between of each individual with other group decisions. Consequently, the two heuristics are adaptive through varying these acceptable values of such two indicators and that for the other consensus measures. Setting these values should be subjected to the viewpoint of the decision analysts, and to be particularly based on the inherent nature and characteristic of the decision problem in hand and that of the combined FESs.

* I have published the above two heuristics in (Aly & Vrana, 2006e).

The illustrative example in the next section, will demonstrate how both heuristics could be utilized.

6.6 An illustrative example

Now, an example will demonstrate how the two proposed consensus-based heuristics could be applied. The following decision problem will be judged using the two heuristics.

Suppose that five FESs of equal-knowledge are involved in the decision making transaction in hand, and it is required to reach at consensus or in other meaning to obtain a finally consolidated group decision. The crisp outputs and the weights are given below:-

$O_1 = 9$, $O_2 = 3$, $O_3 = 6$, $O_4 = 8$, $O_5 = 2$.
$w_1 = 0.2$, $w_2 = 0.1$, $w_3 = 0.4$, $w_4 = 0.05$, $w_5 = 0.25$.

Now, I will show how each heuristic of the proposed two heuristics will reach at consensus and what will be the DCA.

<u>Applying the first heuristic</u>:-

Step 0: Initializing:
The values of GASI, IASI* are set to: 0.5 , 0.75, respectively.

Step 1: Attributing outputs:
Condition 1.1: "Yes" voting class = $\{O_1, O_3, O_4\}$.
Condition 1.2: "No" voting class = $\{O_2, O_5\}$.
Condition 1.3: "Non-biased" voting class = $\{\varphi\}$.

Step 2: Preliminary checking:
The percentages of voting's are:
%YV = 3/5 = 0.6, %NV = 2/5 = 0.4 , %NBV = 0/5 = 0.
The sums of weights of voting's are:
SWYV = 0.65, SWNV = 0.05, SWNBV = 0.6

Then, checking to find whether or not there exist dominance levels:
Condition 2.1: max{%YV, %NV, %NBV}= 0.6 < 0.75, then there is no detected high degree of voting's dominance level. Proceed.
Condition 2.2: max{SWYV, SWNV, SWNBV}= 0.65 < 0.75, then there is no detected high degree of weighting dominance level. Proceed.

Step 3: Computing indicators:
Similarity values between all pairs of decision outputs are computed using equation 6.1, and the individual consensus vectors, $AICV^i$, for all decisions are computed using equation 6.8 through 6.13. The computed values are shown in table 6.1 and table 6.2, respectively.

Table 6.1 Computed similarity values for all pairs.

Pair: (i,j)	(1,2)	(1,3)	(1,4)	(1,5)	(2,3)	(2,4)	(2,5)	(3,4)	(3,5)	(4,5)
Sim(i,j)	0.247	0.89	0.99	0.174	0.419	0.303	0.987	0.942	0.37	0.235

Table 6.2. The computed individual consensus vectors.

Decision (i)	**$IASI^i$**	**$IMDI^i$**
1	0.575	0.247
2	0.489	0.247
3	0.66	0.37
4	0.62	0.235
5	0.442	0.174

As shown in table 6.2, decision number 3 (O_3 = 6) has $AICV^3$ that dominates that of other decisions. Then, the needed indicators' values to apply the heuristic are:-
GASI = 0.556, $IASI^*$ = 0.66 (arg $IASI^*$ = 6).

Step 4: Testing and decision making:
Condition 4.2 applies: GASI $\geq$ 0.5, but $IASI^*$ < 0.75, then the DCA is that is pointed to by AM using ± 0.5 direction thresholds:

AM = (9 + 3+6 + 8 + 2)/5 = 5.6 > 5.5. Then, the DCA is 10, or "Yes". Stop.

Here, it should be noted that only I resort to the AM, when the consensus heuristic fails to detect acceptable level of consensus. This confines the use of the AM to only this circumstance.

<u>Applying the second heuristic</u>:-

Step 0, Step 1 and Step 2, Step 3:
All these steps yield the same results. So, we proceed to step 4.

Step 3: Computing indicators:
The similarities between all pairs of decisions and the individual consensus vectors are same as in table 6.1 and table 6.2, respectively.
Decision number 3 (O_3 = 6) has $AICV^3$ that dominates that of all other decisions. The required indicators values are: GASI = 0.556, $IASI^*$ = 0.66 (arg $IASI^*$ = 6)

Step 4: Testing and decision making:
Condition 4.2 applies: GASI $\geq$ 0.5, but $IASI^*$ < 0.75, then consider consensus facilitation procedure:-

Step 4.2.1: Attribution of the outputs to each decision class has been already performed in step 1 in the application of first heuristic:
"Yes" voting class = $\{O_1, O_3, O_4\}$.
"No" voting class = $\{O_2, O_5\}$.
"Non-biased" voting class = $\{\varphi\}$.

Step 4.2.2: first trial: t =1: considering the removal of the first decision that has the worst AICV, $AICV^w = (^{min}IASI, {}^{min}ISDI)$:
Decision number 5 (O_5 = 2), has the worst $AICV^i$: $AICV^2 = AICV^w$ = (0.442, 0.174). Then, the consensus indicators, and all $AICV^i$ are to be recomputed as follows (table 6.3):

Table 6.3 The individual consensus vectors re-computed.

Decision (i)	**$IASI^i$**	**$IMDI^i$**
1	0.709	0.247
2	0.323	0.247
3	0.75	0.419
4	0.745	0.303

GASI(1) = 0.63 > GASI (0) > 0.556, $IASI^*(1)$ = 0.75 (arg $IASI^*$ = 6)

After re-computing the indicators' values, a check is made to find whether or not the acceptable values of the indicators have been reached. The result is:

Condition 4.2.2 applies: GASI(1) $\geq$ 0.5, and $IASI^*(1) \geq 0.75$. In this case, the DCA is that is pointed to by the current arg $IASI^*$, arg $IASI^*(1)$ = 6, then the DCA is "Yes", 10.

Same result has been obtained through applying the two heuristics. Both heuristics have referred to the same direction opinion "Yes". However, there were different paths employed by each heuristic. In the first heuristic, the agreement level was not sufficient to make a final decision based on consensus, so the resort was to the AM criterion to solve conflict. The second heuristic has attempted to increase the agreement level, through removing disagreeing decisions, which can be called outliers, or problematic decisions. Then, at the satisfaction of acceptable level of the consensus indicators, the heuristic was able to make a decision based on consensus, and after removal of the problematic, far away, decision. The end results of the two heuristic has agreed, which was expected according to similar underlying logic embedded in both heuristics, but this is not necessary, since in some case, there may happen to be a slight difference between the results of the two heuristics. For instance, in the first heuristic, it might has happened that the combined value using the AM does not exceed the threshold for "Yes" direction, then in this case the answer of the first heuristic would be "Non-biased", but, however, in most case it is expected that both heuristics will agree. It should be noted that the selection between both heuristics should be based on the policy of the decision analysts or whoever is responsible for the group decision; it depends on his/her understanding or thinking of how the solution to the problem could be obtained; is it possible to work toward consensus through exclusion of problematic decisions, or the consensus evaluation should be based on all opinions of FESs without exclusion of any.

In this chapter, I have presented two consensus-based heuristics utilizing several types of consensus measures. These types of heuristics has some advantage over the classical combining rules in that more detailed information are exploited, and that they do not have inherent computational characteristics. This is particularly adequate for such type of GDM, where the decision problem is a social choice problem, and the reliance on consensus information is one effective way to obtain a reliable decision. The main benefit of such consensus heuristics is that some useful explanatory data are produced. This was not the case with the simple, blind, stand-alone combining criteria like AM, which merely computes a statistically combined value of the crisp outputs. This combining criterion basically measures the central tendency of a group of numerical values. In most cases it results in smoothing effect and never gives the extreme values. Sometimes its result may be confusing and have more tendencies to be close the non-biased class. Its role in these heuristics was confined to the case when there is no apparent or acceptable degree of consensus. One advantage of these consensus-based heuristics is that they are decisive in concentrating on the direction of correct answer rather than providing a compromising value like in the case of mathematical criteria. This is especially important as decisiveness is an important notion especially in case of yes-or-no type decisions. Also, they do not suffer from the inherent computational characteristics found in the classical combining rules. It has been found that combining rules' performance is affected greatly by the direction of correct answer, such as in the case of AM, which performs better especially if the direction of the correct answer were at the middle (Aly & Vrana, 2006c). Another advantage is that both heuristics are adaptive through adjusting the indicators' acceptable values according to the inherent consensus policy of the decision analysts or the characteristics of the given decision making problem. However, the classical combining criteria and the newly developed MPDI criterion also offer the simplicity, and also the MPDI was been proven a decisive criterion. The choice between utilization of simple combining criteria and the consensus heuristics is determined by the combination philosophy and policy adopted by and suitable for the decision analysts.

The provided illustrative example has showed that the two heuristics can be applied simply and systematically. Further improvement and augmentation can be done for the two heuristics, and some experimentation could consider comparing the simple combining criteria to these heuristics. This experimentation can reveal more information about the inherent strength and weakness of these heuristics as compared to the simple combining criteria.

Different consensus measures have been produced in this study; all of them have been utilized in developing the proposed heuristics, except the standard deviation, which has not been utilized. This attributed to that the indicator GASI was enough to quantify the amount of agreement and disagreement, which is not the case with the standard deviation, which measures merely the dispersion. The notion provided by this indicator could be utilized in other future researches in developing other heuristics. However, in chapter 8, the standard deviation will be utilized to get information about the amount of dispersion existing in the weights of FESs. This information is especially useful to determine whether or not these weights are significant and should be put into consideration when combining/aggregating the outputs of FESs.

In the next chapter, the first heuristic will be utilized to elicit the DCA, upon which the performance of all measurement-level combining criteria, described in chapter 5, will be evaluated.

Chapter 7

Comparing combining criteria

In chapter 5, several classical combining criteria have been described and configured to handle the crisp outputs of FESs according to the established numerical scale. It has been mentioned that the combining criteria at the measurement level are more advantageous than the other abstract and rank level ones, because they enable taking account for the degree of decision answer, and not only the abstract name of the decision class or rank of these classes, and hence more information are processed. Also, new measurement-level criterion was developed, the MPDI, and the logic behind it was explained. In this chapter, I will focus on this type of combining criteria, the measurement level. Important questions arise, that which one of these criteria is considered superior to the others, and which one is considered inferior to the others? Answer to these questions can be obtained if comparisons of these criteria are made based on their performance in an objective experiment. This implies that there must be some datum level for comparison in form of past data containing experts' or ES's judgments and associated with those judgments the known correct decision answer (DCA). These past historical performance data are extremely useful in ill-structured decision situations, where it is difficult to evaluate the solution to the problem before actual occurrence of future outcomes. Two problem arise here; the first is that there is lack of such experts' performance data, because the proposed problem of integrating multiple FESs through objective combination of their final numerical decision outputs have not received much attention in the literature, and that seldom exists such form of integration as long as there has been no objective numerical scale established to realize such integration, even among numerical experts' judgments. The second problem is that any past historical experts' performance data will be specific to particular nature and characteristics of a specific decision making process and decision making environment. This means that if the combining criteria are compared based on this specific data, it implies that we judge the performance of such combining criteria as which of them is closed to the specific rule implicit within such performance data. This means also, that comparison and comparison results should not be considered general, but rather specific. One way to solve this dilemma is to find a more general rule or datum level to compare these combining criteria. One such general rule is a rule that is based on consensus. In this way all combining criteria will be compared to determine which best reflect or adhere to levels of agreement contained in the data and which one does not reflect so. In other meaning, a consensus-based heuristic, which was introduced in chapter 6, section 6.5.1, will be used to determine the direction of correct decision answer (DCA) for a randomly created artificial experts' judgments data. This DCA will be used as a datum level for comparison. This comparison also aims to get more insight about the performance level of the newly developed criterion, MPDI, and comparing this performance to that of previously existing ones. Four measurement-level combining criteria will be compared; namely: AM, GM, HM, and MPDI.

Next section, the comparison experiments and data used will be explained.

7.1 The comparison experiment

In order to compare the performance of MPDI with that of other combining criteria, an experiment was conducted to identify the superior criteria and the inferior ones. Also, it aims to provide insight about which criterion performs well and under which conditions. Because of the lack of actual expertise's' past performance data, I have created artificial random sets of data. Nine sets were created each of which contains different number of experts ranging from 2 to 10 experts participating in judging the binary decision making problem and utilizing the same numerical judgmental scale that has been established. 30 judgment problems (30 points) were created uniformly randomly for each set. For every set of experts' judgments, the DCA was computed utilizing the first consensus-based heuristic described previously in chapter 6 that is used to elicit the DCA based on consensus-level evaluation. Then, the performance of the combining criteria: AM, GM, HM, MPDI, will be measured based on their deviations from such datum level. There are three possible decision directions:

- The value 10 or "Yes" direction.
- The value 5 or "Non-biased" direction.
- The value 0 or "No" direction.

The criteria are compared based on counts best, which is the percentage of times a given criterion stands the best, and counts worst, which is the percentage of times a given criterion stands the worst. Criteria are also compared based on the average performance throughout all different experts' numbers, in terms of the average deviations from the DCA through out the 30 points or test-problems. This average performance is computed for every distinct number of the participating experts.

7.2 The results of the experiment

Appealing results have been obtained that the proposed MPDI criterion has proven superior to all other criteria, on all the three measures of performance, best and worst counts and the average performance bases. Table 7.1 shows that the MPDI always gives the maximum percentage of times being best performing, except for the 9-judgments problem size, where the AM only stood the best by a little difference, approximately 10 % greater than that of MPDI. In contrast, the GM and HM criterion share almost in having the minimum number of times being best. The AM criterion always outperforms the GM and HM in percentage of time being best. These results of best performance are plotted in Figure 7.1, which displays the excellence of MPDI over others. AM can be classified the second best. Table 7.2 shows that the HM criterion absolutely stands the worst throughout all numbers of experts or problem sizes. The MPDI still outperforms the AM criterion in having fewer counts of being worst, except for two problem sizes, 9 and 10 experts; the AM in these two judgments is slightly better than AM. The GM criterion offers the minimum number of times being worst, and accordingly, this means that GM criterion never stands the worst over all other, and also, as has been determined with the count best, never being the best. These results are displayed in figure 7.2. In overall, the MPDI clearly considered superior to all other criteria, in number of times both being best and worst. The AM stands the second most efficient criterion. HM stands the worst always throughout all numbers of experts, and GM never stands the best, and never stands the worst.

On the average performance basis, table 7.3 shows that again MPDI absolutely gives the lowest average deviations under all problem sizes. The HM offers the highest average deviations throughout all numbers of experts participating. The AM stands the second efficient criteria after MPDI in this respect. GM is better than HM, but still inferior to AM. These results are displayed in figure 7.3. As the results say, MPDI and AM have proven superior to the others. HM and GM are clearly inferior to the other two criteria.

Table 7.1 The best performance of the combining criteria at different numbers of experts.

Number of experts	AM	GM	HM	MPDI
2	33.3	6.67†	26.67	70*
3	26.67	6.67†	13.3	63.3*
4	16.67	3.3†	13.3	80*
5	20	10	3.3†	66.67*
6	23.3	6.67	0†	70*
7	30	3.33	0†	66.67*
8	26.67	3.3†	3.3†	66.67*
9	53.3*	3.3	0†	43.3
10	36.67	0†	0†	63.3*

*The maximum of number of times being best.
†The minimum of number of times being best.

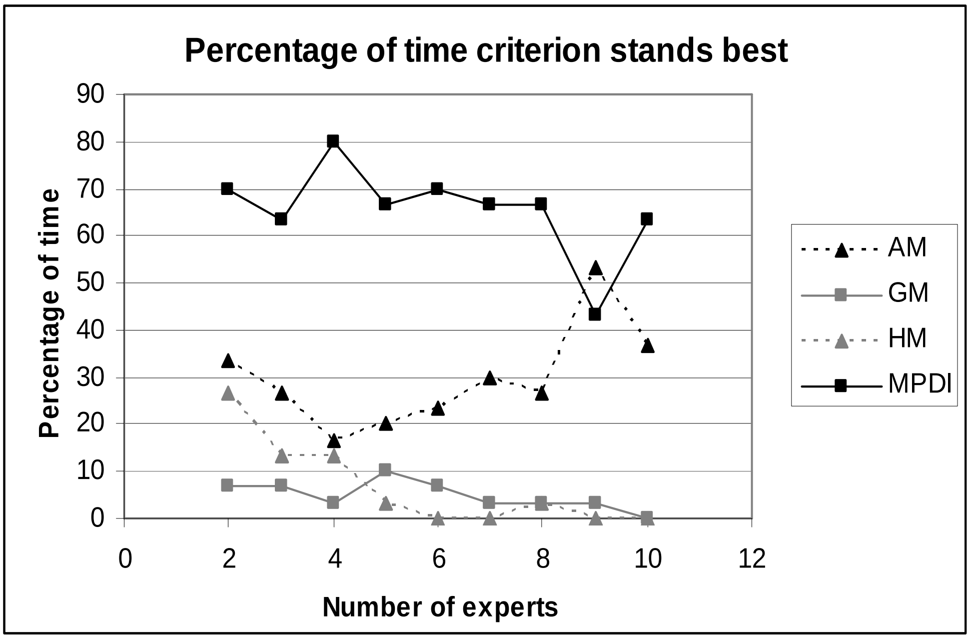

Fig. 7.1 The counts of times a combining criterion stands best performing plotted versus the number of experts.

Table 7.2. The worst performance of the combining criteria at different numbers of experts.

Number of experts	AM	GM	HM	MPDI
2	30	3.3†	60*	10
3	33.3	0†	60*	6.67
4	13.3	0†	76.67*	3.3
5	13.3	0†	76.67*	10
6	30	0†	53.3*	16.67
7	26.67	0†	56.67*	16.67
8	20	0†	63.3*	16.67
9	6.67	0†	56.67*	36.67
10	10	0†	76.67*	13.3

*The maximum of percentage of times being worst.
†The minimum of percentage of times being worst.

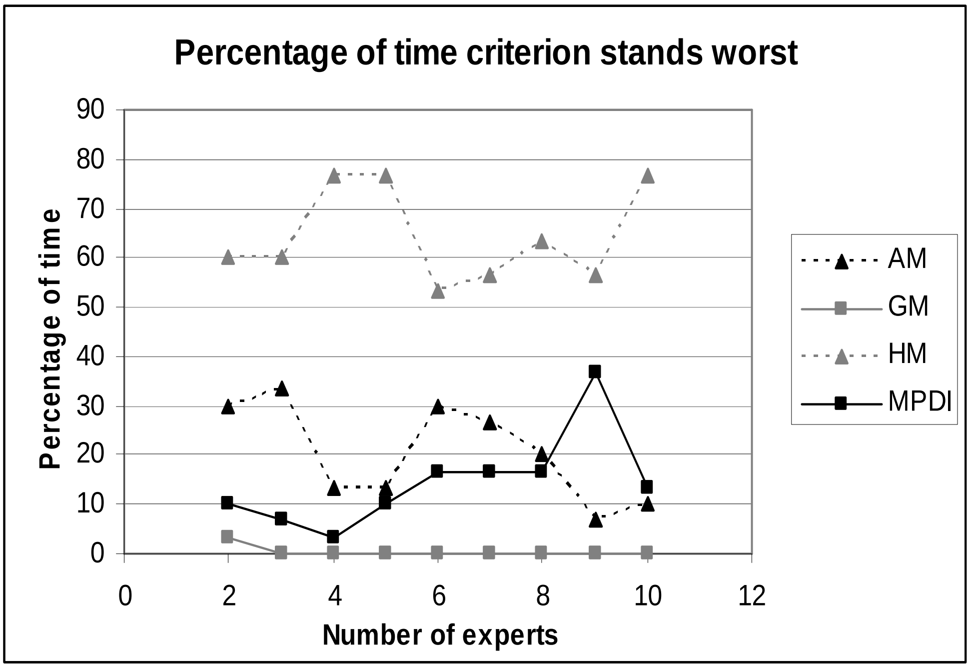

Fig. 7.2. The counts of times a combining criterion stands worst performing plotted versus the number of experts.

Table 7.3 The average deviations of the combining criteria at different number of experts.

Number of experts	ΔAM	ΔGM	ΔHM	ΔMPDI
2	1.97	2.17	2.38*	1.53†
3	2.156	2.33	2.55*	1.235†
4	2.69	3.09	3.56*	0.84†
5	2.25	2.74	3.38*	0.68†
6	2.79	2.96	3.35*	1.134†
7	2.48	2.76	3.23*	0.97†
8	2.78	3.11	3.65*	0.935†
9	1.76	2.33	3.19*	1.16†
10	2.45	3.03	3.85*	0.99†

*The maximum average criterion deviation.
†The minimum average criterion deviation.

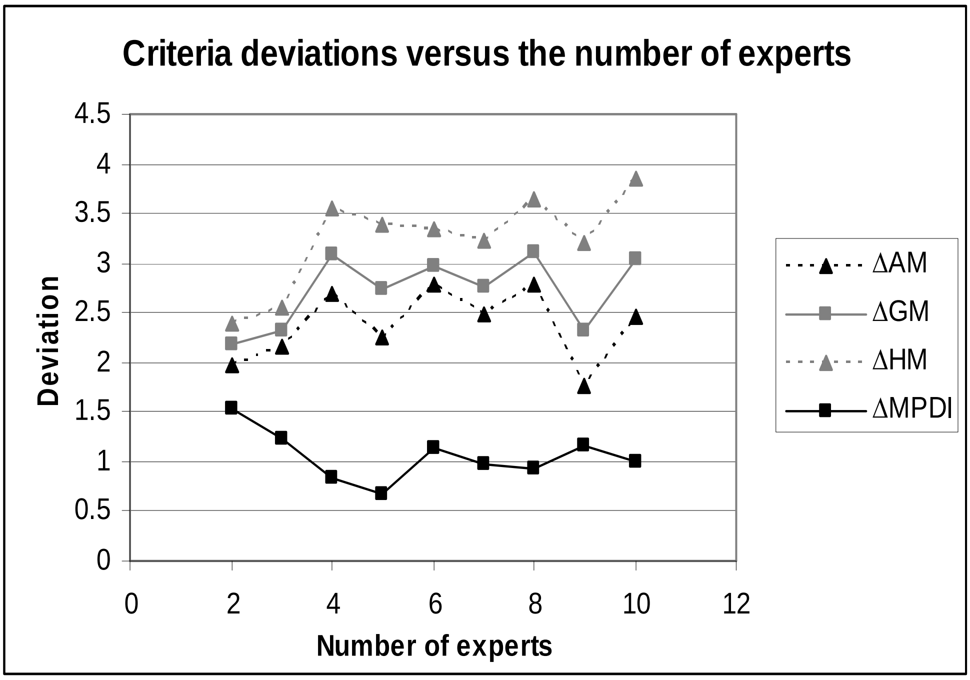

Fig. 7.3 The average performance of combining criteria plotted versus the number of experts.

*I have published the results of this experiment in (Aly & Vrana, 2006c).

7.3 Comment on the obtained experiment results

An experiment was conducted to compare the performance of the developed new criterion, the MPDI, to that of the classical well-known measurement level ones; namely: AM, GM, and HM. A consensus-based heuristic presented in the previous chapter was used to find a datum decision answer to compare the performance of compared criteria upon. This heuristic is extremely useful in ill-structured decision situations, where it is difficult to evaluate the solution to the problem, because not all the influential variables are known, and the actual outcome depends on uncertain future events. In such situations, it is difficult to accurately predict the relationships between inputs and outputs to the decision problem. In addition, the reliance on past historical experts' performance data, which are even lacking, will give a specific results more than general. Hence, a reliance on a more general rule, like combination through evaluating consensus, provides more general basis for comparison; that is measuring which criterion is closely reflecting the agreement or disagreement level within the judgments. The main benefit of the utilized consensus heuristic is the capability to elicit the correct direction of answer based on consensus information embedded in the set of numerical judgments, and it gives representative and non-biased datum level for purpose of criteria comparison. Based on the conducted comparison experiment, an investigation of results obtained reveals clearly that the developed MPDI combining criterion is superior to all other considered criteria, on all bases of comparison, percentage of time being best, percentage of time being worst, and on the average performance bases. This could be attributed to that the developed MPDI is more sensitive and was able to reflect small degrees of consensus levels found within every set of experts' judgments, and always close to the DCA determined via the consensus–based heuristic. This is considered a distinguishing advantage of the MPDI, since consensus base evaluation of FESs outputs is one effective way to ensure reliability of the obtained collective group decision. Another observed advantage of the MPDI is that it does not get affected by where the DCA lies, which has been observed with other criteria like HM and AM. The HM according to its inherent computational characteristics always gives the lowest combined values. Most of the time the HM found best performing; this happens whenever the DCA at the "No" or 0 direction. Similarly, the AM in most cases where it outperforms the MPDI, happens whenever the DCA was at the "Non-biased" or 5. This tells the AM is more compromising than decisive if compared with the MPDI. it has been noted that during computation, in the two problem sizes, 9 and 10 experts judgment, in which the AM has slightly outperformed the MPDI, they were containing most of the DCA's at the middle 5, the "Non-biased", and based on this we can attribute why, in this two-judgment problems, the AM has outperformed the MPDI.

On overall, the excellence of MPDI over the AM criterion and others is due to that it is more decisive in reflecting the agreement of numerical judgments in its computed value, and this is particularly desirable when decisiveness is needed to identify bias to either one of the two extremes. Also, MPDI seldom provides a compromising solution as does AM, especially when there exists a considerable bias toward either decision directions. Both criteria, MPDI and AM are considered superior to other two criteria. The utilization of AM in combining multiple numerical values or judgments should be regarded more as a guide, where AM refers to the degree of bias toward either directions.

Important questions arises at this point, that which criteria to select and when? Which is the most adequate criterion to apply according to a given set of judgments? Whether it is considered appropriate in all cases to apply the measurement-level criteria or to consider some other abstract-level criterion like the Majority Voting (MV)? Whether to consider the weighted version of a selected criterion or to apply its basic version? In the next chapter, all questions will be investigated in attempting to find adequate answers. Only two-measurement criteria the AM and MPDI will be considered, since other criteria were proven inferior to these two.

Chapter 8

Hierarchical fuzzy model for selection of combining criteria

In chapter 5, several classical combining criteria under three levels of combination have been considered for the case of integrating knowledge-equal FESs. They range in sophistication from the Majority Voting (MV) combining criterion, and its weighted version (WMV), which works at the abstract-level of combination, to the Arithmetic Mean (AM) and its weighted version (WAM), which works at the measurement-level of combination. In addition, a new criterion, the MPDI and its weighted version, WMPDI, were proposed. The experimentation conducted in chapter 7 has revealed important information about the performance of the compared measurement-level combining criteria. One important result of such experimentation is that the proposed combining criterion MPDI has outperformed the AM, the second best criterion, on both average and percentage of time comparison bases. However, the AM can be considered superior to the MPDI when the direction of correct answer is pointing to the middle of the used scale (i.e., 5, the non-biased direction). This reinforces the argument that AM is more compromising than decisive. The overall results have indicated that the two criteria: AM and MPDI are considered superior to other the two criteria, the Geometric Mean (GM) and Harmonic Mean (HM). An important question arise that: which criterion is the optimum combining criterion to select? When? In other meaning; what is the best criterion to select, and under which conditions? Another important relevant question is that whether or not a weighted criterion should be used? Whether to select a measurement-level combining criterion like the AM or MPDI, or to use an abstract-level combining criterion like MV? Further, is it more optimal to utilize equal-weights versions or weighted versions of combining criteria?

In this chapter, I am attempting to answer the above questions. One logical step toward answering these questions is to select some candidate reliable combining criteria and then attempt to identify the relevant affecting variables, factors, or characteristics that determine which one of those selected criteria should be adopted and applied. These factors or characteristics should be measured based on the current status of the set of numerical judgments. Then, a model or a method for selecting a combining criterion is to be developed. Three basic combining criteria are considered candidate in this study. The first two criteria are the AM and the MPDI, which are measurement-type. The adoption of these two criteria is based on their performance in the experimentations conducted in chapter 7. A third criterion is the MV, which is an abstract-type criterion. This is due to its popularity and that it is considered reliable particularly when there exists some degree of voting dominance in the input numerical judgments. Associated with the three criteria are three weighted versions: WAM, WMPDI, and WMV. These will be six combining criteria to select a one among them based on the incoming set of FESs' judgments. The selection among these criteria will be guided by a group of adopted factors that influence the selection process. Therefore, the six candidate criteria are:

- AM
- WAM
- MPDI

- WMPDI
- MV
- WMV

Next section will consider the problem of identifying the affecting factors that can help determine which adequate criterion to select.

8.1 Factors influencing criterion selection

Before attempting to identify a set a factors which influence the decision of selecting the adequate combining criterion, it is important to look over the basic and inherent characteristics of the candidate combining criteria. The six combining criteria previously mentioned can be divided into two basic groups. The first group involves criteria that do not utilize weights or in other meaning used for equal weights case, and the second group involves criteria which utilize weights or in other meaning used for different weights case. Consequently, there must be some factors that select between basic combining criteria, and other factors that tell wither to select the criteria itself or its weighted version. Hence, the factors are divided into:

- Factors select among basic combining criteria.
- Factors selects between basic and weighted version of a combining criterion.

Some logical and experimental remarks have been gained from the subjective understanding of the described and presented criteria in chapter 5, and the experimentation conducted in chapter 7. They are as follows:-

- The MV criterion could be considered reliable to judge the GDM problem, providing that there exits some considerable degree of consensus in the multiple numerical judgments; otherwise, it will not have a clear discrimination between the two decision classes.

- The AM criterion only outperforms the MPDI when the direction of correct answer is at or very close to the middle (5). Consequently, it is recommended to utilize it only at or around this situation.

- The MPDI criterion is very sensitive to small degree of consensus or agreement within the combined numerical judgments, since any agreement in either direction is magnified through multiplication.

Based on these remarks, the following conclusions were made:

  - MV criterion should be utilized as long as there is a distinct degree of consensus exists.

  - MPDI is especially useful when the degree of consensus is not so apparent, and hence should be used at these circumstances.

  - AM combining criterion should only be used when there is a consensus toward the non-biased middle, or in other meaning, when the direction of correct answer points to the middle.

After stating these distinctive and inherent characteristics of the selected combining criteria, some relevant factors related to such characteristic should be identified, whose values determine the conditions or circumstances at which these criteria should be selected.

In order to know whether there exists some degree of voting dominance, the percentage of voting's consensus measure that has been defined in chapter 6 is adopted for this purpose. It is composed of three factors, which are:

- %YV : percentage of "Yes" voting's.
- %NV : percentage of "No" voting's.
- %NBV: percentage of "Non-biased" voting's.

These three factors or variables will be used to judge the selection among the basic criteria. Therefore, logically, when there exists some degree of voting's dominance level; that is when there exists some decision option having high percentage of voting's, then, in this case the MV criterion will be selected. When there exists a considerable agreement around the non-biased middle, the AM will be utilized. Otherwise, the MPDI should be used.

In order to decide whether or not to use the weighted version of a criterion, some other factors should be used. These factors should enable to know whether or not there exist apparent differences in weights among FESs, or whether or not there exists some weight dominance level attributed to some decision option. For this purpose, I have adopted two measures. They are:-

- The standard deviation of weights: σ_w.
- The sum of weights of class voting's, which is composed of three factors:-
- SWYV : Sum of weights of "Yes" voting's.
- SWNV : Sum of weights of "No" voting's.
- SWNBV: Sum of weights of "Non-biased" voting's.

Therefore, the complete conceptual model of the criteria selection is shown in figure 8.1. After specifying such model, which defines the input and output factors involved in criteria selection, the problem now is to convert this model into a concrete one, which defines the mathematical relationships or mapping from these input factors into the selection decision. Since there is no exact rule or relationships that converts from the inputs to the output decision, and also no exact value for these input factors that determine the selection decision, the relationship between the inputs and outputs should be viewed as vague, and the problem can be more conveniently controlled using the linguistic If-then decision rules. Hence, a fuzzy model is the most adequate choice in such circumstance, and which enable the use of such human-type control and thinking. Since, the all input factors are not mutually related and can be hierarchically mapped separately, to give a related intermediate outputs, a hierarchical fuzzy model based on the Hierarchical Fuzzy System (HFS) (Raju, 1991) can be constructed to structure the relationship between the input factors and the output decision. This will also offer some advantages that will be described in details in the subsequent section.

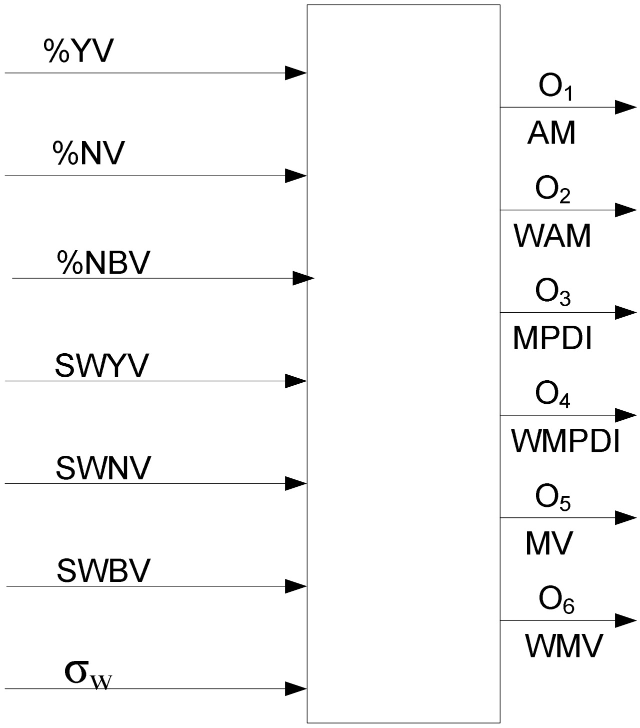

Fig. 8.1 The conceptual model for criteria selection.

Next section, the basic notion of the HFS will be reviewed and its benefits and use, pertaining to the above developed conceptual model, will be explained.

8.2 Fuzzy systems and hierarchical fuzzy systems

Sometimes, when dealing with high-dimensional problems, there exists a common difficulty that is the complexity of the problem increases exponentially with the increase in dimension. This called by researchers "Curse of dimensionality". This notion applies in the context of designing a fuzzy system; as the number of input variables to the problem increases, the number of rules required in the rule base increases exponentially, which constitute heavy burden in maintenance, control, operation of such rule base. One basic solution idea to this problem is to transform the high-dimensional problem into several low-dimensional ones. A HFS, which was first proposed by Raju et al. in 1991 (Raju, 1991), provides a solution for the dimensionality problem based on similar idea. It consists of number of hierarchically connected low-dimensional fuzzy systems LDFS (see figure 8.2). Each LDFS handles a subset of input variables, and these LDFSs are arranged in hierarchical levels in such a way that the output of a predecessor LDFS is used as an input to a successor LDFS in the hierarchy, along with another subset of inputs, which can be either basic input variables or intermediate outputs coming from other LDFSs. With such hierarchical arrangement, it was proven in (Wang, 1998) that the number of rules in a HFS increases linearly with the number of input variables. It greatly reduces the number of rules compared with the standard fuzzy system. In a standard fuzzy systems, the number of rules increase exponentially as the number of input variables increases. Suppose that there are n input variables and m membership functions (i.e., fuzzy sets) for each variable. Then, this needs m^n rules to construct a complete standard fuzzy controller. But, if the n inputs are distributed on a number of (n-1), 2-inputs LDFS within a HFS, then the total number of rules becomes linearly proportional to the number of inputs and equal: (n-1) * m^2. For instance, consider the HFS in figure 8.2…. As n increases, the rule base will quickly overload the memory and make the fuzzy controller difficult to implement. One limitation of the HFS is that the intermediate outputs are artificial in nature in many cases and do not possess physical meaning; consequently it becomes so hard to design the intermediate layers, unless the exact relationships between inputs and intermediate outputs are understood, and that the intermediate outputs physically can be interpreted. However, in the aimed HFS-based model that will be presented in the subsequent section, I will show that this limitation could be overcome, because the physical meaning for the intermediate outputs could be logically interpreted and understood.

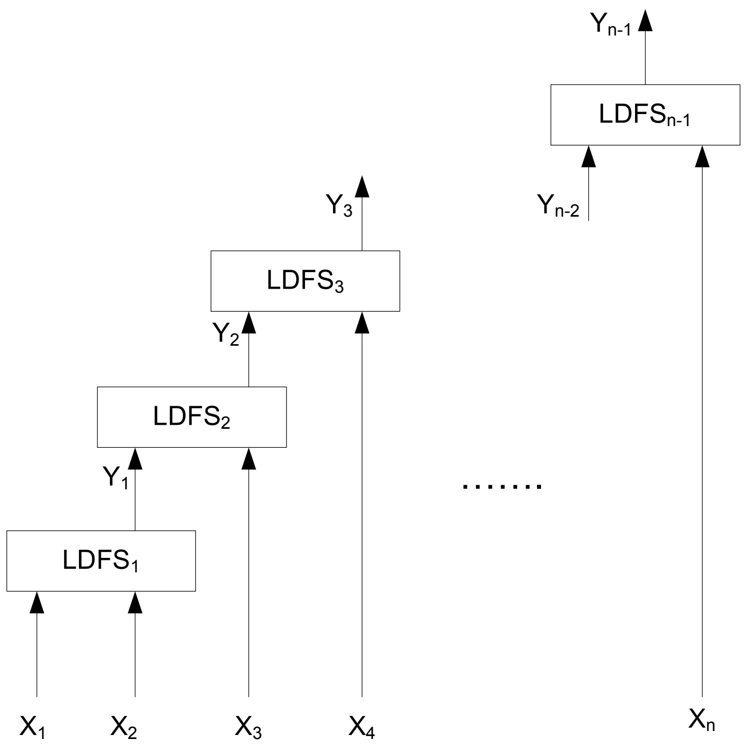

Fig. 8.2 An example of n-input hierarchical fuzzy system comprising (n-1), 2-input LDFS.

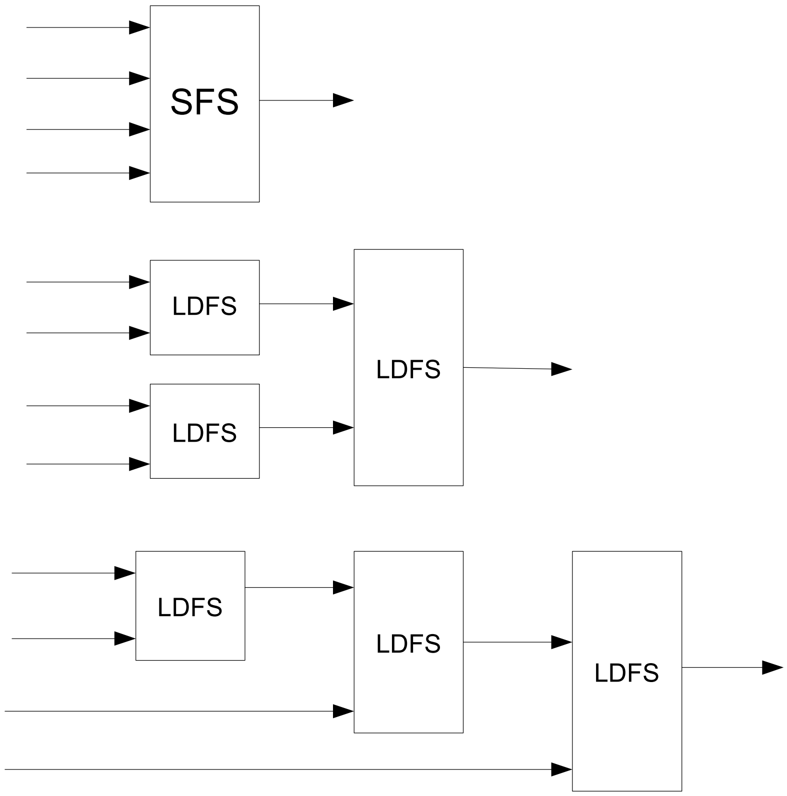

Fig. 8.3(a). Conventional single-layer SFS.
Number of input variables = 4.
Number of fuzzy sets = 5.
Rules = $5^4 = 625$.

Fig. 8.3(b). 2-layer HFS.
Number of input variables = 4.
Number of fuzzy sets = 5.
Rules = $3 * 5^2 = 75$.

Fig. 8.3(c). 3-layer HFS.
Number of input variables = 4.
Number of fuzzy sets = 5.
Rules = $3 * 5^2 = 75$.

Fig. 8.3 Standard fuzzy system (SFS) versus hierarchical fuzzy systems.

Actually, it is not necessary for the LDFS be always of two inputs; it can be of three and more, but the basic idea of HFS is the logical structuring of relationships among the input variables and its contribution to the reduction of total number of rules in the rule-base. In this study, I will demonstrate how this type of hierarchical system will be useful in structuring the relationship between inputs factors and the output selection decision. Also, the limitation of non-interpretability of intermediate outputs will be overcome.

Next section, the conceptual model of criterion selection (figure) presented in section 8.1 will be converted into a concrete HFS-based model, and the necessary fuzzy logics will be specified.

8.3 Designing a HFS-based model for criteria selection

The proposed HFS-based model is shown in figure 8.4. Three basic factors directly affect the decision of which criterion to select. They are:-

- <u>The voting's dominance level (VDL):</u> this factor gives information about whether or not there exists a considerable degree of voting's dominance for the two decision alternatives, "Yes" and "No". The value of this intermediate output is dependent on the two basic inputs: %YV and %NV. The first LDFS, FS_1, is used to map the logical relationship between these inputs and the intermediate output.

- <u>The percentage of Non-biased voting's (%NBV):</u> this factor is used to detect whether or not there exist an agreement toward the non-biased classification.

- <u>The weighting significance (WS):</u> this factor is used to determine whether or not the differences among the computed weights associated with the numerical judgments of FESs are significant significance, and should be included in computation. The value of this intermediate output is dependent on the value of two inputs: the weight dominance level (WDL) and the standard deviation of weights σ_w. The third LDFS, FS_3, is used to map this relationship. The WDL is an intermediate output which determines whether or not there exists a dominance level in weights; that is whether or not there exists some decision option that has considerably high weight computed by summing the weights of its voters among the FESs'. The WDL depends on three input factors: the SWYV, SWNV, and SWBV. The second LDFS, FS_3, is used to map this relationship.

The LDFS, FS_4, is used to map the relationship between the above three basic input factors and the decision of criterion selection. In figure 8.4, there are six outputs associated with the three basic criteria and their weighted versions. The limitation of the HFS is that the intermediate outputs are artificial in nature in many cases and do not possess physical meaning; consequently it becomes so hard to design the intermediate layers, unless the relationships between the inputs and intermediate outputs is understood, and the intermediate outputs can be physically interpreted. This is the case in the developed HFS-model, in which the relationships between the intermediate outputs and the inputs and outputs of the model is logically understood and interpreted, and consequently such limitation is then relaxed.

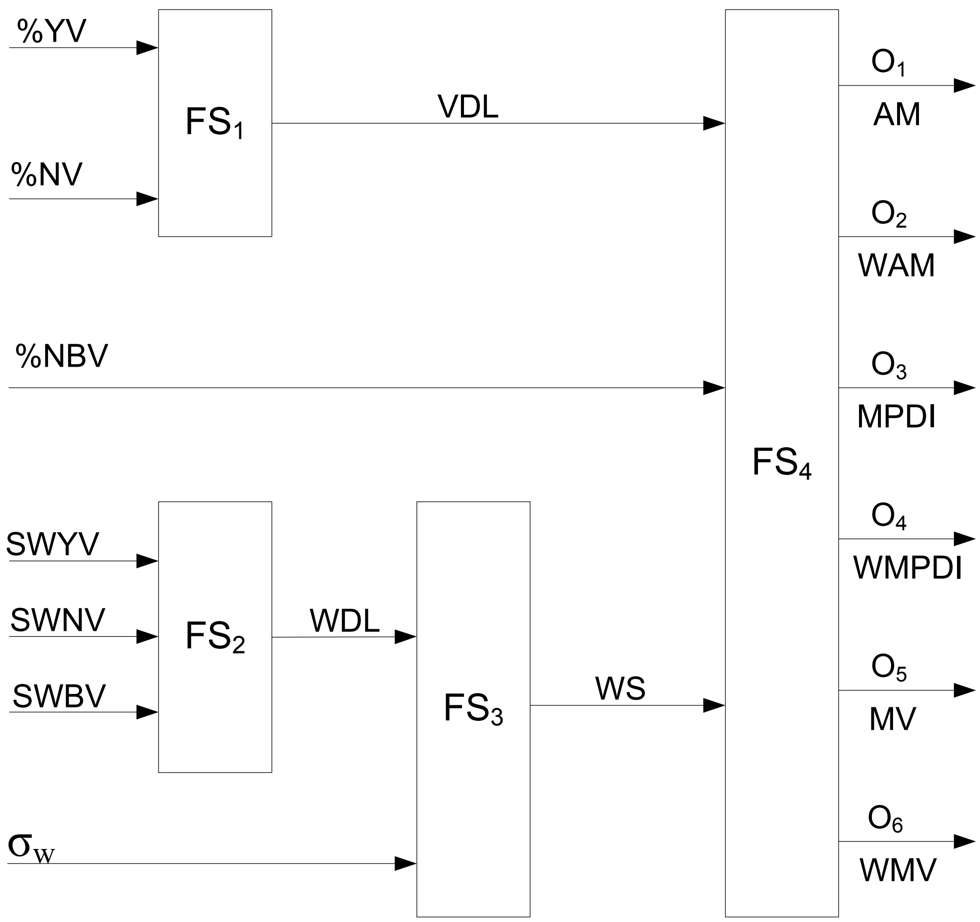

Fig. 8.4 HFS-based model for the selection of the adequate combining criterion.

8.3.1 Specification of the fuzzy logics

Specifying the necessary fuzzy logics to be included in the model involves determining the types of membership functions and the used fuzzy sets for describing the range of values of the input factors and the output decisions, the rules used to fuzzify the values of inputs, the form of the decision rules utilized to map relationships, and the operations used to compute membership values of consequents or the implied fuzzy sets. This information is specified as follows:-

(I) Specifying membership functions:

The standard and popular forms of membership function are the triangular and trapezoidal. They are usually assumed because they are simple to compute with and are efficient as well in approximating fuzzy set concepts. Also, they are used when there is no prior information or no way to empirically compute memberships. For these reason, especially for simplicity, triangular memberships are adopted to describe the range of values of the input factors. Three fuzzy sets like: “Low” (L), “Medium” (M), and “High” (H), could be used to describe the universes of discourse of each input variable. Actually, the adoption of specific fuzzy sets depends on the nature of universe of discourse and accuracy of description available. In most cases it depends on the experts’ or analysts’ viewpoint about which numerical values or range of values should be attributed to which fuzzy set, and what is the mean value of each fuzzy set. Possible variables’ memberships could be as follows (see figures 8.5 through 8.7):-

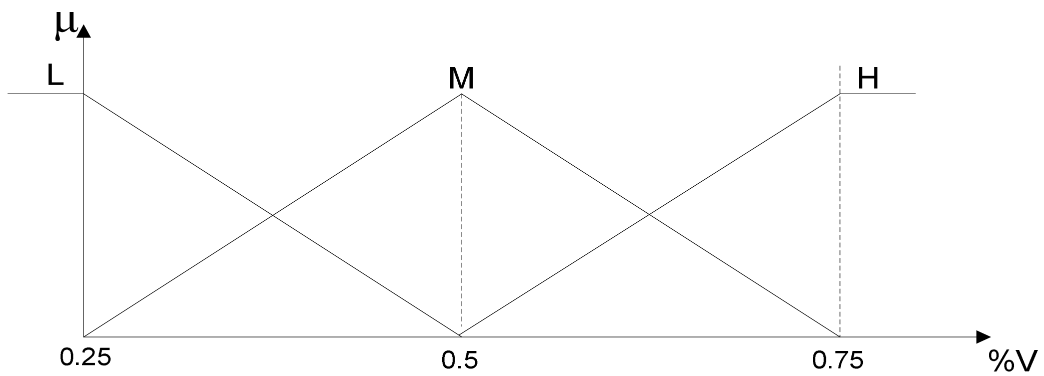

Fig. 8.5 The membership function of the percentages of voting’s (%V).

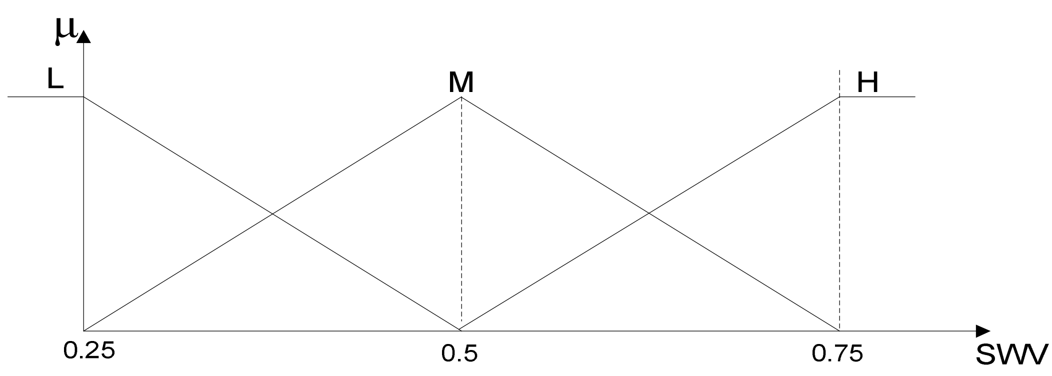

Fig. 8.6 The membership function of the sums of weights of voting’s (SWV).

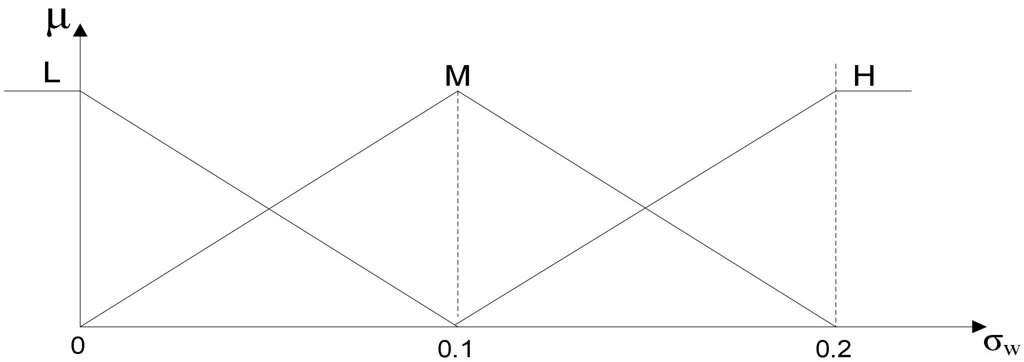

Fig. 8.7 The membership function of the standard deviation of weights (σ_w).

In figure 8.5, the horizontal axis represent the percentage of voting's (%V), which could be %YV, %NV, or %NBV. The universe of discourse for the percentage is naturally from 0 to 1. Using the rule of thumb, the middle values of three fuzzy sets can be specified. If we consider the value of 0.75 percentage of voting's as high value, then correspondingly, the middle value of the fuzzy set "High" is set to 0.75. Similarly, the middle values for fuzzy sets "Low" and "Medium" could be set 0.25, and 0.5 respectively. The same hold for the sum of weights of voting's variable, since the concept of percentage is analogous to the sum of absolute weights, and have the same universe of discourse (see figure 8.6). The membership function for the standard deviation of weights could be specified also using the rule of thumb. If we consider the standard deviation of 0.1 is a medium value of dispersion among the weights of numerical judgments, which means that the average difference between weights is 0.1, then the middle value for the fuzzy set "Medium" can be set to 0.1. Then, the values of 0, 0.2 could represent the middle values for the "Low" and "High" fuzzy sets respectively (see figure 8.7). It should be noted that, the specification of the membership function in any fuzzy model is always subjected to some approximate thinking.

The output of the proposed HFS-based model is a subjective decision, which is the selection of one output out of the six considered outputs; that is the selection of a one combining criterion. This idea of selecting among a set of subjective decision outputs has not been much investigated in the literature concerning the use of fuzzy model. This idea was investigated before by (Aly & Vrana, 2005a) for the selection among a set of several short-term objectives to take up based on some identified influential factors. In this article a psychometric numerical scale was established to express the output objective. This numerical scale was divided among the considered objectives, and each certain possible objective is considered as a fuzzy set or linguistic value on such numerical scale; that is the output decision, which was the selection among a set of objectives, was fuzzified to allow for the use of the defuzzification methods like the Center of Area (Lee, 1990) or the maximum. This was necessary because, there might be several implied sets coming from several fired rules, and that need to be defuzzified. In contrast, in the proposed criterion selection model, only one rule will be fired at the end due to the hierarchical inference structure of the model, and the consequent of this rule will be the final decision of the model, that is the adequate criteria to apply. Thus, there is no need to fuzzify the output decision, which is the selection of a one criterion out of the six possible.

Now, given that the membership functions of basic input factors have been specified, the maximum operator will be used to identify the fuzzy set or label for the actually computed values of the input factors.

(II) Specifying the decision making logic:

The decision rules within the rule base will express the empirical and logical remarks and the conclusion made about the performance of the criteria under consideration, which were previously described in this chapter. These decision making logic will embody these relationships, and consecutively will recommend the value of intermediate and final outputs based on the given situation regarding the values of input factors.

All the rules will take the form of the multi-input single-output (MISO), as follows:-

- If X_1 is A_1 AND X_2 is A_2 AND…….AND X_n is A_n Then Y is B.
- If X_1 is A_1 OR X_2 is A_2 OR………OR X_n is A_n Then Y is B.

The Minimum operator (Mamdani and Assilian, 1975; Mamdani, 1976) will be used to find the consequents' membership value as a minimum of premise's memberships; this is when the connective or conjunction AND is used, as in the first rule. The Maximum operator will be used when the disjunction OR is the connective of the premises memberships, as in the second rule. The decision rules are conveniently tabulated as shown in table 8.1 through 8.4. These rules are based on logical conclusions drawn from experimentation and subjective understanding of how to select the best criterion. Table 8.1 gives the decision rules that map the simultaneous influences of the %YV and %NV on the VDL. This influence relationship is logically connected through the disjunction OR; that is the value of VDL is affected by at least one of he two factors. Table 8.2 gives the decision rules that map the simultaneous influence of the SWYV, SWNV, and SWNBV on the WDL. Also, this influence relationship is logically connected through the disjunction OR. Table 8.3 gives the decision rules that map the simultaneous influence of the WDL, and σ_w on the WS. Also, this influence relationship is logically connected through the disjunction OR, and this naturally can be interpreted as that the significance of weights is decided by either value of WDL or σ_w. For instance, this means that it is enough for the weights to be included in computations, if one of the two values of WDL and σ_w is large enough. Table 8.4 gives the decision rules that map the simultaneous influence of the VDL, %NBV, and WS on the final decision, connected by the conjunction AND. This logically means that the values of VDL, %NBV, and WS are simultaneously necessary to judge which criterion is the most adequate to apply.

Table 8.1 The joint influence of the percentages of votings' (%V) on the voting's dominance level (VDL).

Then VDL		**If %YV**		
		L	**M**	**H**
OR If %NV	**L**	L	M	H
	M	M	M	H
	H	H	H	---

Table 8.2 The joint influence relationship of the sums of weights of voting's (SWV) on the weight dominance level (WDL).

Then WDL		If SWYV								
		L			**M**			**H**		
		OR If SWNBV			**OR If SWNBV**			**OR If SWNBV**		
		L	**M**	**H**	**L**	**M**	**H**	**L**	**M**	**H**
OR If SWNV	**L**	L	M	H	M	M	H	H	H	M
	M	M	M	H	M	L	H	H	H	M
	H	H	H	M	H	H	M	M	M	--

Table 8.3 The joint influence of the weight dominance level (WDL) and the standard deviation (σ_w) on the weight significance (WS).

Then WS		If WDL		
		L	**M**	**H**
OR If σ_w	**L**	L	M	H
	M	M	H	H
	H	M	H	H

Table 8.4 The influence relationship of the VDL, the %NBV, and the WS on the criterion selection decision (O_f).

Then Criterion is (O_f)		If VDL								
		L			**M**			**H**		
		AND If WS			**AND If WS**			**AND If WS**		
		L	**M**	**H**	**L**	**M**	**H**	**L**	**M**	**H**
AND If %NBV	**L**	O_3	O_4	O_4	O_3	O_4	O_4	O_5	O_6	O_6
	M	O_3	O_4	O_4	O_3	O_4	O_4	O_5	O_6	O_6
	H	O_1	O_2	O_2	O_1	O_2	O_2	--	--	--

Now, that the necessary decision logics have been specified, the reasoning process within the HFS is smooth and simple. The maximum operator will be used first to identify the fuzzy set for each input factor's value. Then, the resultant fuzzy sets of the fuzzification process are matched with the premises' of rules within the rule base of the relevant LDFS. Consequently, within each LDFS, one rule must fire, and the output of a LDFS is always the consequent of its fired rule. The membership value of this consequent is the minimum or maximum of the premise's memberships depending on whether the AND conjunction is used or the OR disjunction is used, respectively. The fuzzy output then proceeds from one LDFS to become input to another LDFS until obtaining the final decision output. The maximum operator or center of area defuzzification method could be used when there are several implied fuzzy sets for one output variable; otherwise, as in this proposed model, there will be only one implied fuzzy set for the criterion selection decision, which will be considered the final decision.

More detailed discussion of fuzzy set theory can be found in (Zimmerman, 1987; Mamdani & Gains, 1981; Lee 1990, Zadeh 1965). Also, more details about HFS can be found in (Raju, 1991; Wang, 1998; Wei & Wang, 2000).

Next section, an illustrative example will demonstrate how the proposed HFS-based model could be applied to selecting the best criterion according to the incoming sets of FESs' numerical judgments, and their associated weights.

8.4 An illustrative example

Consider the problem of having five FESs each of which is evaluating a binary decision making problem. Suppose that the crisp numerical outputs obtained from the individual participating FESs were as follows:-

$O_1 = 3$, $O_2 = 7$, $O_3 = 6$, $O_4 = 10$, $O_5 = 2$.

Suppose that their computed associated weights using the AHP or other method were as follows:-

$W_1 = 0.25$, $W_2 = 0.21$, $W_3 = 0.22$, $W_4 = 0.17$, $W_5 = 0.15$.

It is required now to identify which criterion is the most adequate according to these given data.

<u>The solution step will be in three main stages:</u>

(1) Computing the values of input factors:

The values of input factors are computed as follows:-

The outputs are first attributed to their decision classes using 0.5 threshold from the middle, 5, as have been described in chapter 6, and then the percentages of class voting's are computed.

"Yes" voting's class : { O_2, O_3 , O_4 }
"No" voting's class : {O_1, O_2 }.
"Non-biased" voting's class: φ

Then, computing the values of percentages of voting's (%V) and the sums of weights of voting's (SWV):-

%YV = 0.6, %NV = 0.4 , %NBV = 0,
SWYV = 0.6 , SWYV = 0.4, SWYV = 0.

The standard deviation of weights is computed: $\sigma_w = 0.04$.

(2) Fuzzifying the values of input factors:

Using the maximum operator, the computed values of inputs factors is attributed to the convenient fuzzy set. The results of fuzzifying input factors are shown in table 8.5 below.

Table 8.5 Fuzzy sets and associated membership degrees of input values.

Factor name	Fuzzy set	
	label	**μ(degree of membership)**
%YV	Medium	0.6
%NV	Medium	0.6
%NBV	Low	1
SWYV	Medium	0.6
SWNV	Medium	0.6
SWBV	Low	1
$\boldsymbol{\sigma_w}$	Low	0.6

(3) Conducting hierarchical inference:

The fuzzified values of the input factors are first matched with the premises of the rules in each relevant LDFS (tables 8.1 through 8.4). The result is the firing of the following decision rules:-

FS_1: If **%YV** is "**Medium**" (0.6) OR **%NV** is "**Medium**" (0.6) then **VDL** is "**Medium**" (0.6).

FS_2: If **SWYB** is "**Medium**" (0.6) OR **SWNV** is "**Medium**" (0.6) OR **SWNBV** is "**Low**" (1) then **WDL** is "**Medium**" (1).

FS_3: If $\boldsymbol{\sigma_w}$ is "**Low**" (0.6) OR **WDL** is "**Medium**" (1) then **WS** is "**Medium**" (1).

FS_4: If **VDL** is "**Medium**" (0.6) AND **%NBV** is "**Low**" (1) AND **WS** is "**Medium**" (1) then $\mathbf{O_f}$ is "$\mathbf{O_4}$" (0.6).

Based on the hierarchical inference made, the final LDFS has produced the final output of the model, which is the selection of the criterion O_4 (WMPDI) as the most adequate criterion to apply. The formula of WMPDI has been presented in chapter 5. Then, applying this selected criterion to the FESs' crisp outputs gives:

WMPDI = 5.52 > 5.5, then the final decision is "Yes".

The obtained result of selecting the WMPDI criterion to make the outputs' combination can be interpreted as that the built logics in model have identified the significance of weights, and that they also have recognized neither bias toward the non-classified option, the condition necessary for the selection of AM criterion, nor high voting's dominance level, the condition necessary for the selection of the MV criterion, and so this is why the weighted version of the MPDI criterion, WMPDI, was selected. Finally, the proposed HFS-based model has showed its capability to logically structure the relationships among the basic input factors and between the basic input factors and the intermediate derived outputs in a natural and smooth way. Also, the decision logics built within the model have easily incorporated all extracted notions about the performance of the combining criteria under consideration, obtained through actual experimentation, observations, and understanding. And the results of the model could be verified and tested then in comparison with the embedded logics.

Up to now, all the previous analysis and computations were based on handling present or current available information about the FESs' crisp outputs, and that was treated beginning from the fifth chapter up to this chapter. Next chapter will consider combining or aggregating FESs' outputs, based on learning the past historical knowledge and data, which is a one general requirement in this research work that has been stated in chapter 3.

Chapter 9

Handling past expertise's' data and knowledge

This chapter is concerned with combining/aggregating the FESs' outputs based on the experience learned from the available past expertise's' performance data and knowledge. The two words: expertise's and FESs will be interchanged in this chapter, since the past performance data could be of the expertise's modeled within FESs, provided that they are in the form of or converted to the form of the established meaningful output numerical scale, or they could be actually recorded performance of FESs' after a period of implementing these systems.

Past historical knowledge and data are always beneficial in ill-structured, dynamic and uncertain decision making processes. These knowledge and data usually constitute an important guide in understanding the nature and dynamics of such ill-structured decision making processes and environment. In ill-structured decision making problems, which are the type of problems of concern, the complexity in solving these problems arises from the difficulty to evaluate the correct decision solution. This is because first not all the input variables to the problem are fully known; second, some variables are stochastic or vague and consequently no way to evaluate the correct decision solution in advance, unless some future unknown events occur. For this type of situations the reliance on multiple expertise's is one effective way to cope with such inherent complexity and difficulty, and to be able make a reliable decision as much as possible. Therefore, an existence of effective method to utilize such multiple expertise's is crucial to the successfulness of this integrative scheme.

One way to achieve a reliable combination/aggregation of expertise's is to exploit past historical data and knowledge. Usually the past data and knowledge contains very valuable and useful information. These information could help in getting insight about many important characteristics of the decision making process at hand. In the past historical records, the actually realized events and outcomes are known. Historical expertise's' performance data contain expertise's' or expert systems' individual judgments made for every decision making transaction, the collectively made group decision, and the actually recorded correct decision answer known after occurrence of future events. This enables analyzing past performance of the multiple participating expertise's through comparison of their decision solutions with the actually realized outcomes. Information about how the decisions were made in the past can be obtained. Past data and knowledge is of great benefit that enables understanding the relationship between the individual expert systems judgments and the actually recorded outcome or correct answer. This can also provide information about, which implicit or explicit rule had been used to combine or aggregate expertise's' or expert systems' judgments. Information about the difference between the results of this combining or aggregating rule used and actually recorded correct answer can be also extracted. This helps in revising and adjusting the combination or aggregation methods used, or adoption and development of new rules or methods. In addition, questions about which expert system's decision were close to correct decision known, and which one was inferior, can be answered. This knowledge also can give insight about relatedness among individual judgments of expert systems. Explicit information about the relevancy of expert systems to every kind of decision making transactions can be obtained; that is which subset of the available expertise's had participated in judgment of which specific decision making transaction. Also, past performance provides a way to study the behavior of every expert system in response to changing inputs associated with each decision making transaction. This constitutes of course helpful information in analyzing the performance of individual expert systems aiming toward performance optimization. Other valuable information that can be gained from analyzing past data is to study the dynamic change in roles and weights of individual expertise's associated with the dynamic change in the decision making transaction or in decision process. Finally, historical knowledge can help in evaluating and weighting individual expert systems, as has been done in chapter 4.

I shall consider two possibilities of getting benefits from the past historical data, knowledge, and experience. The existence of the two possibilities depends on the type and format of the obtained information. The first possibility is when there are some recorded patterns of numerical judgments of individual expert systems or expertise's accompanied by the corresponding correct decision answers. The correct answers are assumed to be found and recorded after judgments had been made and group decision is executed. The other possibility is when the available knowledge is in form of If-then decision rules comprising linguistically accumulated experience about the relationship between expert systems' judgments and the correct group decision. For these two possibilities, I have adopted and developed two solutions approaches for combining/aggregating expert systems' judgments. A Multi-layer Feed-forward Back-propagation neural network (BPN) is adopted to combine/aggregate FESs' judgments utilizing the past numerical data patterns, and a hierarchical fuzzy system (HFS) combining/aggregating model is proposed to handle the past accumulated IF-then knowledge.

Next section will consider the adoption of the BPN for learning and mapping the relationship between past expertise's' judgments patterns and the actually recorded outcomes.

9.1 Combining/ aggregating the outputs of FESs using BPN

This section proposes the adoption of the BPN to learn the implicit relationship between the expertise's patterns of numerical judgments and the known decision outcomes recorded in the past historical data. These data are in the form of decisions of multiple expertise' or expert systems, and corresponding actually found and recorded correct outcome. These data incorporate valuable and versatile information as it has been mentioned before. Therefore, an adequate modeling tool is needed to handle such expertise's data and to incorporate all such embedded information. Also, it is a restricting requirement for such tool to be able to produce binary output values expressing the "Yes" or "No" subjective decisions. Actually, the utilization of simple combining criteria or consensus-based heuristic presented in chapters 5 and 6 will be deemed inappropriate choice, especially when such past performance data are available. This is because any structured rule like the AM or consensus-based heuristic will not express the ill-structured, and possibly non-linear relationship between expertise's' performance and the true outcome, which most probably is something else rather than those simple structured rules. Therefore, the best choice is to solve this problem specifically for a particular nature of a decision making process and environment, and try to understand the ill-structured relationship existing too between the expertise's judgments and actual true outcome. This will enable implicitly learning which expertise's or FES is the most influential, which subset of expertise's' or expert systems' decision are related, and how this relationships change dynamically throughout the historically chronologically arranged performance patterns. All these complex relationships cannot be grasped only by a single rule based on a simple combining criterion or consensus-based rules. Thus, as long as the past expertise's' performance data of a specific decision process exist, then the most appropriate choice is to attempt to learn and map these relationships and information implicit in these specific data. The neural nets in general and the BPN in particular offer this technology.

In the next section, the reasons for adopting the BPN will be explained.

9.1.1 Reasons behind adoption of BPN

The understanding and modeling of non-linear relationships is still the subject of ongoing researches. In attempting to select an optimal and adequate modeling tool for handling the past recorded patterns of expertise's' or expert systems' judgments, the capabilities of several classification techniques were investigated. These techniques are the artificial neural networks (ANN) and statistical multivariate techniques like multiple linear regression, logistic regression, linear discriminant analysis…, etc. The multivariate modeling techniques were excluded because of the inherent unrealistic assumptions they exhibit. The multiple linear regression assumes nonlinearity, which is not a guaranteed characteristic of the past expertise's' performance data. It assumes also, that the error term is statistically independent and randomly distributed with zero mean. Over and above, the multiple regression is not totally adequate because its dependent variable should be metric or continuous, whereas the decision answers associated with the past recorded patterns are non-metric or binary (i.e., "Yes" and "No" or 1 and 0). Greene, in 1993 (Greene, 1993) pointed out that conventional regression methods are in appropriate when the dependent variable is a discrete outcome (e.g., "Yes" and "No"), and that other techniques are required. The discriminant analysis classification (Fischer, 1939) technique can accommodate binary dependent variable, but it suffers also from more strict and unrealistic assumptions like, normality and linearity of multivariate population and equality of their common covariance matrix. The logistic regression relaxes some of theses assumptions inherent with the discriminant analysis like normality and linearity assumptions, and can be used, but still the ANN a superior classification techniques over all these multivariate ones (see Mak et al., 1996). In reality, these assumptions are often violated, and consequently make the application of these statistical techniques unjustified and their solutions are likely to be unrealistic or inaccurate.

On the other hand, neural nets require no assumptions, and are found to perform best under the conditions of high noise and low sample size (Marquez et al. 1991; Subramanian, Hung, & Hu, 1992), and when the data pattern are complex and nonlinear (Curram & Mingers, 1994). ANN models are powerful modeling approaches and relatively simple compared to mechanistic models. Mechanistic models use mathematical functions to represent processes (Kaul et al., 2004). The difficulty in mathematical modeling is attributed to its stochastic nature and its dependency on a large number of parameters. This is in addition to the fact that assumptions most probably make the resulting decision solution is either inferior or unrealistic. Neural nets could be used for modeling non-linearity, accommodating multivariate and non-parametric data. Neural network approaches, unlike the mechanistic model, is a model-free estimator; they do not require any external manifestation of parametric relationship. Hence, the relationship between the parameters is automatically incorporated into the network model in an implicit manner during the training process. And so, it eliminates the difficulty of extracting the parameters for a mechanistic model (Khazaei et al., 2005). Therefore, based on the above mentioned superiority and advantages of the ANN over the statistical and mathematical techniques, the neural classifier is the adequate choice, and its application will be investigated and exploited in this study.

Because of the capability of ANN models to handle and learn past historical numerical data, a subjective study of these models has been made. The mathematical development of models commenced 7 decades ago. With the work of McCulloch and Pitts (McCulloch & Pitts, 1943), Hebb (Hebb, 1949), Rosenblatt (Rosenblatt, 1959), Widrow and Hoff (Widrow & Hoff, 1960) and others. More recent work by Hopfield (Hopfield, 1982, 1984), Hopfiled and Tank (Hopfield & Tank, 1986.), Rumelhart and McClelland (Rumelhart and McClelland, 1986) Sejnowski and Rosenberg (Sejnowski & Rosenberg, 1986), Feldman and Ballard (Feldman & Ballard, 1982), Grossberg (Grossberg, 1986), and Kohonen (Kohonen, 1984). Lippmann in 1987 (Lippmann, 1987) has classified most well known neural network classifiers, and described details of their topologies, training algorithms, and contexts of applicability. Special focus was on supervised neural classifiers, as the manipulated expertise's performance data contain the correct decision answer. After subjective investigation of several neural network topologies and learning algorithms, the Multi-layer feed-forward neural network trained with Back Error Propagation learning algorithm (BPN) (Rumelhart and McClelland, 1986) was adopted. BPN has a clear termination criterion. This network was preferred over other supervised techniques like Hopfield and Hamming net, because first it can handle continuous valued-inputs not binary as in Hopfield and Hamming net, so it doesn't require conversion of outputs of FESs into binary, and consequently it keeps the detailed, continuous decision degree information in the judgments of individual FESs, and not only the abstract binary choice of specific alternative (i.e., "Yes" or "No"). Second, it has an explicit output which gives the degree by which an input pattern (i.e., FESs' judgments) belongs to each class. This is in addition to its wide applicability, due to its high degree of mapping, adaptability and flexibility to handle numerical input patterns. Also, it has logically understandable topology and training algorithm as well. Over and above, BPN is more deterministic than Hamming which has many arbitrary chosen parameters and functions. This all make BPN is the strongest candidate to classify FESs' crisp outputs into either "Yes" or "No" decision, based upon a previously conducted training course.

Therefore, given the past expertise's' or expert systems' performance data, in form of multiple numerical judgments expressing "Yes" or "No" opinions, and the associated recorded correct decision answer, the BPN will be trained to understand and learn such mapping relationship, and to be used then to classify new FESs' judgments into either "Yes" or "No".

One additional important and desirable feature of the BPN related to the satisfaction of the requirement imposed on the integration problem, is its ability to provide for the most of specific requirements mentioned: preserving extremes, providing for related decision, and providing for veto-type or critical decisions. All these requirements could be easily learned within the input data patterns used to train by the net.

Next section, the suggested network topology will be presented and the BPN training algorithm will be configured for the given inputs format.

9.1.2 BPN topology and training algorithm for combining/aggregating FESs' outputs

In this section, the supervised multi-layer feed-forward network with back-propagation training algorithm is proposed to classify the crisp judgments provided by the FESs into a "Yes" or "No" finally consolidated decision. The net will be provided by the available past expertise's' performance data in form of numerical judgments and the recorded correct decision answer in binary form, 1 if the answer is "Yes", and 0 if it is "No".

A multi-layer feed-forward network structure is composed of a number of interconnected processing elements or neurons indicated by circles as in figures 9.1 and 9.2. Each neuron in the network is able to receive input signals, and to process them into an output signal. Each neuron is connected to at least one neuron, and each connection is evaluated by a real number, called the weight coefficient. The network consists of layers within which neurons are organized. The first layer receives he inputs, and is called input layer. The last layer contains the output neurons, and is called the output layer. The layers between the input and output are called the hidden layers. The weight reflects the degree of importance of the given connection in network. Two configurations or net topologies are possible. The first one (figure 9.1) has only one output node which can takes a real value within [0,1]. The value 1 means complete bias toward “Yes” decision, and the value 0 means complete bias toward “No”. Intermediate values reflect the degree of bias toward either two classes. The second topology uses two outputs (figure 9.2). When the value of the first outputs is 1; in this case the second should be zero; this means complete or clear bias toward “Yes” decision, and when the second output value is 1; in this case the first should be one, and this means complete bias toward “No” decision. The decision in case of other values is decided based on which output has the greatest value. The following rules could be used to attribute the output value to either decision classes:

<u>In case of one-output topology, if:</u>
Output value > 0.5 , then the decision is “Yes”,
Output value = 0.5 , then the decision is “Non-biased”,
Output value < 0.5 , then the decision is “No”.

<u>In case of two-output topology, if:</u>
First output value > second output value, then the decision is “Yes”,
First output value = second output value, then the decision is “Non-biased”,
First output value < second output value, then the decision is “No”.

The selection between the two topologies should be based on experimentation. In both topologies, the number of input nodes is equal to the number of FESs’ outputs to be combined/aggregated. The number of hidden layers and the number of nodes in each hidden layer is determined through experimentation. At least one hidden layer (i.e., three-layer network) with total number of nodes equal $N(2N + 1)$, and using continuously increasing non-linearities can compute any continuous increasing function of N variables (Lorentz, 1976). For a three layers, the number of nodes in the hidden layer is usually set in between 0.5 to1.5 the total sum of neurons in the inputs and output layers (suggested in (Turban et al., 2001)). Also, the size of the learning or training set should be twice the number of hidden units (suggested in Beale & Jackson, 1994; Hassoun, 1995; Haykin, 1994). The whole available set of expertise’s’ performance patterns are separated into a two sets. The patterns in first set is used to train the network and called training or learning set, and the second set of patterns is used to validate the network and is called testing set. Usually 80 percent of the whole set is randomly selected for training, and the remaining 20 percent for testing (suggested in (Turban et al., 2001)). In general the more the data used the more accurate learnability and classification performance of the trained network.

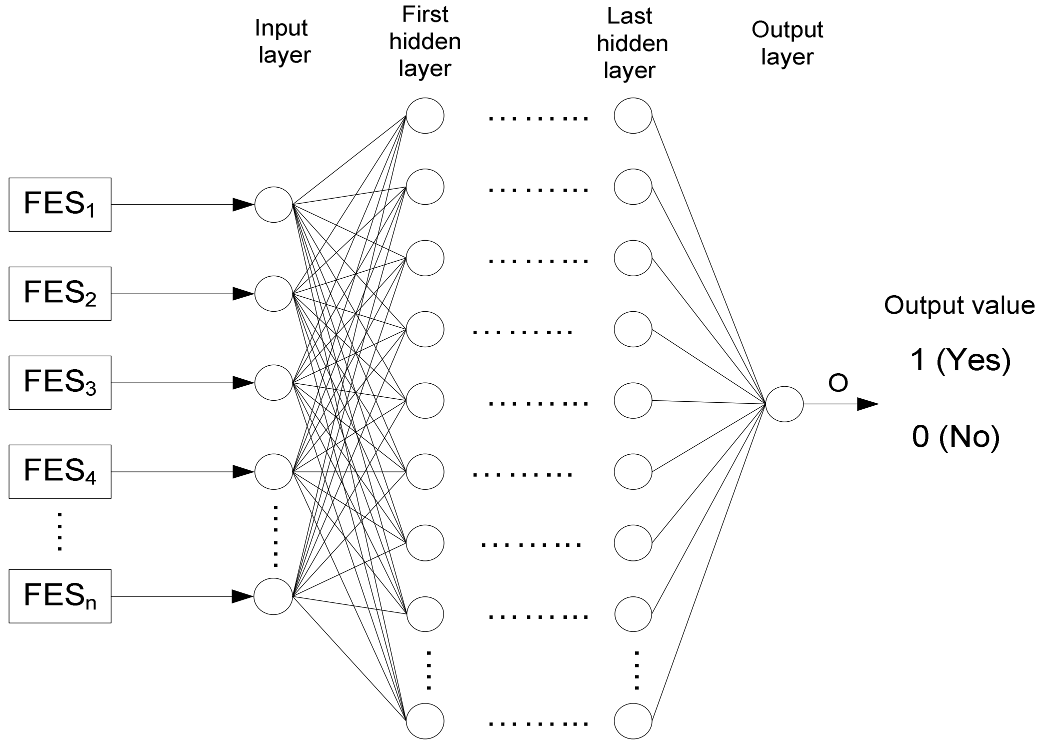

Fig. 9.1 Multi-layer feed-forward neural network for combining/aggregating the outputs of multiple FESs using one output node.

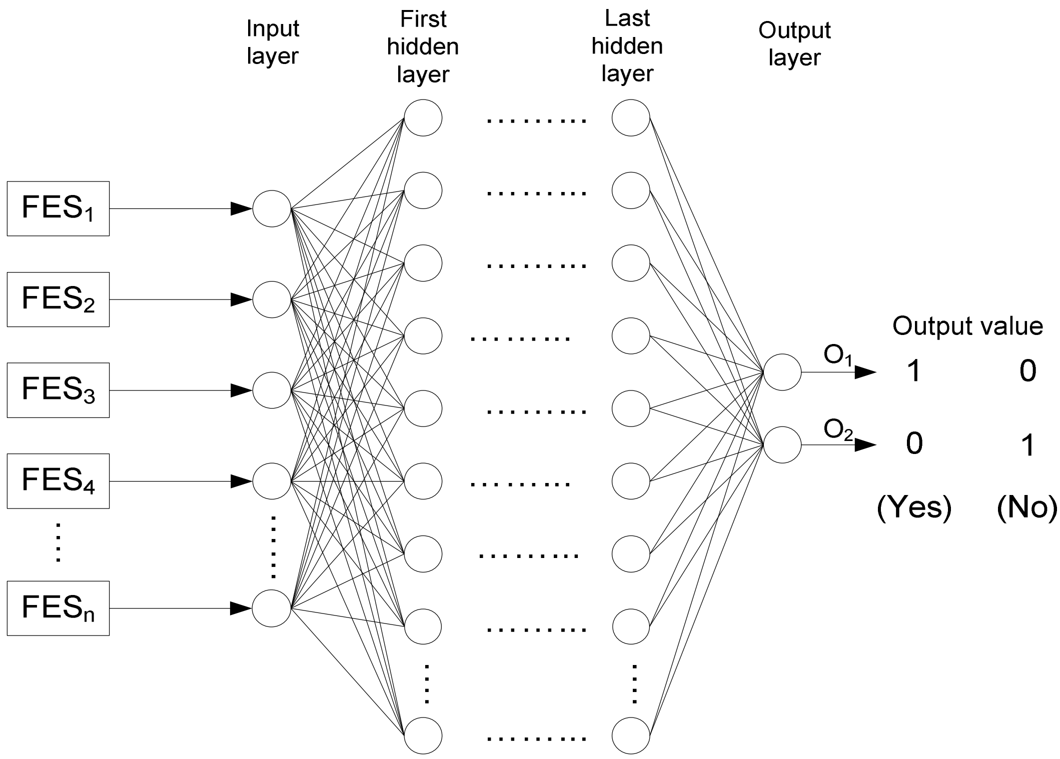

Fig.9.2 Multi-layer feed-forward neural network for combining/aggregating the outputs of multiple FESs using two output nodes.

Any ANN should have specific training rules whereby the weights of connections are adjusted on the basis of the learning data. The BPN learns by examples of known inputs/output sequences. Here, the inputs are the FESs' crisp numerical outputs (i.e., vector of judgments), and the output is the correct recorded decision answer corresponding to each input vector. The BPN learning algorithm is to be used to train either of the two proposed network topologies. Another essential characteristic of the BPN network is the transfer or activation function of a neuron. The transfer functions for the neurons are needed to introduce nonlinearity into the network. Without this nonlinearity, neurons would perform in a linear fashion and the network will not be able to map the non-linear input/output relationships. For the output neuron, the transfer function should be adequate to the distribution of the target. Since the target output limits is 0 and 1, the most popular sigmoid logistic function is utilized, which gives an output between 0 and 1. The equation is as follows:-

$$f(x) = \frac{1}{1+e^{-(x-\theta)}} \tag{9.1}$$

Where,
x : a real input value.
θ : a threshold value (the use of threshold improves the convergence properties of the network).

The Back Propagation (BP) training algorithm is an iterative procedure, which begins first with feed-forward computations, in which an input pattern is applied to the input layer of the network. Then, an activation of the network flows from the input layer through hidden layers to the output layer. During this activation the transfer or activation functions are computed for all neurons in the network, using the initially randomly created set of connections' weights. Then, the values of the outputs of the network are compared to their target levels, and errors are computed. Once the errors have been computed, the connection weights are then updated in back propagation of error mode, in which the change in the network weights is back propagated starting at the output layer and working back toward the input layer. This is done using a gradient descent approach, which aims gradually to minimize the mean squared error of the network. The steps of BP training algorithm, to learn the past expertise's' performance data is described below.

BP training algorithm for learning the past patterns of expertise's or FESs' performance

Step 1: Normalization
The BPN accepts only continuous valued inputs' value within the range [0,1]. Therefore, the input patterns of expertise's judgment should be first normalized using the following formula:

$$X_n = \frac{X_i - X_{\min}}{X_{\max} - X_{\min}} \tag{9.2}$$

Where,
X_n : the normalized input judgment value.
X_i : a numerical input judgment of an expert or a crisp output of a FES.
X_{min}: minimum value of all possible numerical judgments (zero, according to the used scale of judgments).

X_{max}: maximum value of all possible numerical judgments (10, according to the proposed scale of judgments).

Therefore eq. 9.2 reduces to:

$$X_n = \frac{X_i}{10} \tag{9.3}$$

Step 2: Initialization of weights

All weights' values and thresholds are randomly created and initialized to small values.

Step 3: Presenting input vector of expertise's' judgments and corresponding correct decision answer

Present the normalized vector of expertise's' judgments and corresponding desired output value. For the topology which uses one output node. If the decision is "Yes", then the desired output is set to 1, and if "No", then the desired output is set to 0. For the topology which has two outputs, if the decision is "Yes", then the desired value of the first output is set to 1, and the desired value of the second output is set to 0. If the decision is "No", then the opposite is made.

Step 4: Forward computation: compute the actual decision output

Using the sigmoid non-linearity activation function (eq. 9.1), the outputs of all neurons in the successive layers are computed until determining the final outputs of the network (see figure 9.3):

$$f(S_j) = \frac{1}{1+e^{-(S_j-\theta)}} \tag{9.4}$$

$$S_j = \sum_{i=1}^{n} x_i w_{ji} \tag{9.5}$$

Where,
S_j: the weighted input sum at the j^{th} neuron.
x_i: the input from the i^{th} neuron.
w_{ji}: the weight from the i^{th} neuron to the j^{th} neuron.

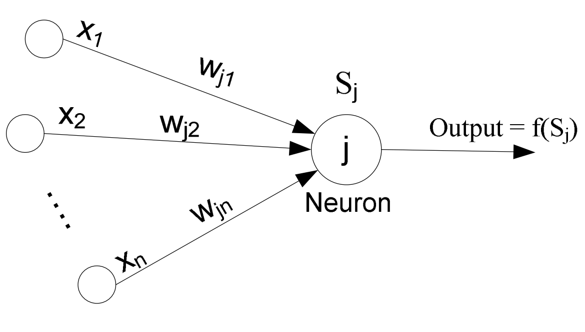

Fig. 9.3 Forward computation of outputs.

Step 5: Adapting weights: computing the actual decision output

First, compare the output value/values obtained in step 4 to its/their desired value/values, and compute the error/errors. Then, use a recursive algorithm to compute the error terms at all hidden neurons starting at the output nodes and working back to the first hidden layer (see figure 9.4). The error terms are computed using the following equation:

$$\delta_j = \left[\sum_{k=1}^{m} \delta_k w_{kj}\right] f'(S_j) \tag{9.6}$$

Where,
δ_j : error term from the j^{th} neuron.
$f'(S_j)$: derivative of function $f(S_j)$, $f'(S_j) = S_j(1 - S_j)$.

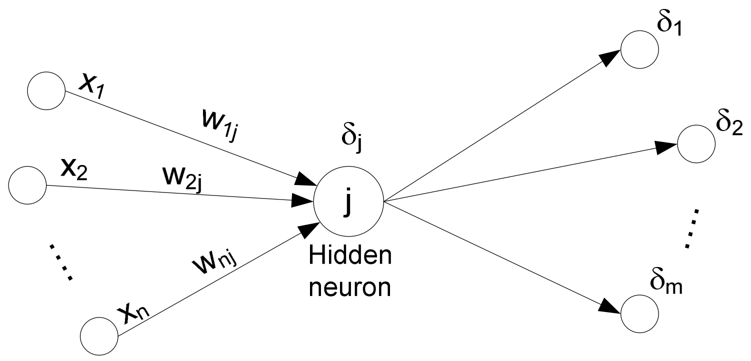

Fig. 9.4 Computing the error for a hidden neuron using back-propagation of errors in the subsequent layer.

Then, adjust the weights by the following equation (see also figure 9.5):

$$w_{ji}(t+1) = w_{ji}(t) + \eta \delta_j S_i \tag{9.7}$$

Where,
η : learning constant or gain term, $0 < \eta < 1$.
S_i: the weighted input sum at the i^{th} neuron.

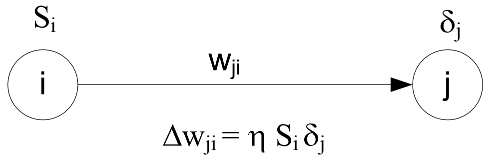

Fig. 9.5 Updating a weight.

Convergence is sometimes faster if a momentum term is added and weight changes are smoothed:

$$w_{ji}(t+1) = w_{ji}(t) + \eta\delta_j S_i + \alpha(w_{ji}(t) - w_{ji}(t-1)) \qquad (9.8)$$

Where,
α: the momentum coefficient , $0 < \alpha < 1$.

Step 6: Repeat by going to step 3

Details for practical hints, computational characteristics, and suggested values of coefficients can be found in (Lippmann, 1987; Beale & Jackson, 1994; Haykin, 1994; Hassoun, 1995; Turban et al., 2001).

Evaluating the performance and accuracy of the BPN:
The performance of the network can be evaluated and compared to other networks during and after training through utilizing the Root Mean Squared Error formula:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(O_i^a - O_i^t)^2} \qquad (9.9)$$

Where,
O_i^a : the actually computed output for the ith input pattern, O_i^a values is within [0,1] .
O_i^t : the target output for the ith input pattern, $O_i^t = 1$ for “Yes” decision answer, and $O_i^t = 0$ for “No” decision answer.

It should be noted that for the second proposed topology of the network, where there are two output nodes, the values of these outputs are always complementing each other and add to 1, and this will be learned by the network as general characteristic of all learned patterns. This means that it is enough to compute RMSE value to either output.

The classification or decision making accuracy of the network is judged for each of the example patterns reserved for testing through the classification error, which is the difference between the actually computed values and it target. Then, the classification as percentage of the full output scale (1) is as follows:

$$\text{Percent classification error (PCE)} = (O_i^a - O_i^t) \times 100 \qquad (9.10)$$

Another measure of accuracy that is adequate for the decision making context of this study can be defined. If we assume that the satisfactory classification is that for which PCE does not exceed certain acceptable value, P. Then, the percentage of satisfactory classification (PSC) within an N testing or validating patterns is:

$$PSC = \frac{NSC^P}{N} \times 100 \qquad (9.11)$$

Where,
NSC^P: the number of satisfactory classifications for which the classification error is less than or equal P %.
N : total number of testing or validating patterns.
P : the acceptable or satisfactory percent classification error.

Note: The P% satisfactory classification error could be changed as convenient, based on the estimation of the analysts or users (P could be set 10, 15, 20,…, etc.).

Actually the main intention is not to optimize either the network topology or the parameters' values for specific experts' performance data, which is beyond the scope of this study, but rather to introduce the multi-layer feed-forward BPN as an adequate tool to combine/aggregate the crisp decision outputs of multiple synergetic FESs, based on learning past performance patterns, when such past performance data exists. This is because the mapping capabilities of the BPN have already been proven. The main focus was to propose the BPN justified by evidences and reasons of superiority and adequacy over the applicable statistical and mathematical multivariate techniques, in general, and other types of ANNs in particular. Other aims are to suggest the possible topology, and describe how the BP training algorithm will manipulate the available data patterns, and finally how the performance of the network could be evaluated during and after learning within the context of binary decision making. The optimization and experimentation with these networks is left to the interested practitioners.

Beside the capability of BPN to learn the past expertise's' performance data, other contribution of BPN in this thesis is that it also could provide for the satisfaction of three specific requirements: preserving extreme FESs' output, related FES's decisions, and veto-type or critical decisions. All these requirements could be learned as input/output relationships by the network.

However, in spite of that the ANNs have proven superior to the statistical and mathematical techniques, especially in classification, it also and exhibits several limitations. First, there are no formal rules specifying how many hidden layers, and how many units per layer should be used (Salchenberger et al., 1992). No exact rule telling how many training patterns should be available. The experimentation and trial and error are only the way to cope with these difficulties. Also, the training of the neural nets may be in some circumstances computationally intensive. In addition, one of the prominent disadvantages of the neural network concerned with that it has no explanation facility; that is no way to understand why particular outputs obtained from certain inputs. However, some of these limitations could be relaxed. For instance, the limitation of intensive computations associated with training neural nets, especially with BPN, is becoming unimportant when offline training is possible.

Next section, a HFS-based model for combining/aggregating FESs' outputs will be presented. This model can manipulate the available past historical knowledge in form of If-then decision rules in order to obtain a finally consolidated output.

9.2 Handling past IF-then knowledge

Past historical knowledge accumulated over time in form of expertise's If-then linguistic rules may be available in many situations. In this case, it is considered an unequaled solution for the ill-structured decision making problems. Under many circumstances, human expert's linguistic rules usually constitute an efficient controller of complex systems. Human expertise and intuition expressed conveniently in natural language can provide good and reliable solution to the ill-structured decision problems.

This section is concerned with utilizing this if-then knowledge in combining/aggregating the crisp outputs of the FESs. This knowledge tells how FESs' output conclusions relate to the final consolidated output of the group. The mechanism proposed to realize such integration is a HFS-based model. In chapter 8, the idea ad functional advantage of the HFS was described. HFSs are used for two purposes, first to help minimize the total number of decision rules necessary to describe system control. Second, it is used also to logically structure the relationships among the input variables. These two notions will be exploited to develop a HFS-based model to combine/aggregate the outputs of FESs in case of existence of such If-then knowledge. Figures 9.6 and 9.7 show the difference between utilizing a standard fuzzy model and utilizing the HFS-based model to structure the relationship between the crisp outputs of FESs and their finally collective decision. In both figures, O_f stands for the final group output, and in figure 9.7, OG_i stands for the output of the i^{th} related subgroup of FESs. In the first case, figure 9.6, the total number of rules is exponentially proportional to the total number of input variables, whereas in the second case, figure 9.7, the total number of rules is linearly proportional to the total number of input variables. As figure 9.7 shows, it is possible to logically structure the relationships among the outputs of FESs to finally obtain the consolidated output. Same as it has been mentioned before in chapter 8, it is not necessary for each low-dimension fuzzy system (LDFS) to be always of two inputs; it can be of three or more. An important implicit notion can be extracted from figure 9.7, is that the decision logic should be able to not only specify how the influential relationships of the FESs' outputs determine the final group output, but also should specify how the influential relationships among the outputs of the subgroups of related FESs determine the final output.

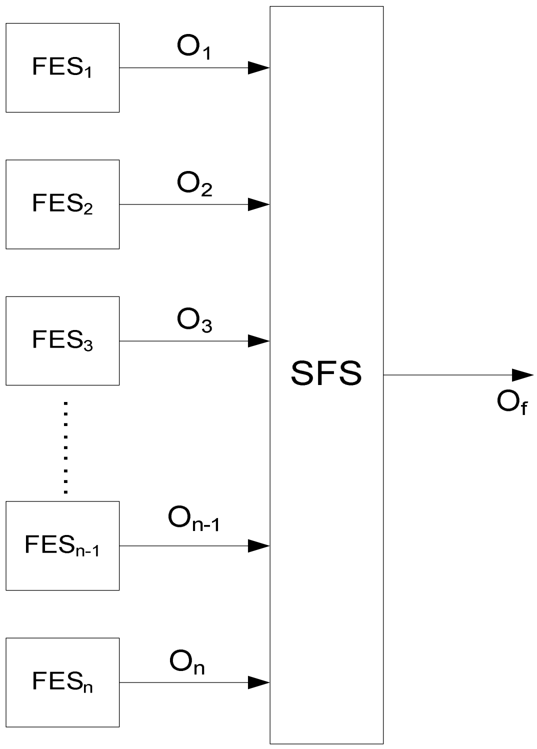

Fig. 9.6 FESs are combined/aggregated using conventional standard fuzzy system.

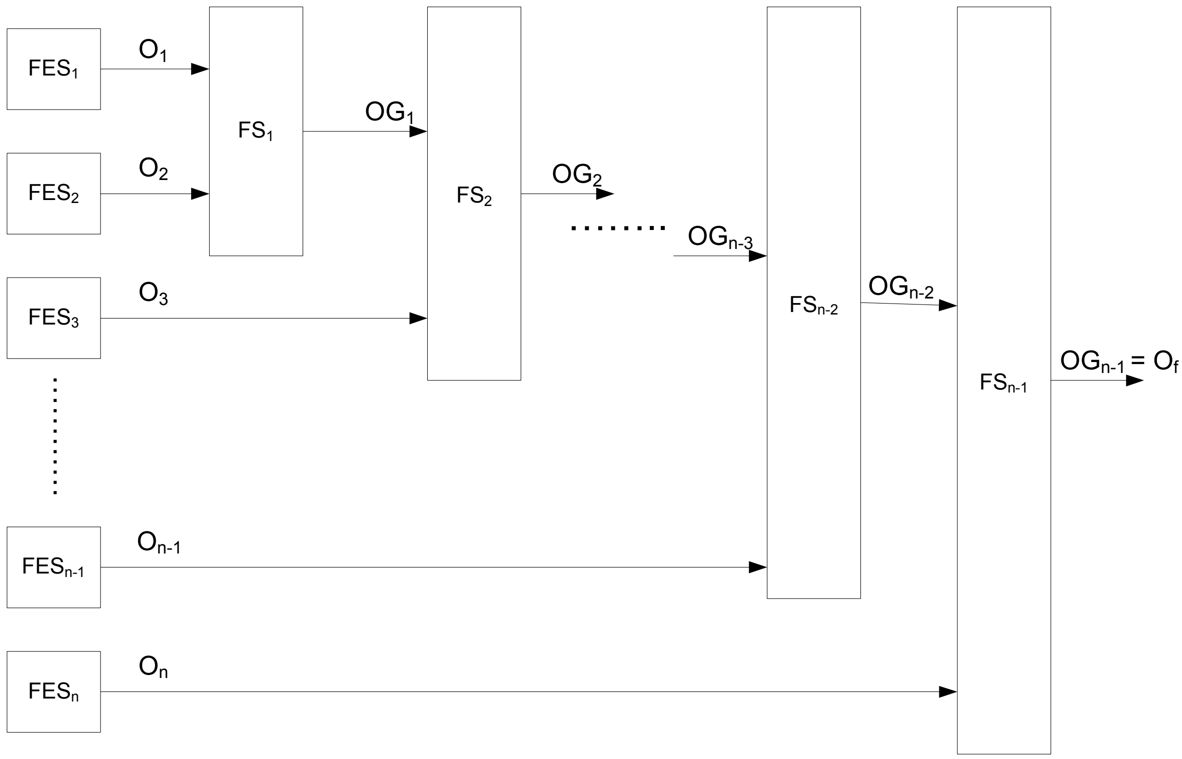

Fig. 9.7 FESs are hierarchically combined/aggregated in the framework of HFS.

In order to use the proposed HFS-based model, it is necessary to specify a set of fuzzy logics. These fuzzy logics involve determining the types of memberships or fuzzy sets for describing the range of values of input factors and the output decisions, the operations used to fuzzify the values of inputs, the type or form of decision rules utilized to map relationships, and the operations used to compute and defuzzify membership values of the consequents or the implied fuzzy sets. As it has been described in chapter 6, the standard membership functions like triangular one could be utilized as a default when there is no knowledge, empirical observations or other methods that can be used to construct memberships.

In the proposed model, all the intermediate outputs will have the same physical meaning as the final collective output. Consequently, only one membership function will be used for inputs (i.e., FESs' crisp outputs), intermediate outputs, and final output when needed. Regarding the operation used in fuzzification of input values, the well-known maximum operator will be used as in chapter 8. The if-then decision logics are the most important input, and are the linguistic formulation of the past knowledge available about how to obtain a collective decision, based on the given status of the FESs' crisp outputs. These If-then rules must also involve knowledge about how the subgroups' outputs relate to the final collective decision. For instance, consider the following three linguistic rules:-

If the output of **FES$_1$** is **High** and the output of **FES$_2$** is **Medium** then the output of the **first group** is **High**.

If the output of the **first group** is **High** and the output of the **second group** is **High**, then the output of the **third group** is **High**.

The first rule is concerned with the two related FESs, and specifies how to covert the two linguistic values of the two FESs into that of the group. The second rule involves two related subgroups of FESs, and specifies how to convert their linguistic output values into an output representing consolidated conclusion of the two groups.

The widely utilized Minimum operator (Mamdani and Assilian, 1975; Mamdani, 1976) will be used to find the consequent membership value as a minimum of premise's memberships; this is when the connective or conjunction AND is used, as in first rule. The Maximum operator will be used when the disjunction OR is the connective of the premise's memberships. The widely utilized Center of Area defuzzification rule (Lee, 1990) is not relevant here in this model of figure 9.2, since it has only one final output. This defuzzification rule is only relevant in case of existence of multiple outputs or implied fuzzy sets obtained through separated or unrelated intermediate fuzzy systems, which is a possible case. For a single output as in figure 9.2, the final output is reached naturally as a result of the consecutively conducted preceding minimum operations. At the final fuzzy system, only one output will be in hand, and the maximum defuzzification rule is then used on a single implied fuzzy set, which gives the its center value as a final group decision.

The next example will demonstrate how this proposed model could be practically utilized.

9.2.1 An illustrative HFS example model

Let us suppose that five FESs are used to decide whether or not a modern car under testing, should undergo an intensive maintenance course. The problem belongs to the class of binary decision making problems. Two decision alternatives are possible: either "Car needs maintenance" or "Car does not need maintenance". The five relevant, participating FESs are:-

- Electronic FES.
- Electric power FES.
- Vibrational FES.
- Structural mechanic FES.
- Chemical FES.

The related sub-groups of FESs are hierarchically structured in the HFS-based model as shown in figure 9.8.

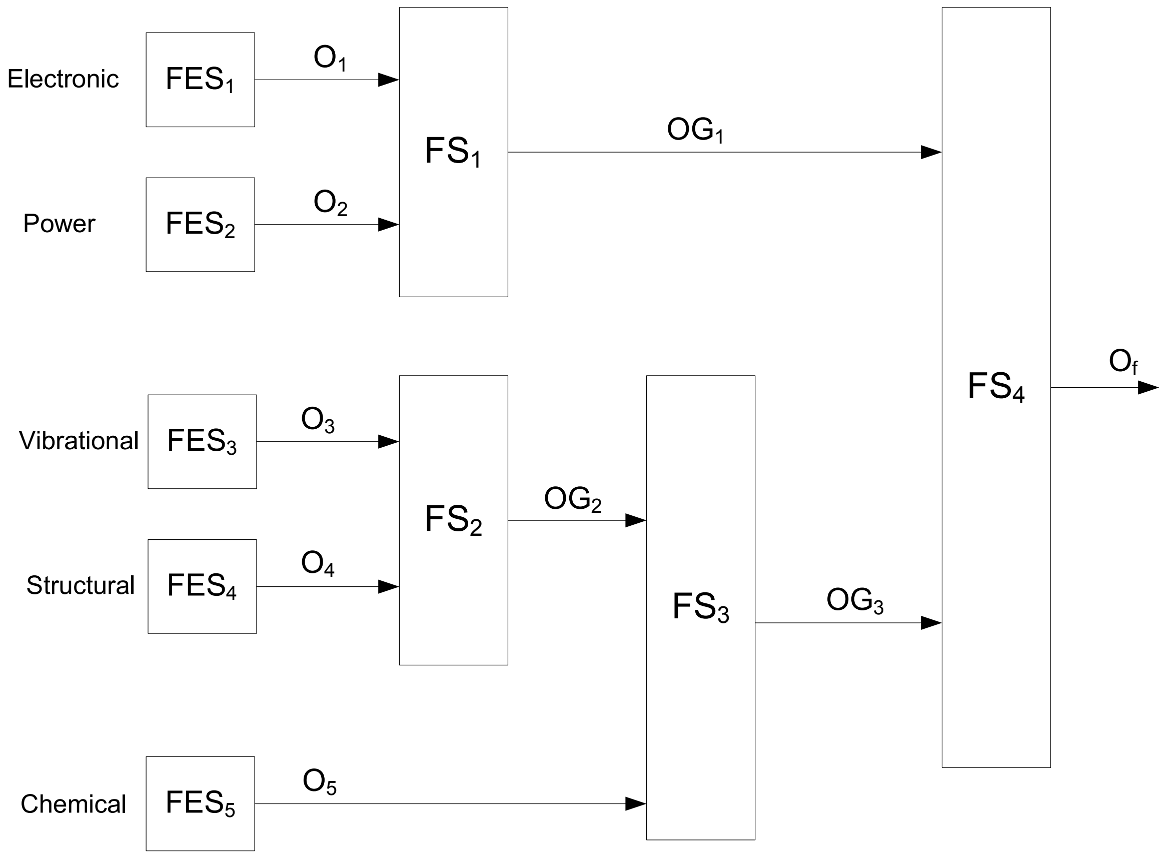

Fig. 9.8 A HFS-model for combining/aggregating the outputs of five FESs of the example problem.

Every FES should provide its crisp output value, O_i, within the range [0,10], expressing the degree of maintenance need; that is the value 0 expresses strongly no need for maintenance, and the value 10 expresses strongly a need for maintenance. Suppose that the crisp numerical outputs of FESs are as follows:

$O_1 = 3$, $O_2 = 7$, $O_3 = 6$, $O_4 = 10$, $O_5 = 2$

Three linguistic values or fuzzy sets are used to describe relationships and control, as available in the past knowledge. Then, the computationally simple triangular membership function could be used to describe the universe of discourse of the output variable (See figure 9.9). Other membership functions could be used as needed and as possible. The numerical values of outputs are fuzzified as shown in table 9.1. Then, given these fuzzified values, the applicable decision rules, table 9.2 through 9.5 are fired.

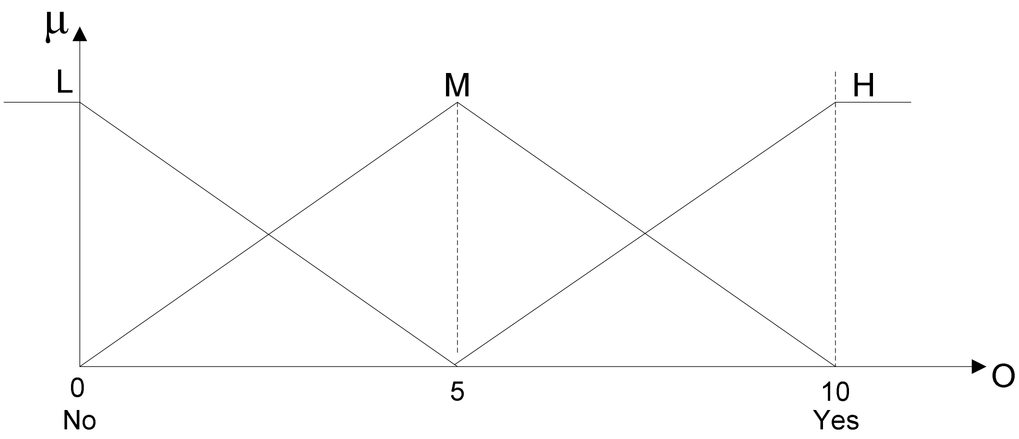

Fig. 9.9 A triangular membership function of a FES's output.

Table 9.1 Fuzzy sets and associated membership values of outputs' values.

Variable name (O_i)	**Fuzzy set**	
	label	**μ(grade of membership)**
O_1	Medium	0.6
O_2	High	0.6
O_3	High	0.8
O_4	Medium	1
O_5	High	0.6

Table 9.2 If-then decision rules for FS_1 defining the joint influence of O_1 and O_2 on the output of the first subgroup, OG_1.

Then (OG_1)		**IF O_1**		
		L	**M**	**H**
And if O_2	**L**	L	L	H
	M	L	M	H
	H	H	H	H

Table 9.3 If-then decision rules for FS_2 defining the joint influence of O_3 and O_4 on the output of the second subgroup, OG_2.

Then (OG_2)		**IF O_3**		
		L	**M**	**H**
And if O_4	**L**	L	H	H
	M	L	H	H
	H	M	H	H

Table 9.4 If-then decision rules for FS_3 defining the joint influence of O_5 and OG_2 on the output of the third subgroup, OG_3.

Then OG_3		IF OG_2 L	M	H
And if O_5	**L**	L	L	M
	M	M	M	H
	H	H	H	H

Table 9.5 If-then decision rules for FS_4 defining the partial influence of OG_1 and OG_3 on the output of the fourth subgroup, the final system's output, O_f.

Then O_f		IF OG_1 L	M	H
And if OG_3	**L**	L	M	H
	M	M	H	H
	H	H	H	H

<u>The fired decision rules are:</u>

FS_1: If $\mathbf{O_1}$ is "**Medium**" (0.6) AND $\mathbf{O_2}$ is "**Medium**" (0.6) then $\mathbf{OG_1}$ is "**Medium**" (0.6)

FS_2: If $\mathbf{O_3}$ is "**Medium**" (0.8) AND $\mathbf{O_4}$ is "**High**" (1) then $\mathbf{OG_2}$ is "**High**" (0.8)

FS_3: If $\mathbf{OG_2}$ is "**High**" (0.8) AND $\mathbf{O_5}$ is "**Low**" (0.6) then $\mathbf{OG_3}$ is "**Medium**" (0.6)

FS_4: If $\mathbf{OG_1}$ is "**Medium**" (0.6) AND $\mathbf{OG_3}$ is "**Medium**" (0.6) then $\mathbf{O_f}$ is "**High**" (0.6)

The output of the FS_4 is the final output, which is "High". Then, the final crisp consolidated decision is the 10, which interpreted as: "Car needs maintenance".

The example has demonstrated that the use of HFS-based model simply enables to logically structure the relationships among FESs, and then to apply the available past knowledge to make a final decision. It should be noted that the importance's or weights of FESs have not been explicitly utilized, because they are implicitly contained in the influence of every crisp output existing in the decision rules.

The developed model could provide for the satisfaction of three specific requirements out of the four described in chapter 3. They are: preserving extreme output values, expressing relatedness among FESs'decisions, and allowing for veto-type or critical decisions. This could be achieved easily through building decision rules in the whole set of the model's decision logics to express these specific relationships and controls.

In this chapter, two approaches have been introduced to provide for the requirement of handling the past historical data and knowledge. The first approach involves the use of the multi-layer feed-forward BPN to learn the numerical data patterns of expertise's' past performance, and the other approach manipulates the If-then past knowledge. Both approaches are practical and computationally feasible and could provide reliable decision solutions. Further, the two approaches provide for the satisfaction of three imposed specific requirements: preserving extreme output values, expressing relatedness among FESs'decisions, and allowing for veto-type or critical decisions. It should be mentioned that as long as these data and knowledge exist specific to a particular decision making process and environment, then in this case attempting to utilize other general combining/aggregating rules like consensus-based heuristics, combining criteria, or an aggregation heuristic, while ignoring these past experience should be deemed erroneous. These data and knowledge should be exploited, and these structured rules or heuristics are most appropriate when these knowledge and data do not exist.

Chapter 10

Satisfying specific requirements

Frequently occurs in many decision making contexts that some specific restrictions or requirements are imposed on the decision solution of the problem. In the context of the integration problem, there have been four requirements imposed. Most of these requirements are actually imposed by the aforementioned practical project, and the others are subjectively elicited as being possible requirements. These specific requirements are more closely viewed as restrictions on combining/aggregating the individual outputs of the FESs, and they are concerning the roles of individual FESs and relatedness among their individual decisions. The first of these requirements is to preserve extreme output values of the FESs. Usually these extreme values constitute special importance especially in this case of binary decision making problem, where the bias toward either extreme is needed. Before describing the importance of extreme output values, it is important to understand what the meaningfulness of extremes produced by a FES. It is known that every FES contains inside a set of If-then decision rules; then when a new set of input values applies, these values are matched with the premises of the in the rule base to identify the firing rules. The inference determines the implied fuzzy set or consequent of each rule. Some of these consequents may be conflicting and some may be agreeing. Then, the defuzzification process combines these rules to give a final crisp output. If these whole process results in an extreme output value pointing to either “Yes” or “No”, then this means that the internal logics existing within each FES clearly showed a bias toward either option. This distinctively biased decision bears more amount of confidence about the biased direction. Consequently these extreme values are of special importance as it shows stronger bias and decisiveness to either decision option. The second requirement is the provision for a null FES’s decision. Under some circumstances, some relevant FES may be unwilling to participate in judgments; either because of uncertainty and incomplete data or knowledge, or because of any other reasons. It is required to allow for this possibility in a way that affect the group decision, and taking into account the null participation of one or a subset of FESs. The third requirement is to provide for related FESs’ decisions. Additional valuable information contributing to increasing the reliability of the obtained group decision could be gained if we consider related expert systems’ decision. If two or more related FESs’ decision agree in their decisions, this gives more confidence about their decision direction. In contrast, if two or more related FESs’ decisions disagree, this decreases the reliability of the two decisions and consequently affects the final group decision. Then, a provision for satisfying this requirement should be made. The last requirement is the provision for critical or veto-type decisions. Some FESs are considered of special importance and may only acquire this veto-type privilege, so if these systems give complete bias toward either decision answer, 0 or 10, then their decision is considered of high priority and has high degree of influence, and only under this circumstance, the group decision should be forced to be in favor of these veto-type decisions. Conveniently, the four requirements can be summarized as follows:-:-

(1) Preserving the extreme values of the FESs’ outputs.
(2) Allowing for null expert systems, who are not willing or do not have sufficient knowledge to participate in the judgment process.
(3) Providing for related expert systems’ decisions.
(4) Providing for critical or veto-type expert systems’ decisions.

This chapter presents a simple heuristic algorithm and practical suggestions for how to integrate multiple parallel FESs while satisfying the above described set of the imposed specific requirements. These heuristics and practical suggestions manipulate present information of the FESs' crisp outputs. The two cases of the integration problem will be considered in satisfying these requirements; the knowledge-equal and knowledge-unique cases.

The next section will introduce an approach and practical suggestions for how to provide for satisfaction of the four requirements in case of knowledge-equal FESs.

10.1 Satisfying the specific requirements for knowledge-equal FESs

In this section, I shall present either a solution approach or practical suggestions to combine the crisp outputs of FESs while satisfying each individual requirement. Also, I shall demonstrate how the established output numerical scale will be especially useful and flexible to help satisfying these requirements. It is assumed that the individual FESs have generally different weights, but however the same approaches and suggestion will be adjusted for the equal-weights case. Then, the requirements will be satisfied as follows:-

Preserving extremes:

Generally, the extreme output values, which are relatively far from the middle value and biased to either "Yes" or "No" decision direction, represent a special importance. This is because as long as the output value is close either extreme, this signifies that the logics within the relevant FES have clearly indicated a bias toward that extreme. This means more confidence and decisiveness. Thus, the care about preserving extreme output values ensue from the relatively high degrees of confidence, reliability, and decisiveness associated with those values. This care requires a suitable combining method or criteria to convey this interest in keeping extremes, and to get benefits from their associated reliability. With this special interest and care, the utilization of combining criteria like the arithmetic mean (AM) or the newly developed MPDI criterion, will not be the correct choice, since this criterion does smoothing effect on all output values, and never gives the extreme values, and this causes the ignorance and loss of such important information and consequently can lead in some cases to incorrect judgment. So, it should be more seriously taken into consideration. In order to combine the outputs of the individual FESs while preserving the extreme values, I have developed a simple extreme-preserving heuristic. It is described below.

First, let us denote:
O_i : the crisp output of the i^{th} FES.
w_i : the weight or priority of the i^{th} FES.
n : the total number of FESs within the predefined matched set.
O_f : the finally consolidated output for all FESs.

Then, the step of the proposed heuristic is as follows:-

Stage 1: Apply AHP to rank the FESs to obtain their priorities:
$w_1, w_2, \ldots. w_i, \ldots, w_n$.

Stage 2: Divides the FES's output values O_i into three distinct groups:

The first group, G_1 : contains all output values or voting's that are above the middle (> 5) (i.e., biased to "Yes"), O_i^+.
The second group, G_2 : contains all output values or voting's that are below the middle (< 5) (i.e., biased to "No"), O_i^-.
The third group, G_3 : contains all output values or voting's that are exactly at the middle (= 5) (i.e., non-biased), O_i^m.

Where,
O_i^+ : the output of i^{th} FES, which is above the middle.
O_i^- : the output of i^{th} FES, which is below the middle.
O_i^m: the output of i^{th} FES, which is equal to the middle.

Stage 3: Sum the weights associated with the outputs of FESs within each group to obtain the groups sums of weights, SWV_j:

$$SWYV = \sum w_i^+, \; SWNV = \sum w_i^-, \; SWNBV = \sum w_i^m.$$

Where,
w_i^+ : the output of i^{th} FES, whose output is above the middle.
w_i^- : the output of i^{th} FES, whose output is below the middle.
w_i^m: the output of i^{th} FES, whose output is equal to the middle.

Note: The magnitudes: SWYV, SWNV, and SWNBV have been defined before in chapter 6, as the sum of weights of "Yes", "No", and "Non-biased" voting's.

Then, find the group G_{max} which has the largest sum of weights, and select its extreme output value among its output values O_k to be the final output. Formally stated through the following two equations:-

$$G_{\max} = \arg\max_j \{SWV_j\} \tag{10.1}$$

$$O_f = \underset{k}{extreme}\{O_k\} \qquad \forall\ O_k \in G_{\max} \tag{10.2}$$

Therefore, the final group decision is "Yes", if this final output value happens to be above the middle (5), and it is "No", if this final output value happens to be below the middle. Otherwise, it is "Non-biased".

If all sums of weights are equal, here only use either the HFS-model described in chapter 8 to select the adequate combining criterion, or the consensus-based heuristics described in chapter 6, depending on the choice and policy of the decision analyst, in order to obtain the finally consolidated group decision output. Thus, the heuristic logically adheres to the group, which has the maximum importance (maximum sum of priorities); accepts its direction opinion, and then it catches the extreme of such group of FESs. Also, if the most important group happens to be that group which contains the middle values, then according to such most important group, the decision solution cannot be attributed either to "Yes" or "No", and group decision is considered as "Non-biased".

The same heuristic applies in case of equal weights, with putting all weights as:

$$w_1 = w_2 = \ldots = w_i = \ldots\ldots = w_n = 1/n.$$

Allowing for null FESs:

In some situation, when some FES has some reasons for not to participate; for instance due to insufficient knowledge or any other reasons, then, the effect of such null or idle FESs situation should be assessed or included the in the final decision. This requirement could be satisfied by considering the output of such FES as exactly at the middle, equal to 5, which means not “Yes” and not “No”. Here, I have allowed for null case and allowed it to influence the decision solution. Then, FESs’ outputs can be combined utilizing the HFS-model described in chapter 8 for selecting an adequate combining criterion, or based on consensus heuristics described in chapter 6. The same holds in case of equal weights.

Providing for related FESs’ decisions:

In certain circumstances, we may find that some FES’s opinions are related. In other words, the domain knowledge of a FES is related to the domain knowledge of another one in making the decision. This means that if some related FESs’ decisions agree, then this should contribute to increasing the reliability of their decision direction, in other meaning increase our confidence about their decision direction. On the other hand, if the decisions of some related FESs disagree, then this should decrease the reliability of both of their decisions. Therefore, in order to provide for this requirement, it may be more useful or necessary to investigate some related FESs’ decisions in a preceding stage, and to combine their outputs before their combined output being manipulated in a subsequent stage with the other remaining outputs in the group. This in order to reflect the effect of their agreement or disagreement on the group decision. Special considerations should be put to this circumstance. My suggestion is to logically reflect the effect of mutually related expert systems and to combine their outputs separately (see figure 10.1). It logically follows that if two or more FESs’ decisions are related, then in case of their agreement, this should reinforce their direction opinions; and if their outputs do not agree, then their direction opinions should be weakened. How to express that in form of numerical values? I suggest simple, logical, and mathematical formulas to express this idea. In case of agreement of two or more FESs’ opinions, there are two possible sub-cases; the first sub-case in which the combined output values are all above the middle value of the chosen output scale, 5 (i.e., toward the “Yes” decision direction). In this sub-case we need a formula which slightly gives a larger resultant values than the combined output values, and that means positively reinforcing or strengthening the agreeing opinions. In the second sub-case of opinions agreement, the combined output values are below the middle value. In this sub-case we need a formula that reinforce the opinions but in the other same direction opinion (i.e., toward the No decision direction); that is the formula which gives a resultant value that is slightly less the combined ones in order to further strengthen their direction. On the other hand, in case of disagreement; that is the FESs’ opinions are split toward disparate direction opinions, what is needed here is to decay or weaken both direction opinions due to the effect of opinions conflict, to obtain a compromising resultant value. Three simple mathematical formulas are used to mathematically approximate and express these notions to introduce the effect of mutual relatedness.

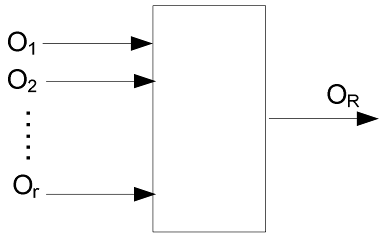

Fig. 10.1 Combining related FESs' decisions.

Simply, for a number r of related ESs: FES_1 , FES_2 ,…… FES_r out of an n total number, and producing outputs O_1 , O_2 , …. O_r respectively, where $2 \le r \le n$, then there are two possible circumstances:

(I) In case of agreement: there are two sub-cases:

Positive reinforcement: when two or more related output values are all above the middle, in this case, I suggest using the Root Mean Square formula (eq. 10.3) to slightly magnify these output values, giving a resultant or combined reinforced output:

$$O_R = \sqrt{(O_1)^2 + (O_2)^2 + \ldots\ldots + (O_r)^2} \tag{10.3}$$

Negative reinforcement: when two or more related output values are all below the middle, in this case, I suggest using the formula in equation 10.4, to slightly decrease these output values giving a resultant or combined reinforced value that always slightly less than all these output values:

$$O_R = \frac{\sqrt{O_1 + O_2 + \ldots. + O_r}}{r} \tag{10.4}$$

In both circumstances, the associated combined weight is the sum of their importance's:

$$w_R = w_1 + w_2 + \ldots\ldots\ldots. + w_r \tag{10.5}$$

(II) In case of disagreement: it is only one circumstance, in which the two related opinions are disparate. In order to combine the two outputs I suggest using the Arithmetic Mean (AM) formula, which slightly attenuates both values giving a resultant compromised output as in the following equation:

$$O_R = \frac{O_1 + O_2}{r} \tag{10.6}$$

The resultant weight is computed again same as in equation 10.5.

Note 1: in case of more than two related opinions that are not all agreeing, then the agreeing sub-groups are combined using equations 10.3 through 10.5, then the resultant combined values of the two sub-groups are then combined using equation 10.6.

Note 2: if there are three related outputs that are found to be disparate; that is one value is above the middle, the second is below the middle, and the third is equal 5, then their combined output is the obtained using the AM of the three values.

After separately combining related decisions, the resultant combined value of these related decisions is dealt with and its associate weight as a new output decision. After, a consensus-based heuristic (chapter 6) or the developed HFS-model (chapter 8) could be used as convenient to combine all the set of FESs' outputs. The same procedure applies in case of equal weights.

Provision for critical or veto-type decisions:

It could happen that one or more FESs within the participating group have some special kind of importance to the group decision. This importance is not only reflected in form of weights, but rather it appears under certain circumstance, when the output values of such critical FESs happen to exhibit complete bias, and be equal either 0 or 10; in this case with that privilege given to these FESs, all other group decisions are forced to accept this veto, or in other meaning the group decision becomes that of the critical FESs. for instance, if only one FES bears that veto-type privilege, then if it happens that the output of such FES becomes equal to 0 or 10, then the group decision becomes "No" or "Yes" respectively depending on the actual case. If there are more than one FES deemed as critical or have veto-type decisions, then if it happens that they are all has a complete bias, either to 0 or 10, then the resort to the rule of majority voting among the critical sub-group to find the decision of high number of voting's, if still a tie, then this tie is broken using the weighted voting rule only among the critical sub-group. If still the tie can not be broken; for instance when two FESs of same weights has both this veto-type privilege, and they are conflicting, that is one has output 10, and the other 0, then their privileges are canceled, and in this case the group decision is made either using consensus-based heuristic or using the aforementioned HFS-model. The use and assignment of this veto-type privilege is based on the subjective choice of the decision analysts. Its contribution to the decision making process is to increase the reliability of the obtained decision as much as possible, and as it deserves so. The same notions apply for the case of equal weights.

Next section, the provision for the satisfaction of these requirements will be considered.

10.2 Satisfying the specific requirements for knowledge-unique FESs

This section is concerned with satisfying the specific requirements for the case of knowledge-unique FESs. The difference between aggregating knowledge-unique FESs and combining knowledge-equal FESs has been described in chapter 3. In the case of having knowledge-unique FESs, every FES represents a unique distinctive knowledge and expertise of a specialization area, and the solution of the complex problem should be reached through investigating and considering all these multiple expertise's or specializations. This is in order to have complete or comprehensive decision solution. The aggregation heuristic presented in chapter 5 aggregates or accumulates the outputs of these unique FESs, through assigning a proportional output scale to each FES, based on its weight proportion. The sum of these output scales gives the total scale. Then, each FES should produce its crisp output expressed within the range of its assigned scale. All individual scales and the total scale are parallel to the established numerical scale within [0,10]; in that the minimum values of each of this scales is 0, and means complete bias toward "No" , and the maximum value of each individual scale refers to complete bias toward “Yes”. Then, all the outputs produced are added or accumulated to give a final group output. If this summed value is above the middle of the total scale, then the group decision is “Yes”; if it is below the middle of the total scale, then the decision is “No”; otherwise the decision is not classified on “Non-biased”. Now, after reviewing the notion of that aggregation heuristic, let us consider the satisfaction of these requirements

Preserving extremes:

The aggregation heuristic presented in chapter 5, and reviewed above does not impose any smoothing effect through averaging like does the AM, which causes loss of extremes. This heuristic only adds the outputs, so the extreme value of each individual scale is preserved through addition. Thus, the use of such aggregation heuristic already provides for satisfaction of this requirement. The heuristic is used for both equal and different weights cases.

Allowing for null FESs:

Similarly, in order to allow for null participation, the output of null FES is set at the middle value of its allocated, corresponding scale. The same holds in case of equal weights.

Providing for related FES’s decisions:

The aggregated FESs essentially represent unique, different domain expertise’s. The mutual relatedness or interactions among them are in most cases non-existing, but if it happens that some domains are mutually related, their output numerical values are still aggregated. If their decision outputs agree, they will be added and accumulated and augment confidence about their direction, and vice versa. It means that the effect of their mutual relatedness is still naturally included. An attempt to develop a mathematical formula to provide for such effect in case of different sizes of allocated scales for domain expertise’s can lead to inconsistencies and complications, but it can be investigated in future researches. For simplicity here, I am assuming that the different, unique expertise’s do not have any mutual decision relatedness; they only participate in judging the same decision problem with no mutual influence or interaction.

Provision for critical or veto-type decisions:

Similarly, as it was possible with the case of knowledge-equal, it could happen that the decisions of one or more FESs within the participating set are deemed critical to the group decision. These critical decisions may acquire veto-type privilege. Similarly, when the output value of such critical FES happen to exhibit complete bias, and becomes equal to either 0 or the maximum of its individual scale; in this case with that privilege given to such FES, all other group decision are forced to accept this veto, or in other meaning the group decision becomes that of the critical FES. for instance, if only one FES bears that veto-type privilege, then if it happens that the output of such FES becomes equal to 0 or scale maximum, then the group decision becomes “No” or “Yes” respectively depending on the actual case. If there are more than one FES deemed as critical or have veto-type decision, then the tie or conflict can be resolved using the same idea described in the previous section for the provision for this requirement in case of knowledge-equal FESs. If still the tie can not be broken, then the privileges are canceled, and in this case the group decision is made using the described aggregation heuristic. The same notions apply for the case of equal weights.

It was clear that the established psychometric numerical scale for FESs' crisp outputs has been flexible to handle and satisfy requirements. It has allowed for getting more detailed information about the crisp outputs, their degree of bias toward either extremes, and has allowed to develop the heuristic that preserve extremes. This not only the contribution of this flexibility in requirement satisfaction, it has also helped to account for null participation of some FESs through simply substituting the crisp output of such FESs with the middle value, and enable this value or status to appear and affect the group decision. It also has allowed for satisfying the requirements of relatedness and permitted for expressing this relatedness through mathematical equations, which would be impossible without utilizing such numerical scale. Also, it has helped in satisfying the last requirement of providing for critical or veto-type FESs. The main desirable feature of the established numerical scale pertaining to requirements satisfaction is its ability to convert the logical or conceptual notions into objective or concrete, practical ones.

Actually, satisfying various imposed requirements and restrictions is necessary for the developed theoretical work to be viable, acceptable, and realistic. Finally, it should be noted that the proposed approaches and suggestions to satisfy the imposed requirements are simple, practical and computationally not expensive.

Next chapter, the overall results of this thesis will be stated, conclusions will be made, and recommendation for future research will be suggested.

Chapter 11

Potential application: automatic on-line analysis of received customer declarations

This study has been essentially inspired by an actual need of a currently held project of developing a decision making system for automatic on-line analysis of received customer declarations managed by the Custom Administration of the Czech Republic. Every custom declaration transaction should be judged as being either suspicious or unsuspicious, based on the information and data found in the documents provided with those transactions by exporters or importers. The custom transactions involve all possible kinds of commodities. Every similar group of commodities requires certain type of criteria, rules and expertise's to judge its validity and correctness. For instance, cars and their accessories constitute one similar sub-group. Similarly, all kinds of cloths and similar materials constitute another sub-group, and so on. There are a vast diversity of business-level criteria and experts' rules so called the Risk Profiles that are used to judge transactions. Each sub-group of these criteria and expertise's is pertaining to certain sub-group or similar kinds of transactions.

The importance of successful and reliable decision making ensues from the two kinds of risks involved and their associated costs. The first kind of risk is involved in the possibility of incorrectly detecting truly wrong, illegal, or invalid transaction as being unsuspicious. The associated cost with this first kind is the loss of customs in terms of money, in addition to other possible subjective hazards of unknown, uncounted, or uncontrolled commodities. The second type of risk is involved in the possibility to judge truly legal, correct or valid transaction as being suspicious. The associated costs with this second type of risks may be in terms of time needed to fully inspect the transaction, the unsatisfied importers or exporters, and may be loss of customs if the transaction is to be rejected without inspection.

During the year 2005, the Custom Administration developed a system for an automatic on-line analysis of received custom declarations-ARCD system. This system is focused mainly to the online analysis of processed declarations but at the same time it allows an automatic adaptation of the analysis process according to the ex-post analysis provided by various data mining tools from the data warehouse. The main drawback of this data mining tools is their lack of explanation facility, such that if a suspicious transaction is found, it is usually already closed and it is not possible to change it. That is there is no way to analyze the reasons of suspiciousness and to tell how to help adjust the transaction after being deemed as suspicious. The system is projected to process all custom declarations accepted in by the Customs Administration. However, it is currently running in a pilot mode and processes only a subset of all declarations. So far, the knowledge base of the system contains several hundreds of those business-level criteria, which vary in complexity from simple one condition criteria to complex algorithm criteria. They currently utilize RETE based inference engine which achieves reasonable response times from several hundreds of milliseconds to several seconds depending on a complexity of the processed declaration, and in the current mode it processes 1000 declarations per day. This amount is projected to grow approximately to 10000 declarations per day. The evaluation is currently based on a crisp decision making.

One potentially planned application of the proposed study is to allow for approximate reasoning appropriate for the nature of such ill-structured type of decision making problems, and to build a fuzzy decision making system for such ARCD system employed through constructing multiple FESs, each of which incorporates a homogenous subset of expertise's, and business-level criteria that are relevant to the judgment of certain subgroups of custom declaration transactions. These multiple FESs should integratively cover the judgment needs of all possible kinds of decision making or custom declaration transactions. The practical reasons for constructing multiple FESs instead of an aggregate large-scale expert system have been described in the introduction (chapter 1) of this thesis. However, the added values expected from the application of this study in the ARCD system can briefly be stated as follows:-

- Manipulating all possible kinds of input variables as needed and as relevant within each FES; quantitative, qualitative, vague, uncertain…,etc. Consequently, it is expected that the decision solutions obtained will be highly realistic due to the inclusion of all kinds of input variable, and at the same time involving all relevant decision aspects and expertise's

- Provision for explanation facilities and analysis generally possible with the utilization of rule-based expert system. This considered one of the prominent added values of the proposed system.

- Exploiting the capability of the intuitive If-then way of knowledge representation, which can pool together the knowledge's and expertise's obtained from multiple sources (e.g., the available experts and the business-level criteria) into unified knowledge chunks in form of If-then format.

- Building cohesive, compact, operationally efficient, and easily maintainable FESs. Compact FESs offer better operating performance and will assist in realizing the planned processing rate of 10000 declarations per day.

- Providing for flexible control through adopting the relevant subset of FESs appropriate for judging certain declaration transactions.

- Provision for more objective and detailed evaluation through the possibility of assessing the degree of the subjective decision options "Yes" and "No", and consequently be more able to evaluate the amount of risks involved, and be able to have more information supporting the final decision making. The utilization of the established psychometric numerical output scale provides for this possibility.

- The capability to consider and satisfy various imposed practical requirements in decision making as needed. Three out of the four specific requirements are real requirements, and were imposed by the need of this project. The solution approaches developed in this thesis have provided for how to satisfy these requirements (chapters 9 and 10).

- The capability to include and change the relative importance's of the utilized expertise's according to different decision making contexts (i.e., different decision making transactions or different sub-group of commodities). The weighting methods adopted in chapter 4 provides for this capability.

- Increasing the reliability of the obtained decision through exploiting multiple relevant expertise’s and knowledge sources.

Actually, it should be noted that there is no restricted application area for this research work. It could be applied in any ill-structured, critical and complex, binary decision making problems. This study could be applied in business, economics, medicine, engineering, military..., etc, as long as the decision making problem is of the aforementioned nature. Some examples possible application of this research are: new product launching decision, food quality tracking, monitoring of suspicious deviation of the business processes from the standard performance, control and logistic of food chains/networks, diagnosis of faults in very sophisticated appliances, critical medical diagnosis, diagnosis of integrity of aero-planes, evaluating Yes-or-No critical military decisions…, etc.

Chapter 12

Overall results and conclusions

The problem of integrating multiple FESs, while satisfying some restricting requirements, has been considered and dealt with in this thesis. The idea of the problem was initially inspired by a practical project carried out at the Custom Administration of the Czech Republic, where it is required to detect suspicious custom declaration transactions, based on the data and information provided by the exporters and importers. This problem belongs to the type of ill-structured, multi-aspect decision making problems, where the reliance on multiple expertise's is the most adequate choice, in order to cope with difficulty and complexity inherent with such types of decision making problems. These multiple expertise's are required in order to be comprehensible in taking account for all relevant decision aspects and to obtain more reliable group decision. The design solution was to construct multiple FESs to model the required expertise's, and to be able to handle, vague, subjective, and uncertain input variables inherent in such ill-structured decision making. Also, the adoption of this design solution was due to some practical reasons, like improving maintainability and control of individual FESs, provision for knowledge cohesion and modularity, avoidance of knowledge interaction and mutual influence, preserving the security of aggregate business knowledge, and improving performance features of individual compact systems.

The integration problem was then generalized into a binary-classification GDM problem. This enables to exploit solution approaches previously developed in the fields of binary classification and GDM, and in order to produce a more generally useful solution to the problem. The integration of FESs was proposed through combining or aggregating the crisp outputs of these systems, and taking into account all imposed general and specific requirements. Then, the main intended aim of this research was formulated, which is to realize objective integration of multiple, relevant FESs, through adoption and development of reliable and effective combining/aggregating heuristics, criteria, methods, or models, and at the same time satisfying the imposed restricting requirements. Based on this main aim, other specific objectives were developed in order to simplify the realization of this main goal.

In this chapter, the results and achievements regarding the realization of the research objectives will be briefly stated. Then, conclusions will be made, pertaining to the features and characteristics of the adopted and developed solution approaches, the results, benefits, and contributions gained from this research work. Finally, recommendations for future researches will be suggested.

Next section, the overall results of this research work pertaining to the realization of objectives will be briefly stated.

12.1 The overall results

In order to summarize the research efforts made in this research work, and to show that all the planned objectives have been achieved, a follow up is to be made, in which each specific objective will be followed by the achievements regarding the realization of such objective. A follow up of objectives realization is as follows:-

A. Structuring the problem

- I have established an objective psychometric numerical scale unified and standardized for output decisions produced by the individual FESs (chapter 3). This scale is used to express the degree of bias toward either “Yes” or “No” decision. This objective numerical scale enable adoption and development of a variety of combining/aggregating approaches ranging from simple to more sophisticated methods.

- I have established a conceptual structure for the integration problem. The problem has been logically structured into combining knowledge-equal FESs and aggregating knowledge-unique FESs (chapter 3). This is in order to adopt and develop adequate solution approaches for both cases.

- I have structured the candidate and possible combination/aggregation approaches and methods according to the general requirements and possible decision making contexts (chapter 3). This has enabled organizing these solution approaches, understand their roles, and integrate them when applicable.

- Based on the established objective outputs’ scale, the combination/aggregation problem was formally stated (chapter 3).

B. Adopting and developing adequate combining/aggregating methods

- I have adopted different methods for weighting the importance’s of the participating FESs, and under different possible decision making contexts (chapter 4). The weighting was based on either the AHP, utilizing current or present evaluation information, or the rate of misclassifications performance of the modeled expertise’s, utilizing past available performance data.

- For the problem case of combining the outputs of knowledge-equal FESs, I have adopted and configured the classical combining criteria like AM, GM, and MV, according to the established meaningful numerical scale (chapter 5). Further, I have developed a more decisive combining criterion, the MPDI, and described its distinct desirable features (chapter 5). Also, I have developed a consensus-analysis framework, and consensus-based heuristics (chapter 6). One of these developed heuristics was used to find a datum level, upon which the new MPDI criterion was compared to the classical existing ones (chapter 7). A HFS-based model was then developed to select the most adequate combining criterion based on the incoming set of FESs’ judgments (chapter 8).

- For the problem case of aggregating the outputs of knowledge-unique FESs, I have developed an adequate and simple aggregation heuristic (chapter 5).

- For the requirement of handling past performance data, I have proposed Multi-layer Feed-forward Back-propagation neural networks to learn the numerical data patterns of the expertise’s’ past performance (chapter 9). Further, I have developed A HFS-based model to combine/aggregate the outputs of FESs, when past If-then knowledge is available about the expertise’s’ performance (chapter 9).

C. Satisfying specific requirements

- I have developed an extreme-preserving heuristic that combines the crisp outputs of FESs, while keeping and adhering to extremes (chapter 10). For the case of aggregating knowledge-unique FESs, the developed aggregation heuristic already preserve extremes of FESs' outputs though addition or accumulation (chapter 5).

- I have specified how to provide for null FESs' outputs (chapter 10).

- I have specified how to provide for related FESs' decisions, and how to express this relatedness mathematically (chapter 10).

- I have provided practical suggestions for how to provide for critical or veto-type decisions.

Based on the above follow up, I have succeeded to realize all the established specific research objectives. Over and above, I have also realized the main aim of this dissertation work, which is realization of objective integration, since all the adopted and developed approaches are practical and viable.

In order to summarize the relationship between the developed solution approaches and the provision for the imposed requirements, the roles and relevancy of these solution approaches to the provision for the imposed requirements are shown in table 12.1. As indicated in the table, the binary combining criteria like MV (chapter 5) and its weighted version, WMV, utilize present information in order to combine abstract level information (i.e., "Yes" or "No") about the FESs' outputs. It is also used to break the tie among several holders of veto-type decisions, as described in chapter 10. The continuous or measurement-level combining criteria, like AM, and MPDI (chapter 5) handle present output information, and manipulate continuous-valued outputs. The developed aggregation heuristic (chapter 5) utilizes present information, manipulates continuous-valued outputs, and aggregates outputs of knowledge-unique FESs. It also preserves extreme output values as has been described in chapter 10. The consensus-based heuristics (chapter 6) utilize present outputs' information, manipulate continuous-valued outputs, and are used in combining outputs of knowledge-equal FESs. The HFS-based model for criteria selection (chapter 8) utilizes present information, manipulates continuous-valued outputs, and is used to select the most adequate combining criteria, based on the incoming FESs' outputs. The proposed Multi-layer Feed-forward Back-propagation networks (chapter 9) utilize past outputs' information for training, and present information for classification or decision making. Also, they manipulate continuous-valued outputs, and can be used for combination or aggregation. Actually, the BPN can provides for the satisfaction of the three specific requirements as indicated, because it can learns them all or they can be embedded easily in form of examples within the training data, and this considered one the prominent advantage of BPN; the flexibility to handle many kinds of imposed specific requirements. The same holds for the HFS-based model (chapter 9) developed for combining or aggregating the FESs' outputs. This model is based on past information, and utilizes the present outputs' information in decision making. It manipulates continuous-valued outputs, and can be used for both combination and aggregation. Also, the model can provide for satisfaction of the three specific requirements as indicated, since the decision logic of the model could incorporates decision rules that work to satisfy these requirements. The extreme-

Table 12.1 The roles and relevancy of the adopted and developed solution approaches to the provision for satisfying requirements.

Solution approach	**General requirements**						**Specific requirements**			
	Information Handled		**Outputs manipulated**		**Case of integration**		**Preserving extremes**	**Provision for related decisions**	**Provision for null decisions**	**Provision for veto-type decisions**
	Present	**Past**	**Binary**	**Continuous**	**Combining**	**Aggregating**				
Binary combining criteria: MV, WMV	ℜ	----	ℜ	----	ℜ	----	----	----	----	ℜ
Continuous combining criteria: AM, WAM, MPDI, WMPDI	ℜ	----	----	ℜ	ℜ	----	----	----	----	---
Aggregation heuristic	ℜ	----	----	ℜ	----	ℜ	ℜ	----	----	----
Consensus-based heuristics	ℜ	----	----	ℜ	ℜ	----	----	----	----	----
HFS-based model for criteria selection	ℜ	----	----	ℜ	ℜ	----	----	----	----	----
BPN	ℜ	ℜ	----	ℜ	ℜ	ℜ	ℜ	ℜ	----	ℜ
HFS-based model for combining/aggregating outputs of FESs	ℜ	ℜ	----	ℜ	ℜ	ℜ	ℜ	ℜ	----	ℜ
Extreme-preserving heuristic	ℜ	----	----	ℜ	ℜ	----	ℜ	----	----	----
Practical suggestions/ mathematical formulas	ℜ	----	----	ℜ	ℜ	ℜ	----	ℜ	ℜ	ℜ

ℜ: Relevant.

preserving heuristic (chapter10) handles present information, manipulates continuous-valued outputs, and used to combine outputs of knowledge-equal FESs, while preserving extreme output values. Finally, the practical suggestions and mathematical formulas presented in chapter 10, utilize present information, works with continuous-valued outputs, and developed for both combination and aggregation. They are mainly intended for satisfying the four specific requirements.

In overall, the presented results provides a theoretical framework toward objectively integrate multiple FESs evaluating binary decision making problems. The developed solution approaches are practical, viable, and improve group decision making process in particular, and bring advantages to the ill-structured binary decision making in general.

Next section, the main conclusions drawn from this research work will be briefly stated.

12.2 Conclusions

In this research study, objective solution approaches have been presented to the problem of integrating multiple FESs. Several conclusions can be drawn from the results of this research. They are discussed briefly as follows:-

- The idea of this research is a novel and a pioneer step toward objectively integrating systems through combining/aggregating their final outputs.

- This research encourages practitioners to take up and implement the integration of knowledge-based systems, as long as there exist objective and practical methods to realize such integration.

- The use of the psychometric numerical scale to express the degree of bias toward subjective “Yes” or “No” decisions offered high flexibility in adopting and developing more sophisticated combining/aggregating methods, and helped in satisfying requirements.

- The newly developed MPDI combining criterion has been proven more decisive than the AM.

- AHP has offered great flexibility, simplicity, and effectiveness in weighting the importance’s of FESs under different decision making circumstances.

- HFS-based models have reduced the complexity associated with building large rule-base and have facilitated logical structuring of both the model for criteria selection and that for combining/aggregating FESs’ decisions.

- Consensus-based heuristics are more reliable and explanatory than simple combining criteria, and add value to the fields of GDM and social choice theory, in form of more objective and thorough methods for evaluating and exploiting consensus information.

- ANN, in general, is the most adequate choice for learning past performance patterns and the BPN in particular is highly adaptive, and offers the ability and flexibility to satisfy majority of the imposed specific requirements.

- There have been no formal techniques for combining/aggregating numerical outputs of several knowledge sources or systems. In this study, I have offered formal and universal techniques that could be used to more objectively integrate multiple knowledge sources or knowledge-based systems.

- The existence of past record of real data is crucial, which enables utilizing the artificial intelligence techniques like BPN, and HFS-based model to learn, understand and map the relationships among FESs' judgments and correct decision, specific to a particular decision making process.

- The accomplished results have broad applicability in dealing with any complex ill-structured, multi-aspect, binary decision making problems.

- The main sources of strength of the proposed FESs' integration are:-

  Realism: more realistic solutions are obtained from integrating multiple FESs representing all relevant qualitative and quantitative aspects together, and representing may be different domain knowledge's and expertise's. Since FESs are able to treat uncertainty, vagueness, and subjective factors within every modeled expertise, the resulting solution is more realistic and comprehensive, than the case with single or conventional non-fuzzy-based expert systems.

  High decision quality: the decision quality and reliability are logically improved through inclusiveness and comprehensibility.

  Flexibility: it is attributed to the ability to selectively meet various problem contexts through complementing and matching several relevant FESs to suit particular needs of the current decision problem.

  Improved maintainability, control and operational efficiency: more compact expert systems are easier to maintain, control and to efficiently operate.

  Reduced risk propensity: the utilization of multiple knowledge's and expertise's clearly contribute to reduction of the risk involved in conducting the decision process.

Next section, recommendations for future research efforts will be suggested.

12.3 Recommendations for future research

The following researches could be conducted in future as extensions to this research work:-

- Neuro-fuzzy hybrid models could be investigated in handling past experts' performance data and knowledge. They could be used as classifiers with added explanatory characteristics and enhanced generalization capability of neural nets. This beside the advantages gained from the capability of fuzzy systems to model vague and inexact relationships associated with the ill-structured decision process. Also, neuro-fuzzy models have the capability of adaptive learning not existing in a stand-alone fuzzy system.

- Cooperative neuro-fuzzy models could be investigated. In such models, Kohonen SOFM is utilized to extract the linguistic If-then rules to explicitly represent the implicit knowledge within the available expertise's' past performance data. These rules will be then fed to a fuzzy system. This contributes to adding the adaptive learning capability to the explanation facility already existing in fuzzy systems.

- The possibility of utilizing Fuzzy Clustering Algorithms to extract decision rules from the experts' performance data could be investigated.

- Clustering approaches like Kohonen SOFM, k-means algorithm, or k-Nearest Neighbor algorithm could be utilized as classifiers to form clusters of similar expert systems judgment vectors found in the available past expertise's' performance data with known decision answers, based on some adequate similarity measures. New judgment vectors are to be classified based on which clusters they are most closely belong.

- The Multi-variate statistical techniques like Factor analysis, and Canonical Correlation, could be utilized to study and analyze the dependency structure found in the past expertise's performance data. This could be useful in determining which expert systems are related and should to be combined separately.

- A selection model could be developed to help in deciding which FESs are currently most relevant to the incoming decision transaction, based on some predetermined identified factors.

- The possibility to develop new other combining and aggregating methods could be investigated. Also, improvement on existing and developed ones could be investigated. The objective may be to search more comprehensive criteria in satisfying requirements or devising new criteria with more desirable feature like simplicity, reliability…etc.

- New improved consensus-based heuristics could be developed to exploit the established consensus-analysis framework. More experimentations and investigations could improve the developed heuristics.

- Future researches could consider the integration of the developed solution approaches in this work, in order to have more inclusive and integrative approach that satisfies all the imposed requirements simultaneously.

- Future researches may consider the generalization of the integration problem from a binary decision making into multi-option or multi-alternative decision making.