韓國國防經營分析學會誌
第 26 券, 第 2 號, 2000. 12. 31.

# A Study on Dynamic Inference for a Knowledge-Based System with Fuzzy Production Rules

Song, Soo Sup*

## Abstract

A knowledge-based with production rules is a representation of static knowledge of an expert. On the other hand, a real system such as the stock market is dynamic in nature. Therefore we need a method to reflect the dynamic nature of a system when we make inferences with a knowledge-based system.

This paper suggests a strategy of dynamic inference that can be used to take into account the dynamic behavior of decision-making with the knowledge-based system consisted of fuzzy production rules. A degree of match(DM) between actual input information and a condition of a rule is represented by a value [0, 1]. Weights of relative importance of attributes in a rule are obtained by the AHP(Analytic Hierarchy Process) method. Then these weights are applied as exponents for the DM, and the DMs in a rule are combined, with the Min operator, into a single DM for the rule. In this way, the importance of attributes of a rule, which can be changed from time to time, can be reflected in an inference with fuzzy production systems.

* Sejong University

# 1. Introduction

Decision-making is a very important managerial activity in any kind of organization. Many quantitative models based on mathematics have been developed to assist decision making and efforts to develop better models still continue. The limits of these quantitative models are that they need mathematical precision of input data and the model itself. However, most human decisions involve many elements that cannot be quantified with mathematical precision. In most decision-making situations, human beings recognize and analyze things based on past experiences and knowledge. In general these recognitions are ambiguous and imprecise.

With the invention of the computer and the proliferation of its use, the concept of a decision support system(DSS) has been introduced and applied in various areas of managerial decision-making. Recent research efforts on artificial intelligence and expert systems (knowledge-based systems) provide another aspect of utilizing computers for decision-making. A knowledge-based system attempts to program experts' experience, knowledge and logical processes which lead to their decisions. Thus, the knowledge-based systems are expected to improve the effectiveness and efficiency of decisions.

A production system, the core of a knowledge-based system, is primarily a formalism for representing knowledge about any particular area of problem solving. A program written as a production system is a collection of "production rules," which takes the form of "If left-hand-side proposition is true, then right-hand-side proposition is true." The production rules may be executed in one or both of two modes. First, it concludes that the right-hand-side proposition is true when the left-hand-side proposition is true, which is referred to as "modus ponens." Second, it tries to find a left-hand-side proposition from the right hand proposition which is the closest to a goal to achieve, which is referred to as "modus tollens."

One of the most important features of a production system is that individual IF-THEN production rules interact with each other by means of the database rather than by means of traditional programming control structures. Production systems in which knowledge is represented by production rule formalism are highly modular.

Therefore knowledge of an expert in the area to be covered by the production system can be elicited in chunks, and subsequently each chunk becomes a production rule which represents the expert' s answer to a "what if" question. This way of organizing knowledge in discrete chunks which interact with each other only by means of database is a very natural way of modeling human cognitive processes, because it comes closer than any other feasible programming technique to mirroring the actual process by which a human expert performs his decision-making task [1].

In the real world, human knowledge includes large portions of uncertainty, and he or she makes decisions with information involving uncertainties and imprecision. Thus the knowledge represented by production rules in a knowledge based system includes various uncertainties. Linguistic variables based on a fuzzy set theory have been utilized as a means of representing and processing these uncertainties. Fuzzy linguistic variables provide a formal structure for translating between imprecise or uncertain numeric measurement scales and words and phrases from a standardized subset of natural language. Because linguistic variables can code natural languages and process them by mathematical rules, they have been used and studied in various domains since Zadeh [2] proposed the concept of linguistic variables. In the fuzzy production system, in which an expert's knowledge is represented and processed with linguistic variables, various methods of inference have been invented. This paper proposes a method of inferencing within the fuzzy production system in which an inference is made after considering the relative importance, with respect to the consequence, of attributes included in the condition portions of production rules.

## 2. Representation of Uncertainty with Fuzzy Set

Most information used by people for decision-making is incomplete or partially true due to uncertainty in the language, or interpretation representing the information and incompatibility of information from different sources, etc. These uncertainties exist in the occurrence of events and causal relations or interrelations among the various events in the real world. Thus human knowledge includes a lot of uncertainties and a decision is made in the circumstances with uncertain and ambiguous rules or facts.

In the knowledge-based system to support the human decisions, we need a scheme representing and processing this uncertain and imprecise knowledge. Knowledge in a knowledge-base is represented by a rule which consists of a consequent and one or many antecedents.

There are multiple sources and types of uncertainty in knowledge-based systems. For example, uncertainty may occur due to less than perfect reliability of measurements and probabilistic information. Uncertainty can also arise due to inferential mechanisms within a knowledge-based system. For example, in a weak implication, antecedents and consequents are only partially correlated. Another source of uncertainty is lexical imprecision [3]. Lexical imprecision is a lack of uniform meanings for the semantic primitives used to represent knowledge.

The fuzzy set theory provides the mathematical tools for developing a common framework for representing different aspects of the reasoning process and the vagueness associated with a partial match between the antecedent part of the rules and the input data, and the uncertainty in asserting the consequent on the basis of the occurrence of the events appearing in the antecedent. A further advantage of the fuzzy set theory is the provision of fuzzy linguistic variables. The fuzzy linguistic variables seem to be the formal tool which deals in the most appropriate way with the adjectives used by human experts. Thus the introduction of linguistic terms can simplify the rules, at least as regards their readability and understandability to the user.

## 2.1 Linguistic variables

The use of linguistic variables for approximate reasoning has attracted great interest since Zadeh's [2] introduction of the concept. A linguistic variable is a fuzzy variable associated with an appropriate linguistic syntax and semantics. In general, human knowledge is neither so precise as to be well represented by ordinary of mathematical variables nor so vague as to be vacuous. The linguistic variable is very useful to represent and process such human knowledge in the computer.

Whereas linguistic variables provide a means of describing or quantifying imprecise or uncertain facts, "approximate reasoning" refers to the need to correspondingly account for uncertainty or imprecision in the "If-Then" type rules upon which knowledge based systems are founded. Beginning with Wenstop [4], many researchers have proposed special purpose linguistic processing systems to support specific applications. General purpose systems include Wenstop's [5] system, Fuzzy and L-Fuzzy [6], Pruf [7], and Fril [8]. Among these systems, Wensop's language system has provided a very useful foundation on which to build operators and vocabularies.

Wenstop's linguistic processing system is based on five primary fuzzy sets "low," "medium," "high," "unknown," and "undefined." It also utilizes 14 hedges such as "above," 6 connectives such as "and," 3 trend mode such as "increasingly," and 3 trend direction such as "falling" and others. The primary terms are characterized by both their membership functions and their names such as "high." The membership function is represented by base values and a corresponding truth-value for each element of base value. For example, if a security analyst wants to express a large number of trade volume of a specific stock in the security market by the term "high" , he can specify a minimum and maximum of the trade volume, divide the range (max-min) into appropriate numbers of intervals, and then assign a truth-value [0,1] for the representative value of each interval.

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Base value | 5 | 7 | 9 | 11 | 13 | 15 | 17 | 19 | 21 | 23 | 25 (hundred thousand) |
| Truth-value | 0 | 0 | 0 | 0 | 0 | 0 | 0.1 | 0.3 | 0.7 | 1 | 1 |

Here the truth-value is the fuzzy membership grade. The linguistic system is greatly simplified by requiring every linguistic term to be represented by a fixed length of a discrete vector.

A truth-value of the linguistic variable represents the possibility of a corresponding base value of the variable. A membership function, in principle, is continuous and the discrete base values are employed to approximate the continuous function. Thus the range and the number of discrete base values can be different according to the field or the attributes of the knowledge to be represented. For example, Gracia et al. [9] employed six discrete base values to represent the linguistic terms for linguistic evaluation in risky situations ( "high" = 0.0/1, 0.2/2, 0.4/3, 0.6/4, 0.8/5, 1.0/6). Membership values in the system with six base values overlap more than those with eleven base values, and thus boundaries among linguistic terms are further blurred. This implies that a linguistic system with six base values represents more uncertain situation than that with eleven base values. Therefore, it is important to select an appropriate number of base values and membership values depending on the field of application for the successful development of knowledge based systems.

# 3. A Knowledge-Based System with Fuzzy Production Rules

A knowledge-based system usually include a inference engine so called "production system". A production system is a collection of production rules which rely primarily on the form "IF (condition) - THEN (action)." Serious interest in the use of production systems to represent human problem solving ability dates from the research of Newell and Simon [1] in which they asserted that human knowledge is organized into 'logical chunks' which operate nearly independently. The IF clause of a production rule corresponds to the conditions under which a person will recall a particular chunk of knowledge, and the THEN clause corresponds to the result of applying that knowledge. The production systems have been used to capture the knowledge of experts in many areas of the so called 'expert system'.

Because classical production systems are rooted in Aristotelian logic, production rules must be executed in an "all or nothing" manner. In effect, all variables are treated as if they were measurable only on nominal scales. To make use of a variable measured on a numerical scale in the classical production system, the scale must be divided into regions and production rules whose left-hand-sides ask whether the variable is or is not located within a given region.

In order to mitigate the problems caused by "all or nothing" execution, some successful applications or production systems, such as MYCIN, introduce various sorts of uncertainty measure on top of the basic production system structure. The fuzzy production systems methodology has been introduced to resolve the "all or nothing" problem without harm to the advantages of the production systems.

The fuzzy production systems utilize many-valued logic instead of Aristotelian logic to execute IF-THEN rules [10,11,12,13]. This means that the system evaluates the agreement between the condition of the IF statement and the input data by the degree of agreement which is expressed in terms of a point on the interval between complete agreement and complete disagreement. Then the system executes the weakened action of the THEN statement, depending on the degree of agreement between the condition and the input data. To make this operation possible, the conditions and a consequent in the fuzzy production rules are consisted of fuzzy propositions that are types of fuzzy set.

In the classical production system, a rule is activated only when the input data perfectly matches the condition of the rule. Because the rules in a classical production system are activated sequentially, the system needs a method of conflict resolution policy when the input data satisfies two rules simultaneously to determine which one takes precedence. In a fuzzy production system, on the other hand, all the rules are executed during each pass through the system and then the result of the execution is expressed in terms of degree of match with strengths ranging from 'not at all' to 'completely' for the consequent of each rule. When the input data disagree perfectly with the conditions of the rules, the result of the execution is null.

Fuzzy production systems utilize linguistic variables to represent expert knowledge. Because the linguistic variable has flexibility in representing human knowledge with any degree of uncertainty, it is useful to express expert knowledge in complicated fields such as management, security analysis, forecasting, etc. Another advantage of the fuzzy production system is that it can apply many rules with overlapping for an inference. Thus we don't have to assign the input data to an interval of the condition of a rule by force to resolve borderline problems. Consequently, each rule can be applied to various situations.

## 4. A Strategy of Dynamic Inference

A decision-making problem can be defined as a classification problem. Because, in general, a decision-making problem can be formulated as the form in which the possible future results of the decision-making are classified as several exclusive and exhaustive classes and then a decision maker makes a decision based on information derived from analyses and forecasts based on a given current situation. For example, a stock investment decision would be made by forecasting whether the specific stock's price will be 'rise,' 'steady,' 'fall' based on analyses of the current stock market situation. Which types of classified results (classes) will occur in the future depend on various attributes affecting the situation, and an expert in the field will have knowledge about causal relationship between the attributes and the future results. Thus, such expert's knowledge can be programmed as a knowledge based system and then can be used to provide consistent advice to the novice.

Knowledge base consisted of production rules can be constructed in the following forms:

IF $X_1$ is $A_{11}$ and $X_2$ is $A_{12}$ ············ and $X_n$ is $A_{1n}$
THEN Class is $C_1$
IF $X_1$ is $A_{21}$ and $X_2$ is $A_{22}$ ············ and $X_n$ is $A_{2n}$
THEN Class is $C_2$
IF $X_1$ is $A_{31}$ and $X_2$ is $A_{32}$ ············ and $X_n$ is $A_{3n}$
THEN Class is $C_3$

.
.
.

IF $X_1$ is $A_{m1}$ and $X_2$ is $A_{m2}$ ············ and $X_n$ is $A_{mn}$
THEN Class is $C_m$

Each rule has n antecedent propositions at the most. Each of the labels $C_1$, $C_2$, ············, $C_m$ is the name of a class such as 'rise,' 'steady,' or 'fall' in a stock investment problem. $X_1$, $X_2$, ············, $X_n$ are attribute names such as 'economic growth rate,' 'the growth rate of industry sales,' 'PER for a individual firm,' etc. and $A_{11}$, ············, $A_{mn}$ are attribute values which are expressed by linguistic variables such as 'high,' 'medium,' 'low,' etc. in the fuzzy production system. A knowledge base consisting of these rules can be used to provide advice to a decision maker by executing the rules with the observed value or the expected value of each attribute.

Dynamic inferences considering degrees of importance of each attribute in the fuzzy production system are as follows.

## 4.1 Calculation of Degree of Match(DM)

There has been proposed a number of methods for evaluating fuzzy production systems to calculate and combine truth-values of multiple conditions in the antecedent of a rule [12,13,14,15]. For inferencing in the actual environment, input values of the attributes match to the condition partially in more cases than cases of complete match. There also exists exact quantitative data such as 'trade volume of stock' and ambiguous data which can not be exactly measured. Thus we need a method to process these partial matches for inferencing in a production system. Cho et. al [13] proposed a method that can be used to calculate the degree of match(DM) in various situations such as whether the condition and consequent part of the rules are precise, imprecise or any combination of both. The DM was expressed in terms of two new parameters, $CERT_+$ and $CERT_-$, defined as follows:

$CERT_+ = Max\{Min(\mu_A, \mu_{A*})\}$

$CERT_- = Max\{Min(1-\mu_A, \mu_{A*})\}$

where $\mu_A$ : the grade of membership of x in A on the rule X is A.

$\mu_{A*}$ : the grade of membership of x in $A^*$ on the input data X is $A^*$

$CERT_+$ measures to what extent A and $A^*$ are matched, and $CERT_-$ measures to what extent 1-A and $A^*$ are matched. Suppose, for an example, there are rule and input data as follows:

Rule: IF X is A then Y is B

Input data: X is $A^*$

where the value: X=[10, 20, 30, 40, 50, 60]

A={0/10, 0.2/20, 0.4/30, 0.6/40, 0.8/50, 1.0/60}

$A^*$={0/10, 0.03/20, 0.11/30, 0.25/40, 1.0/50, 0/60}

Then $CERT_+$ = Max[Min{(0,0), (0.2,0.03), (0.4,0.11), (0.6,0.25), (0.8,1.0), (1.0, 0)}]

= Max[0, 0.03, 0.11, 0.25, 0.8, 0] = 0.8

$CERT_-$ = Max[Min{(1,0), (0.8,0.03), (0.6,0.11), (0.4,0.25), (0.2,1.0), (0,0)}]

= Max[0, 0.03, 0.0.11, 0.25, 0.2, 0] =0.25

The DM is derived from the $CERT_+$ and the $CERT_-$ with respect to three cases. Cho et al [13] proposed two measures of DM; one takes the value of [-1, +1], another takes the value of [0, 1]. The latter gives the same results as Cayrol's method [14] and Leung's method [15]. Since this paper proposes a method evaluating production rules with the degree of importance of information at hand, the DM which takes the value of [0, 1] is employed. The value 1 means the antecedent part is satisfied completely, and 0 means the antecedent part is not satisfied at all, while 0.5 represents neutrality.

If $CERT_+ \geq 0.5$ and $CERT_- \leq 0.5$, DM = $CERT_+$,

If $CERT_+ \geq 0.5$ and $CERT_- \geq 0.5$, DM = $(CERT_+ - CERT_-) + 0.5$,

If $CERT_+ \leq 0.5$ and $CERT_- \geq 0.5$, DM = $(1 - CERT_-)$.

Since the example illustrated above corresponds to the first case, the DM = $CERT_+$ = 0.8. This value means the given information, "X is $A^*$", matches the antecedent part of the rule "X is A" with degree of 0.8 and thus the degree of appropriateness of the conclusion "Y is B" would be 0.8.

When there are many conditions connected by "and" in the antecedent part of a rule, the overall DM of the rule is calculated by $Min\{DM_1, DM_2, DM_3, \cdots, DM_n\}$. And when there exist two or more rules which have the same consequent, the overall DM of the knowledge base is calculated by Max{DM of rule 1, DM of rule 2, ⋯ , DM of rule m}. Therefore when we need to make a rule with conditions connected by "or" in the fuzzy production system, we can make a separate rule for each condition connected by "or". For example a rule "$X_1$ is $A_1$ OR $X_2$ is $A_2$ THEN Y is B" can be separated into two rules such as "$X_1$ is $A_1$ THEN Y is B" and "$X_2$ is $A_2$ THEN Y is B". These two forms of rules return the same result because all the rules are activated in an inference session of a fuzzy production system.

In constructing a knowledge base for an expert in the form of production rules, it is hard to extract detailed rules reflecting all the situations related to the domain of the knowledge base due to the limited memory and information processing capability of the expert. Therefore the inference utilizing partial matches makes possible a simple knowledge base constructed with representative knowledge which can embrace various types of input data (information) of inferencing in a production system.

## 4.2 Assessment of importance of information (input data)

In calculating DM, the input data is a body of collected information or interpreted result of collected information to be used to obtain advice from a knowledge base. We input various information to the knowledge base in order to obtain advice. Among this information, some classes of information affect the results more strongly than other classes in some situations, and different classes affect these more strongly in different times. For example, the stock price level in the stock market is affected by expectations of the investors, and thus different elements could be interpreted as being more important than other elements in different times by situations which form the investors' expectations. Accordingly the importance of information can be vary depending on the time or situation. But the importance of information can not be put into the knowledge base because the importance varies throughout time. Therefore we need a scheme to take into account the varying importance of information in evaluating the knowledge base to obtain more appropriate advice.

AHP(Analytic Hierachy Process) [16,17] has been widely used to derive a degree of importance for an attribute in multi-attribute decision-making situations. AHP derives relative importance of each attribute by repetitive pair-wise comparison of two attributes at a time. This process produces a weight($w_i$) for each attribute which takes a real number of greater than 0. These weights can be used as exponents to the attribute's value of [0, 1], to reflect the importance of attributes by which the greater the exponent the smaller the attribute's value of [0, 1]. This leads to the smallest attribute's value being the overall value in the multi-attribute decision-making situation [18].

Since the DM of each attribute of a rule takes a value of [0, 1] and an overall DM of the rule is $Min\{DM_1, DM_2, DM_3, \cdots, DM_n\}$, AHP can be used to reflect the importance of each attribute in evaluating a fuzzy production system in which multi-attributes are connected by "and" in the antecedent part of a rule.

Lets explain this process by an example. Suppose there is a production rule in a fuzzy production system, and a set of input data is given to evaluate the rule.

Rule: IF $X_1$ is $A_1$ and $X_2$ is $A_2$ and $X_3$ is $A_3$ then Y is B

Input data: $X_1$ is $A_1^*$, $X_2$ is $A_2^*$, $X_3$ is $A_3^*$.

Then the linguistic terms used in the Gracia. et al.[9] are employed as the fuzzy membership values of attributes and input data.

$A_1$(high)= {0.0/10, 0.2/20, 0.4/30, 0.6/40, 0.8/50, 1.0/60}

$A_2$(high)= {0.0/10, 0.2/20, 0.4/30, 0.6/40, 0.8/50, 1.0/60}

$A_3$(low)= {1.0/10, 0.8/20, 0.6/30, 0.4/40, 0.2/50, 0.0/60}

Input data set 1:

$A_1^*$(high) = {0.0/10, 0.2/20, 0.4/30, 0.6/40, 0.8/50, 1.0/60}

$A_2^*$(medium) = {0.1/10, 0.3/20, 1.0/30, 1.0/40, 0.3/50, 0.1/60}

$A_3^*$(rather low) = {0.0/10, 1.0/20, 0.25/30, 0.11/40, 0.03/50, 0.0/60}

The DMs for the above case can be calculated as follows:

Input data set 1: $A_1^*$, $A_2^*$, $A_3^*$

Condition 1($X_1$): $CERT_+ = 1$, $CERT_- = 0.4 \Rightarrow DM_1 = CERT_+ = 1$

Condition 2($X_2$): $CERT_+ = 0.6$, $CERT_- = 0.6 \Rightarrow DM_2 = (CERT_+ - CERT_-) + 0.5 = 0.5$

Condition 3($X_3$): $CERT_+ = 0.8$, $CERT_- = 0.25 \Rightarrow DM_3 = CERT_+ = 0.8$

Overall DM for the rule without importance weight = Min(1, 0.5, 0.8) = 0.5

The AHP can be applied to derive the importance of attributes as follows:

(a). Preparation of an attribute comparison matrix [16,17]

An attribute comparison matrix can be obtained by pair wise comparison of attributes' importance with respect to inferencing, which is the general procedure used in implementing the AHP method. Lets suppose we get a comparison matrix as below for an example.

| Attri-butes | $X_1$ | $X_2$ | $X_3$ |
|---|---|---|---|
| $X_1$ | 1 | 5 | 9 |
| $X_2$ | 1/5 | 1 | 3 |
| $X_3$ | 1/9 | 1/3 | 1 |

1: Equal importance
3: Weak importance of one over another
5: Strong importance
7: Very strong importance
9: Absolute importance
2,4,6,8: intermediate values between two adjacent scale judgements

The above matrix indicates that attribute $X_1$ has a strong importance (5), compare to the attribute $X_2$, and has an absolute importance(9) compared to $X_3$. It also indicates that $X_2$ has weak importances (3), compared to $X_3$. The values on the lower triangle of the matrix are reciprocals of the values on the upper triangle of the matrix.

(b). The next step is deriving the eigenvalue and eigenvector, after which the eigen vector with the largest eigenvalue is normalized, making the sum of the elements of the vector equal to one. Each element of the normalized vector is multiplied by the rank of the matrix to make the sum of weights as the same as the number of attributes, and then this value is considered as the importance weight of the corresponding attribute.

The eigenvalues and eigenvectors of the above matrix are as follows:

The largest eigenvalue: 3.03

Corresponding eigenvector: 0.969, 0.230, 0.091

Normalized eigenvector: 0.751, 0.178, 0.071

Importance weight: 3*(0.751, 0.178, 0.071) = 2.254, 0.535, 0.211

DM of the input data set 1: $\text{Min}((DM_1)^{2.254}, (DM_2)^{0.535}, (DM_3)^{0.211})$

$= \text{Min}(1.0, 0.69, 0.95) = 0.69$

In the previous example, $X_1$ is the most important attribute according to the matrix, and the overall DM of the rule is 0.5 for the given data set. The overall DM of the rule is changed, by applying the importance weight, to 0.69 from 0.5 due to the DM of the most important attribute $X_1(DM_1=1)$ is the largest for the data set. This results show that the importance of the attributes $X_1$ greatly affects the overall DM of the rule. Therefore the importance weights of attributes derived from the AHP method can be used to evaluate the fuzzy production rule with consideration of the degree of importance of attributes, which are changing over time, in the antecedent part of the rule.

To investigate the behavior of the importance weight and the DM of the above rule further, three cases of comparison matrix and 15 cases of input vectors are examined. The linguistic terms and values in the Gracia et. al [9] are employed for the investigation. The results are as follows.

The <Table 1> shows three comparison matrices as Case A, Case B and Case C. The purpose of introducing these three matrices is do investigate the behavior of the DM in different degree of importance of the attributes. In all three cases, $X_1$ is the most important, $X_2$ is the next and $X_3$ is the least important attribute. But the degree of relative importance is decreasing from Case A to Case C. For each case, the largest eigenvalue and corresponding eigenvectors are calculated and shown on the table. Then the eigenvectors are converted to the importance weights by the method explained above. These weights are applied to calculate modified DM in the <Table 2>, <Table 3> and <Table 4>. The 15 sets of input data are examined to investigate the behavior of the modified overall DM for the different combinations of importance weight and the DM of each attributes. There are six sets of input data for overall DM=0.5, six sets for the overall DM=0.2 and three sets for overall DM=0.8 before applying importance weight.

The <Table 2> shows behavior of the modified overall DM for the five sets of data with prior overall DM equals 0.5. Prior to applying importance weights, the overall DMs for the five data sets are the same(0.5). But the modified DMs are changed depending on the individual DM of the most important attribute($X_1$), the next important attribute($X_2$) and the least important attribute($X_3$). In the case 1 and 2, the DM of $X_1$(the most important attribute) is the largest(1.00). Thus modified overall DM is increased to 0.69(case 1) and to 0.89(case 2) from 0.5 for both cases. The difference between case 1 and 2 is the DM of the second most important attribute(0.5 for case 1 and 0.8 for case 2) and the third(0.8 and 0.5). Due to this difference, increment of the modified DM for the case 2(0.89-0.5=0.39) is greater than case 1(0.69-0.5=0.19). The modified overall DMs for the case 1 and 2 are decreased as the weight of $X_1$ is getting smaller(2.254⇒1.945⇒1.636).

<Table 1> Comparison Matrices

| Case | | Attributes/Matrix | | | EigenVal/Vec | Weight |
|---|---|---|---|---|---|---|
| A | | $X_1$ | $X_2$ | $X_3$ | 3.029(Val) | |
| | $X_1$ | 1 | 5 | 9 | 0.969 | 2.254 |
| | $X_2$ | 1/5 | 1 | 3 | 0.230 | 0.534 |
| | $X_3$ | 1/9 | 1/3 | 1 | 0.091 | 0.211 |
| B | | $X_1$ | $X_2$ | $X_3$ | 3.004(Val) | |
| | $X_1$ | 1 | 3 | 5 | 0.928 | 1.945 |
| | $X_2$ | 1/3 | 1 | 2 | 0.329 | 0.689 |
| | $X_3$ | 1/5 | 1/2 | 1 | 0.175 | 0.366 |
| C | | $X_1$ | $X_2$ | $X_3$ | 3.000(Val) | |
| | $X_1$ | 1 | 2 | 3 | 0.857 | 1.636 |
| | $X_2$ | 1/2 | 1 | 1.5 | 0.429 | 0.818 |
| | $X_3$ | 1/3 | 1/1.5 | 1 | 0.286 | 0.545 |

<Table 2> Behavior of Modified DM :(the cases of overall DM is 0.5 with equal weight)

| Case of Input | | Condition of the rule | Input Value | DM | Matrix A | | | Matrix B | | | Matrix C | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Weight | Mod. DM | Over. Mod. DM | Weight | Mod. DM | Over. Mod. DM | Weight | Mod. DM | Over. Mod. DM |
| 1 | $X_1$ | High | High | 1.00 | 2.254 | 1.00 | 0.69 | 1.945 | 1.00 | 0.62 | 1.636 | 1.00 | 0.57 |
| | $X_2$ | High | Medium | 0.50[a] | 0.534 | 0.69 | | 0.689 | 0.62 | | 0.818 | 0.57 | |
| | $X_3$ | Low | Ra. Low | 0.80 | 0.211 | 0.95 | | 0.366 | 0.92 | | 0.545 | 0.89 | |
| 2 | $X_1$ | High | High | 1.00 | 2.254 | 1.00 | 0.89 | 1.945 | 1.00 | 0.78 | 1.636 | 1.00 | 0.69 |
| | $X_2$ | High | Ra. High | 0.80 | 0.534 | 0.89 | | 0.689 | 0.86 | | 0.818 | 0.83 | |
| | $X_3$ | Low | Medium | 0.50[a] | 0.211 | 0.86 | | 0.366 | 0.78 | | 0.545 | 0.69 | |
| 3 | $X_1$ | High | Medium | 0.50[a] | 2.254 | 0.21 | 0.21 | 1.945 | 0.26 | 0.26 | 1.636 | 0.32 | 0.32 |
| | $X_2$ | High | High | 1.00 | 0.534 | 1.00 | | 0.689 | 1.00 | | 0.818 | 1.00 | |
| | $X_3$ | Low | Ra. Low | 0.80 | 0.211 | 0.95 | | 0.366 | 0.92 | | 0.545 | 0.89 | |
| 4 | $X_1$ | High | Ra. High | 0.80 | 2.254 | 0.60 | 0.60 | 1.945 | 0.65 | 0.65 | 1.636 | 0.69 | 0.69 |
| | $X_2$ | High | High | 1.00 | 0.534 | 1.00 | | 0.689 | 1.00 | | 0.818 | 1.00 | |
| | $X_3$ | Low | Medium | 0.50[a] | 0.211 | 0.86 | | 0.366 | 0.78 | | 0.545 | 0.69 | |
| 5 | $X_1$ | High | Medium | 0.50[a] | 2.254 | 0.21 | 0.21 | 1.945 | 0.26 | 0.26 | 1.636 | 0.32 | 0.32 |
| | $X_2$ | High | Ra. High | 0.80 | 0.534 | 0.89 | | 0.689 | 0.86 | | 0.818 | 0.83 | |
| | $X_3$ | Low | Low | 1.00 | 0.211 | 1.00 | | 0.366 | 1.00 | | 0.545 | 1.00 | |
| 6 | $X_1$ | High | Ra. High | 0.80 | 2.254 | 0.60 | 0.60 | 1.945 | 0.65 | 0.62 | 1.636 | 0.69 | 0.57 |
| | $X_2$ | High | Medium | 0.50[a] | 0.534 | 0.69 | | 0.689 | 0.62 | | 0.818 | 0.57 | |
| | $X_3$ | Low | Low | 1.00 | 0.211 | 1.00 | | 0.366 | 1.00 | | 0.545 | 1.00 | |

* 0.50[a] : Overall DM for the rule with equal weight for all attributes($X_1$, $X_2$, $X_3$)

*Mod. DM: Modified Dm with given weight($(DM)^{weight}$)

*Over. Mod. DM: Overall modified DM for the rule(DM for the consequent part of the rule)

* Ra. High = Rather High=(0.0/10, 0.03/20, 0.11/30, 0.25/40, 1.0/50, 0.0/60)

* Ra. Low = Rather Low=(0.0/10, 1.0/20, 0.25/30, 0.11/40, 0.03/50, 0.0/60)

<Table 3>Behavior of Modified DM :(the cases of overall DM is 0.2 with equal weight)

| Case of Input | | Condition of the rule | Input value | DM | Matrix A | | | Matrix B | | | Matrix C | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Weight | Mod. DM | Over. Mod. DM | Weight | Mod. DM | Over. Mod. DM | Weight | Mod. DM | Over. Mod. DM |
| 7 | $X_1$ | High | Ra. High | 0.80 | 2.254 | 0.60 | 0.60 | 1.945 | 0.65 | 0.55 | 1.636 | 0.69 | 0.42 |
| | $X_2$ | High | Medium | 0.50 | 0.534 | 0.69 | | 0.689 | 0.62 | | 0.818 | 0.57 | |
| | $X_3$ | Low | Ra. High | 0.20 | 0.211 | 0.71 | | 0.366 | 0.55 | | 0.545 | 0.42 | |
| 8 | $X_1$ | High | Ra. High | 0.80 | 2.254 | 0.60 | 0.42 | 1.945 | 0.65 | 0.33 | 1.636 | 0.69 | 0.27 |
| | $X_2$ | High | Ra. Low | 0.20 | 0.534 | 0.42 | | 0.689 | 0.33 | | 0.818 | 0.27 | |
| | $X_3$ | Low | Medium | 0.50 | 0.211 | 0.86 | | 0.366 | 0.78 | | 0.545 | 0.69 | |
| 9 | $X_1$ | High | Medium | 0.50 | 2.254 | 0.21 | 0.21 | 1.945 | 0.26 | 0.26 | 1.636 | 0.32 | 0.32 |
| | $X_2$ | High | Ra. High | 0.80 | 0.534 | 0.89 | | 0.689 | 0.86 | | 0.818 | 0.83 | |
| | $X_3$ | Low | Ra. High | 0.20 | 0.211 | 0.71 | | 0.366 | 0.55 | | 0.545 | 0.42 | |
| 10 | $X_1$ | High | Ra. Low | 0.20 | 2.254 | 0.03 | 0.03 | 1.945 | 0.04 | 0.04 | 1.636 | 0.07 | 0.07 |
| | $X_2$ | High | Ra. High | 0.80 | 0.534 | 0.89 | | 0.689 | 0.86 | | 0.818 | 0.83 | |
| | $X_3$ | Low | Medium | 0.50 | 0.211 | 0.89 | | 0.366 | 0.78 | | 0.545 | 0.69 | |
| 11 | $X_1$ | High | Medium | 0.50 | 2.254 | 0.21 | 0.21 | 1.945 | 0.26 | 0.26 | 1.636 | 0.32 | 0.27 |
| | $X_2$ | High | Ra. Low | 0.20 | 0.534 | 0.42 | | 0.689 | 0.33 | | 0.818 | 0.27 | |
| | $X_3$ | Low | Ra. Low | 0.80 | 0.211 | 0.95 | | 0.366 | 0.92 | | 0.545 | 0.89 | |
| 12 | $X_1$ | High | Ra. Low | 0.20 | 2.254 | 0.03 | 0.03 | 1.945 | 0.04 | 0.04 | 1.636 | 0.07 | 0.07 |
| | $X_2$ | High | Medium | 0.50 | 0.534 | 0.69 | | 0.689 | 0.62 | | 0.818 | 0.57 | |
| | $X_3$ | Low | Ra. Low | 0.80 | 0.211 | 0.95 | | 0.366 | 0.92 | | 0.545 | 0.89 | |

*the number in the shadowed cell(0.20): Overall DM for the rule with equal weight for all attributes($X_1$, $X_2$, $X_3$)

In the case 3 and 4, the DM for the second important attribute($X_2$) is the largest(1.00). And the DM for the $X_1$(the most important attribute) is the smallest(0.5) for the case 3 and the medium value(0.8) for the case 4. Accordingly the modified overall DM for the case 3 is decreased to 0.21(0.5⇒0.21) and that for the case 4 is increased to 0.60(0.5⇒0.60).

In the case 5 and 6, the DM for the least important attribute($X_3$) is the largest(1.00). And the DM for the $X_1$(the most important attribute) is the smallest(0.5) for the case 5 and the medium value(0.8) for the case 6. Thus the behavior of the modified overall DM for the case 5 and 6 is similar to the case 3 and 4.

The numbers and pattern shown on the <Table 3> and <Table 4> can be interpreted as the same as the <Table 2>.

<Table 4>Behavior of Modified DM :(the cases of overall DM is 0.8 with equal weight)

| Case of Input | | Condition of the rule | Input value | DM | Matrix A | | | Matrix B | | | Matrix C | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Weight | Mod. DM | Over. Mod. DM | Weight | Mod. DM | Over. Mod. DM | Weight | Mod. DM | Over. Mod. DM |
| 13 | $X_1$ | High | High | 1.00 | 2.254 | 1.00 | 0.89 | 1.945 | 1.00 | 0.86 | 1.636 | 1.00 | 0.83 |
| | $X_2$ | High | Ra. High | 0.80 | 0.534 | 0.89 | | 0.689 | 0.86 | | 0.818 | 0.83 | |
| | $X_3$ | Low | Ra. Low | 0.80 | 0.211 | 0.95 | | 0.366 | 0.92 | | 0.545 | 0.89 | |
| 14 | $X_1$ | Hgih | Ra. High | 0.80 | 2.254 | 0.60 | 0.60 | 1.945 | 0.65 | 0.65 | 1.636 | 0.69 | 0.69 |
| | $X_2$ | High | High | 1.00 | 0.534 | 1.00 | | 0.689 | 1.00 | | 0.818 | 1.00 | |
| | $X_3$ | Low | Ra. Low | 0.80 | 0.211 | 0.95 | | 0.366 | 0.92 | | 0.545 | 0.89 | |
| 15 | $X_1$ | Hgih | Ra. High | 0.80 | 2.254 | 0.60 | 0.60 | 1.945 | 0.65 | 0.65 | 1.636 | 0.69 | 0.69 |
| | $X_2$ | High | Ra. High | 0.80 | 0.534 | 0.89 | | 0.689 | 0.86 | | 0.818 | 0.83 | |
| | $X_3$ | Low | Low | 1.00 | 0.211 | 1.00 | | 0.366 | 1.00 | | 0.545 | 1.00 | |

*the number in the shadowed cell(0.80): Overall DM for the rule with equal weight for all attributes($X_1$, $X_2$, $X_3$)

Through the three cases of comparison matrices for calculating importance weights and 15 cases of input data sets for calculating DM examined above, the importance weights derived by applying AHP method has affected the overall DM, which indicates the degree of accordance for the assertion in the consequent part of a rule. This was done by utilizing the importance weights as exponents to the individual DM of each attribute in the antecedent part of a rule. Therefore this procedure can be used to reflect the importance of attributes by which we can make a dynamic inference in a knowledge-based system.

## 5. Summary

The research efforts on expert systems have been conducted for the purpose of substituting a human decision maker by programming the logical process of making decisions of a decision maker. The fuzzy production rule has proposed to codify and process human's uncertain knowledge as part of the research efforts. An advantage of the knowledge-based system with fuzzy production rules is that the system can describe more realistically a human's logical process of inferencing, in which he interprets collected information and then arrives at a conclusion. However, there is a limit on programming the process with current state of art computer programming, and thus the research efforts have been and will be continuing. The extent of programming is further limited, especially in a complex situation involving many uncertain elements and where the experts' knowledge is imprecise itself, and the dynamic situation in which the meaning of the same information could differ depending on the environment, such as making stock investment decisions.

This paper is also an effort to represent experts' knowledge and to depict inferencing on a knowledge-based system with fuzzy production rules more realistically. The approach proposed in this paper evaluates knowledge-based systems, which consist of fuzzy production rules, while taking into account the importance of attributes which can be changed over time, and thus more realistic advice can be derived from the system. The degree of match between the condition part of a rule and the input data of attributes has a value of range [0, 1], and the importance weights of attributes are derived by the AHP method. Then the weights are used as exponents to the DM. As a results, the dynamic behavior of real world decision-making situations can be reflected in the static knowledge base, and the knowledge base can be simplified and become more manageable.

## References


[1] A. Newell, H. Simon, *Human Problem Solving*, Prentice Hall, Englewood Cliffs, NJ, 1972.

[2] L. A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning I, II, III, *Information Sciences*, 1975, 8:199-249, 302-357, 9:43-80

[3]. S. J. Henkind, M. C. Harrison, An analysis of four uncertainty calculi, *IEEE Transaction on Systems. Man. and Cybernetics*. Vol. 18. No.5. Sep/Oct, 1988, 700-714

[4] F. Wenstop, Deductive verval models of organizations, *Int'l J. of Man-Machine Studies*, 8, 1976.

[5] F. Wenstop, Quantitative analysis with linguistic values, *Fuzzy Sets and Systems*, 4, 1980.

[6] C. Freksa, Linguistic description of human judgements in expert systems and soft sciences, in: M. Gupta and E. Sanchez Eds. *Approximate reasoning in decision analysis*, North-Holland, Amsterdam, 1982.

[7] L. A. Zadeh, Test score semantics for natural languages and meaning representation via PRUF, in: B. Rieger Eds., *Empirical Semantics*, Brockmeyer, Bochum, West Germany, 1981.

[8] J. F. Baldwin, S.Q. Zhou, A Fuzzy relational inference language, *Fuzzy Sets and Systems*, 14, 1982.

[9] J. Garcia, J. Pazos, J. Rios, J. Yague, Methodology of linguistics evaluation in risk situations using fuzzy techniques, *Fuzzy Sets and Systems*, 48, 1992, 185-194.

[10] M. Mizumoto, Fuzzy inference using max composition in the compositional rule of inference, in: M.Gupta and E. Sanchez Eds., *Approximate reasoning in decision analysis*, North-Holland, Amsterdam, 1982.

[11] H. Prade, Approximate and plausible reasoning: the state of art, Communication at he IFAC Symposium on Fuzzy Information, Knowledge Representation, and Decision Analysis, July 1983, Marseille, France

[12] L. Lesmo, L. Saitta, P. Torasso, Evidence combination in expert systems, *Int'l J. of Man-Machine Studies*, 22, 1985, 307-326.

[13] S.Cho, O. Ersoy, M. Lehto, An Algorithm to compute the degree of match in fuzzy systems, *Fuzzy Sets and Systems*, 49, 1992, 285-299.

[14] M. Cayrol, H. Farreny, H. Prade, Fuzzy pattern matching, *Kybernetics* 11, 1982, 103-116.

[15] K. Leung, W. Lam, Fuzzy concepts in expert systems, IEEE Comp., Sept. 1988, 43-56

[16] T. Saaty, *The Analytic Hierachy Process*, McGraw Hill, New York, 1980.

[17] T. Saaty, Rank generation preservation, and reversal in the analytic hierachy decision process, *Decision Sciences*, 18, no.2, 1987, 157-177.

[18] A. Ibrahim, B. Ayyub, Multi-criteria ranking of components according to their priority for inspection, *Fuzzy Sets and Systems*, 48, 1992, 1-14.