Fuzzy Logic - A Mordern Perspective

John Yen, Senior Member, IEEE

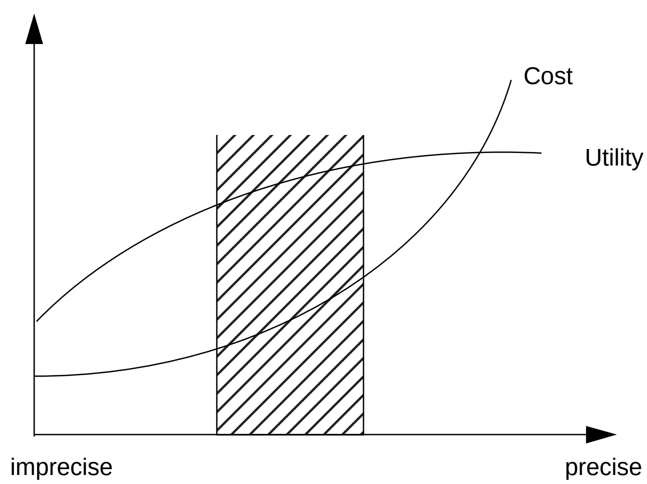
ABSTRACT

Traditionaly, fuzzy logic has been viewed in the AI community as an approach for managing uncertainty. In the 1990's, however, fuzzy logic has emerged as a paradigm for approximating a functional mapping. This complementary mordern view about the technology offers new insights about the foundation of fuzzy logic as well as new challenges regarding the identification of fuzzy models. In this paper, we will first review some of the major milestones in the history of developing fuzzy logic technology. After a short summary of major concepts in fuzzy logic, we discuss a mordern view about the foundation of two types of fuzzy rules. Finally, we review some of the research in addressing various challenges regarding automated identification of fuzzy rule-based models.

1 INTRODUCTION

Fuzzy logic (FL) has been a somewhat controversial technology since its birth. However, the large number of successful industrial fuzzy logic applications in 1990's, especially those developed in Japan, has generated an increasing interest in FL. Fuzzy logic and artificial intelligence (AI) have at least one common objective: to develop computational methods that can perform reasoning and problem solving tasks that require human intelligence. However, fuzzy logic has an additional objective: to explore an effective tradeoff between precision and the cost in developing an approximate model of a complex system or function. While the issue of the cost was not considered an important issue in early AI work, it has become important in the past decade due to an increasing interest in resource-constrained intelligent agents.

FIGURE 1. The Cost-Precision Trade-off



Conceptually, we can use Figure 1 to depict this trade-off for many systems. The horizontal axis represents the degree of precision of the system, while the vertical axis serves the dual-purpose of representing both the cost and the degree of utility. As the precision of a system increases, the cost for developing the system also increases, typically in an exponential manner. On the other hand, the utility (i.e., usefulness) of the system does not increase proportionally as its precision increases -- it usually saturates after a certain point. This insight about the trade-off between precision, cost, and utility inspired Zadeh and his followers to exploit the shaded area in Figure 1, which resulted in a new paradigm for developing approximate solutions that are both cost-effective and highly useful.

Traditionally, fuzzy logic has been viewed as a theory for dealing with uncertainty about complex systems. A modern complementary perspective is, however, to view fuzzy logic as an approximation theory. This perspective on fuzzy logic brings to the surface the underpinning of the theory described above - the cost-precision trade-off. Indeed, providing a cost-effective solution to a wide range of real world problems is the primary reason that fuzzy logic has found so many successful applications in industry to date. Understanding this driving force of the success of fuzzy logic will prevent us from falling into the trap of debating "whether fuzzy logic can accomplish what X can not accomplish" where X is an alternative technology such as probability theory, control theory, etc. Such a debate is usually not fruitful because it ignores one important issue -- cost. A better question to ask is "What is the difference between the cost of a fuzzy logic approach and the cost of an approach based on X to accomplish a certain task?"

2 HISTORICAL DEVELOPMENT OF THE TECHNOLOGY

In early 1960s, Lotfi A. Zadeh, a professor at University of California at Berkley well respected for his contributions to the development of system theories, began to feel that traditional systems analysis techniques were too precise for many complex real-world problems. The idea of grade of membership, which is the concept that became the backbone of fuzzy set theory, occurred to him in 1964 [21], which lead to the publication of his seminal paper on fuzzy sets in 1965 and the birth of fuzzy logic technology [35]. The concept of fuzzy sets and fuzzy logic encountered sharp criticism from the academic community; however, scholars and scientists around the world -- ranging from psychology, sociology, philosophy and economics to natural sciences and engineering -- became Zadeh's followers. B.R. Gaines and L.J. Kohout gave a detailed bibliography of the first decade of fuzzy logic research in [10].

Fuzzy Logic research in Japan started with two small university research groups established in late 1970s: one was lead by T. Terano and H. Shibata in Tokyo, and the other lead by K. Tanaka and K. Asai in Kanasai. Like fuzzy logic researchers in the U.S., these researchers encountered an "anti-fuzzy" atmosphere in Japan during those early days. However, their persistence and hard work would prove to be worthwhile a decade later. These Japanese researchers, their students, and the students of their students would make many important contributions to the theory as well as to the applications of fuzzy logic [11].

In 1974, S. Assilian and E. H. Mamdani in United Kingdom developed the first fuzzy logic controller, which was for controlling a steam generator [19]. In 1976, Blue Circle Cement and SIRA in Denmark developed a cement kiln controller -- which is the first industrial application of fuzzy logic. The system went to operation in 1982.

In the 1980's, several important industrial applications of fuzzy logic was launched successfully in Japan. After eight years of persistent research, development, and deployment efforts, Yasunobu and his colleagues at Hitachi put a fuzzy logic-based automatic train operation control system into operation in Sendai city's subway system in 1987 [29]. Another early successful industrial application of fuzzy logic is a water-treatment system developed by Fuji Electric. These and other applications motivated many Japanese engineers to investigate a wide range of novel fuzzy logic applications. This lead to the fuzzy boom.

The fuzzy boom in Japan was a result of close collaboration and technology transfer between universities and industries. Two large-scale national research projects were established by two Japanese government agencies in 1987: the better known of the two is the Laboratory for International Fuzzy Engineering Research (LIFE). In late January 1990, Matsushita Electric Industrial Co. named their newly developed fuzzy controlled automatic washing machine "Asai-go (beloved wife) Day Fuzzy" and launched a major commercial campaign for the "fuzzy" product. This campaign turns out to be a successfully marketing effort not only for the product, but also for the fuzzy logic technology. A foreign word pronounced "fuzzy" was thus introduced to Japan with a new meaning -- intelligence. Many other home electronics companies followed Panasonic's approach and introduced fuzzy vacuum cleaners, fuzzy rice cookers, fuzzy refrigerators, fuzzy camcorders (for stablizing the image under hand jittering), camera (for smart auto-focus) and others. This resulted in a fuzzy vogue in Japan. As a result, the consumers in Japan all recognized the Japanese word "fuzzy", which won the gold prize for the new word in 1990 [11]. This fuzzy boom in Japan triggered a broad and serious interest in this technology in Korea, Europe, and, to a lesser extent, in the United States, where fuzzy logic was invented.

Fuzzy logic has also found its applications in the financial area. The first financial trading system using fuzzy logic was Yamaichi Fuzzy Fund. It handles 65 industries and a majority of the stocks listed on Nikkei Dow and consists of approximately 800 fuzzy rules. Rules are determined monthly by a group of experts and modified by senior business analysts as necessary. The system was tested for two years, and its performance in terms of the return and growth exceeds the Nikkei Average by over 20%. While in testing, the system recommended "sell" 18 days before the Black Monday in 1987. The system went to commercial operations in 1988.

The first special-purpose VLSI chip for performing fuzzy logic inferences was developed by M. Togai and H. Watanabe in 1986 [27]. These special-purpose VLSI chips can enhance the performance of fuzzy rule-based systems for real-time applications. Several companies were formed (e.g., Togai Infralogic, APTRONIX, INFORM) were formed to commercialize hardware and software tools for developing fuzzy systems. Vendors of conventional control design software also started introducing add-on toolbox for designing fuzzy systems. *The Fuzzy Logic Toolbox* for MATLAB, for instance, was introduced as an add-on component to MATLAB in 1994.

2.1 Learning of Fuzzy Knowledge

The development of fuzzy systems in early days required the manual tuning of the system parameters based on observing the system performance. This drawback has become one of the major criticisms toward fuzzy logic. Even though Mamdani and Baaklini introduced self-adaptive fuzzy logic control as early as 1975, the most common citation to the first work in this area is a paper by T. J. Procyk and E. H. Mamdani published in 1979 [22]. This was followed by Japanese researchers in the 1980's. T. Takagi and his advisor M. Sugeno together took an important step by developing the first approach for constructing (not tuning) fuzzy rules using training data [26]. Their approach learned fuzzy rules for controlling a toy vehicle by observing how a human operator controlled the vehicle. Even though this important work did not gain as much immediate attention as it did later, it laid the foundation for an important subarea in fuzzy logic, which is later referred to as *fuzzy model identification* in the 1990's.

Another trend that contributed to research in fuzzy model identification is the increasing visibility of neural network research in the late 1980's. Because of certain similarities between neural networks and fuzzy logic, researchers began to investigate ways to combine the two technologies. The most important outcome of this trend is the development of various techniques for identifying the parameters in a fuzzy system using neural network learning techniques. A system built this way is called a *neuro-fuzzy system* [15, 18].

The 1990's is an era of new computational paradigms. In addition to fuzzy logic and neural networks, a third non-conventional computational paradigm also became popular -- *evolutionary computing*, which includes *genetic algorithms*, *evolutionary strategies*, and *evolutionary programming*. Genetic algorithms(GA) and evolutionary strategies are optimization techniques that attempt to avoid being easily trapped in local minima by simultaneously exploring multiple points in the search space and by generating new points based on the Darwinian theory of evolution -- survival of the fittest. The popularity of GA in the 1990's inspired the use of GA for optimizing parameters in fuzzy systems [13]. Various synergistic combinations of neural networks, genetic algorithms, and fuzzy logic help people to view them as complementary. To distinguish these paradigms from the conventional methodologies based on precise formulations, Zadeh introduced the term *soft computing* in early 1990's [41].

3 FUZZY SETS, POSSIBILITY DISTRIBUTIONS, AND COMPOSITIONAL RULE OF INFERENCE

Fuzzy sets, linguistic variables, and possibility distributions are three core concepts in fuzzy logic. A fuzzy set is a generalization to classical set to allow objects to take partial membership in vague concepts (i.e., fuzzy sets) [35]. The degree an object belongs to a fuzzy set, which is a real number between 0 and 1, is called the membership value in the set. The meaning of a fuzzy set, is thus characterized by a *membership function* that maps elements of a universe of discourse to their corresponding membership values. The membership function of a fuzzy set A is denoted as μ_A. In addition to membership functions, a fuzzy set is also associated with a linguistically meaningful term (e.g., ‘‘healthy’’ family). Associating a fuzzy set to a linguistic term offers two important benefits. First, the association makes it easier for human experts to express their knowledge using the linguistic terms. Second, the knowledge expressed using linguistic terms is easily comprehensible. This benefit often results in significant savings in the cost of designing, modifying and maintaining a fuzzy logic system.

A *linguistic variable* is a variable whose value can be described (1) qualitatively using an expression involving linguistic terms and (2) quantitatively using a corresponding membership function [39]. The linguistic term is useful for communicating concepts and knowledge with human beings; whereas membership function is useful for processing numeric input data. A linguistic variable is like a composition of a symbolic variable in AI (a variable whose value is a symbol) and a numeric variable (a variable whose value is a number) in science and engineering. Using the notion of linguistic variable to combine these two kinds of variables into a uniform framework is, in fact, one of the main reasons that fuzzy logic has been successful in offering intelligent approaches in engineering and many other areas that deal with continuous problem domains. In general, the value of a linguistic variable can be a linguistic expression involving a set of linguistic terms, modifiers such as “very”, “more or less” (called **hedges**) and **connectives** (e.g., “and”, “or”). For example, the sentence “trading is moderate and not very heavy” can be expressed as assigning the linguistic expression ‘‘Moderate AND NOT VERY Heavy’’ to the linguistic variable TradingQuantity. The meaning of such an expression is governed by several semantic rules about constructing the corresponding membership functions.

When an interval is assigned to a variable with unknown value (e.g., the suspects age is between 20 and 30 years old), it constrains the possible values of the variable. Similarly, when a fuzzy set is assigned to a linguistic variable (e.g., the suspects age is young), it imposes an elastic constrain on the possible values of the variable called the possibility distribution [40]. The main difference between the two is that the notion of possible vs. impossible values becomes a matter of degree.

One of the questions commonly raised about possibility distribution is its relationship with probability distribution. While the subject is too complex for a comprehensive discussion here, it is important enough to deserve some clarification. The best way to understand the relationship between possibility distribution and probability distribution is to compare interval-values with probability. As we have mentioned earlier, an interval-valued assignment constrains the possible value of a variable without indicating the likelihood that the variable takes a specific value in the interval. Similarly, a possibility distribution states the degree of ease (i.e., possibility) for the variable to take a certain value without indicating the likelihood that the variable has such a value. Even though possibility distribution and probability distribution are different, they are also related -- if a value is impossible, it is obviously improbable. In general, a possibility distribution can be viewed as an upper bound on the probability distribution. More importantly, however, fuzzy logic can be viewed as complementary to the probability theory. For instance, the notion of events in probability theory can be generalized to fuzzy events based on fuzzy sets. Such a generalization allows probability theory to better deal with events that do not have a well-defined sharp boundary (e.g., the probability that the stock price of Intel will rise significantly; the probability that the number of heads is much more than the number of tails in 10 coin tossings). A further discussion on the issue of fuzzy logic vs. probability theory can be found in [3].

Fuzzy set theory generalizes the conventional set theory; therefore, its axiomatic foundation is unavoidably different from that of classical set theory. More specifically, it has to violate two fundamental laws of Boolean algebra -- the law of excluded middle $A \cup \bar{A} = U$ and the law of contradiction $A \cap \bar{A} = \phi$ In other words, it is possible for an element to partially belong to both a fuzzy set and the set's complement. For instance, suppose that John is somewhat bald but not completely bald. In fuzzy set theory, John can partially belong to the set of bald people as well as the set of people who are not bald. Due to the fact that the law of excluded middle and the law of contradiction are not axioms of fuzzy set theory, formula equivalents in classical set theory are not necessarily equivalent in fuzzy set theory. Similarly, logically equivalent formula are not necessarily equivalent in fuzzy logic. A potential danger of ignoring such a difference is to reject fuzzy sets (and fuzzy logic in general) based on an inappropriate set of axioms [8].

The set operations *intersection* and *union* correspond to logic operations, *conjunction (and)* and *disjunction (or)*, respectively. There are multiple choices for the fuzzy conjunction and the fuzzy disjunction operators. A common choice is to use min for fuzzy conjunction, max for fuzzy disjunction; another common pair is the algebraic product (for fuzzy conjunction) and algebraic sum (for fuzzy disjunction):

$$\mu_{(A \cap B)}(x) = \mu_A(x) \times \mu_B(x) \qquad \text{(EQ 1)}$$

$$\mu_{A \cup B}(x) = \mu_A(x) + \mu_B(x) - \mu_A(x) \times \mu_B(x) \qquad \text{(EQ 2)}$$

There are an infinite number of other choices.

The choice of a fuzzy conjunction operator could determine the choice of the fuzzy disjunction operator, and vice versa. This is due to the principle of duality between the two operators. More specifically, a fuzzy conjunction operator, denoted as $t(x, y)$ and a fuzzy disjunction operator, denoted as $s(x, y)$, form a dual pair if they satisfy the following condition:

$$1 - t(x, y) = s(1 - x, 1 - y) \qquad \text{(EQ 3)}$$

In fact, this duality condition ensures that

$$\overline{A \cap B} = \bar{A} \cup \bar{B} \qquad \text{(EQ 4)}$$

still holds in fuzzy set theory.

The set of candidate fuzzy conjunction operators, called *triangular norms* or *t-norms,* is defined by a set of axioms. Similarly, the set of candidate fuzzy disjunction operators called *triangular conorms, t-conorms,* or *s-norms* is defined by a set of dual axioms. We formally define t-norms and t-conorms using their axioms below.

DEFINITION 1 A t-norm operator, denoted as $t(x, y)$ is a function mapping from [0, 1] × [0, 1] to [0, 1] that satisfies the following conditions for any $w, x, y, z \in [0, 1]$:
1. $(0, 0) = 0,\ t(x, 1) = t(1, x) = x$
2. $t(x,y) \le t(z,w)$ if $x \le z$ *and* $y \le w$ (monotonicity)
3. $t(x, y) = t(y, x)$ (commutativity)
4. $t(x, t(y, z)) = t(t(x, y), z)$ (associativity)

DEFINITION 2 A t-conorm operator, denoted as *s(x,y)*, is a function mapping from [0, 1] × [0, 1] to [0, 1] the following conditions for any $w, x, y, z \in [0, 1]$.
1. $(1, 1) = 1,\ s(x, 0) = s(0, x) = x$
2. $s(x,y) \le s(z,w)$ if $x \le z$ *and* $y \le w$ (monotonicity)
3. $s(x, y) = s(y, x)$ (commutativity)
4. $s(x, s(y, z)) = s(s((x, y), z))$ (associativity)

A summary of t-norm and t-conorm operators can be found in [14]. An important property about t-norms is that all t-norms are bounded above by min and bounded below by drastic product. Similarly, all t-conorms are bounded above by drastic sum and bounded below by max.

A fuzzy relation generalizes the classical notion of relation into a matter of degree. For instance, the fuzzy relation *Friend* could describe the degree of friendship between two persons. Similarly, a fuzzy relation *Petite* between height and weight of a person describes the degree by which a person with a specific height and weight is considered petite. Formally, fuzzy relation *R* between variables *x* and *y*, whose domains *are X* and *Y*, respectively, is defined by a function that maps ordered pairs in $X \times Y$ to their degree in the relation, which is a number between 0 and 1, i.e., $R: X \times Y \to [0, 1]$. More generally, a fuzzy n-ary relation *R* in $x_1, x_2, \ldots, x_n$, whose domains are $X_1, X_2, \ldots, X_n$, respectively, is defined by a function that maps an n-tuple $<x_1, x_2, \ldots, x_n>$ in $X_1 \times X_2 \times \ldots \times X_n$ to a number in the interval, i.e., $R: X_1 \times X_2 \times \ldots \times X_n \to [0, 1]$. Just as a classical relation can be viewed as a set, a fuzzy relation can be viewed as a fuzzy subset. From this perspective, the mapping above is equivalent to the membership function of a multidimensional fuzzy set.

If the possible values of *x* and *y* are discrete, we can express a fuzzy relation in a matrix form. For example, suppose we wish to express a fuzzy relation *Petite* in terms of the height and the weight of a female. Suppose the range of the height and the weight of interest to us are {5', 5'1", 5'2", 5',3", 5'4", 5'5", 5'6"}, denoted *h*, and {90, 95, 100, 105, 110, 115, 120, 125} (in lb,) denoted *w*, respectively. We can express the fuzzy relation in a matrix form as shown below:

$$
\begin{array}{c|cccccccc}
 & 90 & 95 & 100 & 105 & 110 & 115 & 120 & 125 \\
\hline
5' & 1 & 1 & 1 & 1 & 1 & 1 & 0.5 & 0.2 \\
5'1'' & 1 & 1 & 1 & 1 & 1 & 0.9 & 0.3 & 0.1 \\
5'2'' & 1 & 1 & 1 & 1 & 1 & 0.7 & 0.1 & 0 \\
5'3'' & 1 & 1 & 1 & 1 & 0.5 & 0.3 & 0 & 0 \\
5'4'' & 0.8 & 0.6 & 0.4 & 0.2 & 0 & 0 & 0 & 0 \\
5'5'' & 0.6 & 0.4 & 0.2 & 0 & 0 & 0 & 0 & 0 \\
5'6'' & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}
$$

Each entry in the matrix indicates the degree a female with the corresponding height (i.e., the row heading) and weight (i.e., the column heading) is considered to be petite. Once we define the *Petite* fuzzy relation, we can answer two kinds of questions:

- What is the degree that a female with a specific height and a specific weight is considered to be petite?
- What is the possibility that a petite person has a specific pair of height and weight measures?

In answering the first question, the fuzzy relation is equivalent to the membership function of a multidimensional fuzzy set. In the second case, the fuzzy relation becomes a possibility distribution assigned to a petite person whose actual height and weight are unknown. The second usage of fuzzy relation enables us to reason about the possible height of a petite person given her weight. For instance, we may wish to know the possible weight of a petite female called Michelle who is *about 5'4"* tall where *"about"* indicates impression. The answer to this question can be obtained through an important inference technique in fuzzy logic -- the *compositional rule of inference.* We will use the example above to introduce the foundation of this inference technique.

How do we find out whether it is possible for Michelle to have a specific weight, say 110 lb? We need to consider all possible heights of the person and see if a petite person with such a height can weigh 110 lb. In general, a petitie person is possible to weigh *w* if and only if (1) the person is possible to have height *h* and (2) it is possible for a person with height *h* and weight *w* to be petite. This can be expressed in logic as follows:

$$\forall x \forall w_j \left[Possible\text{-}weight(x, w_j) \leftrightarrow \bigvee_{h_i} (Possible\text{-}height(x, h_i) \wedge Petite(h_i, w_j)) \right] \qquad \text{(EQ 5)}$$

where *Possible-weight* (x, w_j) is a predicate to test whether it is possible for person x to have weight w_j, *Possible-height* (x, h_i), and *Petite*(h_i, w_j) are similar predicates in first order logic, $\wedge$ denotes the conjunction operator, $\vee$ denotes the disjunction operator, and $\leftrightarrow$ denotes bidirectional implication. As we mentioned earlier, both possibilities and relations become matters of degree in fuzzy logic. Hence, we can replace the predicates in Equation 5 with possibility distributions $\Pi_{\text{height}(x)}(h_i)$ and $\Pi_{\text{petite}}(h_i, w_j)$ a to obtain the following formula to infer the possibility of a petite person's weight, denoted $\Pi_{\text{weight}(x)}(w_j)$:

$$\Pi_{\text{weight(x)}}(w_j) = \bigoplus_{h_i}(\Pi_{Height(x)}(h_i) \otimes \Pi_{Petite}(h_i, w_j)) \qquad \text{(EQ 6)}$$

where $\otimes$ and $\oplus$ denote fuzzy conjunction and fuzzy disjunction respectively. We define the compositional rule of inference more formally below.

DEFINITION 3 Let *X* and *Y* be the universes of discourse for variables *x* and *y*, respectively, and x_i and y_j be elements of *X* and *Y*. Let *R* be a fuzzy relation that maps $X \times Y$ to [0, 1] and the possibility distribution of *X* is known to be $\Pi_x(x_i)$. The compositional rule of inference infers the possibility distribution of *Y* as follows:

$$\Pi_Y(y_j) = \bigoplus_{x_i}(\Pi_X(x_i) \otimes \Pi_R(x_i, y_j)) \qquad \text{(EQ 7)}$$

The compositional rule of inference is not uniquely defined. By choosing different fuzzy conjunction and fuzzy disjunction operators, we get different compositional rules of inference. We list two that are commonly used in practice:

1. max-min composition: $\Pi_Y(y_j) = \max_{x_i}(min(\Pi_X(x_i), \Pi_R(x_i, y_j)))$ (EQ 8)

2. max-product composition: $\Pi_Y(y_j) = \max_{x_i}(\Pi_X(x_i) \times \Pi_R(x_i, y_j))$ (EQ 9)

4 FUZZY IF-THEN RULES

Among all the techniques developed using fuzzy sets, fuzzy if-then rules are by far the most visible due to their wide range of successful industrial applications ranging from consumer products, robotics, manufacturing, process control, automotive control, medical imaging, to financial trading. A fuzzy if-then rule associates a *condition* about linguistic variables to a *conclusion*. From a knowledge representation viewpoint, a fuzzy if-then rule is a scheme for capturing knowledge that involves imprecision. The main feature of reasoning using these rules (i.e., fuzzy rule-based reasoning) is its *partial matching* capability, which enables an inference to be made from a fuzzy rule even when the rule's condition is only partially satisfied. The degree the input data matches the condition of a rule is combined with the consequent (i.e., "then" part) of the rule to form a conclusion inferred by the fuzzy rule. The higher is the matching degree, the closer is the inferred conclusion to the rule's consequent.

There are two types of fuzzy rules: 1) *fuzzy mapping rules*, and 2) *fuzzy implication rules*. A fuzzy mapping rule describes a functional mapping relationship between inputs and an output using linguistic terms, while a fuzzy implication rule describes a generalized logic implication relationship between two logic formula involving linguistic variable and imprecise linguistic terms. The foundation of fuzzy mapping rule is fuzzy graph, while the foundation of fuzzy implication rule is a generalization to two-valued logic. The inference of fuzzy mapping rules involves a set of rules whose antecedent conditions form a fuzzy partition of the input space. We call such a collection of fuzzy mapping rules a *fuzzy model*. The inference of fuzzy implication rules are performed individually. Even though the inference results of these rules can be combined, the desired properties of their inference are described in terms of the behavior of individual rules (e.g., generalized modus ponens and modus tollens involving hedges). Consequently, fuzzy mapping rules are designed as a group, whereas fuzzy implication rules are designed individually.

The distinction between fuzzy implication rules and fuzzy mapping rules has not been clear in the literature. Until early 1990's, fuzzy rules used in control systems have been viewed as a special kind of fuzzy implication rule [17]. However, it is difficult to explain the use of conjunction operator in forming the "fuzzy implication relation" of rules and the use of fuzzy disjunction in aggregating the conclusion of rules. This difficulty gradually lead to the crystallization of the fundamental differences between the two types of rules. Zadeh, Kosko, Dubois and Prade and many others contribute to this process through keynote speeches, books [16], conference and journal publications [6].

4.1 Fuzzy Mapping Rules versus Fuzzy Implication Rules

A potential confusing point in distinguishing these two types of fuzzy rules is that they have two similarities. First, both of them can be represented as a fuzzy relation between antecdent variables and consequent variables. Second, both of their inference schemes are based on the compositional rule of inference. These similaries have convinced many scholars that these two types of rules are identical. However, underneath these two similarities are two profound differences. Even though they can be both represented as fuzzy relations, the semantcis of the content of their fuzzy relations differs. Consequently, their inference uses different operators in the compositional rule of inference. We will elaborate on these two points below.

The most fundamental difference between the semantics of fuzzy mapping rules and fuzzy implication rules is in their inference behavior. Even though these two types of rules behave the same when their antecedents are satisfied, they behave differently when their antecedents are not satisfied. We will illustrate this using an example. Suppose x and y are two integer variables taking values from the interval [0, 10]. Suppose we know that ''if x is between 1 and 3, then y is either 7 or 8''. This sentence can describes at least two kinds of knowledge: (1) as a logic implication, and (2) as an association. Assuming that we also know the value of x is 5, the logic implication will infer that y is unknown (i.e., y can be any integer in the interval [0, 10]), but the association knowledge will not make any conclusion regarding y. The two types of knowledge correspond to the two types of fuzzy rules: logic implication is the basis of fuzzy implication rules, while association knowledge is the essence of fuzzy mapping rules. That is, fuzzy implication rules generalize set-to-set implications; whereas fuzzy mapping rules generalize set-to-set associations. The former was motivated to allow intelligent systems to draw plausable conclusions in a way similar to human reasoning; while the latter was motivated to approximate complex relationships (e.g., nonlinear functions) in a cost-effective and easily-comprehensable way.

We will first use an example to illustrate the difference between set-to-set implications and set-to-set mappings. This discussion will then set the stage for us to describe the difference between the theoretical foundation of these two types of fuzzy rules. Suppose x and y are variables taking values from $U = \{a, b, c, d, e, f\}$ and $V = \{r, s, t, u, v\}$, respectively. Suppose further that we know the following implication is true:

$$x \in \{b, c, d\} \Rightarrow y \in \{s, t\} \qquad \text{(EQ 10)}$$

Such a set-to-set implication specifies a set of ***possible** implications* such as $x = b \to y = s$, and a set of **impossible implications** such as $x = b \to y = r$. Notice that if the antecedent is known to be false, the implication is true regardless of *y*'s value. Therefore, the following implications is possible: $x = a \to y = r$. Hence, we can represent the meaning of the set-to-set implication using the following matrix:

$$R(x_i, y_j) = \begin{array}{c} \\ a \\ b \\ c \\ d \\ e \\ f \end{array} \begin{array}{c} \begin{array}{ccccc} r & s & t & u & v \end{array} \\ \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \end{array} \qquad \text{(EQ 11)}$$

where an entry in the relation $R(x_i, y_j)$ represents whether $(x = x_i) \to (y = y_j)$ is possible ("1" means possible, and "0" means impossible). It is also easy to see that such a relation can be constructed by replacing *x* and *y* in Equation 10 with pairs of x_i and y_j and determine whether the resulting (i.e., instantiated) implication is true or false, i.e.,

$$R(x_i, y_j) = \begin{cases} 1 & if \ ((x_i \in \{b, c, d\}) \to (y_j \in \{s, t\})) \\ 0 & if \ \neg((x_i \in \{b, c, d\}) \to (y_j \in \{s, t\})) \end{cases} \qquad \text{(EQ 12)}$$

Now, let us consider a similar set-to-set mapping rule:

$$x \in \{b, c, d\} \text{ map to } y \in \{s, t\} \qquad \text{(EQ 13)}$$

Such a set-to-set mapping specifies a set of possible association between the values of x and y. Such an association can be naturally represented by a relation $R(x_i, y_j)$ that describes whether input-output pair is a possible association based on the mapping rule. Hence, the relation for the mapping rule can be represented as follows:

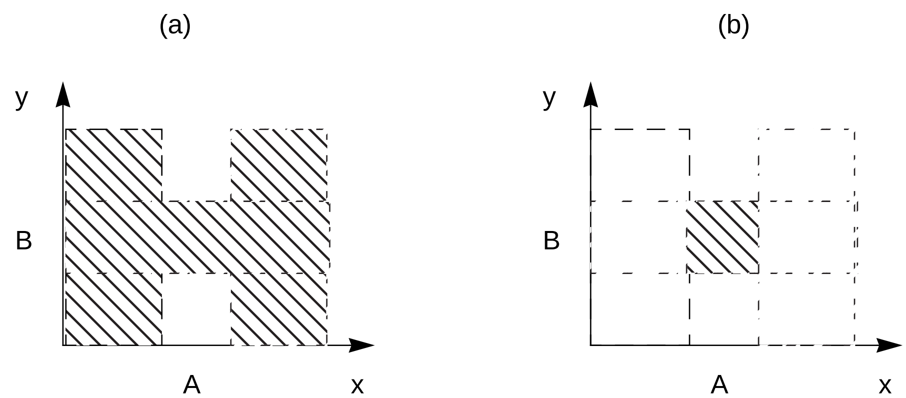
$$R(x_i, y_j) = \begin{cases} 1 & if \ ((x_i \in \{b, c, d\}) \wedge (y_j \in \{s, t\})) \\ 0 & if \ \neg((x_i \in \{b, c, d\}) \wedge (y_j \in \{s, t\})) \end{cases} \qquad \text{(EQ 14)}$$

It may be worthwhile now to point out the difference between the possibility relation of a mapping rule and that of an implication rule. Fig 2 depicts such a difference for the rule "If *x* is *A* then *y* is *B*." The points in shaded areas are possible while those in the white area are not possible. The shaded area in Fig 2 corresponds to the "l"entries in the implications relation in Equation 11.[1] Fig 2 echoes a point we made earlier — implication rules and mapping rules

1. Because the way a fuzzy implication is constructed, we need to turn it counter clockwise 90 degrees to see how Fig 2 (a) is its missor image.

differ in their treatment of the situation that fails to satisfy the if-condition. In Fig 2, these situations are in the two regions outside of the region corresponding to “*x* is *A*.”

FIGURE 2. A Pictoria l View of the Possibility Relation for (a) an Implication Rule and (b) a Corresponding Mapping Rule



Having discussed the meaning of a set-to-set non-fuzzy implication, we can now consider an implication involving fuzzy sets (i.e., fuzzy implication):

$$r_k \qquad (x \text{ is A}) \rightarrow (\text{y is B}) \tag{EQ 15}$$

where *A* and *B* are fuzzy subsets of *U* and *V*, respectively. As in the previous example, this implication also specifies the possibility of various point-to-point implications. The main difference here is that the possibilities are no longer binary. Rather, they become a matter of degree. Therefore, the meaning of the fuzzy implication can be represented by an fuzzy relation *R* defined as

$$R_I(x_i, y_j) = \Pi_I((x = x_i) \rightarrow (y = y_j)) \tag{EQ 16}$$

where Π_I denotes the possibility distribution imposed by the implication. In fuzzy logic, this possibility distribution is constructed from the truth values of the instantiated (i.e., grounded) implications obtained by replacing variables in the implication (i.e., *x* and *y*) with pairs of their possible values (i.e., x_i, y_j):

$$\Pi_{r_k}((x = x_i) \rightarrow (y = y_j)) = t((x_i \text{ is A}) \rightarrow (y_j \text{ is B})) \tag{EQ 17}$$

where *t* denotes the truth value of a proposition. It is easy to see that Equation (17) is a natural extension of Equation (12).

The fuzzy relation of a fuzzy mapping rule in the form of Equation (15) represents the possibility degrees of association between pairs of input and output values. Hence, it extends Equation (14) into the following one:

$$R_{r_k}(x_i, y_j) = \Pi_{r_k}((x = x_i) \wedge (y = y_j)) \tag{EQ 18}$$

The possibility distribution is thus determined from the membership functions of A and B:

$$\Pi_{r_k}((x = x_i) \wedge (y = y_j)) = t((x_i \text{ is A}) \wedge (y_j \text{ is B})) = \mu_A(x_i) \otimes \mu_B(y_j) \quad \text{(EQ 19)}$$

where $\otimes$ denotes a fuzzy conjunction operator.

4.2 Types of Fuzzy Implication Functions

The truth value of the fuzzy implication rule "x_i is $A \to y_j$ is B" in Equation (17) is defined in terms of the truth value of the antecedent proposition "x_i is A" and the truth value of the consequent proposition "y_j is B." For the convenience of our discussion, we will refer to these truth values as α_i and β_j, respectively, i.e., $t(x_i \text{ is } A) = \alpha_i$ $t(y_j \text{ is } B) = \beta_j$. The truth value of the implication (x_i *is* $A \to y_j$ *is B*) is thus a function I of α_i and β_j:

$$t(x_i\ A \to Y_j \text{ is } B) = I(\alpha_i, \beta_j)$$

We call the function *I* an "*implication function*."

Various definitions of implication functions have been developed. To compare and evaluate them, several intuitive criteria of desired inference results of fuzzy implications have been established. These criteria will thus form the basis for evaluating and comparing different fuzzy implication functions. An example of such a criteriia is given below:

Given: x is $A \to y$ is B
x is not A

Infer: y is V (unkown)

Notice that "*y* is *V*" (assigning the entire universe of discourse *V* to *y*) represents "*y* is unknown". Other intuitive criteria are summarized in Table 1..

TABLE 1. Intuitive Criteria for Reasoning Involving Fuzzy Implication x is $A \to y$ is B

Criterion	Given	Infer
I	*x is A*	*y is B*
II-1	*x is very A*	*y is very B*
II-2	*x is very A*	*y is B*
II-2*	x is A' and $A' \subset A$	*y is B*
III -1	*x is more or less A*	*y is more or less B*
III-2	*x is more or less A*	*y is B*
IV	*x is not A*	*y is V (unknown)*
V	*y is not B*	*x is not A*
VI	*y is not (very B)*	*x is not (very A)*
VII	*y is not (more or less B)*	*x is not (more or less A)*
VIII	*y is B*	*x is U (unknown)*
IX	y is $B \to z$ is C	x is $A \to z$ is C

Fuzzy implications can be classified into three families. Each family extends a particular logic formulation of implication in propositional logic. Even though these formulations are equivalent in classical logic, they are not equivalent in fuzzy logic because the law of excluded middle no longer holds in fuzzy logic as we explained in Section 3. Fuzzy implications within a family differ on their choice of the fuzzy conjunction and fuzzy disjunction operators. We briefly describe each family below.

The first family of fuzzy implication is obtained by generalizing material implications in two-valued logic to fuzzy logic. A material implication $p \to q$ is defined as $\neg p \vee q$. Generalizing this to fuzzy logic gives us $t(p \to q) = t(\neg p \vee q)$. More specifically, fuzzy implications in this family can be generically defined as:

$$\begin{aligned} t(x_i \ \ is \ \ A \to y_j \ \ is \ \ B) &= t(\neg(x_i \ \ is \ \ A) \vee (y_j \ \ is \ \ B)) \\ &= ((1-\mu_A(x_i)) \oplus \mu_B(y_j)) \end{aligned} \quad \text{(EQ 20)}$$

An example of fuzzy implication in this familty is Zadeh's arithmetic fuzzy implication:

$$t(x_i \ \ is \ \ A \to y_j \ \ is \ \ B) = 1 \wedge (1-(\mu_A(x_i) + (\mu_B(y_j)) \quad \text{(EQ 21)}$$

which is obtained using a bounded sum operator for fuzzy disjunction.

The second family of fuzzy implication is based on logic equivalence between implications $p \to q$ and $\neg p \vee (p \wedge q)$. Fuzzy implications in this family thus have the following form:

$$\begin{aligned} t(x_i \ \ is \ \ A \to y_j \ \ is \ \ B) &= t(\neg(x_i \ \ is \ \ A) \vee [(x_i \ \ is \ \ A) \wedge y_j \ \ is \ \ B]) \\ &= (1-\mu_A(x_i)) \oplus (\mu_A(x_i) \otimes \mu_B(y_j)) \end{aligned} \quad \text{(EQ 22)}$$

An example of fuzzy implication in this family is Zadeh's maximum fuzzy implication function:

$$t(x_i \ \ is \ \ A \to y_j \ \ is \ \ B) = (1-\mu_A(x_i)) \vee (\mu_A(x_i) \wedge \mu_B(y_j)) \quad \text{(EQ 23)}$$

which is obtained by using min for fuzzy conjunction and max for fuzzy disjunction.

The third family of fuzzy implication generalizes the "standard sequence" of many-valued logic and its variants. The implication in these logic systems is defined to be true whenever the consequent is as true or truer than the antecedent, i.e., $t(p \to q) = 1$ whenever $t(p) \le t(q)$. This is an important property of many multivalued logic systems because it allows the following tautology (i.e., a logic formula that is always true) in two-valued logic to be maintained in multi-valued logic: $f \to f$ where f is any formula. In other words, a logic formula always implies itself, regardless of its truth value. The fuzzy implication function in this family can all be described in the following form:

$$\begin{aligned} t(x_i \ \ is \ \ A \to y_j \ \ is \ \ B) &= \sup\{\alpha | \alpha \in [0,1], \ \alpha \otimes t(x_i \ \ is \ \ A) \le t(y_i \ \ is \ \ B)\} \\ &= \sup\{\alpha | \alpha \in [0,1], \ \alpha \otimes \mu_A(x_i) \le \mu_B(y_j)\} \end{aligned} \quad \text{(EQ 24)}$$

Three fuzzy implication functions in this family are given below:

Standard sequence fuzzy implication

$$t(x_i \;\; is \;\; A \rightarrow y_j \;\; is \;\; B) \; = \begin{cases} 1 & \mu_A(x_i) \le \mu_B(y_j) \\ 0 & \mu_A(x_i) > \mu_B(y_j) \end{cases} \qquad \text{(EQ 25)}$$

Godelian sequence fuzzy implication

$$t(x_i \;\; is \;\; A \rightarrow y_j \;\; is \;\; B) \; = \begin{cases} 1 & \mu_A(x_i) \le \mu_B(y_j) \\ \mu_B(y_j) & \mu_A(x_i) > \mu_B(y_j) \end{cases} \qquad \text{(EQ 26)}$$

Goguen's fuzzy implication

$$t(x_i \;\; is \;\; A \rightarrow y_j \;\; is \;\; B) \; = \begin{cases} 1 & \mu_A(x_i) \le \mu_B(y_j) \\ \dfrac{\mu_B(y_j)}{\mu_A(x_i)} & \mu_A(x_i) > \mu_B(y_j) \end{cases} \qquad \text{(EQ 27)}$$

These three fuzzy implication functions came from, respectively, the standard sequence many-valued logic system (denoted S_n in the literature), a many-valued logic system proposed by Kurt Godel (denoted G_n in the literature), and a many-valued logic system J. A. Goguen introduced in 1969. Fig 3 shows graphically the function surface of the five fuzzy implication functions we discussed.

Even though implication functions in multivalued logic systems can be used for constructing fuzzy implication relations, approximate reasoning in fuzzy logic is fundamentally different from logic inference in multi-valued logic — approximate reasoning infers possible values of a variable, whereas multivalued logic infers the truth values of propositions. The connection between the two was established by Equation 17. Even if we choose to use a fuzzy implication function originated in a multivalued logic system (e.g., standard sequence, Godelian implication, or Goguen's implication), approximate reasoning still benefits from other important concepts and techniques in fuzzy logic such as the compositional rule of inference, fuzzy relations, and possibility distributions. Without them, approximate reasoning would not have been possible.

FIGURE 3. Five Fuzzy Implication Functions: (a) Zadeh's Arithmetic Fuzzy Implication, (b) Zadeh's Maxmin Fuzzy Implication, (c) Standard Sequence Fuzzy Implication, (d) Godelion Fuzzy Implication, and (e) Goguen Fuzzy Implication

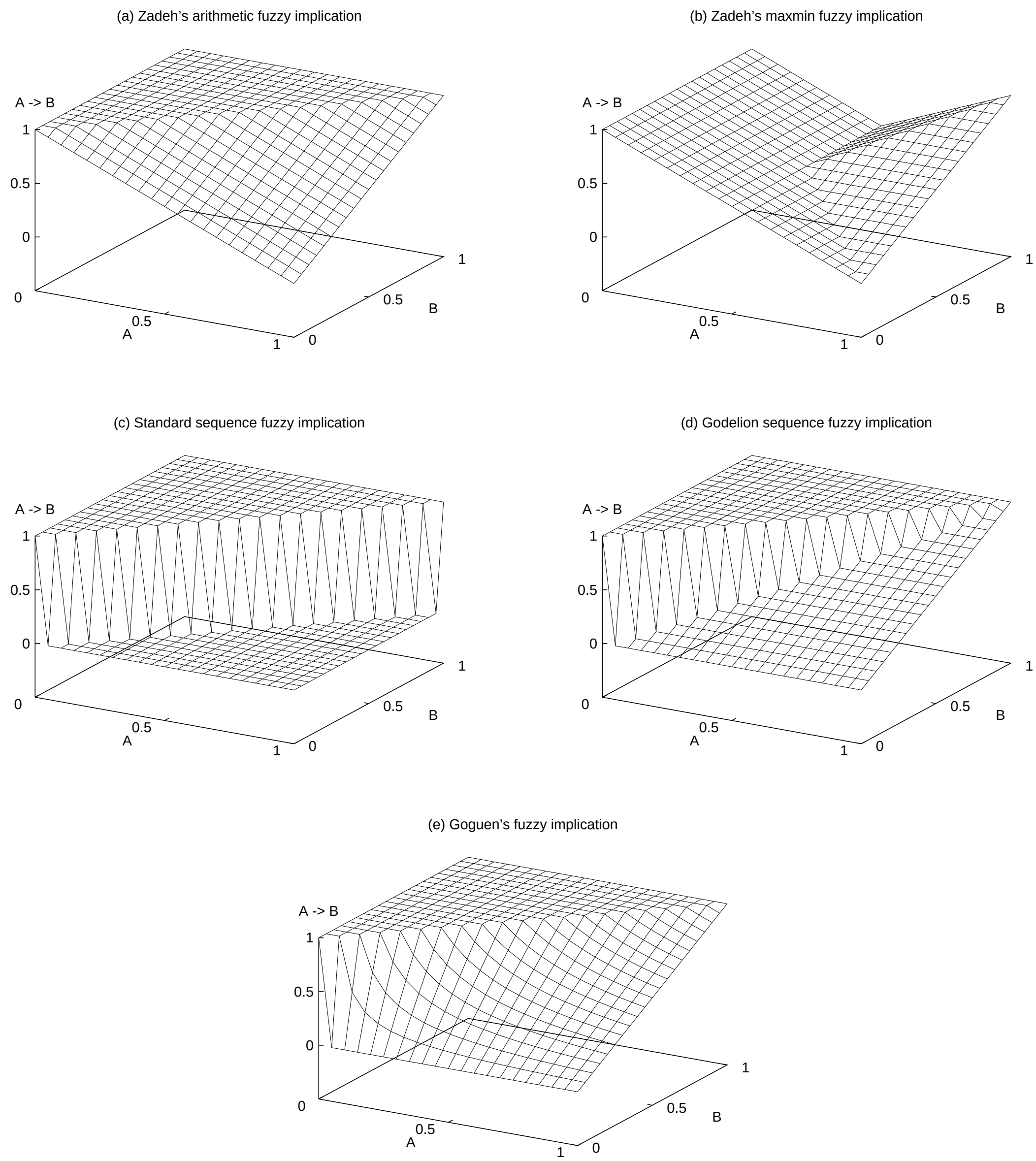


Table 2 summaries how the criteria introduced in Table 1 are satisfied by the five fuzzy implication functions based on sup-min composition (except that the sup-product composition is applied to Goguen's fuzzy implication). We applied the sup-product composition to Goguen's fuzzy implication because the conjunction operator in Goguen's multivalued logic system uses product as its conjunction operator. In other words, the choice of implication function and the choice of disjunction/conjunction operator in the compositional rule of inference are not unrelated.

TABLE 2. Satisfaction of Fuzzy Inference Criteria by Fuzzy Implication Functions *I(x,y)*
X*: denotes that a criterion is not supported by *I(x,y)
and *O* denotes that a criterion is supported by *I(x,y)*

	Arithmetic	Maximum	Standard	Godel	Goguen
I	X	X	O	O	O
II-1	X	X	O	X	X
II-2	X	X	X	O	O
III-1	X	X	O	O	X
III-2	X	X	X	X	X
IV	O	O	O	O	O
V	X	X	O	X	X
VI	X	X	O	X	X
VII	X	X	O	X	X
VIII	O	X	O	O	O
IX	X	X	O	O	O

EXAMPLE 1 Let *U* and *V* be two universes representing numeric ratings (from 1 to 10) of redness and ripeness of tomatoes, respectively. We denote the variable of these two ratings as *x* and *y*, respectively. Let *Red* be a fuzzy subset of *U* defined as

$$Red = 0.25/6 + 0.5/7 + 0.75/8 + 1/9 + 1/10$$

and *Ripe* be a fuzzy subset of *V* defined as

$$Ripe = 0.25/7 + 0.5/8 + 0.75/9 + 1/10$$

We are also given the fuzzy implication

$$tomato \ \ is \ \ Red \rightarrow tomato \ \ is \ \ Ripe$$

We denote fuzzy implication relations obtained from the standard sequence and Goguen's implication as R_s, and R_{gg}, respectively. Suppose we know that a specific tomato is very red, we can express this information as "tomato is VERY Red" where VERY is a hedge that modifies the meaning of a fuzzy set by taking the square of its membership function: (i.e., $\mu_{\text{VERY A}} = (\mu_A)^2$). Applying the sup-min composition to the standard sequence implication, we obtain the following possibility distribution about the ripeness of the tomato:

$$\Pi_{ripeness} = \Pi_{redness} \circ R_s = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0.065 & 0.25 & 0.5625 & 1 \end{bmatrix} \quad \text{(EQ 28)}$$

The inferred possibility distribution of ripeness indicates that the tomato is VERY *Ripe*. If we apply sup-product composition to Goguen's implication for the same problem, we obtain the following result:

$$\Pi_{ripeness} = \Pi_{redness} \circ R_{gg} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{2} & \frac{3}{4} & 1 \end{bmatrix} \qquad \text{(EQ 29)}$$

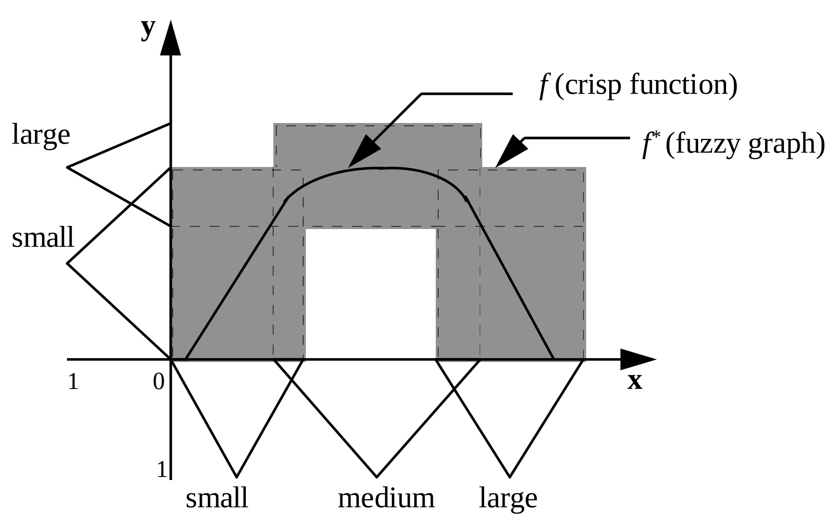
The inferred possibility distribution of the tomato's ripeness is thus *Ripe*.

The example above illustrates that standard sequence implication can satisfy Criterion II-1, while Goguen's implication can satisfy Criterion II-2. However, we should point out that these results require some conditions about the membership functions in the implication rules. Even though these conditions are often satisfied in practice, we should not overlook them. S. Fukami, M. Mizumoto, and K. Tanaka were the first to formally analyze the relationship between various fuzzy implication functions and the intuitive criteria of fuzzy implication rules [9].

5 FOUNDATION OF FUZZY MAPPING RULES -- FUZZY GRAPH

The foundation of fuzzy mapping rules are fuzzy graph[41].

FIGURE 4. Fuzzy Graph Approximation by a Disjunction of Cartesian Products



A fuzzy graph f^* from X to Y is a union of Cartesian products involving linguistic input-output associations (i.e., pairs of "x is A_i" and "y is B_i"). Let f^* be a fuzzy graph described by a set of fuzzy mapping rules in the form of "IF x is A_i then y is B_i." The fuzzy graph can be expressed mathematically as:

$$f^* = \bigcup_i A_i \times B_i \qquad \text{(EQ 30)}$$

In f^*, + denotes the fuzzy disjunction. The Cartesian product of A and B, denoted by $A \times B$, is defined as

$$\mu_{A \times B}(u, v) = \mu_A(u) \otimes \mu_B(v)$$

An expression of the form $A \times B$ where A and B are *words* (fuzzy sets) is referred as a Cartesian granule [42]. Figure depicts a fuzzy graph consisting of three fuzzy mapping rules:

f: IF *X* is small THEN *Y* is small.

IF *X* is medium THEN *Y* is large.

IF *X* is large THEN *Y* is small.

The resulting fuzzy graph is basically a fuzzy relation.

The inference (i.e. interpolative reasoning) of such a set of fuzzy mapping rulesl is also based on *compositional rule of inference* introduced earlier. Given an input "*x* is *A*" to the model, the inferred output of the model is a possibility distribution *B'* of *y*:

$$B' = A' \circ f^* = A' \circ \left(\bigcup_i A_i \times B_i\right) \qquad \text{(EQ 31)}$$

where f^* represents the fuzzy graph of a given fuzzy model, $\circ$ denotes the compositional rule of inferenc.

6 TYPES OF FUZZY MODELS

A set of fuzzy mapping rules form a fuzzy model. Fuzzy models can be classified into two categories: (1) non-additive fuzzy models that aggregate the output of fuzzy rules using the maximum operator, and (2) additive fuzzy models that aggregate the output of rules using the addition operator. The choice of these aggregation operators corresponds to the choice of fuzzy union operator in fuzzy graph (i.e., Equation 30). The most well known non-additive fuzzy model is the Mamdani model, which is named after E. H. Mamdani who developed the first fuzzy logic controller using the model. Most fuzzy control systems developed in the 80's use the Mamdani model. The first additive fuzzy model is the Takagi-Sugeno-Kang (TSK) model, which was first introduced by T. Takagi and M. Sugeno around 1985. The TSK model has drawn much more attention in the 90's, both in the research community and in the industry. One of the main advantages of the TSK model is that it can approximate a function using fewer rules. Another additive fuzzy model is Kosko's standard additive model (SAM) [16]. The inference scheme of the Mamdani model can be derived from Equation 31 by using the min operator for the Cartesian product, by using the max operators for the fuzzy union operation, and by using the sup-min composition. Additive fuzzy models can also be derived by choosing appropriate operators for these operations. For instance, the SAM model uses the product operator for the fuzzy intersection and the cartesian product operation. Before deriving these models from their foundation, however, we discuss a desired property regarding fuzzy graph.

THEOREM 1 (**Distributive Property of Fuzzy Graph**) Suppose a fuzzy model that describes a fuzzy mapping from $U * V$ to W is described by n rules in the form of

$$\text{IF } x \text{ is } A_i \text{ AND } y \text{ is } B_i \text{ THEN } z \text{ is } C_i \tag{EQ 32}$$

where A_i, B_i, and C_i are fuzzy subsets of U, V, and W, respectively. The model's fuzzy graph f^* is expressed as

$$f^* = \bigcup_{i=1}^{n} R_i = \bigcup_{i=1}^{n} (\bar{A}_i \cap \bar{B}_i) \times C_i \tag{EQ 33}$$

where R_i is the fuzzy relation of *i*th. rule, $\bar{A}_i$ and $\bar{B}_i$ denote cylindircal extension of A_i and B_j respectively (i.e., a simple technique to extend these fuzzy sets into the space $U \times V$). Let *A'* and *B'* be fuzzy subsets of *U* and *V* respectively. If we compose *A'* and *B'* with f^* using **sup-min** composition and compute the union in Equation (33) using the **max** operator, then composing inputs with f^* (i.e., the entire fuzzy rule-based model) are equivalent to first composing inputs with individual rules in f^* and then aggregating their composition results.

$$(\bar{A}' \cap \bar{B}') \circ f^* = \bigcup_{i=1}^{n} (\bar{A}' \cap \bar{B}') \circ R_i \tag{EQ 34}$$

The proof of this and remaining theorems can be found in [30].

We now state a theorem that shows how the Mamdani model can be derived from fuzzy graph and Equation 31.

THEOREM 2 (**Mamdani Model**) Suppose a fuzzy rule-based model maps $X \times Y$ to Z using a set of *n* rules in the form of

IF x is A_i AND y is B_i THEN z is C_i $\quad 1 \le i \le n$

and receives inputs in the form of x is *A'* and y is *B'* where *A'* and *B'* are fuzzy subsets of *U* and *V*. Suppose the fuzzy inference of the model is based on a sup-min composition between inputs and a fuzzy graph that is defined using *max* and *min* for all fuzzy disjunctions and fuzzy conjunctions operations. Then, the output of the fuzzy model (before defuzzification), denoted by C', is characterized by the following membership function:

$$\mu_{C'}(z) = \max_{i=1}^{n} (\alpha_i \wedge \mu_{C'_i}(z)) \tag{EQ 35}$$

where $\alpha_i = \sup_x(\mu_{A'}(x) \wedge \mu_{A_i}(x)) \wedge \sup_y(\mu_{B'}(y) \wedge \mu_{B_i}(y))$ and $\wedge$ denotes the min operator.

The foundation of the standard additive model is a fuzzy graph, the sup-product composition, and the use of "addition" as a rule aggregation operator. We formally state this as a theorem below.

THEOREM 3 (Standard Additive Model) Suppose f^* is a fuzzy graph consisting of rules in the form of

IF x is A_i AND y is B_i THEN z is C_i.

If the inference of the model uses sup-product composition, "product" for all fuzzy conjunction, and "addition" for rule aggregation, and centroid defuzzification, then the model's output for crisp inputs $x = x_0$; $y = y_0$ is

$$z = Centroid\left(\sum_{i=1}^{n} \mu_{A_i}(x_0) \times \mu_{B_i}(y_0) \times C_i\right) \quad \text{(EQ 36)}$$

7 FUZZY MODEL IDENTIFICATION

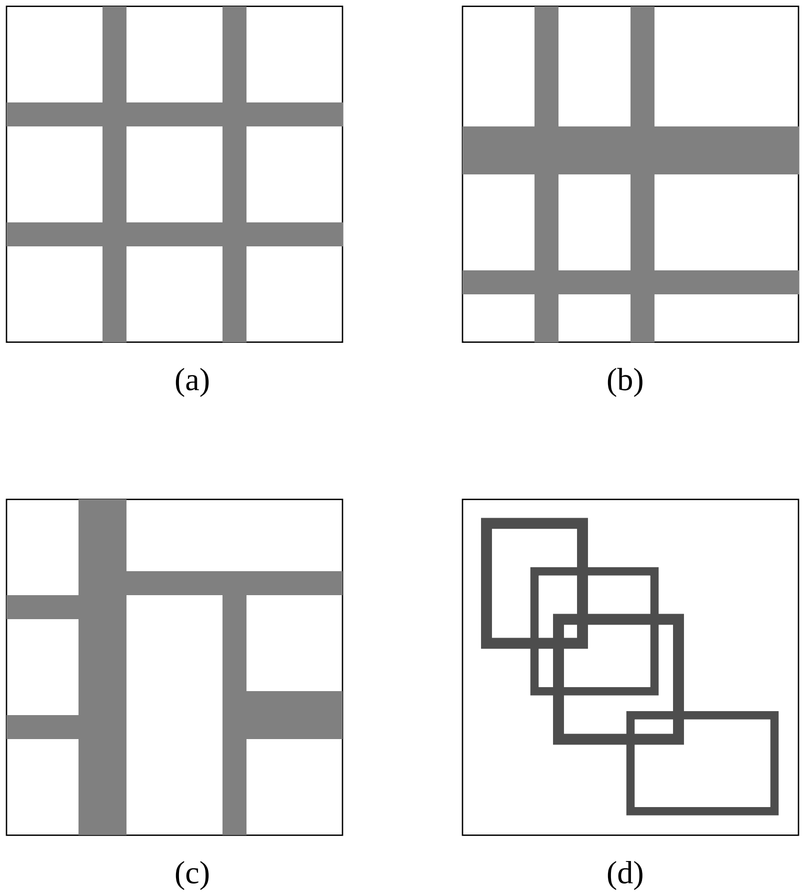
The identification of fuzzy models consists of three basic subproblems: structure identification, parameter estimation, and model validation. *Structure identification* involves finding the important input variables from all possible input variables, specifying membership functions, partitioning input space, and determining the number of fuzzy rules comprising the underlying model. *Parameter estimation* involves the determination of unknown parameters in the model using some optimization method based on both linguistic information obtained from human experts and numerical data obtained from the actual physical system. Structure specification and parameter estimation are interwoven, and either of them cannot be independently identified without resort to another. *Model validation* involves testing the model based on some performance criterion (e.g., accuracy). If the model cannot pass the test, we must modify the model structure and re-estimate the model parameter. This process iterates until a satisfactory model is found. We will discuss two of the most important issues in fuzzy modeling for high dimension problems: (1) identifying the fuzzy partition of the input space, and (2) dealing with the curse of dimensionality.

7.1 Fuzzy Partition of the Input Space

Fuzzy partitions of input space define the antecedent of a fuzzy model. The number of fuzzy partitions usually determines the number of fuzzy rules comprising the underlying model as well. There are three types of fuzzy partitions: (1) the grid partition, (2) the tree partition, and (3) the scatter partition.

The *grid partition* is the most commonly used fuzzy partitioning methods in practice (particularly in control applications). The rule in Equation 32, for instance, partitions the input space this way. Wang and Mendel has used this type of fuzzy partition in their procedure for fuzzy rule extraction from numerical data [28]. Fig. 5(a) illustrates a typical grid partition in a two-dimensional input space. The grids shown in Fig. 5(a) are uniformly partitioned and static, and the performance of the resultant model depends entirely on the initial definition of these grids. An *adaptive* fuzzy grid partition can be obtained if we introduce some learning procedure in constructing the partition. Two typical learning procedures used in practice are the gradient descent method suggested by Jang and genetic algorithms suggested by Karr [12, 13]. Fig. 5(b) gives an example of an adaptive fuzzy grid partition in a two-dimensional input space. Grid partition (both static and adaptive) is convenient to use, but it may encounter serious "rule explosion" problem when the number of input variables is large. This problem is closely related to the so-called “curse of dimensionality”, the well-known problem of exponentially increasing complexity with the number of input variables.

FIGURE 5. Various Methods for Partitioning the Input Space: (a) Grid Partition (Static); (b) Grid Partition (Adaptive); (c) Tree Partition; (d) Scatter Partition



Tree partition is another method used in partitioning input space. Fig. 5(c) gives an example of a tree partition in a two-dimensional input space. A tree partition results from a series of *guillotine cuts*. By a guillotine cut, we mean a cut that is made entirely across the subspace to be partitioned; each of the regions so produced can then be subjected to independent guillotine cutting. At the beginning of the ith iteration step, the input space is partitioned into i regions. Then another guillotine cut is applied to one of the regions to further partition the entire space into (i+1) regions. There are various strategies to decide which dimension to cut and where to cut it at each step. Some are based merely on the distribution of training examples; others take the parameter identification methods into consideration [25]. Tree partition relieves the problem of rule explosion to a great degree, but it is not easy to use in practice. A number of heuristics are usually needed to find a proper tree structure, and there may be difficulties involved in designing an optimal tree partition.

A more flexible method for partitioning input space in high-dimensional modeling problem is the *Scatter partition*. An example of such a partition is shown in Fig. 5(d) which appears in a two-dimensional input space. Instead of covering the whole input space, this method tries to find a subset of the input space that characterizes the fuzzy regions of possible occurrence of training examples. Each fuzzy region is associated with a combination of antecedent membership functions, and the number of the regions defines the number of fuzzy rules. It is usually difficult to find such a subset of input space (and consequently the fuzzy regions) intuitively or manually; instead, some learning or automatic procedure has to be adopted.

7.2 Dealing with the Curse of Dimensionality

As we mentioned in the previous section, a long-standing problem in fuzzy rule-based modeling is the "curse of dimensionality", which occurs because the number of rules increases exponentially as the number of input variables increases [16]. Several approaches using singular value decomposition (SVQ) to overcome this problem have been proposed [20, 34]. Other methods reduce the number of fuzzy rules in a given rule base by *fusing* (rather than eliminating) rules based on some *similarity measure* [25, 4, 18]. An alternative approach is to *augment* a rule base when the rules are not complete [28, 23].

Evaluating fuzzy models purely based on their fitness to training data, like any other modeling paradigm, can potentially lead to the overfitting problem. One way to address this issue is to use information theoretic criteria to evaluate fuzzy models so that the complexity of these models is explicitly taken into account in evaluating their optimality [32]. An important benefit of fuzzy model is that its rules are interpretable because they capture a local relationship between the model's input and output. However, this virtue may be lost if a fuzzy model is evaluated only by its global performance. A remedy to this problem is to combine local learning (i.e., learning individual rules) with globall learning (i.e., learning the entire rule set) [33].

8 SUMMARY

In this paper, we have summarized major concepts and techniques in fuzzy logic. We have also presented a modern perspective about two types of fuzzy rules: fuzzy implication rules and fuzzy mapping rules. The latter have been widely used in fuzzy logic control and other industrial applications. This new perspective not only clarifies the formal foundation of these rules, but also sheds lights on how to deal with various challenges in identifying and learning fuzzy rule-based models for high dimensional problems.

Acknowledgments

We would like to thank Henri Prade and Didies Dubois for fruitful technical exchanges regarding fuzzy logic and artificial intelligence. We also would like to thank Reza Langari and Liang Wang for many fruitful discussions. This research was partially supported by National Science Foundation Young Investigator Award IRI 9257293.

Reference

[1] B. Babuska, M. Setnes, U. Kaymak, and H. R. van Nauta Lemke, Rule base simplification with similarity measures, in *Proc. of the Fifth IEEE Int. Conf. on Fuzzy Systems*, New Orleans, LA, September, 1996, pp. 1642-1647.

[2] J. C. Bezdek. Fuzzy mathematics in pattern classification, *PhD thesis*, Center for Applied Mathematics, Cornell University, Ithaca, New York, USA, 1973.

[3] J. Bezdek. Fuzziness vs. Probability -Again !? *IEEE Transactions on Fuzzy Systems*, Vol. 2, No. 1, pp. 1-3, February 1994.

[4] C. T. Chen, Y. J. Chen, and C. C. Teng, Simplification of fuzzy-neural systems using similarity analysis, *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 26, pp. 344-354, 1996.

[5] D. Dubois and H. Prade. An introduction to possibilistic and fuzzy logics (with discussions). In P. Smets, E. H. Mamdani, D. Dubois and H. Prade (ed.). *Non-Standard Logics for Automated Reasoning* (Academic Press, New York, 1988), pp. 287-326.

[6] D. Dubois and H. Prade, Basic issues on fuzzy rules and their application to fuzzy control, in: Fuzzy Logic and Fuzzy Control (D. Driankov, P.W. Eklund, A.L. Ralescu, eds.), Springer-Verlag, Lecture Notes in Artificial Intelligence, Vol. 833, 1994, 3-14

[7] J. C. Dunn. A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters, *J. Cybernetics*, 3, pp. 32-57, 1973.

[8] C. Elkan et. al., The Paradoxical Success of Fuzzy Logic, (followed by fifteen Responses from L. A. Zadeh et. al.), *IEEE Expert*, Vol. 9, No. 4, pp. 9-49, Aug 1994.

[9] S. Fukami, M. Mizumoto, and K. Tanaka. Some considerations on fuzzy conditional inference. *Fuzzy Sets and Systems*, Vol. 4, pp. 243-273, 1980.

[10] B. R. Gaines and L. J. Kohout. The fuzzy decade : a bibliography of fuzzy systems and closely related topics. In M. M. Gupta, G. N. Saridis and B. R. Gaines, editors, *Fuzzy Automata and Decision Processes*, North-Holland, pp., 403-490, 1977.

[11] K. Hirota, History of Industrial Applications of Fuzzy Logic in Japan. In J. Yen, R. Langari, L. A. Zadeh, editors, *Industrial Applications of Fuzzy Logic and Intelligent Systems*, IEEE Press, pp. 43-54, 1995.

[12] J. -S. R. Jang, ANFIS: adaptive-network-based fuzzy inference systems, *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 23, pp. 665-685, 1993.

[13] C. Karr, Genetic algorithms for fuzzy controllers, *AI Expert*, vol. 6, pp. 26-33, 1991.

[14] G. J. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic: Theory and Applications*, Prentice Hall, 1995.

[15] B. Kosko. *Neural Networks and Fuzzy Systems*. Prentice-Hall, 1**990.**

[16] B. Kosko. *Fuzzy Engineering*. Prentice Hall, 1997.

[17] C. C. Lee. Fuzzy logic in control systems: Fuzzy logic controller, parts i and ii. *IEEE Transactions on Systems, Man, and Cybernetics*, 20, 1990.

[18] C. -T. Lin and C. S. G. Lee, *Neural Fuzzy Systems: A Neuro-Fuzzy Synergism to Intelligent Systems*, Prentice Hall, Upper Saddle River, NJ, 1996.

[19] E. H. Mamdani and S. Assilian. An experiment in linguistic synthesis with a fuzzy logic controller. *International Journal of Machine Studies*, 7(1), 1975.

[20] G. C. Mouzouris and J. M. Mendel, Designing fuzzy logic systems for uncertain environments using a singular-value-QR decomposition method, in *Proc. of the Fifth IEEE Int. Conf. on Fuzzy Systems*, New Orleans, LA, September, 1996, pp. 295-301.

[21] T. S. Perry. *Lotfi A. Zadeh (cover story)*, *IEEE Spectrum*, pp. 32- 35, June 1995.

[22] T. J. Procyk and E. H. Mamdani. A linguistic self organizing process controller. *Automatica*, 15(1), 1979.

[23] T. Sudkamp and R. J. Hammell II, Interpolation, completion, and learning fuzzy rules, *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 24, pp. 332-342, 1994.

[24] M. Sugeno and K. T. Kang. Structure Identification of Fuzzy Model. *Fuzzy Sets and Systems*, Vol. 28, 1988.

[25] C. -T. Sun, Rule-base structure identification in an adaptive-network-based fuzzy inference system, *IEEE Transactions on Fuzzy Systems*, vol. 2, pp. 64-73, 1994.

[26] T. Takagi and M. Sugeno. Fuzzy identification of systems and its application to modeling and control. *IEEE Transactions on Systems, Man, and Cybernetics*, 15(1), 1985.

[27] M. Togai and H. Watanabe. Expert systems on a chip: an engine for real-time approximate reasoning. *IEEE Expert Systems Magazine*, 1(1), 1986.

[28] L. Wang and J. M. Mendel. Fuzzy Basis Functions, Universal Approximation, and Orthogonal Least- Squares Learning. *IEEE Transactions on Neural Networks*. Vol. 3, No. 5, pp. 807-814, Sept. 1992.

[29] S. Yasunobu and S. Miyamoto. Automatic train operation by fuzzy predictive control. In M. Sugeno, editor, *Industrial Applications of Fuzzy Control*. North Holland, 1985.

[30] J. Yen and R. Langari, *Fuzzy Logic: Intelligence, Control, and Information*, Prentice Hall, 1999.

[31] J. Yen, R. Langari, and L. A. Zadeh, *Industrial Applications of Fuzzy Logic and Intelligent Systems*, IEEE Press, 1995.

[32] J. Yen and L. Wang, Application of statistical information criteria for optimal fuzzy model construction, *IEEE Transactions on Fuzzy Systems*, Vol 6, No. 3, Aug. 1998.

[33] J. Yen, L. Wang, and W. Gillespie, Improving the interpretability of TSK fuzzy models by combining global learning and local learning, *IEEE Transactions on Fuzzy Systems*, Vol 6, No. 4, Nov. 1998.

[34] J. Yen and L. Wang, Simplifying fuzzy rule-based models using orthogonal transformation methods, *IEEE Transactions on Systems, Man, and Cybernetics*, Vol 29: Par B, No. 1, Feb. 1999.

[35] L. A. Zadeh. Fuzzy sets. *Information and Control*, 8, 1965.

[36] L. A. Zadeh. Probability measures and fuzzy events. *J. of Math. Analysis and Applications*, Vol 23, No. 2, pp. 421-427, 1968.

[37] L. A. Zadeh. Toward a theory of fuzzy systems. *In Aspects of Network and System Theory*. Rinehart and Winston, New York, pp. 469-490, 1971.

[38] L. A. Zadeh. Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3, 1973.

[39] L. A. Zadeh. The concept of a linguistic variable and its application to approximate reasoning-I, II, III. *Information Sciences*, Vol. 8, pp. 199-249, pp. 301-357, Vol. 9, pp.43-80, 1975.

[40] L. A. Zadeh. Possibility theory and soft data analysis. In Cobb, L. and R. M. Thrall (ed.). *Mathematical Frontiers of the Social and Policy Sciences.* Westview Press: Boulder, Colodardo, pp.69-129, 1981.

[41] L. A. Zadeh. Fuzzy Logic, Neural Networks, and Soft Computing. *Communications of the AC*M, Vol 37(3), pp. 77-84, March 1994.

[42] L. A Zadeh. Fuzzy logic = computing with words. *IEEE Trans. on Fuzzy Systems*, Vol. 4, No. 2, 1996.