

**[Unit: Introduction to fuzzy set theory]**  
**Fuzzy Systems (CSc 613)**

**Jagdish Bhatta**

**Central Department of Computer Science & Information Technology**  
**Tribhuvan University**

### **Probabilistic Reasoning:**

Probabilistic reasoning is used when we consider the diagnostic accuracy of tests in our clinical decisions. It is also called Bayesian reasoning, being based on Bayes' theorem, in which the probability of a hypothesis is modified by further data.

The aim of a **probabilistic logic** (also **probability logic** and **probabilistic reasoning**) is to combine the capacity of probability theory to handle uncertainty with the capacity of deductive logic to exploit structure. The result is a richer and more expressive formalism with a broad range of possible application areas. Probabilistic logics attempt to find a natural extension of traditional logic truth tables: the results they define are derived through probabilistic expressions instead.

Bayes's rule or Bayesian decision theory is a widely used variation of probability theory that analyzes past uncertain situations and determines the probability that a certain event caused the known outcome. This analysis is then used to predict future outcomes.

### **Uncertain Knowledge:**

Let action  $A_t$  = leave for airport  $t$  minutes before flight. Will  $A_t$  get me there on time?

Problems:

1. Partial observability (road state, other drivers' plans, etc.)
2. Noisy sensors (radio traffic reports)
3. Uncertainty in action outcomes (flat tyre, etc.)
4. Complexity of modeling and predicting traffic

Hence a purely logical approach either

1. Risks falsehood: " $A_{25}$  will get me there on time" or
2. Leads to conclusions that are too weak for decision making: " $A_{25}$  will get me there on time if there's no accident on the bridge and it doesn't rain and my tires remain intact etc etc."
3.  $A_{1440}$  might reasonably be said to get me there on time but I'd have to stay overnight in the airport...

*[Refer Russel and Norvig]*

**Handling Uncertainty:**

Instead of providing all condition it can express with degree of beliefs in the relevant sentences.

Example:

Say we have a rule

*if toothache then problem is cavity*

But not all patients have toothaches because of cavities (although perhaps most do). So we could set up rules like

*if toothache and not(gum disease) and not(filling) and .....then problem is cavity*

This gets very complicated! a better method would be to say

*if toothache then problem is cavity with probability 0.8*

A most important tool for dealing with degree of beliefs is probability theory, which assigns to each sentence a numerical degree of belief between 0 & 1.

**Making decisions under uncertainty:**

*Problem: to reach airport on time*

Suppose I believe the following:

$$P(A_{25} \text{ gets me there on time} | \dots) = 0.04$$

$$P(A_{90} \text{ gets me there on time} | \dots) = 0.70$$

$$P(A_{120} \text{ gets me there on time} | \dots) = 0.95$$

$$P(A_{1440} \text{ gets me there on time} | \dots) = 0.9999$$

Which action to choose?

- Depends on my preferences for missing flight vs. length of wait at airport, etc. Utility theory is used to represent and infer preferences

Decision theory = utility theory + probability theory

*The rational decision depends on both the relative importance of various goals and the likelihood that, and degree to which, they will be achieved.*

### **Basic Statistical methods – Probability:**

The basic approach statistical methods adopt to deal with uncertainty is via the axioms of probability:

- Probabilities are (real) numbers in the range 0 to 1.
- A probability of  $P(A) = 0$  indicates total uncertainty in  $A$ ,  $P(A) = 1$  total certainty and values in between some degree of (un)certainly.
- Probabilities can be calculated in a number of ways.

Probability = (number of desired outcomes) / (total number of outcomes)

So given a pack of playing cards the probability of being dealt an ace from a full normal deck is 4 (the number of aces) / 52 (number of cards in deck) which is  $1/13$ . Similarly the probability of being dealt a spade suit is  $13 / 52 = 1/4$ .

**Conditional probability,  $P(A|B)$ , indicates the probability of event  $A$  given that we know event  $B$  has occurred.**

The aim of a **probabilistic logic** (or **probability logic**) is to combine the capacity of probability theory to handle uncertainty with the capacity of deductive logic to exploit structure. The result is a richer and more expressive formalism with a broad range of possible application areas. Probabilistic logic is a natural extension of traditional logic truth tables: the results they define are derived through probabilistic expressions instead. The difficulty with probabilistic logics is that they tend to multiply the computational complexities of their probabilistic and logical components.

### **Random Variables:**

In probability theory and statistics, a **random variable** (or **stochastic variable**) is a way of assigning a value (often a real number) to each possible outcome of a random event. These values might represent the possible outcomes of an experiment, or the potential values of a quantity whose value is uncertain. Intuitively, a random variable can be thought of as a quantity whose value is not fixed, but which can take on different values; normally, a probability distribution is used to describe the probability of different values occurring. Random variables are usually real-valued, but one can consider arbitrary types such as boolean values, complex numbers, vectors, matrices, sequences, trees, sets, shapes, manifolds and functions. The term *random element* is used to encompass all such related concepts.

For example: There are two possible outcomes for a coin toss: heads, or tails. The possible outcomes for one fair coin toss can be described using the following random variable:

$$X = \begin{cases} \text{head}, \\ \text{tail}. \end{cases}$$

and if the coin is equally likely to land on either side then it has a probability mass function given by:

$$\rho_X(x) = \begin{cases} \frac{1}{2}, & \text{if } x = \text{head}, \\ \frac{1}{2}, & \text{if } x = \text{tail}. \end{cases}$$

**Example: A simple world consisting of two random variables:**

**Cavity**– a Boolean variable that refers to whether my lower left wisdom tooth has a cavity

**Toothache**- a Boolean variable that refers to whether I have a toothache or not

We use the single capital letters to represent unknown random variables. P induces a probability distribution for any random variables X.

Each Random Variable has a domain of values that it can take it, e. g. domain of *Cavity* is {true, false}

**Atomic Event:**

An **atomic event** is a complete specification of the state of the world about which the agent is uncertain.

**Example:**

In the above world with two random variables (Cavity and Toothache) there are only four distinct atomic events, one being:

Cavity = false, Toothache = true

Which are the other three atomic events?

**Propositions:**

Think of a proposition as the event (set of sample points) where the proposition is true

Given Boolean random variables A and B:

event  $\alpha$  = set of sample points where  $A(\omega) = \text{true}$

event  $\neg a$  = set of sample points where  $A(\omega) = \text{false}$   
 event  $a \wedge b$  = points where  $A(\omega) = \text{true}$  and  $B(\omega) = \text{true}$

Often in AI applications, the sample points are defined by the values of a set of random variables, i.e., the sample space is the Cartesian product of the ranges of the variables.

With Boolean variables, sample point = propositional logic model  
 e.g.,  $A = \text{true}$ ,  $B = \text{false}$ , or  $a \wedge \neg b$ .

Proposition = disjunction of atomic events in which it is true

$$\text{e.g., } (a \vee b) \equiv (\neg a \wedge b) \vee (a \wedge \neg b) \vee (a \wedge b)$$

$$P(a \vee b) = P(\neg a \wedge b) + P(a \wedge \neg b) + P(a \wedge b)$$

### ***Propositional or Boolean random variables***

e.g., *Cavity*(do I have a cavity?)

### ***Discrete random variables (finite or infinite)***

e.g., *Weather* is one of (*sunny*, *rain*, *cloudy*, *snow*)  
*Weather* = *rain* is a proposition

Values must be exhaustive and mutually exclusive

### ***Continuous random variables (bounded or unbounded)***

e.g., *Temp* = 21.6, also allow, e.g., *Temp* < 22.0.

## **Prior Probability:**

The prior or unconditional probability associated with a proposition is the degree of belief accorded to it in the absence of any other information.

### **Example:**

$P(\text{Weather} = \text{sunny}) = 0.72$ ,  $P(\text{Weather} = \text{rain}) = 0.1$ ,  $P(\text{Weather} = \text{cloudy}) = 0.08$ ,  $P(\text{Weather} = \text{snow}) = 0.1$

**Probability distribution** gives values for all possible assignments:

$$P(\text{Weather}) = (0.72, 0.1, 0.08, 0.1)$$

**Joint probability distribution** for a set of random variables gives the probability of every atomic event on those random variables (i.e., every sample point)

$P(\text{Weather, Cavity})$  = a  $4 \times 2$  matrix of values.

Weather=	sunny	rain	cloudy	snow
Cavity=true	0.144	0.02	0.016	0,02
Cavity=false	0.576	0.08	0.064	0.08

Every question about a domain can be answered by the joint distribution because every event is a sum of sample points.

### **Conditional Probability:**

The conditional probability “ $P(a|b)$ ” is the probability of “a” given that all we know is “b”.

Example:  $P(\text{cavity}/\text{toothache}) = 0.8$  means if a patient have toothache and no other information is yet available, then the probability of patient’s having the cavity is 0.8.

Definition of conditional probability:

$$P(a|b) = P(a \wedge b)/P(b) \text{ if } P(b) \neq 0$$

$$\text{i.e. } P(a \wedge b) = P(a|b)P(b)$$

$$\text{Also, } P(b|a) = P(b \wedge a)/P(a) = P(a \wedge b)/P(a)$$

$$\text{i.e. } P(a \wedge b) = P(b|a)P(a)$$

Hence Product rule gives an alternative formulation:

$$P(a \wedge b) = P(a|b)P(b) = P(b|a)P(a)$$

### **Inference using full joint probability distribution:**

We use the full joint distribution as the knowledge base from which answers to all questions may be derived. The probability of a proposition is equal to the sum of the probabilities of the atomic events in which it holds.

$$P(a) = \sum P(e^i)$$

Therefore, given a full joint distribution that specifies the probabilities of all the atomic events, one can compute the probability of any proposition.

### An example

We consider the following domain consisting of three Boolean variables: Toothache, Cavity, and Catch (the dentist's nasty steel probe catches in my tooth).

The full joint distribution is the following 2x2x2 table:

	toothache		¬toothache	
	catch	¬catch	catch	¬catch
cavity	0.108	0.012	0.072	0.008
¬cavity	0.016	0.064	0.144	0.576

The probability of any proposition can be computed from the probabilities in the table. The probabilities in the joint distribution must sum to 1.

Each cell represents an atomic event and these are all the possible atomic events.

$$\begin{aligned}
 P(\text{cavity or toothache}) &= P(\text{cavity, toothache, catch}) + P(\text{cavity, toothache, } \neg\text{catch}) + \\
 &\quad P(\text{cavity, } \neg\text{toothache, catch}) + P(\text{cavity, } \neg\text{toothache, } \neg\text{catch}) + \\
 &\quad P(\neg\text{cavity, toothache, catch}) + P(\neg\text{cavity, toothache, } \neg\text{catch}) \\
 &= 0.108 + 0.012 + 0.072 + 0.008 + 0.016 + 0.064 = 0.28
 \end{aligned}$$

We simply identify those atomic events in which the proposition is true and add up their probabilities

### Marginalization or summing out:

Distribution over **Y** can be obtained by summing out all the other variables from any joint distribution containing **Y**. This process is called marginalization.

$$P(\mathbf{Y}) = \sum P(\mathbf{Y}, \mathbf{z})$$

Examples:

$$P(\text{cavity}) = 0.108 + 0.012 + 0.072 + 0.008 = 0.2$$



$$P(\neg \text{Toothache}) = 0.072 + 0.008 + 0.144 + 0.576 = 0.8$$

$$P(\text{Cavity}, \neg \text{Toothache}) = 0.072 + 0.008 = 0.08$$

### **Calculating Conditional Probability:**

$$P(\neg \text{cavity} \mid \text{Toothache}) = P(\neg \text{cavity} \wedge \text{Toothache}) / P(\text{Toothache})$$

$$= (0.016 + 0.064) / (0.108 + 0.012 + 0.016 + 0.064)$$

$$= 0.4$$

Again let's calculate

$$P(\text{cavity} \mid \text{Toothache}) = P(\text{cavity} \wedge \text{Toothache}) / P(\text{Toothache})$$

$$= (0.108 + 0.012) / (0.108 + 0.012 + 0.016 + 0.064)$$

$$= 0.6$$

Notice that in above two calculations the term  $1 / P(\text{Toothache})$  remain constant no matter which value of cavity is calculated. This constant term is called normalization constant for the distribution  $P(\text{cavity} \mid \text{Toothache})$ , ensuring that it adds up to 1.

### **Independence:**

A and B are independent iff

$$P(A|B) = P(A) \text{ or } P(B|A) = P(B) \text{ or } P(A, B) = P(A)P(B)$$

Example:

$$P(\text{Toothache}, \text{Catch}, \text{Cavity}, \text{Weather}) = P(\text{Toothache}, \text{Catch}, \text{Cavity})P(\text{Weather})$$

Here weather is independent of other three variables.

### **Bayes' Rule (Theorem) :**

$$P(b|a) = \frac{P(a|b) * P(b)}{P(a)}$$

***Proof of bays rule:***

We know that:

$$P(a|b) = P(a \wedge b) / P(b)$$

$$P(a \wedge b) = P(a|b) P(b) \dots \dots \dots (1)$$

Similarly

$$P(b|a) = P(a \wedge b) / P(a)$$

$$P(a \wedge b) = P(b|a) P(a) \dots \dots \dots (2)$$

Equating 1 and 2

$$P(a|b) P(b) = P(b|a) P(a)$$

$$\text{i.e. } P(b|a) = P(a|b) P(b) / P(a)$$

Bayes' rule is useful in practice because there are many cases where we have good probability estimates for three of the four probabilities involved, and therefore can compute the fourth one.

Useful for assessing diagnostic probability from causal probability:

$$P(\text{Cause}|\text{Effect}) = \frac{P(\text{Effect}|\text{Cause})P(\text{Cause})}{P(\text{Effect})}$$

Diagnostic knowledge is often more fragile than causal knowledge.

**Example of Bayes' rule:**

A doctor knows that the disease meningitis causes the patient to have a stiff neck 50% of the time. The doctor also knows that the probability that a patient has meningitis is 1/50,000, and the probability that any patient has a stiff neck is 1/20.

Find the probability that a patient with a stiff neck has meningitis.

Here, we are given;

$$p(s|m) = 0.5$$

$$p(m) = 1/50000$$

$$p(s) = 1/20$$

Now using Bayes' rule;

$$P(m|s) = P(s|m)P(m)/P(s) = (0.5 * 1/50000)/(1/20) = 0.0002$$

### **Fuzzy Vs. Probability**

**Probability theory is a formal examination of the likelihood (chance) that an event will occur, measured in terms of the ratio of the number of expected occurrences to the total number of possible occurrences.** Probabilistic or stochastic methods describe a process in which imprecise or random events affect the values of variables, so that results can be given only in terms of probabilities.

**Fuzziness is a type of deterministic uncertainty. It describes the event class ambiguity. Fuzziness measures the degree to which an event occurs, not whether it occurs. At issue is whether the event class can be unambiguously distinguished from its opposite. Probability arouses from the question whether or not an event occurs. Moreover, it assumes that the event class is crisply defined and that the law of non contradiction holds.**

**Fuzzy logic is logic of vague, imprecise notions and propositions that may be more or less true. Fuzzy logic is then logic of partial degrees of truth. On the contrary, probability deals with the crisp notions and propositions, propositions that are either true or false.** The probability of a proposition is the degree of belief on the truth of that proposition. If we reserve the word uncertainty to refer to degree of belief, then clearly fuzzy logic does not deal with uncertainty at all.

**Formally speaking, fuzzy logic behaves as a many-valued logic, whereas probability theory can be related to a kind of bi-valued logic.**

The inability of Probability Theory to operate on **perception-based information** is a serious limitation because perceptions have a position of centrality in human cognition. Thus, human have a remarkable capability to perform a wide variety of physical and mental tasks without any measurements and any computations. Everyday examples of such tasks are parking a car, driving in city traffic, playing tennis and summarizing a story.

Basically, a natural language is a system for describing perceptions. Perceptions are intrinsically imprecise, reflecting the bounded ability of sensory organs, and ultimately the brain, to resolve detail and store information. The implication of this point is that bivalent-logic-based methods of natural language processing do not have the capability to deal with perception-based information. This is the basis of the conclusion that Probability Theory does not have the capability to operate on perception-based information

**As an illustration, consider a perception-based version of the balls-in-box problem. A box contains about twenty balls of various sizes. Most are large. There are several times as**

**many large balls as small balls. What is the probability that a ball drawn at random is neither large nor small?**

To enable Probability Theory to deal with problems of this kind, it is necessary to restructure probability theory by replacing bivalent logic on which it is based with fuzzy logic. The rationale for this replacement is that fuzzy logic is, in essence, **the logic of perceptions, while bivalent logic is this logic of measurements.**

### **Fuzzy Logic:**

**Fuzzy logic is a form of many-valued logic; it deals with reasoning that is approximate rather than fixed and exact.** Compared to traditional binary sets (where variables may take on true or false values) fuzzy logic variables may have a truth value that ranges in degree between 0 and 1. Fuzzy logic has been extended to handle the concept of partial truth, where the truth value may range between completely true and completely false. Furthermore, when linguistic variables are used, these degrees may be managed by specific functions. Irrationality can be described in terms of what is known as the fuzzjjective

**The term "fuzzy logic" was introduced with the 1965 proposal of fuzzy set theory by Lotfi A. Zadeh. Fuzzy logic has been applied to many fields, from control theory to artificial intelligence. Fuzzy logics however had been studied since the 1920s as infinite-valued logics notably by Łukasiewicz and Tarski.**

While variables in mathematics usually take numerical values, in fuzzy logic applications, the non-numeric *linguistic variables* are often used to facilitate the expression of rules and facts. **A linguistic variable such as *age* may have a value such as *young* or its antonym *old*.** However, the great utility of linguistic variables is that they can be modified via linguistic hedges applied to primary terms. The linguistic hedges can be associated with certain functions.

Fuzzy logic and probability are different ways of expressing uncertainty. While both fuzzy logic and probability theory can be used to represent subjective belief, **fuzzy set theory uses the concept of fuzzy set membership** (i.e., *how much* a variable is in a set), and **probability theory uses the concept of subjective probability or characteristic function** (i.e., *how probable* do I think that a variable is in a set). While this distinction is mostly philosophical, the fuzzy-logic-derived possibility measure is inherently different from the probability measure, hence they are not *directly* equivalent.

### **Linguistic Variables:**

By a *linguistic variable* we mean a variable whose values are words or sentences in a natural or artificial language. A numerical variable takes numerical values: Age = 65. While, a linguistic variable takes linguistic values: Age is old

**A linguistic value is a fuzzy set. It has a LABEL and a MEANING. The LABEL is a Symbol, Sentence in a Language while MEANING is the Fuzzy Subset of a Universe of Discourse**

All linguistic values form a term set:

$T(\text{age}) = \{\text{young, not young, very young, ...middle aged, not middle aged, ...old, not old, very old, more or less old, ...not very young and not very old, ...}\}$

### **Why Fuzzy Logic?**

More specifically, fuzzy logic may be viewed as an attempt at formalization/mechanization of two remarkable human capabilities. First, the capability to converse, reason and make rational decisions in an environment of imprecision, uncertainty, incompleteness of information, conflicting information, partiality of truth and partiality of possibility – in short, in an environment of imperfect information. And second, the capability to perform a wide variety of physical and mental tasks without any measurements and any computations.

### **Usages in Real world**

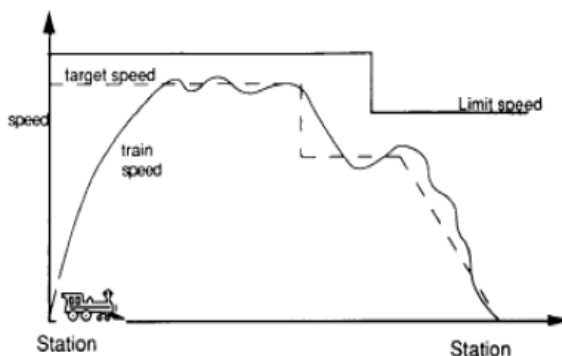
Fuzzy systems were initially implemented in Japan.

- Interest in fuzzy systems was sparked by Seiji Yasunobu and Soji Miyamoto of Hitachi, who in 1985 provided simulations that demonstrated the feasibility of fuzzy control systems for the Sendai railway. Their ideas were adopted, and fuzzy systems were used to control accelerating, braking, and stopping when the line opened in 1987. The Sendai Subway has two operation modes, train speed regulation control and train stopping control. These operate on two separate but communication fuzzy logic controllers and take a number of calculations into account. These include acceleration to a target speed with velocity oscillation. Maintaining target speed (coasting), and stopping at the correct location with velocity control for a no motion feel. Other factors taken into account are safety, energy consumption, traceability, and passenger comfort.

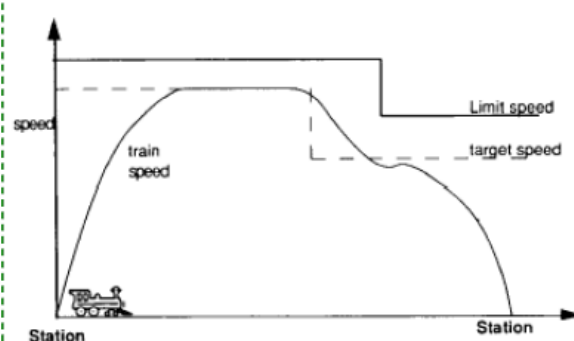
- In 1987, Takeshi Yamakawa demonstrated the use of fuzzy control, through a set of simple dedicated fuzzy logic chips, in an "inverted pendulum" experiment. This is a classic control problem, in which a vehicle tries to keep a pole mounted on its top by a hinge upright by moving back and forth. Yamakawa subsequently made the demonstration more sophisticated by mounting a wine glass containing water and even a live mouse to the top of the pendulum: the system maintained stability in both cases. Yamakawa eventually went on to organize his own fuzzy-systems research lab to help exploit his patents in the field.
- Japanese engineers subsequently developed a wide range of fuzzy systems for both industrial and consumer applications. In 1988 Japan established the Laboratory for International Fuzzy Engineering (LIFE), a cooperative arrangement between 48 companies to pursue fuzzy research. The automotive company Volkswagen was the only foreign corporate member of LIFE, dispatching a researcher for a duration of three years.
- Japanese consumer goods often incorporate fuzzy systems. Matsushita vacuum cleaners use microcontrollers running fuzzy algorithms to interrogate dust sensors and adjust suction power accordingly. Hitachi washing machines use fuzzy controllers to load-weight, fabric-mix, and dirt sensors and automatically set the wash cycle for the best use of power, water, and detergent.
- Canon developed an autofocus camera that uses a charge-coupled device (CCD) to measure the clarity of the image in six regions of its field of view and use the information provided to determine if the image is in focus. It also tracks the rate of change of lens movement during focusing, and controls its speed to prevent overshoot.
- It allowed the camera to consider three possible targets left, right and center ; and focus on most likely. There are series of rules in microcontroller of camera, such as, when left is nearest, the likelihood that the subject is left is high. The camera could then assess how the sensor readings satisfied the rules and choose the target.
- The camera's fuzzy control system uses 12 inputs: 6 to obtain the current clarity data provided by the CCD and 6 to measure the rate of change of lens movement. The output is the position of the lens. The fuzzy control system uses 13 rules and requires 1.1 kilobytes of memory.
- An industrial air conditioner designed by Mitsubishi uses 25 heating rules and 25 cooling rules. A temperature sensor provides input, with control outputs fed to an [inverter](#), a compressor valve, and a fan motor. Compared to the previous design, the fuzzy controller heats and cools five times faster, reduces power consumption by 24%, increases temperature stability by a factor of two, and uses fewer sensors.

- Other applications investigated or implemented include: character and handwriting recognition; optical fuzzy systems; robots, including one for making Japanese flower arrangements; voice-controlled robot helicopters (hovering is a "balancing act" rather similar to the inverted pendulum problem); control of flow of powders in film manufacture; elevator systems; and so on.

<b>Hitachi</b>	Automatic control for subway trains in Sendai yielding smoother, more efficient ride with higher stopping precision
<b>Nippon Electric</b>	Controlling temperatures in glass fusion at the Notogawa and Takatsuki factories
<b>Nissan</b>	Anti-skid brake systems and automatic transmissions for cars
<b>Canon, Minolta, Ricoh</b>	Auto-focusing in cameras by choosing the right subject in the picture frame to focus
<b>Panasonic</b>	Jitter removal in video camcoders by distinguishing between jitters and actual movement of subjects
<b>Matshushita, Toshiba, Sanyo, Hitachi</b>	Vacuum cleaners that use sensors to gather information about floor and dirt conditions and then use a fuzzy expert system to choose the right program
<b>Yamaichi</b>	Computerized trading programs which mimic the approximate reasoning processes of experienced fund managers
<b>Matshushita</b>	A/C that make judgments based on factors such as number of persons in room and optimum degree of comfort



ATO by conventional control



ATO with fuzzy control



## Extension principle

Deduction in fuzzy logic is governed by a collection of rules of deduction which, in the main, are rules that govern propagation and counter-propagation of generalized constraints. The principal rule is the *extension principle*. Extension principle has many versions. The simplest version (Zadeh 1965) is the following. Let  $f$  be a function from reals to reals,  $Y=f(X)$ . What we know is that  $X$  is  $A$ , where  $A$  is a fuzzy subset of the real line. Equivalently, what we know about  $X$  is its granular value, that is, its possibility distribution,  $A$ . What can be said about  $Y$ , that is, what is its granular value or, equivalently, its possibility distribution? In a more general form, (Zadeh 1975)  $X$  is  $A$  is replaced by  $f(X)$  is  $A$ . It is this form that is used in most practical applications. In a form that is used in fuzzy control, what is granulated is  $f$ , resulting in a granular function,  $f^*$ , which is defined by a collection of fuzzy-if-then rules. A simple example is  $f^*$ :

if  $X$  is small then  $Y$  is small

if  $X$  is medium then  $Y$  is large

if  $X$  is large then  $Y$  is small

More generally, the extension principle may be viewed as follows. Let  $Y=f(X)$ , where  $X$  is a real-valued variable. Assume that we can compute  $Y$  for singular values of  $f$  and  $X$ . Basically, the extension principle serves to extend the definition of  $Y$  to granular values of  $f$  and  $X$ .

## History

During much of its early history, fuzzy logic has been an object of skepticism and derision, in part because fuzzy is a word which is usually used in a pejorative sense. Today, fuzzy logic has an extensive literature and a wide variety of applications ranging from consumer products and fuzzy control to medical diagnostic systems and fraud detection (Zadeh 1990; Novak and Perfilieva 2000).

Existing scientific theories are almost without exception based on classical, bivalent logic. What is widely unrecognized is that many scientific theories can be enriched through addition of concepts and techniques drawn from fuzzy logic. In particular, fuzzy logic can add to existing theories NL-capability, that is, the capability to operate on information described in natural language or, equivalently, on perception-based information. In coming years, the issue of NL-capability is likely to grow in visibility and importance, especially in such fields as economics, law, medicine, search, question-answering and, above all, probability theory and decision analysis.

## First Applications

In the early 1970s, the first industrial applications of fuzzy logic were demonstrated in Europe. Among these are:

- control of a steam generator by Ebrahim Mamdani of Queen Mary College in London, England, which he cannot control using conventional techniques
- decision support systems created by Hans Zimmermann of RWTH University of Aachen, Germany
- control of a cement kiln

However, fuzzy logic could not get the broad acceptance of the industry even after the initial applications were introduced. It only gained momentum in decision support and data analysis applications in Europe in the 1980s. Most developments were triggered by empirical research on how well it can model the human decision and evaluation process.

Inspired by the first fuzzy logic applications in Europe, some Japanese companies started applying fuzzy logic in control engineering in the 1980s. But the poor computational performance of the first fuzzy logic algorithms on standard hardware led them to create dedicated fuzzy logic hardware. Some of the first fuzzy logic applications in Japan were:

- a water treatment plant by Fuji Electric in 1983
- a subway system by Hitachi which opened in 1987

## Truth or Falsity

The basic idea underlying all these approaches is that of an intrinsic dichotomy between true and false. This opposition implies the validity of two fundamental laws of classical logic:

- **Principle of excluded middle:** Every proposition is true or false, and there is another possibility.

- **Principle of non-contradiction:** No statement is true and false simultaneously.