Fundamentals of Neural Networks

Chapter 2

In this chapter, we introduce the fundamental concepts of Neural Networks (NN). The biological neuron system, which has been the chief source of inspiration in much of the research work in Neural Networks and the model of an artificial neuron, are first elaborated. Neural Network architectures, their characteristics and learning methods are next discussed. A brief history of Neural Network research and some of the early NN architectures are presented. Finally, some of the application domains where NN architectures have made an impact are listed.

2.1 BASIC CONCEPTS OF NEURAL NETWORKS

Neural Networks, which are simplified models of the biological neuron system, is a massively parallel distributed processing system made up of highly interconnected neural computing elements that have the ability to learn and thereby acquire knowledge and make it available for use.

Various learning mechanisms exist to enable the NN acquire knowledge. NN architectures have been classified into various types based on their learning mechanisms and other features. Some classes of NN refer to this learning process as *training* and the ability to solve a problem using the knowledge acquired as *inference*.

NNs are simplified imitations of the central nervous system, and obviously therefore, have been motivated by the kind of computing performed by the human brain. The structural constituents of a human brain termed *neurons* are the entities, which perform computations such as cognition, logical inference, pattern recognition and so on. Hence the technology, which has been built on a simplified imitation of computing by neurons of a brain, has been termed *Artificial Neural Systems* (ANS) technology or *Artificial Neural Networks* (ANN) or simply *Neural Networks*. In the literature, this technology is also referred to as *Connectionist Networks, Neuro-Computers, Parallel Distributed Processors* etc. Also neurons are referred to as *neurodes, Processing Elements (PEs)*, and *nodes*. In this book, we shall use the terms *Neural Networks* or *Artificial Neural Networks* and *neurons*.

A human brain develops with time and this, in common parlance is known as *experience*. Technically, this involves the 'development' of neurons to adapt themselves to their surrounding environment, thus, rendering the brain *plastic* in its information processing capability. On similar lines, the property of plasticity is also discussed with respect to NN architectures. Further, we are also interested in the *stability* of an NN system; i.e. the adaptive capability of an NN in the face of changing environments. Thus, the *stability–plasticity* issue is of great importance to NN

$$\phi(I) = \frac{1}{1 + e^{-\alpha I}}$$

where, α is the slope parameter, which adjusts the abruptness of the function as it changes between the two asymptotic values. Sigmoidal functions are differentiable, which is an important feature of NN theory. Figure 2.6 illustrates the sigmoidal function.

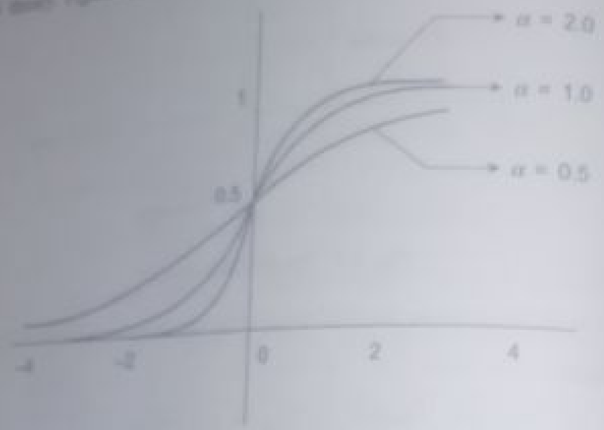


Fig. 2.6 Sigmoidal function.

Hyperbolic tangent function

The function is given by

$$\phi(I) = \tanh(I)$$

and can produce negative output values.

The first mathematical model of a biological neuron was presented by McCulloch and Pitts (1943). The model, known as McCulloch-Pitts model does not exhibit any learning but just serves as a basic building block which has inspired further significant work in NN research. The model makes use of a bias term whose weight is w_0 but with a fixed input of $x_0 = 1$. This is besides the other inputs x_i and weights w_i. The bias is an external parameter for the artificial neuron but serves the purpose of increasing or decreasing the net input of the activation function depending on whether it is positive or negative. Other modelling schemes for the neuron have also been proposed (Macgregor, 1987).

2.4 NEURAL NETWORK ARCHITECTURES

An Artificial Neural Network is defined as a data processing system consisting of a large number of simple highly interconnected processing elements (artificial neurons) in an architecture inspired by the structure of the cerebral cortex of the brain (Tsoukalas and Uhrig, 1997). Generally, an ANN structure can be represented using a directed graph. A graph G is an ordered 2-tuple (V, E) consisting of a set V of vertices and a set E of edges. When each edge is assigned an orientation,

the graph is directed and is called a *directed graph* or a *digraph*. Figure 2.7 illustrates a digraph. Digraphs assume significance in Neural Network theory since signals in NN systems are restricted to flow in specific directions.

The vertices of the graph may represent neurons (input/output) and the edges, the synaptic links. The edges are labelled by the weights attached to the synaptic links.

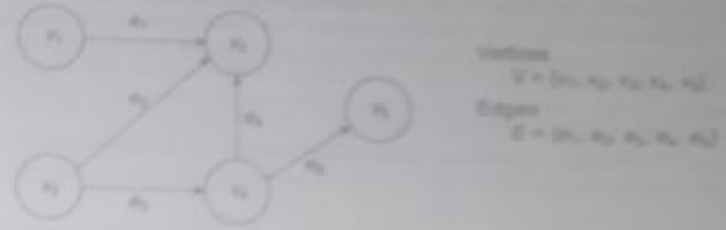


Fig. 2.7 An example digraph.

There are several classes of NN, classified according to their learning mechanisms. However, we identify three fundamentally different classes of Networks. All the three classes employ the digraph structure for their representation.

2.4.1 Single Layer Feedforward Network

This type of network comprises of two layers, namely the *input layer* and the *output layer*. The *input layer neurons* receive the input signals and the *output layer neurons* receive the output signals. The synaptic links carrying the weights connect every input neuron to the output neuron but not vice-versa. Such a network is said to be *feedforward* in type or acyclic in nature. Despite the two layers, the network is termed single layer since it is the output layer, alone which performs computation. The input layer merely transmits the signals to the output layer. Hence, the name *single layer feedforward network*. Figure 2.8 illustrates an example network.

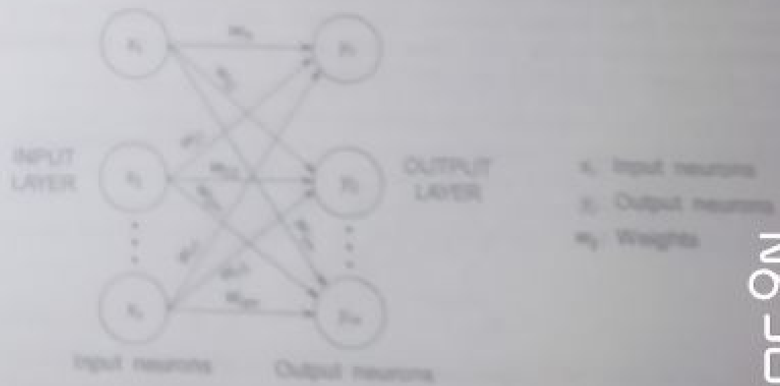


Fig. 2.8 Single layer feedforward network.

2.4.2 Multilayer Feedforward Network

This network, as its name indicates is made up of multiple layers. Thus, architectures of this class besides possessing an input and an output layer also have one or more intermediary layers called hidden layers. The computational units of the hidden layer are known as the hidden neurons or hidden units. The hidden layer aids in performing useful intermediary computations before directing the input to the output layer. The input layer neurons are linked to the hidden layer neurons and the weights on these links are referred to as *input-hidden layer weights*. Again, the hidden layer neurons are linked to the output layer neurons and the corresponding weights are referred to as *hidden-output layer weights*. A multilayer feedforward network with l input neurons, m_1 neurons in the first hidden layer, m_2 neurons in the second hidden layer and n output neurons in the output layer is written as $l - m_1 - m_2 - n$.

Figure 2.9 illustrates a multilayer feedforward network with a configuration $l - m - n$.

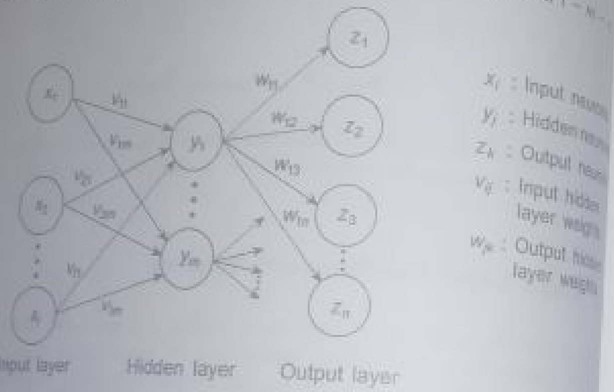


Fig. 2.9 A multilayer feedforward network ($l - m - n$ configuration).

2.4.3 Recurrent Networks

These networks differ from feedforward network architectures in the sense that there is atleast one feedback loop. Thus, in these networks, for example, there could exist one layer with feedback connections as shown in Fig. 2.10. There could also be neurons with self-feedback links, i.e. the output of a neuron is fed back into itself as input.

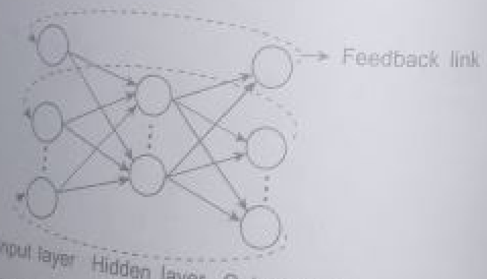


Fig. 2.10 A recurrent neural network.

2.5 CHARACTERISTICS OF NEURAL NETWORKS

(i) The NNs exhibit mapping capabilities, that is, they can map input patterns to their associated output patterns.

(ii) The NNs learn by examples. Thus, NN architectures can be 'trained' with known examples of a problem before they are tested for their 'inference' capability on unknown instances of the problem. They can, therefore, identify new objects previously untrained.

(iii) The NNs possess the capability to generalize. Thus, they can predict new outcomes from past trends.

(iv) The NNs are robust systems and are fault tolerant. They can, therefore, recall full patterns from incomplete, partial or noisy patterns.

(v) The NNs can process information in parallel, at high speed, and in a distributed manner.

2.6 LEARNING METHODS

Learning methods in Neural Networks can be broadly classified into three basic types: *supervised, unsupervised,* and *reinforced.*

Supervised learning

In this, every input pattern that is used to train the network is associated with an output pattern, which is the target or the desired pattern. A teacher is assumed to be present during the learning process, when a comparison is made between the network's computed output and the correct expected output, to determine the error. The error can then be used to change network parameters, which result in an improvement in performance.

Unsupervised learning

In this learning method, the target output is not presented to the network. It is as if there is no teacher to present the desired patterns and hence, the system learns of its own by discovering and adapting to structural features in the input patterns.

Reinforced learning

In this method, a teacher though available, does not present the expected answer but only indicates if the computed output is correct or incorrect. The information provided helps the network in its learning process. A reward is given for a correct answer computed and a penalty for a wrong answer. But, reinforced learning is not one of the popular forms of learning.

Supervised and unsupervised learning methods, which are most popular forms of learning, have found expression through various rules. Some of the widely used rules have been presented below:

Hebbian learning

This rule was proposed by Hebb (1949) and is based on correlative weight adjustment. This is the oldest learning mechanism inspired by biology.

In this, the input-output pattern pairs (X_i, Y_i) are associated by the weight matrix W, k... as the correlation matrix. It is computed as

$$W = \sum_{i=1}^{n} X_i Y_i^T$$

Here, Y_i^T is the transpose of the associated output vector Y_i. Numerous variants of the ... have been proposed (Anderson, 1983; Kosko, 1985; Lippman, 1987; Linsker, 1988).

Gradient descent learning

This is based on the minimization of error E defined in terms of weights and the activa... function of the network. Also, it is required that the activation function employed by the netw... is differentiable, as the weight update is dependent on the gradient of the error E.

Thus, if ΔW_{ij} is the weight update of the link connecting the ith and jth neuron of the... neighbouring layers, then ΔW_{ij} is defined as

$$\Delta W_{ij} = \eta \frac{\partial E}{\partial W_{ij}}$$

where, η is the learning rate parameter and $\partial E/\partial W_{ij}$ is the error gradient with reference to... weight W_{ij}.

The Widrow and Hoffs Delta rule and Backpropagation learning rule are all examples of ... type of learning mechanism.

Competitive learning

In this method, those neurons which respond strongly to input stimuli have their weights upd... When an input pattern is presented, all neurons in the layer compete and the winning ne... undergoes weight adjustment. Hence, it is a "winner-takes-all" strategy.

Stochastic learning

In this method, weights are adjusted in a probabilistic fashion. An example is evident in simul... annealing—the learning mechanism employed by Boltzmann and Cauchy machines, which a... kind of NN systems.

Figure 2.11 illustrates the classification of learning algorithms.

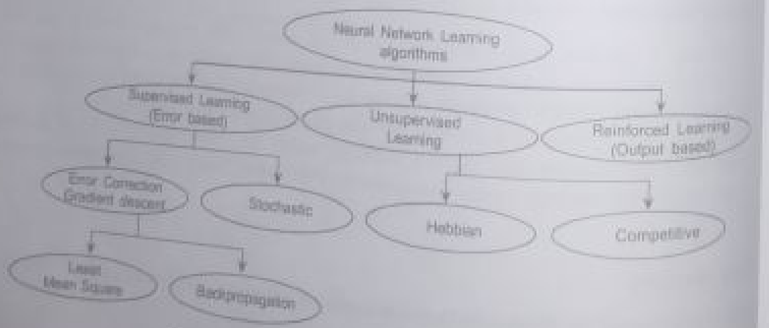


Fig. 2.11 Classification of learning algorithms.

2.7 TAXONOMY OF NEURAL NETWORK ARCHITECTURES

Over the years, several NN systems have evolved. The following are some of the systems that have acquired significance.

- ADALINE (Adaptive Linear Neural Element)
- ART (Adaptive Resonance Theory)
- AM (Associative Memory)
- BAM (Bidirectional Associative Memory)
- Boltzmann Machine
- BSB (Brain-State-in-a-Box)
- CCN (Cascade Correlation)
- Cauchy Machine
- CPN (Counter Propagation Network)
- Hamming Network
- Hopfield Network
- LVQ (Learning Vector Quantization)
- MADALINE (Many ADALINE)
- MLFF (Multilayer Feedforward Network)
- Neocognitron
- Perceptron
- RBF (Radial Basis Function)
- RNN (Recurrent Neural Network)
- SOFM (Self-organizing Feature Map)

Table 2.1 shows the classification of the NN systems listed above, according to their learning methods and architectural types.

Table 2.1 The classification of some NN systems with respect to learning methods and architecture types

TYPE OF ARCHITECTURE	LEARNING METHOD: *Gradient descent*	*Hebbian*	*Competitive*	*Stochastic*
Single-layer feedforward	ADALINE Hopfield Perceptron	AM Hopfield	LVQ SOFM	—
Multilayer feedforward	CCN MLFF RBF	Neocognitron	—	—
Recurrent neural network	RNN	BAM BSB Hopfield	ART	Boltzmann machine Cauchy machine

2.8 HISTORY OF NEURAL NETWORK RESEARCH

The pioneering work of McCulloch and Pitts (1943) was the foundation stone for the gro[wth of] NN architectures. In their paper, McCulloch and Pitts suggested the unification of ne[uro] physiology with mathematical logic, which paved way for some significant results in NN re[search]. Infact, the McCulloch-Pitts model even influenced Von Neumann to try new design techno[logies in] the construction of EDVAC (Electronic Discrete Variable Automatic Computer).

The next significant development arose out of Hebb's book '*The organization of behaviour*'. In this, Hebb proposed a learning rule derived from a model based on syn[aptic] connections between nerve cells responsible for biological associative memory.

The Hebbian rule was later refined by Rosenblatt in 1958, in the Perceptron m[odel] (Rosenblatt, 1958). However, a critical assessment of the Perceptron model by Minsky in 1[969] (Minsky and Papert, 1969) stalled further research in NN. It was much later in the 1980s that [there] was a resurgence of interest in NN and many major contributions in the theory and application of [NN] were made.

The only important contribution made in the 1970's was the Self Organizing Map Archi[tecture] based on Competitive learning (Will Shaw and Von der Malsburg, 1976). Some of the well kn[own] architectures which turned out to be milestones in NN research have been listed in Table 2.2.

Table 2.2 Some milestones in NN research

Name of the neural network	Developer	Year (development and growth)	Remarks
Adaptive Resonance Theory (ART) networks	• Carpenter, Grossberg and others	1980 and onwards	The networks employ a new principle of self organization called Adaptive Resonance Theory based on Competitive learning. The general complexity of the network structures is a li[mitation].
Backpropagation networks	• Rumelhart, Hinton, Williams • Werbos • Parker	1985 1974 1985	The Backpropagation learning rule is applica[ble] on any feedforward network architecture. Slow rate of convergence and local minima problem are its weaknesses.
Bidirectional Associative Memory (BAM)	• Bart Kosko	1988	These are two--layer recurrent, hetero associative networks that can store pattern pairs and ret[rieve] them. They behave as content addressable mem[ories].
Boltzmann and Cauchy machines	• Hinton, Sejnowski • Szu H., E Hartley	1983, 1985 1986 1987	These are stochastic networks whose states are governed by the Boltzmann distribution/Cauchy distribution. The heavy computational load is a drawback.
Brain-state-in-a-box	• James Anderson	1977	A recurrent auto associative network which ma[kes] use of Hebbian/Gradient descent learning.
Counter propagation network	• Robert Hecht Nielsen	1987	The network belongs to the category of self-organization networks and functions as stati[stically] optimal self-programming look up table. The w[eight] adjustments between the layers follow Koho[nen's] unsupervised learning rule and Grossberg's supervised learning rule.
Hopfield network	• John Hopfield	1982	Single layer recurrent network which makes use of Hebbian learning or Gradient Descent lea[rning].

Table 2.2 Some milestones in NN research (cont.)

Name of the neural network	Developer	Year (development and growth)	Remarks
MADALINE network	• Bernard Widrow	1960 1988	It is created by a combination of ADALINE networks spread across multiple layers with adjustable weights. The network employs a supervised learning rule called MADALINE adaptation Rule (MR) based on 'minimal disturbance principle'.
Neocognitron	• Kunihiko Fukushima	1982	A hybrid hierarchical multilayer feedback/forward network which closely models a human vision system, the network employs either supervised or unsupervised learning rules.
Perceptron	• Frank Rosenblatt	1958	A single layer or multilayer feedforward network best understood and extensively studied. However, the network is able to obtain weights, only for linearly separable tasks.
Self-organizing Feature Map networks	• Kohonen	1982	The network is a simplified model of the feature-to-localized-region mapping of a brain. It is a self-organizing network employing competitive learning.

2.9 EARLY NEURAL NETWORK ARCHITECTURES

2.9.1 Rosenblatt's Perceptron

The perceptron is a computational model of the retina of the eye and hence, is named 'perceptron'. The network comprises three units, the *Sensory unit S*, *Association unit A*, and *Response unit R* (refer Fig. 2.12).

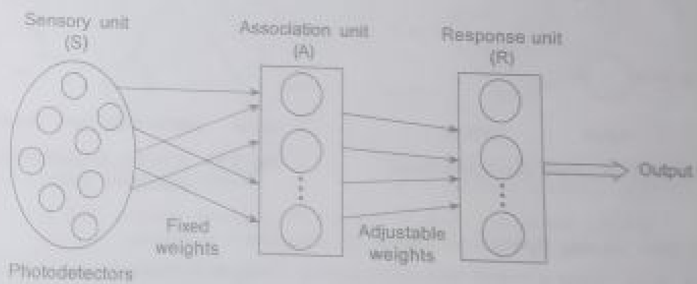


Fig. 2.12 Rosenblatt's original perceptron model.

The S unit comprising 400 photodetectors receives input images and provides a 0/1 electric signal as output. If the input signals exceed a threshold, then the photodetector outputs 1 else 0.

Chapter 6

Fuzzy Set Theory

Problems in the real world quite often turn out to be complex owing to an element of uncertainty either in the parameters which define the problem or in the situations in which the problem occurs.

Although probability theory has been an age old and effective tool to handle uncertainty, it can be applied only to situations whose characteristics are based on random processes, that is, processes in which the occurrence of events is strictly determined by chance. However, in reality, there turn out to be problems, a large class of them whose uncertainty is characterized by a nonrandom process. Here, the uncertainty may arise due to partial information about the problem, or due to information which is not fully reliable, or due to inherent imprecision in the language with which the problem is defined, or due to receipt of information from more than one source about the problem which is conflicting.

It is in such situations that *fuzzy set theory* exhibits immense potential for effective solving of the uncertainty in the problem. *Fuzziness* means 'vagueness'. Fuzzy set theory is an excellent mathematical tool to handle the uncertainty arising due to vagueness. Understanding human speech and recognizing handwritten characters are some common instances where fuzziness manifests.

It was Lotfi A. Zadeh who propounded the fuzzy set theory in his seminal paper (Zadeh, 1965). Since then, a lot of theoretical developments have taken place in this field. It is however, the Japanese who seem to have fully exploited the potential of fuzzy sets by commercializing the technology. More than 2000 patents have been acquired by the Japanese in the application of the technique and the area spans a wide spectrum, from consumer products and electronic instruments to automobile and traffic monitoring systems.

6.1 FUZZY VERSUS CRISP

Consider the query, "Is water colourless?" The answer to this is a definite *Yes/True*, or *No/False*, as warranted by the situation. If "yes"/"true" is accorded a value of 1 and "no"/"false" is accorded a value of 0, this statement results in a 0/1 type of situation. Such a logic which demands a binary (0/1) type of handling is termed *crisp* in the domain of fuzzy set theory. Thus, statements such as "Temperature is 32°C", "The running time of the program is 4 seconds" are examples of crisp situations.

On the other hand, consider the statement, "Is Ram honest?" The answer to this query need not be a definite "yes" or "no". Considering the degree to which one knows Ram, a variety of answers spanning a range, such as "extremely honest", "extremely dishonest", "honest at ... "very honest" could be generated. If, for instance, "extremely honest" were to be accorded ... of 1, at the high end of the spectrum of values, "extremely dishonest" a value of 0 at the ... of the spectrum, then, "honest at times" and "very honest" could be assigned values of ... 0.85 respectively. The situation is therefore so fluid that it can accept values between 0 and ... contrast to the earlier one which was either a 0 or 1. Such a situation is termed *fuzzy*. Fig... shows a simple diagram to illustrate fuzzy and crisp situations.

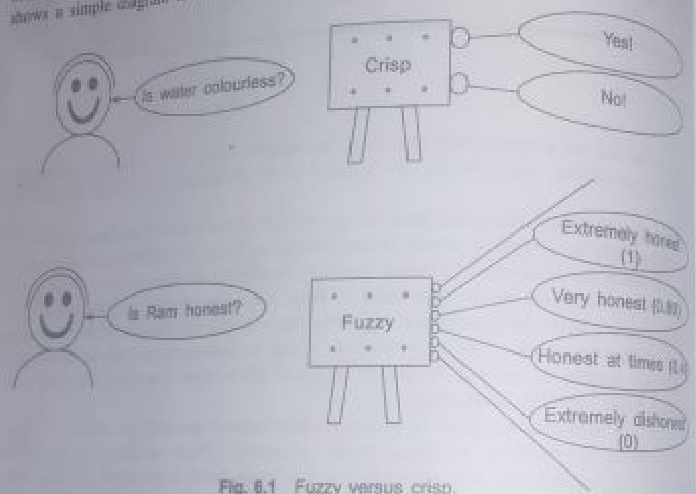


Fig. 6.1 Fuzzy versus crisp.

Classical set theory also termed *crisp set theory* and propounded by George Cantor ... fundamental to the study of fuzzy sets. Just as Boolean logic had its roots in the theory of ... sets, fuzzy logic has its roots in the theory of fuzzy sets (refer Fig. 6.1).

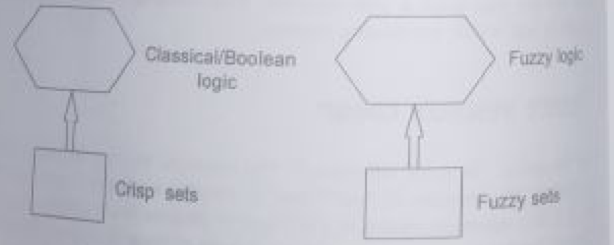


Fig. 6.2 Crisp sets and fuzzy sets.

We now briefly review crisp sets and its operations before a discussion on fuzzy sets ... undertaken.

6.2 CRISP SETS

Universe of discourse

The *universe of discourse* or *universal set* is the set which, with reference to a particular context, contains all possible elements having the same characteristics and from which sets can be formed. The universal set is denoted by E.

Example

(i) The universal set of all numbers in Euclidean space.

(ii) The universal set of all students in a university.

Set

A set is a *well defined* collection of objects. Here, well defined means the object either belongs to or does not belong to the set (observe the "crispness" in the definition).

A set in certain contexts may be associated with its universal set from which it is derived.

Given a set A whose objects are $a_1, a_2, a_3, \ldots, a_n$, we write A as $A = \{a_1, a_2, \ldots, a_n\}$. Here, $a_1, a_2, \ldots, a_n$ are called the *members* of the set. Such a form of representing a set is known as *list form*.

Example

$A = \{\text{Gandhi, Bose, Nehru}\}$

$B = \{\text{Swan, Peacock, Dove}\}$

A set may also be defined based on the properties the members have to satisfy. In such a case, a set A is defined as

$$A = \{x \mid P(x)\} \tag{6.1}$$

Here, $P(x)$ stands for the property P to be satisfied by the member x. This is read as 'A is the set of all X such that $P(x)$ is satisfied'.

Example

$A = \{x \mid x \text{ is an odd number}\}$

$B = \{y \mid y > 0 \text{ and } y \bmod 5 = 0\}$

Venn diagram

Venn diagrams are pictorial representations to denote a set. Given a set A defined over a universal set E, the Venn diagram for A and E is as shown in Fig. 6.3.

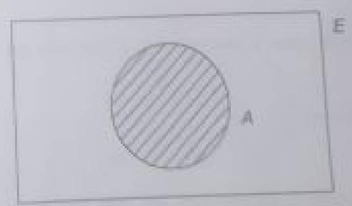


Fig. 6.3 Venn diagram of a set A.

Example

In Fig. 6.3, if E represents the set of university students then A may represent the set of female students.

Membership

An element x is said to be a member of a set A if x belongs to the set A. The membership is indicated by '$\in$' and is pronounced "belongs to". Thus, $x \in A$ means x belongs to A and $x \notin A$ means x does not belong to A.

Example

Given $A = \{4, 5, 6, 7, 8, 10\}$, for $x = 3$ and $y = 4$, we have $x \notin A$ and $y \in A$.

Here, observe that each element either belongs to or does not belong to a set. The concept of membership is definite and therefore crisp (1—belongs to, 0—does not belong to). In contrast, as we shall see later, a fuzzy set accommodates membership values which are not only 0 or 1 but anything between 0 and 1.

Cardinality

The number of elements in a set is called its cardinality. Cardinality of a set A is denoted as $n(A)$ or $|A|$ or $\#A$.

Example

If $A = \{4, 5, 6, 7\}$ then $|A| = 4$

Family of sets

A set whose members are sets themselves, is referred to as a family of sets.

Example

$A = \{\{1, 3, 5\}, \{2, 4, 6\}, \{5, 10\}\}$ is a set whose members are the sets $\{1, 3, 5\}$, $\{2, 4, 6\}$, and $\{5, 10\}$.

Null Set/Empty Set

A set is said to be a *null set* or *empty set* if it has no members. A null set is indicated as $\emptyset$ or $\{\}$ and indicates an impossible event. Also, $|\emptyset| = 0$.

Example

The set of all prime ministers who are below 15 years of age.

Singleton Set

A set with a single element is called a *singleton set*. A singleton set has cardinality of 1.

Example

If $A = \{a\}$, then $|A| = 1$

Subset

Given sets A and B defined over E the universal set, A is said to be a *subset* of B if A is fully contained in B, that is, every element of A is in B.

Denoted as $A \subset B$, we say that A is a subset of B, or A is a proper subset of B. On the other hand, if A is contained in or equivalent to that of B then we denote the subset relation as $A \subseteq B$. In such a case, A is called the *improper subset* of B.

Superset

Given sets A and B on E the universal set, A is said to be a superset of B if every element of B is contained in A.

Denoted as $A \supset B$, we say A is a superset of B or A contains B. If A contains B and is equivalent to B, then we denote it as $A \supseteq B$.

Example

Let $A = \{3, 4\}$, $B = \{3, 4, 5\}$ and $C = \{4, 5, 3\}$

Here, $A \subset B$, and $B \supset A$

$C \subseteq B$, and $B \supseteq C$

Power set

A *power set* of a set A is the set of all possible subsets that are derivable from A including null set.

A power set is indicated as $P(A)$ and has cardinality of $|P(A)| = 2^{|A|}$

Example

Let $A = \{3, 4, 6, 7\}$

$P(A) = \{\{3\}, \{4\}, \{6\}, \{7\}, \{3, 4\}, \{4, 6\}, \{6, 7\}, \{3, 7\}, \{3, 6\}, \{4, 7\},$
$\{3, 4, 6\}, \{4, 6, 7\}, \{3, 6, 7\}, \{3, 4, 7\}, \{3, 4, 6, 7\}, \emptyset\}$

Here, $|A| = 4$ and $|P(A)| = 2^4 = 16$.

6.2.1 Operations on Crisp Sets

Union ($\cup$)

The union of two sets A and B ($A \cup B$) is the set of all elements that belong to A or B or both.

$$A \cup B = \{x \mid x \in A \text{ or } x \in B\} \tag{6.2}$$

Example

Given $A = \{a, b, c, 1, 2\}$ and $B = \{1, 2, 3, a, c\}$, we get $A \cup B = \{a, b, c, 1, 2, 3\}$

Figure 6.4 illustrates the Venn diagram representation for $A \cup B$

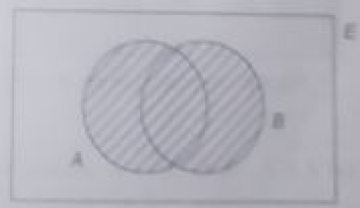


Fig. 6.4 Venn diagram for $A \cup B$.

Intersection ($\cap$)

The intersection of two sets A and B ($A \cap B$) is the set of all elements that belong to A and B

$$A \cap B = \{x / x \in A \text{ and } x \in B\} \quad (6.3)$$

Any two sets which have $A \cap B = \varnothing$ are called *Disjoint Sets*.

Example

Given $A = \{a, b, c, 1, 2\}$ and $B = \{1, 2, 3, a, c\}$, we get $A \cap B = \{a, c, 1, 2\}$

Figure 6.5 illustrates the Venn diagram for $A \cap B$

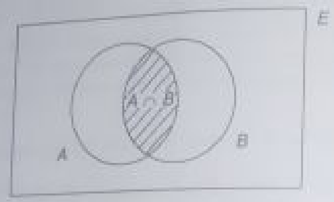


Fig. 6.5 Venn diagram for $A \cap B$.

Complement (^{c})

The complement of a set A ($\bar{A} | A^c$) is the set of all elements which are in E but not in A.

$$A^c = \{x / x \notin A,\ x \in E\} \quad (6.4)$$

Example

Given $X = \{1, 2, 3, 4, 5, 6, 7\}$ and $A = \{5, 4, 3\}$, we get $A^c = \{1, 2, 6, 7\}$

Figure 6.6 illustrates the Venn diagram for A^c.

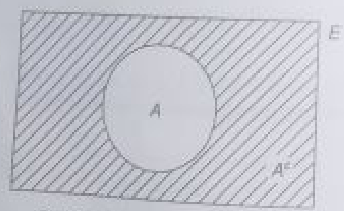


Fig. 6.6 Venn diagram for A^c

Difference ($-$)

The difference of the set A and B is $A - B$, the set of all elements which are in A but not in B.

$$A - B = \{x \mid x \in A \text{ and } x \notin B\} \quad (6.5)$$

Example

Given $A = \{a, b, c, d, e\}$ and $B = \{b, d\}$, we get $A - B = \{a, c, e\}$

Figure 6.7 illustrates the Venn diagram for $A - B$.

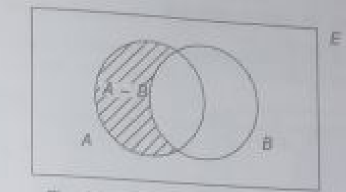


Fig. 6.7 Venn diagram for $A - B$.

6.2.2 Properties of Crisp Sets

The following properties of sets are important for further manipulation of sets.

Commutativity:
$$A \cup B = B \cup A$$
$$A \cap B = B \cap A \quad (6.6)$$

Associativity:
$$(A \cup B) \cup C = A \cup (B \cup C)$$
$$(A \cap B) \cap C = A \cap (B \cap C) \quad (6.7)$$

Distributivity:
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$
$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \quad (6.8)$$

Idempotence:
$$A \cup A = A$$
$$A \cap A = A \quad (6.9)$$

Identity:
$$A \cup \varnothing = A$$
$$A \cap E = A$$
$$A \cap \varnothing = \varnothing$$
$$A \cup E = E \quad (6.10)$$

Law of Absorption:
$$A \cup (A \cap B) = A$$
$$A \cap (A \cup B) = A \quad (6.11)$$

Transitivity: If $A \subseteq B$, $B \subseteq C$ then $A \subseteq C$ (6.12)

Involution: $(A^c)^c = A$ (6.13)

Law of the Excluded Middle: $A \cup A^c = E$ (6.14)

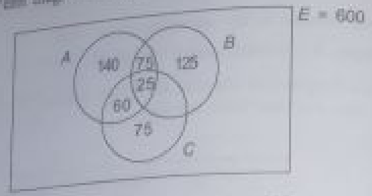
Law of Contradiction: $A \cap A^c = \varnothing$ (6.15)

De Morgan's laws:
$$(A \cup B)^c = A^c \cap B^c$$
$$(A \cap B)^c = A^c \cup B^c \quad (6.16)$$

All the properties could be verified by means of Venn diagrams.

Solution

From the given data, the Venn diagram obtained is as follows:



(i) No. of female students $|A^c| = |E| - |A| = 600 - 300 = 300$

(ii) No. of students who are not bowlers $|B^c| = |E| - |B| = 600 - 225 = 375$

(iii) No. of students who are not batsmen $|C^c| = |E| - |C| = 600 - 160 = 440$

(iv) No. of female students who can bowl $|A^c \cap B| = 125$ (from the Venn diagram)

6.2.3 Partition and Covering

Partition

A *partition* on A is defined to be a set of non-empty subsets A_i, each of which is pairwise disjoint and whose union yields the original set A.

Partition on A indicated as $\Pi(A)$, is therefore

(i) $A_i \cap A_j = \varnothing$ for each pair $(i, j) \in I$, $i \neq j$ (6.17)

(ii) $\bigcup_{i \in I} A_i = A$

The members A_i of the partition are known as blocks (refer Fig. 6.8).

Fig. 6.8 Partition of set A.

Example

Given $A = \{a, b, c, d, e\}$, $A_1 = \{a, b\}$, $A_2 = \{c, d\}$ and $A_3 = \{e\}$, which gives

$$A_1 \cap A_2 = \varnothing,\ A_1 \cap A_3 = \varnothing,\ A_2 \cap A_3 = \varnothing$$

Also,

$$A_1 \cup A_2 \cup A_3 = A = \{a, b, c, d, e\}$$

Hence, $\{A_1, A_2, A_3\}$, is a partition on A.

Covering

A *covering* on A is defined to be a set of non-empty subsets A_i whose union yields the original set A. The non-empty subsets need not be disjoint (Refer Fig. 6.9).

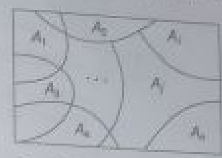


Fig. 6.9 Covering of set A.

Example

Given $A = \{a, b, c, d, e\}$, $A_1 = \{a, b\}$, $A_2 = \{b, c, d\}$, and $A_3 = \{d, e\}$. This gives

$$A_1 \cap A_2 = \{b\}$$
$$A_1 \cap A_3 = \varnothing$$
$$A_2 \cap A_3 = \{d\}$$

Also,

$$A_1 \cup A_2 \cup A_3 = \{a, b, c, d, e\} = A$$

Hence, $\{A_1, A_2, A_3\}$ is a covering on A.

Rule of Addition

Given a partition on A where A_i, $i = 1, 2, \ldots, n$ are its non-empty subsets then,

$$|A| = |\bigcup_{i=1}^{n} A_i| = \sum_{i=1}^{n} |A_i| \qquad (6.18)$$

Example

Given $A = \{a, b, c, d, e\}$, $A_1 = \{a, b\}$, $A_2 = \{c, d\}$, $A_3 = \{e\}$, $|A| = 5$, and $\Sigma_{i=1}^{3} |A_i| = 2 + 2 + 1 = 5$

Rule of Inclusion and Exclusion

Rule of addition is not applicable on the covering of set A, especially if the subsets are not pairwise disjoint. In such a case, the rule of inclusion and exclusion is applied.

Example

Given A to be a covering of n sets $A_1, A_2, \ldots, A_n$,

for $n = 2$, $$|A| = |A_1 \cup A_2| = |A_1| + |A_2| - |A_1 \cap A_2| \qquad (6.19)$$

for $n = 3$, $$|A| = |A_1 \cup A_2 \cup A_3| = |A_1| + |A_2| + |A_3| - |A_1 \cap A_2| - |A_2 \cap A_3| - |A_1 \cap A_3| + |A_1 \cap A_2 \cap A_3| \qquad (6.20)$$

Generalizing

$$|A| = \left|\bigcup_{i=1}^{n} A_i\right| = \sum_{i=1}^{n}|A_i| - \sum_{i=1}^{n}\sum_{\substack{j=1\\ i\neq j}}^{n}|A_i \cap A_j|$$

$$+\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{\substack{k=1\\ i\neq j\neq k}}^{n}|A_i \cap A_j \cap A_k| \ldots (-1)^{n+1}\left|\bigcap_{i=1}^{n} A_i\right| \quad (6.22)$$

Example 6.4

Given $|E| = 100$, where E indicates a set of students who have chosen subjects from different streams in the computer science discipline. It is found that 32 study subjects chosen from the Computer Networks (CN) stream, 20 from the Multimedia Technology (MMT) stream, and 45 from the Systems Software (SS) stream. Also, 15 study subjects from both CN and SS streams, 7 from both MMT and SS streams, and 30 do not study any subjects chosen from either of the three streams.

Find the number of students who study subjects belonging to all three streams.

Solution

Let A, B, C indicate students who study subjects chosen from CN, MMT, and SS streams respectively. The problem is to find $|A \cap B \cap C|$.

The no. of students who do not study any subject chosen from either of the three streams = 30.

Let $|A' \cap B' \cap C'| = 30$

$\Rightarrow$ $|(A \cup B \cup C)'| = 30$ (using De Morgan's laws)

$\Rightarrow$ $|E| - |A \cup B \cup C| = 30$

$\Rightarrow$ $|A \cup B \cup C| = |E| - 30$

$= 100 - 30 = 70$

From the principle of inclusion and exclusion,

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |B \cap C| - |A \cap C| + |A \cap B \cap C|$$

$$\Rightarrow |A \cap B \cap C| = |A \cup B \cup C| - |A| - |B| - |C| + |A \cap B| + |B \cap C| + |A \cap C|$$

$$= 70 - 32 - 20 - 45 + 15 + 7 + 10$$

$$= 5$$

Hence, the no. of students who study subjects chosen from all the three streams is 5.

6.3 FUZZY SETS

Fuzzy sets support a flexible sense of membership of elements to a set. While in crisp set theory, an element either belongs to or does not belong to a set, in fuzzy set theory many degrees of membership (between 0 and 1) are allowed. Thus, a membership function $\mu_{\tilde{A}}(x)$ is associated with a

fuzzy set $\tilde{A}$ such that the function maps every element of the universe of discourse X (or the reference set) to the interval [0, 1].

Formally, the mapping is written as $\mu_{\tilde{A}}(x) : X \to [0, 1]$

A *fuzzy set* is defined as follows:

If X is a universe of discourse and x is a particular element of X, then a fuzzy set A defined on X may be written as a collection of ordered pairs

$$A = \{(x, \mu_{\tilde{A}}(x)), x \in X\} \quad (6.23)$$

where each pair $(x, \mu_{\tilde{A}}(x))$ is called a singleton. In crisp sets, $\mu_{\tilde{A}}(x)$ is dropped.

An alternative definition which indicates a fuzzy set as a union of all $\mu_{\tilde{A}}(x)/x$ singletons is given by

$$A = \sum_{x_i \in X} \mu_{\tilde{A}}(x_i)/x_i \quad \text{in the discrete case} \quad (6.24)$$

and

$$A = \int_X \mu_{\tilde{A}}(x)/x \quad \text{in the continuous case} \quad (6.25)$$

Here, the summation and integration signs indicate the union of all $\mu_{\tilde{A}}(x)/x$ singletons.

Example

Let $X = \{g_1, g_2, g_3, g_4, g_5\}$ be the reference set of students. Let $\tilde{A}$ be the fuzzy set of "smart" students, where "smart" is a fuzzy linguistic term.

$$\tilde{A} = \{(g_1, 0.4)(g_2, 0.5)(g_3, 1)(g_4, 0.9)(g_5, 0.8)\}$$

Here $\tilde{A}$ indicates that the smartness of g_1 is 0.4, g_2 is 0.5 and so on when graded over a scale of 0–1.

Though fuzzy sets model vagueness, it needs to be realized that the definition of the sets varies according to the context in which it is used. Thus, the fuzzy linguistic term "tall" could have one kind of fuzzy set while referring to the height of a building and another kind of fuzzy set while referring to the height of human beings.

6.3.1 Membership Function

The membership function values need not always be described by discrete values. Quite often, these turn out to be as described by a continuous function.

The fuzzy membership function for the fuzzy linguistic term "cool" relating to temperature may turn out to be as illustrated in Fig. 6.10.

Fig. 6.10 Continuous membership function for "cool".

A membership function can also be given mathematically as

$$\mu_{\tilde{A}}(x) = \frac{1}{(1+x)^2}$$

The graph is as shown in Fig. 6.11.

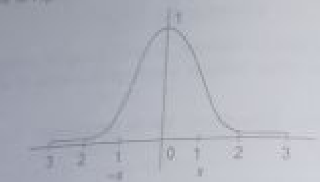


Fig. 6.11 Continuous membership function dictated by a mathematical function.

Different shapes of membership functions exist. The shapes could be triangular, trapezoidal, curved or their variations as shown in Fig. 6.12.

Fig. 6.12 Different shapes of membership function graphs.

Example

Consider the set of people in the following age groups

0–10	40–50
10–20	50–60
20–30	60–70
30–40	70 and above

The fuzzy sets "young", "middle-aged", and "old" are represented by the membership function graphs as illustrated in Fig. 6.13.

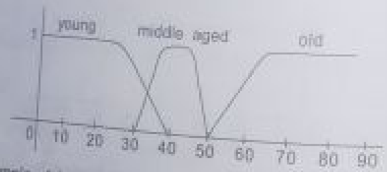


Fig. 6.13 Example of fuzzy sets expressing "young", "middle-aged", and "old".

6.3.2 Basic Fuzzy Set Operations

Given X to be the universe of discourse and $\tilde{A}$ and $\tilde{B}$ to be fuzzy sets with $\mu_A(x)$ and $\mu_B(x)$ as their respective membership functions, the basic fuzzy set operations are as follows:

Union

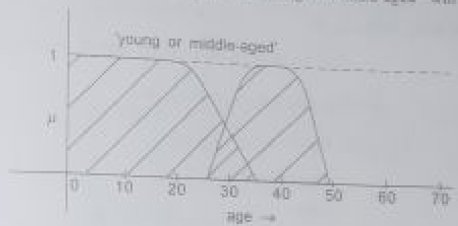
The union of two fuzzy sets $\tilde{A}$ and $\tilde{B}$ is a new fuzzy set $\tilde{A} \cup \tilde{B}$ also on X with a membership function defined as

$$\mu_{\tilde{A} \cup \tilde{B}}(x) = \max(\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x)) \tag{6.26}$$

Example

Let $\tilde{A}$ be the fuzzy set of young people and $\tilde{B}$ be the fuzzy set of middle-aged people as illustrated in Fig. 6.13. Now $\tilde{A} \cup \tilde{B}$, the fuzzy set of "young or middle-aged" will be given by

$\tilde{A} \cup \tilde{B}$:



In its discrete form, for x_1, x_2, x_3

if $\tilde{A} = \{(x_1, 0.5), (x_2, 0.7), (x_3, 0)\}$ and $\tilde{B} = \{(x_1, 0.8), (x_2, 0.2), (x_3, 1)\}$

$$\tilde{A} \cup \tilde{B} = \{(x_1, 0.8), (x_2, 0.7), (x_3, 1)\}$$

since,

$$\mu_{\tilde{A} \cup \tilde{B}}(x_1) = \max(\mu_{\tilde{A}}(x_1), \mu_{\tilde{B}}(x_1))$$
$$= \max(0.5, 0.8)$$
$$= 0.8$$

$$\mu_{\tilde{A} \cup \tilde{B}}(x_2) = \max(\mu_{\tilde{A}}(x_2), \mu_{\tilde{B}}(x_2)) = \max(0.2, 0.7) = 0.7$$

$$\mu_{\tilde{A} \cup \tilde{B}}(x_3) = \max(\mu_{\tilde{A}}(x_3), \mu_{\tilde{B}}(x_3)) = \max(0, 1) = 1$$

Intersection

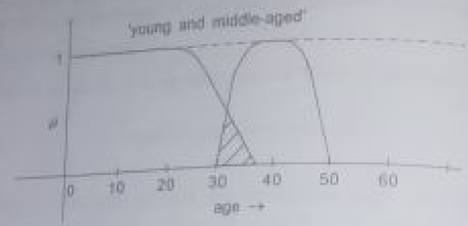
The intersection of fuzzy sets $\tilde{A}$ and $\tilde{B}$ is a new fuzzy set $\tilde{A} \cap \tilde{B}$ with membership function defined as

$$\mu_{\tilde{A} \cap \tilde{B}}(x) = \min(\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x)) \tag{6.27}$$

Example

For $\tilde{A}$ and $\tilde{B}$ defined as "young" and "middle-aged" as illustrated in previous examples.

$\tilde{A} \cap \tilde{B}$



In its discrete form, for x_1, x_2, x_3

if $\tilde{A} = \{(x_1, 0.5), (x_2, 0.7), (x_3, 0)\}$ and $\tilde{B} = \{(x_1, 0.8), (x_2, 0.2), (x_3, 1)\}$

$$\tilde{A} \cap \tilde{B} = \{(x_1, 0.5), (x_2, 0.2), (x_3, 0)\}$$

since,

$$\mu_{\tilde{A} \cap \tilde{B}}(x_1) = \min(\mu_{\tilde{A}}(x_1), \mu_{\tilde{B}}(x_1))$$
$$= \min(0.5, 0.8)$$
$$= 0.5$$

$$\mu_{\tilde{A} \cap \tilde{B}}(x_2) = \min(\mu_{\tilde{A}}(x_2), \mu_{\tilde{B}}(x_2))$$
$$= \min(0.7, 0.2)$$
$$= 0.2$$

$$\mu_{\tilde{A} \cap \tilde{B}}(x_3) = \min(\mu_{\tilde{A}}(x_3), \mu_{\tilde{B}}(x_3))$$
$$= \min(0, 1)$$
$$= 0$$

Complement

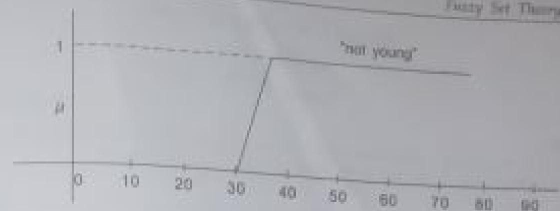
The complement of a fuzzy set $\tilde{A}$ is a new fuzzy set $\tilde{A}^c$ with a membership function

$$\mu_{\tilde{A}^c}(x) = 1 - \mu_{\tilde{A}}(x) \qquad (6.28)$$

Example

For the fuzzy set $\tilde{A}$ defined as "young" the complement "not young" is given by $\tilde{A}^c$. In its discrete form, for x_1, x_2, and x_3

$\tilde{A}^c$



if $\tilde{A} = \{(x_1, 0.5)\ (x_2, 0.7)\ (x_3, 0)\}$

then, $\tilde{A}^c = \{(x_1, 0.5)\ (x_2, 0.3)\ (x_3, 1)\}$

since,

$$\mu_{\tilde{A}^c}(x_1) = 1 - \mu_{\tilde{A}}(x_1)$$
$$= 1 - 0.5$$
$$= 0.5$$

$$\mu_{\tilde{A}^c}(x_2) = 1 - \mu_{\tilde{A}}(x_2)$$
$$= 1 - 0.7$$
$$= 0.3$$

$$\mu_{\tilde{A}^c}(x_3) = 1 - \mu_{\tilde{A}}(x_3)$$
$$= 1 - 0$$
$$= 1$$

Other operations are,

Product of two fuzzy sets

The product of two fuzzy sets $\tilde{A}$ and $\tilde{B}$ is a new fuzzy set $\tilde{A} \cdot \tilde{B}$ whose membership function is defined as

$$\mu_{\tilde{A} \cdot \tilde{B}}(x) = \mu_{\tilde{A}}(x)\, \mu_{\tilde{B}}(x) \qquad (6.29)$$

Example

$$\tilde{A} = \{(x_1, 0.2), (x_2, 0.8), (x_3, 0.4)\}$$
$$\tilde{B} = \{(x_1, 0.4)\ (x_2, 0), (x_3, 0.1)\}$$
$$\tilde{A} \cdot \tilde{B} = \{(x_1, 0.08)\ (x_2, 0)\ (x_3, 0.04)\}$$

Since

$$\mu_{\tilde{A} \cdot \tilde{B}}(x_1) = \mu_{\tilde{A}}(x_1) \cdot \mu_{\tilde{B}}(x_1)$$
$$= 0.2 \cdot 0.4 = 0.08$$

$$\mu_{\tilde{A} \cdot \tilde{B}}(x_2) = \mu_{\tilde{A}}(x_2) \cdot \mu_{\tilde{B}}(x_2)$$
$$= 0.8 \cdot 0 = 0$$

$$\mu_{\tilde{A}\cdot\tilde{B}}(x_3) = \mu_{\tilde{A}}(x_3)\cdot\mu_{\tilde{A}}(x_3)$$
$$= 0.4 \cdot 0.1$$
$$= 0.04$$

Equality

Two fuzzy sets $\tilde{A}$ and $\tilde{B}$ are said to be equal $(\tilde{A} = \tilde{B})$ if $\mu_{\tilde{A}}(x) = \mu_{\tilde{B}}(x)$ (6.30)

Example

$$\tilde{A} = \{(x_1, 0.2)(x_2, 0.8)\}$$
$$\tilde{B} = \{(x_1, 0.6)(x_2, 0.8)\}$$
$$\tilde{C} = \{(x_1, 0.2)(x_2, 0.8)\}$$
$$\tilde{A} \neq \tilde{B}$$

since $\mu_{\tilde{A}}(x_1) \neq \mu_{\tilde{B}}(x_1)$ although

$$\mu_{\tilde{A}}(x_2) = \mu_{\tilde{B}}(x_2)$$

but $\tilde{A} = \tilde{C}$

since $\mu_{\tilde{A}}(x_1) = \mu_{\tilde{C}}(x_1) = 0.2$

and $\mu_{\tilde{A}}(x_2) = \mu_{\tilde{C}}(x_2) = 0.8$

Product of a fuzzy set with a crisp number

Multiplying a fuzzy set $\tilde{A}$ by a crisp number a results in a new fuzzy set product $a \cdot \tilde{A}$ with the membership function

$$\mu_{a\cdot\tilde{A}}(x) = a \cdot \mu_{\tilde{A}}(x) \tag{6.31}$$

Example

For $\tilde{A} = \{(x_1, 0.4), (x_2, 0.6), (x_3, 0.8)\}$

$$a = 0.3$$
$$a \cdot \tilde{A} = \{(x_1, 0.12), (x_2, 0.18), (x_3, 0.24)\}$$

since,

$$\mu_{a\cdot\tilde{A}}(x_1) = a \cdot \mu_{\tilde{A}}(x_1)$$
$$= 0.3 \cdot 0.4$$
$$= 0.12$$
$$\mu_{a\cdot\tilde{A}}(x_2) = a \cdot \mu_{\tilde{A}}(x_2)$$
$$= 0.3 \cdot 0.6$$
$$= 0.18$$

$$\mu_{a\cdot\tilde{A}}(x_3) = a \cdot \mu_{\tilde{A}}(x_3)$$
$$= 0.3 \cdot 0.8$$
$$= 0.24$$

Power of a fuzzy set

The α power of a fuzzy set $\tilde{A}$ is a new fuzzy set A^{α} whose membership function is given by

$$\mu_{A^{\alpha}}(x) = (\mu_{\tilde{A}}(x))^{\alpha} \tag{6.32}$$

Raising a fuzzy set to its second power is called *Concentration* (CON) and taking the square root is called *Dilation* (DIL).

Example

$$\tilde{A} = \{(x_1, 0.4), (x_2, 0.2), (x_3, 0.7)\}$$

For $\alpha = 2$

$$\mu_{\tilde{A}^2}(x) = (\mu_{\tilde{A}}(x))^2$$

Hence, $(\tilde{A})^2 = \{(x_1, 0.16), (x_2, 0.04), (x_3, 0.49)\}$

Since

$$\mu_{\tilde{A}^2}(x_1) = (\mu_{\tilde{A}}(x_1))^2 = (0.4)^2 = 0.16$$
$$\mu_{\tilde{A}^2}(x_2) = (\mu_{\tilde{A}}(x_2))^2 = (0.2)^2 = 0.04$$
$$\mu_{\tilde{A}^2}(x_3) = (\mu_{\tilde{A}}(x_3))^2 = (0.7)^2 = 0.49$$

Difference

The difference of two fuzzy sets $\tilde{A}$ and $\tilde{B}$ is a new fuzzy set $\tilde{A} - \tilde{B}$ defined as

$$\tilde{A} - \tilde{B} = (\tilde{A} \cap \tilde{B}^c) \tag{6.33}$$

Example

$$\tilde{A} = \{(x_1, 0.2), (x_2, 0.5), (x_3, 0.6)\};\ \tilde{B} = \{(x_1, 0.1), (x_2, 0.4), (x_3, 0.5)\}$$
$$\tilde{B}^c = \{(x_1, 0.9), (x_2, 0.6), (x_3, 0.5)\}$$
$$\tilde{A} - \tilde{B} = \tilde{A} \cap \tilde{B}^c$$
$$= \{(x_1, 0.2)(x_2, 0.5)(x_3, 0.5)\}$$

Disjunctive sum

The disjunctive sum of two fuzzy sets $\tilde{A}$ and $\tilde{B}$ is a new fuzzy set $\tilde{A} \oplus \tilde{B}$ defined as

$$\tilde{A} \oplus \tilde{B} = (\tilde{A}^c \cap \tilde{B}) \cup (\tilde{A} \cap \tilde{B}^c) \tag{6.34}$$

Example

$$\tilde{A} = \{(x_1, 0.4)(x_2, 0.8)(x_3, 0.6)\}$$

$$\tilde{B} = \{(x_1, 0.2)(x_2, 0.6)(x_3, 0.9)\}$$

$$\tilde{A}^c = \{(x_1, 0.6)(x_2, 0.2)(x_3, 0.4)\}$$

Now,

$$\tilde{B}^c = \{(x_1, 0.8)(x_2, 0.4)(x_3, 0.1)\}$$

$$\tilde{A}^c \cap \tilde{B} = \{(x_1, 0.2)(x_2, 0.2)(x_3, 0.4)\}$$

$$\tilde{A} \cap \tilde{B}^c = \{(x_1, 0.4)(x_2, 0.4)(x_3, 0.1)\}$$

$$\tilde{A} \oplus \tilde{B} = \{(x_1, 0.4)(x_2, 0.4)(x_3, 0.4)\}$$

6.3.3 Properties of Fuzzy Sets

Fuzzy sets follow some of the properties satisfied by crisp sets. In fact, crisp sets can be thought of as special instances of fuzzy sets. Any fuzzy set $\tilde{A}$ is a subset of the reference set X. Also, the membership of any element belonging to the null set $\varnothing$ is 0 and the membership of any element belonging to the reference set is 1.

The properties satisfied by fuzzy sets are

Commutativity:

$$\tilde{A} \cup \tilde{B} = \tilde{B} \cup \tilde{A}$$

$$\tilde{A} \cap \tilde{B} = \tilde{B} \cap \tilde{A} \quad (6.35)$$

Associativity:

$$\tilde{A} \cup (\tilde{B} \cup \tilde{C}) = (\tilde{A} \cup \tilde{B}) \cup \tilde{C}$$

$$\tilde{A} \cap (\tilde{B} \cap \tilde{C}) = (\tilde{A} \cap \tilde{B}) \cap \tilde{C} \quad (6.36)$$

Distributivity:

$$\tilde{A} \cup (\tilde{B} \cap \tilde{C}) = (\tilde{A} \cup \tilde{B}) \cap (\tilde{A} \cup \tilde{C})$$

$$\tilde{A} \cap (\tilde{B} \cup \tilde{C}) = (\tilde{A} \cap \tilde{B}) \cup (\tilde{A} \cap \tilde{C}) \quad (6.37)$$

Idempotence:

$$\tilde{A} \cup \tilde{A} = \tilde{A}$$

$$\tilde{A} \cap \tilde{A} = \tilde{A} \quad (6.38)$$

Identity:

$$\tilde{A} \cup \varnothing = \tilde{A}$$

$$\tilde{A} \cup X = \tilde{A}$$

$$\tilde{A} \cap \varnothing = \varnothing$$

$$\tilde{A} \cup X = X \quad (6.39)$$

Transitivity: If $\tilde{A} \subseteq \tilde{B} \subseteq \tilde{C}$, then $\tilde{A} \subseteq \tilde{C}$ (6.40)

Involution:

$$(\tilde{A}^c)^c = \tilde{A} \quad (6.41)$$

De Morgan's laws:

$$(\tilde{A} \cap \tilde{B})^c = (\tilde{A}^c \cup \tilde{B}^c)$$

$$(\tilde{A} \cup \tilde{B})^c = (\tilde{A}^c \cap \tilde{B}^c) \quad (6.42)$$

Since fuzzy sets can overlap, the laws of excluded middle do not hold good. Thus,

$$\tilde{A} \cup \tilde{A}^c \neq X \quad (6.43)$$

$$\tilde{A} \cap \tilde{A}^c \neq \varnothing \quad (6.44)$$

Example 6.5

The task is to recognize English alphabetical characters (F, E, X, Y, I, T) in an image processing system.

Define two fuzzy sets $\tilde{I}$ and $\tilde{F}$ to represent the identification of characters I and F.

$$\tilde{I} = \{(F, 0.4), (E, 0.3), (X, 0.1), (Y, 0.1), (I, 0.9), (T, 0.8)\}$$

$$\tilde{F} = \{(F, 0.99), (E, 0.8), (X, 0.1), (Y, 0.2), (I, 0.5), (T, 0.5)\}$$

Find the following.

(a) (i) $\tilde{I} \cup \tilde{F}$ (ii) $(\tilde{I} - \tilde{F})$ (iii) $\tilde{F} \cup \tilde{F}^c$

(b) Verify De Morgan's Law, $(\tilde{I} \cup \tilde{F})^c = \tilde{I}^c \cap \tilde{F}^c$

Solution

(a) (i) $\tilde{I} \cup \tilde{F} = \{(F, 0.99), (E, 0.8), (X, 0.1), (Y, 0.2), (I, 0.9), (T, 0.8)\}$

(ii) $\tilde{I} - \tilde{F} = (\tilde{I} \cap \tilde{F}^c)$

$= \{(F, 0.01), (E, 0.2), (X, 0.1), (Y, 0.1), (I, 0.5), (T, 0.5)\}$

(iii) $\tilde{F} \cup \tilde{F}^c = \{(F, 0.99), (E, 0.8), (X, 0.9), (Y, 0.8), (I, 0.5), (T, 0.5)\}$

(b) De Morgan's Law

$$(\tilde{I} \cup \tilde{F})^c = \tilde{I}^c \cap \tilde{F}^c$$

$$\tilde{I} \cup \tilde{F} = \{(F, 0.99), (E, 0.8), (X, 0.1), (Y, 0.2), (I, 0.9), (T, 0.8)\}$$

$$(\tilde{I} \cup \tilde{F})^c = \{(F, 0.01), (E, 0.2), (X, 0.9), (Y, 0.8), (I, 0.1), (T, 0.2)\}$$

$$\tilde{I}^c = \{(F, 0.6), (E, 0.7), (X, 0.9), (Y, 0.9), (I, 0.1), (T, 0.2)\}$$

$$\tilde{F}^c = \{(F, 0.01), (E, 0.2), (X, 0.9), (Y, 0.8), (I, 0.5), (T, 0.5)\}$$

and

$$\tilde{I}^c \cap \tilde{F}^c = \{(F, 0.01), (E, 0.2), (X, 0.9), (Y, 0.8), (I, 0.1), (T, 0.2)\}$$

Hence,

$$(\tilde{I} \cup \tilde{F})^c = \tilde{I}^c \cap \tilde{F}^c$$

Example 6.6

Consider the fuzzy sets $\tilde{A}$ and $\tilde{B}$ defined on the interval $X = [0, 5]$ of real numbers, by the membership grade functions

$$\mu_{\tilde{A}}(x) = \frac{x}{x+1}, \ \mu_{\tilde{B}}(x) = 2^{-x}$$

Determine the mathematical formulae and graphs of the membership grade functions of each of the following sets

(a) A^c, B^c

(b) $A \cup B$

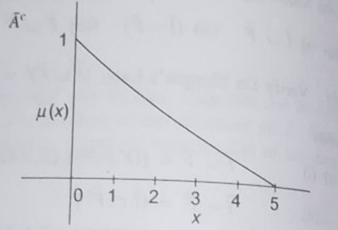
(c) $A \cap B$

(d) $(A \cup B)^c$

Solution

(a) $\mu_{\tilde{A}^c}(x) = 1 - \mu_{\tilde{A}}(x) = 1 - \frac{x}{x+1}$

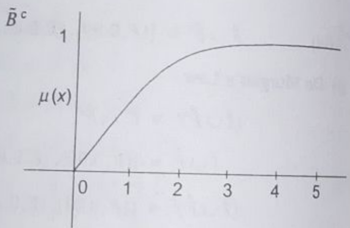
$= \frac{1}{x+1}$



$\mu_{\tilde{B}^c}(x) = 1 - \mu_{\tilde{B}}(x)$

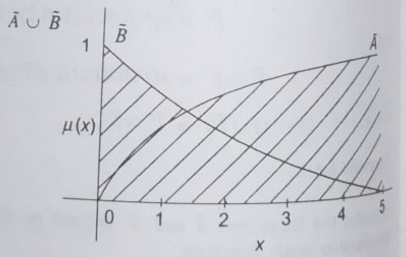
$= 1 - 2^{-x}$

$= \frac{2^x - 1}{2^x}$



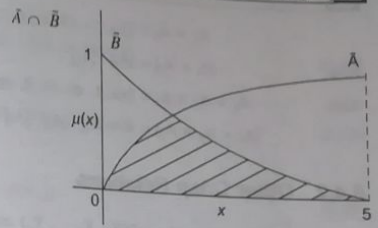
(b) $\mu_{\tilde{A} \cup \tilde{B}}(x) = \max(\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x))$

$= \max\left(\frac{x}{x+1}, 2^{-x}\right)$



(c) $\mu_{\tilde{A} \cap \tilde{B}}(x) = \min(\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x))$

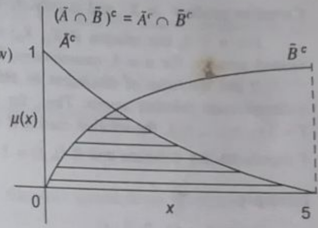
$= \min\left(\frac{x}{x+1}, 2^{-x}\right)$



(d) $\mu_{(\tilde{A} \cup \tilde{B})^c}(x) = \mu_{\tilde{A}^c \cap \tilde{B}^c}(x)$ (Θ De Morgan's law)

$= \min(\mu_{\tilde{A}^c}(x), \mu_{\tilde{B}^c}(x))$

$= \min\left(\frac{1}{x+1}, \frac{2^x - 1}{2^x}\right)$



6.4 CRISP RELATIONS

In this section, we review crisp relations as a prelude to fuzzy relations. The concept of relations between sets is built on the Cartesian product operator of sets.

6.4.1 Cartesian Product

The *Cartesian product* of two sets A and B denoted by $A \times B$ is the set of all ordered pairs such that the first element in the pair belongs to A and the second element belongs to B.

i.e. $$A \times B = \{(a, b)/a \in A, \ b \in B\}$$

If $A \neq B$ and A and B are non-empty then $A \times B \neq B \times A$.

The Cartesian product could be extended to n number of sets

$$\mathop{\times}_{i=1}^{n} A_i = \{(a_1, a_2, a_3, \ldots, a_n)/a_i \in A_i \text{ for every } i = 1, 2, \ldots, n\} \quad (6.45)$$

Observe that $$\left|\mathop{\times}_{i=1}^{n} A_i\right| = \prod_{i=1}^{n} |A_i| \quad (6.46)$$

Example

Given $A_1 = \{a, b\}, A_2 = \{1, 2\}, A_3 = \{\alpha\}$,

$$A_1 \times A_2 = \{(a, 1), (b, 1), (a, 2), (b, 2)\}, |A_1 \times A_2| = 4, \text{ and } |A_1| = |A_2| = $$

Here, $|A_1 \times A_2| = |A_1| \cdot |A_2|$

Also, $A_1 \times A_2 \times A_3 = \{(a, 1, \alpha), (a, 2, \alpha), (b, 1, \alpha), (b, 2, \alpha)\}$

$$|A_1 \times A_2 \times A_3| = 4 = |A_1| \cdot |A_2| \cdot |A_3|$$

6.4.2 Other Crisp Relations

An *n-ary* relation denoted as $R(X_1, X_2, \ldots, X_n)$ among crisp sets $X_1, X_2, \ldots, X_n$ is a subset of the Cartesian product $\overset{n}{\underset{i=1}{\times}} X_i$ and is indicative of an association or relation among the tuple elements.

For $n = 2$, the relation $R(X_1, X_2)$ is termed as a *binary* relation; for $n = 3$, the relation is termed *ternary*; for $n = 4$, *quarternary*; for $n = 5$, *quinary* and so on.

If the universe of discourse or sets are finite, the *n-ary* relation can be expressed as an *n*-dimensional *relation matrix*. Thus, for a binary relation $R(X, Y)$ where $X = \{x_1, x_2, \ldots, x_n\}$ and $Y = \{y_1, y_2, \ldots, y_m\}$, the relation matrix R is a two dimensional matrix where X represents the rows, Y represents the columns and $R(i, j) = 1$ if $(x_i, y_j) \in R$ and $R(i, j) = 0$ if $(x_i, y_j) \notin R$.

Example

Given $X = \{1, 2, 3, 4\}$,

$$X \times X = \begin{Bmatrix} (1,1)(1,2)(1,3)(1,4)(2,1)(2,2)(2,3)(2,4) \\ (3,1)(3,2)(3,3)(3,4)(4,1)(4,2)(4,3)(4,4) \end{Bmatrix}$$

Let the relation R be defined as

$$R = \{(x, y)/y = x + 1,\ x, y \in X\}$$

$$R = \{(1, 2)(2, 3)(3, 4)\}$$

The relation matrix R is given by

$$R = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

6.4.3 Operations on Relations

Given two relations R and S defined on $X \times Y$ and represented by relation matrices, the following operations are supported by R and S

Union: $R \cup S$

$$R \cup S(x, y) = \max(R(x, y), S(x, y)) \tag{6.47}$$

Intersection: $R \cap S$

$$R \cap S(x, y) = \min(R(x, y), S(x, y)) \tag{6.48}$$

Complement: $\overline{R}$

$$\overline{R}(x, y) = 1 - R(x, y) \tag{6.49}$$

Composition of relations: $R \circ S$

Given R to be a relation on X, Y and S to be a relation on Y, Z then $R \circ S$ is a composition of relation on X, Z defined as

$$R \circ S = \{(x, z)/(x, z) \in X \times Z,\ \exists\ y \in Y \text{ such that } (x, y) \in R \text{ and } (y, z) \in S\} \tag{6.50}$$

A common form of the composition relation is the *max-min composition.*

Max-min composition:

Given the relation matrices of the relation R and S, the max-min composition is defined as

For $$T = R \circ S$$

$$T(x, z) = \max_{y \in Y}(\min(R(x, y), S(y, z))) \tag{6.51}$$

Example

Let R, S be defined on the sets $\{1, 3, 5\} \times \{1, 3, 5\}$

Let $R: \{(x, y) \mid y = x + 2\}$, $S: \{(x, y) \mid x \le y\}$

$$R = \{(1, 3)(3, 5)\}, \qquad S = \{(1, 3)(1, 5)(3, 5)\}$$

The relation matrices are

$$R: \begin{array}{c} \\ 1 \\ 3 \\ 5 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 3 & 5 \end{array} \\ \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \end{array} \qquad S: \begin{array}{c} \\ 1 \\ 3 \\ 5 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 3 & 5 \end{array} \\ \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \end{array}$$

Using max-min composition

$$R \circ S = \begin{array}{c} \\ 1 \\ 3 \\ 5 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 3 & 5 \end{array} \\ \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \end{array}$$

$$R \circ S\,(1, 1) = \max\{\min(0, 0), \min(1, 0), \min(0, 0)\}$$

since

$$= \max(0, 0, 0) = 0.$$

$$R \circ S\,(1, 3) = \max\{0, 0, 0\} = 0$$

$$R \circ S\,(1, 5) = \max\{0, 1, 0\} = 1.$$

$$R \circ S\,(3, 1) = 0.$$

Similarly,

$$R \circ S\,(3, 3) = R \circ S\,(3, 5) = R \circ S\,(5, 1) = R \circ S\,(5, 3) = R \circ S\,(5, 5$$

$R \circ S$ from the relation matrix is $\{(1, 5)\}$.

Also,

$$S \circ R = \begin{matrix} & 1 & 3 & 5 \\ 1 & 0 & 0 & 1 \\ 3 & 0 & 0 & 0 \\ 5 & 0 & 0 & 0 \end{matrix}$$

6.5 FUZZY RELATIONS

Fuzzy relation is a fuzzy set defined on the Cartesian product of crisp sets $X_1, X_2, \ldots, X_n$ where n-tuples $(x_1, x_2, \ldots, x_n)$ may have varying degrees of membership within the relation. Th membership values indicate the strength of the relation between the tuples.

Example

Let R be the fuzzy relation between two sets X_1 and X_2 where X_1 is the set of diseases a X_2 is the set of symptoms.

X_1 = {typhoid, viral fever, common cold}

X_2 = {running nose, high temperature, shivering}

The fuzzy relation R may be defined as

	Running nose	*High temperature*	*Shivering*
Typhoid	0.1	0.9	0.8
Viral fever	0.2	0.9	0.7
Common cold	0.9	0.4	0.6

6.5.1 Fuzzy Cartesian Product

Let $\tilde{A}$ be a fuzzy set defined on the universe X and $\tilde{B}$ be a fuzzy set defined on the universe the Cartesian product between the fuzzy sets $\tilde{A}$ and $\tilde{B}$ indicated as $\tilde{A} \times \tilde{B}$ and resulting i fuzzy relation $\tilde{R}$ is given by

$$\tilde{R} = \tilde{A} \times \tilde{B} \subset X \times Y \quad (6.52)$$

where $\tilde{R}$ has its membership function given by

$$\mu_{\tilde{R}}(x, y) = \mu_{\tilde{A} \times \tilde{B}}(x, y) = \min(\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(y)) \quad (6.53)$$

Example

Let $\tilde{A} = \{(x_1, 0.2), (x_2, 0.7), (x_3, 0.4)\}$ and $\tilde{B} = \{(y_1, 0.5), (y_2, 0.6)\}$ be two fuzzy sets defined on the universes of discourse $X = \{x_1, x_2, x_3\}$ and $Y = \{y_1, y_2\}$ respectively. Then the fuzzy relation $\tilde{R}$ resulting out of the fuzzy Cartesian product $\tilde{A} \times \tilde{B}$ is given by

$$\tilde{R} = \tilde{A} \times \tilde{B} = \begin{matrix} & y_1 & y_2 \\ x_1 & 0.2 & 0.2 \\ x_2 & 0.5 & 0.6 \\ x_3 & 0.4 & 0.4 \end{matrix}$$

since,

$$\tilde{R}(x_1, y_1) = \min(\mu_{\tilde{A}}(x_1), \mu_{\tilde{B}}(y_1)) = \min(0.2, 0.5) = 0.2$$
$$\tilde{R}(x_1, y_2) = \min(0.2, 0.6) = 0.2$$
$$\tilde{R}(x_2, y_1) = \min(0.7, 0.5) = 0.5$$
$$\tilde{R}(x_2, y_2) = \min(0.7, 0.6) = 0.6$$
$$\tilde{R}(x_3, y_1) = \min(0.4, 0.5) = 0.4$$
$$\tilde{R}(x_3, y_2) = \min(0.4, 0.6) = 0.4$$

6.5.2 Operations on Fuzzy Relations

Let $\tilde{R}$ and $\tilde{S}$ be fuzzy relations on $X \times Y$.

Union

$$\mu_{\tilde{R} \cup \tilde{S}}(x, y) = \max(\mu_{\tilde{R}}(x, y), \mu_{\tilde{S}}(x, y)) \quad (6.54)$$

Intersection

$$\mu_{\tilde{R} \cap \tilde{S}}(x, y) = \min(\mu_{\tilde{R}}(x, y), \mu_{\tilde{S}}(x, y)) \quad (6.55)$$

Complement

$$\mu_{\tilde{R}^c}(x, y) = 1 - \mu_{\tilde{R}}(x, y) \quad (6.56)$$

Composition of relations

The definition is similar to that of crisp relation. Suppose $\tilde{R}$ is a fuzzy relation defined on $X \times Y$, and $\tilde{S}$ is a fuzzy relation defined on $Y \times Z$, then $\tilde{R} \circ \tilde{S}$ is a fuzzy relation on $X \times Z$. The fuzzy max-min composition is defined as

$$\mu_{\tilde{R}\circ\tilde{S}}(x,z) = \max_{y\in Y}(\min(\mu_{\tilde{R}}(x,y),\ \mu_{\tilde{S}}(y,z)))$$

Example

$X = \{x_1, x_2, x_3\}$ $Y = \{y_1, y_2\}$ $Z = \{z_1, z_2, z_3\}$

Let $\tilde{R}$ be a fuzzy relation

$$\begin{array}{c} \\ x_1 \\ x_2 \\ x_3 \end{array}\begin{array}{c} \begin{matrix} y_1 & y_2 \end{matrix} \\ \begin{bmatrix} 0.5 & 0.1 \\ 0.2 & 0.9 \\ 0.8 & 0.6 \end{bmatrix} \end{array}$$

Let $\tilde{S}$ be a fuzzy relation

$$\begin{array}{c} \\ y_1 \\ y_2 \end{array}\begin{array}{c} \begin{matrix} z_1 & z_2 & z_3 \end{matrix} \\ \begin{bmatrix} 0.6 & 0.4 & 0.7 \\ 0.5 & 0.8 & 0.9 \end{bmatrix} \end{array}$$

Then $R \circ S$, by max-min composition yields,

$$R \circ S = \begin{array}{c} \\ x_1 \\ x_2 \\ x_3 \end{array}\begin{array}{c} \begin{matrix} z_1 & z_2 & z_3 \end{matrix} \\ \begin{bmatrix} 0.5 & 0.4 & 0.5 \\ 0.5 & 0.8 & 0.9 \\ 0.6 & 0.6 & 0.7 \end{bmatrix} \end{array}$$

$$\mu_{\tilde{R}\circ\tilde{S}}(x_1,z_1) = \max(\min(0.5, 0.6), \min(0.1, 0.5))$$
$$= \max(0.5, 0.1)$$
$$= 0.5$$

$$\mu_{\tilde{R}\circ\tilde{S}}(x_1,z_2) = \max(\min(0.5, 0.4), \min(0.1, 0.8))$$
$$= \max(0.4, 0.1)$$
$$= 0.4$$

Similarly,

$$\mu_{\tilde{R}\circ\tilde{S}}(x_1,z_3) = \max(0.5, 0.1) = 0.5$$

$$\mu_{\tilde{R}\circ\tilde{S}}(x_2,z_1) = \max(0.2, 0.5) = 0.5$$

$$\mu_{\tilde{R}\circ\tilde{S}}(x_2,z_2) = \max(0.2, 0.8) = 0.8$$

$$\mu_{\tilde{R}\circ\tilde{S}}(x_2,z_3) = \max(0.2, 0.9) = 0.9$$

$$\mu_{\tilde{R}\circ\tilde{S}}(x_3,z_1) = \max(0.6, 0.5) = 0.6$$

$$\mu_{\tilde{R}\circ\tilde{S}}(x_3,z_2) = \max(0.4, 0.6) = 0.6$$

$$\mu_{\tilde{R}\circ\tilde{S}}(x_3,z_3) = \max(0.7, 0.6) = 0.7$$

Example 6.7

Consider a set $P = \{P_1, P_2, P_3, P_4\}$ of four varieties of paddy plants, set $D = \{D_1, D_2, D_3, D_4\}$ of the various diseases affecting the plants and $S = \{S_1, S_2, S_3, S_4\}$ be the common symptoms of the diseases.

Let $\tilde{R}$ be a relation on $P \times D$ and $\tilde{S}$ be a relation on $D \times S$

For,

$$\tilde{R} = \begin{array}{c} \\ P_1 \\ P_2 \\ P_3 \\ P_4 \end{array}\begin{array}{c} \begin{matrix} D_1 & D_2 & D_3 & D_4 \end{matrix} \\ \begin{bmatrix} 0.6 & 0.6 & 0.9 & 0.8 \\ 0.1 & 0.2 & 0.9 & 0.8 \\ 0.9 & 0.3 & 0.4 & 0.8 \\ 0.9 & 0.8 & 0.1 & 0.2 \end{bmatrix} \end{array} \qquad \tilde{S} = \begin{array}{c} \\ D_1 \\ D_2 \\ D_3 \\ D_4 \end{array}\begin{array}{c} \begin{matrix} S_1 & S_2 & S_3 & S_4 \end{matrix} \\ \begin{bmatrix} 0.1 & 0.2 & 0.7 & 0.9 \\ 1 & 1 & 0.4 & 0.6 \\ 0 & 0 & 0.5 & 0.9 \\ 0.9 & 1 & 0.8 & 0.2 \end{bmatrix} \end{array}$$

Obtain the association of the plants with the different symptoms of the diseases using max-min composition.

Solution

To obtain the association of the plants with the symptoms, $R \circ S$ which is a relation on the sets P and S is to be computed.

Using max-min composition,

$$R \circ S = \begin{array}{c} \\ P_1 \\ P_2 \\ P_3 \\ P_4 \end{array}\begin{array}{c} \begin{matrix} S_1 & S_2 & S_3 & S_4 \end{matrix} \\ \begin{bmatrix} 0.8 & 0.8 & 0.8 & 0.9 \\ 0.8 & 0.8 & 0.8 & 0.9 \\ 0.8 & 0.8 & 0.8 & 0.9 \\ 0.8 & 0.8 & 0.7 & 0.9 \end{bmatrix} \end{array}$$

SUMMARY

- Fuzzy set theory is an effective tool to tackle the problem of uncertainty.
- In crisp logic, an event can take on only two values, either a 1 or 0 depending on whether its occurrence is true or false respectively. However, in fuzzy logic, the event may take a range of values between 0 and 1.
- Crisp sets are fundamental to the study of fuzzy sets. The basic concepts include universal set, membership, cardinality of a set, family of sets, Venn diagrams, null set, singleton set, power set, subset, and super set. The basic operations on crisp sets are union, intersection, complement, and difference. A set of properties are satisfied by crisp sets. Also, the concept of partition and covering result in the two important rules, namely rule of addition and principle of inclusion and exclusion.
- Fuzzy sets support a flexible sense of membership and is defined to be the pair $(x, \mu_{\tilde{A}}(x))$ where $\mu_{\tilde{A}}(x)$ could be discrete or could be described by a continuous function. The membership functions could be triangular, trapezoidal, curved or its variations.

Fundamentals of Genetic Algorithms

Chapter 8

Decision-making features occur in all fields of human activities such as scientific and technological and affect every sphere of our life. Engineering design, which entails sizing, dimensioning, and detailed element planning is also not exempt from its influence.

For example an aircraft wing can be made from aluminium or steel and once material and shape are chosen, there are many methods of devising the required internal structure. In civil engineering also, designing a roof to cover large area devoid of intermediate columns requires optimal designing.

The aim is to make objective function a maximum or minimum, that is, it is required to find an element X_0 in A if it exists such that

$$F(X_0) \le F(X) \text{ for minimization}$$
$$F(X) \le F(X_0) \text{ for maximization} \quad (8.1)$$

The following major questions arise in this process

- Does an optimal solution exist?
- Is it unique?
- What is the procedure?
- How sensitive the optimal solution is?
- How the solution behaves for small changes in parameters?

Since 1940, several optimization problems have not been tackled by classical procedures including:

1. Linear programming
2. Transportation
3. Assignment
4. Nonlinear programming
5. Dynamic programming
6. Inventory
7. Queuing
8. Replacement
9. Scheduling

The classification of optimization techniques is shown in Fig. 8.1. Basically, we have been following traditional search technique for solving nonlinear equations. Figure 8.2 shows the classes of both traditional and nontraditional search techniques. Normally, any engineering problem will have a large number of solutions out of which some are feasible and some are infeasible. The designer's task is to get the best solution out of the feasible solutions. The complete set of feasible solutions constitutes feasible design space and the progress towards the optimal design involves some kind of search within the space (combinatorial optimization). The search is of two kinds, namely deterministic and stochastic.

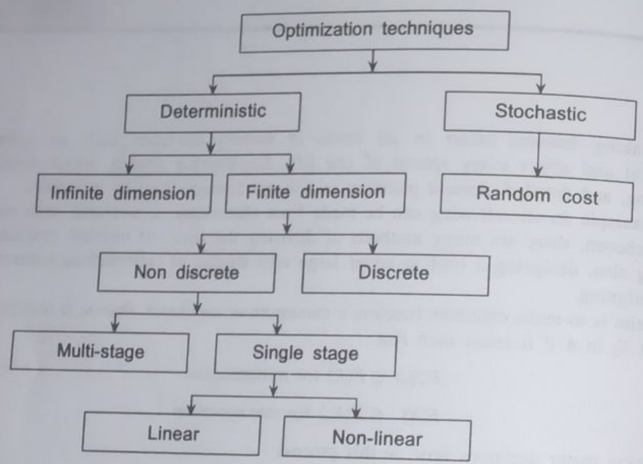


Fig. 8.1 Classification of optimization techniques.

In the case of *deterministic search*, algorithm methods such as steepest gradient methods are employed (using gradient concept), whereas in *stochastic approach*, random variables are introduced. Whether the search is deterministic or stochastic, it is possible to improve the reliability of the results where reliability means getting the result near optimum. A transition rule must be used to improve the reliability. Algorithms vary according to the transition rule used to improve the result.

Nontraditional search and optimization methods have become popular in engineering optimization problems in recent past. These algorithms include:

1. Simulated annealing (Kirkpatrik, et al. 1983)
2. Ant colony optimization (Dorigo and Caro, 1999)
3. Random cost (Kost and Baumann, 1999)
4. Evolution strategy (Kost, 1995)
5. Genetic algorithms (Holland, 1975)
6. Cellular automata (Wolfram, 1994)

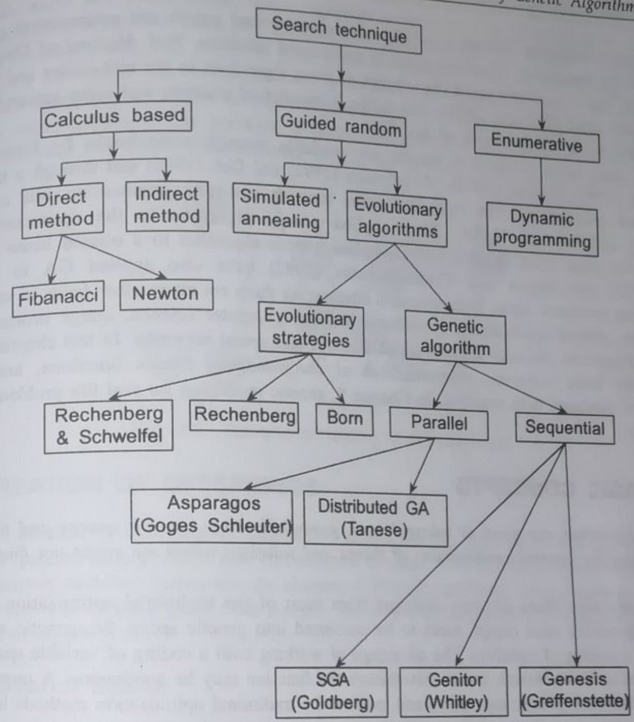


Fig. 8.2 Classes of search techniques.

Simulated annealing mimics the cooling phenomenon of molten metals to constitute a search procedure. *Genetic algorithm* and *evolutionary strategies* mimic the principle of natural genetics and natural selection to construct search and optimization procedures. The collective behaviour that emerges from a group of social insects such as ants, bees, wasps, and termites has been dubbed as *Swarm intelligence*. The foraging of ants has led to a novel algorithm called *Ant colony optimization* for rerouting network traffic in busy telecommunication systems. This method was originally developed by Deneubourg and extended by Dorigo (1999) of Brussels. Random cost method is a stochastic algorithm which moves as enthusiastically uphill as down-hill. The method has no severe problems in escaping from a dead end and is able to find the optima. In this chapter, we discuss the fundamentals of genetic algorithms.

8.1 GENETIC ALGORITHMS: HISTORY

The idea of evolutionary computing was introduced in 1960 by I. Rechenberg in his work

Evolutionary strategies. Genetic algorithms are computerized search and optimization algorithms based on the mechanics of natural genetics and natural selection. Prof. Holland of University of Michigan, Ann Arbor, envisaged the concept of these algorithms in the mid-sixties and published his seminal work (Holland, 1975). Thereafter, a number of students and other researchers have contributed to the development of this field.

To date, most of the GA studies are available through some books by Davis (1991), Goldberg (1989), Holland (1975), Michalewicz (1992) and Deb (1995) and through a number of conference proceedings. The first application towards structural engineering was carried by Goldberg and Samtani (1986). They applied genetic algorithm to the optimization of a ten-member plane truss. Jenkins (1991) applied genetic algorithm to a trussed beam structure. Deb (1991) and Rajeev and Krishnamoorthy (1992) have also applied GA to structural engineering problems. Apart from structural engineering there are many other fields in which GAs have been applied successfully. It includes biology, computer science, image processing and pattern recognition, physical science, social sciences and neural networks. In this chapter, we will discuss the basic concepts, representatives of chromosomes, fitness functions, and genetic inheritance operators with example. In Chapter 9, genetic modelling for real life problems will be discussed.

8.2 BASIC CONCEPTS

Genetic algorithms are good at taking larger, potentially huge, search spaces and navigating them looking for optimal combinations of things and solutions which we might not find in a life time.

Genetic algorithms are very different from most of the traditional optimization methods. Genetic algorithms need design space to be converted into genetic space. So, genetic algorithms work with a coding of variables. The advantage of working with a coding of variable space is that coding discretizes the search space even though the function may be continuous. A more striking difference between genetic algorithms and most of the traditional optimization methods is that GA uses a population of points at one time in contrast to the single point approach by traditional optimization methods. This means that GA processes a number of designs at the same time. As we have seen earlier, to improve the search direction in traditional optimization methods, transition rules are used and they are deterministic in nature but GA uses randomized operators. Random operators improve the search space in an adaptive manner.

Three most important aspects of using GA are:

1. definition of objective function
2. definition and implementation of genetic representation
3. definition and implementation of genetic operators.

Once these three have been defined, the GA should work fairy well beyond doubt. We can, by different variations, improve the performance, find multiple optima (species if they exist) or parallelize the algorithms.

8.2.1 Biological Background

All living organisms consist of cells. In each cell, there is a set of chromosomes which are strings

of DNA and serve as a model for the whole organism. A chromosome consists of genes on blocks of DNA as shown in Fig. 8.3. Each gene encodes a particular pattern. Basically, it can be said that each gene encodes a trait, e.g. colour of eyes. Possible settings of traits (bluish brown eyes) are called *alleles*. Each gene has its own position in the chromosome search space. This position is called *locus*. Complete set of genetic material is called *genome* and a particular set of genes in genome is called *genotype*. The genotype is based on organism's phenotype (development after birth), its physical and mental characteristics such as eye colour, intelligence and so on.

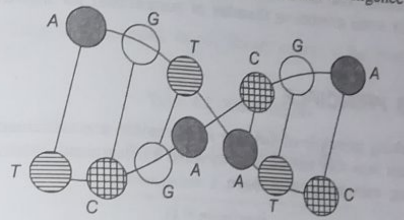


Fig. 8.3 Genome consisting of chromosomes.

8.3 CREATION OF OFFSPRINGS

During the creation of offspring, recombination occurs (due to cross over) and in that process genes from parents form a whole new chromosome in some way. The new created offspring can then be mutated. *Mutation* means that the element of DNA is modified. These changes are mainly caused by errors in copying genes from parents. The *fitness* of an organism is measured by means of success of organism in life.

8.3.1 Search Space

If we are solving some problems, we work towards some solution which is the best among others. The space for all possible feasible solutions is called *search space*. Each solution can be marked by its value of the fitness of the problem. 'Looking for a solution' means looking for extrema (either maximum or minimum) in search space. The search space can be known by the time of solving a problem and we generate other points as the process of finding the solution continues (shown in Fig. 8.4).

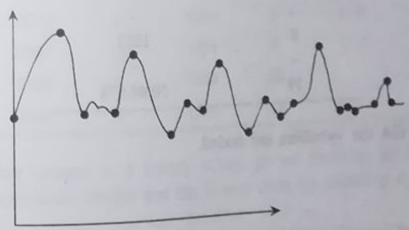


Fig. 8.4 Examples of search space.