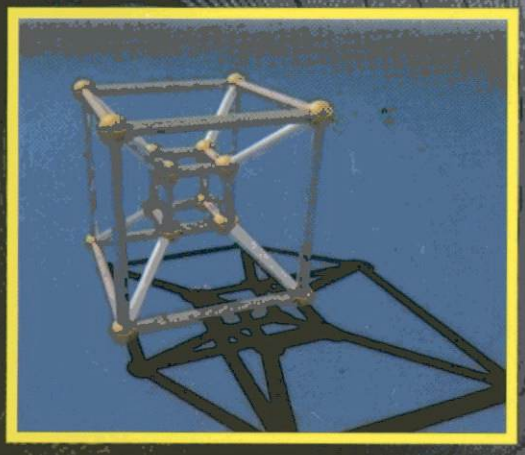
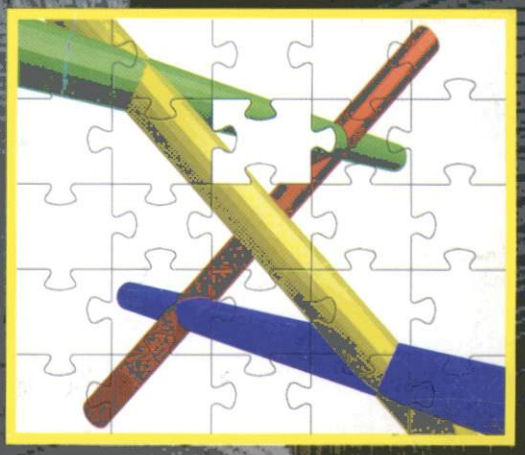
ಕರ್ನಾಟಕ ರಾಜ್ಯ ಮುಕ್ತ ವಿಶ್ವವಿದ್ಯಾನಿಲಯ
ಮಾನಸಗಂಗೋತ್ರಿ, ಮೈಸೂರು – ೫೭೦ ೦೦೬

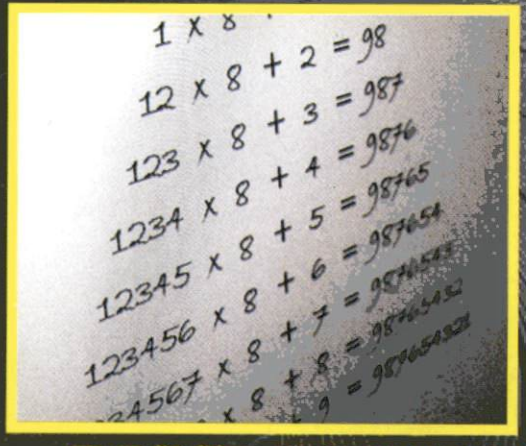
KARNATAKA STATE OPEN UNIVERSITY
Manasagangotri, Mysore - 570 006

M.Sc. Computer Science

First Semester







COURSE : 1 DISCRETE MATHEMATICS
MODULE : 1 - 6
MSCS - 501

ಉನ್ನತ ಶಿಕ್ಷಣಕ್ಕಾಗಿ ಇರುವ ಅವಕಾಶಗಳನ್ನು ಹೆಚ್ಚಿಸುವುದಕ್ಕೆ ಮತ್ತು ಶಿಕ್ಷಣವನ್ನು ಪ್ರಜಾತಂತ್ರೀಕರಿಸುವುದಕ್ಕೆ ಮುಕ್ತ ವಿಶ್ವವಿದ್ಯಾನಿಲಯ ವ್ಯವಸ್ಥೆಯನ್ನು ಆರಂಭಿಸಲಾಗಿದೆ.

ರಾಷ್ಟ್ರೀಯ ಶಿಕ್ಷಣ ನೀತಿ 1986

The Open University System has been initiated in order to augment opportunities for higher education and as instrument of democrating education.

National Educational Policy 1986

ವಿಶ್ವ ಮಾನವ ಸಂದೇಶ

ಪ್ರತಿಯೊಂದು ಮಗುವು ಹುಟ್ಟುತ್ತಲೇ – ವಿಶ್ವಮಾನವ, ಬೆಳೆಯುತ್ತಾ ನಾವು ಅದನ್ನು 'ಅಲ್ಪ ಮಾನವ'ನನ್ನಾಗಿ ಮಾಡುತ್ತೇವೆ. ಮತ್ತೆ ಅದನ್ನು 'ವಿಶ್ವಮಾನವ'ನನ್ನಾಗಿ ಮಾಡುವುದೇ ವಿದ್ಯೆಯ ಕರ್ತವ್ಯವಾಗಬೇಕು.

ಮನುಜ ಮತ, ವಿಶ್ವ ಪಥ, ಸರ್ವೋದಯ, ಸಮನ್ವಯ, ಪೂರ್ಣದೃಷ್ಟಿ ಈ ಪಂಚಮಂತ್ರ ಇನ್ನು ಮುಂದಿನ ದೃಷ್ಟಿಯಾಗಬೇಕಾಗಿದೆ. ಅಂದರೆ, ನಮಗೆ ಇನ್ನು ಬೇಕಾದುದು ಆ ಮತ ಈ ಮತ ಅಲ್ಲ; ಮನುಜ ಮತ. ಆ ಪಥ ಈ ಪಥ ಅಲ್ಲ ; ವಿಶ್ವ ಪಥ. ಆ ಒಬ್ಬರ ಉದಯ ಮಾತ್ರವಲ್ಲ; ಸರ್ವರ ಸರ್ವಸ್ತರದ ಉದಯ. ಪರಸ್ಪರ ವಿಮುಖವಾಗಿ ಸಿಡಿದು ಹೋಗುವುದಲ್ಲ; ಸಮನ್ವಯಗೊಳ್ಳುವುದು. ಸಂಕುಚಿತ ಮತದ ಆಂಶಿಕ ದೃಷ್ಟಿ ಅಲ್ಲ; ಭೌತಿಕ ಪಾರಮಾರ್ಥಿಕ ಎಂಬ ಭಿನ್ನದೃಷ್ಟಿ ಅಲ್ಲ; ಎಲ್ಲವನ್ನು ಭಗವದ್ ದೃಷ್ಟಿಯಿಂದ ಕಾಣುವ ಪೂರ್ಣದೃಷ್ಟಿ.

ಕುವೆಂಪು

Gospel of Universal Man

Every Child, at birth, is the universal man. But, as it grows, we trun it into "a petty man". It should be the function of education to turn it again into the enlightened "universal man".

The Religion of Humanity, the Universal Path, the Welfare of All, Reconciliation, the Integral Vision - these ***five mantras*** should become View of the Future. In other words, what we want henceforth is not this religion or that religion, but the Religion of Humanity; not this path or that path, but the Universal Path; not the well-being of this individual or that individual, but the Welfare of All; not turning away and breaking off from one another, but reconciling and uniting in concord and harmony; and above all, not the partial view of a narrow creed, not the dual outlook of the material and the spiritual, but the Integral Vision of seeing all things with the eye of the Divine.

Kuvempu

Karnataka State Open University

Manasagangothri, Mysore-570006

M. Sc. (Computer Science)

MSC-501: DISCRETE MATHEMATICS

Course Design and Editorial Committee

Prof. K.S.Rangappa
Vice-Chancellor & Chairperson
Karanataka State Open University
Manasagangotri, Mysore – 570 006

Prof. Jagadeesha
Dean (Academic) & Convenor
Karnataka State Open University
Manasagangotri, Mysore– 570 006

Head of the Department - Incharge

Prof. Jagadeesha
DOS in Commerce
and Management
Karnataka State Open University
Manasagangothri
Mysore-570 006

Course Co-Ordinator

Smt. Sumati. R. Gowda
BE(CS & E)., MSc(IT)., MPhil (CS).,
Lecturer
DOS in Information Technology
and Computer Science
Karnataka State Open University
Manasagangothri
Mysore-570 006

Course Writers — **Module 1 - 6** — **Units 1-24**

Dr. Narasimhamurthy S.K
Associate Prof.
Kuvempu University. Shimoga

Modules: 1 & 2

Dr Lalitha Rangarajan
Reader, DOS in CS,
UOM, Mysore

Modules: 4

Dr Sharada
DOS in CS,
UOM, Mysore

Modules: 3 ,5 & 6

Publisher

Registrar
Karnataka State Open University
Manasagangotri, Mysore - 6.

Developed by Academic Section, KSOU, Mysore
Karnataka State Open University, 2010

Further information on the Karnataka State Open University Programmes may obtained from the University's office at Manasagangotri, Mysore-6

Printed and Published on behalf of Karnataka State Open University. Mysore-6 by the **Registrar (Administration)**

Preface

All advanced texts in Computer Science contain discussions in brief on topics of Mathematics necessary for understanding of the subject. These discussions may be either in the beginning chapters of the text or in the appendices. However the discussions in the Computer Science texts are very brief and are barely adequate. Concepts of Mathematics are important for a good understanding of Computer Science subjects. This course on Discrete Mathematics covers almost all concepts recommended to be learnt by a Computer Science professional.

The course is organized into six modules each with four units. Each module covers a key area of Discrete Mathematics.

Module 1 is about Mathematical Logic and Set Theory. The module begins with basics of logics right from well formed formulae, conversion rules from one normal form to other, to discussions on predicate calculus. Later part of the module introduces set theory and module is concluded with a very important result of Computer Science namely Pigeon Hole principle.

Module 2 is on relations covering equivalence, compatibility and composition of relations and Warshall's algorithm to find transitive closure of a relation. .

Module 3 is on recurrence relation and generating functions. Concept of recurrence is important in Computer Science. This module contains discussions on linear recurrence relations, its solutions, functions and their types such as permutation, hashing and recursive functions.

Many applications on Computer Science are based on Graph Theory. In module 4 Graph Theory is introduced. Representations of graphs, concepts like paths, circuits, trees and binary search trees are discussed here.

Groups are important algebraic structures useful for Computer Scientists. Module 5 discusses Groups and some key results such as sub groups, homomorphism and isomorphism. This module is concluded with discussions on error detection and correction of binary codes.

In the last module of this course, namely module 6, elementary concepts in probability theory are discussed. Distributions of random variables such as discrete, continuous and joint are elaborated in this module.

List of Notations and Symbols:

Symbols	Meaning
$p \wedge q$	p and q
$p \vee q$	p or q
$\sim p$	not p
$p \rightarrow q$	if p then q
$p \leftrightarrow q$	p if and only if q
$P \equiv Q$	P is logically equivalent to Q
$\forall$	for all
$\exists$	there exits
$p \oplus q$	p xor q(that is,p or q but not both)
p/q	p nand q (that is, not both p and q)
$pNORq$	not p or q
$\therefore$	therefore
$\{....\}$	'The set of....'
$x \in A$	x is an element of A
ϕ	null set
$A \subseteq B$	A is subset of B
$A \cap B$	intersection of A and B
$A \cup B$	union of B
$\bar{A}$	complement of A
$\mid A \mid$	cardinality of A
$P(A)$	power of a set A
$A \times B$	cartesian product of A and B
A^n	n-fold cartesian product of A itself
$E(x)$	equivalent class containing x
$\{x_1....x_n\}$	order n-tuple

Module - 1: MATHEMATICAL LOGIC AND SET THEORY

Unit 1: MATHEMATICAL LOGIC

Structure

1.0 Objectives

After studying this unit you will be able to:

- Understand the basic concepts & Notations of Mathematical Logic.
- Know the different types of connections & their uses.
- Study the importance of WFF, Tautology & logical implication.
- Study the application of logical equivalences in computer science.
- Define the Duality structures & Normal Forms.

1.1 Introduction

The dictionary meaning of the term ***"Logic"*** is the Science of reasoning correctly. The rules of logic is precise meaning to mathematical statements. These rules are used to distinguish between valid and invalid mathematical arguments. Logical reasoning is used in mathematics to prove theorems, in Computer Science to verify the correctness of program and to prove theorems, in natural and physical sciences to draw conclusions from experiments and in the Social Sciences and in our everyday lives to solve a multitude of problems. Include, we are constantly using reasoning.

A symbolic language has been developed over the past two centuries to express the principles of logic in precise and unambiguous terms. Logic expressed in such a language has come to be called "*Symbolic Logic*"or "*Mathematical Logic*".

Greeks were first to study and develop the subject which was subsequently modified and used in teaching throughout the middle ages. The process of reaching general theory/conclusions based on the results of several individual case studies is termed as 'Inductive Reasoning' or 'Inductive Logic' in Philosophy. The basic foundation of social, natural and physical sciences underline in such a process. It may be noted, however that the absolute truth of the general theory/conclusions based on inductive reasoning in unlikely to be an tested as it is subjected to continuous modification or alteration.

The reverse process of arguing from a general conclusion to a specific one is termed as 'Deductive Logic'. In deductive logic, results/conclusions on ideas based on observations, experiments, theorems and definitions from a system or a structure which does not preclude the existence of deductive test for its verification. Such a procedure is used by mathematics is proving mathematical statements or theorems, the truth of which could be further tested. In Mathematics, a mathematical statement can be either true or false but not both.

Based on such ideas *George Boole*(1815-1864) founded mathematical theory of logic which along with pioneering work of Claude Shanon on switching circuits and new technological developments resulted in the design of a digital computer. Here we discuss some of basic ideas of mathematical logic and formulate the algebra propositions.

1.2 Mathematical logic

Mathematical logic is the Science dealing with the method of reasoning. Reasoning plays an important role in every area of knowledge, particularly mathematics.

Mathematical Logic has now become a core subject of study for every student of mathematical sciences. Some basic notations of the subject are introduced in this unit. The student is already familiar with the terminology and notation employed in this subject.

Mathematical Logic finds applications in many areas of computing. The laws of logic are employed in the design of the digital circuitry in a computer. Logical expressions occur as conditions in the control structures in algorithms and computer programs and databases. Expert systems employing knowledge-based software use rules of logical inference to draw conclusions from known facts.

1.3 Statements

The fundamental object we work with in arithmetic are numbers. In a similar way, the fundamental objects in logic are propositions.

Definition 1.3.1. *A proposition is a statement(declaration) which, in a given context, can be said to be either true or false but not both.*

Propositions are usually denoted by small letters such as p,q,r,s,....

Examples:

1. The following are propositions,
   (a) Latha likes Science. (b) 3+4=5
2. The following are not propositions
   (a)Let me go! (Exclamation)
   (b)x+3=5 (x is unknown)

Definition 1.3.2 Truth value: *The truth or the falsity of a proposition is called its truth value.* If a proposition is true, we will indicate its truth value by the symbol T and if it is false by the symbol F.

Definition 1.3.3. Truth Table: *A table showing the truth values of a statement is called a truth table.* It is a compact way of listing symbols to show all possible truth values for a set of sentences.

Example: If we denote the proposition "The number 3 is a prime number" by p, then the truth value of p is T. Similarly, we denote the proposition "Every rectangle is a square" by q, then the truth value of q is F.

1.4 Notations

Statements can be connected by the words like 'not', 'and', 'or', 'conditional', 'bi-conditional' etc. These words are known as logical connectives. The statements which do not contain any of the connectives are called atomic statements or simple statements.

The common connectives are : 'negation', 'and', 'or', 'if....then', 'if and only if' and 'equivalence'. We use the following notations to the corresponding connections.

Connectives	Notations
Negation	$\sim$
and	$\wedge$
or	$\vee$
if-then	$\rightarrow$
if and if only if	$\leftrightarrow$
equivalence	$\equiv$ (*or* $\Leftrightarrow$)

1.5 Connectives

Definition 1.5.1. Negation ($\sim$): *A proposition obtained by inserting the word 'not' at an appropriate place in a given proposition is called the negation of the given proposition.*
The negation of a proposition p is denoted by $\sim$p(read not p), the symbol ' $\sim$' denoting the word 'not'.

Examples:

1. Let the proposition "2 is a prime number" be denoted by p, i.e.,
p: 2 is a prime number.
Then $\sim$ p: 2 is not a prime number.

2. p: Every rectangle is a square. Then $\sim$ p: Not every rectangle is a square.

Truth table for Negation

p	$\sim$p
T	F
F	T

Definition 1.5.2. Conjunction ($\wedge$): *Conjunction is a compound statement formed by using the word 'and' to combine two simple sentences.*
If p and q represent two simple statements, the conjunction of p and q is written symbolically as $p \wedge q$.

Examples: Consider the statements,

1. p: Socrates was a Greek philosopher. q: Euclid was a mathematician.
Their conjunction is given by
$p \wedge q$: Socrates was a Greek philosopher and Euclid was a mathematician.

2. p: Pradeep likes Algebra. q: Aveesha likes Physics.
$p \wedge q$: Pradeep likes Algebra and Aveesha likes Physics.

The following rule is adopted in deciding the truth value of a conjunction.

Truth table for Conjunction

p	q	$p \wedge q$
T	T	T
T	F	F
F	T	F
F	F	F

Definition 1.5.3. Disjunction($\vee$): *Disjunction is a compound sentence formed by using the word 'or' to combine two simple statements.*

Thus when p and q represent simple statements, the disjunction p or q is written symbolically as $p \vee q$.

Examples :

1. p: Vasantha likes pepperoni pizza for lunch. q: Vasantha likes mushroom pizza for lunch.

Their disjunction is given by,

$p \vee q$: Vasantha likes pepperoni pizza for lunch or Vasantha likes mushroom pizza for lunch.

Note 1: Vasantha could like pepperoni pizza or mushroom pizza, or both for lunch. In other words, the connective 'or' is used in the inclusive sense and / or to mean at least one, may be both. Such a disjunction is an *Inclusive disjunction.*

2. Consider the statements

p: Ajith will play basketball at 3pm today. q: Ajith will go to a matinee at 3pm today. Then

$p \vee q$: Ajith will play basketball at 3pm today or Ajith will go to a matinee at 3pm today.

Note 2: Ajith cannot play basketball and go to a matinee at the same time. So the word 'or' is used in the exclusive sense to mean at least one, but not both. Such a disjunction is an *exclusive disjunction.*

The following rule is adopted in deciding the truth value of disjunction.

Truth table for Disjunction

p	q	$p \vee q$
T	T	T
T	F	T
F	T	T
F	F	F

Definition 1.5.4 Conditional($\rightarrow$): *A compound proposition obtained by combining two given propositions by using the words 'if' and 'then' at appropriate places is called a conditional proposition or just a conditional.*

Examples:

1. Let p: I study. q: I will pass the test.

Then, $p \rightarrow q$: If I study, then I will pass the test.

2. Let p: Kavya learns discrete mathematics. q: Kavya will find a good job.

Then, $p \rightarrow q$: If Kavya learns discrete mathematics then she will find a good job.

A conditional is sometimes called an *implication.* Thus, we may also read the symbol for conditional $p \rightarrow q$ as p implies q.

The antecedent usually follows the word 'if' and the consequent usually follows the word 'then'.

The truth table for the conditional $p \rightarrow q$ is given below.

Truth table for Conditional

p	q	$p \rightarrow q$
T	T	T
T	F	F
F	T	T
F	F	T

Definition 1.5.5. Bi conditional($\leftrightarrow$): *Let p and q be two propositions. Then the conjunction of the conditionals $p \rightarrow q$ and $q \rightarrow p$ is called the bi conditional of p and q, it is denoted by $p \leftrightarrow q$.*

Thus $p \leftrightarrow q$ is the same as $(p \rightarrow q) \wedge (q \rightarrow p)$. As such $p \leftrightarrow q$ is read as 'if p then q and if q then p'.

The following is the truth table for $p \leftrightarrow q$

Truth table for Bi conditional

p	q	$p \rightarrow q$	$q \rightarrow p$	$p \leftrightarrow q$
T	T	T	T	T
T	F	F	T	F
F	T	T	F	F
F	F	T	T	T

Examples:

1. 4+4=8 if and only if 2+5=7.

2. Bangalore in India if and only if 2+8=10.

3. Let p: You can take the flight. q: You buy a ticket.

Then $p \leftrightarrow q$: You can take the flight if and only if you buy a ticket.

Definition 1.5.6. Converse, Inverse and Contrapositive Propositions: *If $p \rightarrow q$ is a conditional statement then*

$q \rightarrow p$ *is called converse.*

$\sim p \rightarrow \sim q$ *is called its inverse.*

$\sim q \rightarrow \sim p$ *is called its contrapositive.* Truth values of these propositions are given below

Table for converse,

p	q	$p \to q$	$q \to p$
T	T	T	T
T	F	F	T
F	T	T	F
F	F	T	T

Table for Inverse,

p	q	$\sim p$	$\sim q$	$\sim p \to \sim q$
T	T	F	F	T
T	F	F	T	F
F	T	T	F	F
F	F	T	T	T

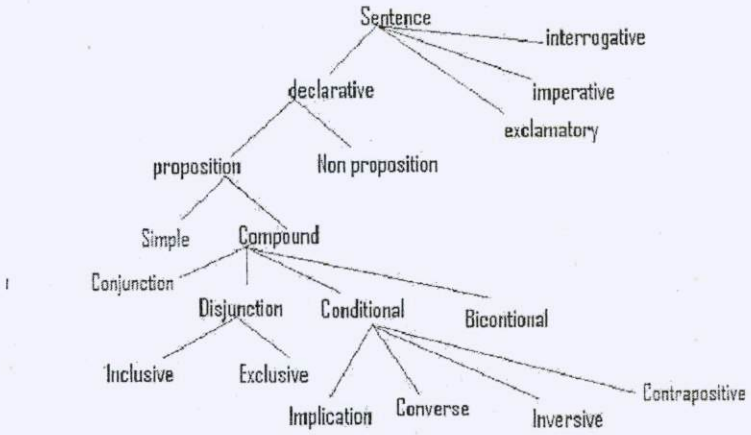
Table for Contrapositive

p	q	$\sim p$	$\sim q$	$\sim q \to \sim p$
T	T	F	F	T
T	F	F	T	F
F	T	T	F	T
F	F	T	T	T

Example: Let p: It's -6^0. q: It's cold.

Then the converse of $p \to q$ is $q \to p$: If it is cold, then it's -6^0.

Various types of sentences and propositions can be summarized like this-



Definition 1.5.7. Other Connectives: *We now introduce the connectives NAND, NOR which have useful applications in the design of computers.*

The word NAND is a combination of NOT and AND where NOT stands for negation and AND stands for the conjunction. It is denoted by the symbol $\uparrow$.

If P and Q are two formulas then $p \uparrow q \to \sim (p \wedge q)$. The connective $\uparrow$ has the following equivalence.

$$p \uparrow p \leftrightarrow (\sim (p \wedge p)) \leftrightarrow (\sim p \vee \sim p) \Leftrightarrow (\sim p)$$
$$(p \uparrow q) \uparrow (p \uparrow q) \leftrightarrow \sim (p \uparrow q) \leftrightarrow (p \wedge q)$$
$$(p \uparrow p) \uparrow (q \uparrow q) \leftrightarrow \sim (p \uparrow \sim q) \Leftrightarrow \sim (\sim p \wedge \sim q) \leftrightarrow (p \vee q)$$

The connective NAND is commutative but not associative i.e, $P \uparrow Q \leftrightarrow Q \uparrow P$

NAND is not associative, since $P \uparrow (Q \uparrow R) \leftrightarrow \sim P \vee (Q \wedge R)$ and $(P \uparrow Q) \uparrow R \leftrightarrow \sim (P \wedge Q) \vee \sim R)$.

Connective NOR is a combination of "NOT" and "OR", where NOT stands for negation and OR stands for the disjunction. The connective NOR is denoted by the symbol $\downarrow$, and is called joint $P \downarrow Q$ is read as "Neither P nor Q".

The connective NOR has following equivalence,

$$P \downarrow P \leftrightarrow (\sim (p \vee P)) \leftrightarrow (\sim P \wedge \sim P) \Leftrightarrow (\sim P)$$
$$(P \downarrow Q) \downarrow (P \downarrow Q) \leftrightarrow \sim (P \downarrow Q) \leftrightarrow (P \vee Q)$$

$$(P \downarrow P) \downarrow (Q \downarrow Q) \leftrightarrow \sim P \uparrow \sim Q \Leftrightarrow (P \wedge Q)$$

The connective $\downarrow$ is commutative but not associative.

The connective $\wedge$, $\vee$ and $\sim$ can be expressed in terms of the connective as follows-

(a) $\sim p \equiv p \downarrow p$

(b) $\sim q \equiv q \downarrow q$

(c) $p \wedge q \equiv (p \downarrow p) \downarrow (q \downarrow q)$

(d) $p \vee q \equiv (p \downarrow p) \downarrow (q \downarrow q)$

Examples:

1. Write a formula which is equivalent to the formula $P \wedge (Q \leftrightarrow R)$ and contains the connective NAND ($\uparrow$) only obtain an equivalent formula which contains the connective NOR ($\downarrow$) only.

Solution:

$$\begin{aligned}
P(Q \leftrightarrow R) &\equiv P \wedge ((Q \rightarrow R) \wedge R \rightarrow Q) \\
&\equiv P \wedge ((\sim Q \vee R) \wedge (\sim R \vee Q)) \quad (because\ P \rightarrow Q \equiv \sim P \vee Q) \\
&\equiv P \wedge (\sim Q \vee R) \wedge \sim (R \vee \sim Q) \quad (By\ De-Morgan's\ law) \\
&\equiv P \wedge ((Q \uparrow \sim R) \vee \uparrow (R \uparrow \sim Q)) \\
&\equiv P \sim ((Q \uparrow \sim R) \uparrow (R \uparrow \sim Q)) \\
&\equiv \sim (P \uparrow \sim ((Q \uparrow \sim R) \uparrow (R \uparrow \sim Q))) \quad P \wedge (Q \leftrightarrow R) using\ NOR\ only \\
P \wedge (Q \leftrightarrow R) &\equiv P \vee ((Q \rightarrow R) \wedge (R \rightarrow q)) \\
&\equiv P \wedge ((\sim Q \vee R) \wedge (\sim R \vee Q)) \\
&\equiv P(\sim (Q \downarrow R) \wedge \sim (\sim R \wedge Q)) \\
&\equiv P \wedge \sim (\sim Q \downarrow R) \vee (\sim R \downarrow Q)) \\
&\equiv \sim (\sim P \vee (\sim (\sim Q \downarrow R)))) \downarrow (\sim R \downarrow Q) \\
&\equiv \sim (P \downarrow P) \vee (\sim (\sim Q \downarrow R) \downarrow (\sim R \downarrow Q)) \\
&\quad (P \downarrow P) \downarrow \sim (\sim Q \downarrow R) \downarrow (\sim R \downarrow Q).
\end{aligned}$$

2. Show that the connectives $\sim$ and $\wedge$ may be expressed interms of the connective $\downarrow$ as follows,

(i) $\sim p \equiv ? \downarrow p$ (ii) $p \wedge q \equiv (p \downarrow q) \downarrow (q \downarrow q)$

Solution: (i) We have

p	$\sim p$	$p \downarrow q$
T	F	F
F	T	F

form the above table, it is clear that $\sim p$ and $p \downarrow q$ have same truth values. *therefore* $\sim p \equiv p \downarrow q$

(ii)

p	q	$p \wedge q$	$p \downarrow q$	$q \downarrow q$	$(p \downarrow q) \downarrow (q \downarrow q)$
T	T	T	F	F	T
T	F	F	F	F	F
F	T	F	T	F	F
F	F	F	T	T	F

Since the columns $p \wedge q$ and $(p \downarrow q) \downarrow (q \downarrow q)$ have same truth values,

$\therefore p \wedge q \equiv (p \downarrow p) \downarrow (p \downarrow q)$.

Solved Examples:

1. Let us consider the following propositions-

p: $\sqrt{2}$ is an irrational number. q: 9 is a prime number. r: All triangles are equilateral. Then,

$p \wedge q$: $\sqrt{2}$ is an irrational number and 9 is a prime number.

$q \wedge r$: 9 is a prime number and all triangles are equilateral.

$r \wedge p$: All triangles are equilateral and $\sqrt{2}$ is an irrational number.

Here, p is true and q and r are false, since p is true and q is false, $p \wedge q$ is false. since both q and r false, $q \wedge r$ is false, since r is false and p is true, $r \wedge p$ is false.

If we consider one more proposition, s: All squares are rectangles.

$p \wedge s$: $\sqrt{2}$ is an irrational number and all squares are rectangles.

Here both p and s are true. Hence $p \wedge s$ is true.

2. What are the contrapositive, converse and inverse of the conditional statements "The home team wins whenever it is raining?"

Solution: Because "q whenever p" is one of the ways to express the conditional statement $p \rightarrow q$ the original statement can be written as,

"If it is raining, then the home team wins"

consequently, *contrapositive* is,

"If the home team does not win, then it is not raining".

The *converse* is,

"If it is not raining, then the home team does not win".

The *inverse* is,

"If it is not raining, then the home team does not win".

Only contrapositive is equivalent to the original statement.

3. Show that the truth values of the following statements are independent of the truth values of their components

(i) $(p \wedge (p \to q)) \to q$ (ii) $(p \to q) \leftrightarrow (\sim p \vee q)$

p	p	$q \to q$	$r = p \wedge (p \to q)$	$r \to q$
T	T	T	T	T
T	F	F	F	T
F	T	T	F	T
F	F	T	F	T

p	p	$r = p \to q$	$\sim p$	$s = \sim p \wedge q$	$r \to s$
T	T	T	F	T	T
T	F	F	F	F	T
F	T	T	T	T	T
F	F	T	T	T	T

The last columns in both of the tables shows that the truth value of the given (corresponding) statement is T irrespective of what the truth values of its components are. This is what we had to show.

1.6 Well Formed Formulas

Statement formulas contain one or more simple statements and some connectives. If p and q are any two statements, then $p \vee q$, $p \wedge \sim p \wedge q$ are some statement formulas derived from the statements variables p and q where p and q are called components of the statement formulas. A statement formula has no truth value. It is only when the statement variables in a statement formula are replaced by definite statement that we get a statement, which has a truth value that depends upon the truth values of its statements used in replacing the variables. A statement formula is a string consisting of variables, parentheses and connective symbols. A statement formula is called a *Well formed formulas*(WFF) if it can be generated by the the following rules-

1. A statement variable p standing alone is a well formed formulas.
2. If p is a well formed formula.
3. If p and q are well formed formulas, then $p \wedge q$, $p \vee q$, $p \to q$ and $p \leftrightarrow q$ are well formed formulas.
4. A string of symbols is a well formed if and only if it is obtained by finitely many applications of the rules 1, 2 and 3.

According to the above recursion definition of a well formed formula, the formulas, $\sim (p \vee q)$, $\sim p \wedge q$, $p \to (p \wedge q)$ are well formed formulas.

A statement formula is not a statement and has no truth values. But if we substitute definite statements in place of variables in a given formula we get a statement. The truth value of this resulting statement depends upon the truth values of the statement substituted

for the variables, which appears as one of the entries in the final column of the truth table constructed. Therefore, the truth value of a well formed formula is the summary of truth values of the resulting statements for all possible assignments of truth values of the variables appearing in the formula. The final column entries of the truth table of a well formed formula gives the truth value of the formulas.

1.7 Tautology and Contradiction

Propositional logic is to study of propositions and the propositional connectives. It is the study not only of one particularly interpretation of a formula but also of what can be deduced about all interpretations of a formula of particular interest are those formulas that are true "by virtue of pure logic". Definitions 1.7.1 and 1.7.2 captures the notion of "true by virtue of logic", at least as closely as is possible from the standpoint of propositional logic.

Definition 1.7.1 *A statement formula (well formed formula) whose truth value is T for all possible assignments of truth values to the propositional variables is a Tautology.*

Examples:

1. $(p \wedge q) \rightarrow (q \vee p)$ is a tautology.

2. $p \vee \sim p$ is a tautology.

3. $(p \wedge \sim q) \wedge (\sim p \vee \sim r \vee q) \wedge (p \vee \sim p)$ is not a Tautology.

Definition 1.7.2. *A statement formula whose truth value is F for all possible assignments of truth values to the propositional variables is called a contradiction or absurdity.* **Example:**

1. $p \wedge \sim p$ is contradiction.

Note: If α is a well formed formula, $-\alpha$ is a tautology then α is an absurdity. i.e, α is an absurdity if and only if $\sim \alpha$ is a tautology.

Definition 1.7.3. *A statement formula which is neither a tautology nor a contradiction is called a contingency.*

Examples:

1. $(p \rightarrow q) \wedge p \wedge q$ is a contingency.

2. $q \vee (\sim q \wedge p)$ is a contingency

3. $(q \wedge p) \vee (q \vee \sim p)$ is a contingency.

Solved Examples:

1. Prove that any propositional $p \wedge \sim p$ is a tautology and the compound proposition $p \wedge \sim p$ is a contradiction.

Solution: Let us construct the following truth table giving the truth values of $p \vee \sim p$ and $p \wedge \sim p$ for all possible truth values of p (*and* $\sim p$)

p	q	$p \vee \sim p$	$p \wedge \sim p$
T	F	T	F
F	T	T	F

We note that $p \wedge \sim p$ is always false, hence it is a contradiction.

2. Prove that for any propositions p and q, the compound proposition $(\sim q) \wedge (p \rightarrow q)] \sim p$ is a tautology.

Solution: Let us prepare the following truth table

p	q	$p \rightarrow q$	$r = (\sim q) \wedge (p \rightarrow q)$	$r \rightarrow \sim p$
T	T	T	F	T
T	F	F	F	T
F	T	T	F	T
F	F	T	T	T

We observe that the proposition $r \rightarrow \sim p$ where $r = (\sim q) \wedge (p \rightarrow q)$ is always true. Therefore, this proposition is a tautology.

1.8 Logical Implication

We state the following theorem,

Theorem 1.8.1. *Let $P(p_1, p_2....)$ and $Q(p_1, p_2....)$ be two propositions. Then the following conditions are equivalent.*

1. $\sim P(p_1, p_2....) \vee Q(p_1, p_2....)$ is a tautology.
2. $P(p_1, p_2....) \wedge Q(p_1, p_2....)$ is a contradiction.
3. $P(p_1, p_2....) \rightarrow Q(p_1, p_2....)$ is a tautology.

Definition 1.8.1 *A proposition $P(p_1, p_2, ...)$ is said to logically imply a proposition $Q(p_1, p_2....)$ if one of the conditions in the Theorem 1.8.1 holds.*

If $P(p_1, p_2....)$ logically implies $Q(p_1, p_2, ...)$, then we symbolically denote it by writing $P(p_1, p_2, ...) \Rightarrow Q(p_1, p_2, ...)$.

Examples:

1. $(p \wedge q \wedge (p \vee q))$ is a contradiction.

Hence $p \wedge q \Rightarrow p \vee q$

2. $(p \wedge q) \wedge (q \rightarrow r) \rightarrow (p \rightarrow r)$ is tautology.

Hence $(p \rightarrow q) \wedge (q \rightarrow r) \Rightarrow (p \rightarrow r)$.

Theorem 1.8.2. *The relation in proposition defined by $P(p_1, p_2, ...) \Rightarrow Q(p_1, p_2, ...)$ is reflexive, antisymmetric and transitive.*

From the section 1.5.4 we may recall that no restrictions were imposed on the choice of p and q while defining $p \rightarrow q$. In other words, $p \rightarrow q$ can be constructed for any two propositions p and q. As such, we may consider, for example, the propositions.

p: A square is a rectangle. q: 3 is a prime number.
and get the conditional,

$p \rightarrow q$: If a square is a rectangle, then 3 is a prime number.

We note that here p is true and q is true, hence by the truth table for conditional Table 1.5.1 $p \rightarrow q$ is also true. But the question, is does this conditional $p \rightarrow q$ make any sense? The answer is no. Because there is no consistency in the statement $p \rightarrow q$ (all though it is logically true!).
As another example, consider the propositions

p: 4 is an odd number. q: Bangalore is not in Karnataka.
Both of which are false. We note normally deal with conditions such as the ones conditional

$p \rightarrow q$: If 4 is an odd number, then Bangalore is not in Karnataka, makes no sense, but it is logically true!.

We do not normally deal with conditionals such as the ones considered above. Our major interest lies in conditionals $p \rightarrow q$ where p and q are related in some way so that the truth value of q depends on the truth value of p or viceversa. Such conditionals are called hypothetical or implicative statements, or implications. In the implication $p \rightarrow q$ the component p is called the prime, antecedent or hypothesis and the component q is called the consequent or conclusion.

Following are some examples of implication.

1. If a triangle is equilateral, then it is isosceles.
2. If a polygon is a square, then it is not a triangle.
3. If a number x is a multiple of 6, then it is multiple of 3.

When the implication $p \rightarrow q$ is true, we say that p(logically) implies q. This is symbolically written as $p \Rightarrow q$, the symbol $\Rightarrow$ denoting the word implies. When the implication $p \rightarrow q$ is false we say that p does not imply q. This is symbolically written as $p \Rightarrow q$, the symbol $\Rightarrow$ denoting the phrase does not imply.

It is to be emphasized that $p \Rightarrow q$ is always a true statement, and this statement means that $p \rightarrow q$ is a tautology.

1.9 Logical Equivalence

Definition 1.9.1: *Two expressions (composed of the same variables) are logical equivalent if they have the same truth values for every combination of the variables.*

Definition 1.9.2: *Two propositions P and Q are said to be logically equivalent or simply equivalent if $P \rightarrow Q$ is a tautology.*

Logically equivalence, on the other hand, is a relationship between two logically expressions. The two concepts are related in the following way, two expressions A and B are logically equivalent if and only if the $A \leftrightarrow B$ is a tautology.

Some important questions about logically equivalence arise when we consider expressions of the form $p \to q$. Such expressions are called implications.

If "P is equivalent to Q", then we can represent equivalence by writing "$P \Leftrightarrow Q$". Which can also be written as $P \Leftrightarrow Q$. The symbol "$\Leftrightarrow$" is not a connective. Usually we drop quotation marks. If P and Q are logically equivalent we also write $P \equiv Q$ and read "P is equivalent to Q". There are a number of fundamental equivalences which are useful in proofs.

1.9.1. Laws of Logic:

1. Idempotent laws: (a) $p \vee p \equiv p$ (b) $p \wedge p \equiv p$
2. Commutative laws: (a)$p \vee q \equiv q \vee p$ (b)$p \wedge q \equiv q \wedge p$
3. Associative laws: (a) $(p \vee q) \vee r \equiv p \wedge (q \vee r)$ (b) $(p \wedge q) \wedge r \equiv p \vee (q \wedge r)$
4. Distributive laws: (a) $p \vee (q \wedge r) \equiv (p \vee q) \wedge (p \vee r)$ (b) $p \wedge (q \vee r) \equiv (p \wedge q) \vee (p \wedge r)$
5. Identity laws: (a) (i) $p \vee f \equiv p$ (ii) $p \vee t \equiv t$ (b) (i) $p \wedge f \equiv f$ (ii) $p \wedge t \equiv p$
6. Complement laws: (a) (i) $p \vee f \equiv p$ (ii)$p \vee \sim p \equiv t$ (b)(i) $p \wedge \sim p \equiv p$ (ii) $\sim t \equiv f, \sim f \equiv t$
7. De-Morgan's laws: (a) $\sim (p \vee q) \equiv \sim p \wedge \sim q$ (b) $\sim (p \wedge q) \equiv \sim p \vee \sim q$

where t and f are used to denote the variables which are restricted to the truth values true and false respectively.

Solved Example:

1. Show that $(p \to q) \equiv p \wedge \sim q$

Solution: Let us construct the truth table for the given propositions

p	q	$p \to q$	$\sim (p \to q)$	$\sim q$	$p \wedge \sim q$
T	T	T	F	F	F
T	F	F	T	T	T
F	T	T	F	F	F
F	F	T	F	T	F

form the table it is clear that the values of $\sim (p \to q)$ and $p \wedge \sim q$ are identical.
Therefore, $(p \to q) \equiv p \wedge \sim q$

2. Show that $\sim (p \vee (\sim p \wedge q))$ and $\sim p \wedge q$ are logically equivalent.

Solution: Consider $\sim (p \vee (\sim p \wedge q))$

$$
\begin{aligned}
\sim (p \vee (\sim p \wedge q)) &\equiv \sim p \wedge \sim (\sim p \wedge q) \quad (from\ De-Morgan's\ law) \\
&\equiv \sim p \wedge (\sim (\sim p) \vee \sim q) \quad (from\ De-Morgan's\ law) \\
&\equiv \sim p \wedge (p \vee \sim q) \quad (since\ \sim (\sim p) = p\ by\ double\ negation) \\
&\equiv (\sim p \wedge p) \vee (\sim p \wedge \sim q) \\
&\equiv f \vee (\sim p \wedge \sim q)) \quad (since\ (\sim p \wedge p)\ is\ false\ and\ (\sim p \wedge p) \equiv f) \\
&\equiv \sim p \wedge \sim q
\end{aligned}
$$

$\sim (p \rightarrow q)$ and $p \wedge \sim q$ are logically equivalent.

1.10 Duality

The principle of duality states that any established result involving statement formulas and connectives $\vee$ and $\wedge$ give a corresponding dual by replacing $\wedge$ by $\vee$ and $\vee$ by $\wedge$. If the formula contains special variables 't' and 'f', then corresponding dual is obtained by replacing 't' by 'f' and 'f' by 't'. The connectives $\wedge$ and $\vee$ are called duals of each other.

Definition 1.10.1. *Two statement formulas p and p^* are said to be duals of each other if one can be obtained from the other by replacing $\wedge$ by $\vee$ and $\vee$ by $\wedge$.*

Example: Write the duals of (a) $(p \wedge q) \vee r$ (b) $(p \wedge q) \vee r$ (c) $\wedge(p \wedge q)$
Solution: The duals are (a) $(p \vee q) \wedge r$ (b) $(p \vee q) \wedge r$ (c) $\vee(p \vee q)$

Theorem 1.10.1. *Let A and A^* be dual formulas and let$p_1, p_2, ...p_n$ be all the atomic variables that occur in A and A^*. That is to say, we may write A as $A(p_1, p_2, ...p_n)$ and A^* as $A^*(p_1, p_2, ...p_n)$. Then through the use of De-Morgan's laws,*

$$P \wedge Q \Leftrightarrow \sim (\sim P \vee \sim Q)$$
$$P \vee Q \Leftrightarrow \sim (\sim P \wedge \sim Q)$$

We can show $\sim A(p_1, p_2, ...p_n) \Leftrightarrow A^*(p_1, p_2, ...p_n)$

Thus the negation of a formula is equivalent to its dual in which every variable is replaced by its negation. As a consequence of this fact, we also have, $\sim A(p_1, p_2, ...p_n) \Leftrightarrow \sim A^*(p_1, p_2, ...p_n)$

Example: Verify equivalence if A(P, Q, R) is $\sim P \wedge \sim (Q \vee R)$
Solution: Now $A^*(\sim P, \sim Q, \sim R)$ is $\sim\sim P \vee \sim (\sim Q \wedge \sim R) \Leftrightarrow P \vee (Q \vee R)$.
On the other hand, $\sim A(P, Q, R)$ is $\sim (\sim P \wedge \sim (Q \vee R)) \Leftrightarrow P \vee (Q \vee R)$.

We shall now prove an interesting theorem which states that if any two formulas are equivalent, then their duals are also equivalent to each other In other words, if $A \Leftrightarrow B$ then $A^* \Leftrightarrow B^*$.

Theorem 1.10.2. *Let $P_1, P_2, ...P_n$ be all the atomic variables appearing in the formulas A and B. Given that "$A \leftrightarrow B$ is a tautology" then the following are also tautologies.*

$$A(P_1, P_2, ...P_n) \leftrightarrow B(P_1, P_2, ...P_n)$$
$$A(\sim P_1, \sim P_2, ... \sim P_n) \leftrightarrow B(\sim P_1, \sim P_2, ... \sim P_n)$$

Using theorem 1.10.1 we get $\sim A^*(P_1, P_2, ...P_n) \leftrightarrow \sim B^*(P_1, P_2, ...P_n)$
Hence $A^* \Leftrightarrow B^*$.

Example: Show that $\sim (P \wedge Q) \rightarrow (\sim P \vee (\sim P \vee Q)) \Leftrightarrow (\sim P \vee Q)$

Solution: $\sim (P \wedge Q) \rightarrow (\sim P \vee (\sim P \vee Q))$

$$\begin{aligned}
&\Leftrightarrow (P \wedge Q) \vee (\sim P \vee (\sim P \vee Q)) \\
&\Leftrightarrow (P \wedge Q) \vee (\sim P \vee Q \\
&\Leftrightarrow (P \wedge Q) \vee \sim P \vee Q \\
&\Leftrightarrow ((P \vee \sim P) \wedge (Q \vee \sim P)) \vee Q \\
&\Leftrightarrow (Q \vee \sim P) \vee Q \Leftrightarrow Q \vee \sim P \Leftrightarrow \sim P \vee Q
\end{aligned}$$

From (1.10.3) it follows that,

$\sim (P \wedge Q) \rightarrow (\sim P \vee (\sim P \vee Q)) \Leftrightarrow (\sim P \vee Q)$

1.11 Normal Form

Although two formulas may be logically equivalent, one may be "easier" for someone to understand or to manipulate. For example, in one formula is satisfiable. It may be fairly obvious that one formula is a tautology but quite difficult to conclude that from the other form of the same formula. In this section, we discuss two special forms or representations for formulas logically equivalent to a given formula. These forms and conjunctive normal forms. Formulas in conjunctive normal form are easy to use when asking whether a formula is a tautology. These special forms have assumed prominence in Computer Science, in both theoretical and applied areas. The famous $P \neq NP$ problems deals with conjunctive normal forms to find representations of combinatorial circuits.

1.11.1. Disjunctive Normal Form: Consider the following two formulas,

$\phi = (p \rightarrow (q \vee r)) \leftrightarrow (q \rightarrow p)$ and

$\psi = (p \wedge q) \vee (p \wedge q) \vee (\sim p \wedge \sim q)$

The truth table for ϕ and ψ would show that these two formulas are logically equivalent. By some measures, ψ is more complicated. For example ϕ has four propositional connectives. Whether ψ has nine. Nevertheless, many people find ψ to be far easier to understand. The formula ψ explicitly lists three cases in which the formula is true.

1. p and q are both T.
2. p and r are T and q is F.
3. p and q are both F.

For all other interpretation of p, q and r the truth value of ψ is F. It is not nearly so obvious what ϕ "says". Although ϕ is shorter, it also sums to be more complex.

A formula like ψ that is just a list of cases that make the formula have a truth value of T is called a disjunctive normal form(DNF). Each of the three cases $(p \wedge q)$, $(p \wedge \sim q \wedge r)$ and $(\sim p \wedge \sim q)$ is called a term. One might think of each term as describing single case. The

entire disjunctive normal form formula is just a disjunct of terms that make the formula T. (The words term and disjunctive normal form will be defined formally bellow)

The difference in comprehensibility is even more extreme if the formula ϕ is negated. The formula

$$\sim ((p \rightarrow (q \vee r)) \leftrightarrow (q \rightarrow p))$$

is logically equivalent to the disjunctive normal form formula.

$$(\sim p \wedge q) \vee (p \wedge \sim q \wedge \sim r)$$

The disjunctive normal form is a disjunction of only two terms, which makes it particularly easy to understand.

Definition 1.11.1 *Let p be a proposition letter. Then, p is positive literal, and $\sim p$ is a negative literal. A literal is a positive literal or a negative literal.*

Definition 1.11.2 *Let $\lambda_1, \lambda_2 ..... \lambda_m$ be a set of m literals with $m \in N$. A term is a conjunction. $\lambda_1 \wedge \lambda_2 \wedge ..... \wedge \lambda_m$ of m literals a formula ϕ is in DNF if it is a disjunction $\phi_1 \vee \phi_2 \vee .... \vee \phi_k$ of k terms where $k \in N$.*

The disjunction of zero formulas is F. The conjunction of zero formulas is T. This is analogous to defining the sum of zero numbers to be zero and the product of zero numbers to be 1. For example, $F \vee p \leftrightarrow p$ is analogous to 0+x=x.

Examples :

1. (a)$(a \wedge b \wedge \sim c)$ is a term.

(b)The formula $(a \wedge b \wedge c) \vee (\sim a \wedge \sim b \wedge \sim c) \vee (a \wedge \sim c \wedge q)$ is in disjunctive normal form.

(c) T is a term. It is a conjunction of zero literals.

(d) a is a term.It is a conjunction of zero literals.

(e) $(a \wedge b \wedge \sim c)$ and T are disjunctive normal form.

Each is a disjunction normal form. It is disjunction of zero term.

2 Let $\psi = (\sim (p \rightarrow q)) \rightarrow (q \wedge \sim r)$. Determine a DNF for ψ.

Solution: A formula may have several equivalent formulas in DNF, but we want a systematic way to find one. The first step in finding a DNF for all the interpenetrations of ψ, as shown in the following

Abbreviated truth table for ψ

Interpretation	p	q	r	$(\sim(p \to q)) \to (q \wedge \sim r)$
I_0	T	T	T	T
I_1	T	T	F	T
I_2	T	F	T	F
I_3	T	F	F	F
I_4	F	T	T	T
I_5	F	T	F	T
I_6	F	F	T	T
I_7	F	F	F	T

The next step is construct, for each interpretation I, $0 \leq 1 \leq 7$, a team that is T in that interpretation and F in all other interpretation.

Truth terms in the interpretation

Interpretation	Matching Terms
I_0	$p \wedge q \wedge r$
I_1	$p \wedge q \wedge \sim r$
I_2	$p \wedge \sim q \wedge r$
I_3	$p \wedge \sim q \wedge \sim r$
I_4	$\sim p \wedge q \wedge r$
I_5	$\sim p \wedge q \wedge \sim r$
I_6	$\sim p \wedge \sim q \wedge r$
I_7	$\sim p \wedge \sim q \wedge \sim r$

The reader should observe that these terms have desired properties. That these terms have the desired properties. That is, I_0 satisfies $p \wedge q \wedge r$, and all seven other interpretation do not satisfy $p \wedge q \wedge r$.

1.11.2.Conjunctive Normal Form: Consider again the formula as a motivating example for DNFS:

$$(p \to (q \vee r)) \leftrightarrow (q \to p)$$

The formula is logically equivalent to the formula,

$$(p \vee \sim q) \wedge (\sim p \vee q \vee r)$$

This logically equivalent formula is in conjunctive normal form(CNF). It consists of a conjunction of two formulas that are disjunction of literals. In this examples, it is the conjunct of $(p \vee \sim q)$ and $(\sim p \vee q \vee r)$. Each disjunction of zero or more literals can be though of as a restriction on when the formula can be T. The first restriction is that at least one of p and $\sim q$ of $\sim p$, q and r must be T. This can be thought of as a list of rules that must all be met for the formula to be satisfied. Thus, CNF formulas are often easy to understand.

Definition 1.11.3 *Let $\lambda_1, \lambda_2 .... \lambda_m$ be a set of m literals with $m \in \mathbf{N}$. A clause is a disjunction. $\lambda_1 \vee \lambda_2 \vee .... \vee \lambda_m$ of m literals. A formula ϕ is in CNF if it is a conjunction $\phi_1 \wedge \phi_2 \wedge .... \wedge \phi_k$ of k clauses $\phi_1 \phi_2 .... \phi_k, k \in \mathbf{N}$.*

Examples:

1. (a) $a \vee b \vee \sim c$ is a clause.

(b) T is in CNF. It is a conjunction of zero clauses.

(c) F is a clause. It is a disjunction of zero literals.

(d) a is a clause. It is a disjunction of one literal.

(e) $(a \vee b \vee \sim c)$ and F are in CNF. Each is a conjunction of one clause.

2. Find the conjunctive normal form for the formula $\psi = (\sim (p \rightarrow q)) \rightarrow (q \wedge \sim r)$

Solution: The process starts by finding a formula in DNF that is equivalent to $\sim \psi$.

The following is an abbreviated truth table for $\sim \psi$. We will misuse the word interpretation exactly as we did in Example 2 in section 1.11.1.

Interpretation	p	q	r	$\sim ((\sim (p \rightarrow q)) \rightarrow (q \wedge \sim r))$
I_0	T	T	T	F
I_1	T	T	F	F
I_2	T	F	T	T
I_3	T	F	F	T
I_4	F	T	T	F
I_5	F	T	F	F
I_6	F	F	T	F
I_7	F	F	F	F

$\phi_{\sim\psi} = \phi_2 \vee \phi_3 = (p \wedge \sim q \wedge r) \vee (p \wedge \sim q \wedge \sim r)$

So, ψ is logically equivalent to

$$\sim ((p \wedge \sim q \wedge r) \vee (p \wedge \sim q \wedge \sim r))$$

push the negations inside, first past the $\vee$ using De-Morgan's law

$$(\sim p \vee \sim q \wedge r) \wedge (p \wedge \sim q \wedge \sim r)$$

the past the internal $\wedge$'s, again using De-Morgan's law.

$$(\sim p \vee \sim\sim q \vee \sim r) \wedge (\sim p \vee \sim\sim q \vee \sim\sim r)$$

and finally, eliminate the double negations:

$$(\sim p \vee q \vee \sim r) \wedge (\sim p \vee q \vee \sim r)$$

Since $\phi_{\sim\psi}$ was in DNF, negating and pushing the negations inside creates a formula in CNF logically equivalent to ψ.

Solved Examples:

1. Obtain the disjunctive normal form of $(p \vee (\sim p \rightarrow q \rightarrow \sim r))$.

Solution:

$$\begin{aligned} p \vee (\sim p \rightarrow (q \vee (q \rightarrow \sim r))) & \\ \equiv\ & p \vee (\sim p \rightarrow (q \vee (\sim q \rightarrow \sim r)) \\ \equiv\ & p \vee (p \vee (\sim q \vee \sim r)) \\ \equiv\ & p \vee p \vee q \vee \sim q \vee \sim r \\ \equiv\ & p \vee q \vee \sim q \vee \sim r. \end{aligned}$$

2. Obtain the principal disjunctive normal form of $\sim p \vee q$.
Solution:

$$\begin{aligned} \sim p \vee q \equiv\ & (\sim p \wedge (q \vee \sim q)) \vee (q \wedge (p \vee \sim p)) \\ \equiv\ & (\sim p \wedge q) \vee (\sim p \wedge \sim q) \vee (q \vee p) \vee (q \wedge \sim p) \\ \equiv\ & (\sim p \wedge q) \vee (\sim p \wedge \sim q) \vee (p \wedge q) \vee (\sim p \wedge q) \\ \equiv\ & (\sim p \wedge q) \vee (\sim p \wedge \sim q) \vee (p \wedge q) \vee (\sim p \wedge q) \\ \equiv\ & (\sim p \wedge q) \vee (\sim p \wedge \sim q) \vee (p \wedge q) \end{aligned}$$

Hence $(\sim p \wedge q) \vee (\sim p \wedge \sim q) \vee (p \wedge q)$ is in the required principal disjunctive normal form.

Check Your Progress

(1) Explain different types of connectives with example.
(2) Define logical equivalence with example.
(3) Explain types of normal formes.

1.12 Summary

- Mathematical logic is the science dealing with the method of reasoning.
- Mathematical logic finds applications in many areas.
- A proposition is a statement which in a given context can be said to be either true or false but not both.
- Communally used connections and Notations are negation ($\sim$), and ($\wedge$), or($\vee$), if and only if ($\leftrightarrow$) etc.
- Converse, Inverse and Contrapositive propositions are very important in mathematical logic.
- A statement formula is a string consisting of variables, parentheses and connective symbols. A statement formula is called a well formed formulas.
- Two expressions are logical equivalent if they have the have the same truth values for every combination of the variables.
- We concluded with solved examples and exercises

1.13 Key Words

Mathematical logic, Notations, Statements, Connectives, Converse, logical equivalence, Truth table, Truth values.

1.14 Answers Check Your Progress

(1) 1.5 (2) 1.9 (3) 1.11

1.14 Exercise and Answers

1) Given that p is true and q is false find the truth values of the following:

(a) $\sim p \wedge q$ (b) $\sim (p \rightarrow \sim q)$ (c) $(p \wedge q) \rightarrow (p \vee q)$

(d)$\sim (p \wedge q) \vee \sim (q \leftrightarrow p)$ e) $(p \rightarrow q) \wedge \sim (p \leftrightarrow \sim q)$.

2) Given $p : \sqrt{2}$ is an irrational number and q: Every rectangle is a parallelogram, indicate the truth values of the following propositions:

(1) $p \rightarrow q$ (2) $q \rightarrow p$ (3) $p \rightarrow \sim q$ (4) $\sim p \rightarrow q$ (5) $q \rightarrow \sim p$ (6) $p \leftrightarrow q$

(7) $\sim p \leftrightarrow q$ (8) $p \leftrightarrow \sim q$ (9) $\sim p \leftrightarrow \sim q$.

3) Construct the truth tables for the following

(1) $(p \vee q) \wedge (\sim p)$ (2) $\sim (p \wedge \sim q)$ (3) $\sim (p \wedge q) \vee \sim (q \leftrightarrow p)$

(4)$[[p \wedge q] \vee (\sim r)] \leftrightarrow p$ (5) $[p \rightarrow (\sim q \vee r)] \wedge \sim [q \vee (p \leftrightarrow \sim r)]$

4) Show that $p \wedge q \equiv (p \downarrow q) \downarrow (p \downarrow q)$.

5) Construct a truth table for $p \downarrow q$.

6) Prove that for any proposition p, the compond proposition $p \vee \sim q$ is a tautology and the compound proposition $p \wedge \sim p$ is a contradiction.

7) Prove that, for any proposition p, q, r the proposition $[(p \rightarrow q) \wedge (q \rightarrow r)] \rightarrow (p \rightarrow r)$ is a tautology.

8) Construct the truth tables for the following Compound propositions and check these are either tautology contradiction as contingency.

(a) $p \wedge \sim q$ (b) $p \rightarrow \sim q$ (c) $(p \wedge q) \rightarrow \sim r$ (d) $p \wedge (q \rightarrow \sim r)$.

9) Determine whether the following are tautologies:

(a) $p \wedge [\sim (p \wedge q)]$ (b) $(p \vee q) \vee \sim p$ (c) $p \rightarrow (p \wedge q)$ (d) $p \rightarrow (p \vee q)$.

10) Show that $p \rightarrow (q \rightarrow r) \Leftrightarrow p \rightarrow (\sim q \vee r) \Leftrightarrow (p \wedge q) \rightarrow r$.

Hint: On recalling that $\Leftrightarrow$ is the same as $\equiv$.

11) Prove that $p \rightarrow q \equiv \sim p \vee q$.

12) Show that $p \leftrightarrow q \equiv (p \leftarrow q) \wedge (q \rightarrow p)$ are equivalent.

13) Show that $[\sim p \wedge (\sim q \wedge r)] \vee (q \wedge r) \vee (p \wedge r) \equiv r$.

14) Obtain the disjunctive normal form of $(p \wedge \sim (q \wedge r)) \vee (p \rightarrow q)$

15) Find the principal conjunctive normal form of $p \vee (q \rightarrow r)$.

16) Find the principal conjunctive normal form of $(p \wedge q) \vee (\sim p \wedge r)$.

17) Obtain the principal disjunctive normal form of $(\sim p \vee \sim q) \rightarrow (\sim p \vee r)$.

Answers:

1) a.F, b.F, c.T, d.T, e.F, 2)1.T, 2.T, 3.F, 4.T, 5.F, 6.T, 7.F, 8.F, 9.T

3)(1)

p	q	$(p \vee q) \vee (\sim p)$
T	T	F
T	F	F
F	T	T
F	F	F

(2)

p	q	$\sim (p \vee \sim q)$
T	T	F
T	F	F
F	T	T
F	F	F

(3)

p	q	$\sim (p \wedge q) \vee \sim (q \leftrightarrow p)$
T	T	F
T	F	T
F	T	T
F	F	T

(4)

p	q	r	$[(p \wedge q) \vee (\sim r)] \leftrightarrow p$
T	T	T	T
T	T	F	T
T	F	T	F
T	F	F	T
F	T	T	T
F	T	F	F
F	F	T	T
F	F	F	F

(5)

p	q	r	$[p \rightarrow (\sim q \vee r)] \wedge \sim [q \vee (p \leftrightarrow \sim r)]$
T	T	T	F
T	T	F	F
T	F	T	T
T	F	F	F
F	T	T	F
F	T	F	F
F	F	T	F
F	F	F	T

1.15 Suggested Readings

- Kolman, Barby, Ross *Discrete Mathematical Structures, 4th* Edition, Prentice Hall of India, 2003.
- G. Shankar Rao, *Mathematical Foundations of Computer Science,* I.K. International Publishing House Private Limited, 2006.
- Thomas Koshy, *Discrete Mathematics with Application,* Academic Press An Imprint of Elsevier, 2004.

UNIT-2: Predicate Calculus

Structure

2.0 Objectives

After studying this unit you will be able to:

- Study the important properties & methods of Mathematical logic.
- Discuss the concept of predicate calculus.
- Study rules of Inference, arguments & their consistence.
- Study the Quantified statements different methods of proofs.
- Study the Negations of quantified predicates.

2.1 Introduction

In this unit we describe the process of derivation by which one demonstrates that a particular formula in a valid consequence of given set of premises. The method of derivation involving predicate statement calculus and also certain additional rules which are required to deal with the formulae involving quantifiers. The rules P and T, regarding the introduction of a premise at any stage of derivation, and the introduction of any formula which follows logically from the formulae. If the conclusion is given in the form of conditional, we shall

also use the rule of the conditional proof called CP. Occasionally, we may use the indirect method of proof in introducing the negation of the conclusion as an additional premise in order to arrive at a contradiction.

The equivalence and implication of the statement calculus can be used in the process derivation as before, expect that the formulae involved are generalized to predicates.

2.2 Rules of Inference

Logic is the study of inferences. An inference is a judgement derived from another judgement or from other judgements. The terms inference and reasoning are used as synonyms. Reasoning is of two types. Reasoning, which starts from some assertions or statements and seeks to unfold their implications is formal, and the reasoning which starts from observed facts to discover their character is inductive. The term reasoning is applied for both inference and proof. The term inference and proof appear almost synonyms in the definition of logic.

Definition 2.2.1. *Logic is the branch of mathematics that whether an argument is valid or invalid. In every argument there are two things.*

i) The premises (i.e., series of statements)

ii)The conclusion (i.e., the inference)

We start from the premises and proceed to the conclusion the process is called inference.

2.2.1 : Rules of Inference: Law of Detachment: The law of detachment states that when two given premises are true; one a conditional and the other the hypothesis of that conditional, it then follows the conclusion of the conditional is true.

The law of detachment is also known as the modus pones symbolically, we can write the law of detachment as follows: $p \rightarrow q$ therefore q.

The word therefore only when we reach a conclusion. In symbols, we can also write the law of detachment as

$(p \wedge (p \rightarrow q)) \rightarrow q)$, where q is called valid conclusion.

"If n is greater than 5 then n^2 is greater than 25" is a symbol or the word therefore only when we reach a conclusion. In symbols,we can also write the law of detachment as, $(p \wedge (p \rightarrow q)) \rightarrow q$

where q is called valid conclusion. $(p \wedge (p \rightarrow q) \rightarrow q)$ is a tautology.

"If n is a greater than 25 " is a true statement. Then by modus ponens, it follows that n^2 is greater than 25.

Example. Assume that the following propositions are true.

p: Two triangles are similar

q: If two triangles are similar then their corresponding sides are proportional.

By modus ponens: the conclusion is q. The corresponding are proportional is true.

Rule Of Transitivity(Law of the syllogism or chain rule): If $p \rightarrow q$ and $q \rightarrow r$ are accepted as true then $p \rightarrow r$ must be accepted as true i.e $p \rightarrow q, \quad q \rightarrow r, \quad so, \quad p \rightarrow r.$ we can extend the transitive rule as follows: $p \rightarrow q, \quad q \rightarrow r, \quad r \rightarrow s, \quad so, \quad p \rightarrow s.$

Example: we use the chain rule to solve equations:
In algebra, consider if 3x+7=16 then 3x=9; If 3x=9 then x=3
we get the conclusion as follows, $p \rightarrow q$: if 3x+7=16 then 3x=9

$$\frac{q \rightarrow r: \ if \ 3x = 9 \ then \ x = 3}{p \rightarrow r: \ if 3x + 7 = 16 \ then \ x = 3}$$

Symbolically, the law can be written as $\frac{p \rightarrow q}{\sim q \rightarrow \sim p}$

Law of contrapositive: The law of contrapositive states that when a conditional premise is true, it follows that the contrapositive of the premise is also true. Symbolically, the law can be written as:
$\frac{p \rightarrow q}{\sim q \rightarrow \sim p}$ is the tautology.
$(p \rightarrow q) \leftrightarrow (\sim q \rightarrow \sim p)$ is the basis of the rule of inference.
Example: Use a contrapositive argument to verify the following valid inference.

$$\frac{w \rightarrow (r \rightarrow s)}{(w \wedge r) \rightarrow s}$$

$$Suppose \sim s \longrightarrow \sim (w \wedge r)$$

$$\begin{aligned} Now \sim s \longrightarrow \sim (w \wedge r) &\equiv \sim (\sim s) \vee (w \wedge r) \\ &\equiv s \vee \sim (w \wedge r) \\ &\equiv \sim (w \wedge r) \vee s \\ &\equiv w \wedge r \longrightarrow s \end{aligned}$$

Rules of Inference

Implications		
I_1	$P \wedge Q \Rightarrow P$	(Simplification)
I_2	$P \wedge Q \Rightarrow Q$	(Simplification)
I_3	$P \Rightarrow P \vee Q$	(Addition)
I_4	$Q \Rightarrow P \vee Q$	(Addition)
I_5	$\sim P \Rightarrow P \rightarrow Q$	
I_6	$Q \Rightarrow P \rightarrow Q$	
I_7	$\sim (P \rightarrow Q) \Rightarrow P$	
I_8	$\sim (P \rightarrow Q) \rightarrow \sim Q$	
I_9	$P : Q \Rightarrow P \wedge Q$	
I_{10}	$\sim P, P \vee Q \Rightarrow Q$	(Disjunctive Syllogism)

$I_{11}P$	$P \to Q \Rightarrow Q$	
I_{12}	$\sim Q, P \to Q \Rightarrow \sim P$	(Modus Tollens)
I_{13}	$P \to R, Q \to R \Rightarrow P \to R$	(Hypothetical Syllogisn)
I_{14}	$P \vee Q, P \to R, Q \to R \Rightarrow R$	(Dilemma)

Equivalences:		
E_1	$\sim\sim \Leftrightarrow P$	Double negation
E_2	$P \wedge Q \Leftrightarrow Q \Leftrightarrow P$	(Commutative Law)
E_3	$P \vee Q \Leftrightarrow Q \vee P$	(Commutative Law)
E_4	$(P \wedge Q) \wedge R \Leftrightarrow P \wedge (Q \wedge R)$	(Associative Law)
E_5	$(P \vee Q) \wedge R \Leftrightarrow P \vee (Q \wedge R)$	(Associative Law)
E_6	$P \wedge (Q \vee R) \Leftrightarrow (P \wedge Q) \vee (P \wedge R)$	(Distributive Law)
E_7	$P \vee (Q \wedge R) \Leftrightarrow (P \vee Q) \wedge (P \vee R)$	(Distributive Law)
E_8	$\sim (P \wedge Q) \Leftrightarrow \sim P \vee \sim Q$	(Demorgan's Law)
E_9	$\sim (P \vee Q) \Leftrightarrow \sim P \wedge \sim Q$	(Demorgan's Law)
E_{10}	$P \vee P \Leftrightarrow P$	
E_{11}	$P \wedge P \Leftrightarrow P$	
E_{12}	$\vee(P \wedge \sim P) \Leftrightarrow R$	
E_{13}	$R \vee (P \vee P) \Leftrightarrow R$	
E_{14}	$R \vee (P \vee \sim P) \Leftrightarrow T$	
E_{15}	$R \wedge (P \wedge \sim P) \Leftrightarrow F$	
E_{16}	$P \to Q \Leftrightarrow \sim P \wedge Q$	
E_{17}	$\sim (P \to Q) \Leftrightarrow \sim P \vee Q$	
E_{18}	$(P \to Q) \Leftrightarrow \sim Q \to \sim P$	
E_{19}	$P \to (Q \to R) \Leftrightarrow (P \wedge Q) \to R$	
E_{20}	$\sim (P \rightleftharpoons Q) \Leftrightarrow P \leftrightarrow \sim Q$	
E_{21}	$P \rightleftharpoons Q \Leftrightarrow (P \to Q) \wedge (Q \wedge P)$	
E_{22}	$P \rightleftharpoons Q \Leftrightarrow (P \wedge Q) \vee (\sim P \wedge \sim Q)$	

2.3 Argument

An argument is an assertion that, given a set of propositions $p_1, p_2, ...p_n$ called premises has as a consequence another proposition q. It is denoted by $p_1, p_2, ..., p_n \vdash q$ (or by $(p_1 \wedge p_2 \wedge ... \wedge p_n) \to q$). If $p_1, p_2, ...p_n$ yields q, then q is called the conclusion (or consequence).

In this case we say that q logically follows from $p_1, p_2, ...p_n$. Arguments may be categorical, hypothetical, disjunctive or relational.

2.3.1. Valid argument: An argument $p_1, p_2, ...p_n \vdash q$ is true if q is true, whenever all the premises $p_1, p_2, ...p_n$ are true. An argument which is true is called a valid argument.

2.3.2. False argument: If an argument is not true, then it is said to be false argument.

2.3.3. Fallacy: If an argument is false, then it is called fallacy.

Fallacies are invalid arguments. They violate some logical principle. There are fifteen types of fallacies. Aristotle arranged them in two groups. Namely

(i) Verbal fallacies

and (ii)Non-Verbal fallacies.

In this, we arrange the fallacies in three groups.

1. The fallacy of affirming the consequent or affirming the converse. It can be expressed symbolically as $p \rightarrow q \ \frac{q \ fallacy}{p}$

2. The fallacy denying antecedent. This can be expressed as, $p \rightarrow q, \sim p, \ fallacy, \ therefore \sim q$

3. Non-sequitur fallacy: It takes the form $p, \ therefore \ q$

2.3.4 : Rules P and T:

Rule P: A premise may be introduced at any point in the derivation.

Rule T: A formula S may be introduced in a derivation, if S is a Tautology implied by anyone or more of the preceding formulas in the derivation.

The above rules are used in the process of derivation by which one demonstrates that particular formula is a valid consequence of a gives set of premises.

The most commonly used proof in logic involves the column. The first column consists of statements and the second column consists of reasons that follows us to make the statements.

Examples:

1. Prove, $\sim q, p \rightarrow q \rightarrow \sim p$.

Solution:

statements	Reasons
$1) p \rightarrow q$	rule P
$2) \sim q \rightarrow \sim p$	$ruleT, (1) \ and \ E_{18}$
$3) \sim q$	rule P
$4. \sim p$	$ruleT, (2), (3) \ and \ L_{11}$

2. Given the premises $p \rightarrow r, \ \sim p \rightarrow q, \ \sim r$

Prove the conclusion q.

Solution:

No.	statements	Reasons
1.	$p \longrightarrow r$	given
2.	$\sim p \longrightarrow q$	given
3.	$\sim r$	given
4.	$\sim p$	The law of modus tollens,(1),(3)
5.	q	The law of detachment,(2),(4)

2.4 Consistency

Consistency is an extremely important notation in mathematical logic.

A collection of statement is consistent if the statements can all be true simultaneously .

We use the notation inconsistency of in a method of proof called proof by contradiction.

Example: Show that $(R \rightarrow \sim Q)$, $R \vee S$, $S \rightarrow \sim Q$, $P \rightarrow Q \leftrightarrow \sim P$ are consistent.

Solution:

1.	P	(assumed)
2.	$P \rightarrow Q$	Rule P
3.	Q	P
4.	$S \rightarrow \sim Q$	P
5.	$Q \rightarrow \sim S$	$P \rightarrow Q \leftrightarrow \sim Q \rightarrow \sim P$
6.	$\sim S$	(3),(5)
7.	$R \vee S$	P
8.	$\sim R \rightarrow S$	$P \rightarrow Q \leftrightarrow \sim P \vee \sim P$
9.	$\sim S \rightarrow R$	$P \rightarrow Q \leftrightarrow \sim Q \rightarrow \sim P$
10.	R	(6),(9)
11.	$R \rightarrow \sim Q$	P
12.	$\sim Q$	(10),(11)
13.	$Q \wedge \sim Q$	(3),(12)

inconsistent.

Automatic Theorem Proving:

In this section we describe a procedure of derivation which can be conducted mechanically. The formulation of proof is based on the work of Hao wang.

Hao Wang Rules:

1) **Variables:** The capital letters $A, B, C, ..., P, Q, R....$ are used as statement variables.

2) **Connectives:** The connectives $\sim, \wedge, \vee, \longrightarrow$ and $\Longleftrightarrow$ appears in the formulas with the order of procedure as given.

3) **A string of formulas is defined as follows:**

i) A formula is a string of formulas.

ii) If α and β are strings of formulas.Then α, β are strings of formulas.

iii) Only those strings which are obtained by steps.

where i) and ii) are strings of formulas, with the exception of the empty strings which is also a string of formulas.

4) **Sequence:** If α and β are strings of formulas, then $\alpha \rightarrow \beta$ is called a sequent in which α is denoted the antecedent and β the consequent of sequent. A sequent $\alpha \rightarrow \beta$ is true if only if either at least one of the formulas of an antecedent is false or at least one of the formulas of the consequent is true.

5) **Axiom Schema:** If α and β are strings of formulas such that every formula in both α and β is a variable only,the sequent $\alpha \rightarrow \beta$ is an axiom if and only if α and β have at least one variable in common.

If $\alpha \rightarrow \beta$ is an axiom,then $\alpha \Rightarrow \beta$

6) **Theorem:** The following sequent are theorem of the systems.

i) Every axiom is a theorem.

ii) If a sequent α is a theorem. If a sequent β results from α through the use of one of the rules of the system which are given in 0, then β is also a theorem.

iii) The sequent obtained i) and ii) are the only theorem.

7)**Rules:** If $\alpha, \beta, \gamma,.....$ denote strings of formulas to which the connectives are applied, then the following rules are used to combine formulas within strings by introducing connectives.

Antecedent Rules:

Rule $\sim\rightarrow$: If $\alpha, \beta \rightarrow$ x, y, then $\alpha, \sim x, \beta \rightarrow \gamma$

Rule $\wedge \rightarrow$: If x, y, $\alpha, \beta \rightarrow \gamma$, then $\alpha, x \wedge y, \beta \rightarrow \gamma$

Rule $\vee \rightarrow$: If x, $\alpha, \beta \rightarrow \gamma$ and y, $\alpha, \beta \rightarrow \gamma$, then $\alpha, x \vee y, \beta \rightarrow \gamma$.

Rule $\rightarrow\rightarrow$: If y, $\alpha, \beta \rightarrow \gamma$, and $\alpha, \beta \rightarrow x, \gamma$, then $\alpha, x \rightarrow y, \beta \rightarrow \gamma$.

Rule $\leftrightarrow\Longrightarrow$: If x, y, $\alpha, \beta \rightarrow \gamma$ and $\beta \rightarrow x, y, \gamma$ then α, x$\leftrightarrow$y, $\beta \rightarrow y$

Consequent rules:

Rule $\rightarrow\sim$: If x, $\alpha \rightarrow \beta, \gamma$ then $\alpha \rightarrow \beta, \sim x, \gamma$

Rule $\rightarrow \wedge$: If $\alpha \rightarrow x, \beta, \gamma$ and $\alpha \rightarrow y, \beta, \gamma$ then $\alpha \rightarrow \beta, x \wedge y, \gamma$

Rule $\rightarrow \vee$: If $\alpha \rightarrow x, y, \beta, \gamma$ then $\alpha \rightarrow \beta, x \vee y, \gamma$.

Rule $\Rightarrow\rightarrow$: If x, $\alpha \Rightarrow y, \beta, \gamma$ then $\alpha \Rightarrow \beta, x \rightarrow y, \gamma$.

Rule $\Rightarrow\leftrightarrow$: If x, $\alpha \Rightarrow y, \beta, \gamma$ and y, $\alpha \Rightarrow x, \beta, \gamma$ then $\alpha \Longrightarrow \beta, x \Longleftrightarrow y\gamma$.

In the above rules the symbol " $\Rightarrow$ " is a generalization of the connective $\rightarrow$ to strings of formulas. And the symbol " $\Rightarrow$ "is applied to strings of formulas as a generalization of the symbol " $\rightarrow$ "

$\alpha \rightarrow \beta$ means"α" implies "β " or $\alpha \rightarrow \beta$ is a tautology and $\alpha \rightarrow \beta$ means that $\alpha \rightarrow \beta$ is true. Also, $A, B, C \rightarrow D, E, F$ is true if $A \wedge B \wedge C \rightarrow D \vee E \vee F$ is true.

In the method of derivation explained earlier, we showed a conclusion C follows from the premises $H_1, H_2, ...H_n$ we introduced at various stages by using rule P. In the new formulation : to show that C follows H1, H2,...Hn,
we establish that

$$(\rightarrow H_1 \rightarrow (H_2 \rightarrow (H_3...(H_n \rightarrow C)...))) \qquad (2.4.1)$$

is a theorem. That is

$$(\rightarrow H_1 \rightarrow (H_2 \rightarrow (H_3...(H_n \rightarrow C)...)))C)...))) \tag{2.4.2}$$

and show that the assumption (2.4.2) is justified then (2.4.1) is not a theorem. This task is accomplished by working backward from (2.4.2), using the rules and showing that (2.4.2) holds if some simpler sequent is a theorem. We continue to work backward we arrive at the simplest possible sequent. If these sequents are axioms, then we have justified our assumption of (2.4.2) is not justified and the conclusion C does not follow from the premises $H_1, H_2, ..., H_n$.

Example: Show that P does not follow from $P \vee Q$.

Solution: Assume,

i) $\Rightarrow (P \vee Q) \rightarrow P$

ii) If (ii) $P \vee Q \Rightarrow P(\Rightarrow\rightarrow)$

iii) If iii) $P \Rightarrow P$ and iv) $Q \Rightarrow P(v \Rightarrow)$

iv) Is an axiom

but iv) is not an axiom.

Hence P does not follows from $P \vee Q$.

i.e $\rightarrow (P \rightarrow Q) \rightarrow P$ is not a theorem.

2.5 Methods of Proof

In this section we briefly discuss direct proof, indirect proof, proof by counter example and proof by cases.

2.5.1. Direct Proof: We assume that P is true ; and from the available information the conclusion Q is shown to be true by valid reference. In this method of proof, we construct a chain of statements $P, P_1, P_2..., P_n, Q$ where is either a hypothesis of the theorem or an axiom and to each of the implication $P \Rightarrow P_1, P_1 \Rightarrow P_2, ..., P_n \Rightarrow Q$ is either an axiom or is implied by the implication preceding it.

Example: If x and y are odd integers then $x + y$ is an even integer.

Solution: An odd integer is of the form $2m + 1$, where m is an integer.

x and y are odd integers, $x = 2m_1 + 1, y = 2m_2 + 1$ for some integers m_1 and m_2.

$$\begin{aligned} \therefore\ x + y &= (2m_1 + 1) + (2m_2 + 1) \\ &= 2m_1 + 1 + 2m_2 + 1 \\ &= 2m_1 + (2m_2 + 1) \\ &= 2(m_1 + m_2 + 1) \end{aligned}$$

$x + y$ *is even.*

2.5.2. Method of Contraposition(Indirect Proof): This method of proof is very useful and is powerful at all levels of the subject of mathematics. Indirect method follows from the tautology $(p \rightarrow q) \leftrightarrow ((\sim q) \rightarrow (\sim p))$. This states that the implication $p \Rightarrow q$ is equivalent to $\sim q \Rightarrow \sim p$.

To prove $p \Rightarrow q$ indirectly, we assume that q is false and then show that p is false.

Example: Prove If α^2 is an even integer then α is an even integer.

Solution: Let $P : \alpha^2$ is an even integer.

$q : \alpha$ is an even integer.

Let $\sim q$ be true.

α is not an even integer. Therefore α must be odd and α is of the form $\alpha = 2m+1$ for some integer m. Now

$$\begin{aligned} \alpha &= 2m+1 \\ \alpha^2 &= (2m+1)^2 \\ &= 4m^2+4m+1 \\ &= 2(2m^2+2m)+1 \end{aligned}$$

Therefore α^2 is of the form $2n+1$ where $n = 2m^2+2m$.

Therefore α^2 is odd. Thus we have $\sim q \Rightarrow \sim p$

Hence by contraposition, α is even.

2.5.3. Proof by Contradiction: This method is based on the tautology $(p \rightarrow q) \wedge (\sim q \rightarrow \sim p)$. In this method, the proof can be constructed as follows:

1. Assume $p \wedge (\sim q)$ is true.
2. On the basis of assumption, find some conclusion that is false.
3. The contradiction is step(2) leads to the conclusion that $p \wedge \sim q$ is false which proves that $p \rightarrow q$ is true. The method is known as reduction or absurdum or proof by contradiction.

Example: Show that $\sqrt{2}$ is not a rational number.

Solution: Let us assume that $\sqrt{2}$ is rational.

Then we can find integers such that

$$\sqrt{2} = \frac{p}{q}$$

where p and q have no common factor.

Squaring on both sides

$$2 = \frac{p^2}{q^2}, \Rightarrow p^2 = 2q^2$$

$\Rightarrow p^2$ is even. $\Rightarrow p$ is even. $\Rightarrow p = 2m$ for some integer m.

$$\begin{aligned} \Rightarrow (2m)^2 &= 2q^2 \\ \Rightarrow 4m^2 &= 2q^2 \\ \Rightarrow q^2 &= 2m^2 \Rightarrow q \textit{ is even.} \end{aligned}$$

Now, p is even and q is even.
$\Rightarrow p$ *and* q have 2 as the common factor which is contradiction to the statement that p and q have no common factor.

Hence our assumption $\sqrt{2}$ is rational leads to a contradiction.
Thus $\sqrt{2}$ is irrational.

2.5.4. Proof by Counter Example: To show that $\forall x, P(x)$, it is sufficient to give specific example k, in the universe such that $P(k)$ is false, where the object k is called a counter example to the assertion $\forall x, P(x)$.

Example: Prove or disprove the statement
If x and y are real numbers then $(x^2 = y^2) \Leftrightarrow (x = y)$.
Solution: $-5, 5$ are real numbers such that
$(-5)^2 = 5^2$ but $-5 \neq 5$
Therefore the result is false, hence the implication is false.

2.5.5. Proof by Cases: To prove $p \rightarrow q$ by cases, we take p to be in the form $p_1 \vee p_2 \vee ... \vee p_n$ by proving separately each of the following $p_1 \rightarrow q, p_2 \rightarrow q, .., p_1 \rightarrow q$ we can establish $(p_1 \vee p_2 \vee ... \vee p_n) \rightarrow q$.
Example: Prove the conclusion
If Ali does not study, then he will fail $\frac{\textit{Ali did not fail}}{\textit{Ali studied}}$
Solution: Let p :Ali Studied q :Ali failed.

Statement	Reasons
1.$\sim p \rightarrow q$	Premise
2.$\sim q$	Premise
3.$\sim q \rightarrow p$	(1), law of contrapositive
4.p	(2),(3) law of detachment

Therefore Ali studied
It is of the form

$$\sim p \rightarrow q, \ \sim q, So\ p$$

a valid conclusion.

Example: If the product of two integers a and b is even, then show that a is even or b is even.

Solution: Let $p : ab$ is even, $q : a$ is even or b is even.
We have to prove $p \rightarrow q$.
We assume $\sim q$ and prove the contrapositive $\sim q \rightarrow \sim t$. Assume $\sim q$.
We have $\sim q : a$ is odd and b is odd (by De Morgan's law)
$\Rightarrow a = 2m + 1$, and $b = 2n + 1$ where m and n are integers.
Therefore

$$\begin{aligned} ab &= (2m+1)(2n+1) \\ &= 4mn + 2m + 2n + 1 \\ &= 2(2mn + m + n) + 1 \end{aligned}$$

$\Rightarrow ab$ is odd.
Therefore $\Rightarrow \sim p$.
Therefore $\sim q \rightarrow \sim p$, which implies $p \rightarrow q$.

2.6 Predicate

A predicate is a statement containing one or more variables. If values are assigned to all variables in a predicate, the resulting statement is a proposition.

2.6.1. n- Place Predicates: A predicate which involves n variables is called an n Place Predicate.

A predicate which involves one variable is called a one Place Predicates and a predicate which involves two variables is called a two-Place Predicates

Predicates are generally denoted by capital letters such as P, Q, R etc and following the name of a Predicate.

2.6.2. Satisfiable predicate: An n-Place Predicate is said to be satisfiable if there exists an n tuple which satisfies it.

2.6.3. Valid predicate: An n-place predicate is said to be valid if all n tuples satisfy it.

2.6.4. Truth set of a predicate: Let $P(x)$ be a predicate on a set S. Then the set of all these values of $x \in S$, which makes $P(x)$ a true proposition is called the truth set of $P(x)$. It is denoted by $T[(px)]$.

2.6.5. Equivalent Predicates: Two predicates are said to be equivalent, if they have same truth value for possible values of their variables.

Thus, two predicates $P(x)$ and $Q(x)$ are equivalent if $P(x) \leftrightarrow Q(x)$.

Examples:

1. $x + y = 8$

2. The sum of the first natural numbers is $n(n + 1)/2$.

3. If $x < y$ then $x^3 < y^3$ are all predicates.

2.7 Negations of Quantified Predicates

Consider the universally quantified predicate $\forall x P(x)$. The negation of $\forall x,\ P(x)$ is $\sim \forall x,\ P(x)$

Clearly $\sim \forall x,\ P(x)$ is equivalent to $\exists x \sim P(x)$ i.e $\sim \forall x,\ P(x) \equiv \exists x \sim P(x)$

Example: The negation of "No Rectangle is a square" is "Some Rectangles are squares".

Similarly the negation of the existential quantified predicate $\exists x P(x)$ is $\exists x P(x)$.

$\sim \exists x,\ P(x)$ equivalent to $\forall x,\ \sim P(x)$. i.e $\sim \exists x,\ P(x) \equiv \forall x,\ \sim P(x)$

The negation of $\exists x,\ (x^2 = 5)$ is equivalent to $\forall x,\ \sim (x^2 = 5)$, This can be written as $\forall x,\ (x^2 \neq 5)$

2.7.1. Nested Quantifiers: Quantifiers that occur with the scope of other quantifiers are called Nested quantifiers.

Example:

$\forall x,\ \exists y,\ (x + y = 1)$: where the universe discourse is the set of all real numbers.

Equivalence involving Quantifiers

l_1'	Distributivity of $\exists$ over $\vee$, $\exists x\ (P(x) \vee Q(x)) \equiv \exists X\ (P(x) \vee \exists x\ Q(x))$, $\exists x\ (P \vee Q(x)) \equiv P \exists X\ Q(x)$
l_2'	Distributivity of $\forall$ over $\wedge$,$\forall x\ (P(x) \wedge Q(x)) \equiv \forall x\ P(x) \wedge \forall x\ Q(x)$, $\forall x\ (P \wedge Q(x)) \equiv P \wedge \forall (xQ(x))$
l_3'	$\sim (\exists x\ P(x)) \equiv \forall x - P(x)$
l_4'	$\sim (\forall x\ P(x)) \equiv \exists x\ \sim (P(x))$

Example : Prove that the sum of any two even numbers is an even number.

Solution: Let x and y be even numbers.

Then x=2m for some integer m and y=2n for some integer, n.

Therefore , x+y=2m+2n=2(m+n). Since m and n are integers, so is m+n,

So 2(m+n) is even.

Hence, $x + y$ is even.

Solved Examples:

1. Find the truth values of each of the following statements R is the Universe of discourse

(i) $\forall x,\ |x| = x$ (ii) $\exists x,\ x^2 = x$

(iii) $\forall x,\ x + 2 > x$ (iv) $\forall x,\ x + 5 = x$

Solution :

i) False, since $-5 = 5$, i.e when $x = -5$, $x = x$ does not hold.

ii) True

iii) True

iv) False, since the equation $x + 5 = x$ has no solution.

2. Translate the following symbolic form.

All men are mortal. Socrates is a man. Using these prove that Socrates is mortal.

Solution: Let $P(x) = x$ is a man, $Q(x) = x$ is mortal.

$P(x) \wedge Q(x)$ = Man is mortal. $\forall x,\ [P(x) \wedge Q(x)]$: All men are mortal.

Let x: Socrates,

then $P(x)$: Socrates is a man. $Q(x)$: Socrates is mortal.

then $P(x) \wedge Q(x)$:Socrates is Mortal.

3. Determine the truth value of each of these following statements where $x = 1, 2, 3$ is the universal set.

i) $\exists x,\ \forall y, x^2 < y + 1$ ii) $\exists x,\ \forall y, x^2 + y^2 < 12$ iii) $\forall x,\ \forall y, x^2 + y^2 < 12$

Solution: i) T ii) T iii) F

Check Your Progress

(1) Write different types of arguments

(2) Explain automatic theorem proving.

(3) Write the different methods of proof.

(4) Write short note on Predicate.

2.8 Summary

- The equivalence and implications of statement calculus can be used in the process derivation as before, expect that the formulae involved are generalized to call it as predicates.
- In every argument there are two types, one is premises and conclusion.
- An argument is an assertion that, given a set of propositions called premisses has a consequence another proposition.
- A collection of statement is consistent if the statements can all be true simultaneously.
- We are explained different methods of proofs.

2.9 Key Words

Inference, Predicates, Arguments, Consistency, Nested quantifiers.

2.10 Answers Check Your Progress

(1) 2.3 (2) 2.4 (3) 2.5 (4) 2.6

2.11 Exercise and Answers

1) Write down the following propositions in symbolic form:

(a) Some integers are divisible by (b) There exists a matrix whose transpose is itself.

(c) Every element of a group has an inverse.

2) Write the negation of the following:

a) $\{\forall x,\ p(x)\} \wedge \{\forall x,\ \sim q(x)\}$ b) $\{\exists x,\ \sim p(x)\} \wedge \{\forall x,\ q(x)\}$

c) $\forall x,\ p(x) \rightarrow q(x)$ d) $\{\exists x,\ p(x)\} \rightarrow \{\exists x,\ \sim q(x)\}$

3) Write down the validity of the following arguments, for which the premises are given on the left and conclusion on the right.

a) $\sim (p \wedge \sim q),\quad \sim q \vee r,\quad \sim r \qquad \sim p$

b) $(p \wedge q) \rightarrow r,\quad \sim r \vee s,\quad \sim s \qquad \sim p \vee \sim q$

c) $(p \rightarrow q) \rightarrow r, p \wedge s,\quad q \wedge t \qquad r$

4) Derive the following, using rule CP if necessary.

a) $\sim p \vee q,\quad \sim q \vee r,\quad r \rightarrow s \Rightarrow p \rightarrow s$

b) $p \rightarrow (q \rightarrow r) \Rightarrow (p \wedge q) \rightarrow r$

c) $p \rightarrow (q \rightarrow r),\quad q \rightarrow (r \rightarrow s) \Rightarrow p \rightarrow (q \rightarrow s)$

5) show the following set of premises are inconsistent.

a) $p \rightarrow q,\quad p \rightarrow r,\quad q \rightarrow \sim r,\quad p$

b) $p \rightarrow (q \rightarrow r),\quad s \rightarrow (q \wedge \sim r),\quad p \wedge s$

Hence show that $p \rightarrow q,\quad q \rightarrow r,\quad q \rightarrow \sim r,\quad p \Rightarrow m$ and $p \rightarrow (q \rightarrow r),\quad s \rightarrow (q \wedge \sim r),\quad p \wedge s \Rightarrow p$

6) If the universe of discourse is the set $\{a, b, c\}$, eliminate the quantifiers in the following formulas.

a) $(x)P(x)$ b) $(x)R(x) \wedge (x)S(x)$

c) $(x) \sim P(x) \vee (x)P(x)$ d) $(x)(P(x) \rightarrow Q(x))$

7) Give a direct proof for each of the following statements:

(a) The square of an even integer is even integer., (b) The sum of two odd numbers is odd.

(c) If a, b, c are positive integers such that a divides b and b divides c, then a divides c.

8) Prove the following statements by the method of contradiction:

(a) If n^2 is an even number, then n is an even number.

(b) If there are 13 persons in a room, two or more of these have their birthday in the same month.

(c) If the bases and the heights of two triangles are equal, then the areas of the triangles are equal.

2.12 Suggested Readings

- Kolman, Barby, Ross *Discrete Mathematical Structures,* 4*th* Edition, Prentice Hall of India, 2003.
- G. Shankar Rao, *Mathematical Foundations of Computer Science,* I.K. International Publishing House Private Limited, 2006.
- Thomas Koshy, *Discrete Mathematics with Application,* Academic Press An Imprint of Elsevier, 2004.

UNIT-3: Set Theory

Structure

3.0 Objectives

After studying this unit you will be able to:

- Understand the most important concept in mathematics, i.e., "Set"
- Understand the concept of sets in different notations.
- Study the properties & different operations on sets.
- Study the collection of sets i.e., indexed family of sets and cartesian products of sets.
- Study the concept of mathematical induction based on set theory.

3.1 Introduction

The concept of sets in mathematics is the fundamental or basic ideas for the development of higher mathematical concepts. The word set is used in mathematics to mean any well defined collection of items. The items in a set are called the elements of the set. For example,

we can refer to the set of all the employees of a particular company, the set of all ASCII character, the set of all the integers that are divisible by 5. Sets are used to group the objects together. Often the objects in a set have similar properties. We now provide a definition of a set. This definition is an intuitive definition which is not part of a formal theory of set. The description of a set as a collection of objects based on the intuitive notion of an object was stated by German mathematician George Cantor in 1895.

George Cantor (1845-1918) was considered the father of set theory, his contribution in this area including the discovery that the set of real numbers is uncountable. He also noted for his many important contributions to analysis.

3.2 The Concept of a Set

Definition 3.2.1: *A set is a collection of well-defined distinct objects.*
By a *well-defined* collection of objects we mean that it is possible to say, without ambiguity whether a particular object to the collection or not, the objects in a set are distinct, we do not repeat an object over and over in a set.
Examples: 1. Set of natural numbers. **2.** Set of all prime numbers between 1 to 10.
Definition 3.2.2: *A set is denoted by capital letters* $A, B, X, Y, \ldots\ldots$ *and its elements are denoted by small letters* $a, b, x, y \ldots$ *are called the members of that set.*
If an element x belongs to a set A then we write $x \in A$ which is read as "x is an element of A" or "x belongs to a set A ", or "$x \in A$". If there exists an object y which does not belong to the set A, then we express this fact as $y \notin A$.
Which is equivalent to the negation of the statement "y is in A". i.e., $\sim (y \in A) \Leftrightarrow$ *y does not belong to A.*

3.3 Method of describing a set

A specific set can be defined in two ways . If there are only a few elements then they can be listed individually by writing them between bracescurly and placing commas in between.
Examples:
1. The set of all positive odd numbers less than 10 can be written in the following way, $\{1, 3, 5, 7, 9\}$ thus $\{1, 3, 5, 7, 9\}$ represents a set.

Another method of defining a set is by a description of some attribute or characteristic of elements of the set, this method is more general and involves a description of the property. $A = \{x/x \textit{ has the property } P\}$, read it as The set A of all objects x such that x has the property P. The vertical bar "/" is read as such that, the set can also be written as, $A = \{x : x \textit{ has the property } P\}$ in which the symbol ':' is read as such that, this notation

is called as set-builder notation.

2. (a) $A = \{x : \ x \text{ is a complex number. } \}$

(b) $B = \{x : \ x \text{ is a positive integer greater than } 5\}$

we describe a set by its characteristics function.

3.

$$A(x) = \begin{cases} 1 & if\ x \in A \quad (if \ \ x \in A) \\ 0 & if\ x \notin A \quad (if x \in A) \end{cases}$$

In another method describe the set by a recursive formula.

3.3.1. Subset:

Definition 3.3.1 : *If every element of a set A is also an element of a set B, then A is called a subset of the set B.*

If A is a subset of the set B, then we say that A is contained in B and we write $A \subseteq B$ symbolically, if $x \in A \Rightarrow x \in B$, then $A \subseteq B$.

Examples:

1. $A = \{1, 5, 7\}$ is subset of $B = \{1, 2, 3, 4, 5, 6, 7, 8\}$

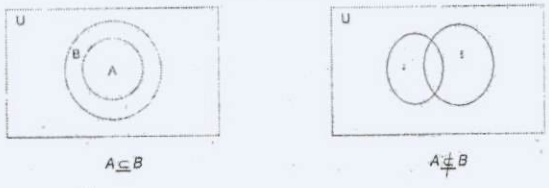
A is not a subset of B is symbolically written as $A \nsubseteq B$, the symbol $\nsubseteq$ denoting is not a subset of or not contained in.

2. If $A = \{x : \ x \text{ is a positive integer with } x^2 < 10\}$, $B = \{1, 2, 3, 4, 5\}$ and $C = \{2, 3, 4, 5, 6\}$ then $A \subseteq B$ and $A \nsubseteq C$.

In the above example, we observe that the B which contains A as a subset possesses elements that are not in A (namely the elements 4 and 5). In such a situation we say that the set A is properly contained in B or A is a proper subset of the set B.

Thus, a set A is a proper subset of set B if $A \subseteq B$ and B possesses at least one element that is not in A, in this situation we write $A \subset B$ here the symbol $\subset$ stand for is a proper subset of. Relationship between sets are often depicted in diagrams called Euler-Venn diagrams (or) simply Venn diagrams for a grasp of the situation figure.

Venn Diagrams:



$A \subseteq B$ $A \nsubseteq B$

Theorem 3.3.1 : *For every set S, (a) $\phi \subseteq S$, (b)$S \subseteq S$.*

Proof: We prove (a) Let S be a set, to show that $\phi \subseteq S$, we must show that $\forall x \ (x \in \phi \rightarrow x \in S)$ is true. Because the empty set contains no elements, it follows that $x \in \phi$ is always false. It follows that the conditional statement $x \in \phi \rightarrow x \in S$ is always true because its hypothesis is always false and a conditional statement with false hypothesis is true, that is $\forall x \ (x \in \phi \rightarrow x \in S)$ is true. This completes the proof.

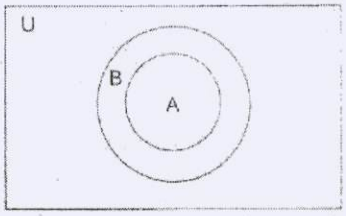
Theorem 3.3.2 : *The total number of subsets of a given set containing n element is* 2^n.
Proof: Let A be an arbitrary set containing n elements. Then, one of its subsets is the empty set. Apart from this, the number of singleton subset of A=n= ${}^nc_1$. The number of subsets of A, each containing 2 elements $={}^nc_2$, the number of subsets of A, each containing 3 elements ${}^nc_3 \cdots$
The number of subsets of A, each containing $(n-1)$ elements ${}^nc_{n-1}$
The number of subsets of A, each containing n elements $={}^n c_n$ total number of subsets of A.

$1 +^n c_1 +^n c_2 + \ldots .^n c_{n-1} +^n c_n$

$(1+1)^n = 2^n$ [*using Binomial Theorem*]

3.3.2. Proper Subset: Set A is called a proper subset of the set B, if
(1) A is subset of B, and
(2) B is not a subset of A i.e., A is said to be a proper subset of B, if every element of A belongs to the set B but there is at least one element of B, which is not in A. If A is proper subset of B, then we denote it by $A \subset B$.

Venn Diagram



(f) $A \subset B$

Note: "Every set A is a subset to itself and ϕ is a subset of A. The subset ϕ and A are called improper subsets of A".

3.3.3. Equal Sets: If A and B are sets such that every element of A is an element of B, and every element of B is also element of A, then A and B are said to be equal sets. If A and B are equal then we write $A = B$, is read as A and B are identical.
Examples: 1. $A = \{2, 4, 6\}$, $B = \{x \in Z : 1 \leq x \leq 7,\ x\ is\ even\}$.
2. If $B = \{5, 7, 9, 11\}$,and $C = \{5, 7, 9, 11\}$ then B=C or C=B.

3.3.4. Super Set: If A is subset of B, then B is called a super set of A.
Example: Let $A = \{0, 1, 3\}$ $B = \{0, 1, 2, 3, 4\}$, B is a super set of A.

3.3.5.Null Set: The set with no element in it is called an empty set or null set. A null set is denoted by the symbol ϕ.
The null set is a subset of every set that is if A is any set then $\phi \subseteq A$.
Example: The set of real roots of $x^2 + 7 = 0$.

3.3.7. Singleton Set: A set having only one element is called a singleton set.
Example: A=$\{5\}$.
Theorem 3.3.3 : *Two sets A and B are equal if and only if $A \subseteq B$ and $B \subseteq A$.*
Proof: If A=B then every element of A is a member of B and every element of B is a member of A, $\therefore A \subseteq B$ and $B \subseteq A$
conversely, let us suppose that $A \subseteq B$, $B \subseteq A$ and $A \neq B$. Since
$A \neq B$ there is an element of B. that is either an element of A. i.e, not in B are there not an element of B, i.e, not in A.
But $A \subseteq B$, so every element of A is in B since $B \subseteq A$, every element of B is in A. Therefore, our assumption that $A \neq B$, leads to a contradiction. Hence, $A = B$.

3.3.7. Finite Set: A set is said to be finite if it has a finite number of elements.
Example: $A = \{2, 4, 6\}$.
If A is a finite set with n elements in it, then n is called the cardinality of A. If n is the cardinality of A, we write $|A| = n$.

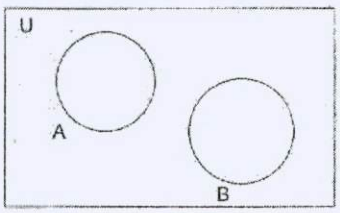
3.3.8. Infinite Set: A set is infinite, if it is not finite.
Example: The set of natural numbers.

3.3.9. Universal Set: In many discussions all the sets under consideration to be the subsets of one particular set. This set is called universal set for that discussion and is denoted by U. The universal set is also denoted by $\mu(or\ X)$
Example: If $A = \{0, 1, 3\}$, $B = \{2, 4, 5\}$, $C = \{6, 7, 8, 9\}$ then the universal set is $U = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$.

3.3.10. Power Set: The set of all subsets of A is called the power set of A. The power set of A is denoted by $P(A)$ or by 2^n. If A has n elements in it, then $P(A)$ has 2^n elements.
Example: If $A = \{1, 2\}$, then $P(A) = \{\phi, \{1\}, \{2\}, A\}$. A set is never equal to its power set.

3.3.11. Equivalent Sets: The number of distinct elements contained by a finite set A is called the cardinal number of A, denoted by n(A). Two finite sets are said to be equivalent, if they have the same number of distinct elements. Equivalent sets are not always equal, but equal sets are always equivalent.
Example: A=$\{1, 2, 3\}$ and B=$\{a, b, c\}$ are equivalent, but not equal.

3.3.12. Disjoint Sets: Two sets A and B are said to be disjoint, if they have no elements in common.



A, B disjoint

Example: The sets A=$\{1, 2, 3\}$, and B=$\{2, 4, 6\}$ are not disjoint sets.

3.4 Properties of Sets Containment

Theorem 3.4.1 : *If A is any set then A contains A.*
Proof: If $x \in A$, then by the repetition of the statement, A contains A.
Theorem 3.4.2 : *If $A \subseteq B$ and $B \subseteq C$, then $A \subseteq C$ where A, B and C are sets.*
Proof: Let $x \in A$, $\therefore x \in A$

$\Rightarrow x \in B$ (because $A \subseteq B$)

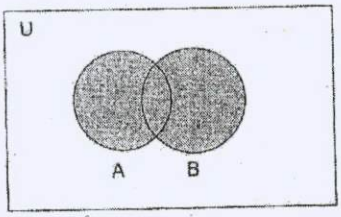
$\Rightarrow x \in C$ (because $B \subseteq C$)

$x \in A \Rightarrow x \in C$

$\therefore A \subseteq C$.

3.5 Operation on Set

3.5.1. Union of sets: Let A and B are sets. The union of sets A and B is the set that contain those elements that are either in A or B or in both. The union of the sets A and B is denoted by $A \cup B$ and read as "A union B".
Symbolically: $A \cup B = \{x : x \in A (or) x \in B\}$ or
$A \cup B = \{x : x \in A \vee x \in B\}$



$A \cup B$ (shaded)

Example: If A=$\{2, 5, 7\}$ B=$\{1, 3, 5, 8\}$ then $A \cup B = \{1, 2, 3, 5, 7, 8\}$

3.5.2. Union of more than Two Sets: If A_1,A_2.....A_n denote sets, then the union

of sets denoted by $\bigcup_{i=1}^{n} A_i$ is defined as the set $\bigcup_{i=1}^{n} A_i = \{x : x \in A_i \text{ for at least one } i\}$

3.5.3. Properties of Union Operation:

Theorem 3.5.1 : *The following properties hold for the union of sets A and B*

(a) $A \cup A = A$ *(Idempotent law)*

(b) $A \cup B = B \cup A$ *(Commutative law)*

(c) $(A \cup B) \cup C = A \cup (B \cup C)$ *(Associative law)*

(d) $A \cup \phi = A$ *(identity law)*

(e) $A \cup U = U$ *(identity law)*

Proof: (a) $A \cup A = \{x : x \in A \vee x \in A\}$

$= \{x : x \in A\} = A$

$\therefore A \cup A = A.$

(b) $A \cup B = \{x : x \in A \vee x \in B\}$

$= \{x : x \in B \vee x \in A\}$

$\therefore A \cup B = B \cup A$ that is, union of sets is commutative.

(c) $A \cup (B \cup C) = \{x : x \in A \text{ or } x \in B \cup C\}$

$= \{x : x \in A \text{ or } (x \in B \text{ or } x \in C)\} = \{x : x \in A \text{ or } x \in B \text{ or } x \in C\}$

$= \{x : (x \in A \text{ or } x \in B) \text{ or } x \in C\}$

$= \{x : x \in A \cup B \text{ or } x \in C\}$

$= (A \cup B) \cup C$

$\therefore$ union of sets is associative.

(d) $A \cup \phi = \{x : x \in A \text{ or } x \in \phi\}$

$= \{x : x \in A\} = A$

$\therefore A \cup \phi = A$

(e) clearly $A \cup U \subseteq U$:

Let $x \in U$ then $x \in U \Longrightarrow x \in A$ or $x \in U$

$\Longrightarrow x \in A \cup U$

$\therefore U \subseteq A \cup U.$

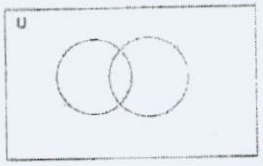
Hence $A \cup U = U.$

3.5.4. Intersection of Sets: The intersection of sets A and B is the set whose elements are all of the elements common to both A and B. The intersection of the sets A and B is denoted by $A \cap B$.

Symbolically: $A \cap B = \{x : x \in A \text{ and } x \in B\}$ or

$A \cap B = \{x : x \in A \wedge x \in B\}$

$A \cap B$ is read as A intersection B.



$A \cap B$ (shaded portion)

If$A \cap B = \phi$, then A and B are called disjoint sets.

3.5.5. Intersection of more than two sets: If $A_1, A_2, ...A_n$ denote sets, then the intersection of these sets denoted by $\bigcap_{i=1}^{n} A_i$ is defined as follows:
$\bigcap_{i=1}^{n} A_i = \{x : x \in A_i \ for \ every \ i \ (i = 1, 2, \ldots n)$
Example 1: If $A = \{1, 2, 4, 6\}$, $B = \{2, 4, 6, 8\}$ then $A \cap B = \{2, 4\}$.

3.5.6. Properties of sets Intersection:
Theorem 3.5.2 : *The following properties hold for the intersection of the sets.*

(a) $A \cap A = A$ *(Idempotent law).*
(b)$A \cap B = B \cap A$ *(Commutative law).*
(c)$(A \cap B) \cap C = A \cap (B \cap C)$ *(Associative law).*
(d) $A \cap \phi = \phi$ *(Identity law)*
(e) $A \cap U = A$ *(Identity law).*

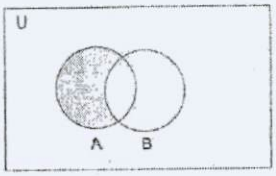
Where A and B are sets, U is the universal set.

3.5.7. Distributive Laws:
Theorem 3.5.3 : *The following properties hold for operation of set union and intersection.*

(a) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
(b)$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

for all sets A,B and C. These properties are called distributive properties.

3.5.8. Difference of Sets: If A and B are subsets of the universal set U, then the relative complement of B in A is the set of all elements in A which are not in A. It is denoted by $A - B$.
$\therefore A - B = \{x : x \in A \ and \ x \notin B\}$.



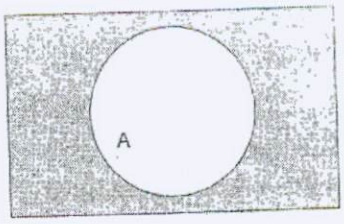
$A - B$ (shaded portion)

Example: Let $A = \{1, 2, 3\}$, $B = \{2, 3, 4, 5, 6\}$ then $A - B = \{1\}$.

3.5.9. Compliment Set: Let U be the universal set. The complement of the set A, denoted by A^c, is the complement of the set A with respect to U. In other words, the complement of the set A is $U - A$.

$$A^c = \{x : \ x \notin A\}$$

The complement of A is also denoted by $\overline{A}$



The set $\overline{A}$ (shaded portion)

3.5.10. Properties of Complementation:

Theorem 3.5.4 : *If A and B are subsets of a universal set U, then the following properties hold:*

(a) $\bar{U} = \phi$ *(b)* $\bar{\phi} = U$

(c) $A \cup \bar{A} = U$ *(d)* $A \cap \bar{A} = \phi$

(e) $\bar{A}A = A$ *(f)* $A \subseteq B \Rightarrow \bar{B} \subseteq \bar{A}$

(g) $A \bar{\cap} B = \bar{A} \cup \bar{B}$ *(De Morgan's law)* *(h)* $A \bar{\cup} B = \bar{A} \cap \bar{B}$

3.5.11. Properties of Difference:

Theorem 3.5.5 *If A and B are subsets of a universal set U, then*

(a) $A - B = A \cap \bar{B}$

(b) $\bar{A} = U - A$

(c) $A - A = \phi$

(d) $A - \phi = A$

(e) $A - B = B - A$, *if and only if* $A = B$

(f) $A - B = A$, *if and only if* $A \cap B = \phi$

(g) $A - B = \phi$, *if and only if* $A \subseteq B$

3.5.12. Symmetric Difference: The symmetric difference of the sets A and B is the set containing those elements either A or B, but not in both A and B.

In other words, the symmetric difference of the sets A and B is the relative complement of $A \cup B$ in $A \cap B$. The symmetric difference of A and B is denoted by $A \triangle B$ (or by $A \oplus B$)

Symbolically: $A \triangle B = \{x : x \in A \cup B, \text{ and } x \notin A \cap B\}$

from the definition, we have $A \triangle B = (A - B) \cup (B - A)$

Example: Let A=$\{1,2,3,4,5,6,7\}$, B=$\{3,4,a,b,c,d\}$
then $A \cup B = \{1,2,3,4,5,6,7,a,b,c,d\}$, $A \cap B = \{3,4\}$
$\therefore A \triangle B = \{1,2,5,6,7,a,b,c,d\}$

3.5.13. Properties Of Symmetric Difference:
Theorem 3.5.7 : *If A and B are sets, then*

(a) $A \triangle B = \phi$
(b) $A \triangle B = B \triangle A$
(c) $A \triangle \phi = A$
(d) $(A \triangle B) \triangle C = A \triangle (B \triangle C)$
(e) $A \triangle B = (A \cup B) - (A \cap B) = (A - B) \cup (B - A)$

3.5.14. Principle Of Duality: The principle of duality states that any established result involving sets and complements and operation of union and intersection give a corresponding dual result by replacing $\cup$ by ϕ, and $\cup$ by $\cap$, and vice versa.
Example: Consider $A \cup \bar{A} = U$, then applying the principle of duality, we can write $A \cap \bar{A} = \phi$
Solved Problems:
1. Show that the empty set is unique.
Solution : If possible let ϕ_1 and ϕ_2 be two empty sets. Empty set is a subset of every set, therefore $\phi_1 \subseteq \phi_2$ *and* $\phi_2 \subseteq \phi_1$ Hence, $\phi_1 = \phi_2$
2. Prove that $A - (A - B) = A \cap B$
Solution: We have,

$$\begin{aligned} A - (A - B) &= A - (A \cap \bar{B}) \\ &= A \cap (\bar{A} \cup \bar{B}) \\ &= A \cap (\bar{A} \cup B) \\ &= (A \cap \bar{A} \cup (A \cap B)) \\ &= \phi \cup (A \cap B) \\ &= A \cap B \\ \therefore A - (A - B) &= A \cap B \end{aligned}$$

3.5.15. Set Inclusion: Let A and B be two sets. We say A is a subset of B (or A is contained in B, or B contains A,or B includes) written as $A \subseteq B$ *or* $B \supseteq A$ if every member of A is a member of B formally
$A \subseteq B \Leftrightarrow \forall x, (x \in A \Rightarrow x \in B)$
the symbol '$\subseteq$' is called the set inclusion symbol. If A is not a subset of B we write $A \nsubseteq B$ instead of the longer statement $\sim (A \subseteq B)$

Examples:

1. $\{1,3\} \subseteq \{1, \Pi, 3\}$

2. $\{\phi\} \subseteq \{\phi, 5\}$

3. $\mathbf{N} \subseteq \mathbf{Z} \subseteq \mathbf{Q} \subseteq \mathbf{R}$.

3.6 Indexed Family of Sets

Consider a non empty set $\wedge$ to each element $a \in \wedge$, associate a set denoted by A_α then the family of sets $\{A_\alpha : \alpha \in \wedge\}$ or $\{A_\alpha\}_{\alpha \in \wedge}$ called indexed family of sets, indexed by the set $\wedge$. This set is also called the arbitrary family of sets. The set $\wedge$ is called index set. In particular, if the index set is the set of natural numbers then the indexed family of sets is given by

$$\{A_i : i \in N\} \ or \ \{A_i\}_i \in N$$

Examples:

1. Let $\wedge = \{1,2,3\}$ for each $\alpha \in \wedge$, let $A_\alpha = \{1, \frac{1}{2}, \frac{1}{3}, ..., \frac{1}{\alpha}\}$.
The index family of sets is given by $\{A_1, A_2, A_3\}$.
Here $A_1 = \{1\}, A_2 = \{1, \frac{1}{2}\}, A_3 = \{1, \frac{1}{2}, \frac{1}{3}\}$

2. Let $\wedge = N$ =Set of natural numbers. Now each $n \in N$,
A_n = the set of all rational numbers with denominator n,
then the indexed family of sets is $\{A_1, A_2, ...A_n....\}$ where $A_1 = \{0, \pm 1, \pm 2, \pm 3, ... \pm n..\}$
$A_2 = \{o, \pm\frac{1}{2}, \pm\frac{2}{2}, \pm\frac{3}{2}, ...\}$ etc.,

3. Let the set N of all natural number be the index set and A_n be the set of all real numbers in the interval $(-n, n)$ then the indexed family of set is given by

$$A_n : n \in NA_1, A_2, ...A_n$$

Where $A_1 = \{(-1,1), \ A_2 = (-2,2), \ A_3 = (-3,3), \ldots\}$.

3.6.1. Set Operations On Indexed Family Of Sets: The operation like union, intersection can also be defined on indexed family of sets.
Let $A_\alpha, \ \alpha \in \wedge$ be an arbitrary family of sets.

3.6.2. Union of indexed family of the sets: The union of the family of sets $\{A_{\alpha\alpha\in\wedge}$ is denoted by $A_\alpha\}$, read as arbitrary union of sets A_α for $\alpha \in \wedge$ and defined as the set,
$\bigcup_{\alpha\in\wedge} A_\alpha = x : x \in A_\alpha \ for \ some \ a \in \wedge$
i.e., $x \in \bigcup_{\alpha\in\wedge} A_\alpha \Rightarrow \exists \ a \in \wedge$ such that $x \in A_\alpha$
$\Rightarrow$ x belongs to at least one of the member of the family. Further note that $x \notin \bigcup_{\alpha\in\wedge} A_\alpha \Rightarrow$ x is not an element of the any of the member of the family.
i.e., $x \notin \bigcup_{\alpha\in\wedge} A_\alpha \Rightarrow x \notin A_\alpha, \forall\alpha \in \wedge$

Example: Let the set N of natural number be the index set and for $n \in N$, Let A_n be the set of all rational number with denominator n then clearly one can see
$\bigcup_{n\in N} A_n = Q$ = set of all rational numbers.

3.6.3. Intersection of Indexed Family of Sets: The intersection of family of sets $\{A_\alpha\}_{\alpha\in\wedge}$ is denoted by $\cap A_\alpha$, read as arbitrary intersection of sets A_α, $\alpha \in \wedge$ and defined as the set
$\cap_{\alpha\in\wedge} A_\alpha = \{x : x \in A_\alpha \text{ for every } \alpha \in \wedge\}$ i.e., $x \in \cup_{\alpha\in\wedge} A_\alpha \Rightarrow \forall\alpha \in \wedge,\ x \in A_\alpha$
$\Rightarrow$ x belongs to every member of the family. Further note that $x \notin \cap_{\alpha\in\wedge} A_\alpha$
$\Rightarrow\ \exists\ \alpha \in \wedge$ such that $x \notin A_\alpha$
$\Rightarrow x$ is not an element of at least one of the member of the family.
i.e., $x \notin \bigcup_{\alpha\in\wedge}$
$A_\alpha \Rightarrow x \notin A_\alpha$ for some $\alpha \in \wedge$.
Example : Let the set N of natural numbers be the index and for each $n \in N$, det A_n be the set of all rational numbers with denominator n. then clearly

$$\cap_{n\in N} A_n = Z = \text{Set of all integers.}$$

3.7 Cartesian product

Definition 3.7.1 : *If A and B are nonempty sets, then the cartesian product of A and B is the set of all ordered pairs (a, b) where $a \in A$ and $b \in B$.*
The product $A \times B$ *of set* A *and* B *is the sets* $(a, b) : a \in A$ *and* $B \in B$.
Examples:
1. If $A = \{a, b\}, B = \{p, q, r\}$ then
$A \times B = \{(a, p), (a, q), (b, p), (b, q), (b, r)\}$
2. If $A = \{0, 1, 3, 5\}, B = \{2, 3, 4\}$
then $A \times B = \{(0, 2), (0, 3), (0, 4), (1, 2), (1, 3), (1, 4), (3, 2), (3, 3), (3, 4), (5, 2), (5, 3), (5, 4)\}$
3. If A and B are both Sets of real numbers, then $A \times B$ is the cartesian plane.
Theorem 3.7.1 : *If A B and C are nonempty sets then*

i) $A \times (B \cup C) = (A \times B) \cup (A \times C)$
ii) $A \times (B \cap C) = (A \times B) \cup (A \times C)$

Proof: i)

$$\begin{aligned}
A \times (B \cup C) &= \{(a, b) : a \in A, \text{and } b \in (B \cup C)\} \\
&= \{(a, b) : a \in A, \text{and } (b \in B \text{ or } b \in C)\} \\
&= \{(a, b) : (a \in A, \text{ and } b \in B) \text{ or } (a \in A \text{ and } b \in B) \in C\} \\
&= \{(a, b) : (a, b) \in A \times B \text{ or } (a, b) \in (A \times C)\} \\
&= \{(a, b) : (A \times B) \cup (A \times C)\}
\end{aligned}$$

$$= (A \times B) \cup (A \times C)$$

ii)

$$\begin{aligned} A \times (B \cap C) &= \{(a,b) : a \in A,\ b \in (B \cap C)\} \\ &= \{(a,b) : a \in A,\ b \in B \text{ and } b \in C\} \\ &= \{(a,b) : a \in A,\ b \in B \text{ or } a \in A,\ b \in C\} \\ &= \{(a,b) : (a,b) \in A \times B \text{ and } (a,b) \in A \times C\} \\ &= \{(a,b) : (a,b) \in (A \times B) \cap (A \times C)\} \\ &= (A \times B) \cap (A \times C) \end{aligned}$$

Theorem 3.7.2 : *If A, B and C are nonempty sets, then $A \subseteq B \Rightarrow A \times C \subseteq B \times C$.*
Proof: Let (a,b) be any element of $A \times B$, then

$(a,b) \in A \times C \Rightarrow a \in A,\ b \in C$
$\Rightarrow a \in B,\ b \in C (\because A \subseteq B)$
$\Rightarrow (a,b) \in B \times C$
$\therefore A \times C \subseteq B \times C$

3.7.1. Cartesian Product of Sets: Let $A_1, A_2, ...A_n$ denote n sets where $n \geq 2$, then the Cartesian product $A_1 \times A_2..... \times A_n$ is the set of all n-tuples of the form $a_1, a_2, ...a_n$ where $a_1 \in A, a_2 \in A_2....a_n \in A_n$.
From the definition, $A_1 \times A_2..... \times A_n = (a_1, a_2, ...a_n) : a_1 \in A_1, a_2 \in A_2....a_n \in A_n$

3.8 Mathematical Induction

Mathematical induction is a powerful technique for proving statements about large sets, such as the set N of natural numbers n has number. If we can prove that every natural number n has a certain property p, then we have in effect proved infinitesimally many problem $p_1, p_2, ...p_n$. The miracle of mathematical of mathematical induction is that is enable use to carry out such a task in only a finite number of steps. Induction is closely linked to the notion of recursion which is structural frame work for the definitions of many mathematical structure. Many mathematical statements assert that a property is true for all positive integers.

Examples of such statements are that for every positive integer n : $n! \leq n^n.n^3 - n$ is divisible by 3; a set with n element s has 2^n subsets and sum of the first n positive integer is $\frac{n(n+1)}{2}$. A major goal of this is to give understanding of mathematical induction which is used to prove results of this kind. The first known use of mathematical induction is in the work of 16^{th} century mathematician Francesco-Maurolico (1494-1475). Maurolico wrote extensively on the works of classical mathematics and made many contribution to Geometry

and Optics. Maurolico presented a variety of properties of the integers together with proofs of these properties. To prove some of these properties he derive the method of mathematical induction.

3.8.1. The principle of Mathematical Induction: Suppose there is a given statement $p(n)$ involving the natural number n such that

(a) The statement is true for $n = 1$ i.e, $p(1)$ is true and

(b) If the statement is true for $n = k$ (where k is some positive integer) then the statement is also true for $n = k + 1$.

i.e, truth of $p(k)$ implies the truth of $p(k + 1)$ then, $p(n)$ is true for all natural number n.

Property (a) is simply a statement of fact there may be situation when statement is true for all $n \geq 4$. In this case, step will start from $n = 4$ and we shall verify the result for $n = 4$ i.e, $p(4)$.

Property (b) is a conditional property, it does not assert that the given statement is true for $n = k$, but only that if it is true for $n = k$, then it is also true for $n = k + 1$, so to prove that the property holds only prove that conditional propositional. If the statement is true for $n = k$, then it is also true for $n = k + 1$.

This is sometimes is true for $n = k$, then it is also true for $n = k + 1$.

This is some times reffered to as the inductive step. The assumption that the statement is true for $n = k$ in this inductive step is called the inductive hypothesis.

3.9 Proving summation Formulas

We begin by using mathematical induction to prove several different summation formulae. As we will see, mathematical induction is particularly well suited for proving that such formulae are valid. However summation formulae can be proved in other ways. This is not surprising because there are often different way to prove a theorem. The major disadvantage of this use of mathematical induction is that you can not use it to derive a summation formula. That is you must already have the formula before you attempt to prove it by mathematical induction. We begin by using mathematical induction to prove a formula for the sum of the smallest n positive integer.

Examples:

1. Show that if n is a positive integer, then $1 + 2 + ... + n = \frac{n(n+1)}{2}$.

Solution: Let $p(n)$ be the proposition that the sum of the first n positive integers is $\frac{n(n+1)}{2}$. We must do two things to prove that $p(n)$ is true for $n = 1, 2, ...$ namely we must show that $p(1)$ is true and that the conditional statement $p(k)$ implies $p(k + 1)$ is true for $k = 1, 2, ...$

Basic Step:- $p(1)$ is true because $1 = \frac{1(1+1)}{2}$

Inductive Step:- For the inductive hypothesis we assume that $p(k)$ holds for an arbitrary

positive integer k that is we assume that $1+2+.....+k=\frac{k(k+1)}{2}$
under this assumption, it must be shown that $p(k+1)$ i.e.,

$$1+2+.....+k+(k+1)=\frac{(k+1[(k+1)+1])}{2}$$
$$=\frac{(k+1)(k+2)}{2}$$

is also true when we add $k+1$ to both sides of the equation in $p(k)$ we obtain

$$1+2+.....+k+(k+1)=\frac{k(k+1)}{2}+(k+1)$$
$$=\frac{k(k+1)+2(k+1)}{2}$$
$$=\frac{(k+1)(k+2)}{2}$$

This last equation shown that $p(k+1)$ is true under the assumption that $p(k)$ is true. This completes the inductive step.

We have completed the step and inductive step, so by mathematical induction we known that $p(n)$ is true for all positive integer n that is we have proved that $1+2+.....+n=\frac{n(n+1)}{2}$ for all positive integers n.

2. Prove that $2^n > n$ for all positive integer n.

Solution: Let $p(n) : 2 > n\ 2^2$¿2

when $n=1, 2^1>1$.

Hence $p(1)$ is true

assume that $p(k)$ is true for any positive integer k $2^k > k$

we shall now prove that $p(k+1)$ is true whenever $p(k)$ is true.

Multiplying both sides of (1) by 2 we get

$2.2^k > 2k$

i.e, $2^{(k+1)} > 2k = k+k > k+1$

therefore $p(k+1)$ is true when $p(k)$ is true. Hence by principle of mathematical induction, $p(n)$ is true for every positive integer n.

Check Your Progress

(1) Describe different types of set with example.
(2) State and prove, distributive laws of set.
(3) Explain properties of complimentation.
(4) Prove that $2^n > n$ for all positive integers n.

3.10 Summary

- We introduced the notation of sets: A set is a collection of well defined distinct objects.
- We are study the different methods of describing a set.
- Different types of sets like proper set, equal sets, super set, null set, finite set, infinite set, universal set, power set, disjoint sets were defined.

- We are defining operation on sets: Union, intersection, properties of intersection were studied.
- We define the cartesian of two sets.
- Explain the concept of mathematical induction and proving summation formulas.

3.11 Key Words

Sets, Types of sets, Union, Intersection, cartesian product, mathematical induction.

3.12 Answers Check Your Progress

(1) 3.3, (2) 3.5.7, (3)3.5.10, (4)3.8 (example 2).

3.13 Exercise and Answers

1) Suppose $A \subseteq B$ Show that a) $(A \cap C) \subseteq (B \cap C)$ b) $(A \cup C) \subseteq (B \cup C)$

2) Prove that : a) If $B \subseteq A$, then $(B - C) \subseteq (A - C)$ b) If $B \subseteq A$, then $A - (A - B) = B$

c) $A \cup B = A \cap B$ if and only if $A = B$ d) $(A - B) \cup B = A$ if and only if $A \supseteq B$.

3) Prove that $A \times B = \Phi$ if $A = \phi$ or $B = \phi$.

4) Suppose $A = 1, 2, 3$ and $B = a, b$ then show that $A \times B \neq B \times A$.

5) List the members of the sets

a) $\{x/x$ is a real number $\ni x^2 = 1\}$.

b) $\{x/x$ is a positive integer less than $12\}$.

c) $\{x/x$ is the square of an integer and $x < 100\}$.

6) Use a Venn diagram to illustrate the set of all months of the year whose names donot contain the letter R in the set of all months of the year.

7) Use a Venn diagram to illustrate the relationship $A \subseteq B$ and $B \subseteq C$.

8) Can you conclude that $A = B$ if A and B are two sets with the same power set?

9) Let A, B, C be sets, show that $\overline{A \cup (B \cap C)} = (\overline{C} \cup \overline{B}) \cap \overline{A}$

10) Does every set have proper subset? Explain with example.

11) Show that If A and B are disjoint, then $A - B = A$ and $B - A = B$.

12) Prove that: a) $A \triangle \phi = A$ b) $A \triangle (B \triangle C) = (A \triangle B) \triangle C$.

Answers:

5. a)$\{1, 2\}$, b)$\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11\}$, c) $\{1, 4, 9, 16, 25, 36, 49, 64, 81\}$

3.14 Suggested Readings

- Kolman, Barby, Ross *Discrete Mathematical Structures, 4th* Edition, Prentice Hall of India, 2003.
- G. Shankar Rao, *Mathematical Foundations of Computer Science,* I.K. International Publishing House Private Limited, 2006.
- Thomas Koshy, *Discrete Mathematics with Application,* Academic Press An Imprint of Elsevier, 2004.

UNIT-4: Counting Principle

Structure

4.0 Objectives

After studying this unit you will be able to:

- Understand the concept of counting the objects.
- Understand basic principles of counting.
- Study the concept of permutation & combinations which are very much useful in computer applications.
- Study the repetitions of permutations & combinations.
- Study the concept of pigeonhole principle and it's applications.

4.1 Introduction

Combinatorics, the study of arrangements of objects is an important part of discrete mathematics. This subject was studied as long ago as the seventeenth century. When combination questions arose in the study of gambling games, Enumeration, the counting of objects with certain properties, is an important part of combinatorics. We must count object

to solve many different type of problems for instance. Counting is used to determine the complexity of algorithms. Counting is also required to determine whether there are enough telephone numbers or internet protocol addresses to meet demand. Further more, counting techniques are used extensively when probabilities of events are computed.

The basic rule of counting, which we will study can solve a tremendous variety of problems. For instance, we can use these rules to enumerate the different phone numbers on a computer system and the distinct order in which the runners in a race can finish, on other important combinatorial tool is the "**Pigeonhole Principle**"

We use the counting principle to determine how many, one can choose / do certain events. We can phrase many counting problems in terms of ordered (*or*) unordered arrangements of the objects of a set. These arrangements are called permutations and combinations are used in many counting problems.

4.2 Basic Counting Principle

We will present two basic counting principle. The product rule and the sum rule.

4.2.1. The Product Rule: Suppose that a procedure can be broken down into a sequence of two tasks. If there are n_1 ways to do the first task and n_2 ways to do the second task after the first task has been done. Then there are $n_1 n_2$ ways to do the procedure.

Example 1: In how many ways can this diagram be coloured subject to the following condition?

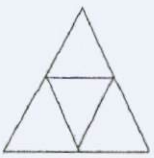


FIGURE 1

(1) Each of the smaller triangle is to be painted with one of the three colours red, blue (or) green.

(2) No two adjacent regions receive the same colour.

Solution: In order to satisfy the condition we first of all paint the central triangle by any of the three colours. Then the remaining triangles can be painted by the two remaining colours.

Associating by the principle of counting, the required number of ways $= 3 \times 2 \times 2 = 24$.

4.2.2. Counting Rules:

Rule 1: If any one of K mutually exclusive and exhaustive events can occur of each of N trials, there are K^N different sequence that may result from a set of such trials example. Flip a coin three times finding the number of possible sequences $N = 3; K = 2$,
therefore $K^N = 2^3 = 8$.

Rule 2: If $K_1, K_2, ...K_N$ are the numbers of distinct events that can occur on trials 1,....N

in a series,the number of different sequence of N events that can occur is $(K_1)(K_2)....(K_N)$
Example: Flip a coin and roll a die, finding the number of possible sequence. Therefore, $(K_1)(K_2) = (2)(6) = 12$.
Rule 3: The number of different ways that N distinct things may be arranged in order is $N! = (1)(2)(3).....(N-1)(N)$, where $0! = 1$.. An arrangement in order is called a permutation, so that the total number of permutations of N objects is $N!$ (the symbol $N!$ is called N-factorial)
Example: Arrange 10 items in order, finding the number of possible ways. Therefore, $10! = 10 \times 9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 3,628,800$.
Rule 4: The number of ways, N, of selecting and arranging r objects from among N distinct objects is: $N!/(N-r)!$, or as seen on calculators, ${}^nP_r$
Example: Pick 3 things from 10 in any order where $N = 10, r = 3$,

$$\therefore \quad \frac{10!}{3!7!} = \frac{720}{6} = 120.$$

Example :(*The telephone numbering plan*)
The format of telephone number in North America is specified by a numbering plan. A telephone number consists of 10 digits. Which are split into a three-digit, area code, and a four digit station code. Because of signaling considerations, there are certain restrictions on some of these digits, to specify the allowable format. Let X denotes a digit that can take any of the value 0 through 9. Let N denotes a digit that can take any of the value 2 through 9 and Y denote a digit that must be $o(or)1$. Two numbering plan, which will be called the old plan and the new plan, but the recent rapid growth in demand of new plans, which for mobile phone and devices make even this new plan obsolete, in this example the letters used to represent digits follow the conventions of the North American numbering plan. As will be shown, the new plan allows the use of more numbers in the old plan, the formate of the area code, office code, and station code are NYX, NXX and $XXXX$, respectively.

So that telephone numbers had the form $NYX - NNX - XXXX$. In the new plan the format of these codes are NXX, NXX and $XXXX$, respectively. So that telephone numbers have the form $NXX - NXX - XXXX$
How many deferent North American telephone numbers are possible under the old plan and under the new plan?
Solution: By the product rule, there are $8.2.10 = 160$ area codes with format NYX and $8.10.10 = 800$ are codes with format NXX. Similarly, by the product rule, there are $8.8.10. = 640$ office codes with format NNX. The product rule also shows that there are $10 \times 10 \times 10 = 10,000$ station codes with format $XXXX$.

Consequently, applying the product rule again, it follows that, under the old plan there are $160 \times 640 \times 10,000 = 1,024,000,000$ different numbers available in North America, under the new plan there are $800 \times 800 \times 10.000 = 6,400,000,000$ different numbers available.

4.2.3. Product Rule For Differentiation: Let $y = u.v$, where both u and v are

differentiable function.

Derivative of the product of two functions =first function × derivative of second function + second function× derivative of first function i.e,

$$\text{y=u.v}$$
$$\Rightarrow (u.v)' = u.v' + v.u'$$

4.2.4. The Sum Rule: If a first task can be done in n_1 ways and a second tasks can be done in n_2 ways and if these task cannot be done in the same time then there are $n_1 + n_2$ ways to do one of these tasks.

Example : A student can choose a computer project from one of three lists. The three contain 23,15 and 19 possible project respectively. How many possible projects are due to choose from?

Solution: The student can choose a projects from the first list in 23 ways, from the second list in 15 ways and from the third list in 19 ways.

Hence there are $23 + 15 + 19 = 57$ project to choose from.

4.3 Permutations

4.3.1 Permutations: *A permutation of n distinct elements $x_1, ..., x_n$ is an ordering of n elements $x_1, ..., x_n$.*

(or) *A permutation of a set of distinct objects is an ordered arrangements of these objects. We also are interested in ordered arrangements of some of the elements of a set. An ordered arrangement of r elements of sets is called an r-permutations.*

Theorem 4.3.1: There are $n!$ permutation of n elements.

Proof: We use the multiplication principle. A permutation of n elements can be constructed in n successive steps: select the first element: select the second element:....:select the least element. The first element can be selected in n ways. Once the first elements has been select, the second element can be selected in $n - 1$ ways. Once the second element can be selected, the third element can be selected in $n - 2$ ways and so on. By the product principle, there are

$n(n-1)(n-2)\ldots 2.1 = n!$ permutations of n elements.

Example: How many permutation of the letters $ABCDEF$ contain the letters DEF together in any order?

We can solve the problems by a two step procedure: select an ordering of the lettersDEF: construct a permutation of $ABCDEF$ containing the given ordering of the letters DEF. By Theorem(4.3.1), the first step can be done in $3! = 6$ ways and the second step can be done in 24 ways. By the product principle, the number of permutations of the letters $ABCDEF$ containing the letters DEF together in any order is

$6 \times 24 = 144$.

Theorem 4.3.2: If n is a positive integer and r is an integer with $1 \leq r \leq n$, then there are $p(n,r) = n(n-1)(n-2)...(n-r+1)$ r-permutations of a set with n distinct elements.
Proof: We will use product rule to prove that this formulas is correct. The first element of the permutation can be chosen in n ways because there are n elements in the set. There are $n-1$ ways to choose the second element of the permutation, because there are $n-1$ elements left in the set after using the element picked for the first position. similarly there are $n-2$ ways to choose the third element, and so on until there are exactly $n-(r-1) = n-r+1$ ways to choose the r^{th} element. Consequently, by the product rule, there are $n(n-1)(n-2)....(n-r+1)$ r-permutation of the set. Note that, $p(n,0) = 1$ whenever n is a non-negative integer because there is exactly one way to order zero elements. That is, there is exactly one list with no elements in it, namely the empty list.
Theorem 4.3.3: The number of r-permutations of a set of n (distinct) element is given by

$$p(n,r) = \frac{n!}{n-r}!$$

.

Proof: Since there are n elements, the first element can be chosen in n ways. Now $n-1$ elements are left; so the second element can be chosen in $n-1$ ways. Continue like this until the r^{th} element is ready to be chosen. At this point there are $n-r+1$ elements left. Consequently, the r^{th} elements can be chosen in $n-r+1$ ways. Thus by the product principle

$$p(n,r) = n(n-1)(n-2)....(n-r+1)$$
$$= \frac{n(n-1)(n-2)....(n-r+1)(n-r)...2.1}{(n-r)...2.1} = \frac{n!}{(n-r)!}$$

suppose we let r=n in theorem. Then $p(n,n) = \frac{n!}{(n-n)!} = \frac{n!}{0!} = \frac{n!}{1!} = n!$
Examples:
1. A photographer would like to arrange 10 cats for a television commercial. How many ways can she arrange them in a row?
Solution: Since all the cats have to be in the commercial at the same time $r = n = 10$. Therefore, the number of possible arrangement is $p(10,10) = 10! = 3,628,800$.
2. Find the number of words that can be formed by scrambling the letters of the word SCRAMBLE (Remember a word is just an arrangement of symbol: it need not make sense)
Solution: The word SCRAMBLE contains eight distinct letters. Therefore, the number of words that can be formed in the word that can equals the number or arrangement of the letters in word, namely $p(8,8) = 8! = 40,320$.
3. In how many ways can 8 persons be seated at a round table if they can set any where?

4.3.2 Permutation with like elements:

Theorem 4.3.4: The number of permutation of n objects of which are $q_1!$ are alike $q_2!$ are alike.... $q_r!$are alike's

$$p(nq_1q_2q_3....q_r) = \frac{n!}{q_1!q_2!q_3!....q_r!}$$

where $n = q_1 + q_2 + q_3 + ... + q_r$

Proof: The number of permutations be x, if the q_1 like objects are unlike, then for each of these x arrangements the q_1 like objects could be rearranged among themselves in $q_1!$ ways without altering the positions of the objects.

Therefore, the number of permutations would be $xq_1!$. Similarly, if all the objects were unlike, the number of permutations would be $xq_1!q_2!q_3!....q_r!$

But if all the objects were unlike, the number of permutation with the n objects would be $n!$. Hence

$$xq_1!q_2!q_3!....q_r! = n!$$

$$x = \frac{n!}{q_1!q_2!q_3!....q_r!}$$

$$p(nq_1q_2q_3....q_r) = \frac{n!}{q_1!q_2!q_3!....q_r!}$$

Example: There are 4 black, 3 green and 5 red balls. In how many ways can they be arranged in a row?

Solution: Total number of balls 4 $black + 3$ $green + 5$ $red = 12$

The black balls are alike

The green balls are, and the red balls unlike

Therefore the number of ways in which the balls can be arranged in a row

$$= \frac{12!}{4!3!5!} = 27,720$$

4.3.3 Circular Permutations: A circular permutation of n objects is an arrangement of the objects around a circle. In circular arrangement we have to consider the relative positions of the different things. The circular permutations are different only when the relative order of the objects is changed otherwise they are same.

4.3.4 Number of Circular Permutations: Let n distinct objects be given. If the n objects are to be arranged round a circle we take an objects and fix it in one position.

Now the remaining $(n-1)$ objects can be arranged to fill the $(n-1)$ positions in the circle $(n-1)!$ ways.

Therefore the number of circular permutation of n different objects= $(n-1)!$

4.3.5 Number of Different Circular Permutations: We consider the order, clockwise (ro anticlockwise) of objects around a circle as the same circular permutation. Every arrangement with n objects around is counted twice in $(n-1)!$ Circular permutations. The total number of different permutation of n distinct objects is $= \frac{(n-1)!}{2}$.

Examples:

1. In how many ways can a party of 9 persons arrange themselves around a circular table?
Solution: one person can set at any place in circular table. The other 8 persons can be arranged themselves in 8! ways that is the 9 persons can be arranged among themselves round the table in $(9-1)! = 8!$ ways.
2. In how many ways 5 gents and 4 ladies dine at a round table, if no two ladies are to sit together ?
Solution: Since no two ladies are to sit together they should seat themselves in between gents (That is a lady is to seated in between two gents).The 5 gents can sit round the circular table in 5 positions (marked G in fig(5)). They can be arranged in $(5-1)! = 4!$ ways. The ladies can sit in the 4 out of 5 seats (marked X in the fig(5)), this can be done in $p(5,4)$ ways. Therefore the required number of ways in which 5 gents and 4 ladies can sit round a table

$$= 4!\ p(5,4) = (4 \times 3 \times 2 \times 1) \times (5 \times 4 \times 3 \times 21) = 2,880.$$

4.3.6. Cyclic Permutations: In how many different ways can you place 5 beads on a necklace ? The answer is not $5! = 120$. But for less. Since it contains a lot of duplicate arrangements. For instance, the tow circular arrangements show in below fig(7) are identical (look at the relative positions of the beads p_1 through p_5)each circular arrangement is a cyclic permutation.
Before we find the number of cyclic permutations of the five beads in below Example1 the following general result will be useful to prove.
Theorem 4.3.5: The number of cyclic permutations of n (distinct) items is $(n-1)!$.
Proof: To avoid duplicates, let us assign a fixed position to the first item a_1 around the circle (see fig(8)) now $n-1$ positions are left. So the second item a_2 can be placed in any one of the the $n-1$ positions. Now $n-2$ positions.

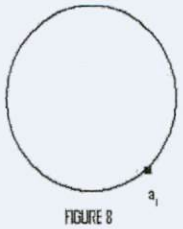


FIGURE 8

are left therefore the third item a_3 can be placed in any of the $n-2$ positions. continue like this until all items have been placed. Thus, by the product principle, the number of cyclic permutations is $1.(n-1)(n-2)....2.1 = (n-1)!$
The next example illustrates this result.
Example: Find the number of different ways five zinnias can be planted in a circle.
Solution: *Number of ways of planting five zinneas in a circle*

$$= \textit{Number of cyclic permutation of five items}$$

$= (5-1)! = 4! = 24$

4.4 Combinations

A combinations of n objects taken at a time is an unordered selection of r of the n objects $(r \leq n)$

A combination of n objects taken r at a time is also called r-combination of n objects.

Examples:

1. The two combination of a, b, c taken two at a time are ab, ac, ad, bc, bd and cd.

2. Consider the objects a, b, c from which the selection are to be made by taking 2 objects at a time.

Solution: The 2-combinations of n objects taken r at a time is denoted by c(n,r). The symbols c(n,r) (or) $c_{n,r}$ (or) ${}^nC_r$; $\begin{pmatrix} n \\ r \end{pmatrix}$ and $c_{n,r}$ are also denote r-combination of n objects.

Definition 4.4.1

Suppose we are interested in selecting (choosing) r-objects from a set $n \leq r$ objects without regard to order. The set of r objects being selected is traditionally called a combination of r objects.

The r-combination of a set of n elements, where $0 \leq r \leq n$ is a subset containing r elements.

The number of r-combinations of a set with n-elements is denoted by $C(n, r)$ (or) $\begin{pmatrix} n \\ r \end{pmatrix}$. Both notations frequently appear in combination. The number of combinations is also called the 'binomial co-efficient.

Before deriving a formula for ${}^nC_r$, Let as study the following :

Example: Find the number of r-combinations of the set $\{a, b, c\}$ where $r = 0, 1, 2, 3$.

Solution:

(*) Exactly one subset contains zero elements ; the null set. Number of 0-combinations $= C(3, 0) = 1$.

(*)Three subsets contain one element each; (a),(b) and (c). Number of 1-combinations $= C(3, 1) = 3$.

(*) Three subsets containing two elements each; $\{a, b\}$, $\{b, c\}$ *and* $\{c, a\}$ number of 2-combinations

$= C(3, 2) = 3$

(*) Finally, exactly one subset contains three elements; the set itself. Number of 3-combinations $= C(3, 3) = 1$.

We now derive a formula for $C(n, r)$

Theorem 4.4.1: (Number of r-combinations without repetitions) The number of r-combinations of n-objects taken r at a time is

$$C(n,r) = \frac{P(n,r)}{r!} = \frac{n!}{r!(n,r)!}; (1 \leq r \leq n).$$

Proof: r-combinations means a selection of r-objects from the n-objects, in which order of the object does not matter, each r-combination contains r-objects and these r-objects can be arranged among themselves in $r!$ hence each r-combination gives rise to $r!$ permutations.

Therefore $C(n,r)$ combinations will gives rise to $C(n,r).r!$ permutation. But the number of n-objects is $P(n,r)$.

$$C(n,r).(r)! = P(n,r) = \frac{(n)!}{(n-r)!(r)!} = \frac{(n)!}{(r)!(n-r)!}$$

Theorem 4.4.2: If n and r are nonnegative integers such that $r \leq n$, to prove that $C(n,r) + C(n,r-1) = C(n+1,r)$.

Proof:

$$\begin{aligned}
LHS &= C(n,r) + C(n,r-1) \\
&= \frac{n!}{r!(n-r)!} + \frac{n!}{(r-1)!(n-r+1)!} \\
&= \frac{n!}{r(r-1)!(n-r)!} + \frac{n!}{(r-1)!(n-r+1)!(n-r)!} \\
&= \frac{n!}{(r-1)!(n-r)!}[\frac{1}{r} + \frac{1}{n-r+1}] \\
&= \frac{n!}{(r-1)!(n-r)!}[\frac{n+1}{r(n-r+1)}] = \frac{n!(n+1)}{r(r+1)!(n-r)!(n-r+1)(n-r)!} \\
&= \frac{(n+1)!}{r!(n-r+1)!} = C(n+1,r) = RHS
\end{aligned}$$

Theorem 4.4.3: To prove that $nC(n-1,r-1) = (n-r+1)\ C(n,r-1)$ for all $1 \leq rn$.

Proof:

$$\begin{aligned}
LHS &= nC(n-1;r-1) \\
&= \frac{n(n-1)!}{(r-1)!(n-1)-(r-1)!} \\
&= \frac{n(n-1)!}{(r-1)!(n-r)!} \\
&= \frac{n!}{(r-1)!(n-r)!} \\
RHS &= (n-r+1)C(n,r-1) \\
&= (n-r+1)\frac{n!}{(r-1)!(n-r+1)(n-r)!} \\
&= \frac{n!}{(r-1)!(n-r)!} = LHS
\end{aligned}$$

Corollary 4.4.4: If $C(n,x) = C(n,y)$ then either $x = y$ (or) $xy = n$
Proof:

$$\begin{aligned} C(n,x) &= C(n,y) \\ \Rightarrow x &= y \\ \Rightarrow C(n,x) &= C(n,y) \\ \Rightarrow C(n,x) &= C(n,n-y) \\ \Rightarrow x &= n-y \\ \Rightarrow x+y &= n. \end{aligned}$$

Hence either $x = y$ (or) $x + y = n$

$$c(n,r) = \frac{P(n,r)}{r!} = \frac{n!}{r!(n-r)!} \left(P(n,r) = \frac{n!}{n-r!} \right) 1 \leq r \leq n$$

Corollary 4.4.5: $C(n,n) = \frac{n!}{n!(n-1)!} = \frac{n!}{n!0!} = 1$

Corollary 4.4.6: $C(n,0) = \frac{n!}{0!(n-0)!} = \frac{n!}{0!1!} = 1$

Corollary 4.4.7: $C(n,r) = C(n,n-r)$
Proof:

$$\begin{aligned} C(n,n-r) &= \frac{n!}{(n-r)!n-(n-r)!} \\ &= \frac{n!}{(n-r)!(n-n-r)!} \\ &= \frac{n!}{(n-r)!r!} \end{aligned}$$

These are called complementary combinations.
Corollary 4.4.8: If $C(n,a) = C(n,b)$ then either $a = b$ (or) $n = a + b$
Proof: Since $C(n,b)$
Then either $a = b$ (or) $b = n - a$ *i.e.*, $n = a + b$
Examples:
1. If $C(n,7) = C(n,5)$ find $C(n,4)$
Solution: We know that if $C(n,a) = C(n,b)$ then $n = a + b$

$$\begin{aligned} C(n,7) &= C(n,5) \\ n &= 7+5 = 12 \\ C(n,4) &= C(12,4) \\ &= \frac{12!}{8!4!} = 495. \end{aligned}$$

2. Prove that $\sum_{r=1}^{5} c(5,r) = 31$

Solution:

$$\begin{aligned} L.H.S &= \sum_{r=1}^{5} C(5,r) \\ &= C(5,1) + C(5,2) + C(5,3) + C(5,4) + C(5,5) \\ &= C(5,1) + C(5,2) + C(5,2) + C(5,1) + 1 \\ &= 2C(5,1) + 2C(5,2) + 1 \\ &= 2\frac{5!}{4!1!} + 2\frac{5!}{3!2!} + 1 \\ &= 31 \end{aligned}$$

3. How many ways are there to select a committee to develop a discrete mathematics course at a school if the committee is to consist of three faculty members from the mathematics department and four the computer science department, if there are nine faculty members of the mathematics department and 11 of the computer science department?.

Solution: By the product rule, the answer is the product of the number of 3-combinations of a set with nine elements and the number of 4-combinations of a set with 11 elements. By Theorem (4.7.1), the number of ways to select the committee is

$$C(9,3)\ C(11,4) = \frac{9!}{3!6!} \cdot \frac{11!}{4!7!} = 84,330 = 27,721.$$

4.5 Permutations with Repetitions

The permutations and combinations examined so for involved un repeated items. If the items repeat, then computations becomes a bit more complicated. This section explores such permutations and combinations.

4.5.1. Permutations with Repetitions: Consider the word REFERENCE. If we swap the second E with the fourth E in the word. We do not get a new word. How can we compute the number of permutations in such cases?

Example : Find the number of different arrangement of the letters of the word REFERENCE.

Solution: The word REFERENCE contains letters. If they were all distinct, the answer would be $9! = 362,880$. But since duplicate letters exists, the answer is indeed much less.

Let N denote the number of different words. We shall find the value of N in an indirect way.

The word REFERENCE contains two R's and four E's the remaining letters are distinct. Think of the two R's distinct letters. R_1 and R_2 and the four E's as four distinct letters.

E_1 through E_4 The letters R_1 and R_2 can be arranged in 2! ways and the four E's in 4! ways.Therfore, if all the letters were distinct, there would be a total of 2!4! N different words

Thus 2!4! = 9! so

$$N = \frac{2!}{4!} = 7560.$$

Theorem 4.5.1: The number of permutation of n items of which n_1 items are of one type n_2 are of a second type,...,and n_k are of a k^{th} type, is $\frac{n!}{n_1!n_2!...n_k!}$.

Proof: Let N denote the total number of permutation as in Example(1) we shall find the value of N indirect.

Let $A_1, ...A_n$ denote the items of the first type, $B_1, ...B_{n_2}$ items of the second type... and $z_1, ..., z_{n_k}$ items of the k^{th} type. If all items were distinct, the total would be $n!$ permutations.

If items $A_1, ...A_{n_1}$ are distinct they can be arranged in $n_1!$ ways. Items $B_1, ...B_{n_2}$ if distinct arranged in $n_2!$ ways, so on, items $z_1, ..., z_{n_k}$, if distinct, can be arranged in $n_k!$ ways, and thus by the product principle, if all items are distinct, there would be $(n_1!n_2!, ..., n_k\ )\ N$ permutations.

so n!=$(n_1!n_2!, ..., n_k\ )N$. Thus

$$N = \frac{n!}{n_1!n_2...n_k!}$$

This theorem works well in solving the next two problems.

Examples :

1. Find the number of bytes containing exactly three $0's$.

Solution : Number of bytes containing exactly three $o's$

=(number bytes containing three $o's$ and five $1's$)

=(number of permutations of eight symbols of which three are alike($0's$) and five are alike (1's))

$$= \frac{8!}{3!5!} = 56$$

2.(Lattice-Walking) figure shows portion of a city map.

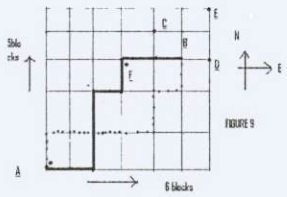


FIGURE 9

Suppose you would like to travel from point A to point B covering exactly 8 blocks, you can travel in the easterly or northerly direction only. Two possible rates are shown in fig(9). How many such rates are possible ?

Solution:

The heavy rate in the fig(9) can be represented by the string $EENNENEE$; it means travel 2 block east, 2 block north, 1 block east,1 bock north and 2 block east, the second route

shown $NEEEENEN$ can be interpreted similarly .
Every route from A and B can be represented by an eight-latter word of which five letters are alike (E's) and three are alike (N's).therefore,
total number of paths from A to B

=(total number of 8-letters words of which five letters are alike and the other three are alike)

$=\frac{8!}{5!3!}$ =56.

4.6 Combinations with Repetitions

Just as permutations can deal with repeated elements. so can combinations (called selections) for example suppose five friends go to a local restaurant for beverages ;iced tea, hot tea or coffee. The waitress puts all five requests on the same order. How many different order for the table are possible? The order in which the beverage need to be selected by more than one person. Also not every beverage need to be selected is immaterial and the same beverages can be selected by more than one person. Also not every beverages need be selected
Before returning to this problem in Example(2), let us study a couple of simple ones.
Example: Five friends would like to order beverages with their dinner at a local restaurant that serves iced tea , hot tea (or) coffee find the number of beverage orders possible.
Solution: A convenient notation will prevent confusion.

Denote each type of beverage by a dash and separate them by using two slashes as shown below

$$\overline{icedtea}/\overline{hottea}/\overline{coffee}$$

mark each person's selected by an X in the appropriate area.
for instance, the distribution $XX/X/XX$ indicates that two people selected iced tea, one selected hot tea and two selected coffee:The distribution $XXX//XX$ means. Three peoples selected iced tea none ordered hot tea and two selected coffee.

Thus the number of possible beverage orders equals the number of permutation of seven items (five X's and two /'s) of which five are alike (X's) and the other two are alike (/'s)

$$\frac{7!}{5!2!} = 21$$

This solution strategy produces the following Theorem.
Theorem 4.6.1: The number of r-combination with repetition from a set of n elements is $c(n+r-1,r)$.
Proof: Each r-combination with repeated elements from a set of n elements can be considered at a string of $rX's$ and $(n-1)$ slashes, as in Example(2) each string contains $r+n-1 = n+r-1$ symbols, of which r are alike $(X's)$ and $n-1$ are alike (slashes). Therefore, by Theorem(4.8.1), the number of such strings, that is r-combination, equals

$$\frac{n+r-1}{r!(n-1)!} = c(n+r-1, r)$$

Example : There are five types of soft drinks at a fast food restaurant: coke classic, diet coke, root beer, pepsi and sprite. Find the number of beverage orders 11 guests can make.
Solution: Since there are five type of soft drinks, $n = 5$, each beverage order is a selection containing 11 items, that is, 11-combination with repeated elements. Therefore by Theorem (4.8.2), the number of possible beverage orders equals

$$\begin{aligned} C(n+r-1, r) &= c(5+11-1, 11) \\ &= c(15, 11) \\ \frac{151}{11!4!} &= 1365. \end{aligned}$$

4.7 Pigeonhole Principle (*Dirichlet's Drawer Principle*)

Attributed to Peter *Gustav Lejeune Dirichlet* (1805 – 59), the pigeonhole Principle observes that if n objects are put in P boxes and $n < p$ then at least one box receives two (or) more objects.

Suppose that a flock of 20 pigeon flies into a set of 19 pigeonholes to roost. Because there are 20 pigeons but only 19 pigeonholes must have at least two pigeons in it. To see why this is true, note that if each pigeonhole at most one pigeon in it, at most 19 pigeons, one per hole, could be accommodated. This illustrates a general principle called the **"Pigeonhole Principle"**, which states that if there are more pigeons than pigeonholes, then there must be least one pigeonhole with at least two pigeons in it.

By G.lejeune *Dirichlet* was born in a *french* family living near cologne, Germany. He studied at the university of *paris* and held positions at the university of *Berslow* and the university of *berlin* in 1855 he was chosen to succeed Gauss at the university of *gottingen Dirichlet* is said to be the first person to master Gauss's *Disquisitiones* Arithmeticae, which appeared 20 years earlier.

Dirichlet made many important discoveries in number theory, including the theorem that there are infinitely many prime in arithmetical progressions $a_n + b$ when a and b are relatively prime.

Theorem 4.7.1: The pigeonhole principle. If K is a positive integer and $K+1$ or more objects are placed into K boxes, then there is at least one box containing two (or) more of the objects.

Proof: We will prove the pigeonhole principle using a proof by contraposition. Suppose that none of the k boxes contains more than one object, then the total number of objects would be at most K, this is a contradiction, because there are at least $K+1$ objects. The pigeonhole principle is also called the *Dirichlet drower* principle, after the nineteenth century German mathematician *Dirichlet*, who often used this principle in his work.

The pigeonhole principle can be used to prove a useful corollary about function.

corollary 4.7.2: A function $'f'$ from a set with $K+1$ or more elements to a set with K elements is not one-to-one.

Proof: Suppose that for each element y in the codomain of $'f'$ we have a box that contains all elements x of the domain of $'f'$ such that $f(x) = y$. Because the domain contain $K+1$ or more elements and the codomain contains only K elements, the pigeonhole principle tells us that one of this boxes contains two ro more elements x of the domain. This means that f cannot be one-to-one.

4.7.1. The following example shows how the pigeonhole principle used:

Examples:

1. Among any group of 367 people, there must be at least two with the same birthday because there are only 366 possible birthdays.

2. In any group of 27 English words, there must be at least two that being with the same letter, since there are 26 letters in the English alphabet.

3. How many students must be in a class to guarantee that at least two students receive the same score of the final exam, if the exam, is graded on a scale from 0 to 100 points?

Solution: There are 101 possible scores in the final. The pigeonhole principle shows that among any 102 students there must be at least 2 students with the same score.

4.7.2. The Generalized Pigeonhole Principle: The pigeonhole principle states that there must be at least two objects in the box when there are more objects than boxes. However, even more can be said when the number of objects exceeds a multiple of the number of boxes, for instance, among any set of 21 decimal digits there must be 3 that are the same . This follows because when 21 objects are distributed into 10 box must have more then 2 objects.

Theorem 4.7.3:The Generalized Pigeonhole Principle. If N objects are placed into K box containing at least $[N/K]$ objects.

Proof: We will use a proof by contradiction. Suppose that none of the boxes contains more than $[N/K]-1$ objects. Then the total number of objects is at most

$$K([N/k]-1) < ((N/K+1)-1) = N$$

where the inequality $[N/K] < (N/K)+1$ has been used. This is a contradiction because there are a total of N objects.

Examples :

1. What is the minimum number of students required in a discrete mathematics class to be sure that are at least six will receive the same grade if there are five possible grades A, B, C, D and F?

Solution: The minimum number of students needed to ensure that at least six students receive the same grade is the smallest integer N such that $[N/5] = 6$. The smallest such integer is $N = 5$. $5 + 1 = 26$ if you have, only 25 students, it is possible for there to be five who have received each grade so that no six students same grade 26 is the minimum number of students needed to ensure that at least six students will receive the same grade.

2. (a) How many cards must be selected deck of 52 cards to guarantee that at least three cards of the same suit are chosen?

(b) How many must be selected to guarantee that at least three heart are selected?

Solution: (a) Suppose there are four boxes one for each suit and as cards are selected they are placed in the box reserved for cards of the suit. Using generalized pigeonhole principle, we see that if N cards are selected, there is at least one box containing at least $[N/4]$ cards,consequently we know that at least three cards of one suit are selected $[N/4] \geq 3$ is $N = 2.4+1=9$ nine cards suffice. Note that if 8 cards are selected, it is possible to have two cards of each suit, so more than 8 cards are needed. Consequently, 9 cards selected to guarantee that at least 3 cards of one suit chosen. One good way to think about this is to note that after 8^{th} card is chosen, there is no way to avoid having a third card of some suit.

(b) We do not use the generalized pigeonhole principle to answer this question, because we want to make sure that there are three hearts, not just three cards of one suit. Note that in the worst case, we can select all the clubs, diamonds, and spades, 39 cards in all, before we select a single heart, the next three cards will be all hearts, so we may need to select 42 cards to get three hearts.

4.7.3. The following theorem is a generalization(extension) of the pigeonhole principle:

Theorem 4.7.4: If n pigeons are assigned to m pigeonholes, then one of the pigeonholes will contain at least $[\frac{n-1}{m} + 1]$ pigeons.

Proof: We prove the theorem by the method of contradiction. Assume that every pigeonhole will contain no more than $\frac{n-1}{m}$ pigeons. Then the total number of pigeons in the m pigeonholes will be at most equal to $m(\frac{n-1}{m}) = n - 1$. This is a contradiction (because the number of pigeons is equal to n). Hence our assumption is wrong. This means that at least one pigeonhole will contain more than $(\frac{n-1}{m})$ pigeons. This is equivalent to saying that(at least) one pigeonhole will contain at least $(\frac{n-1}{m} + 1)$ pigeons. This completes the proof.

Examples :

1. ABC is an equilateral triangle whose sides are of length 3 *cms* each. If we select 5 points inside the triangle. Prove that at least two of these points are such that the distance between them is less than 1cm.

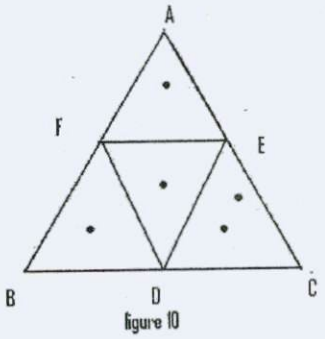


figure 10

Solution: Consider the triangle DEF formed by the midpoints of side BC, CA and AB of the given triangle ABC see above figure. Then the triangle ABC is partitioned into four small equilateral triangles (portions) each of which as sides equal to 1.5 cms. Treating each of these four portions as a pigeonhole and five points chosen inside the triangle as pigeons. We find by using the pigeonhole principle that at least one portion must contain two or more points. Evidently, the distance between such points is less than 1 cm.

2. Suppose that a patient is given a prescription of 45 pills with instructions to take at least one pill per day for 30 days. Prove that there must be a period of consecutive days during which the patient take a total of exactly 14 pills.

Solution: Let a_i be the number of pills. The patient has taken through the end of the i^{th} day, since the patient taken at least one pill per day and at most 45 pills in 30 days, we have $1 \leq a_1 \leq a_2 \leq a_3 ..... \leq a_{30} \leq 45$.

these inequalities give

$a_1 + 14 < a_2 + 14 < a_3 + 14 < .... < a_{30} + 14 \leq 45 + 14 (= 59)$

we have 60 positive integers $a_1, a_2 ... a_{30}, a_1 + 14, a_2 + 14 ... a_{30} + 14$

all of which lie in between 1 and 59 (both inclusive). (that is we have 60 pigeons in 59 pigeonhole) So two of these number must be equal. Since $a_1, a_2, ... a_{30}$ are all different and $a_1 + 14, a_2 + 14 .... a_{30} + 14$ are all different, it must be that one of $a_1, a_2, ... a_{30}$, is equal to one of $a_1 + 14, a_2 + 14 ... a_{30} 14$. This means that there are i and j such that $a_i = a_j + 14$. Thus between days i and j, the patient takes exactly 14 pills.

Theorem 4.7.5: Every sequence of $n^2 + 1$ distinct real numbers contains a subsequence of length $n + 1$ that is either strictly increasing (or) strictly decreasing.

Example: Assume you have a drawer containing a random distribution of a dozen brown socks and a dozen black socks. It is dark, so how many socks do you have to pick to be sure that among them there is a matching point?

Solution: There are two type of socks, so if you pick at least 3 socks, there must be either at least two brown socks or at least black socks.

Generalized pigeonhole principle: $[\frac{3}{2}] = 2$.

4.7.4. Some Elegant Applications of the pigeonhole Principle:

In many interesting application of the pigeonhole principle, the objects to be placed in boxes must be chosen in a clever way. A few such application will be described here.

Example: During a month with 30 days a baseball team plays at least one game a day, but no more than 45 game show that must be a period of some number of consecutive day during which the term must play exactly 14 games.

Solution: Let a_j be the number of games played on or before the j^{th} day of the month. Then $a_1, a_2, ...a_{30}$ is an increasing sequence of distinct positive integers, with $1 \leq a_j \leq 45$. Moreover $a_1+14, a_2+14, ..., a_{30}+14$ is also an increasing sequence of distinct positive integers with $15 \leq a_j + 14 \leq 59$.

The 60 positive integers $a_1, a_2, ...a_{30}, a_1 + 14, a_2 + 14, ...a_{30} + 14$ are all less than or equal to 59. Hence by pigeonhole principle two of these integers are equal, since the integers a_j, j=1,2,...30 are all distinct and the $a_j + 14$, j=1,2,..30 are all distinct, there must be indices i and j with $a_i = a_j+14$. This means that exactly 14 games were played from day $j+1$ to day i.

Check Your Progress

(1) Explain basic counting principle.

(2) Prove that the number of r-permutation of a set of n element is given by $p(n,r) = \frac{n!}{(n-r)!}$

(3) If n and r are nonnegative integers such that $r \leq n$, then prove that $C(n,r)+C(n,r-1) = C(n+1,r)$.

(4) If n pigeons are assigned to m pigeonholes, then one of the pigeonholes will contain at least $[\frac{n-1}{m} + 1]$ pigeons.

4.8 Summary

- Counting principle is very important to determine whether there are telephone numbers or internet protocol addresses to meed demand. Further more counting techniques are used extensively when probabilities of events are computed.
- The basic counting principle are product rule and sum rule.
- Permutation and Combinations are very much useful to count the different objects in many ways.
- Pigeonhole is explained and proved in this section.

4.9 Key Words

Counting principle, Product rule, Sum rule, Permutations, Combinations, Pigeonhole principle.

4.10 Answers Check Your Progress

(1)4.2, (2)Theorem 4.3.3 (3) Theorem 4.4.2 (4)Theorem 4.7.3

4.11 Exercise and Answers

1) Find the number of permutations of the word BANANA.

2) Ramesh has 6 friends. In how many ways can he invite one or more of them at a dinner?

3) If ${}^nC_12 = {}^n C_8$, then find n.

4) Three persons enter a railway compartment. If there are 5 seats vacant, in how many ways can they take these seats?

5) The sides AB, BC, CA of a triangle ABC have 3, 4 and 5 interior points respectively on them. Find out the total number of triangles that can be constructed by using these points as vertices?

6) There are three identical red balls and four identical blue balls in a bag. Three balls are drawn. Find the number of different colour combinations.

7) There are six roads between A and B and four roads between B and C. Find the number of ways that one can drive: (a) from A to C by a way of B; (b) round trip from A to C way of B; (c) round trip from A to C way of B without using the same road more than once.

8) Find the number of ways in which six people can ride a toboggan if one of a subset of three must drive.

9) Find the number of ways in which five large books, four medium size books and three small books can be placed on a shelf so that all books of the small size are together.

10) (a) Find the number of ways in which five persons can sit in a row .

(b)How many ways are there if two of the person insist on sitting next to one another?

(c) Solve part (a) assuming they sit around a circular table.

(d) Solve part (b) assuming they sit around a circular table.

11) (a) In how many ways can three boys and two girls sit in a row?

(b) In how many ways can they sit in a row if the boys and girls are each to sit together?

(c) In how many ways can they sit in a row if just the girls are to sit together?

12) A woman has 11 close friends of whom six are women.

(a) In how many ways can she invite three or more to a party?

(b) In how many ways can she invite three or more of them if she wants the same number of men as women(including herself)?

13) A student is to answer 10 out of 13 questions on an exam.

(a) How many choices has he?

(b) How many if he must answer the first two question?

(c) How many if he must answer the first or second question but not both?

(d) How many if he must answer at least three of the first five questions?

14) In how many ways can 10 students be divided into three, one containing four students and the others three?

15) In how many ways can 14 people be partitioned into six committees where two of the committees contain three members and the others two?

Answers:

1)60, 2)63, 3)20, 4)60, 5)120, 6)35, 7) a. 24, b.576, c.360

8) 360, 9)103680, 10) a.120, b.48, c.24, d.12

11) a.120, b. 24, c.48, 12) a. 1981, b.325

13) a. 286, b.165, c.170, d.276, 14) 2100, 15)3153150

4.12 Suggested Readings

- Seymour Lipschutz & Marc Lars Lipson, *Discrete Mathematics,* Tata McGraw-Hill Publishing Company Limited, 2008.
- Kolman, Barby, Ross *Discrete Mathematical Structures,* 4*th* Edition, Prentice Hall of India, 2003.
- G. Shankar Rao, *Mathematical Foundations of Computer Science,* I.K. International Publishing House Private Limited, 2006.

References:

(1) J.P Tremblay and R. Manohar, *Discrete Mathematical Structure with Applications to Computer Science,* Tata McGraw-Hill, 1997.

(2) Seymour Lipschutz & Marc Lars Lipson, *Discrete Mathematics,* Tata McGraw-Hill Publishing Company Limited, 2008.

(3) Kolman, Barby, Ross *Discrete Mathematical Structures,* 4*th* Edition, Prentice Hall of India, 2003.

(4) G. Shankar Rao, *Mathematical Foundations of Computer Science,* I.K. International Publishing House Private Limited, 2006.

(5) Thomas Koshy, *Discrete Mathematics with Application,* Academic Press An Imprint of Elsevier, 2004.

(6) Joe L. Moff, Abraham Kandel, Theodone P. Baker, *Discrete Mathematics for Computer Scientists and Mathematicians,* 2^{nd} Edition, Prentice Hall of India Private Limited, 2000.

(7) J.K Tress, *Discrete Mathematics for Computer Science,* 2^{nd} Edition, Pearson Education Private Limited, 2000.

(8) C.L. Liu, *Elements of Discrete Mathematics,* 2^{nd} Edition Tata McGraw-Hill Publishing Company Limited, 2000.

(9) Kenneth Rosen, *Discrete Mathematics and Its Application,* 6^{th} Edition, Tata McGraw-Hill Education Private Limited, 2009.

(10) R.C. Penner, *Discrete Mathematics Proof Techniques and Mathematical Structures,* Alied Publication, 2003.

(11) Gary Haggard, John Schlipf, Sue Whites, *Discrete Mathematics for Computer Science,* Thomson Aria Private Limited, 2005.

(12) W.D Wallis, *A Begginner Guide to Discrete Mathematics,* Springer Private Limited, 2004.

M. Sc. (Computer Science)

MSC-501: DISCRETE MATHEMATICS

Module - 2: RELATIONS

UNIT-1: Relation and its Properties

Structure

1.0 Objectives

After studying this unit you will be able to:

- Study properties of relations.
- Study relation matrix and its uses.
- Undestand the basic concepts of graph, Digraph of a relation.
- Solved some examples which are related to communication networks, project scheduling - etc.

1.1 Introduction

Relationships between elements of sets occur in many contexts everyday we deal with relations such as the relations of father to son, brother to sisters etc., Familiar examples in arithmetic are relations such as 'greater than'(¿), 'less than'(¡), or that of equality between two real numbers. Relationships such as that between a computer language and a valid statement in this language often arise in computer science.

Relationships between elements of sets are represented using the structure called a relation, which is just a subset of the Cartesian product of the sets. Relations can be used to solve problems such as determining which pairs of cities are linked by airline flights in a network or producing a useful way to store information in computer data bases.

In this section we introduce the basic terminology used to describe relations and properties of relations.

1.2 Product Sets:

Definition 1.2.1. *Let A and B be two non-empty sets. Then the set of all ordered pairs (a, b), where $a \in A$ and $b \in B$, is called the cartesian product or product set of A and B and is denoted by $A \times B$. Thus $A \times B = \{(a, b) |\ a \in A \text{ and } b \in B\}$*

We note that, in general, the product set $A \times B$ is not the same as the product set $B \times A$, that is $A \times B \neq B \times A$, in general, because $B \times A = \{(b, a)/b \in B,\ a \in A\}$ and $(a, b) \neq (b, a)$ in general.

Example : Let $A = \{1, 2\}$ and $B = \{a, b, c\}$.
Then $A \times B = \{(1, a), (1, b), (1, c), (2, a), (2, b), (2, c)\}$
$B \times A = \{(a, 1), (a, 2), (b, 1), (b, 2), (c, 1), (c, 2)\}$
$A \times A = \{(1, 1), (1, 2), (2, 1), (2, 2)\}$

The idea of a product of sets can be extended to any finite number of sets. For any sets $A_1, A_2, A_3 \ldots, A_n$, then set of all ordered n-tuples $(a_1, a_2, a_3, \ldots, a_n)$, where $a_1 \in A_1, a_2 \in A_2, \ldots, a_n \in A_n$ is called the product of the sets $A_1, \ldots, A_n$ and is denoted by

$$A_1 \times A_2 \times A_3 \times \ldots \times A_n$$

Just as we write A^2 instead of $A \times A$, so we write A^n instead of $A \times A \times \ldots \times A$, where there are n factors all equal to A.

Examples :

1. Find x and y if $(2x, x + y) = (6, 1)$

Solution: We note that $(2x, x + y) = (6, 1)$ if and only if $2x = 6$ and $x + y = 1$. These conditions yield $x = 3$ and $y = 1 - x = -2$.

2. If A, B, C, D are non-empty sets such that $A \subseteq C$ and $B \subseteq D$, prove that $A \times B \subseteq C \times D$.

Solution: Take any $(x, y) \in A \times B$. Then $x \in A$ and $y \in B$. Since $A \subseteq C$ and $B \subseteq D$, it follows that $x \in C$ and $y \in D$. Therefore, $(x, y) \in (c, D)$.
Thus, $(x, y) \in A \times B \Rightarrow (x, y) \in C \times D$.
Hence $(A \times B) \subseteq (C \times D)$.

3. For any non-empty sets A, B and C, prove that $A \times (B \cup C) = (A \times B) \cup (A \times C)$.

Solution: Take any $(x, y) \in A \times (B \cup C)$.

Then (i) $x \in A$ and (ii) $y \in B \cup C$; that is is (i) $x \in A$ and (ii) $y \in B$ or $y \in C$. This means that (i) $x \in A$ and $y \in B$ or (ii) $x \in A$ and $y \in C$. Hence $(x, y) \in (A \times B)$ or $(x, y) \in (A \times C)$ i.e., $(x, y) \in (a \times B) \cup (a \times C)$. This proves that

$$A \times (B \cup C) \subseteq (A \times B) \cup (A \times C) \qquad (1.2.1)$$

Conversely, take any $(x, y) \in (A \times B) \cup (A \times C)$. Then $(x, y) \in A \times B$ or $(x, y) \in A \times C$; i.e., (i) $x \in A$ and $y \in B$, or (ii) $x \in A$ and $y \in C$. This means that (i) $x \in A$ and (ii) $y \in B$ or $y \in C$; i.e., $x \in A$ and $y \in (B \cup C)$. Hence $(x, y) \in A \times (B \cup C)$.

$$(A \times B) \cup (A \times C) \subseteq A \times (B \cup C) \qquad (1.2.2)$$

From the equations (1.2.1) and (1.2.2) , it follows that,

$$A \times (B \cup C) = (A \times B) \cup (A \times C)$$

1.3 Relations:

Relation is a very familiar word, some of the relationships that exist between human beings are that of an uncle, aunt, father, mother etc. Before attempting to give the definition of relation, we will look at some of these familiar relationships from a mathematical viewpoint.

Consider the following sets.

$$X = \{Rama, Raju, Sita, Krishna\}$$
$$Y = \{Govinda, Vishwam, Usha, Mallikarjuna\}$$

Assume that Rama is the father of Vishwam and Govind, Raju is the father of Mallikarjuna, Krishna is the father of usha. This relation 'father of' can be represented by forming a set of ordered pairs.

$R = \{(Rama, Vishwam), (Rama, Govind), (Raju, Mallikarjuna), (Krishna, usha)\}$

The relation 'father' of is given the name R, we note that the ordered pair (x, y) belongs to R if and only if x is the father of y. Since Krishna is not the father of Govind, the ordered pair (Krishna, Govind) doesn't belong to R. Since Sita is the mother of Usha, Vishwam cannot be the mother of Sita. Hence the ordered pair (Sita,Usha) belongs to R but (Sita, Vishwam) doesnot belong to R.

Definition 1.3.1. *Let A and B be sets. A binary relation from A to B is a subset of $A \times B$ Binary relations represent relationships between the elements of two sets..*

Example: Let $A = \{0, 1, 2\}$ and $B = \{a, b\}$. Then, $\{(0, a), (0, b), (1, a), (2, b)\}$ is a relation from A to B.

This means, for instance, that 0 related to a, but 1 is not related to b.

1.3.1 Domain and Range: Let R be relation form a set A to a set B. Then the set of all first elements of the ordered pairs which belong to R is called the *domain* of R, denoted by Dom(R), and the set of all second elements is called the *range* of R, denoted by Ran(R). That is,

$$Dom(R) = \{a \in A | (a,b) \in R \ for \ some \ b \in B\}$$
$$Ran(R) = \{b \in B | (a,b) \in R \ for \ some \ a \in A\}$$

Example : Let $A = \{1,2,3,4,6\}$ and $B = \{1,3,4,5\}$, and a relation R from A to B be defined by aRb if and only if $a < b$. Find Dom(R) and Ran(R). Also, determine R(3), R(6) and $R(\{2,3,4\})$.

Solution: We first note that, here,

$$R = \{(1,3),(1,4),(1,5),(2,3),(2,4),(2,5),(3,4),(3,5),(4,5)\}$$

From this we find that

$$Dom(R) = \{1,2,3,4\},$$
$$Ran(R) = \{3,4,5\},$$
$$R(3) = \{4,5\},\ R(6) = \phi,\ R(\{2,3,4\}) = \{3,4,5\}$$

1.4 Properties of Relations

There are several properties that are used to classify relations on a set.

Definition 1.4.1. *A relation R on a set A is called reflexive if* $(a,b) \in R$ *for every element* $a \in A$

Example : Consider the following relations on $\{1,2,3,4\}$.

R_1= {(1,1),(1,2)(2,1)(2,2),(3,4),(4,1),(4,4)},

$$R_2 = \{(1,1),(1,2)(2,1)\},$$
$$R_3 = \{(1,1),(1,2),(1,4),(2,1),(2,2)(3,3),(4,1),(4,4)\},$$
$$R_4 = \{(2,1),(3,1),(3,2),(4,1),(4,2),(4,3)\},$$
$$R_5 = \{(1,1),(1,2),(1,3),(1,4),(2,2)(2,3),(2,4),(3,3),(3,4),(4,4)\}$$

$R_6 = \{(3,4)\}$.

Which of the these relations are reflexive?

Solution: The relations R_3 and R_5 are reflexive because they contain all pairs of the form (a,a) namely, $(1,1),(2,2),(3,3)$ *and* $(4,4)$. The other relations are not reflexive because they do not contain all of these ordered pairs. In particular R_1, R_2, R_4 and R_6 are not reflexive because $(3,3)$ is not in any of these relations.

Definition 1.4.2 *A relation on a set A is said to be irreflexive if $(a, a) \notin R$ for any $a \in A$. That is, a relation R is irreflexive if no element of A is related to itself by R.*

For example, the relation "is less than" and "is greater than" are irreflexive on the set of real numbers.

Definition 1.4.3 *A relation R on a set A is called symmetric if $(b, a) \in R$ whenever $(a, b) \in R$, for all $a, b \in A$.*

Definition 1.4.4 *A relation R on a set A is called antisymmetric if $(a, b) \in R$ and $(b, a) \in R$, $\forall\ a, b \in A$ then $a = b$.*

A relation is symmetric if and only if a is related to b implies that b is related to a. A relation is antisymmetric if and only if there are no pairs of distinct elements a and b with a related to b and b related to a. A relation cannot be both symmetric and antisymmetric if it contains same pair of the form (a, b) where $a \neq b$.

Example Which of the relations from Example 1 are symmetric and which are antisymmetric?

Solution: The relations R_2 and R_3 are symmetric, because in each case (a, b) belongs to the relation R_2, the only thing to check is that both $(2, 1)$ and $(1, 2)$ are in the relation. For R_3, it is necessary to check that both $(2, 1)$ and $(1, 2)$ belongs to the relation and $(1, 4)$ and $(4, 1)$ belong to the relation.

R_4, R_5 and R_6 are all antisymmetric. For each of these relations there is no pair of elements a and b with $a \neq b$ such that (a, b) and (b, a) belong to the relation.

Definition 1.4.5 *A relation R on a set A is called transitive if whenever $(a, b) \in R$ and $(b, c) \in R$ then $(a, c) \in R$, $\forall\ a, b, c \in A$.*

Example : For each of these relations on the set $\{1, 2, 3, 4\}$, decide whether it is reflexive, symmetric, antisymmetric, transitive?

a) $\{(2, 2), (2, 3), (2, 4), (3, 2), (3, 3), (3, 4)\}$

b) $\{(1, 1), (1, 2), (2, 1), (2, 2), (3, 3), (4, 4)\}$

c) $\{(2, 4), (4, 2)\}$

d) $\{(1, 2), (2, 3), (3, 4)\}$

e) $\{(1, 1), (2, 2), (3, 3), (4, 4)\}$

f) $\{(1, 3), (1, 4), (2, 3), (2, 4), (3, 1), (3, 4)\}$

Solution: (a) is transitive and reflexive.

(b) is reflexive, symmetric and transitive.

(c) is symmetric.

(d) is antisymmetric

(e) is symmetric and transitive.

1.5 Representing Relations using Matrices

A relation between finite sets can be represented using a zero-one matrix. Suppose that R is a relation from $A = a_1, a_2, a_3, ..., a_m$ to $= b_1, b_2, b_3, ..., b_n$, the relations R can be represented by the matrix $M_R = [M_{ij}]$, where $M_{ij} = \begin{cases} 1 \ if \ (a_i, b_j) \in R \\ 0 \ if \ (a_i, b_j) \notin R \end{cases}$

In other words, the zero-one matrix representing R has a 1 as its (i, j) entry when a_i is related to b_j.

Example: Consider the sets $A = \{0, 1, 2\}$ *and* $B = \{p, q\}$ and the relation R from A t o B defined by, $R = \{(0, p), (1, q), (2, p)\}$ here, $A = \{a_1, a_2, a_3\}$ and $B = \{b_1, b_2\}$ with $a_1 = 0$, $a_2 = 1$, $a_3 = 2$, $b_1 = p$, $b_2 = q$.

We note that

$$M_{11} = (a_1, b_1) = (0, p) = 1, \text{ because } (0, p) \in R.$$
$$M_{12} = (a_1, b_2) = (0, q) = 0, \text{ because } (0, q) \notin R.$$
$$M_{21} = (a_2, b_1) = (1, p) = 0$$
$$M_{22} = (a_2, b_2) = (1, q) = 1$$
$$M_{31} = (a_3, b_1) = (2, p) = 1$$
$$M_{32} = (a_3, b_2) = (2, q) = 0$$

$\therefore$ Here the matrix of the relation R is

$$M_R = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}$$

1.6 Digraph of a Relation

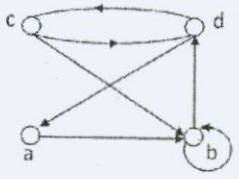
Let R be a relation on a finite set A. Then R can be represented pictorially as described below.

Draw a small circle for each element of A and label the circle with the corresponding element of A. These circles are called vertices. Draw an arrow called an edge, from a vertex x to a vertex y if and only if xRy. The resulting pictorial representation of R is called a directed graph or diagraph of R.

If a relation is pictorially represented by a digraph, the number of edges terminating at a vertex is called the in-degree of that vertex and the number of edges leaving a vertex is called the out degree of that vertex.

Example : Consider the set $A = \{a, b, c, d\}$ and the relation

$R = \{(a,b),(b,b),(b,d),(c,b),(c,d),(d,a),(d,c)\}$ defined on A. The digraph of this relation is as shown below.



Check Your Progress

(1) Explain the properties of relation
(2) Describe the concept of digraph of a relation with example.

1.7 Summary

- Relationships between elements of sets are represented using the structure called a relation, which is just a subset of the cartesian product of the sets.
- A relation R on a set A is called reflexive if $(a,b) \in R$ for any element $a \in A$.
- A relation R on a set A is called symmetric if $(b,a) \in R$ whenever $(a,b) \in R$, $\forall a,\ b \in A$.
- A relation is symmetric if and only if a is related to b implies that b is related to a.
- A relation R on a set A is called transitive if $(a,b) \in R, (b,c) \in R$, then $(a,c) \in R, \forall a,b,c \in A$
- Relation between sets can be represented using matrices.

1.8 Key Words

Relations, refluxive, symmetric, transitive, matrices, diagraph.

1.9 Answers Check Your Progress

(1) 1.4
(2) 1.6

1.10 Exercise and Answers

1) Give an example of a relation which is neither reflexive nor irreflexive.

2) Give an example of a relation which is both symmetric and antisymmetric.

3) Show whether the following relations are transitive:

(a) $R = \{(1,1)\}$

(b) $S = \{(1,2),(2,2)\}$

(c) $T = \{(1,2),(2,3),(1,3),(2,1)\}$

4) Given $S = \{1,2,3,\ldots,10\}$ and a relation R on S where $R = \{(a,b)|\ a+b=10\}$. What are the properties of the relation R?

5) Let $A = \{a,b,c\}$ and $B = \{0,1\}$, and $R = \{(a,0),(b,0),(c,1)\}$ be a relation from A to B. Write the matrix of this relation.

6) Determine the relation R from a set A to a set B as represented by the following matrix:

$$M_R = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

7) Let $A = \{1,2,3,4\}$ and $R = \{(1,1),(1,2),(2,1),(2,2),(3,4),(4,1)\}$ be a relation on A. write the digraph of R.

8) If $R = \{(x,y)|\ x > y\}$ is a relation defined on the set $A = \{1,2,3,4\}$, write the matrix and the digraph of R.

9) Let $A = \{2,4,5,7\}$, and let R be the relation on A having the matrix

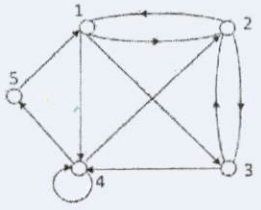
$$M_R = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

Construct the digraph of R and indicate the in-degrees and out-degrees of the vertices.

10) Find the relation R on the set A and write down its digraph, given that $X = \{a,b,c,d,e\}$ and the matrix of R is

$$M_R = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

11) Find the relation R determined by the digraph given below. Also write the matrix of the relation. Determine the in-degrees and out-degrees of the vertices.



1.11 Suggested Readings

- J.P Tremblay and R. Manohar, *Discrete Mathematical Structure with Applications to Computer Science,* Tata McGraw-Hill, 1997.
- Seymour Lipschutz & Marc Lars Lipson, *Discrete Mathematics,* Tata McGraw-Hill Publishing Company Limited, 2008.
- Thomas Koshy, *Discrete Mathematics with Applications, Academic press,* (2004).

UNIT-2: Equivalence Relation

Structure

2.0 Objectives

After studying this unit you will be able to:

- Study relations with a particular combination of properties that allows them to be used to relate objects that are similar in some way.
- Study compatibility of relations.
- Study compatibility of Binary relations.
- Discuss some problems of equivalence relations which are important throughout mathematics and computer science.

2.1 Introduction

In unit-1, we are introducing relations that are reflexive, symmetric and transitive. Naturally we can now ask: Are there relations that simultaneously manifest all three properties? The answer is yes; For instance the relation is logically equivalent to on the set of propositions has all these properties. Such as a relation is an equivalence relation.

For example in some programming languages the names of variables can contain an unlimited number of characters that are checked when a compiler determines whether two variables are equal.

2.2 Equivalence Relation

Definition: 2.2.1 *A subset R of AxA is called an equivalence relation on A if R satisfies the following conditions:*

(i) $(a, a) \in R$ *for all* $a \in A$ *(R is reflexive)*

(ii) If $(a, b) \in R$ *then* $(b, a) \in R$ *, then* $(a, b) \in R$ *(R is symmetric)*

(iii) If $(a, b) \in R$ *and* $(b, c) \in R$*, then* $(a, c) \in R$ *(R is transitive)*

Whenever R is an equivalence relation and x and y are elements of A such that $(x, y) \in R$, we say that x is equivalent to y. The symbol for the equivalence relation R is '$\sim$', then the properties can be restated as follows:

i) $\forall\ a \in A,\ a \sim a$.

ii) $\forall a,\ b \in A$, if $a \sim b$, then $b \sim a$.

iii) $\forall\ a, b, c \in A$, if $a \sim b$ and $b \sim c$,then $a \sim c$.

Examples:

1) The relation has the same color hair as on the set of people is reflexive, symmetric and transitive. So it is an equivalence relation.

2) Equality of numbers on a set of real numbers.

3) Equality of subsets of a universal set.

4) Similarity of triangles on the set of triangles.

5) Relation of lines being parallel on a set of lines in a plane.

6) Relation of living in the same town on the set of persons living in America.

7) Relation of statements being equivalent in the set of statements.

Solved Examples

1. Let T be a set of triangles in a plane, and define R as the set

$$\{R = \{(a, b) | a, b \in T, a,\ is,\ congruent,\ to,\ b\}\}.$$

Then show that R is an equivalence relation.

Solution: If a and b are triangles in a plane then $(a, b) \in R$ if and only if, a is congruent to b (two triangles are congruent if the area of one triangle is equal to the area of the area of the other i.e, area of the triangle 'a'= area of the triangle 'b' or $\triangle a = \triangle$).

We see that:

i) Since $\triangle a = \triangle a\ \forall\ a \in T$, a is congruent to a i.e., $(a, a) \in R$. Hence R is reflexive.

ii) $(a, b) \in R \Rightarrow$ a is congruent to b.

$\Rightarrow \triangle a = \triangle b$

$\Rightarrow \triangle b = \triangle a \Rightarrow b$ is congruent to b.

$\triangle(b, a) \in R$.

Hence R is symmetric.

iii) $(a,b) \in R$, $(b,c) \in R \Rightarrow$ a is congruent to b and b is congruent to c.

$\Rightarrow \triangle a = \triangle b$ and $\triangle b = \triangle$ c.

$\Rightarrow \triangle a = \triangle c$

$\Rightarrow a$ is congruent to c

$\Rightarrow (a,c) \in R$.

Hence R is transitive. Thus R is an equivalence relation.

2. Let S be the set of all points in a plane. Let R be a relation such that for any two points a and b, $(a,b) \in R$ if b is within one inch from a.

Examine if R will be an equivalence relation.

Solution: i) Since the point a overlaps a, i.e., the distance between a and a is zero which is obviously less than one inch. Thus $\forall a \in S$, $(a,a) \in R$. Hence R is reflexive.

ii) $(a,b) \in R \Rightarrow b$ is within one inch from a

$\Rightarrow a$ is within one inch from b

$\Rightarrow (b,a) \in R$.

Hence R is symmetric.

iii) $(a,b) \in R$ and $(b,c) \in R \Rightarrow b$ is within one inch from a and c is within one from b which doesn't implies c is within one inch from a.

Hence R is not transitive.

This proves R is not an equivalence relation.

3. Which of the following relations in the set of real numbers are equivalence relations

a) $R = \{(a,b)|a| = |b|\}$

b) $R = \{(a,b)|a \geq b\}$?

Solution:a) We have to verify the following properties:

i) Since $\forall a \in R$, the set of real numbers,$|a| = |a|$, $(a,a) \in R$. Hence R is reflexive.

ii) $(a,b) \in R \Rightarrow |a| = |b|$

$\Rightarrow |b| = |a|$

$\Rightarrow (b,a) \in R$.

Hence R is symmetric.

iii) $(a,b) \in R, (b,c) \in R \Rightarrow |a| = |b|$ and $|b| = |c|$

$\Rightarrow |a| = |c|$

$\Rightarrow (a,c) \in$ R.

Hence R is transitive.

Thus R is an equivalence relation.) For this relation we have

i) $a \geq a$ $\forall a \in$ R, the set of real numbers, $(a,a) \in R$. Hence R is reflexive.

ii) (a,b) $\in R \Rightarrow a \geq b$ does imply $b \geq a$.

$\Rightarrow (b, a)$ doesn't belongs to R.

Hence R is not symmetric.

iii) $(a, b) \in R$, $(b, c) \in R$

$\Rightarrow a \geq b,\ b \geq c$

$\Rightarrow a \geq c$

$\Rightarrow (a, c) \in R.$

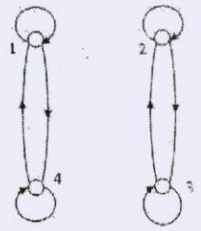
Hence R is transitive.

But R is not equivalence relation.

4. Let $X = \{1, 2, 3, 4\}$ and $R = \{(1, 1), (1, 4), (4, 1), (4, 4), (2, 2), (2, 3), (3, 2), (3, 3)\}$. Write the matrix of R and sketch its graph.

Solution: The matrix and graph of R are given below:

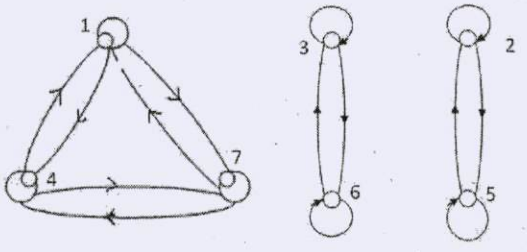
$$M_R = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}$$



It is clear from the above that R is an equivalence relation.

5. Let $X = \{1, 2, 3, ..., 7\}$ and $R = \{(x, y)—x - y \ \textit{is divisible by}\ 3\}$. Show that R is an equivalence relation. Draw the graph of R.

Solution:



From the above figure it is clear that R is an equivalence relation. It is possible to prove this statement without using the graph of the relation in following manner:

i) For any $a \in X$, a - a is divisible by 3; hence aRa or R is reflexive.

ii) For any $a, b \in X$, if a - b is divisible by 3, then b - c is also divisible by 3, that is,$aRb \Rightarrow bRa$. Thus R is symmetric.
iii) For $a, b, c \in X$, if aRb and bRc, then both $a - c = (a - b) + (b - c)$ is also divisible by 3 and hence aRc. Thus R is transitive.

Definition 2.2.2 *Let R be an equivalence relation on a set X. For any $x \in X$, the set $[x]_R \subseteq X$ given by $[x]_R = \{y | y \in X \wedge xRy\}$ is called an R-equivalence class generated by $x \in X$. Accordingly, the set $[x]_R$ consists of all the R - relatives of x in the set X. Sometimes $[x]_R$ is also written as x/R.*

2.2.1 Some properties of the equivalence classes generated by the elements of X.

1) For any element $x \in X$, we have xRx because R is reflexive; therefore $x \in [x]_R$.
2) Let $y \in X$ be any other element such that xRy, so that $y \in [x]_R$. Because of the symmetry of R, yRx and $x \in [y]_R$. Now if there is an element $Z \in [y]_R$, then Z must be in $[x]_R$ because yRz, along with xRy, implies xRz. Thus $[y]_R \subseteq [x]_R$. By symmetry we must also have $[x]_R \subseteq [y]_R$. Finally, from $[y]_R \subseteq [x]_R$ and $[x]_R \subseteq [y]_R$, we have $[x]_R = [y]_R$.
3) In step 2 it is shown that if xRy, then $[x]_R = [y]_R$. We now show that if x is not related to y, then $[x]_R$ and $[y]_R$ must be disjoint. This demonstration can be done by assuming that there is at least one element $z \in [x]_R$ and also $z \in [y]_R$; that is xRz and yRz, but this would imply zRy, and then from transitivity, xRy, which is contradiction.

The above result shows that the R-equivalence class generated by any element $y \in X$ is equal to the R-equivalence class generated by $x \in X$ provided that $y \in [x]_R$. Otherwise the R-equivalence classes generated by x and y are disjoint. Further, each element of X generates an R-equivalence class which is non-empty. Therefore the R-equivalence classes generated by the elements of X cover X, that is, their union is the set X. Since the R-equivalence classes generated by any two elements are either equal or disjoint, we can say that the family of R-equivalence classes generated by the elements of X defines a partition of X. Such a partition is unique because an R-equivalence class of any element of X is unique.

Theorem 2.2.1 *Every equivalence relation on a set generates a unique partition of the set. The blocks of this partition correspond to the R-equivalence classes.*

As we have denoted the R-equivalence class generated by an element $x \in X$ by $[x]_R$ or x/R, we shall denote the family of equivalence classes by X/R, which is also written as X modulo R, or in short as X mod R. X/R is called the quotient set of X by R. Note that the elements of X/R are the equivalence classes which are themselves sets. They are, in fact, subsets of X or elements of the power set ρ(x).

Now we consider two special equivalence relations on a set X. The first such relation is $R_1 = X \times X$ and every element of X is in R_1-relation to all the elements of X. In this case the quotient set of X by R_1 is the set $\{X\}$. The other relation R_2 is such that every element of X is related to itself and to no other element. Such a relation is called an identity relation. An identity relation is an equivalence relation, and the quotient set of X by R_2 consists of sets in which each contain a single element. Thus R_2 generates the largest partition of X.

Examples

1. Let Z be the set of integers and let R be the relation called Congruence modulo 3 defined by

$R = \{(x, y) | x \in z \wedge y \in z \wedge (x - y)$ *is divisible by* $3\}$. Determine the equivalence classes generated by the elements of z.

Solution: The equivalence classes are

$$[0]_R = \{....., -6, -3, 0, 3, 6, ...\}$$
$$[1]_R = \{....., -5, -2, 1, 4, 7, ...\}$$
$$[2]_R = \{....., -4, -1, 2, 5, 8, ...\}$$
$$Z/R = \{[0]_R, [1]_R, [2]_R\}$$

2. Let S be the set of all statement functions in n variables and let R be the relation given by $R = \{(x, y) x \in X \wedge y \in S \wedge x \Longleftrightarrow y\}$.

Discuss the equivalence classes generated by the elements of S.

Solution: The number of possible distinct truth tables for statement functions which depend upon n statement variables is 2^{2^n}. Thus there are 2^{2^n} R-equivalence classes generated by the elements of X.

3. Let $X = \{a, b, c, d, e\}$ and let $C = \{\{a, b\}, \{c\}, \{d, e\}\}$. Show that the partition C defines an equivalence relation on X.

Solution: $R = \{(a, b), (b, b), (a, b), (b, a), (c, c), (d, d), (e, e), (d, e), (e, d)\}$.

2.3 Compatibility Relations

Definition 2.3.1 *A relation R in A is said to be a compatibility relation if it is reflexive and symmetric.*

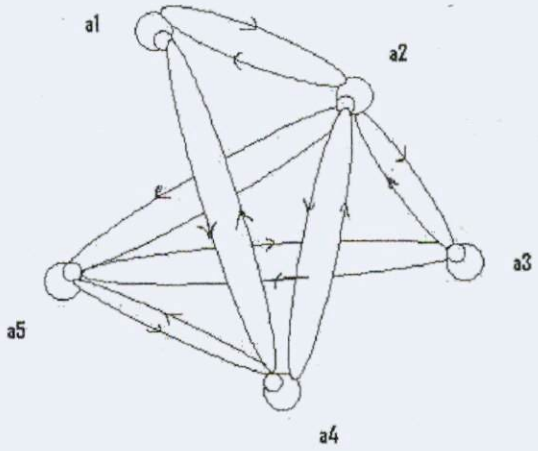
Obviously all equivalence relations are compatibility relations. We shall, however be concerned with those compatibility relations which are not equivalence relations. The following is an example of a compatibility relation.

Let A = {ball, bed, dog, let, egg} and let the relation R be given by

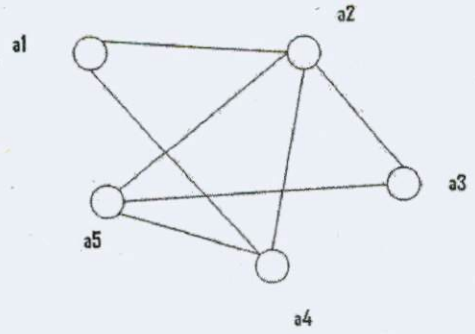
$R = \{(a, b)/a, b \in A \wedge aRb$ if a and b contain some common letter$\}$.

Then R is a compatibility relation and a,b are called compatible if $a \sim b$. The compatibility

relation is sometimes denoted by $\approx$. Note that ball $\approx$ bed, bed $\approx$ egg, but ball not approximately equal to egg. Thus $\approx$ is not transitive.
Denoting ball by a_1, bed by a_2 ,dog by a_3, let by a_4, and egg by a_5, the graph of the compatibility relation is given in the figure.



Since $\approx$ is a compatibility relation, it is not necessary to draw the loops at each element nor is it necessary to draw both $a \sim b$ and $b \sim a$. Thus we can simplify the graph as



Although an equivalence relation on a set defines a partition of the set into equivalence classes, a compatible relation does not necessarily define a partition. However, a compatibility relation does define a covering of the set.

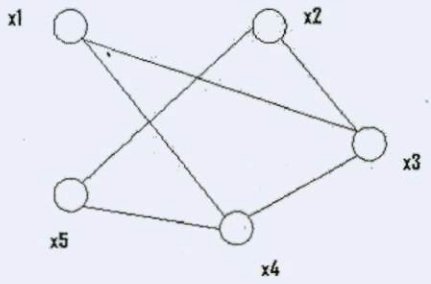
Definition 2.3.2 *Let X be set and $\approx$ a compatibility relation on X. A subset $A \subseteq X$ is called a maximal compatibility block if any element of A is compatibility to every other element of A and no element of X - A (i.e, relative complement of A in X) is compatible to all the elements of A.*

Examples:

1. Let the compatibility relation on a set $\{x_1, x_2, ...x_4\}$ be given by the matrix

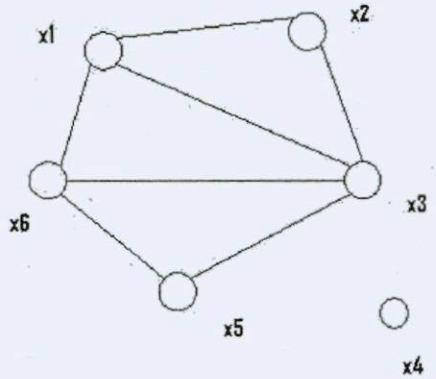
X2	0			
X3	1	1		
X4	1	0	1	
X5	0	1	0	1
	X1	X2	X3	X4

Draw the graph and find the maximal compatibility blocks of the relation.



$\{1,3,4\}$, $\{2,3\}$, $\{4,5\}$, $\{2,5\}$.

X2	1			
X3	1	1		
X5	0	0	1	
X6	1	0	1	1
	X1	X2	X3	X5



$\{1,2,3\}$, $\{1,3,6\}$, $\{3,5,6\}$, $\{4\}$.

2.4 Composition of Binary Relations:

Definition 2.4.1 *Let R be a relation from A to B and S be a relation from B to C. Then a relation written as $R \circ S$ is called a Composite relation of R and S where*
$R \circ S = \{(a, c) | a \in A \wedge c \in C \wedge (\exists b) b \in B \wedge (a, b) \in R \wedge (b, c) \in S\}$
The operation of obtaining $R \circ S$ from R and S is called composition of relations.
Note: $R \circ S$ is empty if the intersection of the range of R and the domain of S is empty. $R \circ S$ is non empty if there is at least one ordered pair $(a, b) \in R$ such that the second member in an ordered pair in S. For the relation $R \circ S$, the domain is a subset of X and the range is a subset of Z. In fact, the domain is a subset of the domain of R and its range is a subset of the range of S.

The operation of composition is a binary operation on relations, and it produces a relation from two relations. The same operation can be applied again to produce other relations.
Examples:
1. Let R be a relation from A to B, S a relation from B to C, and P a relation from C to D. Then $R \circ S$ is relation from A to C. We can form $(R \circ S) \circ P$ which is a relation from A to D. Similarly, we can also form $R \circ (S \circ P)$ which again is a relation from A to D.

Let us assume that $(R \circ S) \circ P$ is non empty and let $(a, b) \in R$, $(b, c) \in S$, and $(c, d) \in P$. This assumption means $(a, c) \in R \circ S$and $(a, d) \in (R \circ S) \circ P$. Of course, $(b, d) \in S \circ P$ and $(a, d) \in R \circ (S \circ P)$ which shows that $(R \circ S) \circ P = R \circ (S \circ P)$.
This result can be stated that the operation of composition on relation is associative, so that $(R \circ S) \circ P = R \circ (S \circ P) = R \circ S \circ P$.
2. Let $R = \{(1, 2), (3, 4), (2, 2)\}$ and $S = \{(4, 2), (2, 5), (3, 1), (1, 3)\}$.
Find $R \circ S$, $S \circ R$, $R \circ (S \circ R)$,$(R \circ S) \circ R$, $R \circ R$, $S \circ S$ and $R \circ R \circ R$.
Solution: $R \circ S = \{(1, 5), (3, 2), (2, 5)\}$
$S \circ R = \{(4, 2), (3, 2), (1, 4)\} \neq R \circ S$.
$(R \circ S) \circ R = \{(3, 2)\}$
$R \circ (S \circ R) = \{(3, 2)\} = (R \circ S) \circ R$
$R \circ R = \{(1, 2), (2, 2)\}$
$S \circ S = \{(4, 5), (3, 3), (1, 1)\}$
$R \circ R \circ R = \{(1, 2), (2, 2)\}$.
3. Let R and S be two relations on a set of positive integer I: $R = \{(x, 2x) | x \in I\}$ and $S = \{(x, 7x) | x \in I\}$.

Find $R \circ S, R \circ R, R \circ R \circ R$ *and* $R \circ S \circ R$.
Solution: $R \circ S = \{(x, 14x) | x \in I\} = S \circ R$
$R \circ R = \{(x, 4x) | x \in I\}$

$$R \circ R \circ R = \{(x, 8x) | x \in I\}$$
$$R \circ S \circ R = \{(x, 28x) | x \in I\}.$$

Definition 2.4.2 *Given a relation R from X to Y, a relation $\tilde{R}$ from Y to X is called the converse of R, where the ordered pairs of $\tilde{R}$ are obtained by interchanging the members in each of the ordered pairs of R. This means, for $x \in X$ and $y \in Y$, that $xRy \iff y\,\tilde{R}\,x$.*

From the definition of $\tilde{R}$ it follows that $\tilde{\tilde{R}}$ = R. The relation matrix $M_{\tilde{R}}$ of $\tilde{R}$ can be obtained by simply interchanging the rows and columns of M_R.

Therefore, $M_{\tilde{R}}$ = transpose of M_R.

The graph of $\tilde{R}$ is also obtained from that of R by reversing the arrows of each arc.

Now consider the converse of a composite relation. For this purpose, let R be a relation from X to Y and S be a relation from Y to Z. Obviously, $\tilde{R}$ is a relation from Y to X, $\tilde{S}$ from Z to Y; R o S is a relation from X to Z, and R $\tilde{o}$ S is a relation from Z to X. Also the relation $\tilde{S}$ o $\tilde{R}$ is from Z to X.

Now we show that R $\tilde{o}$ S = $\tilde{S}$ o $\tilde{R}$

I xRy and ySz, then x(R o S) z and z (R $\tilde{o}$ S) x. But z $\tilde{S}$y and y $\tilde{R}$ x, so that z($\tilde{S}$ o $\tilde{R}$) x. This is true for any x $\in$ X and z $\in$ Z; Hence the required result.

Note: The same rule can be expressed in terms of the relation matrices by saying that the transpose of M_{RoS} is the same as the matrix $M_{\tilde{S}o\tilde{R}}$. The matrix $M_{\tilde{S}o\tilde{S}}$ can be obtained from the matrices $M_{\tilde{S}}$ and $M_{\tilde{S}}$, which in turn can be obtained from the matrices M_S and M_R.

Example: 1. Given the relation matrices M_R and M_S, find M_{RoS}, $M_{\tilde{R}}$, $M_{\tilde{S}}$, $M_{R\tilde{o}S}$ and show that $M_{R\tilde{o}S} = M_{\tilde{S}o\tilde{R}}$.

$$M_R = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \quad M_S = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \end{pmatrix}$$

Solution:

$$M_{\tilde{R}} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} = \text{transpose of } M_R$$

$$M_{\tilde{S}} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \text{transpose of } M_S$$

$$M_{RoS} = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \quad M_{R\tilde{o}S} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

$$M_{\tilde{S}o\tilde{R}} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} = M_{R\tilde{o}S}$$

Check Your Progress

(1) Define equivalence relation with example.
(2) Explain compatibility of a relation.
(3) Explain composition of binary relation.

2.5 Summary

- In the unit-1 we have study the properties of relations, i.e reflexive, symmetric and transitive properties.
- Equivalence relation is the combination above properties,
- A relation R is said to be equivalence relation if it satisfies the reflexive,symmetric and transitive conditions simultaneously.
- Every equivalence relation on a set generates a unique partition of the set. The blocks of this partitions corresponds to the R-equivalence class.
- A relation R in A is said to be a compatibility relation if i t is reflexive and symmetric.

2.6 Key Words

Equivalence relation, Partition, Equivalence class, R-equivalence class, Compatibility, Binary relation.

2.7 Answers Check Your Progress

(1) 2.2
(2) 2.3

(3) 2.4

2.8 Exercise and Answers

1) Determine whether each of the following relations R is an equivalence relation.

a) Let {(x, y)|x - y = 2n, x, y and n ∈ Z}

b) Let R = {(x, y) | x - y = 5n, x, y and n ∈ Z}

c) Let S be the set of books in a specific college library. If x, y ∈ S, then xRy if the books x and y have the same number of pages in them.

d) Let x, y ∈ Q and let xRy if and only if x- y > 0.

2) Show that the relation of congruence modulo m a ≡ (mod m) in the set Z is an equivalence relation. That is the relation R = {(a, b)/ a - b = k m for some fixed integer m and a, b, k ∈ Z} is an equivalence relation.

3) Show that the relation of congruence modulo m has m distinct equivalence classes.

4) Show that equality of numbers is an equivalence relation for the set of real numbers.

5) Let $R = \{(a,a),(a,b),(b,c),(c,c)\}$ and $S = \{(a,a),(b,b),(b,c),(c,a)\}$ on $\{a,b,c\}$,. Then find $R \circ S$, $S \circ R$, R^2 and S^3

6) Let $M_R = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$ $M_S = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix}$

Find (i) $M_{R \circ S}$ (ii) $M_{S \circ R}$ (iii) M_{R^4}.

7) Let the compatibility relation on a set $\{X1, X2, X3, X4, X5, X6\}$ be given by the following matrix

X2	1				
X3	1	1			
X4	0	0	1		
X5	0	0	1	1	
X6	1	0	1	0	1
	X1	X2	X3	X4	X5

Draw the graph and find the maximal compatibility blocks of the relation.

Answers:

5. $R \circ S = \{(a,a),(a,b),(b,b),(b,c),(c,c),(c,a)\}$, $S \circ R = \{(a,a),(a,b),(b,b),(b,c),(c,c)\}$
$R^2 = \{(a,a),(a,b),(b,a),(b,c),(c,b),(c,c)\}$, $S^3 = \{(a,a)\}$

6. i) $M_{R \circ S} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix}$, ii), $M_{S \circ R} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, iii) $M_{R^4} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$

7. $\{1,2,3,4\}$, $\{2,5\}$, $\{3,6\}$, $\{5,6\}$

2.9 Suggested Readings

- J.P Tremblay and R. Manohar, *Discrete Mathematical Structure with Applications to Computer Science,* Tata McGraw-Hill, 1997.
- Seymour Lipschutz & Marc Lars Lipson, *Discrete Mathematics,* Tata McGraw-Hill Publishing Company Limited, 2008.
- Thomas Koshy, *Discrete Mathematics with Applications, Academic press,* (2004).

UNIT-3: Manipulation of Relations

Structure

3.0 Objectives

After studying this unit you will be able to:

- Study the various operations such as union, intersection, complement and inverse on relations.
- Study important properties of partitions.
- Study the applications of covering on relations.
- Solve some problems & examples based on above topics.

3.1 Introduction

We are very familiar with the concept of relations, operations of relations, properties of relations and equivalence class of relations. In this unit, we also consider the various manipulation, such as union, intersection, compliment and inverse that can be performed on relations. Also we proved the fundamental ideas of partition and covering.

3.2 Manipulation of Relations:

Union of relation: Consider two sets A and B. Let R and S be two relations from A to B. Then R and S are subsets of $A \times B$. Hence we can form their union R $\cup$ S. It is the set

of all those elements of A x B which are either in R or in S. Since R ∪ S is a subset of A x B; it is also a relation from A to B. That is *If $a \in A$ and $b \in B$, then a $(R \cup S)$ b if and only if either aRb or aSb.*

Intersection of relation: We can also form the intersection of the two sets R and S. It is the set of all those elements of A x B which belong to both R and S. Since R ∩ S is a subset of A x B, it is also a relation from A to B. That is *If $a \in A$ and $b \in B$, then a $(R \cap S)$ b if and only if aRb and aSb.*

Comliment of relation: Since R is a subset of A x B, we can form its complement R^c. It is the set of all those elements of A x B which does not belong to R. Since R^c is a subset of A x B, it is also a relation from A to B. That is *If $a \in A$ and $b \in B$, then a R^c b if and only if a is not related to b through R.*

Inverse of relation: We define a relation R^{-1} from B to A as *If $b \in B$ and $a \in A$, then b R^{-1} a if and only if aRb.*
In other words, (b, a) ∈ R^{-1} if and only if (a, b) ∈ R.

Example: Let $A = \{1, 2, 3, 4\}$, $R = \{(1,1), (1,3), (2,3), (2,4), (2,2), (3,3), (4,4), (3,1)\}$ and
$S = \{(1,1), (2,3), (2,4), (1,4), (2,2), (3,3), (4,4)\}$.
Then $R \cup S = \{(1,1), (1,3), (1,4), (2,4), (2,3), (2,2), (3,3), (3,1)), (4,4)\}$
$R \cap S = \{(1,1), (2,3), (2,4), (2,2), (3,3), (4,4)\}$
$R^c = \{(1,2), (1,4), (2,1), (3,2), (3,4), (4,1), (4,2), (4,3)\}$
$R^{-1} = \{(1,1), (3,1), (3,2), (4,2), (2,2), (3,3), (4,4), (1,3)\}$.

Theorem 3.2.1 *If R and S are both reflexive relations on a set A, then $R \cup S$ and $R \cap S$ are also reflexive relations on A. If R and S are both symmetric relations on A then $R \cup S$ and $R \cap S$ are also symmetric relations on A.*

Proof: Suppose R and S are both reflexive. We will prove that $R \cup S$ is also reflexive.
Let $a \in A$. Since R is reflexive $(a, a) \in R$ and since S is reflexive $(a, a) \in S$. Hence $(a, a) \in R \cup S$ and this proves that $R \cup S$ is reflexive. Since $(a, a) \in R \cap S$ for all $a \in A$ and hence $R \cap S$ is reflexive.

To prove that $R \cup S$ is symmetric, assume that $(a, b) \in R \cup S$. We have to prove that $(b, a) \in R \cup S$. From $(a, b) \in R \cup S$, we can conclude that either $(a, b) \in R$ or $(a, b) \in S$. If $(a, b) \in R$, then since R is symmetric. It follows that $(b, a) \in R$ and hence $(b, a) \in R \cup S$.
Similarly we can argue that if $(a, b) \in S$.
If $(a, b) \in R \cup S$ then $(a, b) \in R$ and $(a, b) \in S$. Since both R and S are symmetric, $(b, a) \in R$

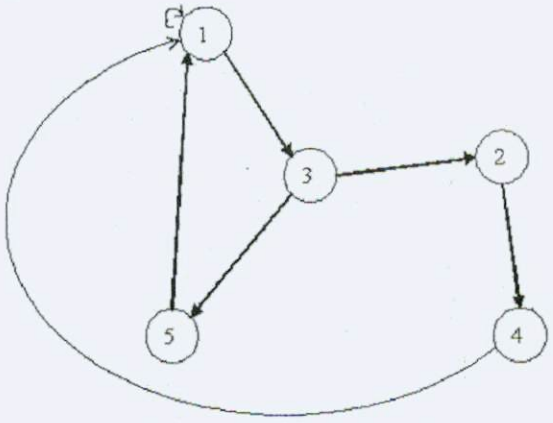
and $(b, a) \in S$ and hence $(b, a) \in R \cap S$.
This proves that $R \cap S$ is symmetric if R and S are both symmetric.

Let R be a relation on a set A and let $a, b \in A$. A path of length n from a to b is a sequence $x_0 = a, x_1, x_2, x_3, ..., x_{n-1}, x_n = b$ such that $aRx_1, x_1Rx_2, ..., x_{n-1}Rb$. Since aRx_1, in the corresponding diagraph of R there is a directed edge from a to x_1. Since x_1Rx_2, there is a directed edge from $x_1 to x_2$. Finally, there exists a directed edge from x_{n-1} to b. Combining all these directed edges, we obtain a directed path from a to b.

Example: If R and S are equivalence relations in the set X, prove that R ∩ S is an equivalence relation.
Solution: We've to verify that R ∩ S is reflexive, symmetric and transitive.
i) $\forall a \in X, (a, a) \in R$ and $(a, a) \in S$, since R and S equivalence relations.
Hence $\forall a \in X$, $(a, a) \in (a, a) \in R \cap S$.
$S \circ R \cap S$ is reflexive.
ii) $(a, b) \in R \cap S \Rightarrow (a, b) \in R$ and $(a, b) \in S$
$\Rightarrow (b, a) \in R$ and $(b, a) \in S$
Since R and S are symmetric being equivalence relations
$\Rightarrow (b, a) \in R \cap S$
Hence $R \cap S$ is symmetric.
iii) $(a, b) \in R \cap S$, $(b, c) \in R \cap S$
$\Rightarrow (a, b) \in R, (b, c) \in R$ and $(a, b) \in S, (b, c) \in S$
$\Rightarrow (a, c) \in R$ and $(a, c) \in S$.
Since R and S are transitive being equivalence relation $\Rightarrow (a, c) \in R \cap S$.
Hence $R \cap S$ is transitive.
This proves $R \cap S$ is an equivalence relation.

Definition 3.2.1 *Let R be a relation on a set A and let $a, b \in A$. We say $aR^n b$ if there exist $x_1, x_2, ..., x_{n-1}$ in A such that $aRx_1, x_1Rx_2, ..., x_{n-2}Rx_{n-1}$ and $x_{n-1}Rb$.*
In particular, when n = 2, we obtain a R^2 b if there exists $x_1 \in$ A such that aRx_1 and x_1Rb i.e., if there exists a directed edge from a to x_1 and a directed edge from $x_1, x_2 \in A$ such that $aRx_1x_1Rx_2$ and x_2Rb.

Example: Let A = {1, 2, 3, 4, 5} and consider a relation whose diagraph is given below.



Some of paths in the above diagraph are $P_1 : 1, 1$ which is a path from 1 to 1 P_2 :1, 3, 2, 4, 1, which is again a path (of length 4) from 1 to 1, and P_3: 3, 5, 1, which is a path of length 2 from 3 to 1. We note that |R|, $|R^4|$ and 3 R^2 1 but 3 R 1 does not exist (no directed edge from 3 to 1).

Paths such as P_1 and P_2 which start from a vertex and end at the same vertex are called cycles.

We say that a R^∞ b if there exists a path of arbitrary length from a to b. We've

R^n (x) = { y $\in$ A | there is a path of length n from x to y}.

R^∞ (x) = {y $\in$ A | there is a path from x to y}.

3.3 Partition and Covering

Definition 3.3.1 *Let S be a non-void set. Let P be a collection $\{A_\alpha\}$ of non-void subsets of S indexed by A. Then P is a partition of S if and only if,*

(a) U_α $A_{\alpha \in A}$ = S (the union of the subsets in P is S)

(b) If $A_\alpha \neq A_\beta$ then $A_\alpha \cap A_\beta = \Phi$ for all $\alpha, \beta \in A$ (the intersection of two distinct subsets in P is empty).

Theorem 3.3.1 *Let S be non-void set and P a partition of S. Let P be indexed by A. Let R be the relation in S x S given by (a, b) $\in$ R if and only if, $a \in A_\alpha$ implies $b \in A_\alpha$, where $A_\alpha \in P$. That is, a, b belong to the same subset in P. Then R is an equivalence relation in S x S.*

Proof: We have to show that R is reflexive, symmetric and transitive. Since every element a is in the same subset A_α of S in P as itself, we have (a, a) $\in$ R for all a $\in$ S. If a and b belong to S also belong to $A_\alpha \in$ P, then b and a belong to $A_\alpha \in$ P. That is, if (a, b) $\in$ R, then (b, a) $\in$ R. If a and b belong to A_α, that is , if (a, b) $\in$ R and (b, c) $\in$ R then (a, c) $\in$ R. Hence the proof.

Theorem 3.3.2 *Let S be non-void set and R an equivalence relation in S x S. Let P be*

a collection of subsets of S called equivalence sets indexed by S and given as follows. Let x $\in$ *S. Then the equivalence set of x,* $A_x = \{y \in S \mid (x, y) \in R\}$. *Then the collection of* $P = \{A_x\}_x \subseteq S$ *is a partition of S.*

Proof: We've to show that P possess properties (a) and (b) of definition of partition. The first of these $U_{x \in S}\ A_x$ = S. Since the union of a collection of subsets of S is contained in S, we've $U_{x \in S}\ A_x \subseteq$ S. Let $a \in S$. Since R is reflexive (a, a) $\in$ R. By the definition of A_α, a $\in$ A_α. Therefore, S $\subseteq U_{\alpha \in S}\ A_x$. Hence we've $S = U_{x \in S}\ A_x$.

Now we shall prove that P satisfies the property (b) of definition of partition.

Let u, v $\in$ S and A_u, $A_v \in$ P, we must show $A_u \cup A_v = \phi$ or A_u, A_v are not different subsets of S. If $A_u \cap A_v = \phi$, then we've finished. If $\exists$ w $\in$ S such that w $\in A_u \cap A_v$, then $w \in A_u$ and $w \in A_v$. By the definition of A_u (u, w) $\in$ R and by the definition of A_v, (v, w) $\in$ R. By symmetric property, if (v, w) $\in$ R, then (w, v) $\in$ R. By transitive property if (u, w) $\in$ R and (w, v) $\in$ R, then (u, v) $\in$ R.

Therefore, u is equivalent to v. It proves that the equivalence set of u $\in$ S is equal to the equivalence set of v $\in$ S that is $A_u = A_v$. This completes the proof of the theorem.

Example: Let Z be the set of integers; $J = \{0, \pm 1, \pm 2, \pm 3, ...\}$. Let n be non-negative integer. Then the relation R_n is the subset of all pairs $(a, b) \in Z \times Z$ such that $a - b = k_n$ for some $k \in Z$, is an equivalence relation.

Solution: We must show that is reflexive, symmetric and transitive. Let $a \in Z$, then $a - a = 0.n$ for $0 \in Z$ that is $(a, a) \in R_n$. It shows that R_n is reflexive.

Now let $(a, b) \in R_n$, then $\exists k \in Z$ such that $a - b = k.n$. Then $b - a = (-k)n$. Here $-k \in Z$. Therefore $(b, a) \in R_n$.This prove is R_n symmetric.

To prove R_n is transitive, let $(a, b) \in R_n$, and $(b, c) \in R_n$, that is $\exists k_1$, $k_2 \in Z$ such that $a - b = k_1 n$ and $b - c = k_2 n$. But $(a - b) + (b - c) = k_1 n + k_2 n$ or $(a - c) = (k_1 + k_2)n$, where $k_1 + k_2 \in Z$. This shows that $(a, c) \in R_n$. Hence R_n satisfies the transitive property and R_n is an equivalence relation.

Definition 3.3.2 *The partition* R_n *induced on the set Z by is* R_n *called the set of integers modulo n.*

Example: Let $n = 3$, for two integers a and b, $(a, b) \in R_3$ if and only if $\exists k \in Z$ such that $a - b = k.3$. The pair $(3, 18) \in R_3$, since $3 - 18 = (-5)3$.
Similarly $(7, -2), (2, 13), (-5, -2)$ belong to R_3.
Because $7 - (-2) = 3.3, -2 - 13 = (-5)3, -5 - (-2) = (-1)3$. There are pairs that do not belong to R_3; $(2, 3), (-17, 11)$ are examples.

Now we form the equivalence sets of elements 0, 1, 2. Therefore

$$[0] = \{y \in Z | 0 \sim y\}$$

$$[0] = \{y \in Z | (0, y) \in R_3\}$$
$$[0] = \{x \in Z | 0 - y = k.3, \text{ for } y \in z\}$$
$$\text{i.e., } [0] = \{0, \pm 3, \pm 6, \pm 9, ...\}$$
$$\text{Similarly, } [1] = \{y \in Z | (1, y) \in R_3\}$$
$$[1] = \{...., -5, -2, 1, 4, 7, ...\} \text{ and}$$
$$[2] = \{....., -4, -1, 2, 5, 8, ...\}$$

There is no need to go further, because $[3] = [0], [4] = [1], [5] = [2], ...$
Observe that if $a, b \in [0]$ then $(a, b) \in R_3$, if e, f $\in$ [1], (e, f) $\in R_3$ and if x, y $\in$ [2] then (x, y) $\in R_3$.

Also notice that [0] $\cup$ [1] $\cup$ [2] = Z. Therefore the collection $Z_3 = \{[0], [2]\}$ is the partition of Z induced by R_3. This set $Z_3 = \{[0], [1], [2]\}$ is called the set of integers modulo 3.

Check Your Progress

(1) If R and S are both reflexive relations on a set A, then R $\cup$ S and R $\cap$ S are also reflexive relations on A. If R and S are both symmetric relations on A then R $\cup$ S and R $\cap$ S are also symmetric relations on A.

(2) Let S be non-void set and R an equivalence relation in S x S. Let P be a collection of subsets of S called equivalence sets indexed by S and given as follows. Let x $\in$ S. Then the equivalence set of x, $A_x = \{y \in S | (x, y) \in R\}$. Then the collection of $P = \{A_x\}_x \subseteq S$ is a partition of S.

(3) Explain partition and covering of a relation.

3.4 Summary

- Different types of operations such as union, intersection, compliment, and inverse were stated.
- The union of two relations R and S denoted by $R \cup S$ on two sets A and B is a subset of $A \times B$, i.e the set of all these elements of $A \times B$ which are either in R or i S.
- The intersection of two relations R and S denoted by $R \bigcap S$ on two sets A and B, is the set of all those elements of $A \times B$ which belongs to both R and S, since $R \bigcup S$ is a subset $A \times B$.
- The compliment of relation R is denoted by R^c and defined as set of all those elements of $A \times B$ which does not belongs to R. since R^c is a subset of $A \times B$.
- The inverse of relation R is denoted by R^-1 from B to A is defined as, if $b \in B$ and $a \in A$, then $bR^{-1}b$ if and only if aRb.

3.5 Key Words

Union of relations, Intersection of relations, Complement of relations, Inverse, partition, Covering.

3.6 Answers Check Your Progress

(1) Theorem 3.2.1
(2) Theorem 3.3.2

3.7 Exercise and Answers

1)Show that a partition of a set S determines an equivalence relation in S.
2) If $\{a, c, e\}$, $\{b, d, f\}$ is a partition of $A = \{a, b, c, d, e, f\}$, determine the corresponding equivalence relation R.
3) Let $R = \{(a, a), (a, b), (b, c), (b, d)\}$ and $S = \{(a, b), (b, b), (b, c), (c, a), (d, a)\}$ on $\{a, b, c, d\}$. Then find $R \cup S$ and $R \cap S$.
4) Let $R = \{(a, 1), (b, 2), (b, 3)\}$ and $S = \{(a, 2), (b, 1), (b, 2)\}$ from $\{a, b\}$ and $\{1, 2, 3\}$, find the following:
(a)$\tilde{R}$ (b) R^{-1} (c) $\tilde{\tilde{R}}$ (d) $(R^{-1})^{-1}$ e) $\tilde{R} \cap \tilde{S}$ (f) $(R \tilde{\cup} S)$ g) $R^{-1} \cup S^{-1}$ h) $(R \cap S)^{-1}$.
5) Using the matrices of relations $R = \{(a, a), (a, b), (b, c)\}$ and $S = \{(a, a), (a, c), (b, b), (b, c), (c, c)\}$, prove that $M_{R \cup S} = M_R \vee M_S$ and $M_{R \cap S} = M_R \wedge M_S$.
6) Let E be the identity relation on a set X and R be any relation in X; Show that $E \cup R \cup \tilde{R}$ is a compatibility relation.
7) Let S = {1, 2, 3, 4, 5} have a partition consisting of the sets {1, 3, 5} and {2, 4}. Show that this partition determines an equivalence relation.
8) What is the ranges of the relations $S = \{(x, x^2) | x \in N\}$ and $T = \{(x, 2x) | x \in N\}$ where $N = \{0, 1, 2, 3, \ldots\}$. Find $S \cup T$ and $S \cap T$.

Answers:
3. $R \cup S = \{(a, a), (a, b), (b, b), (b, c), (b, d), (c, a), (d, a)\}$, $R \cap S = \{(a, a), (a, b), (b, c)\}$
8. $S \cup T = \{(x, x^2, 2x) | x \in N\}$, $S \cap T = \{x | x \in N\}$.

3.8 Suggested Readings

- J.P Tremblay and R. Manohar, *Discrete Mathematical Structure with Applications to Computer Science,* Tata McGraw-Hill, 1997.
- Seymour Lipschutz & Marc Lars Lipson, *Discrete Mathematics,* Tata McGraw-Hill Publishing Company Limited, 2008.
- Thomas Koshy, *Discrete Mathematics with Applications, Academic press,* (2004).

UNIT-4: Warshalls Algorithm

Structure

4.0 Objectives

After studying this unit you will be able to:

- Understand the basic concepts of Transitive closures on relations.
- Study the warshall's algorithm used to find the transitive closures.
- Solve the related problems based on warshalls algorithm.

4.1 Introduction

The connectivity relation of a relation R is closure associated with its transitive closure. A relation R may not have a desired property such as reflexive, symmetry, transitivity. Suppose it is possible to find a relation containing R and having the desired property. The smallest such relation is the closure of R with respect to the property.

Warshall's algorithm, named after Stephen Warshall, who describe this algorithm in 1960, is an efficient method for computing the transitive closure of a relation. In general, algorithm can find the transitive closure of a relation on a set with n elements using $2n^3(n = -1)$ bit operations.

4.2 Transitive Closure

Definition: 4.2.1 *Let R be a relation on a set A. The transitive closure of R is defined to be the smallest transitive relation containing R.*

Theorem 4.2.1 *The transitive closure of R is R^∞.*

Proof: We will prove the following.

(i) The relation R^∞ is transitive.

(ii) R is contained in R^∞.

(iii) If S is any transitive relation such that $R \subseteq S$, then $R^\infty \subseteq S$.

If we prove the above, then it will follow that R^∞ is the smallest transitive relation containing R.

To prove R^∞ is transitive, Assume that $xR^\infty y$ and $yR^\infty z$.

Then there is a path in R from x to y, say

$$x = x_1, x_2, \ldots, x_{m-1}, x_m = y.$$

Also, there is a path in R from y to z, say

$$y = y_1, y_2, \ldots, y_{n-1}, y_n = z.$$

Combining these two paths we obtain

$$x = x_1, x_2, \ldots, x_{m-1}, y, y_2, \ldots, y_{n-1}, y_n = z,$$

which is a path from x to z.

This proves that $xR^\infty z$ and R^∞ is transitive.

To prove $R \subseteq R^\infty$, Assume that $(a, b) \in R$.

Then there is a directed edge from a to b which is also a path from a to b.

Since a path from a to b exists, it follows that $(a, b) \in R^\infty$ proving that

$$R \subseteq R^\infty.$$

To prove (iii), Assume that S is a transitive relation such that $R \subseteq S$.

We first note that $R^\infty \subseteq S^\infty$.

For, if $(a, b) \in R^\infty$, then there is a path in R from a to b.

Since $R \subseteq S$, this path in R from a to b is also a path in S from a to b,

proving that $(a, b) \in S^\infty$.

We will now prove that $S^\infty \subseteq S$, from which it will follow that $R^\infty \subseteq S$.

To prove $S^\infty \subseteq S$, Assume that $(a, b) \in S^\infty$.

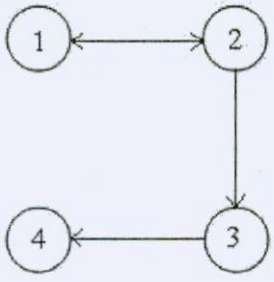
Then there is a path in S from a to b, say

$$a = a_1, a_2, \ldots, a_{k-1}, a_k = b.$$

Then $(a_1, a_2) \in S$, $(a_2, a_3) \in S$, $\ldots$, $(a_{k-1}, b) \in S$.
Since S is transitive, it follows that $(a, b) \in S$.
This proves that $S^\infty \subseteq S$.

Example:(1) Let A={1, 2, 3, 4}, and $R = \{(1,2), (2,3), (3,4), (2,1)\}$. Find the transitive closure of R.

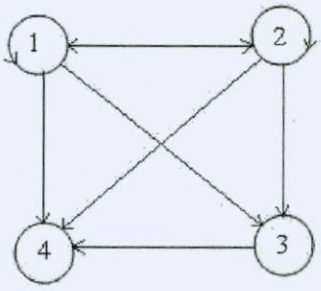
Solution: We observe that $(2,3) \in R$ and $(3,4) \in R$ but $(2,4)$ does not belong to R. Therefore R is not transitive. The digraph of R is as shown below:



By examining the above digraph we note the following:
From the vertex 1, there are paths to 1, 2, 3, 4. Hence, $(1,1) \in R^\infty$, $(1,2) \in R^\infty$, $(1,3) \in R^\infty$, $(1,4) \in R^\infty$.
From the vertex 2, there are paths to 1, 2, 3, 4. Hence, $(2,1) \in R^\infty$, $(2,2) \in R^\infty$, $(2,3) \in R^\infty$, $(2,4) \in R^\infty$.
From the vertex 3, there is a path to 4. Hence $(3,4) \in R^\infty$.
From the vertex 4, there is no path to any vertex. Thus,

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

This is the transitive closure of R.
Its digraph is shown below:



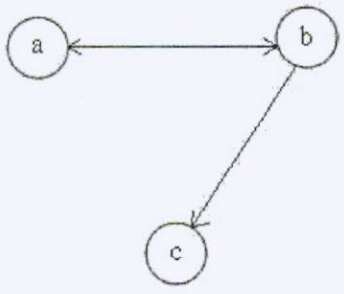
Example:(2) Find the transitive closures of the relations

(i) $R = \{(a,b), (b,a), (b,c)\}$,

(ii) $S = \{(a,a), (b,b), (c,c)\}$,

(iii) $T = \emptyset$, on $\{a, b, c\}$

Solution:(i) We observe that $(a, b) \in R$ and $(b, c) \in R$ but (a, c) does not belong to R. Therefore, R is not transitive. The digraph of R is as shown below:

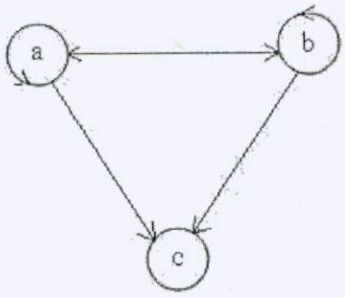


$Fig(A)$

By examining the above digraph we note the following:
From the vertex a, there are paths to a, b, c.
Hence, $(a, a) \in R^{\infty}$, $(a, b) \in R^{\infty}$, $(a, c) \in R^{\infty}$.
From the vertex b, there are paths to a, b, c.
Hence, $(b, a) \in R^{\infty}$, $(b, b) \in R^{\infty}$, $(b, c) \in R^{\infty}$.
From the vertex c, there is no path to any vertex.Thus,

$$R^{\infty} = \{(a, a), (a, b), (a, c), (b, a), (b, b), (b, c)\}.$$

this is the transitive closure of R.
Its digraph is as shown below:



$Fig(B)$

(ii) The relation S is transitive , by default
So transitive closure of S is S.
(iii) The transitive closure of $\emptyset$ is itself.

The transitive closure of the relation R in example (2)(i) has practical applications. Suppose the relation indicates the communication links in a network of computers a, b, c, as in Fig (A). The transitive closure of R shows the possible ways one computer can communicate with another, perhaps through intermediaries. For instance, computer a cannot communicate directly with c, but it can through b.

Fig(B) displays the transitive closure of R.

Theorem 4.2.2 *If R is a relation on a finite set A with $|A| = n$, then*

$$R^{\infty} = R \cup R^2 \cup R^3 \cup \ldots \cup R^n.$$

Proof: Let $a, b \in A$ and suppose that $\{a, x_1, x_2, \ldots, x_m, b\}$ is a path from a to b in R. For any x_i, x_j (with $i \leq j$) in this path, the path can be divided into three sections: First, a path from a to x_i, then a path from x_i to x_j and finally a path from x_j to b.

If $x_i = x_j$, then the middle path is a cycle and so we can leave it out and put the first two paths together. This gives a shorter path from a to b. In this way, we can go on finding shorter and shorter paths from a to b, leaving one x at a time.

Now, let $\{a, x_1, x_2, \ldots, x_k, b\}$ be the shortest path from a to b.

If $a \neq b$, then all of the vertices $a, x_1, x_2, \ldots, x_k, b$ are distinct (-otherwise, we would find a shorter path by selecting vertices that are equal), and, since $|A| = n$, the length of this path is at most $(n-1)$.

Similarly, if a=b, then the length of the path is atmost n.
This means that, if $aR^{\infty}b$ then aR^kb for some k with $1 \leq k \leq n$. Accordingly,

$$R^{\infty} = R \cup R^2 \cup R^3 \cup \ldots \cup R^n,$$

and the proof is complete.

Remark: The result proved in the above theorem shows that powers of R greater than n are not needed to compute R^{∞} for a set with n elements.

Example:(3) Let $A = \{1, 2, 3\}$ and $R = \{(1, 2), (2, 3), (3, 3)\}$. Find the transitive closure of R by computing R^2 and R^3.
Solution: We note that

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

Therefore, the transitive closure of R is

$$R^{\infty} = R \cup R^2 \cup R^3 = \{(1, 2), (1, 3), (2, 3), (3, 3)\}.$$

Example:(4) Let $A = \{1, 2, 3\}$ and $R = \{(1, 1), (1, 2), (2, 3), (1, 3), (3, 1), (3, 2)\}$. Find the transitive closure of R by computing R^2 and R^3.
Solution: We note that

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

Therefore, the transitive closure of R is

$$R^{\infty} = R \cup R^2 \cup R^3 = \{(1,1),(1,2),(1,3),(2,1),(2,2),(2,3),(3,1),(3,2),(3,3)\}.$$

4.3 Warshall's Algorithm

Let A be a set consisting of n elements say

$$A = \{a_1, a_2, \ldots a_n\}$$

and let R be a relation on A. Consider a path $x_1, x_2, \ldots x_m$ in R. Here x_1 and x_m are called end vertices of the path. All other vertices are called interior vertices. Consider any integer k $(1 \leq k \leq n)$. We will define a boolean matrix W_k as follows.

The i^{th} row, j^{th} column entry of W_k is 1 if and only if there is a path from a_i to a_j in R whose interior vertices come only from the set $\{a_1, a_2, \ldots a_k\}$.

We see that the i^{th} row and j^{th} column entry of W_n is 1 if and only if there is a path from a_i to a_j in R whose interior vertices come from the set $A = \{a_1, a_2, \ldots a_n\}$.
Since all interior vertices have to come only from the set $\{a_1, a_2, \ldots a_n\}$, it follows that the i^{th} row, j^{th} column entry of W_n is 1 if and only if there is a path from a_i to a_j in R if and only if the i^{th} row and j^{th} column entry of M_R^{∞} is 1. This means $W_n = M_R^{\infty}$.

We take $W_0 = M_R$ and show how to obtain W_k if W_{k-1} is known. This will enable us to obtain $W_1, W_2, \ldots W_{n-1}, W_n$. Finally, W_n will give the transitive closure of R.

Let $W_{k-1} = (q_{ij})$ and $W_k = (p_{ij})$. Suppose $p_{ij} = 1$. Then by the definition of W_k, there is a path from a_i to a_j whose interior vertices come from the set $\{a_1, a_2, \ldots a_k\}$. Suppose a_k is not an interior vertex of the path. Then this path from a_i to a_j has its interior vertices coming from the set $\{a_1, a_2, \ldots a_{k-1}\}$, proving that $q_{ij} = 1$. Conversely, we can show that if $q_{ij} = 1$, then $p_{ij} = 1$.

Suppose a_k is an interior vertex. Then the path from a_i to a_j passes through a_k and can be split in to two parts: One from a_i to a_k and the other from a_k to a_j. The path from a_i to a_k has all its interior vertices coming from the set $\{a_1, a_2, \ldots a_{k-1}\}$ and the path from a_k to a_j also has all its interior vertices coming from the same set.

$$\text{Hence } q_{ik} = 1 \text{ and } q_{kj} = 1$$

Conversely, if $q_{ik} = 1$ and $q_{kj} = 1$, then $p_{ij} = 1$.

In other words, to obtain the i^{th} row and j^{th} column entry of W_k, we look at the i^{th} row and j^{th} column entry of W_{k-1}. If the latter entry is 1, then 1 is placed in the i^{th} row, j^{th} column entry of W_k. If it is not 1, then we look at the i^{th} row, k^{th} column entry and k^{th} row, j^{th} column entry of W_{k-1}. If both are 1, we place 1 in the i^{th} row, and j^{th} column position

of W_k. All those entries of W_k where there are no $1s$ are made $0s$.

Example:(1) Let $A = \{a, b, c\}$ and $R = \{(a, b), (b, a), (b, c)\}$. Compute the transitive closure of R using Warshall's algorithm.

Solution: Step(1) Find W_0.

$$W_0 = M_R = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

Step(2) Find W_1.

If the ij^{th} element of W_0 is 1, the ij^{th} element of W_1 is also 1. In other words, every 1 in W_0 stays in W_1. To find the remaining $1's$ in W_1, locate the $1's$ in column 1 $(= k)$; there is just one 1; it occurs in position $i = 2$. Now locate the $1's$ in row 1 $(= k)$. Again, there is just one 1, namely in position $j = 2$. Therefore, the ij^{th} entry in W_1 should be 1, where $i = 2$ and $j = 2$.

Thus

$$W_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

Step(3) Find W_2.

Again, all the $1's$ in W_1 stay in W_2.

To find the other $1's$, if any, locate the $1's$ in column 2 $(= k)$ and row 2 $(= k)$. They occur in positions 1 and 2 of column 2 and in positions 1, 2, and 3 of row 2, so the ij^{th} entry of W_2 must be 1, where $i = 1, 2$ and $j = 1, 2, 3$. So change the $0's$ in such locations of W_1 to $1's$. Thus

$$W_2 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

Step(ı) Find W_3.

All the $1's$ in W_2 remain in W_3.

To find the remaining $1's$, if any, locate the $1's$ in column 3- namely, positions 1 and 2- and the $1's$ in row 3. Because no $1's$ appear in row 3, we get no new $1's$. so $W_3 = W_2$.

Since A Contains three elements,

$$W_3 = M_{R^\infty}.$$

Thus

$$W_3=\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

Observe that no new $1's$ are inserted in W_3; that is,

$$W_3 = W_2.$$

From the matrix $W_3 = M_{R^\infty}$, we find that the transitive closure R^∞ of the relation R is

$$R^\infty = \{(a,a),(a,b),(a,c),(b,a),(b,b),(b,c)\}.$$

Example:(2) Let $R = \{(1,2),(2,3),(3,4),(2,1)\}$ be a relation on $A = \{1,2,3,4\}$. find the transitive closure of R by Warshall's method.

Solution: We have

$$W_0 = M_R=\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Since W_0 has a 1 in the locations(positions) $(1,2),(2,1),(2,3)$and$(3,4)$, W_1 too will have a 1 in the corresponding locations.

In the first column of W_0, there is a 1 in the second position; that is in location $(2,1)$. In the first row of W_0 there is a 1 in the second position; that is in location $(1,2)$. Therefore, there will be a 1 in location $(2,2)$ of W_1.

In all other locations of W_1 there will be $0's$. Thus,

$$W_1=\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Since W_1 has a 1 in locations $(1,2),(2,1),(2,2),(2,3)$ and $(3,4)$, W_2 will also have a 1 in the corresponding locations.

In the second column of W_1, there is a 1 in the first and second positions; that is in locations $(1,2)$ and $(2,2)$. In the second row of W_1, there is a 1 in the first, second and third positions; that is in locations $(2,1),(2,2),(2,3)$. Therefore, W_2 will have a 1 in the locations $(1,1),(1,2),(1,3),(2,1),(2,2),(2,3)$.

In all other locations, W_2 has $0's$.

Thus,

$$W_2=\begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Since W_2 has a 1 in locations $(1,1),(1,2),(1,3),(2,1),(2,2),(2,3),(3,4)$, too will have a 1 in the corresponding locations.

In the third column of W_2 ,there is a 1 in $(1,3)$ and $(2,3)$ positions. In the third row of W_2, there is a 1 in the $(3,4)$ position. Therefore, W_3 will have a 1 in $(1,4)$ and $(2,4)$ positions.

In all other positions, W_3 will have $0s$.

Thus,

$$W_3=\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Since W_3 has a 1 in the first two rows and the fourth location in the third row, W_4 too will have 1 in the corresponding locations.

Since there is no 1 in the fourth row of W_3, there will not be any more $1's$ in W_4.

Thus,

$$W_3 = W_4 = M_{R^\infty}$$

Accordingly, the transitive closure of R is

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

Example:(3) Let $A = \{1,2,3\}$ and $R = \{(1,1),(1,2),(2,3),(1,3),(3,1),(3,2)\}$. Compute the transitive closure of R using Warshall's algorithm.

Solution: We have

$$W_0 = M_R=\begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

Applying Warshall's algorithm, we obtain

$$W_1=\begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}; \quad W_2=\begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix};$$

and

$$W_3=\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}=M_R^\infty.$$

Check Your Progress

(1) If R is a relation on a finite set A with $|A| = n$, then

$$R^{\infty} = R \cup R^2 \cup R^3 \cup \ldots \cup R^n.$$

(2) Describe Warshall's algorithm.

(3) Let $A = \{a, b, c\}$ and $R = \{(a, b), (b, a), (b, c)\}$. Compute the transitive closure of R using Warshall's algorithm.

4.4 Summary

- Warshall's algorithm is used to compute the transitive closure. In general we find the transitive closure of a relation on a set with n elements 2^n bit operations.
- The transitive closure of a relation R on a set A is defined to be the smallest transitive relation consisting R.
- Warshall's algorithm is described in the last section.

4.5 Key Words

Transitive closure, Warshall's algorithm, bit operater, Relatons.

4.6 Answers Check Your Progress

(1) Theorem 4.2.2

(2) 4.3

(3) 4.3 example

4.7 Exercise and Answers

1) Let $A = \{a, b, c\}$ and $R = \{(a, a), (a, b), (b, c), (c, c)\}$. Find the transitive closure of R by computing R^2 and R^3.

2) Find the transitive closure of each relation on $A = \{a, b, c\}$.

(i) $\{(a, b), (b, a)\}$

(ii) $\{(a, b), (b, c), (c, a)\}$

(iii) $\{(b, a), (b, c), (c, b)\}$

(iv) $\{(a,a),(a,c),(b,c),(c,a)\}$.

3) Find the transitive closure of each relation on $A=\{a,b,c,d\}$.

(i) $\{(a,a),(a,b)\}$ (ii)$\{(a,b),(b,c),(c,a)\}$

4) Let $A=\{a,b,c,d,e\}$, and $R=\{(a,a),(a,b),(b,c),(c,d),(c,e),(d,e)\}$. Find the transitive closure of R by

(i) computing R^2, R^3, R^4, and R^5,

(ii) using Warshall's algorithm.

5) Using Warshall's algorithm, find the transitive closure of each relation R on $\{a,b,c\}$ in

$$\text{(i)}\begin{pmatrix}0&1&1\\0&0&1\\0&0&0\end{pmatrix};\quad \text{(ii)}\begin{pmatrix}1&0&1\\0&1&0\\0&1&1\end{pmatrix};\quad \text{(iii)}\begin{pmatrix}0&1&0\\0&1&1\\1&0&0\end{pmatrix}.$$

6) Let $A=\{1,2,3\}$ and $R=\{(1,2),(2,3),(3,3)\}$. Find the transitive closure of R by Warshall's method.

7) Using Warshall's algorithm, find the transitive closure of each relation R on $\{a,b,c,d\}$ in

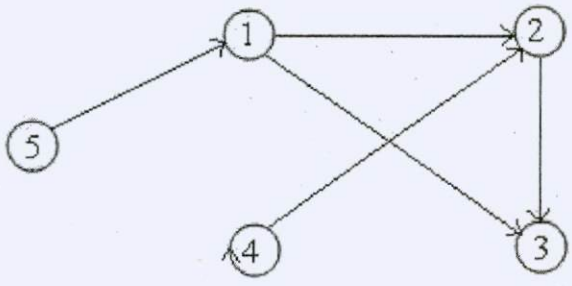
$$\text{(i)}\begin{pmatrix}0&1&0&0\\0&0&1&0\\1&0&1&0\\0&1&0&1\end{pmatrix}\quad \text{(ii)}\begin{pmatrix}1&0&0&0\\0&1&0&1\\0&1&0&1\\0&1&0&1\end{pmatrix}\quad \text{(iii)}\begin{pmatrix}0&1&0&1\\1&0&1&0\\0&0&0&1\\1&0&0&1\end{pmatrix}$$

8) Let R be a relation on $A=\{1,2,3,4\}$ with

$$M_R=\begin{pmatrix}0&1&0&0\\1&0&1&0\\0&0&0&1\\0&0&0&1\end{pmatrix}.$$

Find the transitive closure of R by Warshall's method.

9) On the set $A=\{1,2,3,4,5\}$, the relation R is represented by the following digraph. Obtain the transitive closure of R, by using Warshall's algorithm.



10) Using Warshall's algorithm, find the transitive closure of the relatior

$R=\{(a,a),(a,b),(a,d),(b,a),(c,b),(c,c),(d,b),(d,c),(d,d)\}$ on $\{a,b,c,d\}$.

11) Let $A=\{1,2,3,4\}$. Use Warshall's algorithm to find the transitive closures of the relations

$(i)\{(1,2),(2,1),(2,3),(3,4),(4,1)\}$

$(ii)\{(2,1),(2,3),(3,1),(3,4),(4,1),(4,3)\}$
$(iii)\{(1,2),(1,3),(1,4),(2,3),(2,4),(3,4)\}$
$(iv)\{(1,1),(1,4),(2,1),(2,3),(3,1),(3,2),(3,4),(4,2)\}$.

12) Let $A = \{a,b,c,d,e\}$. Use Warshall's algorithm to find the transitive closures of the relations

$(i)\{(a,c),(b,d),(c,a),(d,b),(e,d)\}$
$(ii)\{(b,c),(b,e),(c,e),(d,a),(e,b),(e,c)\}$
$(iii)\{(a,b),(a,c),(a,e),(b,a),(b,c),(c,a),(c,b),(d,a),(e,d)\}$
$(iv)\{(a,e),(b,a),(b,d),(c,d),(d,a),(d,c),(e,a),(e,b),(e,c),(e,e)\}$.

Answers:

1.$R^{\infty} = \{(a,a),(a,b),(a,c),(b,c),(c,c)\}$, **3.** $\{(a,a),(a,b)\}$

bf 4.$R^{\infty} = \{(a,a),(a,b),(a,c),(a,d),(a,e),(b,c),(b,d),(b,e),(c,d),(c,e),(d,e)\}$.

6. $R^{\infty}\{(1,2),(1,3),(2,3),(3,3)\}$

7. $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix}$, $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix}$

10. $W_4 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$

11. i) $\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$, ii) $\begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{pmatrix}$

iii) $\begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$, iv) $\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$

4.8 Suggested Readings

- J.P Tremblay and R. Manohar, *Discrete Mathematical Structure with Applications to Computer Science,* Tata McGraw-Hill, 1997.
- Seymour Lipschutz & Marc Lars Lipson, *Discrete Mathematics,* Tata McGraw-Hill Publishing Company Limited, 2008.
- Thomas Koshy, *Discrete Mathematics with Applications, Academic press,* (2004).

References:

(1) J.P Tremblay and R. Manohar, *Discrete Mathematical Structure with Applications to Computer Science,* Tata McGraw-Hill, 1997.

(2) Seymour Lipschutz & Marc Lars Lipson, *Discrete Mathematics,* Tata McGraw-Hill Publishing Company Limited, 2008.

(3) Kolman, Barby, Ross *Discrete Mathematical Structures,* 4*th* Edition, Prentice Hall of India, 2003.

(4) G. Shankar Rao, *Mathematical Foundations of Computer Science,* I.K. International Publishing House Private Limited, 2006.

(5) Thomas Koshy, *Discrete Mathematics with Application,* Academic Press An Imprint of Elsevier, 2004.

(6) Joe L. Moff, Abraham Kandel, Theodone P. Baker, *Discrete Mathematics for Computer Scientists and Mathematicians,* 2^{nd} Edition, Prentice Hall of India Private Limited, 2000.

(7) J.K Tress, *Discrete Mathematics for Computer Science,* 2^{nd} Edition, Pearson Education Private Limited, 2000.

(8) C.L. Liu, *Elements of Discrete Mathematics,* 2^{nd} Edition Tata McGraw-Hill Publishing Company Limited, 2000.

(9) Kenneth Rosen, *Discrete Mathematics and Its Application,* 6^{th} Edition, Tata McGraw-Hill Education Private Limited, 2009.

(10) R.C. Penner, *Discrete Mathematics Proof Techniques and Mathematical Structures,* Alied Publication, 2003.

(11) Gary Haggard, John Schlipf, Sue Whites, *Discrete Mathematics for Computer Science,* Thomson Aria Private Limited, 2005.

(12) W.D Wallis, *A Begginner Guide to Discrete Mathematics,* Springer Private Limited. 2004.

Karnataka State Open University

Manasagangothri, Mysore-570006

M. Sc. (Computer Science)

MSC-501: DISCRETE MATHEMATICS

MODULE	UNITS
3	1 to 4

Unit 1: Introduction to Recurrence Relation

Structure

1.0 Objectives

After going through this lesson you will be able to

- Explain the meaning of a recurrence relation;
- Evaluate the recurrence relation by the method of backtracking.
- Evaluate the recurrence relation by the method of characteristic equation.

1.1 Introduction

A recurrence relation is an equation that recursively defines a sequence: each term of the sequence is defined as a function of the preceding terms. The term difference equation sometimes refers to a specific type of recurrence relation. Note however that "difference equation" is frequently used to refer to *any* recurrence relation. Some simply defined recurrence relations can have very complex (chaotic) behaviors, and they are a part of the field of mathematics known as nonlinear analysis. Solving a recurrence relation means obtaining a closed-form solution: a non-recursive function of n.

1.2 Recurrence Relation

Definition: An equation that expresses a_n i.e. general term of the sequence $\{a_n\}$ in terms of one or more of the previous terms of the sequence, namely $a_0, a_1, \ldots, a_{n-1}$ for all integers n with $n \geq n_0$ where n_0 is a non-negative integer is called a recurrence relation for $\{a_n\}$ or a difference relation.

Every recurrence relation must have an initial value.

Examples:

(i) $x_n = 5x_{n-1} + 8x_{n-2}, x_1 = 4.$

(ii) $x_n = 2x_{n-1} + 5, x_1 = -2.$

Definition: A recurrence relation of the form $C_0a_n + C_1a_{n-1} + C_2a_{n-2} + \ldots + C_K a_{n-K} = f(n)$ is called a linear recurrence relation of degree K with constant coefficients where $C_0, C_1, \ldots, C_K$ are real numbers and $C_K \neq 0$.

Note:

1) The recurrence relation is called linear because each a_n is raised to the power 1.

2) The degree of the recurrence relation is the difference between the greatest and least subscripts of the members of the sequence occurring in the recurrence relation.

3) If $f(n) = 0$ the recurrence relation is said to be homogeneous; otherwise it is said to be non-homogeneous.

Examples:

1) A relation, $C_n = -2\ C_{n-1}$ is a linear homogeneous recurrence relation of degree 1.

2) A relation $x_n = 4\ x_{n-1} + 5$ is a linear non-homogeneous recurrence relation because second term in RHS is a constant. It does not contain x_{n-2} factor.

3) A relation $x_n = x_{n-1} + 2\ x_{n-2}$ is a linear homogeneous relation of degree 2.

4) A relation $x_n = x^2_{n-1} + x_{n-2}$ is a non-linear, non-homogeneous relation, because the first term in RHS is a second degree term.

1.3 Method of Solving Recurrence Relations

The important methods to solve a recurrence relation are:

(i) Backtracking method

(ii) Characteristic equation method.

(iii) Generating function method.

1.4 Backtracking method

This is a suitable method for linear non-homogeneous recurrence relation of the type $x_n = rx_{n-1} + s$. This method is explained below using an example.

Example 1.4.1: Using backtrack method solve the recurrence relation

$x_n = x_{n-1} + 5,\ x_1 = 5.$

Solution:

$x_n = x_{n-1} + 5, x_1 = 2.$ .....(1)

Put $n = n-1$ in (1)

$x_{n-1} = x_{n-2} + 5$ .....(2)

Using (2) in (1)

$x_n = (x_{n-2} + 5) + 5$

$x_n = x_{n-2} + 2 \times 5$ .....(3)

Put $n = n-2$ in (1)

$x_{n-2} = x_{n-3} + 5$ .....(4)

Using (4) in (3)

$x_n = (x_{n-3} + 5) + 2 \times 5$

$x_n = x_{n-3} + 3 \times 5$

.....

.....

Repeating for $(n-1)$ times, we get

$x_n = x_1 + 5(n-1), x_1 = 5.$

$x_n = 5 + 5(n-1).$

$x_n = 5n.$

Example 1.4.2: Solve $C_n = C_{n-1} + n$, $C_1 = 5$, by the method of backtracking.

Solution:

Given, $C_n = C_{n-1} + n$, $C_1 = 5$ .....(1)

Put $n = n-1$ in (1)

$C_{n-1} = C_{n-2} + (n-1)$ .....(2)

Using (2) in (1)

$C_n = C_{n-2} + (n-1) + n$ .....(3)

Put $n = n-1$ in (1)

$C_{n-2} = C_{n-3} + (n-2)$ .....(4)

Using (4) in (3),

$C_n = C_{n-3} + [\,(n-2) + (n-1) + n]$

.....

.....

$C_n = C_1 + [\,1 + 2 + 3 + \ldots\ldots + (n-2) + (n-1) + n]$

$C_n = C_1 + \frac{n(n-1)}{2}$

$C_n = 5 + \frac{n(n-1)}{2}$.

Example 1.4.3 Using backtrack method solve $e_n = e_{n-1} - 2$, $e_1 = 2$.

Solution:

$e_n = e_{n-1} - 2,\ e_1 = 2$ .....(1)

Put n=n-1 in (1)

$e_{n-1} = e_{n-2} - 2$ .....(2)

Using (2) in (1)

$e_n = e_{n-2} - 2\text{-}2$

$e_n = e_{n-2} - 2(2)$ .....(3)

Put n=n-2 in (1)

$e_{n-2} = e_{n-3} - 2$ .....(4)

Using (4) in (3),

$e_n = e_{n-3} - 2\text{-}2(2)$

$e_n = e_{n-3} - 2(3)$

.....

.....

$e_n = e_1 - 2(n-1)$

$e_n = 2 - 2(n-1)$

$2 - 2n + 2$

$e_n = 4 - 2n$.

1.5 Characteristic equation method

The method involves the computation of Complementary Function and Particular Integral.

Given the recurrence relation $aU_{n+2} + bU_{n+1} + cU_n = f(n)$ with $U_0 = \alpha$, $U_1 = \beta$ we proceed as follows.

Step 1: Rule to find the complementary function.

Frame the auxiliary (characteristic) equation

$$am^2 + bm + c = 0$$

Depending on the nature of the roots of above, *CF* is got as follows:

Nature of Roots	***CF***
Real distinct λ, μ	$c_1 \lambda^n + c_2 \mu^n$
Real repeated λ, λ	$(c_1 + c_2 n) \lambda^n$
Complex roots $r e^{\pm i\theta}$	$r^n(c_1 \cos n\theta + c_2 \sin n\theta)$

Step 2: Finding *PI* by the method of indeterminant coefficients.

The *PI* is formed for each term or group of terms on RHS. A suitable trial *PI* function U_n is chosen from the table below, substituted in the recurrence relation and the unknowns are found.

The choice of trial PI function depends on the type of term in $f(n)$.

Table of trial function

Type of term	Trial function
a^n	$A a^n$ $An a^n$ *if* a^n is in *CF* $An^2 a^2$ *if* a^n and $n a^n$ are in *CF*
$a^n(\alpha n^2 + \beta n + c)$	$a^n(An^2 + Bn + C)$ Take higher order polynomial if needed
$\alpha n^2 + \beta n + c$	$An^2 + Bn + C$ Take higher order if needed
$\alpha \cos nA + \beta \sin nA$	$A \cos nA + B \sin nA$
$a^n(\alpha \cos nA + \beta \sin nA)$	$a^n(A\cos nA + B \sin nA)$

Step 3 Using the initial conditions C_1, C_2 are evaluated to get the original solution.

Example 1.5.1 Solve $a_n = 4a_{n-1} - 4a_{n-2} + (n + 1)2^n$ by the characteristic equation method.

Solution: By rearranging the given equation we get,

$$a_n - 4a_{n-1} + 4a_{n-2} = (n + 1)2^n$$

The auxiliary equation is

$$m^2 - 4m + 4 = 0$$

$$m = 2, 2$$

$$CF = (c_1 + c_2 n)2^n$$

We assume the particular solution as

$$a_n = n^2(c_1 + c_2 n)2^n$$

$$a_{n-1} = (n-1)^2 (c_1 + c_2(n-1))2^{n-1}$$

$$a_{n-2} = (n-2)^2 (c_1 + c_2(n-2))2^{n-2}$$

$$n^2(c_1 + c_2 n)2^n - 4(n-1)^2(c_1 + c_2(n-1))2^{n-1} + 4(n-2)^2 (c_1 + c_2(n-2))2^{n-2} = (n + 1)2^n$$

$$n^2(c_1 + c_2 n)2^n - 4(n^2-2n+1)(c_1 + c_2 n - c_2) 2^{n-1} + 4(n^2-4n+4)(c_1 + c_2 n - 2c_2) 2^{n-2} = n2^n + 2^n$$

Equating the coefficient of n^2 on both sides

$$n^2(c_1 + c_2 n) - 2(n^2-2n+1)(c_1 + c_2 n - c_2) + (n^2-4n+4)(c_1 + c_2 n - 2c_2) = n + 1$$

$$c_1 - 2c_1 + 2c_2 + 4c_2 + c_1 - 2c_2 - 4c_2 = 0$$

Equating the coefficient of n, $4c_1 - 4c_2 - 2c_2 - 4c_1 + 8c_2 + 4c_2 = 1$

$$6c_2 = 1$$

$$c_2 = 1/6$$

Equating constant terms

$$-2c_1 + 2c_2 + 4c_1 - 8c_2 = 1$$

$$2c_1 - 6c_2 = 1$$

$$2c_1 - 1 = 1$$

$$c_1 = 1$$

$$a_n = (c_1 + c_2 n)2^n + n^2[1 + (n/6)]2^n$$

$$= [1 + (1/6)n]2^n + n^2[1 + (n/6)]2^n$$

Example 1.5.2. Solve $a_{n+2} - 3a_{n+1} + 2a_n = 2^n$; $n \geq 0$ given $a_0 = 3$, $a_1 = 6$.

Solution: The characteristic equation is

$$m^2 - 3m + 2 = 0$$

$$m = 1, 2$$

$$CF = c_1 1^n + c_2 2^n = c_1 + c_2 2^n$$

Assume the trial solution as

$a_n = An2^n$

$a_{n+1} = A(n+1)2^{n+1} = 2(An+A)2^n = 2An2^n + 2A2^n$

$a_{n+2} = A(n+2)2^{n+2} = 2^2(An+2A)2^n = 4An2^n + 3A2^n$

$$4An2^n + 8A2^n - 3(2An2^n + 2A2^n) + 2An2^n = 2^n$$

$$4An + 8A - 6An - 6A + 2An = 1$$

Equating coefficient of n, $4A - 6A + 2A = 0$

Equating constants c, $2A = 1$, $A = ½$

$a_n = c_1 + c_2 2^n + (1/2)n2^n = c_1 + c_2 2^n + n2^{n-1}$

put $n=0$, $a_0 = c_1 + c_2$ $\Rightarrow$ $c_1 + c_2 = 3$

$n=1$, $a_1 = c_1 + 2c_2 + 1$ $\Rightarrow$ $c_1 + 2c_2 = 5$

$c_2 = 2$ and $c_1 = 1$

$\therefore$ $a_n = 1 + 2.2^n + n2^{n-1}$

$a_n = 1 + 2^{n+1} + n2^{n-1}$

Example 1.5.3. Solve $a_{n+1} - 2a_n = 5$; $n \geq 0$ $a_0 = 1$

Solution: The auxiliary equation is

$$m - 2 = 0$$

$$m = 2$$

$$CF = c_1 2^n$$

Assume the trial particular solution as

$a_n = A$; $a_{n+1} = A$

$A - 2A = 5$

$-A = 5$; $A = -5$

$a_n = c_1 2^n - 5$

put $n=0$, $a_0 = c_1 - 5$

$c_1 = 6$

$a_n = 6(2^n) - 5$

Example 1.5.4. Solve $a_n - 2a_{n-1} = 2n^2$; $n \geq 1$ given $a_0 = 4$.

Solution: The auxiliary equation is

$m - 2 = 0$; $m = 2$; $CF = c_1 2^n$

Assume the trial particular solution as

$a_n = An^2 + Bn + C$

$a_{n-1} = A(n-1)^2 + B(n-1) + C$

$An^2 + Bn + C - 2An^2 + 4An - 2A - 2Bn + 2B - 2C = 2n^2$

Equating coefficient of n^2, $A - 2A = 2$

$A = -2$

Equating coefficient of n, $B + 4A - 2B = 0$

$4A - B = 0$

$-8 - B = 0$

$B = -8$

Equating constants, $C - 2A + 2B - 2C = 0$

$-2A + 2B - C = 0$

$C = -2(-2) + 2(-8)$

$C = -12$

$a_n = -2n^2 - 8n - 12$

$\therefore$ $a_n = c_1 2^n - 2n^2 - 8n - 12$

put $n=1$, $a_1 = 2c_1 - 2 - 8 - 12$

$4 = 2c_1 - 22$

$\Rightarrow 2c_1 = 26$; $c_1 = 13$

$a_n = 13(2^n) - 2(n^2 + 4n + 6)$

Example 1.5.5. Solve $a_{n+2} - 7a_{n+1} - 8a_n = n(n-1)2^n = (n^2 - n)2^n$

Solution: The auxiliary equation is

$m^2 - 7m - 8 = 0$

$(m - 8)(m + 1) = 0$

$m = 8, -1$

$CF = c_1 8^n + c_2(-1)^n$

Assume the trial particular solution as

$$a_n = (An^2 + Bn + C)2^n$$

$$a_{n+1} = [A(n^2 + 2n + 1) + B(n + 1) + C]2^{n+1}$$

$$a_{n+2} = [A(n^2 + 4n + 4) + B(n + 2) + C]2^{n+2}$$

$$(An^2 + 4An + 4A + Bn + 2B + C)2^{n+2} - 7(An^2 + 2An + A + Bn + B + C)2^{n+1} - 8(An^2 + Bn + C)2^n = (n^2 - n)2^n$$

$$4An^2 + 16An + 16A + 4Bn + 8B + 4C - 14An^2 - 28An - 14A - 14Bn - 14B - 14C - 8An^2 - 8Bn - 8C) \; C) = n^2 - n$$

Equating coefficients of n^2,

$$4A - 14A - 8A = 1$$

$$A = -(1/18)$$

Equating coefficient of n,

$$16A + 4B - 28A - 14B - 8B = -1$$

$$-12A - 18B = -1$$

$$18B = 1 - 12A$$

$$= 1 - 12[\tfrac{1}{8}(1/18)]$$

$$= 1 + (2/3) = 5/3$$

$$B = 5/54$$

Equating constant,

$$16A + 8B + 4C - 14A - 14B - 14C - 8C = 0$$

$$2A - 6B - 18C = 0$$

$$18C = 2A - 6B$$

$$9C = A - 3B$$

$$= (-1/18) - 3(5/54)$$

$$= -(1/18) - (5/54)$$

$$= -(6/18) = -1/3$$

$$C = -1/54$$

$$\therefore \quad a_n = \left(\frac{-1}{18}n^2 + \frac{5}{54}n - \frac{1}{54}\right)2^n$$

$$a_n = c_1 8^n + c_2(-1)^n - (1/54)(3n^2 - 5n + 1)2^n.$$

1.6 Summary

A recurrence relation is an equation that recursively defines a sequence: each term of the sequence is defined as a function of the preceding terms.

A recurrence relation of the form $C_0a_n + C_1a_{n-1} + C_2a_{n-2} + \ldots + C_K a_{n-K} = f(n)$ is called a linear recurrence relation of degree K with constant coefficients where $C_0, C_1, \ldots, C_K$ are real numbers and $C_K \neq 0$.

The important methods to solve a recurrence relation are:

(i) Backtracking method

(ii) Characteristic equation method.

(iii) Generating function method.

Backtracking is a suitable method for linear non-homogeneous recurrence relation of the type $x_n = rx_{n-1} + s$.

The characteristic equation method involves the computation of complementary function and particular integral to solve linear recurrence relation.

1.7 Keywords

Recurrence relation, sequence, auxiliary equation.

1.8 Supplementary Problems

1.8.1. Solve the following recurrence relations by the method of backtracking.

i) $b_n = 3b_{n-1} + 1, b_1 = 7$

ii) $a_n = a_{n-1} + 2n, a_1 = 5.$

iii) $e_n = e_{n-1} - 2, e_1 = 2.$

iv) $x_n = 2x_{n-1} + 2, x_1 = 4.$

1.8.2. Solve the following recurrence relations by the characteristic equation method.

i) $a_n = 3\ a_{n-1} - 2\ a_{n-2}$, $a_0 = 5$ and $a_1 = 4$.

ii) $a_n = 4a_{n-1} - 4a_{n-2} + 2^n$, $a_0 = 3$, $a_1 = 6$.

iii) $a_{n+2} - 7a_{n+1} - 12\ a_n = n(n-1)\ 2^n$

iv) $a_{n+1} - 8a_n = 5$; $n \geq 0$ $a_0 = -2$.

1.9 References:

1. Discrete mathematics, by P. Geetha, Scitech publications.
2. Discrete mathematics for computer science, by Kolman, Busby and Ross, PHI publications.
3. Discrete Mathematical Structures, by N. G. Goudru, Himalaya Publishing House.

Unit 2: Generating Functions

Structure

Unit-2: Generating Functions

2.0 Objectives

After going through this lesson you will be able to

- Explain the meaning of the generating functions;
- Evaluate homogeneous recurrence relation by the method of generating functions.
- Evaluate non-homogeneous recurrence relation by the method of generating functions.

2.1 Introduction

A generating function is a formal power series in one indeterminate, whose coefficients encode information about a sequence of numbers that is indexed by the natural numbers. Generating functions were first introduced by Abraham de Moivre in 1730, in order to solve the general linear recurrence problem.

Generating functions are often expressed in closed form (rather than as a series), by some expression involving operations defined for formal power series. Indeed, the closed form expression can often be interpreted as a function that can be evaluated at (sufficiently small) concrete values of x, and which has the formal power series as its Taylor series; this explains the designation "generating functions". Generating functions are not functions in the formal sense of a mapping from a domain to a codomain; the name is merely traditional, and they are sometimes more correctly called generating series.

The particular generating function, if any that is most useful in a given context will depend upon the nature of the sequence and the details of the problem being addressed.

2.2 Generating Function

Definition: The generating function of a sequence $a_0, a_1, a_2, \ldots$ is the expression

$$G(x) = a_0 + a_1 x + a_2 x^2 + \ldots$$
$$= \sum_{n=0}^{\infty} a_n x^n.$$

2.3 Solved problems:

2.3.1. Solve the recurrence relation $a_n = 3a_{n-1} + 1$, $n \geq 1$; given that $a_0 = 1$, by the method of generating function.

Solution:

Let the generating function of $\{a_n\}$ be $G(x) = \sum_{n=0}^{\infty} a_n x^n$

Multiply the given RR by x^n and sum n=1 to ∞

$$= \sum_{n=1}^{\infty} a_n x^n = 3 \sum_{n=1}^{\infty} a_{n-1} x^n + \sum_{n=1}^{\infty} x^n$$

$$G(x) - a_0 = 3x \sum_{n=1}^{\infty} a_{n-1} x^{n-1} + x + x^2 + \ldots\ldots \infty$$

$$G(x) - a_0 = 3x\, G(x) + x(1-x)^{-1}$$

$$(1-3x)G(x) = 1 + \frac{x}{1-x} = \frac{1-x+x}{1-x} = \frac{1}{1-x}$$

$$G(x) = \frac{1}{(1-x)(1-3x)} = \frac{-\frac{1}{2}}{1-x} + \frac{\frac{3}{2}}{1-3x}$$

$$= \frac{-1}{2}(1-x)^{-1} + \frac{3}{2}(1-3x)^{-1}$$

$$\sum_{n=0}^{\infty} a_n = \frac{-1}{2} \sum_{n=0}^{\infty} x^n + \frac{3}{2} \sum_{n=0}^{\infty} 3^n x^n$$

$\therefore a_n$ = Coefficient of x^n in $G(x)$

$$= \frac{-1}{2} + \frac{3}{2} 3^n = \frac{1}{2}(3^{n+1} - 1).$$

2.3.2. Use the method of generating function to solve the recurrence relation $a_n = 4a_{n-1} - 4a_{n-2} + 4^n$; $n \geq 2$, given that $a_0 = 2$, $a_1 = 8$.

Solution: Let $G(x) = \sum_{n=0}^{\infty} a_n x^n$

Multiply the Recurrence Relation by x^n and sum n=2 to ∞

$$\sum_{n=2}^{\infty} a_n x^n = 4 \sum_{n=2}^{\infty} a_{n-1} x^n - 4 \sum_{n=2}^{\infty} a_{n-2} x^n + \sum_{n=2}^{\infty} 4^n x^n$$

$$G(x) - a_0 - a_1 x = 4x \sum_{n=2}^{\infty} a_{n-1} x^{n-1} - 4x^2 \sum_{n=2}^{\infty} a_{n-2} x^{n-2} + [(4x)^2 + (4x)^3 + \ldots]$$

$$= 4x\,(G(x) - a_0) - 4x^2 G(x) + (4x)^2(1-4x)^{-1}$$

$$G(x) - 2 - 8x = 4x\,(G(x) - 2) - 4x^2 G(x) + (4x)^2(1-4x)^{-1}$$

$$(4x^2 - 4x - 1)\, G(x) = 2 + 8x - 8x + \frac{16^2}{1-4x}$$

$$= 2+\frac{16^2}{1-4x} = \frac{2-8x+16x^2}{1-4x}$$

$$G(x)= \frac{2-8x+16x^2}{(1-4x)(1-2x)^2}$$

Consider

$$\frac{2-8x+16x^2}{(1-2x)(1-2x)^2} = \frac{A}{1-2x} + \frac{B}{(1-2x)^2} + \frac{C}{1-4x}$$

$$2 - 8x + 16x^2 = A\,(1-2x)\,(1-4x) + B(1-4x) + C(1-2x)^2$$

Put $x = \frac{1}{2}$ $\quad -B = 2-4+4$

$B=-2$

Put $x = \frac{1}{4}$ $\quad \frac{C}{4} = 2-2+1 = 1$

$C=4$

Put $x=0$ $\quad A+B+C=2$

$A=0$

$$G(x) = \frac{2-8x+16x^2}{(1-4x)(1-2x)^2}$$

$$= \frac{4}{1-4x} - \frac{2}{(1-2x)^2}$$

$$= 4(1-4x)^{-1} - 2(1-2x)^{-2}$$

$$= 4(1+4x+(4x)^2+ \ldots. +(4x)^n + \ldots \infty) - 2(1+2(2x)+3(2x)^2+ \ldots. +(n+1)(2x)^n+$$

$$\ldots\ldots. \infty)$$

$$a_n = 4.4^n - 2.2^n(n+1)$$

$$= 4^{n+1}-2^{n+1}(n+1)$$

2.3.3. Use the method of generating function to solve $a_n - a_{n-1} - 2a_{n-2} = 0$ with $a_0 = 5$, $a_1 = -3$.

Solution: Let $G(x) = \sum_{n=0}^{\infty} a_n\, x^n$

Consider, $a_n - a_{n-1} - 2a_{n-2} = 0$

$a_n x^n - a_{n-1}x^n - 2a_{n-2}x^n = 0.$ [By multiplying by x^n]

$$\sum_{n=2}^{\infty} a_n x^n - \sum_{n=2}^{\infty} a_{n-1}x^n - 2\sum_{n=2}^{\infty} a_{n-2}x^n = 0$$ [By taking summation]

$$[G(x) - a_0 - a_1x] - x[G(x) - a_0] - 2x^2G(x) = 0$$

$(1-x-2x^2)\ G(x) - 5 + 3x + 5x = 0$

$(1-x-2x^2)\ G(x) = 5 - 8x$

$$G(x) = \frac{5-8x}{1-x-2x^2}$$

$$= \frac{8x-5}{(2x-1)(x+1)}$$

Consider, $\frac{8x-5}{(2x-1)(x+1)} = \frac{A}{2x-1} + \frac{B}{x+1}$

$$8x - 5 = A(x+1) + B(2x-1)$$

$$A + 2B = 8$$

$$A - B = -5$$

$$3B = 13$$

$$B = 13/3$$

$$A = B - 5 = (13/3) - 5$$

$$A = -2/3$$

$$G(x) = \frac{8x-5}{(2x-1)(x+1)}$$

$$= \frac{-2/3}{(2x-1)} + \frac{13/3}{(x+1)}$$

$$= \frac{2}{3}\,\frac{1}{1-2x} + \frac{13}{3}\,\frac{1}{1+x}$$

$$= \frac{2}{3}(1-2x)^{-1} + \frac{13}{3}(1+x)^{-1}$$

$$= \frac{2}{3}[1 + (2x) + (2x)^2 + \ldots + (2x)^n + \ldots \infty]$$

$$+ \frac{13}{3}[1 - x + x^2 - \ldots + (-1)^n x^n + \ldots \infty]$$

$$\therefore \quad \sum_{n=0}^{\infty} a_n x^n = \frac{2}{3}\sum_{n=0}^{\infty} 2^n x^n + \frac{13}{3}\sum_{n=0}^{\infty}(-1)^n x^n = 0$$

$$\therefore \quad a_n = \frac{2}{3}.\,2^n + \frac{13}{3}(-1)^n.$$

2.3.4. Solve u_{n+1} -3 u_n =7.2^n with $u_0 = 1$

Solution: Let $G = \sum_{n=0}^{\infty} u_n\ z^n$ be the generating function of the sequence u_n.

Given, u_{n+1} -3 u_n =7.2^n .....(1)

Multiply (1) by z^{n+1} and sum from n=0 to ∞

$$\sum_{n=0}^{\infty} u_{n+1}\ z^{n+1} - 3\sum_{n=0}^{\infty} u_n\ z^{n+1} = 7\sum_{n=0}^{\infty} 2^n z^{n+1}$$

$$= 7z \sum_{n=0}^{\infty} 2^n z^n$$

$$G - u_0 - 3zG = 7z\,(1-2z)^{-1} = \frac{7z}{1-2z}$$

$$(1-3z)G = 1 + \frac{7z}{1-2z}$$

$$= \frac{1-2z+7z}{1-2z} = \frac{1+5z}{1-2z}$$

$$G = \frac{1+5z}{(1-2z)(1-3z)} = \frac{A}{1-2z} + \frac{B}{1-3z}$$

$$\frac{1+5z}{(1-2z)(1-3z)} = \frac{A}{1-2z} + \frac{B}{1-3z}$$

$$1 + 5z = A(1-3z) + B(1-2z)$$

Equating coefficient of z, $-3A - 2B = 5$ $\quad -3A - 2B = 5$ $\quad B = 1 - A$

Equating constant c, $A + B = 1$ $\quad 2A + 2B = 2$ $\quad B = 1 + 7$

$-A = 7$ $\quad B = 8$

$A = -7$

$$G = \frac{-7}{1-2z} + \frac{8}{1-3z}$$

$$u_n = -7(2)^n + 8(3)^n$$

check $u_0 = -7 + 8 = 1$

2.3.5. Solve u_{n+2} -5u_{n+1} + 6u_n =n+2 with $u_0 = 0$, u_1=1.

Solution: Let $G = \sum_{n=0}^{\infty} u_n\ z^n$ be the generating function of the sequence u_n.

Given u_{n+2} -5u_{n+1} + 6u_n =n+2 .....(1)

Multiply (1) by z^{n+2} and sum from n=0 to ∞

$$\sum_{n=0}^{\infty} u_{n+2}\ z^{n+2} - 5\sum_{n=0}^{\infty} u_{n+1}\ z^{n+2} + 6\sum_{n=0}^{\infty} u_n\ z^{n+2} = \sum_{n=0}^{\infty}(n+2)\ z^{n+2}$$

$$G-(u_0+u_1z)-5z(G-u_0)+6z^2G=z^2\sum_{n=0}^{\infty}n\,z^n+2z^2\sum_{n=0}^{\infty}\quad z^n$$

$$G-z-5zG+6z^2G=z^2[z+2z^2+3z^3+\ldots.]+2z^2[\,1+z+z^2+\ldots\ldots\,]$$

$$(1-5z+6z^2)G-z=z^3(1-z)^{-2}+2z^2(1-z)^{-1}$$

$$=\frac{z^3}{(1-z)^2}+\frac{2z^2}{1-z}$$

$$(1-5z+6z^2)G=z+\frac{2z^2}{1-z}+\frac{z^3}{(1-z)^2}$$

$$=\frac{z(1-z)^2+2z^2(1-z)+z^3}{(1-z)^2}$$

$$=\frac{z(1+z^2-2z)+2z^2-2z^3+z^3}{(1-z)^2}$$

$$=\frac{z+z^3-2z^2+2z^2-2z^3+z^3}{(1-z)^2}=\frac{z}{(1-z)^2}$$

$$[6z^2-5z+1=0$$

$$z=\frac{5\pm\sqrt{25-24}}{12}$$

$$=\frac{5\pm1}{12}=\frac{1}{2},\frac{1}{3}$$

$$(2z-1)\,(3z-1)=0]$$

$$G=\frac{z}{(1-z)^2(1-5z+6z^2)}$$

$$=\frac{z}{(1-z)^2(2z-1)(3z-1)}$$

$$\frac{z}{(1-z)^2(2z-1)(3z-1)}=\frac{A}{1-z}+\frac{B}{(1-z)^2}+\frac{C}{2z-1}+\frac{D}{3z-1}$$

$$z=A(1-z)\,(2z-1)\,(3z-1)+Bz(2z-1)(3z-1)+C(1-z)^2(3z-1)+D(1-z)^2(2z\text{-}1)$$

Put z=1, $\quad 1=2B \Rightarrow B=½$

Put z=0, $\quad 0=A-C-D$

$$A-D=4$$

$$A+(9/4)=4 \quad \Rightarrow A=7/4$$

Put z=1/2 $\quad (1/2)=C.\,(1/8)$

$$C=4$$

Put $z = 1/3$ $\quad (1/3) = (-4/27)D$

$D = (-27/4)\,(1/3) = -9/4$

$$G = \frac{7/4}{1-z} + \frac{1z/2}{(1-z)^2} + \frac{4}{2z-1} - \frac{9/4}{3z-1}$$

$$= \frac{7}{4}\frac{1}{(1-z)} + \frac{1}{2}\frac{z}{(1-z)^2} - \frac{4}{1-2z} + \frac{9}{4}\frac{1}{1-3z}$$

$$u_n = \frac{7}{4}(1)^n + \frac{7}{4}n - 4(2)^n + \frac{9}{4}(3)^n$$

Check:

$$u_0 = \frac{7}{4} - 4 + \frac{9}{4} = \frac{7-16-9}{4} = 0$$

$$u_1 = \frac{7}{4} + \frac{1}{2} - 8 + \frac{27}{4} = \frac{7+2-32+27}{4} = \frac{4}{4} = 1$$

2.3.6. Solve $y_{n+2} - 6y_{n+1} + 5y_n = 0$ with $y_0 = 2$, $y_1 = 6$.

Solution: Let $G(z) = \sum_{n=0}^{\infty} y_n\, z^n$ be the generating function of the sequence y_n

Given, $y_{n+2} - 6y_{n+1} + 5y_n = 0$ .....(1)

Multiply (1) by z^{n+2} and sum from $n=0$ to ∞

$$\sum_{n=0}^{\infty} y_{n+2}\, z^{n+2} - 6\sum_{n=0}^{\infty} y_{n+1}\, z^{n+2} + 5\sum_{n=0}^{\infty} y_n z^{n+2} = 0$$

$$G(z) - [y_0 + y_1 z] - 6z[G(z) - y_0] + 5z^2 G(z) = 0$$

$$G(z) - [2 + 6z] - 6z[G(z) - 2] + 5z^2\, G(z) = 0$$

$$(1 - 6z + 5z^2)\, G(z) - 2 - 6z + 12z = 0$$

$$(1 - 6z + 5z^2)\, G(z) = 2 - 6z$$

$[5z^2 - 6z + 1 = 0$

$$z = \frac{6 \pm \sqrt{36-20}}{10}$$

$$= \frac{6 \pm 4}{10};\ 1,\ 1/5$$

$(z - 1)(5z - 1) = 0]$

$$(1 - 6z + 5z^2)G = 2 - 6z$$

$$G(z) = \frac{2-6z}{5z^2-6z+1} = \frac{A}{z-1} + \frac{B}{5z-1}$$

$$2 - 6z = A(5z - 1) + B(z-1)$$

$$5A + B = -6$$

$-A - B = 2$

$4A = -4$; $A = -1$ and $B = -1$

$$G(z) = \frac{-1}{z-1} - \frac{1}{5z-1}$$

$$= \frac{1}{1-z} + \frac{1}{1-5z}$$

$y_n = 1 + 5^n$

2.4 Summary

The generating function of a sequence (a_n) is the expression

$$G(x) = \sum_{n=0}^{\infty} a_n \, x^n.$$

Generating functions were first introduced by Abraham de Moivre in 1730, in order to solve the general linear recurrence problem. Generating functions are not functions in the formal sense of a mapping from a domain to a co-domain; the name is merely traditional, and they are sometimes more correctly called generating series.

2.5 Keywords

Generating function, sequence, recurrence relation.

2.6 Supplementary Problems

2.6.1. Solve the recurrence relation $a_n = 5a_{n-1} + 2$, $n \geq 1$; given that $a_0 = 3$, by the method of generating function.

2.6.2. By the generating function method, solve $u_n - 2u_{n-1} - u_{n-2} = 0$ with $u_0 = 3$, $u_1 = -5$.

2.6.3. Solve $u_{n+1} - 3\, u_n = 7.2^n$ with $u_0 = -2$.

2.6.4. Solve the recurrence relation $a_n = 2a_{n-1} - 2a_{n-2} + 3^n$; $n \geq 2$, given that $a_0 = 2$, $a_1 = 8$.

2.6.5. Solve $y_{n+2} - 6y_{n+1} + 5y_n = 0$ with $y_0 = 2$, $y_1 = 6$.

2.7 References

1. Discrete mathematics, by P. Geetha (Scitech publications).

2. Discrete mathematics for computer science, by Kolman, Busby and Ross, PHI publications.

Unit 3: Introduction to Functions

Structure

3.0 Objectives

After going through this lesson you will be able to

- Explain the meaning of a function;
- Differentiate the various types of functions.
- Analyse the composition of functions.
- Evaluate the characteristic functions

3.1 Introduction

A relation is mainly a correspondence between the members of two sets, associating members of the first set with those of the second. It is possible that a given relation associates with any member of the first set several different members of the second set. It is possible that some elements of the first set are not associated with any from the second.

A special type of relation is that which associates with each member of the first set only one member of the second. Such a relation or correspondence is called a function from one set into the other. Thus a function is only a special type of relation or correspondence.

3.2 Basic Terminology

Definition: Let X and Y are any two sets. A relation f from X to Y is called a function if for every $x \in X$ there is a unique $y \in Y$ such that $(x, y) \in f$.

Note:

1) A function f from X to Y is an assignment of exactly one element of Y to every element of X.

2) If $y=f(x)$, then y is called the image of x and x is called the pre-image of y under f.

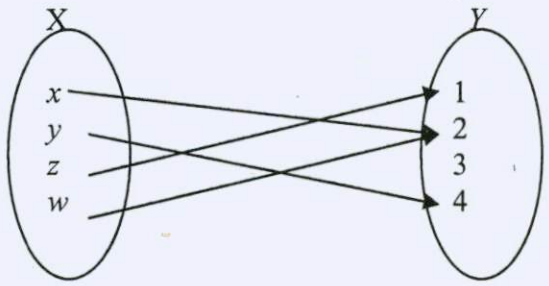
3) The set X is called the domain of f denoted by *Dom* (f) and Y is called the co domain of f.

4) The set of the images of all elements of X is called the range of f denoted by *Ran* (f).

$$\text{Ran}\,(f) = \{f(x): x \in X\}.$$

5) *Ran*(f) is a subset of Y.

Example : Let $X = \{x, y, z, w\}$ and $Y = \{1, 2, 3, 4\}$. If $Dom\ (f) = \{x, y, z, w\}$ and $f(x) = 2$, $f\ (y) = 4$, $f\ (z) = 1$, $f\ (w) = 2$, then the pictorial representation of f is



A function from *A* to *B*

$Ran(f) = \{ 1, 2, 4 \}$ which is a subset of the co domain Y.

3.3 Types of Functions

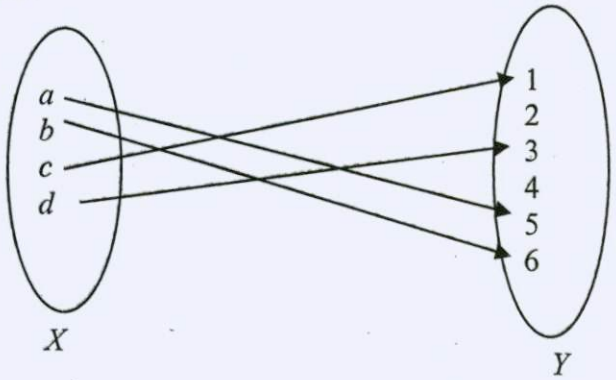
Definition: A function $f: X \rightarrow Y$ is called one-to-one or injective if distinct elements of X are mapped into distinct elements of Y.

In other words, f is one-to-one if and only if

$f(x_1) \neq f(x_2)$ whenever $x_1 \neq x_2$ or equivalently

$f(x_1) = f(x_2)$ whenever $x_1 = x_2$.

Example : The following function is one-to-one, since distinct elements of X are mapped into distinct elements of Y.



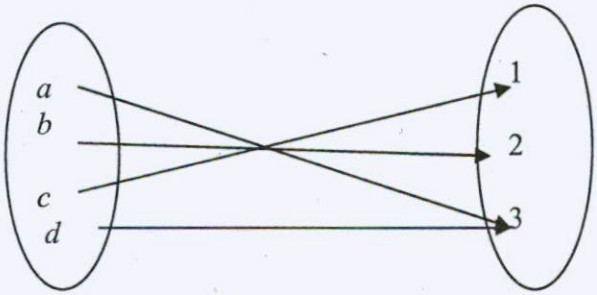
The function in Example 1 is not one-to-one, since

$f(x) = f(w) = 2$ but $x \neq w$.

Definition: A function $f: X \rightarrow Y$ is called onto or surjective if the range $Ran(f) = Y$. Otherwise it is called into.

In other words, a function f is onto, iff for every element $y \in Y$, there is an element $x \in X$ such that $f(x) = y$.

Example :

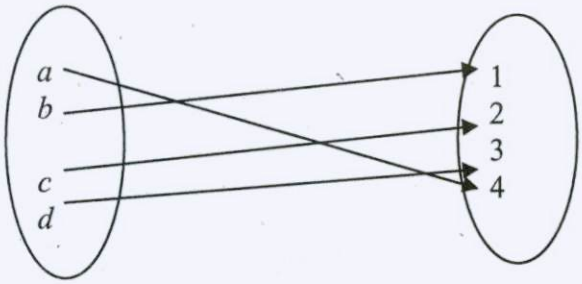


The above two examples are not onto.

Definition: A function $f: X \rightarrow Y$ is called bijective or bijection or one-to-one correspondence if it is both one-to-one and onto.

Obviously if X and Y are finite such that $f: X \rightarrow Y$ is bijective, then X and Y have the same number of elements.

Example :



3.4 Composition of Functions

Definition: If $f: A \rightarrow B$ and $g: B \rightarrow C$ then the composition of f and g is a new function from A to C denoted by $g \circ f$ given by:

$$(g \circ f)(x) = g\{f(x)\} \text{ for all } x \in A$$

3.4.1. Properties:

1. Composition of functions is associative.
   i.e., If $f: A \rightarrow B$, $g: B \rightarrow C$ and $h: C \rightarrow D$ are functions then $h \circ (g \circ f) = (h \circ g) \circ f$.

2. If $f: A \rightarrow B$ and $g: B \rightarrow C$ are functions, then $gof: A \rightarrow C$ is an injection; surjection or bijection accordingly as f and g are injections, surjections or bijections.

3.5 Identity Function

The function $f: A \rightarrow A$ where $f(x) = x$, $x \in A$ is called the identity function on A. In other words, the identity function is the function that assigns each element of A to itself and is denoted by I_A or simply I. The function I_A is a bijection.

3.6 Inverse of a Function

Definition: If $f: A \rightarrow B$ and g: B→A, then the function g is called the inverse of the function f, if g o f = I_A and f o g = I_B.

In other words, if x∈A and y∈B then, the function g: B→A is called the inverse of f: A→B if x = g(y) whenever y = f(x).

The inverse of f is denoted by f^{-1}. Thus if f^{-1} is the inverse of f then $x = f^{-1}(y)$ where y = f(x).

3.6.1. Properties:

1. The inverse of a function f, if exists is unique.

2. The necessary and sufficient conditions for the function f : A→B to be invertible is that f is one- to-one and onto.

3. If f : A→B and g : B → C are invertible function, then $g \circ f$: A → C is also invertible and $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

i.e., the inverse of the composition of two functions is equal to the composition of the inverse of the function in the reverse order.

3.7 Characteristic function of a set

Definition: If A is a subset of a universal set $\mathfrak{U}$, the characteristic function f_A of A is defined as the function from $\mathfrak{U}$ t o the set {0,1} such that

$$f_A(x) = \begin{cases} 1 & if\ x \in A \\ 0 & if\ x \notin A \end{cases}$$

Example: If $\mathfrak{U} = \{1, 2, 3, 4, 5\}$ and A ={2, 4} then

$f_A(1) = 0 = f_A(3) = f_A(5)$ and

$f_A(2) = f_A(4) = 1$ since 2, 4 ∈ A and 1,3,5 ∉ A

3.7.1. Properties of characteristic functions

1. If A is a subset of $\mathfrak{U}$ then $f_{\bar{A}}(x) = 1 - f_A(x)$ for all xε $\mathfrak{U}$

Proof :

$$f_{\bar{A}}(x) = 1 \quad \Leftrightarrow x \in \bar{A}$$

$$\Leftrightarrow x \notin A$$

$$\Leftrightarrow f_A(x) = 0$$

$$\Leftrightarrow 1 - f_A(x) = 1$$

$$\therefore\ f_{\bar{A}}(x) = 1 - f_A(x),\ \ for\ x \notin A.$$

$f_{\bar{A}}(x) = 0 \quad \Leftrightarrow x \notin \bar{A}$

$\Leftrightarrow x \varepsilon A$

$\Leftrightarrow f_A(x) = 1$

$\Leftrightarrow 1 - f_A(x) = 0$

$\therefore \quad f_{\bar{A}}(x) = 1 - f_A(x)$.

2. If A and B are any two subsets of $\mathfrak{U}$ then $f_{A\cap B}(x) = f_A(x) f_B(x)$, *for all* $x \in \mathfrak{U}$.

Proof: $f_{A\cap B}(x) = 1 \Leftrightarrow x \in A \cap B$.

$\Leftrightarrow x \in A$ and $x \in B$

$\Leftrightarrow f_A(x) = 1$ and $f_B(x) = 1$

$\Leftrightarrow f_A(x)\, f_B(x) = 1$

$\therefore f_{A\cap B}(x) = f_A(x) f_B(x)$, *when* $x \in A \cap B$.

$f_{A\cap B}(x) = 0 \Leftrightarrow x \notin A \cap B$.

$\Leftrightarrow x \notin A$ and $x \notin B$.

$\Leftrightarrow f_A(x) = 0$ and $f_B(x) = 0$

$\Leftrightarrow f_A(x)\, f_B(x) = 0$

$\therefore f_{A\cap B}(x) = f_A(x) f_B(x)$, *when* $x \notin A \cap B$.

Hence, $f_{A\cap B}(x) = f_A(x) f_B(x)$, *for all* $x \in \mathfrak{U}$.

3. If A and B are any two subsets of $\mathfrak{U}$ then

$f_{A\cup B}(x) = f_A(x) + f_B(x) - f_{A\cap B}(x)$, *where* $x \in \mathfrak{U}$.

Proof: $f_{A\cup B}(x) = 1 \Leftrightarrow x \in A \cup B$.

$\Leftrightarrow x \in A$ or $x \in B$

$\Leftrightarrow f_A(x) = 1$ or $f_B(x) = 1$

$\Leftrightarrow f_A(x) + f_B(x) - f_A(x) f_B(x) = 1$

$\Leftrightarrow f_A(x) + f_B(x) - f_{A\cap B}(x) = 1$

$\therefore f_{A\cup B}(x) = f_A(x) + f_B(x) - f_{A\cap B}(x)$, *where* $x \in A \cup B$.

$f_{A\cup B}(x) = 0 \Leftrightarrow x \notin A \cup B.$

$\Leftrightarrow x \notin A \text{ or } x \notin B$

$\Leftrightarrow f_A(x) = 0 \text{ or } f_B(x) = 0.$

$\Leftrightarrow f_A(x) + f_B(x) - f_A(x) f_B(x) = 0$

$\Leftrightarrow f_A(x) + f_B(x) - f_{A\cap B}(x) = 0$

$\therefore f_{A\cup B}(x) = f_A(x) + f_B(x) - f_{A\cap B}(x), \text{ where } x \notin A \cup B.$

$f_{A\cup B}(x) = f_A(x) + f_B(x) - f_{A\cap B}(x), \text{ where } x \in \mathfrak{U}.$

4. Using characteristic functions, prove that

$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

Proof: $f_{(A \cup B)}(x) \cdot f_{(AUC)}(x) = [f_A(x)+f_B(x)- f_{A\cap B}] \; [f_A(x)+f_C(x)- f_{A\cap C}(x)]$

$= f_A(x)\, f_A(x)+ f_A(x)\, f_C(x) - f_A(x)\, f_{A\cap C}(x)+ f_B(x)\, f_A(x)+ f_B(x)\, f_C(x) - f_B(x)\, f_{A\cap C}(x)-$

$f_{A\cap B}(x)\, f_A(x)- f_{A\cap B}(x)\, f_C(x) + f_{A\cap B}(x)\, f_{A\cap C}(x)$

$= f_A(x) + f_{A\cap C}(x)- f_{A\cap C}(x)+f_{A\cap B}(x)+ f_{B\cap C}(x)- f_{A\cap B\cap C}(x)- f_{A\cap B}(x)- f_{A\cap B\cap C}(x)+ f_{A\cap B\cap C}(x)$

$= f_A(x)+ f_{B\cap C}(x)- f_{A\cap B\cap C}(x)$

$= f_{AU(B \cap C)}(x)$ by property (3)

Hence the result.

3.8 Solved Problems

3.8.1. Determine whether or not each of the following relations is a function with domain {1, 2, 3, 4}. If any relation is not a function, explain why?

a) $R_1 = \{(1, 1), (2, 1), (3,1), (4,1), (3, 3)\}$

b) $R_2 = \{(1, 2), (2,3), (4,2)\}$

c) $R_3 = \{(1, 1), (2, 1), (3, 1), (4,1)\}$

d) $R_4 = \{(1, 4), (2, 3), (3,2), (4, 1)\}$

Solution:

a) R_1 is not a function since there are 2 pairs (3, 1) and (3, 3) which means that the image of the element 3 is not unique.

b) R_2 is not a function since there is no image for the element for the element 3 of the domain.

c) R_3 is a function even though the images of 1, 2, 3, 4 of the domain are one and the same element 1.

d) R_4 is a function.

3.8.2. If $f: R \rightarrow R$ and $g: R \rightarrow R$ are functions defined by $f(x) = x^2 + 3x + 1$ and $g(x) = 2x-3$, find $f \circ g$, $f \circ f$, $g \circ g$.

Solution:

$$(f \circ g)(x) = f[g(x)] = f(2x-3)$$
$$= (2x-3)^2 + 3(2x-3) + 1$$
$$= 4x^2 - 6x + 1$$

$$(g \circ f)(x) = g[f(x)] = g[x^2 + 3x + 1]$$
$$= 2(x^2 + 3x + 1) - 3$$
$$= 2x^2 + 6x - 1.$$

$$(f \circ f)(x) = f[f(x)] = f[x^2 + 3x + 1]$$
$$= (x^2 + 3x + 1)^2 + 3(x^2 + 3x + 1) + 1$$
$$= x^4 + 6x^3 + 14x^2 + 15x + 5.$$

$$(g \circ g)(x) = g[g(x)] = g[2x-3]$$
$$= 2(2x-3) - 3$$
$$= 4x - 9$$

3.8.3. If $f: R \rightarrow R$ and $g: R \rightarrow R$ are defined as $f(x) = x^2 - 2$ and $g(x) = x + 4$. Find $g \circ f$ and $f \circ g$ and state whether these functions are injective, surjective and bijective.

Solution.

$$(g \circ f)(x) = g[f(x)]$$
$$= g(x^2-2)$$
$$= x^2 - 2 + 4 = x^2 + 2.$$

$$(f \circ g)(x) = f[g(x)]$$
$$= f(x + 4)$$

$= (x+4)^2 - 2 = x^2 + 8x + 14$

Here $g \circ f \neq f \circ g$.

Given that $f(x) = x^2-2$

For one to one

Suppose, $f(x) = f(y)$

$\Rightarrow x^2 - 2 = y^2 - 2$

i. e. $x^2 = y^2$

i. e. $x = \pm y$

$\therefore f$ is not 1- 1 and also it is not onto.

Now consider, $g(x) = x+4$.

For 1-1:

$g(x) = g(y)$

$x+4 = y+4$

$x=y$.

$\therefore$ g is 1-1.

Let $y \in R$ (co domain), suppose $x \in R$ (Domain) such that

$g(x)=y$

$x+4=y$

$x=y-4$

For all $y \in R$, we know that $y - 4 \in R$.

Hence, g is an onto map.

3.8.4. Check whether the function $f(x) = x^2 - 11$ from R to R is 1- 1? Onto or both? Justify.

Solution. Given that $f(x) = x^2 - 11$, $x \in R$

For 1-1:

Suppose, $f(x) = f(y)$, then

$x^2 - 11 = y^2 - 11$

$$x^2 = y^2$$

$$x = \pm y$$

$\therefore$ f is not 1-1

For onto: For all $y \in R$., we have to show that there exists x such that $f(x) = y$.

$$x^2 - 11 = y$$

$$x^2 = y + 11$$

$X = \sqrt{y+11}$ which is not in R for various values of y

$\therefore$ f is not onto.

3.8.5. If S = {1, 2, 3, 4, 5} and if $f, g, h: S \to S$ are given by

f = {(1, 2), (2, 1), (3, 4), (4, 5), (5, 3)}

g = {(1, 3), (2, 5), (3, 1), (4, 2), (5, 4)}

h = {(1, 2), (2, 2), (3, 4), (4, 3), (5, 1)}

a). Verify whether $f \circ g = g \circ f$

b). Verify whether f, g and h have inverses..

c). Find f^{-1} and g^{-1}.

d). Show that $(f \circ g)^{-1} = g^{-1} \circ f^{-1} \neq f^{-1} \circ g^{-1}$.

Solution:

a) $(f \circ g)(1) = f[g(1)] = f(3) = 4$

$(f \circ g)(2) = f[g(2)] = f(5) = 3$

$(f \circ g)(3) = f[g(3)] = f(1) = 2$

$(f \circ g)(4) = f[g(4)] = f(2) = 1$

$(f \circ g)(5) = f[g(5)] = f(4) = 5$

$\therefore$ $f \circ g$ = {(1, 4) , (2, 3) , (3, 2) , (4, 1), (5, 5)}

Similarly $g \circ f$ = {(1, 5), (2, 3), (3, 2), (4, 4), (5, 1)}

$\therefore$ $f \circ g \neq g \circ f$.

b) Both f and g are 1-1 and onto

∴ They are invertible

$h(1) = h(2) = 2$

But $1 \neq 2$

∴ h is not 1-1.

Also range (h) = {1, 2, 3, 4} ≠S.

∴ h is also not onto.

Hence, the inverse of h does not exist.

c) f^{-1} is obtained by reversing the elements in all the ordered pairs of f.

f^{-1}= {(2, 1), (1, 2), (4, 3), (5, 4), (3, 5)}.

It is easy to verify that

$fof^{-1} = f^{-1}of$ = {(1, 1), (2, 2), (3, 3), (4, 4), (5, 5)} = I.

Similarly, g^{-1} = {(3, 1), (5, 2), (1, 3), (2, 4), (4, 5)}.

d) From f o g,

$(fog)^{-1}$ = {(4, 1), (3, 2), (2, 3), (1, 4), (5, 5)}.

From f^{-1} and g^{-1}

$g^{-1}o\ f^{-1}$ = {(2, 3), (1, 4) , (4, 1), (5, 5), (3, 2)}.

But, $f^{-1}o\ g^{-1}$={ (3, 2), (5, 1), (1, 5), (2, 3), (4, 4)}.

Therefore, $(f \circ g)^{-1} = g^{-1} \circ f^{-1} \neq f^{-1} \circ g^{-1}$.

3.8.6. Show that the function f: R → R defined by $f(x) = \frac{x}{x+4}$ is one-to-one and onto and hence find the inverse.

Solution: Given that f: R→ R is defined as $f(x) = \frac{x}{x+4}$

f is one-one:

Let f(x) = f(y) then x/(x+4) = y/(y+4)

$x(y+4) = y(x+4)$

i. e. $4x = 4y$

i. e. $x = y$.

Therefore, f is one-to-one

f is onto: if for every $y \in R$ there is a pre- image $x \in R$, such that $f(x) = y$.

i.e, $y = f(x) = \frac{x}{x+4}$

$y(x+4) = x$

$xy + 4y = x$

$x(y - 1) = -4y$

$x = \frac{4y}{1-y}$

Therefore, $f^{-1}(x) = \frac{4x}{1-x}$ is the inverse function.

3.8.7. Show that the function $f(x) = x^3$ and $g(x) = x^{1/3}$ for $x \in R$ are inverse of each other.

Solution.

Given that $f(x) = x^3$ and $g(x) = x^{1/3}$.

To Prove: $f = g^{-1}$ or $g = f^{-1}$

Its enough if we prove

$(fog)x = Ix$ or $(gof)(x) = Ix$

Consider, $(fog)x = f[g(x)] = f[x^{1/3}] = x = Ix$

Therefore, $fog = I$ (1)

Consider, $(gof)(x) = g[f(x)] = g[x^3] = x = Ix$

$gof = I$ (2)

From (1) and (2) we get

$f = g^{-1}$ or $g = f^{-1}$.

3.9. Summary

Definition: Let X and Y are any two sets. A relation f from X to Y is called a function if for every $x \in X$ there is a unique $y \in Y$ such that $(x, y) \in f$.

Definition: A function $f: X \to Y$ is called one-to-one or injective if distinct elements of X are mapped into distinct elements of Y.

Definition: A function $f: X \to Y$ is called onto or surjective if the range $Ran(f) = Y$.

Definition: If f: $A \rightarrow B$ and g: $B \rightarrow C$ then the composition of f and g is a new function from A to C denoted by $g \circ f$ given by:

$$(g \circ f)(x) = g\{f(x)\} \text{ for all } x \in A$$

Definition: If f: $A \rightarrow B$ and g: B→A, then the function g is called the inverse of the function f, if g o f = I_A and f o g = I_B.

Definition: If A is a subset of a universal set $\mathcal{U}$, the characteristic function f_A of A is defined as the function from $\mathcal{U}$ t o the set {0,1} such that

$$f_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$

3.10 Keywords

Function, one-one, onto, inverse, composite function.

3.11 Supplementary Problems

3.11.1. Let A = {a, b, c}, B = {x, y, z}. Determine whether or not each relation below is a function from A to B. Find the range if the relation is a function.

(i) f = {(a, y), (c, z)}

(ii) g = {(a, y),(b, z),(c, x),(c, z)}

(iii) h = {(a, z), (b, z), (c, x)}

3.11.2. State which of the following are injections, surjections or bijections from R into R, where R is the set of all real numbers.

i) f(x) = -2x ii) $g(x) = x^2 - 1$.

3.11.3. Let X={1, 2, 3, 4} and a mapping f: X→ X be given by f={(1, 2), (2, 3), (3, 4), (4, 1)}. Form the composite functions f^2, f^3, f^4.

3.11.4. If A = {1,2,3} and f,g,h are functions from A to A given by f = [(1,2), (2,3), (3,1)}, g={(1,2),(2,1), (3,3)} and h= {(1,1), (2,2), (3,1)}, find (i)fog, (ii) fohog, (iii) gof.

3.11.5. Show that the function $f(x) = x^3$ and $g(x) = x^{1/3}$ for x ∈R are inverse of each other.

3.11.6. Show that the function f: R – {3} → R –{1} given by f(x)= $\frac{x-2}{x-3}$ is a bijection and find its inverse.

3.11.7. If A and B are any two subsets of $\mho$ then prove that $f_{A-B}(x)= f_A(x)[1-f_B(x)]$

3.11.8. Using characteristic function, prove that

$f_{A\cap B}(x) = f_A(x)f_B(x)$, *for all* $x \in \mho$.

3.12 References

1. Discrete mathematics, by P. Geetha (Scitech publications).
2. Discrete mathematics for computer science, by Kolman, Busby and Ross, PHI publications.
3. Discrete mathematical structures with applications to computer science, by Tremblay and Manohar (McGraw-Hill publications).

Unit4: Hashing, Recursive and Permutation Functions

Structure

4.0 Objectives

After going through this lesson you will be able to

- Explain the Hashing function;
- Anlalyse the recursive function.
- Evaluate the permutation function.

4.1 Introduction

In a database huge amount of data is stored in the form of records. Each record contains a field called a key to that record. The key has a value that identifies a record in the computer storage. The address of a record in the storage is obtained by performing some reproducible arithmetic or logical operation on the internal bit representation of its key. Any transformation which maps the internal bit representation of the set of keys to a set of addresses is called a hashing function. Various hashing functions are available. We discuss the division method.

Recursion is a method of defining functions in which the function being defined is applied within its own definition; specifically it is defining an infinite statement using finite components. The term is also used more generally to describe a process of repeating objects in a self-similar way. Many useful recursively defined functions have domains that are inductively defined sets.

The notion of permutation is related to the act of permuting objects or values. A permutation of a set S is defined as a bijection from S to itself. Such a map f, is associated with the rearrangement of S in which each element s takes the place of its image f(s).

4.2 Hashing Functions

Definition: If n is the number of available memory locations and k is non-negative integer representing the key, the hashing function h(k) representing the address of the memory cell in which k is stored is defined as:

$$h(k) = k(\text{mod } n).$$

i.e., h(k) is simply the remainder when k is divided by n and it takes values from the set {0, 1, 2, …, n-1} known as address set.

A hashing function quite often maps different keys to the same address. In general a collision for a hash function occurs if $h(k_1)=h(k_2)$ but $k_1 \neq k_2$. It is necessary to provide storage space for and also a method of finding the colliding records. There are many techniques called collision resolution techniques for this purpose. The method called open addressing inserts the colliding record at the first empty location found.

4.3 Recursive Functions

Definition: A partial function f: X→Y is a rule which assigns to every element of X atmost one element of Y.

Definition: A total function f: X→Y is a rule which assigns to every element of X a unique element of Y.

Example: The function $f(r) = +\sqrt{r}$ is a partial function since f(r) is defined only for the positive real numbers and not for negative numbers.

Note:

A partial function can be made a total function if we restrict the domain of the function only to those values for which function value is defined.

Definition: The initial functions over N are (i) zero function, (ii) successor function, (iii) projection function which are defined by

(i) Zero function Z defined by $Z(x) = 0$

(ii) Successor function S defined by $S(x) = x+1$

(iii) Projection function U^n_i defined by $U^n_i\ (x_1, x_2, ..., x_n) = x_i$

Note:

As $U_1'(x) = x$ for every x in N, U_1' is simply the identify function on N.

Definition: If $f_1, f_2, ..., f_k$ are partial functions of n variables and g is a partial function of k variables, then the composition of g with $f_1, f_2, ..., f_k$ is a partial function h of n variables defined by

$h(x_1, x_2, ..., x_n) = g(f_1(x_1, x_2, ..., x_n), f_2(x_1, x_2, ..., x_n), ..., f_k(x_1, x_2, ..., x_n))$.

Example:

Let $f_1(x, y) = x+y$, $f_2(x, y) = 2x$, $f_3(x, y) = xy$ and $g(x, y, z) = x+y+z$.

Then,

$g(f_1(x, y), f_2(x, y), f_3(x, y)) = g(x+y, 2x, xy) = x+y+2x+xy$.

Thus the composition of g with f_1, f_2, f_3 is given by a function h defined by

$h(x, y) = x+y+2x+xy$.

Definition: The following operation which defines a function $f(x_1, x_2, ..., x_n, y)$ of n+1variables by using two other functions $g(x_1, x_2, ..., x_n)$ and $h(x_1, x_2, ..., x_n, y, z)$ of n and n+2 variables, respectively, is called recursion.

$f(x_1, x_2, ..., x_n, 0) = g(x_1, x_2, ..., x_n)$

$f(x_1, x_2, ..., x_n, y+1) = h(x_1, x_2, ..., x_n, y, f(x_1, x_2, ..., x_n, y))$

Definition: A function f is called primitive recursive if and only if it can be obtained from the initial functions by a finite number of operations of composition and recursion.

Definition: Let $g(x_1, x_2, ..., x_n, y)$ be a total function over N. g is said be a regular function if there exists some $y_0 \in N$ such that $g(x_1, x_2, ..., x_n, y_0) = 0$ for all n-tuples $(x_1, x_2, ..., x_n)$ in N^n.

Example: $G(x, y) = \min(x, y)$ is a regular function since $g(x, 0) = 0$ for all $x \in N$.

Definition: A function $f(x_1, x_2, ..., x_n)$ is said to be defined from a total function $g(x_1, x_2, ..., x_n, y)$ by minimization if

$$f(x_1, x_2, ..., x_n) = \begin{cases} \mu_y(g(x_1, x_2, ..., x_n, y) = 0) & \text{if there is such a } y \\ \text{undefined} & \text{otherwise} \end{cases}$$

where μ_y means the least y greater than or equal to zero.

Definition: A function is said to be recursive if and only if it can be obtained from the initial functions by a finite number of applications of the operations of composition, recursion and minimization over regular functions.

4.4 Permutation Function

Definition: A bijective function from A to A is called a permutation function from A to A.

Definition: The set of all bijective functions from A to A is called the set of permutation functions from A to A.

4.5 Solved Problems

4.5.1. A company has 10,000 customers. Each customer id is an eight digit number. The hashing function takes the first four digits as one number and the last four digits as another number, add them and then applies (mod 64) function to assign an address to the customer record. Determine the address assigned to the following numbers.

(a) 27266036 (b) 35674690

Solution:

(a) 27266036

$2726 + 6036 = 8762$

$h(8762) = 8762 \pmod{64}$

$\therefore \quad h(8762)=58.$

The number 27266036 is stored in the address 58.

(b) 35674690

$3567 + 4690 = 8257$

$h(8257) = 8257 \pmod{64}$

$\therefore \quad h(8257)=1.$

The number 35674690 is stored in the address 1.

4.5.2. Compute the addresses of 6 memory cells in which the integers 23, 38, 46, 55, 67 and 71 are to be stored, assuming there are 6 records in the file.

Solution:

Let n = 7, then the address of the memory cells are given by the hashing function $h(k) = k \pmod 7$.

The address set is {0, 1, 2, 3, 4, 5, 6}.

When k= 23, 38, 46, 55 the value of h(k)= 2, 3, 4, 6 respectively. The integers 23, 38, 46 and 55 are stored in the memory cells with addresses 2, 3, 4, 6.

h(k)	0	1	2	3	4	5	6
K	71	-	23	38	46	67	55

The next integer to be stored is 67. When k = 67, h(k)= 4.

i.e., 67 must be stored in the cell with address 4. But this cell with address 4 has been already occupied by 46. So, a collision has occurred.

By collision resolution policy, the first empty cell that follows the already occupied cell is used to store the current value of k.

The first unoccupied cell that follows the memory cell numbered as 4 is that with address 5. The integer 67 is thus stored in this cell. The last integer 71 is then stored in the cell with address 0. The cell with address 1 will remain as an unoccupied cell.

4.5.3. For the hashing function h(x) = x (mod 17) show how the following data would be interested in the order in given initially empty cells. Use the collision resolution policy of inserting the number in the next higher unoccupied cell. Cells are indexed from 0 to 16.

Given data: 714, 681, 26, 373, 775, 906, 509, 2032, 42, 4,136, 1028

Solution:

h(k)	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
k	714	509	631	136	4	906	-	-	42	26	775	2032	1028	-	-	-	373

0 = 714(mod 17) 10 =775 (mod 17) 8 =42 (mod 17)

2 = 631(mod 17) 5 = 906 (mod 17) 4 = 4(mod 17)

9 = 26(mod 17) 16 = 509(mod 17) 0 = 136(mod 17)

16 = 373(mod 17) 9 = 2032(mod 17) 8 = 1028(mod 17)

4.5.4. For the hashing function $h(x) = x^2 \pmod{11}$ show how the following data would be inserted in the order given initially empty cells. Use the collision resolution policy of inserting the number in the next higher occupied cell. All cells are indexed from 0 to 10.

Data: 53, 13, 281, 743, 377, 20, 10, 796.

Solution:

$4 = 53^2 \pmod{11}$ $3 = 281^2 \pmod{11}$ $9 = 377^2 \pmod{11}$

$1 = 10^2 \pmod{11}$ $4 = 13^2 \pmod{11}$ $3 = 743^2 \pmod{11}$

$4 = 20^2 \pmod{11}$ $5 = 796^2 \pmod{11}$

h(x)	0	1	2	3	4	5	6	7	8	9	10
x^2	-	10^2	-	281^2	53^2	13^2	743^2	796^2	-	377^2	20^2

4.5.5. For the hashing function $h(x) = x \pmod{11}$, show how the corresponding data given would be inserted in the order given in initially empty cells. Use the usual collision resolution policy to resolve collision. Cells indexed 0 to 10.

Data: 5, 15, 132, 102, 32, 257, 53.

Solution:

$5 = 5 \pmod{11}$ $4 = 15 \pmod{11}$ $0 = 132 \pmod{11}$

$3 = 102 \pmod{11}$ $10 = 32 \pmod{11}$ $4 = 257 \pmod{11}$

$9 = 53 \pmod{11}$

h(x)	0	1	2	3	4	5	6	7	8	9	10
X	132	-	-	102	15	5	257	-	-	53	32

4.5.6. Show that $f(x, y) = x+y$, $x, y \in N$ is primitive recursive.

Solution:

Note that $x+(y+1) = (x+y)+1$ (1)

L.H.S. of (1) can be expressed in terms of f. R.H.S. of (1) can be expressed in terms of the successor function S.

That is f(x, y+1)=f(x, y)+1 =S(f(x, y)).

Also, f(x, 0) = x.

Define f(x, y) as

$$f(x, 0) = x = U_1'(x)$$

$$f(x, y+1) = S(U_3^3(x, y, f(x, y)).$$

Now U_1^1, U_3^3, S are initial functions.

Thus, f is got by applying recursion for the functions U_1^1, U_3^3 and S. Hence f is primitive recursive.

4.5.7. Show that f(x, y) = x*y is a primitive recursive function.

Solution:

$$f(x,0) = x*0 = 0 \quad (1)$$

$$f(x, y+1) = x*(y+1) = (x*y)+x = f(x, y) + x \quad (2)$$

Comparing (1) & (2) with definition we can write

$$f(x, 0) = z(x) \quad (3)$$

$$f(x, y+1) = f_1(U_3^3(x, y, f(x, y)), U_1^3(x, y, f(x, y)) \quad (4)$$

where $f_1(x, y) = x+y$ which is primitive recursive.

Taking g = Z and h defined by $h(x, y, z) = f_1(U_3^3(x, y, z), U_1^3(x, y, z))$, we see that (3), (4) define f_1 by recursion. As Z is an initial function of g = Z is primitive recursive.

As h is defined using composition of f_1, which is primitive recursive, U^3_3, U^1_3 which are initial functions, h is primitive recursive. Hence f_2, obtained from g and h, using recursion is primitive recursive.

4.5.8. Show that $f(x, y) = x^y$ is primitive recursive .

Solution:

Let $f(x, 0)= x^0 = 1$

$$f(x, y+1) = x^{y+1}+1 = x * x^y = x * f(x, y)$$

Define f(x, 0) = 1

$$f(x, y+1) = x* f(x, y)$$

$= U_1^3((x, y, f(x, y)) * U_3^3 (x, y , f(x, y))$

Now $f(x, 0) = S(Z(x))$ (S o Z is primitive recursive)

$f(x, y+1) = h(x, y, f(x, y))$

where $h(x, y, z) = U_1^3(x, y, z) * U_3^3(x, y, z)$.

U_1^3, U_3^3 are initial functions and $f_2(x, y) = x*y$ is primitive recursive, we see that f is defined by applying recursion to primitive recursive functions S(z(x)) and h. hence f is primitive recursive.

4.5.9. Find all permutation of A = {1, 2, 3}.

Solution: The permutation of A are

$$P_1=\begin{bmatrix}1 & 2 & 3\\ 1 & 2 & 3\end{bmatrix} \quad P_2=\begin{bmatrix}1 & 2 & 3\\ 1 & 3 & 2\end{bmatrix} \quad P_3=\begin{bmatrix}1 & 2 & 3\\ 2 & 1 & 3\end{bmatrix}$$

$$P_4=\begin{bmatrix}1 & 2 & 3\\ 3 & 2 & 1\end{bmatrix} \quad P_5=\begin{bmatrix}1 & 2 & 3\\ 2 & 3 & 1\end{bmatrix} \quad P_6=\begin{bmatrix}1 & 2 & 3\\ 3 & 1 & 2\end{bmatrix}$$

4.5.10. Let A = {1, 2, 3, 4}, f : A → A be defined by f(1) = 2, f(2) = 1, f(3) = 4, f(4) = 3. Write this in permutation notation.

Solution:

$$\begin{bmatrix}1 & 2 & 3 & 4\\ 2 & 1 & 4 & 3\end{bmatrix}$$

4.5.11. Find the inverse of permutation $\begin{bmatrix}1 & 2 & 3\\ 2 & 3 & 1\end{bmatrix}$

Solution:

Inverse permutation is $\begin{bmatrix}1 & 2 & 3\\ 3 & 1 & 2\end{bmatrix}$

4.5.12. If $p_1=\begin{bmatrix}1 & 2 & 3\\ 2 & 1 & 3\end{bmatrix}$, $p_2=\begin{bmatrix}1 & 2 & 3\\ 3 & 1 & 2\end{bmatrix}$., find p_2 o p_1.

Solution:

$$p_1: \begin{matrix}1 & 2 & 3\\ \downarrow & \downarrow & \downarrow\\ 2 & 1 & 3\end{matrix}$$

$$p_2: \begin{matrix}\downarrow & \downarrow & \downarrow\\ 1 & 3 & 2\end{matrix}$$

Hence, $p_1 \circ p_2 = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{bmatrix}$.

4.6 Summary

Definition: If n is the number of available memory locations and k is non-negative integer representing the key, the hashing function h(k) representing the address of the memory cell in which k is stored is defined as:

$$h(k) = k(\text{mod } n).$$

Definition: A function f is called primitive recursive iff it can be obtained from the initial functions by a finite number of operations of composition and recursion.

Definition: A function is said to be recursive iff it can be obtained from the initial functions by a finite number of applications of the operations of composition, recursion and minimization over regular functions.

Definition: The set of all bijective functions from A to A is called the set of permutation functions from A to A.

4.7 Keywords

Hashing function, recursion, permutation.

4.8 Supplementary Problems

4.8.1. For each hashing function, show how the corresponding data given would be inserted in the order given in initially empty cells. Use the usual collision resolution policy to resolve collision.

(i) $H(x)=(x^2+x)(\text{mod}17)$; cells indexed 0 to 16; data: 714, 631, 26, 373, 775, 906, 509, 2032, 47, 4, 136, 1028.

(ii) $H(x) = x+5(\text{mod } 11)$; cells indexed 0 to 10; data: 53, 13, 281, 743, 377, 20, 10, 796.

4.8.2. Show that the following functions over N are primitive recursive

(i) Constant function over N.

(ii) Predecessor function.

(iii) Proper subtraction function.

(iv) Zero test function.

(v) Odd and even parity function.

4.8.3. Show that if f(x, y) defines the remainder upon division of y by x, then it is a primitive function.

4.8.4. Let A={1,2,3,4,5,6} and $p_1 = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & 1 & 2 & 6 & 5 \end{bmatrix}$, $p_2 = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 1 & 5 & 4 & 6 \end{bmatrix}$,

$p_3 = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 3 & 2 & 5 & 4 & 1 \end{bmatrix}$, find

(i) p_2^{-1},

(ii) $(p_2 \circ p_1) \circ p_2$,

(iii) $p_1 \circ (p_3 \circ p_2^{-1})$,

(iv) $p_1^{-1} \circ p_2^{-1}$,

(v) $p_3 \circ (p_2 \circ p_1)^{-1}$.

4.9 References

1. Discrete mathematics, by P. Geetha, Scitech publications.
2. Discrete mathematics for computer science, by Kolman, Busby and Ross, PHI publications.
3. Discrete mathematical structures, by N. G. Goudru, Himalaya Publishing House.
4. Discrete mathematical structures with applications to computer science, by Tremblay and Manohar (McGraw-Hill publications).

Karnataka State Open University

Manasagangothri, Mysore 6

M.Sc Computer Science - Semester 1

MSC-501: DISCRETE MATHEMATICS

Module 4: Graph Theory

Contents

Introduction

A course on Discrete Mathematics studied by Computer Science students usually covers Graph theory at introductory level, as the subject of graph theory has many applications in Computer Science.

This module introduces the concept of graph and lists some important and interesting applications of graphs. Matrix representation of graphs is important from the view point of development of algorithms. Concepts like walks, paths, circuits lay foundation to the discussion on Hamiltonian circuit and travelling salesman problem. These are fundamental problems in the study of algorithm design and complexity issues. Trees are another class of graphs which are equally important as circuits. Tree is a very useful data structure. Binary search tree is often used to perform searching efficiently in many implementations that perform the task of searching. Spanning trees and weighted spanning trees have abundant practical applications.

All these concepts are discussed in a simple manner, as this is just one of the modules of the course, Discrete Mathematics.

We hope you will enjoy learning Graph Theory. The practice problems given at the end of each unit are classified as Easy and Challenging. Easy problems are solvable just by learning the concepts discussed in the notes provided. As many examples are covered in the text, the easy problems can be solved without any difficulty. The other kind of problems namely 'Challenging' are tough ones. You may probably have to go through the chapters in the reference texts, the details of which may be found at the end of each unit. Attempting to solve the exercise is as important as reading the notes. The learning is incomplete without the attempt towards solving problems.

Unit 1: Basic concepts

Structure

1.0: Objectives

After going through this unit you will be able to

- Explain graphs and its diagrammatic representation
- Explore some interesting applications
- Identify the various types of graphs, classify vertices and edges
- Make a matrix representation of graphs

1.1 Definition and representation

Here we discuss the definition of a graph and representation of a graph.

Definition

A graph G = (V, E) consists of a set of vertices V = {v_1, v_2, ...} and a set of edges E = {e_1, e_2, ...}. Each edge e_k is denoted by a pair of vertices (v_i, v_j). These vertices v_i, v_j are end vertices of edge e_k.

Representation

Most common representation is a diagram such as the one given in figure 1.1.

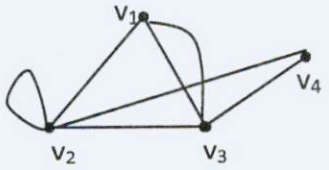


Fig 1.1: Example graph

There are 4 vertices (shown as dots) v_1, v_2, v_3, v_4 and 7 edges (shown in the figure as lines/curves/loops). Observe there are 2 edges connecting v_1 and v_3 and these edges are called parallel edges. Also there is an edge that begins and ends at v_2. This is called self loop. You may observe that edge connecting v_2 and v_4 intersects two other edges but not every intersection of edges is a vertex. Such intersecting edges should be thought of as being in different planes and thus have no common points.

Note 1: Vertices are also called as nodes, points, junctions. Edges can also be called as arcs, branches, lines.

Note 2: The definition of graph does not impose restriction on the length or shape of the edge, nor does it prescribe the order among the list of vertices or the positioning of the vertices. Therefore for a given graph there is no unique diagram which represents the graph. This point may be clear by the following example.

The graph G=(V, E) where V={1,2,3,4} and E={(1,1), (1,2), (1,3), (2,3), (2,4), (4,1)} can be represented by one of the following figures. Observe that many more such representations are possible.

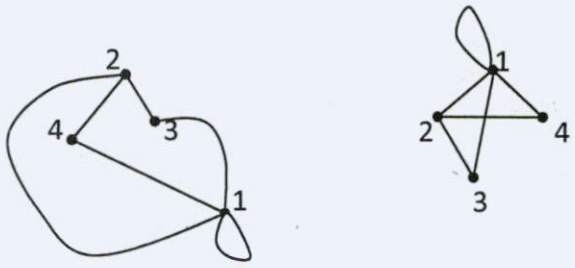


Fig 1.2: Alternate diagrams for G

1.2 Applications

Application fields are diverse. Here we mention some key fields such as physical, social, engineering, computer science where graphs are used commonly.

1.2.1 Origin of graph theory

Königsberg bridge problem:

Figure 1.3 below shows two islands C and D formed by the river Pregel (in Königsberg-present name Kaliningrad, Russia) and the seven bridges connecting the two banks A and B. An interesting question here is whether or not a person can start from a land area (A, B, C or D) and walk over all the bridges exactly once and return to the starting place. Such a tour is called Euler line.

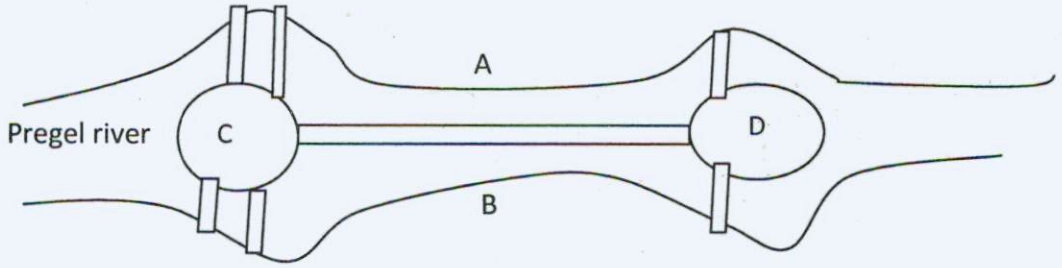


Fig 1.3: Bridges (shown as rectangles) on the river Pregel

Euler (a renowned mathematician of 18th century) published his first ever paper in the then new subject 'Graph Theory' and proved that above tour is not possible. Euler converted this problem to a simple graph where each land area is a vertex and bridges are edges connecting the vertices. The graph of Königsberg bridge problem is given in figure 1.4.

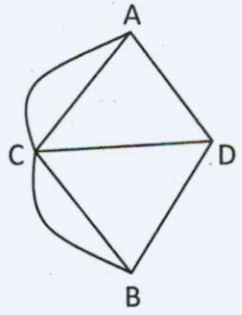


Fig 1.4: Graph of Königsberg bridge problem

Königsberg bridge problem is the same as drawing the graph in figure 3 without lifting the pen and without retracing any edge.

1.2.2 Utilities problem (application for physical)

There are 3 houses H1, H2, H3 and 3 utilities Water (W), Electricity (E) and Gas (G) are to be provided to each of the house by means of conduits. Is it possible to make such connections without any crossovers of the conduits?

This problem can be represented by the graph below (figure 1.5).

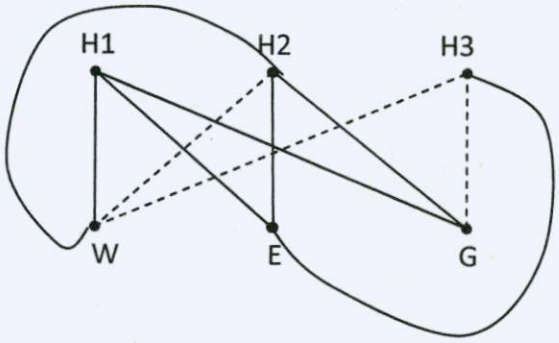


Fig 1.5: Graph of utilities problem

In the graph above, houses, utilities are vertices and conduits are edges. Note that in the graph some crossovers are unavoidable. The crossovers are shown in dotted lines.

1.2.3 Network problem (application in the field of engineering/computer science)

The interconnection of computing systems (LAN, WAN) can be represented using a graph. Nodes (vertices) are systems and edges are connections between systems. Some types of connections are shown below in figure 1.6.

Star connection

Bus connection

Fig 1.6: Interconnections of computing systems

When a connection fails, it is equivalent of removal of an edge. Strength of a network, connectivity between two systems etc can be studied using concepts of graph theory.

1.2.4 Club meet problem (social application of graph theory)

Six members of club meet daily for lunch together at a round table. They decide to sit with new neighbor during each day. How many days can this arrangement last?

Equivalent graph: Vertex can be used to represent members and edges represent neighbor relationship. Figure 1.7 shows two such seating arrangements.

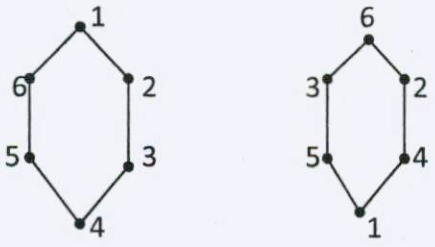


Fig 1.7: Two seating arrangements with different neighbors

The number of such arrangements can be found using by graph theoretic considerations.

1.3 Additional terminologies

In 1.1 we discussed about special types of edges namely parallel and self loops. Here we learn more about vertices, edges and graphs. We underline the new terminologies as and when we introduce them to you.

Simple graph is one which does not have any self loop or parallel edges. Referring to figure 1.1 if the self loop connecting v_2 to v_2 and one of the edges connecting v_1 and v_3 are removed it becomes simple graph. The graph after deletion of the two edges is shown in figure 1.8.

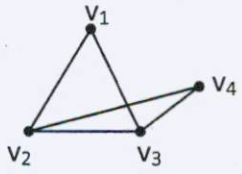


Fig 1.8: A simple graph

Fig 1.9: An infinite graph

So far we discussed graphs having finite number of vertices and edges (called finite graph). A graph can also have infinite number of vertices and edges which is then called infinite graph. Figure 1.9 shows portion of one such infinite graph. Assume every intersection lines to be a vertex.

A simple graph with edges connecting every pair of vertices is called complete graph. Figure 1.10 is an illustration of a complete graph of 4 vertices.

If an edge e_k connects v_i to v_j we say that e_k is incident on v_i and v_j. The degree of a vertex is the number of edges incident on it. Note the in case of complete graph of n vertices the degree of every vertex is n-1. Figure 1.10 shows a graph and the degree of each vertex indicated.

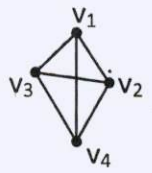


Fig 1.10: Complete graph of 4 vertices. Degree of every vertex is 3

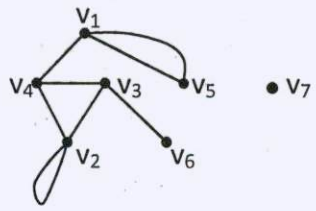


Fig 1.11: Degrees of vertices 1 to 7, in order, are 3, 4 (self loops are incident twice), 3, 3, 2, 1, 0

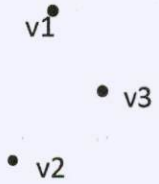


Fig 1.12: Null graph of 3 vertices

A vertex having no edge is called an isolated vertex. In figure 1.11 v_7 is an isolated vertex. Isolated vertex has degree 0. A vertex of degree one is called a pendant vertex. In figure 1.11 v_6 is a pendant vertex.

Number of vertices is order of a graph and the number of edges is its size. For example the order and size of graphs in figures 1.10 and 1.11 are 4, 6 and 7, 8 respectively.

A graph with no edges is a null graph. This graph has only vertices and no edges. All vertices are isolated vertices. Figure 1.12 above is a null graph of three vertices.

Adjacent edges are those which are incident on a common vertex. For example in figures 1.10 and 1.11 (v_1, v_2), (v_1, v_3) and (v_2, v_3), (v_3, v_6) are adjacent edges.

Vertices are said to be adjacent if they are end vertices of an edge. For example in figures 1.10 and 1.11 vertices v_2, v_4 and v_1, v_5 are adjacent.

A graph with some parallel edges is called multi-graph. Graphs in figures 1.1, 1.4 and 1.11 are multi-graphs.

A graph is said to be regular graph if all the vertices are of same degree. Figures 1.13 and 1.14 below illustrate regular and not a regular graph.

Fig 1.13: 3 vertex regular graph

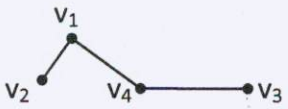


Fig 1.14: Not regular

A graph with weights attached to edges is said to be weighted graph. Graph below (figure 1.15) is a weighted graph.

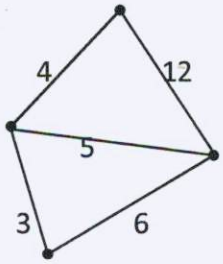


Fig 1.15: Weighted graph

Applications of weighted graphs are plenty. For instance, a graph representing a system of pipe lines through which commodities are transferred can become a weighted graph, where weights can be capacities of pipe lines. A graph of roads may be a weighted graph where weights of edges could be length of roads or width of roads or the volume of traffic in the roads.

Graphs discussed so far contain no direction for edges. Such graphs are called are undirected graphs. In these graphs edges (1,2) and (2,1) are one and the same edge. Digraph (or directed graph) is a graph where edges have directions. In a digraph edge (1,2) is an edge with initial vertex 1 and terminal vertex 2. Edges (1,2) and (2,1) are different in digraphs. If both (1,2) and (2,1) are present in a digraph, then it can be replaced by an undirected edge (1,2). Graphs can have directed and undirected edges. Such graphs are mixed graphs. Figure 1.16 shows diagrams of digraph and mixed graph.

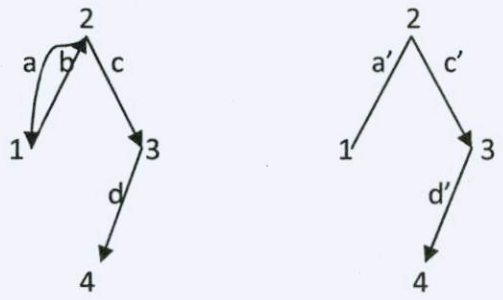


Fig 1.16: Digraph and equivalent mixed graph

Note: Most discussions in this module are focused on undirected graphs. Unless otherwise mentioned, by graph we mean undirected graph.

1.4 Matrix representation

Although pictorial representation is simple and convenient, other representations are better for computer processing. Graphs can also be represented in the form of a matrix. Many derivations are easy with matrix representation. There are two types of matrix representation. Incidence matrix or vertex edge incidence matrix is used to represent undirected graphs with no self loops. This is vertex by edge binary matrix. The number of rows is same as number of vertices and number of columns is same as number of edges. Given below are an undirected graph and its incidence matrix.

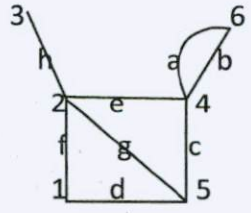


Fig 1.17: An undirected graph G

	a	b	c	d	e	f	g	h
1	0	0	0	1	0	1	0	0
2	0	0	0	0	1	1	1	1
3	0	0	0	0	0	0	0	1
4	1	1	1	0	1	0	0	0
5	0	0	1	1	0	0	1	0
6	1	1	0	0	0	0	0	0

Incidence matrix A(G)

Some observations on incident matrix:

1. Each column has two 1s, since each edge is incident on two vertices. (No incident matrix representation for graphs with self loops). Observe that column 1 has two ones at rows 4 and 6. (edge a is incident on vertices 4 and 6)

2. The number of 1s in a row is degree of the vertex. Observe that row has two 1s at column d and f. (edges d and f are incident on vertex 1)
3. A row with only 0s represent isolated vertex.
4. Parallel edges produce two identical columns. (columns 1 and 2 corresponding to parallel edges a and b are identical)

As an alternative to incidence matrix, it is sometimes more convenient to use another representation called adjacency matrix or connection matrix. Adjacency matrix can be used for undirected graphs and digraphs. This cannot represent graphs with parallel edges. Self loops can be represented. It is vertex by vertex matrix and also binary like incidence matrix. Adjacency matrix X is defined thus: $x_{ij} = 1$ if there is an edge from i to j and this entry is 0 in absence of the edge.

The adjacency matrices corresponding to the graphs in figures 1.2 (undirected graph) and 1.16 (mixed graph) is given below.

	1	2	3	4
1	1	1	1	1
2	1	0	1	1
3	1	1	0	0
4	1	1	0	0

Adjacency matrix of graph in figure 1.2

	1	2	3	4
1	0	1	0	0
2	1	0	1	0
3	0	0	0	1
4	0	0	0	0

Adjacency matrix of graph in figure 1.16

Observe that adjacency matrix of undirected graph is symmetric.

Entries in diagonal position are 1 only if there is a self loop. Degree of a vertex is number of 1s in the row or column. For example, number of 1s in row 2 of matrix (on the left) is 3 (= degree of vertex 2 in the graph of figure 1.2). In case of self loop the degree is number of 1s in the row plus one. (Degree of vertex 1 is 5 = number of 1s in row 1 +1).

We close the section with frequently used yet very simple results.

Theorem 1.1

The sum of the degrees of all vertices is twice the number of edges.

Proof:

Each edge is incident on two vertices. Hence each edge contributes to two degrees. If total number of edges is e, then the sum of all degrees is 2e.

For example consider the graph in figure 1.15. Total number of edges is 5. The degrees of vertices of the graph are 2,3,2,3 (beginning from top and in clockwise direction) their sum being 2+3+2+3=10 = 2*number of edges.

Considering the graph in figure 1.11, the number of edges is 8. The degrees of vertices v_1 to v_7 in order are 3, 4, 3, 3, 2, 1, 0. The sum of degrees is 16 which is 2*8=2*number of edges.

Theorem 1.2

The number of vertices of odd degree in a graph is always even.

Proof:

Suppose that d(x) denotes the degree of the vertex x in the graph. Group the vertices into even and odd degrees. Then $\Sigma_{all}\, d(x) = \Sigma_{even}\, d(y) + \Sigma_{odd}\, d(z)$. From previous theorem left hand side of the equation is an even number. Also first term on the right hand side is even. Hence the second summation on the right hand side should also be even. But each d(z) in the summation is odd. Hence total number of terms in the summation, $\Sigma_{odd}\, d(z)$, should be even to make the sum, $\Sigma_{all}\, d(x)$ even. Therefore the number of odd degree vertices is even.

For example in the graph of figure 1.16, (directed or mixed graph) all 4 vertices are of odd degree. In figure 1.17, 2 vertices namely 3 and 5 are of odd degree.

Theorem 1.3

Maximum number of edges in a simple graph of n vertices is n(n-1)/2

Proof:

You may observe that a simple graph with maximum number of edges is nothing but a complete graph.

Let us draw simple graphs of 1, 2 3, 4 vertices, with maximum number of edges.

The maximum number of edges in the these graphs of 1, 2, 3, 4 vertices are 0, 1, 3, 6 … Thus in an n vertex graph the maximum number of edges is n(n-1)/2. (This is nothing but the n^{th} term of the series 0, 1, 3, 6 …)

1.5 Summary

To summarize, in the first section of this unit we started with definition of a graph and a diagrammatic representation, came across interesting applications in diverse fields in section 2, learnt some simple and basic concepts such as (i) types of graphs – simple, multi-graph, complete, regular, null, infinite, weighted; (ii) concept of incidence and degree of vertices; (iii) isolated, pendant vertices (iv) order and size of graph (v) adjacent vertices and edges in section 3. Matrix representation of graphs is important from the point of view of computer processing and this is detailed in section 4. Also in section 4 some interesting results are presented with proofs.

1.6 Key words

Diagrammatic representation of graphs, graphs-various types, incidence matrix, adjacency matrix

1.7 Practice problems

Easy to do: Based on the text material

1. Draw a graph with 6 vertices having self loops, parallel edges, pendant vertices, isolated vertices. Indicate each of these.
2. Mention some applications of graph theory.
3. Describe the origin of graph theory.
4. Define complete graph, simple graph, and regular graph. Illustrate each one of these. Also provide illustration for not complete, not simple and not regular graph. Give reasons for each of the illustration.
5. What do you mean by infinite, null and weighted graphs?

6. Find the degree of each of the vertices in the illustrations of problem 2.
7. Find order and size of the graph for every illustration of problem 2.
8. Define adjacency of vertices and edges.
9. Draw digraphs and mixed graphs with 3, 4, 5 vertices.
10. Find incidence and adjacency of all graphs (whenever possible) you have drawn.

Challenging: Requires the support of reference texts

1. Convince yourself that an infinite graph with a finite number of edges must have infinite number of isolated vertices.
2. Convince yourself that an infinite graph with finite number of vertices should have infinite number of parallel edges (or parallel loops) connecting at least two vertices.
3. Show that the maximum degree of any vertex in a simple graph of n vertices is n-1.
4. Show that sum of in degrees of all nodes of a simple digraph is equal to the sum of out degrees of all its nodes and that this sum equal to number of edges of the graph.

1.8 References

1. Narsingh Deo, Graph Theory with applications to Engineering and Computer Science, PHI, Chapter 1 (Sections 1.1 to 1.5)
2. J.P.Tremblay, R.Manohar, Discrete Mathematical Structures with applications to Computer Science, TATA McGRAW-HILL, Chapter 5 (Section 5.1.1, 5.1.3)
3. Dr. N.G.Goudru, Discrete Mathematical Structures, Himalaya Publishing House, chapter 7 (Section 7.1, 7.11)
4. Bernard Kolman, Robert C. Busby, Sharon Cutler Ross, Discrete Mathematical Structures, PEARSON Education, Chapter 8 (Section 8.1)

Unit 2: Paths and circuits

Structure

2.0 Objectives

After a detailed study of this unit, you will be able to

- Solve some problems introduced in the previous unit
- Describe isomorphism, sub-graphs, connectivity and components and identify matrix form for these
- Explain concepts like path, walk, circuit, Euler graph
- Discuss Hamiltonian paths, circuits and travelling salesman problem - an important application of graph theory

2.1 Isomorphism

Isomorphism is similar to the concept of 'congruent' or 'equivalent' in geometry. Two graphs G and G' are called isomorphic if there is a one to one correspondence between their vertices and their edges preserve incidence relationship. In other words if an edge e is incident on vertices v_1 and v_2 in G, then the corresponding edge e' in G' must be incident on the vertices v_1' and v_2' that correspond to v_1 and v_2. Given below (figure 2.1) is a pair of isomorphic graphs.

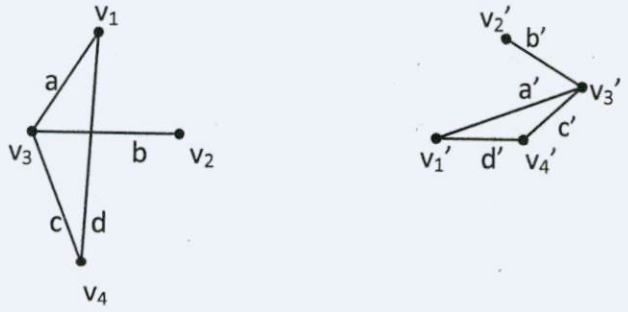


Fig 2.1: Isomorphic graphs

The correspondence between vertices and edges in the graphs of figure 2.1 are as follows: Vertices v_1, v_2, v_3, v_4 correspond to v_1', v_2', v_3', v_4'and edges a, b, c, d correspond to a', b', c', d'.

Given below are more graphs in figures 2.2 and 2.3.

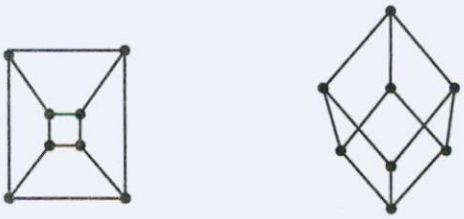


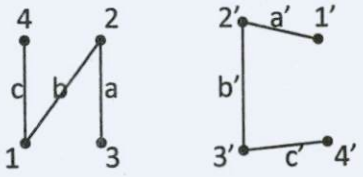
Fig 2.2: Isomorphic graphs

Fig 2.3: Not isomorphic graphs

Except for relabeling of vertices and edges, isomorphic graphs are the same, perhaps drawn differently. It is not easy to discover isomorphism. Definition of isomorphism is as follows: Isomorphic graphs must have: (i) The same number of vertices (ii) The same number of edges (iii) An equal number of vertices with a given degree. However these are not sufficient. For example the graphs in figure 2.3 satisfy all these and yet not isomorphic. This is because, there are two pendant vertices that are adjacent to a vertex of degree three in the graph on the left whereas there is only one pendant vertex adjacent to the vertex of degree 3 in the graph to the right. Hence these graphs are not isomorphic.

We now discuss the concept of isomorphism with the aid of matrix representation of graphs. Graphs G and G' are isomorphic if and only if their incidence matrices A(G) and A(G') differ by permutations of rows and columns.

As an example let us consider the following graphs of 4 vertices.



These are isomorphic with correspondence as follows: Vertices 1,2,3,4 correspond to 3', 2', 1', 4'. Edges a, b, c correspond to a', b', c'. Let us examine their incidence matrices, which are given below.

	a	b	c
1	0	1	1
2	1	1	0
3	1	0	0
4	0	0	1

	a′	b′	c′
1′	1	0	0
2′	1	1	0
3′	0	1	1
4′	0	0	1

Observe that after exchange of rows 1' and 3' we see that matrix on right is same as that on the left. Hence G and G' are isomorphic.

2.2 Sub-graphs

A graph g is said to be a sub-graph of a graph G if all the vertices and all the edges of g are present in G. A graph of 6 vertices and its sub-graph with 3 vertices are given in figure 2.4.

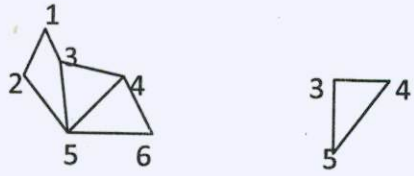


Fig 2.4: Graph and a sub-graph

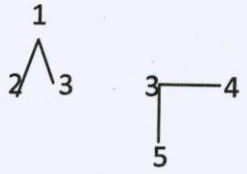


Fig 2.5: Edge disjoint sub-graphs

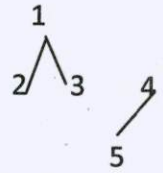


Fig 2.6: Vertex disjoint sub-graphs

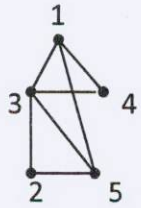
The concept of sub-graph is similar to that of subset. Sometimes the symbol "⊂" ($g \subset G$) is used instead of the word sub-graph.

The following observations can be made immediately and without any difficulty.

1. Every graph is its own sub-graph.
2. A sub-graph of a sub-graph of G is also a sub-graph of G.
3. A single vertex in a graph is a sub-graph.
4. A single edge together with its end vertices is a sub-graph.

Two or more sub-graphs that do not have any edge in common are called edge disjoint sub-graphs. However they can have common vertices. Refer figures 2.4 and 2.5 for such examples.

Coming to matrix form, g is a sub-graph of G if A(g) is a sub-matrix of A(G). Sub-matrix is one which is obtained after deletion of some rows and columns. As an example let us consider the following graph and its sub-graph.



Graph G

Sub-graph g

	a	b	c	d	e	f	g
1	1	1	1	0	0	0	0
2	0	0	0	1	1	0	0
3	1	0	0	1	0	1	1
4	0	1	0	0	0	1	0
5	0	0	1	0	1	0	1

	b	f
1	1	0
3	0	1
4	1	1

Incidence matrices of G and g

In the graph G assume that a, b, c, d, e, f, g are the edges (1,3), (1,4), (1,5), (2,3), (2,5), (3,4), (3,5) respectively.

It is obvious that second (on the right) matrix is a sub-matrix of first one on the left.

2.3 Walks, paths and circuits

Walk

Walk is an alternating sequence of vertices and edges beginning and ending with vertices. Also each edge in the sequence is such that its beginning vertex is the ending vertex of the previous edge. No edge can appear more than once in a walk. Length of the walk is the number of the edges in the walk.

Refer the graph in figure 2.4 with 6 vertices.

Examples:

1. 1, (1,2), 2, (2,5), 5, (5,3), 3 is a walk. It begins at 1 and ends at 3. Terminal vertices of this walk are 1 and 3. These vertices are distinct. This walk is called open walk. Length of this walk is 3.
2. 1, (1,2), 2, (2,5), 5, (5,4), 4, (4,3), 3, (3,1), 1 is also a walk. Length is 5. Terminal vertices are same. This is called closed walk.
3. 3, (3,5), 5, (5,6), 6, (6,4), 4, (4,5), 5, (5,2), 2 is another walk. This is also an open walk. Observe that vertex 5 is visited twice in the walk. Length of this walk is 5.
4. 1, (1,2), 2, (2,5), 5, (5,2), 2 is not a walk. This is because the same edge (2,5) (note that (2,5) and (5,2) are the same edge) is traversed twice.

Path

Path is an open walk with no vertex repetition in the sequence. That is no vertex is revisited. Length of the path is the number of edges in the sequence.

Example 1 above is a path. The length of this path is 3.

Example 2 is not a path. Since terminal vertices are the same and hence not open walk. Example 3 although is an open walk is not a path since vertex 5 is revisited for a second time in the walk.

Notice that walks and paths are sub-graphs. In a path all vertices except the beginning and ending are of degree two. The terminal vertices are of degree one.

Circuit

A closed walk in which no vertex repeats more than once is called a circuit. In other words circuit is a closed non-intersecting walk.

For example referring to the graph of six vertices in figure 2.4,

1, (1,2), 2, (2,5), 5, (5,3), 3, (3,1), 1 is a circuit of 4 vertices and 4 edges.

5, (5,6), 6, (6,4), 4, (4,3), 3, (3,1), 1, (1,2), 2, (2,5), 5 is another circuit of length 6

Observe that the degree of every vertex in a circuit is two. Other names for a circuit are cycle, elementary cycle, circular path, loop. Note that every self loop is a circuit.

The definitions in this section are summarized in the figure 2.7 below. The arrows are in the direction of increasing restriction.

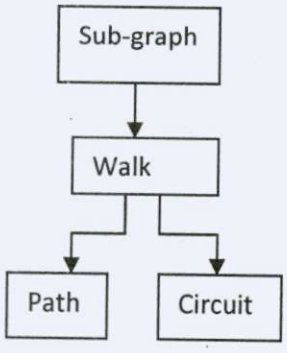


Fig 2.7: Walks, paths and circuits as sub-graphs

2.4 Connected graphs and components

The concept of connectedness is obvious. A graph is connected if we can reach any vertex from any other by travelling along the edges. A formal definition of connectedness is as follows:

A graph is said to be connected if there is at least one path between every pair of vertices. Otherwise the graph is disconnected. The graphs in figures 2.8 and 2.9 are examples of connected and disconnected graphs.

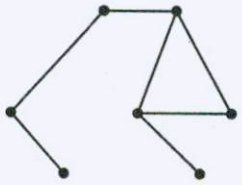


Fig 2.8: 7 vertex connected graph

Fig 2.9: 7 vertex disconnected graph

Disconnected graphs consist of two or more connected sub-graphs. The graph in figure 2.9 has two connected sub-graphs. Each of these connected sub-graphs is called a component. An easy way to find a component is to find all vertices that are reachable from a vertex v_i. Vertex v_i and all the vertices of the graph that have paths to v_i, together with all the edges incident on them form a component. It is evident that a component itself is a graph. To be precise, a component is a sub-graph of the given one. We now discuss some important theorems on connectivity of a graph.

Theorem 2.1

A graph G is disconnected if and only if its vertex set V can be partitioned into two nonempty, disjoint subsets V1 and V2 such that there is no edge connecting a vertex in V1 to any vertex in V2.

Proof:

If: Assume a partition as described exists. We need to show that the graph is disconnected.

Consider two vertices x and y, where $x \in$ V1 and $y \in$ V2. Let us examine if there can be a path from x to y. Note that a path from x to y requires connection between a vertex in V1 and some vertex in V2. By assumption there is no such connection. Hence it is evident that there is no path from x to y. That is the graph G is disconnected.

Only if: Assume G is a disconnected graph. We need to prove that the described partition exists in the graph.

Consider a vertex x in G. Let V1 be the set of vertices can be reached from x. Since G is disconnected V1 does not include all vertices (follows from the definition of connectedness). The remaining set of vertices will form a set V2. No vertex in V1 is joined to any vertex in V2. Thus the partition V1 and V2 is found.

Theorem 2.2

If a graph (connected or disconnected) has exactly two vertices x and y of odd degree, there must be a path joining these two vertices.

Proof:

Let G be graph with all even vertices except the odd vertices x and y. From theorem 1.2, which is true for every graph and therefore for any component, no graph can have odd number of odd degree vertices. Therefore in G, x and y must belong to the same component and hence there should be a path between them.

Theorem 2.3

A simple graph (a graph having no parallel edges or self loops) G with n vertices and k components can have at most (n-k)(n-k+1)/2 edges.

Proof:

Let $n_1, n_2, \ldots n_k$ be the number of vertices in the k components. Then $n_1+n_2+\ldots n_k=n$. The proof of the theorem depends on the following algebraic inequality.

$$\Sigma n_i^2 \leq n^2 - (k-1)(2n-k) \qquad \text{---------- (1)}$$

From the theorem 1.3, we know that maximum number of edges in the i^{th} component is $n_i(n_i-1)/2$. Therefore the maximum number of edges in G is

$$\begin{aligned} \Sigma n_i(n_i-1)/2 &= \Sigma(n_i^2)/2 - n/2 \\ &\leq [n^2 - (k-1)(2n-k)]/2 - n/2 \qquad \text{(using inequality (1) above)} \\ &= (n-k)(n-k+1)/2 \end{aligned}$$

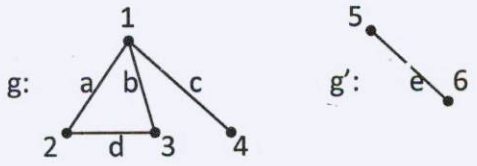
Observe that this theorem is generalization of theorem 1.3.

We close this section after a discussion of matrix form of disconnected graphs.

If a graph G is disconnected (with no common vertex and common edge) with two components g and g′ then A(G) will be in the form

$$\begin{pmatrix} A(g) & 0 \\ 0 & A(g') \end{pmatrix}$$

Example: Consider the disconnected graph G with components g and g′ below.



The incidence matrix is given by,

$$\begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array} \begin{array}{c} \begin{array}{cccc} a & b & c & d \end{array} \\ \left(\begin{array}{cccc:c} 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right) \end{array} = \begin{pmatrix} A(g) & 0 \\ 0 & A(g') \end{pmatrix}$$

2.5 Euler graphs

Euler is the founder of Graph theory. Euler solved the bridge problem, which was when Graph theory became a subject of study. In fact in that paper, on the problem about the bridge, he posed a more general problem. His problem goes as follows:

In what type of graph G is it possible to find a closed walk running through every edge exactly once?

Such a walk is now called Euler line (Euler walk). The graph which contains an Euler line is called an Euler graph. In other words if some closed walk in a graph contains all edges of the graph, then the walk is called an Euler line and the graph an Euler graph. By definition walk is always connected. Since Euler line contains all edges of the graph, an Euler graph is always connected, except for any isolated vertex the graph may have. As isolated vertices do not

contribute anything to concept of Euler line or graph, henceforth we assume that Euler graphs are always connected.

Next we prove an important theorem which will enable us to conclude if the graph is Euler graph or not.

Theorem 2.4

A given connected graph G is an Euler graph if and only if all vertices of G are of even degree.

Proof:

If: Suppose that G is an Euler graph. G contains an Euler line. When the walk hits a vertex it goes through two new edges; one we traversed to reach the vertex and the other through which we exit through. In this process we encounter two new edges incident on a vertex, each time we pass through. This is also true of terminal vertices. Thus every vertex is of even degree.

Only if: Assume all vertices of G are of even degree. Start the walk from an arbitrary vertex v. Go to a neighboring vertex. Since every vertex is even, when we enter a new vertex x there is an edge to exit from this vertex x. When you get back to v you have completed a closed walk h. See if all edges are traversed. If so the graph G is Euler graph. If not, remove from G all those edges which are part of h and obtain a sub-graph G'. Since G and h have all vertices with even degrees, vertices of G' are also of even degree. Moreover h and G' must have a common vertex, since G is connected. Find a closed walk j in G' starting from a vertex w (this is possible since vertices of G' are all even). Construct a new walk combining h and w. This walk has more edges than h or j. If all edges are covered then G is Euler. If not the above process is repeated until we obtain a closed that traverses all edges of G. Thus G is an Euler graph.

Example: Follow the arrows to find a walk in the graphs below.

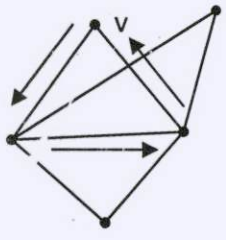


Fig 2.10(a): Closed walk h in G starting at v

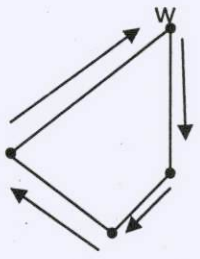


Fig 2.10(b): Sub-graph G' and the closed walk j starting at w

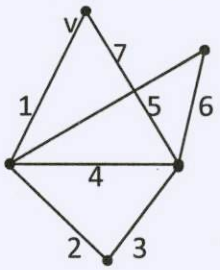


Fig 2.10(c): Combining h and j. All edges covered. Numbers indicate order of traversal of edges

Now coming back to Königsberg bridge problem recall figure 1.3, which is the graph corresponding to the problem. Observe that degrees of vertices C and D (5 and 3 respectively) are odd. Hence the graph is not Euler. Therefore a closed walk covering all edges does not exist.

One often encounters Euler graphs in various puzzles. The problem common to these puzzles is to find how to draw a picture in one continuous line without retracing and without lifting the pencil from the paper.

In defining Euler line some authors drop the requirement that the walk be closed. For example the traversal 4, (4,6), 6, (6,5), 5, (5,4), 4, (4,3), 3, (3,5), 5, (5,2), 2, (2,1), 1, (1,3), 3 in the graph of figure 2.4 includes all edges just once but starting and ending vertices are different. This kind of walk is called open Euler line or unicursal line. A connected graph that has a unicursal line is called unicursal graph. It is clear that by adding an edge between starting and ending vertices of unicursal line we get an Euler line. Thus a connected graph is unicursal if and only if there are exactly two vertices that are of odd degree. The generalization of this statement is the following theorem. An interested reader can refer the text by Narsingh Deo for a detailed proof.

Theorem 2.5

In a connected graph G with exactly 2k odd vertices, there exist k edge disjoint sub-graphs such that they together contain all edges of G and that each is a unicursal graph.

2.6 Hamiltonian circuits and paths

An Euler line is characterized by closed walk covering all edges exactly once. Hamiltonian circuit is a closed walk going through each vertex exactly once except that first and last vertex in the tour is the same (recall definition of closed walk).

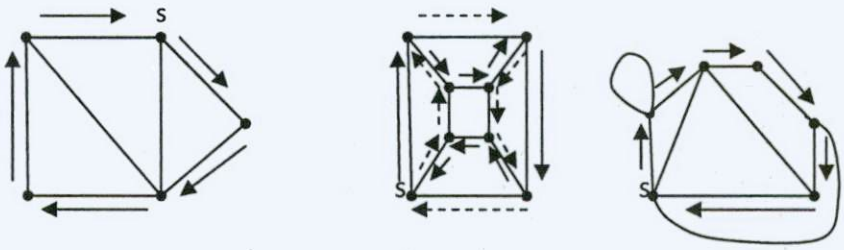


Fig 2.11: Hamiltonian circuits

The three graphs in the figure 2.11 are with vertices 5, 8 and 6 in number. s is the starting vertex of the Hamiltonian circuit in each graph. Follow the arrow marks for the Hamiltonian circuits in these graphs. Recall the definition of circuit in section 3. It is a closed walk where no vertex repeats. In addition if the circuit includes every vertex it is Hamiltonian circuit. A Hamiltonian circuit of n vertex graph consists of n edges. A graph can have many Hamiltonian circuits. For example second graph in figure 2.11 has one more circuit indicated in dashed arrows.

Not every graph will have Hamiltonian circuit. Two graphs (with vertices numbered) in the figure next (figure 2.12) do not have Hamiltonian circuit.

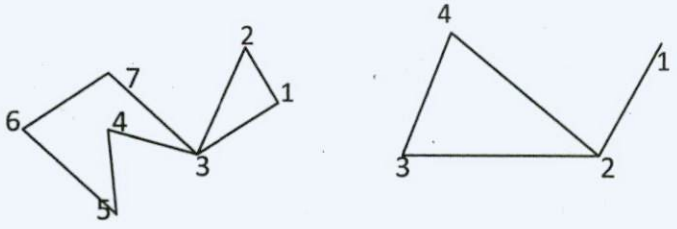


Fig 2.12: Graphs with no Hamiltonian circuit

Hamiltonian circuits and Euler lines are different. Hamiltonian circuit is much more complex. The problem of necessary and sufficient condition in a graph for the presence of Hamiltonian circuit is still unsolved. This problem was first posed by Sir William Rowan Hamilton and hence the name Hamiltonian circuit.

Hamiltonian path:

If we remove the last edge from a Hamiltonian circuit, we get a Hamiltonian path. Hamiltonian path is sub-graph of Hamiltonian circuit, every graph that has Hamiltonian circuit should have a Hamiltonian path. However not vice versa. That is there may be graphs with Hamiltonian paths but yet no Hamiltonian circuits. For example both the graphs in figure 2.12 have Hamiltonian paths. Follow the vertices beginning from 1 in the increasing order to get the Hamiltonian paths. These are 1 to 2 to 3 to 4 to 5 to 6 to 7 and 1 to 2 to 3 to 4 respectively. Note that length of Hamiltonian path is n-1 in an n vertex graph.

In considering existence of Hamiltonian circuits or paths we need only consider simple graphs. This is because a Hamiltonian circuit or path traverses each vertex only once. Hence it cannot include parallel edges or self loops. Thus it may be sensible to remove parallel edges and self loops before looking for a Hamiltonian circuit. Finally not all graphs have Hamiltonian paths. Graphs below in figure 2.13 are examples of graphs that do not have Hamiltonian paths.

Fig 2.13: Graphs not having Hamiltonian paths

Complete graph was discussed in unit 1. It is nothing but a simple graph with edges connecting every pair of vertices. This is also referred to as universal graph or a clique. Every vertex is joined to every other vertex in a complete graph. Hence the degree of every vertex in complete graph of n vertices is n-1. Also the number of edges in a complete graph of n vertices is n(n-1)/2 (refer theorem 1.3).

It is easy to construct a Hamiltonian circuit in a complete graph. Let the vertices be numbered 1 to n. Traverse the vertices in the order 1 to 2 to 3 … n-1 to n to 1. A graph may contain many Hamiltonian circuits. Number of edge disjoint Hamiltonian circuits is an unsolved problem. However something can be said about this number in some graphs.

Theorem 2.6

In a complete graph with n vertices there are (n-1)/2 edge-disjoint Hamiltonian circuits, if n is odd and ≥ 3.

Proof:

In a complete graph of n vertices there are n(n-1)/2 edges. A Hamiltonian circuit in an n vertex graph has n edges. Hence number of edge-disjoint Hamiltonian circuits cannot exceed (n-1)/2. That there are (n-1)/2 edge disjoint Hamiltonian circuits if n is odd can be shown as follows:

Here we propose an informal proof for the above statement. Let us find the number of Hamiltonian circuits in complete graphs of 3, 5, 7 vertices. Later we conclude by extrapolation for n vertices.

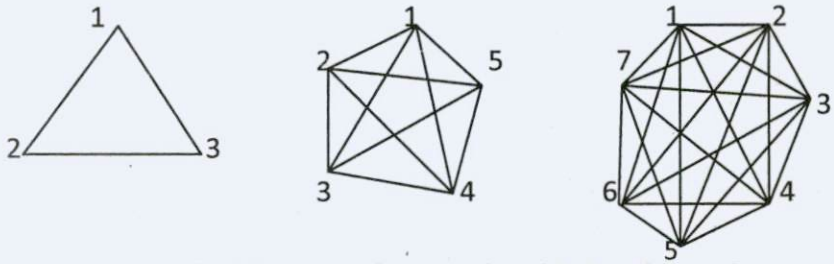


Fig 2.14: Complete graphs of 3, 5 and 7 vertices

The edge-disjoint Hamiltonian cycles in these graphs are:

3 vertex graph: 1, (1,2), 2, (2,3), 3, (3,1), 1 (only one [(3-1)/2] Hamiltonian cycle)

5 vertex graph:

1, (1,2), 2, (2,3), 3, (3,4), 4, (4,5), 5 (5,1), 1;

1, (1,3), 3, (3, 5), 5, (5,2), 2, (2,4), 4, (4,1), 1 (two [(5-1)/2] cycles

1, (1,2), 2, (2,3), 3,(3,4), 4, (4,5), 5, (5,6), 6, (6,7), 7, (7,1),1;

1, (1,3), 3, (3,5), 5, (5,7), 7, (7,2), 2, (2,4), 4, (4, 6), 6, (6,1), 1;

1, (1,4), 4, (4,7), 7, (7,3), 3, (3, 6), 6, (6,2), 2, (2, 5), 5, (5, 1),1 (three [(7-1)/2] cycles)

Hence number of edge disjoint cycles in an n vertex complete graph is (n-1)/2, if n is odd and greater than 1.

Travelling salesman problem:

A problem closely related to the question of Hamiltonian circuits is travelling salesman problem, stated as follows: A salesman is required to visit a number of cities during his trip. Given the

distances between the cities, in what order should he travel so that he visits each city exactly once and return home, with the minimum distance travelled?

Representing the cities by vertices and the roads between cities as edges, we get a graph. In this graph, edges are weighted, weights being distances between cities. In our problem, if each of the cities has a road to every other city, we have a complete graph and there are numerous Hamiltonian circuits. The tour of the salesman is after all a Hamiltonian circuit beginning at starting city and ending at the same city. We are to pick that cycle whose sum of distances is minimum. The total number of Hamiltonian cycles (not necessarily edge disjoint ones) in complete graph is (n-1)!/2. This follows from the fact that there are n-1 choices of cities at first (starting) city, n-2 at the second city and so on. These being independent choices, we get (n-1)! possible number of choices. This number is to be divided by 2 since each cycle is counted twice. (Remember the graph is not directed. The cycle 1 to 2 to 3 … n to 1 is same as 1 to n to n-1 to … 3 to 2 to 1).

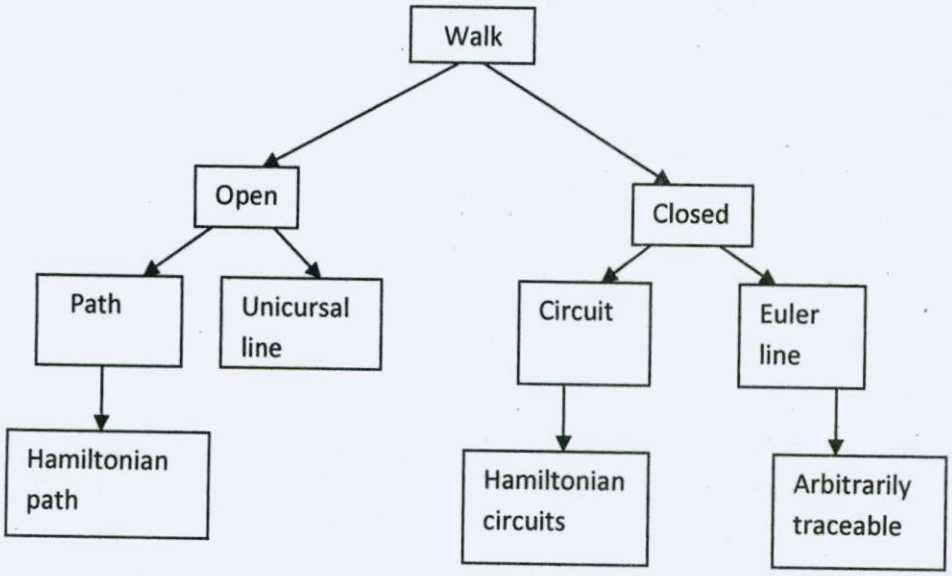
Theoretically the travelling salesman problem can be solved by finding all cycles and its distances and the minimum distance cycle can be chosen. However for a large value of n, it is too tedious to find all cycles. The problem is to find a manageable algorithm to find a solution in reasonable time. No such algorithms exist. This being a very important problem in Operations Research, many attempts have been made. There are some good heuristic algorithms available. These heuristic methods may not give us an optimal solution. But will definitely get one very close to optimal solution, called near optimal solutions.

2.7 Summary

In this unit we started with the concept of isomorphism. Isomorphic graphs are structure wise identical. That is edge-vertex relationship is preserved. There may be relabeling of vertices and the shape may be different. Later we learnt about sub-graphs and edge disjoint sub-graphs. Walks, paths circuits are discussed in detail with examples. We also discussed the concept of connectedness, components and dis-connectivity. Matrix representations are covered during discussions on isomorphism, sub-graphs and connectivity. We moved onto Euler graphs and derived conditions for a graph to be Euler graph. Finally we closed this section after providing a proof for the famous Königsberg bridge problem. A slightly different circuit (different from Euler line) is discussed in the following section. This circuit is Hamiltonian circuit (again named

after the person who posed the problem) and an important application, the travelling salesman problem is discussed in the end of that section.

Various types of walks discussed in this unit are classified as follows.



2.8 Key words

Isomorphism, sub-graph, walk, path, circuit, connected graphs, Euler graphs, Hamiltonian circuit, travelling salesman problem

2.9 Practice problems

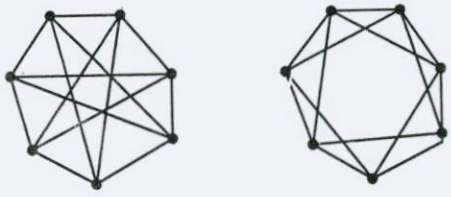
Easy problems: Based on the notes of this unit

1. Draw graphs of 4 and 5 vertices that are isomorphic.
2. State the conditions for graphs to be isomorphic.
3. Draw graphs that satisfy all 3 conditions of isomorphism and yet not isomorphic. State the reasons why they are not isomorphic.
4. Define sub-graph and provide examples.
5. Draw a graphs and sub-graphs that are edge disjoint and that are vertex disjoint.
6. Define walks, paths, circuits. Illustrate.
7. Illustrate closed and open walk.
8. Distinguish circuit and walk. Give examples.
9. Comment about the degrees of vertices in walks, paths and circuits.

10. Define connected graph. Give examples of connected and disconnected graphs with various numbers of vertices.
11. What is component?
12. State and prove the necessary and sufficient conditions for a graph to be dis-connected.
13. If a graph has exactly two vertices that of odd degree, prove that there must be path connecting these two.
14. What is the maximum number of edges in a simple graph of n vertices and k components? Justify your answer. Also draw some graphs and verify.
15. What is Euler graph? Draw one that is Euler and one that is not.
16. State and prove the theorem on Euler graphs.
17. What is Hamiltonian circuit? Why are they called so?
18. Distinguish Hamiltonian circuit and Euler line. Give examples.
19. Define unicursal line. Illustrate.
20. Comment on the maximum number of edge-disjoint Hamiltonian circuits. Verify your statement in some graphs.
21. What do you mean by Hamiltonian path? Give an example.
22. What can you say about presence of Hamiltonian circuit and path?
23. State the problem of travelling salesman. How is this similar to Hamiltonian circuit?
24. Why is travelling salesman problem difficult to solve by enumeration?

Challenging problems: Consult the reference texts and solve.

1. Show that the graphs below are isomorphic.



2. Construct an example to show the conditions given in section 2.1 is not sufficient for graphs to be isomorphic.
3. Prove that connected graphs with every vertex of degree two are isomorphic.

4. Prove that a simple graph with n vertices with more than (n-1)(n-2)/2 edges is connected. (Hint: use theorem 2.3)
5. A connected graph remains connected even after removing edge e if and only if e is in a circuit.
6. Draw a connected graph that becomes disconnected when a single edge is removed.
7. Label vertices and edges in the graphs of problem 1. List 3 paths and 3 circuits of various lengths.
8. Can there be path longer than Hamiltonian path (if it exists) in a graph? Justify.
9. Draw two 5 vertex graphs one which has Hamiltonian path but not a circuit and another which has both.

2.10 References

1. Narsingh Deo, Graph Theory with applications to Engineering and Computer Science, PHI, Chapter 2 (Sections 2.1, 2.2, 2.4, 2.5, 2.6, 2.9, 2.10)
2. J.P.Tremblay, R.Manohar, Discrete Mathematical Structures with applications to Computer Science, TATA McGRAW-HILL, Chapter 5 (Section 5.1.2)
3. Dr. N.G.Goudru, Discrete Mathematical Structures, Himalaya Publishing House, chapter 7 (Section 7.2 to 7.6)
4. Bernard Kolman, Robert C. Busby, Sharon Cutler Ross, Discrete Mathematical Structures, PEARSON Education, Chapter 8 (Section 8.1)

Unit 3: Trees

Structure

3.0 Objectives

When you have learnt the concepts discussed in this unit it will be possible for you to

- Define a tree and understand the significance of trees in Computer Science
- Understand important properties of trees
- Identify distance, radius and center of a tree
- Explain rooted trees and path lengths
- Work on binary search tree, which is a very important concept in Computer Science.

3.1 Definition and applications

Tree is a special graph. The concept of tree is very important in graph theory and also in many applications of graph theory. Also tree is an important data structure in computer science.

Definition

Tree is a connected graph without any circuits. The graphs in figure 3.1 are all trees with varying number of vertices.

Fig 3.1: Trees with different number of vertices

Parallel edges and self loops form circuits. Hence, it is obvious that trees are simple graphs. Trees can be infinite as graphs. But our discussions focus on finite trees only.

Applications

Trees are useful in describing any structure which involves hierarchy. Familiar examples of such trees are family trees, decimal classification of books in a library, the hierarchy of positions in an organization, an algebraic expression involving operations which comes with precedence, sorting of mails according to PIN code etc. Figure 3.2 shows mail sorting process as a tree diagram. All mails arrive at a local office N. PIN codes of mails are read at N. Mails are first sorted using the most significant digit in the PIN code and are divided into 10 piles N0, N1, ...N9. Each pile is

then divided into 10 piles and this continues for 4 more times (the number of digits in a PIN code is 6).

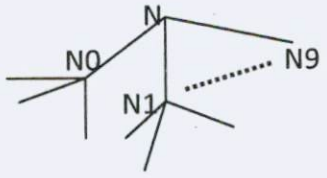


Fig 3.2: Mail sorting process

3.2: Properties of trees

We state some interesting properties of trees in this section and provide a simplified proof for them.

Theorem 3.1

There is only one path between every pair of vertices in a tree T.

Proof:

Since T is a connected graph there is a path between every pair of vertices. If the number of paths between two vertices x and y are two, then there is a circuit. But T being a tree cannot have a circuit. Hence there must be one and only path between all pairs of vertices.

Theorem 3.2 (Converse of theorem 3.1)

If in a graph there is one and only path between every pair of vertices then the graph is a tree.

Proof:

There is a path between all pairs of vertices. So the graph is connected. Because there is only one path between all pairs of vertices, there is no circuit. So the graph is tree.

Theorem 3.3

A tree with n vertices has n-1 edges.

Proof:

We use induction in the proof process. It is easy to see that the theorem is true for 1, 2 and 3 vertices. Refer figure 3.1 for various such trees. Assume that theorem holds for n-1 vertices. That is a tree T of n-1 vertices has n-2 edges. Consider a tree T' of n vertices. Suppose that the new

and n^{th} vertex is v_n. Because T' is connected there should be an edge e_k from some v_i in the tree T of n-1 vertices to v_n. If e_k is removed then the graph T' becomes disconnected (one component is T and the other is vertex v_n). Otherwise it means that there is an alternate path from v_i to v_n, which is not possible. Therefore the number of edges in the connected tree T' is n-2 (those in T) plus the 1 (edge e_k is a new edge). This number is n-1. Hence the theorem is proved.

Following three theorems are stated without proof. Interested reader can refer to the text by Narsingh Deo in the reference list for a formal proof.

Theorem 3.4: Any connected graph with n vertices and n-1 edges is a tree.
Theorem 3.5: A graph is a tree if and only if it is minimally connected (removal of one edge will disconnect the graph).
Theorem 3.6: A graph with n vertices and n-1 edges, and no circuits is a tree.

The results of previous 6 theorems are summarized as follows:
A graph G with n vertices is called a tree if

1. G is connected and is circuitless, or
2. G is connected and has n-1 edges, or
3. G is circuitless and has n-1 edges, or
4. There is exactly one path between every pair of vertices, or
5. G is minimally connected graph.

Another interesting property we state here again without proof.
A tree (of two or more vertices) has at least two pendant vertices.

3.3 Distance and centers

You will definitely agree that vertex b is center of the tree in figure 3.3 below.

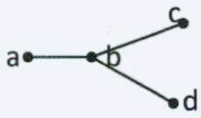


Fig 3.3: Tree with center b

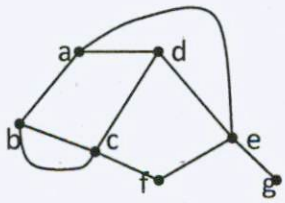


Fig 3.4: Graph G

This is intuitive since b is centrally located. Inherent in the concept of center is the concept of distance. Distance between vertices is the minimum path length between those vertices. Refer the graph G in the figure 3.4 above. There are at least four paths from vertex a to vertex f: (i) a,d,e,f (ii) a,d,c,f (iii) a,b,c,f (iv) a,e,f (note that in these paths only vertices are listed). The minimum path length is 2. Hence the distance between these vertices is 2.

Another concept in the understanding of center is the eccentricity of a vertex. Eccentricity of a vertex v denoted by E(v), is the distance from v to the farthest vertex in G. For example in the graph of figure 3.4, E(a) = 2 and E(g) = 3.

Finally a vertex v with minimum eccentricity is said to be a center of the graph. Going back to the tree of figure 3.3, E(a)=E(c)=E(d)=2 and E(b)=1. Hence b is the center. Given in figure 3.5 are more trees and centers.

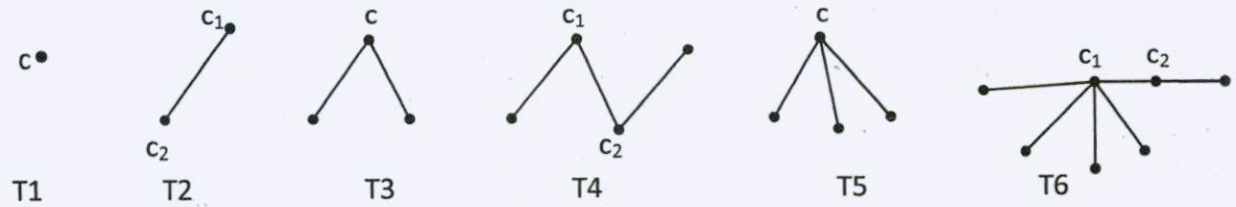


Fig 3.5: Trees and centers denoted by c(s)

Note that trees can have one or two centers. Trees T4 and T6 in the figure above has two centers. Also if a tree has two centers they are adjacent.

We close this section pointing out an application of tree centers. Suppose that the communication among a group of 7 persons is represented by the last graph in the figure 3.5. Vertex in the figure

represents person and edge represents communication link. If closeness to members is required to lead the group, who could be the leaders? Obviously the persons at the centers are leaders, because they are closest to team members.

Radius of a tree is defined as eccentricity of the center. In the graphs of figure 3.5 above, radius of all trees (excluding single vertex tree T1) T2 to T5 is 1 and that of T6 is 2.
Diameter is the length of the longest path in the tree. For example, the trees in the figure 3.5 (excluding single vertex tree T1) have diameters 1, 2, 3, 1 and 3 respectively.

3.4 Rooted trees and binary trees

A tree in which one vertex is distinguished from all others is called rooted tree. The root of the tree is marked differently than other vertices. Trees without roots are referred to as trees, non-rooted trees or free trees. Given below in figure 3.6 some rooted trees.

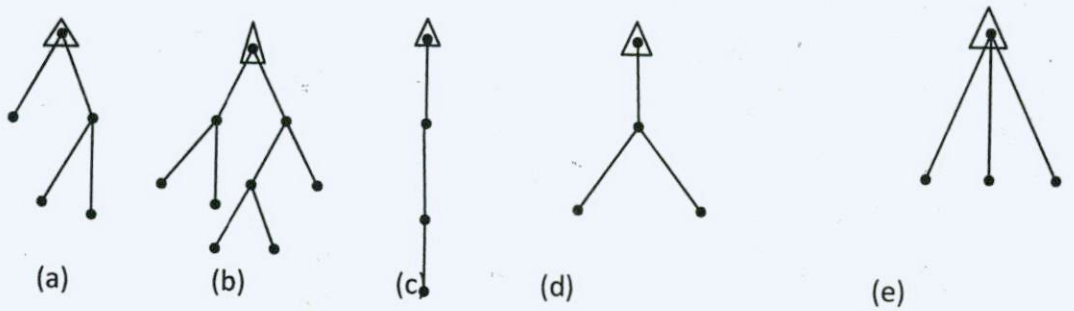


Fig 3.6: Examples of rooted trees (roots enclosed in a triangle)

Binary tree is one in which exactly one vertex is of degree 2 and the rest are of degree 3 or 1. Evidently we are referring to trees of vertices 3 or more. Since the vertex of degree is unique, it is called root of the binary tree. Binary tree has lot of applications in computer based search methods. In the next section we will discuss one such method. In figure 3.6 above, first two trees are binary trees. The following observations are easy to derive in a binary tree:

1. The number of vertices in a binary tree is always odd. This is because there is exactly one vertex is of even degree and remaining vertices are of odd degree. Recall a result in unit 1 which says that "the number of vertices of odd degree is even". With this even number of

vertices, if we add the root of the tree, we get odd number of vertices. In the binary trees in figure 3.6 (a) and 3.6 (b), the numbers of vertices are 5 and 9 (both odd).

2. Suppose that we have a binary tree of n vertices and p be the number of pendant vertices. Then n-p-1 is the number of vertices of degree three. Hence the number of edges is [p+3(n-p-1)+2]/2 (recall the result "number of edges is half of sum of the degrees of all vertices").

   Number of edges = [p+3(n-p-1)+2]/2 = n-1 $\Rightarrow$ p=(n+1)/2

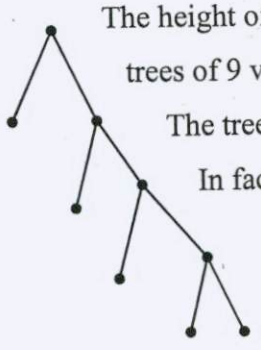
   Refer tree in figure 3.6 (a). Here n=5, p=3 [(5+1)/2], number of vertices of degree three = 1 [5-3-1= (n-p-1)].

   Refer tree in figure 3.6 (b). Here n=9, p=5= (9+1)/2, number of vertices of degree three = 3 = 9-5-1 (n-p-1).

3. A non pendant vertex is called internal vertex. The number of internal vertices is n-p = n-(n+1)/2 = (n-1)/2 = (n+1)/2 -1 = p-1. The level of a vertex in binary tree is the distance of the vertex from the root. For example in the binary tree in figure 3.6 (a), the number of vertices in level 1 is 2 and the number of vertices in level 2 is 2. In the binary tree next, the number of vertices in levels 1, 2 and 3 are 2, 4 and 2.

4. The maximum level of any vertex is called height of the tree denoted by l_{max}. It can be shown that minimum height of n vertex binary tree is (proof beyond the scope of the discussion) $\min(l_{max}) = \lceil \log_2(n+1) - 1 \rceil$, where $\lceil x \rceil$ is the smallest integer $\geq x$.

   Let us go back to binary tree in figure 3.6 (b). n=9, n+1=10, $\lceil \log_2(n+1) - 1 \rceil = 3$ and l_{max} is 3 for that tree in the figure and this is min of l_{max}, 9 vertex binary trees. Another 9 vertex binary tree is shown below.



The height of the tree to the left is 4. Construct other binary trees of 9 vertices and verify that $\min(l_{max}) = 3$.

The tree on the left has maximum possible levels.

In fact max $(l_{max}) = (n-1)/2$ (= 4 for 9 vertex binary tree)

5. In many algorithms we are interested in computing the sum of the levels of all pendant vertices. This is called path length of the tree. Going back to example trees in figure 3.6 (a) and 3.6 (b) the path lengths of the two binary trees are 1+2+2=5 and 2+2+2+3+3 = 12. The importance of path length is that it is often related to execution time of an algorithm. The tree above has greater path length (1+1+2+2+3+3+4+4=20) when compared to tree in figure 3.6 (b).

3.5 Binary search tree

Here we learn a very efficient way of searching an element in an array. Searching is a fundamental operation and very common task in Computer Science. Many higher level tasks perform searching routinely and repeatedly. Hence searching algorithms should be as efficient as possible.

The straight forward procedure to search for a given element is to compare the given element and the array element one at a time. Search is halted as soon as given element is found in the array, in which case success reported. Search can also come to halt when there are no more elements in the array to compare the given element with, which means it is the case of failure (to find the given element in the array). This procedure is called 'linear search' and requires no preprocessing of array elements. But it takes time. This is not good especially when you have to search often.

There is an alternate procedure called 'binary search' which is efficient. But this requires elements to be arranged in an order (increasing or decreasing). This preprocessing time is small, compared to the total time spent on many executions of 'linear search' procedure. Hence this is often used in many implementations.

Binary search is also not very useful, when elements are inserted and deleted often in data set. A different data structure called 'binary search tree' is used for this case of frequent updating of the data set. This is also good if we want to know the smallest or biggest from time to time.

We first discuss the data structure 'binary search tree' and then the associated procedures to insert, delete and finding maximum/minimum. Each node in the tree has a label. The labels are nothing but elements of the set.

Definition

A binary search tree for a set of elements is such that, for each node v, elements in the left sub-tree are less than v and elements in the right sub-tree are greater than v.

Examples

1. Suppose the set A= {3, 7, 1, 4, 8}. Two binary search trees for the set is given below in figure 3.7

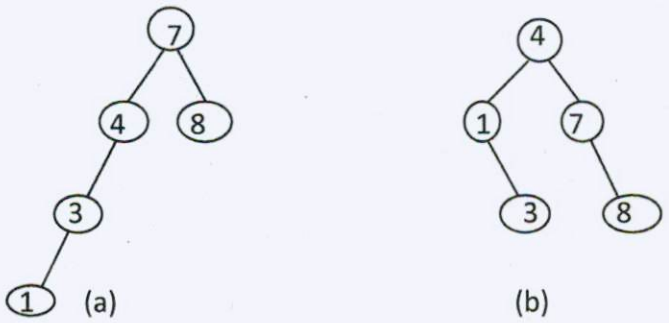


Fig 3.7: Two binary search trees for the set A

l_{max} is 3 and 2 for the trees (a) and (b)

2. Suppose the set B= {begin, else, end, if, then}. Binary search trees for B could be the ones in figure 3.8

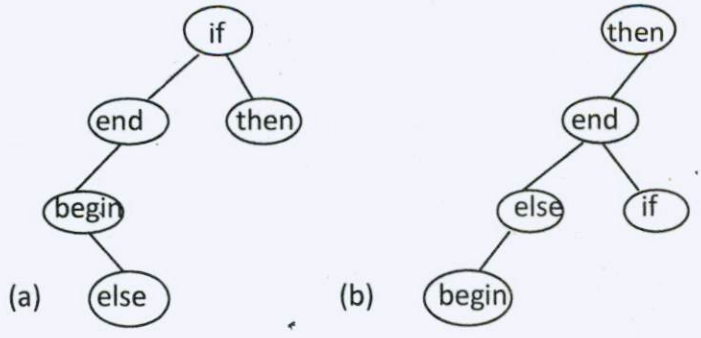


Fig 3.8: Two binary search trees for the set B

l_{max} is 3 for both trees

Searching: To determine if an element x is in the set or not, we first compare x with root. If x is root return success and halt. If not, compare x and root. If x < root go to left sub-tree and if x >

root go to right sub-tree. Repeat the procedure above, until you find the element or there are no more nodes in the tree.

Let us execute search with the tree in the figure 3.7 (a). Let x= 5.

1. Compare x and 7 (root). x is smaller. Move to left sub-tree.
2. Compare x and 4. x is bigger. Move to right sub-tree.
3. Right sub-tree is empty. Report failure and halt.

Let search element x=3.

1. Compare x and 7. x is smaller. Move to left sub-tree.
2. Compare x and 4. x is smaller. Move to left sub-tree.
3. Compare x and 3. Equal. Report success and halt.

Now let us trace the execution in the binary search tree of figure 3.7 (b) for x=5.

1. Compare x and 4. x is bigger. Move to right sub-tree.
2. Compare x and 7. x is smaller. Move to left sub-tree.
3. Left sub-tree empty. Report failure and halt.

When x=3, the execution is as follows:

1. Compare x and 4. x is smaller. Move to left sub-tree.
2. Compare x and 1. x is bigger. Move to right sub-tree.
3. Compare x and 3. Equal. Report success and halt.

Let us execute search with the tree in the figure 3.8 (a). Let x= repeat.

1. Compare x and if (root). x is bigger. Move to right sub-tree.
2. Compare x and then. x is smaller. Move to left sub-tree.
3. Left sub-tree is empty. Report failure and halt.

Let search element x be else

1. Compare x and if. x is smaller. Move to left sub-tree.
2. Compare x and end. x is smaller. Move to left sub-tree.

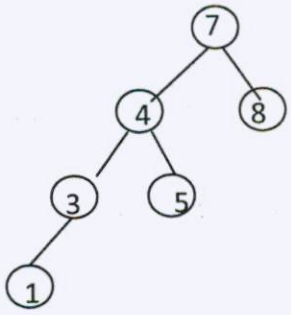
3. Compare x and begin. x is bigger. Move to right sub-tree
4. x is equal to else. Report success and halt.

Insertion: Start with current node as root. If an element x is to be inserted traverse the tree using comparisons with current node and insert when the sub-tree at any time is empty.

Execution of insertion with x=5 in figure 3.7 (a)

1. Compare x and 7 (root). x is smaller. Move to left sub-tree.
2. Compare x and 4. x is bigger. Move to right sub-tree.
3. Right sub-tree is empty. Inset 5 as right child of node 4. Halt.

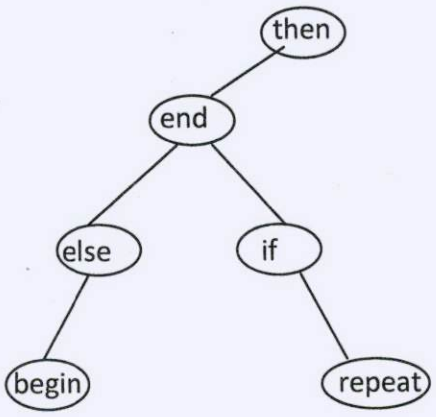
The tree after insertion is as follows:



Execution of insertion with x=repeat in figure 3.8 (b)

1. Compare x and then (root). x is smaller. Move to left sub-tree.
2. Compare x and end. x is bigger. Move to right sub-tree.
3. Compare x and if. x is bigger. Move to right sub-tree.
4. Right sub-tree is empty. Insert repeat as right child of if. Halt.

The tree after insertion is as follows:



Deletion: Locate the element x in the tree. If it is the leaf node, remove it from tree. If x is at root or any other position then the largest element in the left sub-tree is located and placed at the position of x.

Let us trace how deletion works when x=3 is deleted from tree 3.7 (b)

1. Locate x. This is done by comparisons with various nodes in the tree.
2. It is at leaf. Remove the node.

Let us see how 7 is deleted.

1. Locate 7. This is done by comparisons with various nodes in the tree.
2. It is in the middle of the tree. Find the maximum of left sub-tree.
3. The left sub-tree is empty. Move the right sub-tree to position of node x.

Consider deletion of 4.

1. Locate 4. This is done by comparisons with various nodes in the tree.
2. It is at root. Find the maximum of left sub-tree.
3. It is node 3. Make this as root.

The tree after each of these deletions is shown below.

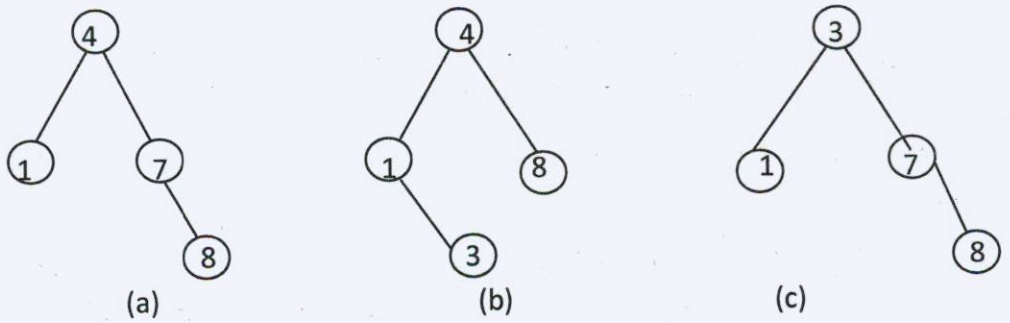


Fig 3.9: Tree 3.8 (b) after deletion of 3, 7, 4.

We conclude this unit after discussing the operation of locating maximum or minimum of set elements. To find maximum, keep moving right until right-sub-tree is empty. The last node visited is the maximum. Similarly to find minimum, keep moving left until left sub-tree is empty. The last node visited is the minimum.

For example, consider the tree of figure 3.9 (a). To find maximum, start from root 4 and visit the nodes on right namely, 7 and 8. The right sub-tree is empty below 8. The last node visited namely 8 is maximum. If in tree 3.9 (b) minimum is to be found, start from root 4, move left to 1 and find the left sub-tree to be empty. The last node visited is 1 and this is the minimum.

3.6 Summary

In this unit we learnt about trees, starting from definition, applications and some key properties of trees. The concept of center, radius is introduced after distance and eccentricity is defined. Later we described in length rooted trees and binary trees. Searching being an important operation in Computer Science, we studied in detail the concept of binary search tree. Extensive illustrations of all operations in binary search trees are also provided in the end of the unit.

3.7 Key words

Tree, centre and radius of a tree, rooted tree, binary tree, binary search tree

3.8 Practice problems

Easy to do: Based on the material

1. Define trees. Mention some applications of trees.
2. Draw trees of various numbers of vertices (ranging from 2 to 6) and shapes. For each tree determine center(s), radius and diameter.
3. For each tree you drew in problem 1, verify all the properties we discussed in this unit.
4. Explain binary trees. Draw some (with vertices ranging from 3 to 11) and verify all the results we stated on the binary trees.
5. Discuss some search methods.
6. Explain the importance of binary search tree.
7. What are the operations on a binary search tree? How do you execute these?
8. Demonstrate the operations on binary search trees.

Challenging: Requires additional study from the reference texts

1. How many trees of 5 vertices (non-isomorphic) can be drawn? Prove. Demonstrate.
2. Prove theorems 3.4, 3.5 and 3.6.
3. Prove that trees can have only one or two centers.
4. Show that a tree has at least two pendant vertices.

3.9 References

1. Narsingh Deo, Graph Theory with applications to Engineering and Computer Science, PHI, Chapter 3 (Sections 3.1, 3.2, 3.4, 3.5)
2. J.P.Tremblay, R.Manohar, Discrete Mathematical Structures with applications to Computer Science, TATA McGRAW-HILL, Chapter 5 (Section 5.1.4)
3. Dr. N.G.Goudru, Discrete Mathematical Structures, Himalaya Publishing House, chapter 7 (Section 7.9)
4. Bernard Kolman, Robert C. Busby, Sharon Cutler Ross, Discrete Mathematical structures, PEARSON Education, Chapter 7 (Section 7.1)

Unit 4: Spanning trees

Structure

4.0 Objectives

When you learnt the ideas described in this unit you will be able to

- To know the concept of spanning tree and some applications of spanning tree
- To understand weighted trees
- Use the algorithms to find minimum weight spanning tree

4.1 Spanning tree

In unit 3 we discussed graphs which are trees. In this unit we identify a tree in a graph.

Definition

Let G be a graph. A spanning tree is a tree connecting all vertices of G.

Examples

Given below in figure 4.1 are some graphs and spanning trees for each graph,

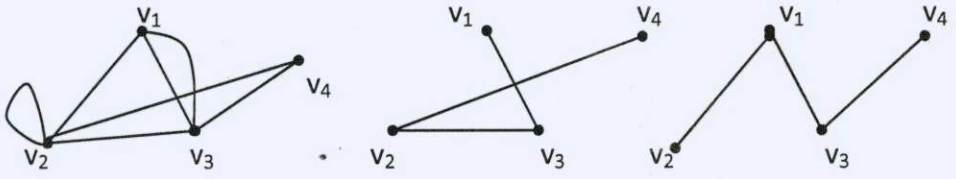


Fig 4.1(a): Graph Spanning tree T_1 Spanning tree T_2

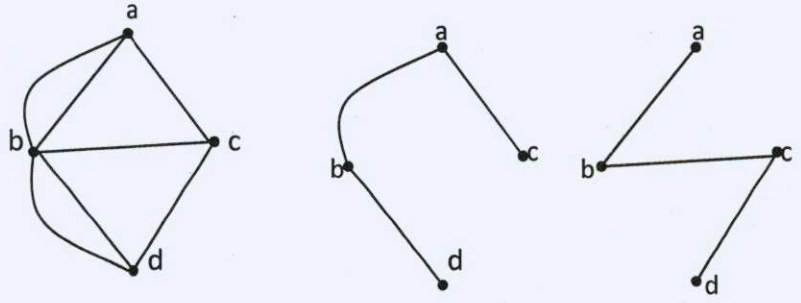


Fig 4.2 (b): Graph Spanning tree T_1 Spanning tree T_2

Finding a spanning tree:

Finding a spanning tree in graph is easy. If G has no circuit, then it is a spanning tree. If G has a circuit, delete an edge from the circuit. This will still leave the graph connected. If there are more circuits, repeat the above operation until there are no more circuits. As the resulting graph is connected and has no circuits, it is a tree and has all vertices. Thus a spanning tree in G is found.

Let us go back to the graph in figure 4.1 (a). Self loop at v_2 is a circuit. Remove this. Edges (v_1,v_2), (v_2, v_3) and (v_3, v_1) form a circuit. Remove an edge say (v_1,v_2) from this circuit. There is a set of parallel edges between v_2 and v_3. This is also a circuit. Remove one of the edges. Edges (v_4, v_2), (v_2, v_3) and (v_3, v_4) form a circuit. Remove an edge say (v_3,v_4) from this circuit. What we have is spanning tree T_1 in figure 4.1 (a).

We now discuss some elementary properties of a spanning tree.

Theorem 4.1

Every connected graph has at least one spanning tree.

This is evident from the previous discussion of finding a spanning tree in a graph.

An edge in a spanning tree is called as branch and the edges not in the spanning tree are called chords. Refer figure 4.1(a) and the spanning tree T_1. The branches of this spanning tree are (v_4, v_2), (v_2, v_3) (one of the parallel edges) and (v_3, v_2). Chords are self loop at v_2, (v_1, v_2), (v_3, v_4) and one of the parallel edges between v_1 and v_3. The number of vertices in the graph is 4. The number of edges is 7. The number of branches is 3 and the chords are 4 in number. The following theorem is a formal statement of these details.

Theorem 4.2

With respect to any of its spanning trees, a connected graph of n vertices and e edges has n-1 tree branches and e-n+1 chords.

This has been verified in figure 4.1 (a) in the previous paragraph.

4.2 Weighted trees

As discussed in the previous section a spanning tree is a minimally connected sub-graph of a graph. If there are real numbers associated with edges we call such a graph as weighted graph. We already came across weighted graphs in unit 1. A spanning tree of a weighted graph is a weighted tree. The weight of the tree is the sum of weights of all edges in the tree. We already discussed that a graph can have many spanning trees and each having different weights. In this section we give examples of weighted graphs and find some weighted spanning trees of the graphs. Figure 4.3 below has some weighted graphs and spanning trees with associated weights.

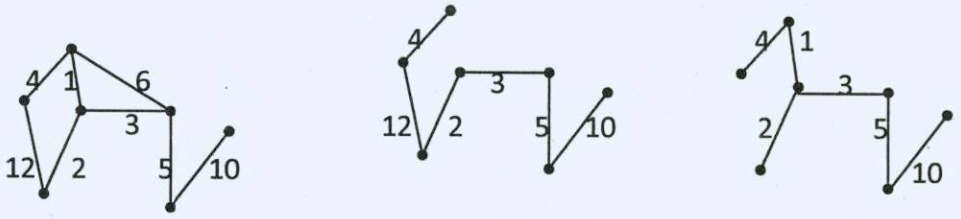


Fig 4.3 (a): Weighted graph (b): Spanning tree T_1 (c): Spanning tree T_2

The spanning tree in figure 4.3 (b) has weight 4+12+2+3+5+10 = 36 and the spanning tree in 4.3 (c) has weight 4+1+2+3+5+10 = 25. There may be many more spanning trees each with their own weight. A spanning tree with minimum weight is called minimum weight spanning tree or shortest spanning tree or shortest distance spanning tree or minimal spanning tree. Often this minimal spanning tree is of interest. One possible application of the shortest spanning tree is as follows: Suppose that we are to connect n cities 1, 2, 3, …, n through a network of roads. The cost of building road between city i and city j be c_{ij}. These are weights of the network of roads, which is our graph. The question is 'what is the minimum expense of connecting all cities by a network of roads?' This can be solved by finding minimal spanning tree. This sub-graph connects all cities and at the same time cost of this network is least.

4.3 Algorithms to find minimum weight spanning tree

Here we discuss two algorithms to find minimum weight spanning tree one proposed by Kruskal and the other by Prim.

Kruskal's method:

List all edges of the graph in the order of non-decreasing weights. Pick edges (one at a time from this list) corresponding to these weights and add them to the partial tree if it does not form a cycle with previously selected edges. Stop selection once you have collected n-1 edges (assuming graph has n vertices). The edges selected will form a spanning tree and has minimum weight.

Prim's method:

This algorithm does not require ordering of edges according to weights. Also a check on whether a cycle is formed is not needed, as proposed by Kruskal. Suppose that the graph has n vertices.

Prim's method requires you to tabulate the weights of edges in an n x n array (like adjacency matrix; but entries are not binary. They are weights of edges) called weight matrix. Diagonals in this matrix are blank. Note that even if there is self loop with least weight we won't select it in the tree because it forms cycle. So the diagonal elements in the weight matrix are left blank. Also it is evident that the matrix is symmetric. Set the weights of the non-existent edges to be infinity. Start from vertex 1 and connect to its nearest vertex say i. The vertices in the tree are 1 and i and the only edge is (1,i). Next find a vertex nearest to 1 or i, say j (assume (i,j) is smallest weights of all edges incident on 1 and i) and make a connection from i to j in the tree. Now, the tree has vertices 1,i and j and edges (1,i) and (i,j). As we always add a new vertex to the tree, checking for cycle formation is not needed. Continue this process until all vertices are included in the tree. What we have then is a minimum weight spanning tree.

Let us run (hand simulate) these algorithms on a graph given in figure 4.4.

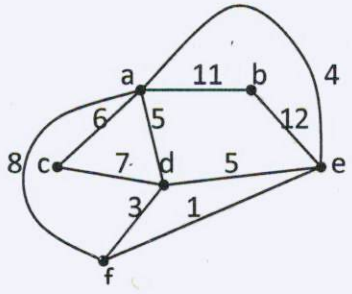


Fig 4.4: Graph G of 6 vertices a to f and 10 edges

Kruskal's method:

As a first step we order the edges according to their weights. The order is {(e,f), (d,f), (a,e), (a,d), (d,e), (a,c), (c,d), (a,f), (a,b), (b,e)} with weights {1, 3, 4, 5, 5, 6, 7, 8, 11,12}

The table below shows selection of edges from the list above that makes up the spanning tree. One edge from the list is added to the tree at a time.

Edge	Weight	Total weight so far
(e,f)	1	1
(d,f)	3	4
(a,e)	4	8
(a,d) – not selected.	-	8
(d,e)-not selected	-	8
(a,c)	6	14
(c,d)-not selected	-	14
(a,f)-not selected	-	14
(a,b)	11	25

Spanning tree is complete as five edges (the graph has 6 vertices) are selected already and its weight is 25. The edges in the spanning tree are (e,f), (d,f), (a,e), (a,c) and (a,b). The spanning tree is shown in figure 4.5.

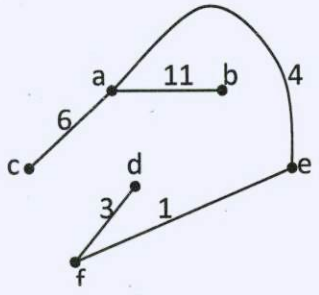


Fig 4.5: Kruskal's spanning tree

Prim's method:

Here we need to prepare a matrix of weights.

	a	b	c	d	e	f
a	-	11	6	5	4	8
b	11	-	∞	∞	12	∞
c	6	∞	-	7	∞	∞
d	5	∞	7	-	5	3
e	4	12	∞	5	-	1
f	8	∞	∞	3	1	-

Begin the tree with vertex a and add one new vertex to the tree at a time. The order of addition of vertices is shown in the table below.

Vertices in the tree	Min weight edge and weight	Total weight so far
a	(a,e): 4	4
a,e	(e,f): 1	5
a,e,f	(d,f): 3	8
a,e,f,d	(a,c): 6	14
a,e,f,d,c	(a,b): 11	25

Figure 4.6 is the spanning tree of Prim's method.

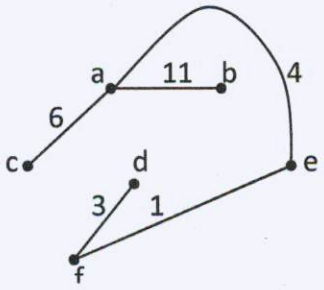


Fig 4.6: Spanning tree of Prim's method

Note that both the methods have generated the same spanning tree; this need not be so in all cases. However, the total weight of both the trees will be the same.

4.4 Degree constrained shortest spanning tree

In a shortest spanning tree of the preceding section we can have any number of edges from a vertex. In some practical cases an upper limit on the degree of a vertex in the tree has to be imposed. For instance, in an electrical wiring problem, one may be required to wire together pins, using as little wire as possible (hence a minimum weight spanning tree to be found) with not more than three wires wrapped around any individual pin. Thus degree of vertices in the tree cannot exceed 3. Such spanning trees are called degree constrained shortest spanning trees.

In general, the problem may be stated as follows: Given a weighted connected graph G, find a shortest spanning tree T in G such that degree of any vertex is $\leq k$. If k=2, this problem, in fact,

reduces to the problem of finding the shortest Hamiltonian path, as well as the travelling salesman problem (without the salesman returning to his home base). So far, no efficient method to find degree constrained shortest spanning tree is available.

4.5 Summary

In this unit we discussed spanning tree of a graph and listed some simple and useful properties. We then discussed weighted graphs, shortest spanning trees and applications. Algorithms to find minimum weight spanning tree in graph G is given with illustrations. The unit is concluded with a type of shortest spanning tree where, degree of vertices in the tree is bounded above.

4.6 Keywords

Spanning trees, weighted graph, minimum weight spanning tree

4.7 Practice problems

Easy problems: Based on the notes

1. Define spanning trees. Draw a graph of 4, 5 and 6 vertices and show spanning trees (at least 3) for each graph. Find the number of branches and chords in each tree.
2. Draw weighted graphs and find spanning trees (multiple) and find their weights.
3. Discuss some applications of minimum weight spanning tree.
4. Explain Kruskal's and Prim's methods.
5. Illustrate Kruskal's and Prim's methods with an example of graph of 6 or more vertices with twice the number of edges.
6. Explain the concept of degree constrained shortest spanning tree.

Challenging problems: Refer the texts listed in the references

1. Show that a Hamiltonian path is a spanning tree.
2. Let T and T' be two spanning trees. If edge e is in T but not in T', prove that there is an edge f in T' but not in T. Also show that $(T-e)\cup f$ and $(T'-f)\cup e$ are spanning trees.
3. Show that shortest tree should always contain the edge with minimum weight.

4.8 References

1. Narsingh Deo, Graph Theory with applications to Engineering and Computer Science, PHI - Chapter 3 (Section 3.7, 3.10)
2. J.P. Tremblay, R.Manohar, Discrete Mathematical structures, with applications to Computer Science, TATA McGraw Hil - Chapter 5 (5.1.4)
3. N.G.Goudru, Discrete Mathematical structures, Himalaya Publishing House - Chapter 7 (Section 7.10)
4. Bernard Kolman, Robert C. Busby, Sharon Cutler Ross, Discrete Mathematical structures, PEARSON Education, Chapter 7 (Section 7.5)

Karnataka State Open University

Manasagangothri, Mysore-570006

M. Sc. (Computer Science)

MSC-501: DISCRETE MATHEMATICS

MODULE	UNITS
5	1 to 4

Unit 1: Introduction to Group Theory

Structure

1.0 Objectives

After going through this lesson you will be able to

- Explain semigroup, monoid, group and subgroup;
- Give an account of the properties of a group;
- Analyze some theorems on subgroup.

1.1 Introduction

Semigroups are the simplest algebraic structures which satisfy the properties of closure and associativity. They are very important in the theory of sequential machines, formal languages and in certain applications relating to computer arithmetic such as multiplication.

A monoid in addition to being a semigroup, also satisfies the identity property. Monoids are used in a number of applications, but most particularly in the area of syntactic analysis and formal languages.

Groups are monoids which also possess the inverse property. The application of group theory is important in the design of fast adders and error correcting codes.

Application of subgroups is in the construction of computer modules which perform group operations. Such modules are constructed by joining various subgroup modules that do operations in subgroups.

1.2 Semigroups and monoids

Definition: An n-ary operation is a mapping $f: X^n \rightarrow X$ and n is called the order of the operation. For $n=1$, $f: X \rightarrow X$ is called a unary operation. For $n=2$ it is called a binary operation.

Definition: Let X be a set and f be a mapping from $X \times X$ to X *i. e.* $f: X \times X \rightarrow X$. Then f is called a binary operation on X.

In other words, a binary operation is a rule that assigns to each ordered pair of elements of X, a unique element of X. A binary operation is denoted by the symbol *.

For $x, y \in X$, if $x*y \in X$, then * is a binary operation.

1.2.1 Properties of Binary Operation:

- A binary operation * is said to be commutative if for every $x, y \in X$, $x*y=y*x$.
- A binary operation * is said to be associative if for every $x, y, z \in X$, i.e., $(x*y)*z=x*(y*z)$
- Let * be a binary operation on X. If there exists an element $e \in X$ such that $a*e = e*a = a$, $\forall a \in X$, then e is called the identity element.
- Let * be a binary operation on X with the identity e. If there exists an element $a^{-1} \in X$ such that $a*a^{-1} = a^{-1}*a = e$, then a^{-1} is called the inverse element of a.

Definition: A non-empty set G together with an associative binary operation * is called a semigroup. It is denoted by $(G, *)$. In other words, $(G, *)$ is called a semigroup if for all $a, b, c \in G$, $(a * b)*c = a*(b*c)$.

Example 1: Let the binary operation * be defined by $x*y= xy$ on the set of all integers Z. Show that $(Z, *)$ is a semigroup.

Solution: We know that $Z= \{\ldots -2,-1, 0, 1, 2\ldots.\}$

For $-3,-2,-1 \in Z$

$(-3*-2)*(-1) = 6*(-1) = -6 \in Z$

$-3*(-2*-1) = -3*(2) = -6 \in Z$

Thus, for any $x, y, z \in Z$, $(x*y)*z = x*(y*z)$

So, the binary operation * is associative

Thus, $(Z, *)$ is a Semi group.

Example 2: Let the binary operation * be defined by $x*y= xy +2y$ on the set of all real numbers R. Is $(R, *)$ a semigroup?

Solution: For $-2,-3,-4 \in R$

$(-2*-3)*(-4) = ((-2)(-3) + 2(-3))*(-4)$

$= (6-6)*(-4)$

$= 0*(-4)$

$= (0)(-4) + 2(-4)$

$= 0-8$

$(-2*-3)*(-4) = -8$ ---------- (1)

$(-2)*(-3*-4) = (-2)*((-3)(-4)+2(-4))$

$= (-2)*(12-8)$

$= (-2)*(4)$

$= (-2)(4) + 2(4)$

$= -8+8$

$(-2)*(-3*-4) = 0$ ---------- (2)

From (1) & (2)

$(-2*-3)*(-4) \neq (-2)*(-3*-4)$

For $x, y, z \in R$, $x*(y*z) \neq (x*y)*z$

So the binary operation * is not associative

Hence $(R, *)$ is not a semi-group.

Example 3: Show that the set of all natural numbers is a semigroup under the binary operation addition '+'.

Solution: We know that the set of all natural numbers is $N = \{1, 2, 3, \ldots\}$

For $3,2,1 \in N$

$(3 + 2) + 1 = 5 + 1 = 6 \in N$

$3 + (2 + 1) = 3 + 3 = 6 \in N$

Thus, for any $x, y, z \in N$, $(x*y)*z = x*(y*z)$

So, the binary operation * is associative

Thus, $(N, +)$ is a Semi group.

Definition: A **monoid** is a semigroup with identity. In other words a non-empty set G, together with a binary operation * is called a monoid if

(i) * is associative

i. e. $(a * b)*c = a*(b*c)$, for all $a, b, c \in G$

(ii) $\exists$ an element $e \in G$ such that

$a * e = e * a = a, \quad \forall a \in G$

Example 4: Show that the set of all integers is a monoid under ordinary multiplication "."

Solution: We know that $Z = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$

(i) For $-3, -2, 1 \in Z$

$(-3 \,.\, -2) \,.\, 1 = 6 \,.\, 1 = 6 \in Z$

$(-3) \,.\, (-2 \,.\, 1) = (-3).(-2) = 6 \in Z$

Thus, for any $x, y, z \in N$, $(x*y)*z = x*(y*z)$

So, the binary operation * is associative.

(ii) 1 is the identity element

$1.\, x = x.1 = x, \quad \forall x \in Z.$

Therefore $(Z, .)$ is a monoid.

Example 5: Let Z^+ denote the set of all positive integers and the binary operation * be defined by $x * y = \max\{x, y\}$. Is $(Z^+, *)$ a monoid?

Solution: We know that $Z^+ = \{1, 2, 3\ldots.\}$

For $x, y, z \in Z^+$

$(x * y)*z = (\max\ \{x, y\})*z$

$(x * y) *z = \max\ \{x, y, z\} \rightarrow (1)$

$x*(y*z) = x*(\max\{y, z\})$

$x*(y*z) = \max\{x, y, z\}$

Hence $(x*y)*z = x*(y*z)$

So the binary operation * is associative

Now $1 \in Z^+$ acts as identity element

For $x \in Z^+$, $1 * x = \max\ \{1, x\} = x$

$x * 1 = \max\ \{x, 1\} = x$

Hence 1 is identity element

Therefore $(Z^+, *)$ is a Monoid.

Note: If binary operation is addition +, then 0 is the identity element.

1.3 Groups

Definition: A non empty set G together with a binary operation * is called a group if the following axioms are satisfied.

(i) Associative axiom.

$(a*b)*c = a*(b*c)$ for all $a, b, c \in G$

(ii) Identity axiom.

There exists an element $e \in G$ such that

$a * e = e * a = a, \quad \forall\ a \in G$

(iii) Inverse axiom.

For each $a \in G$, there exists $a^{-1} \in G$ such that

$a * a^{-1} = a^{-1} * a = e$

1.3.1 Commutative group:

A group $(G, *)$ is called abelian (commutative)

if $a*b = b*a \quad \forall\ a, b \in G.$

Example 6: Show that the set of all integers is an abelian group under addition '+'.

Solution: We know that $Z = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$

(i) For $-3, -2, 1 \in Z$

$((-3)+(-2)) + 1 = (-5) + 1 = -4 \in Z$

$(-3) + ((-2)+ 1) = (-3)+(-1) = -4 \in Z$

Thus, for any $x, y, z \in Z$, $(x + y) + z = x + (y + z)$

So, the binary operation + is associative.

(ii) 0 is the identity element

$0 + x = x + 0 = x, \quad \forall\ x \in Z.$

(iii) For any $x \in Z$, there exists an element $-x \in Z$ such that

$x + (-x) = (-x) + x = 0.$

(iv) For $x, y \in Z$, $x + y = y + x$.

Therefore $(Z, +)$ is an abelian group.

Example 7: Show that the set of all positive integers is not a group under ordinary multiplication.

Solution: We know that $Z^+ = \{1, 2, 3, \ldots\}$

(i) For any $x, y, z \in Z^+$, $(x.y).z = x.(y.z)$

So, the binary operation . is associative.

(ii) 1 is the identity element

$1.x = x.1 = x, \quad \forall\ x \in Z^+$.

(iii) For $2 \in Z^+$, ½ does not belong to Z^+.

Thus for any $x \in Z^+$ with $x \neq 1$, $1/x \notin Z^+$.

So, inverse axiom is not satisfied.

Therefore $(Z^+, .)$ is not a group.

Example 8: Show that the set of all non-zero real numbers R is an abelian group under the binary operation * defined by $a * b = ab/2$.

Solution:

(i) For any $x, y, z \in R$, $(x*y)*z = (xy/2)*z = xyz/4$.

$x*(y*z) = x*(yz/2) = xyz/4$.

Hence $(x*y)*z = x*(y*z)$.

So, the binary operation * is associative.

(ii) $2 \in R$ is the identity element

For, $2 * x = 2x/2 = x$.

$x * 2 = x2/2 = x, \quad \forall\ x \in R$.

(iii) If $x \in R$, then $4/x$ is the inverse of x.

For, $x*4/x = x(4/x)/2 = 2$

$4/x*x = (4/x)x/2 = 2$.

Therefore $(R, *)$ is a Group.

For any $x, y \in R$, $x*y = xy/2 = yx/2 = y*x$.

Therefore the binary operation * is commutative.

Thus $(R, *)$ is an abelian group.

1.3.2 Properties:

If $(G, *)$ is a group, then

(1) The identity of G is unique.

(2) For each $a \in G$, a^{-1} is unique.

(3) $(a^{-1})^{-1} = a$, for $a \in G$.

(4) $(a*b)^{-1} = b^{-1} * a^{-1}$.

Proof:

(i) If possible, let e_1 and e_2 be two identities of G.

If e_1 is the identity then $e_2 * e_1 = e_1 * e_2 = e_2$

If e_2 is the identity then $e_1 * e_2 = e_2 * e_1 = e_1$

$\therefore e_1 = e_1 * e_2 = e_2$

$\therefore e_1 = e_2$

(ii) Suppose a and b are two inverses of c. Let e be the identity of G. Then,

$a * c = c * a = e$ -------(1)

$b*c = c* b = e$ --------(2)

$b = b*e$, by identity law

$b = b* (c*a)$, by (1)

$b = (b*c)* a$, by associative law

$= e * a$, by (2)

$= a$, by identity law.

(iii) Now,

$(a^{-1})^{-1} * a^{-1} = e$	[By definition]
$((a^{-1})^{-1} * a^{-1}) * a = e*a$	[By multiplying on the right by a]
$(a^{-1})^{-1} * (a^{-1} * a) = a$	[By associative and identity axioms]
$(a^{-1})^{-1} * e = a$	[By inverse axiom]
$(a^{-1})^{-1} = a$	[By identity axiom]

$\therefore$ Inverse of a^{-1} is a.

(iv) Let $a, b \in G$. Let a^{-1} and b^{-1} be inverses of a and b.

Consider,

$(a*b) * (b^{-1}*a^{-1}) = a*(b*(b^{-1}*a^{-1}))$ [By associative axiom]

$= a*((b*b^{-1})*a^{-1})$ [By associative axiom]

$= a*(e*a^{-1})$ [By inverse axiom]

$= a*a^{-1}$ [By identity axiom]

$= e$ [By inverse axiom]

$\therefore (a*b)^{-1} = b^{-1} * a^{-1}$

1.3.3 Theorem: If a, b, c are elements of a group G, then

i) $ab = ac$ implies $b = c$ (left cancellation law)

ii) $ba = ca$ implies $b = c$ (right cancellation law)

Proof:

i) Suppose that $ab = ac$

Multiplying on the left by a^{-1}, we get

$a^{-1}(ab) = a^{-1}(ac)$

$(a^{-1}a)b = (a^{-1}a)c$ [By associative axiom]

$(e)b = (e)c$ [By inverse axiom]

$b = c$ [By identity axiom]

ii) Suppose that $ba = ca$

Multiplying on the right by a^{-1}, we get

$(ba)a^{-1} = (ca)a^{-1}$

$b(aa^{-1}) = c(aa^{-1})$ [By associative axiom]

$b(e) = c(e)$ [By inverse axiom]

$b = c$ [By identity axiom]

1.3.4 Theorem: If a and b are elements of a group G, then

i) The equation $ax = b$ has a unique solution in G

ii) The equation $ya = b$ has a unique solution in G.

Proof: (i) We observe that $a(a^{-1}b) = (aa^{-1})b = eb = b$

So, $x = a^{-1}b$ is a solution of $ax = b$.

To prove the uniqueness, let x_1 and x_2 be two solutions of the equation $ax = b$.

Thus, $ax_1 = b$ and $ax_2 = b$.

So, $ax_1 = ax_2$

$x_1 = x_2$, by left cancellation law.

Hence, the equation $ax = b$ has a unique solution.

(ii) Similarly, we observe that $(ba^{-1})a = b(a^{-1}a) = be = b$

So, $y = ba^{-1}$ is a solution of $ya = b$.

To prove the uniqueness, let y_1 and y_2 be two solutions of the equation $ya = b$.

Thus, $y_1a = b$ and $y_2a = b$.

So, $y_1a = y_2a$.

$y_1 = y_2$, by right cancellation law.

Hence, the equation $ya = b$ has a unique solution.

Example 9: Let Z be set of all integers with the binary operation * defined by $a*b = a + b + 1$ for $a, b \in Z$. Then show that $(Z, *)$ is an abelian group.

Solution: For $a, b, c \in Z$, $(a*b)*c = (a+b+1)*c$

$= (a+b+1) + c + 1$

$= a+b+c+2.$

$a*(b*c) = a*(b+c+1)$

$= a+(b+c+1)+1$

$= a+b+c+2$

$\therefore (a*b)*c = a*(b*c)$

$\therefore$ * is associative.

Let $e \in Z$ be such that $e*a = a$

$\Rightarrow e + a + 1 = a$

$\Rightarrow e = -1$

$\therefore$ Identity of $(Z, *)$ is -1

For $a \in Z$, let $b \notin Z$ be the inverse of a, then

$a*b = -1$

$\Rightarrow a + b + 1 = -1$

$\Rightarrow b = -a-2$

$\therefore$ Inverse of a is $-a-2$

Also, $a*b = a + b + 1$

$= b + a + 1$

$= b * a$

$\therefore (Z, *)$ is an abelian group.

Example 10: Determine whether the set $G = \{1, w, w^2\}$, where w is the cube root of unity forms a group under multiplication.

Solution: The multiplication table for G is given by

$\times$	1	w	w^2
1	1	w	w^2
w	w	w^2	1
w^2	w^2	1	w

All the entries in the table belong to the set G. Therefore closure axiom holds.

For $a \in G$, $a*1 = 1* a = a$. Therefore 1 is the identity element.

Inverse of 1, w, w^2 are 1, w^2, w respectively.

All the entries are symmetric with respect to the diagonal. Therefore the commutative axiom is satisfied.

Thus $(G, \times)$ is an abelian group.

Example 11: Let $G = \{1,-1\}$. Determine whether G is a group under multiplication of real numbers.

Solution: Construct the multiplication table,

$\times$	1	-1
1	1	-1
-1	-1	1

Clearly from the table,

(i) '×' is associative.

(ii) 1 is the identity element

(iii) The inverse of 1 is itself; -1 is the inverse of itself.

Hence, $(G, \times)$ is a group.

Example 12: Is the set $\{1, 2, 3, 4, 5\}$ a group under addition modulo 6?

Solution: Construct the addition table,

+6	1	2	3	4	5
1	2	3	4	5	0
2	3	4	5	0	1
3	4	5	0	1	2
4	5	0	1	2	3
5	0	1	2	3	4

From the table, given composition is not binary, as $5 + 1 = 0$ but $0 \notin \{1, 2, 3, 4, 5\}$.

Therefore the given set is not a group.

Example 13: Let $G = \left\{ \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} : a \neq 0 \in R \right\}$. Is G a group under matrix multiplication?

Solution: Let $A = \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} b & 0 \\ 0 & 0 \end{pmatrix} \in G$ with $a, b \neq 0$ in R.

Consider,

$AB = \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} b & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} ab & 0 \\ 0 & 0 \end{pmatrix} \in G$ as $ab \neq 0$.

Therefore G is closed w.r.t. matrix multiplication.

Similarly matrix multiplication is associative.

$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ is the identity element. For $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix}$ for any $\begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix}$ in G.

Consider, $\begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$

$$\begin{pmatrix} ax & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

$$\Rightarrow ax = 1;\ x = 1/a$$

$\therefore \begin{pmatrix} 1/a & 0 \\ 0 & 0 \end{pmatrix}$ is the inverse of $\begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix}$

Also, $AB = \begin{pmatrix} ab & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} ba & 0 \\ 0 & 0 \end{pmatrix} = BA$

$\therefore$ G is an abelian group.

1.4 Subgroup

Definition: A non-empty set H of a group G is called a subgroup if H itself is a group under the operation defined in G.

Any group G has at least two subgroups namely $\{e\}$, the set containing the identity element e and G itself. These two subgroups are called trivial subgroups.

Examples:

- Let $G = \{1, -1, i, -i\}$, a multiplicative group. Then $H = \{1,-1\}$ is subgroup of G.
- Set of all integers Z is a subgroup of the set of all rationals Q under addition.

1.4.1 Theorem:

A non-empty subset H of a group G is a subgroup of G if and only if

1) $a, b \in H$ implies that $ab \in H$.

2) $a \in H$ implies that $a^{-1} \in H$.

Proof: Let H be subgroup of G and $a, b \in H$

Now $a \in H$ and $b \in H \Rightarrow ab \in H$. (by the closure axiom in H, being a subgroup)

Since H is a subgroup, for any $a \in H \Rightarrow a^{-1} \in H$.

Conversely, suppose condition (1) and (2) holds. To prove that H is a subgroup of G, it is enough to prove that associative and identity axioms hold in H.

Each element of H is an element of G. Since associative axiom holds good in G, being a group. Thus associative law holds good for elements of H also.

Now for any $a \in H$ by condition (2), $a^{-1} \in H$ and by (1), $e = a.a^{-1} \in H$.

$$\therefore e \in H.$$

$\therefore H$ is a group. Since $H \subseteq G$, H is a subgroup of G.

1.4.2 Theorem: (Necessary and sufficient condition for the subgroup)

A non-empty set H of a group G is subgroup of G, if and only if $a, b \in H \Rightarrow ab^{-1} \in H$

Proof: Let H be subgroup of G and $a, b \in H$

Since H is a subgroup, for any $b \in H \Rightarrow b^{-1} \in H$.

Now $a \in H$ and $b^{-1} \in H \Rightarrow ab^{-1} \in H$. (by the closure axiom in H, being a subgroup)

Thus for $a, b \in H \Rightarrow ab^{-1} \in H$

Conversely, suppose $a, b \in H \Rightarrow ab^{-1} \in H$.

We prove that H is a subgroup of G.

Let $a \in H$ be arbitrary.

Given, $a \in H, b \in H \Rightarrow ab^{-1} \in H$.

Choose $b = a$, then $a \in H, a \in H \Rightarrow aa^{-1} = e \in H$.

Therefore identity element exists.

Now $e \in H, a \in H \Rightarrow e\,a^{-1} = a^{-1} \in H$

Hence the inverse of every element of H exists and belongs to H.

Let $a, b \in H \Rightarrow a \in H, b^{-1} \in H$.

$\Rightarrow a\,(b^{-1})^{-1} \in H$.

$\Rightarrow ab \in H$.

Therefore closure axiom is satisfied.

Since the binary operation is associative in G, it is associative in H also.

Thus H is a subgroup of G.

Example 14: Let $G = \{1, -1, i, -i\}$ be a multiplicative group. Show that $H = \{1, -1\}$ is a subgroup of G.

Solution: Clearly $H \subseteq G$ and the multiplicative identity $1 \in H$

Now i) For $1, -1 \in H \Rightarrow 1(-1) = -1 \in H$.

For $1, 1 \in H \Rightarrow 1(1) = 1 \in H$.

$\therefore$ Condition (1) of theorem is satisfied.

ii) $1.1 = 1$, $\therefore$ the inverse of 1 is 1

$(-1) \cdot (-1) = 1 \therefore$ the inverse of -1 is -1

$\therefore$ for every $a \in H$, $a^{-1} \in H$

$\therefore$ Condition (2) of theorem is satisfied.

$\therefore H$ is a subgroup of G

Example 15: Prove that the set of all integers Z, is a subgroup of set of all rationals Q, under addition.

Solution: Clearly $Z \subseteq Q$

The identity element 0 of Q belongs to Z

Now i) For $a, b \in Z \Rightarrow a+b \in Z$

ii) For every $a \in Z$, there exists $-a \in Z$ such that

$$a + (-a) = (-a) + a = 0$$

$\therefore$ By the theorem, Z is a subgroup of Q.

Example 16:

$$\text{Let } G = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \;\middle/\; \begin{matrix} a,b,c,d \in R \\ ad - bc \neq 0 \end{matrix} \right\}$$

$$\text{and } H = \left\{ \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \;\middle/\; \begin{matrix} a,b,d \in R \\ ad \neq 0 \end{matrix} \right\} \text{ then show that}$$

H is a subgroup of G under multiplication,

Solution: Let $A = \begin{pmatrix} a_1 & b_1 \\ 0 & d_1 \end{pmatrix}$, $B = \begin{pmatrix} a_2 & b_2 \\ 0 & d_2 \end{pmatrix}$ be two elements of H.

$$\text{Consider } AB = \begin{pmatrix} a_1 & b_1 \\ 0 & d_1 \end{pmatrix} \begin{pmatrix} a_2 & b_2 \\ 0 & d_2 \end{pmatrix}$$

$$= \begin{pmatrix} a_1 a_2 & a_1 b_2 + b_1 d_2 \\ 0 & d_1 d_2 \end{pmatrix} \in H$$

Since, $a_1 d_1 \neq 0$ and $a_2 d_2 \neq 0$

$\Rightarrow (a_1 a_2)(d_1 d_2) = (a_1 d_1)(a_2 d_2) \neq 0$

For any $A=\begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \in H$

$$\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}\begin{pmatrix} x & y \\ z & w \end{pmatrix}=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} ax+bz & ay+bw \\ dz & dw \end{pmatrix}=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$\Rightarrow$ $ax+bz=1$, $ay+bw=0$, $dz=0$, $dw=1$

Now, $d\neq 0$ as $ad \neq 0$

$\therefore dz=0 \quad \Rightarrow \quad z=0$

$\therefore ax+bz=1 \quad \Rightarrow \quad ax=1 \Rightarrow x=1/a$

$dw=1 \quad \Rightarrow w=1/d$

$\therefore ay+bw=0 \Rightarrow \quad y=-bw/a$

$= -b/a(1/d)$

$\therefore \quad y=-b/ad$

Inverse of $\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$ is $\begin{pmatrix} 1/a & -b/ad \\ 0 & 1/d \end{pmatrix}$

and, $ad \neq 0 \Rightarrow 1/ad \neq 0$

$$\therefore \begin{pmatrix} 1/a & -b/ad \\ 0 & 1/d \end{pmatrix} \in H$$

$\therefore$ H is subgroup.

1.5 Solved Problems

1.5.1. If $(G,*)$ is an abelian group, then for all $a, b \in G$ show that $(a*b)^n=a^n*b^n$.

Solution: We prove the result by induction on n.

If $n=1$, then $(a*b)^1=a*b$ [trivial]

If $n=2$, then $(a*b)^2 = (a*b)*(a*b)$

$= a*(b*a)*b$ [By associative axiom]

$= a*(a*b)*b$ [G is abelian]

$= (a*a)*(b*b)$ [By associative axiom]

$= a^2*b^2$

Result is true for n=2.

We assume that the result is true for n=m.

i.e. $(a*b)^m = a^m * b^m$

Consider ,

$(a*b)^{m+1}=(a*b)^m * (a*b)$

$=(a^m * b^m) * (a*b)$ [By the assumption]

$=(a^m * b^m) * (b*a)$ [G is abelian]

$=a^m * (b^m * b)*a$ [By associative axiom]

$=a^m * (b^{m+1}) * a$

$=a^m * (b^{m+1}*a)$ [By associative axiom]

$=a^m*(a*b^{m+1})$ [G is abelian]

$=(a^m*a)*b^{m+1}$ [By associative axiom]

$=a^{m+1}*b^{m+1}$

$\therefore$ Result is true for n=m+1. Thus by induction result is true for all n.

1.5.2. Show that in a group $(G,*)$, if for $a, b \in G$, $(a*b)^2=a^2*b^2$ then $(G,*)$ must be abelian.

Solution: Given: $(a*b)^2 = a^2*b^2$

$\Rightarrow (a*b) * (a*b) = (a*a)*(b*b)$

$\Rightarrow a*(b*a)*b = a*(a*b)*b$ [By associative axiom]

$\Rightarrow b*a = a*b,$ [By cancellation law]

$\therefore$ G is abelian.

1.5.3. Show that if every element in a group is its own inverse, then the group G must be abelian.

Solution: Let $a, b \in G \Rightarrow ab \in G$

We have, $a=a^{-1}$, $b=b^{-1}$, $ab=(ab)^{-1}$ [Given: Every element is its own inverse]

$\Rightarrow a^2 = e$, $b^2=e$, $(ab)^2=e$

Consider, $(ab)^2 = e$

$\Rightarrow (ab).(ab)=e$

$\Rightarrow a(ba)b=e$ [By associative axiom]

$\Rightarrow a(a(ba)b)b= aeb$ [Multiplying by a on the left and by b on the right]

$\Rightarrow (aa)(ba)(bb) = ab$ [By associative axiom]

$\Rightarrow a^2(ba)b^2 = ab$

$\Rightarrow e(ba)e=ab$

$\therefore ba=ab$

$\therefore G$ is abelian.

1.5.4 If H_1 & H_2 are subgroups of G, show that $H_1 \cap H_2$ is also a subgroup of G. Show that in general $H_1 \cup H_2$ need not be subgroup of G except when $H_1 \subseteq H_2$ or $H_2 \subseteq H_1$.

Solution: Let $a, b \in H_1 \cap H_2$.

$\Rightarrow a, b \in H_1$ and $a, b \in H_2$.

$\therefore ab^{-1} \in H_1$ and $ab^{-1} \in H_2$ [as H_1 & H_2 are subgroups].

$\therefore ab^{-1} \in H_1 \cap H_2$, $\forall a, b \in H_1 \cap H_2$

$\therefore H_1 \cap H_2$ is a subgroup.

We now give an example to show that $H_1 \cup H_2$ is not a subgroup of G.

Let $H_1=(2Z, +)$, $H_2=(3Z, +)$, $G=(Z, +)$. Then H_1 and H_2 are subgroups of G.

Clearly $2 \in 2Z$ and $3 \in 3Z$

But $3-2=1 \notin 2Z \cup 3Z$

$\therefore 2Z \cup 3Z$ is not a subgroup.

If $H_1 \subseteq H_2 \Rightarrow H_1 \cup H_2 = H_2$, a subgroup of G

If $H_2 \subseteq H_1 \Rightarrow H_1 \cup H_2 = H_1$, a subgroup of G

Suppose that $H_1 \cup H_2$ is a subgroup and $H_1 \not\subset H_2$ and $H_2 \not\subset H_1$, then

$H_1 \not\subset H_2 \Rightarrow \exists a \in H_1$ such that $a \notin H_2$

$H_2 \not\subset H_1 \Rightarrow \exists b \in H_2$ such that $b \notin H_1$

But, $a, b \in H_1 \cup H_2$

$\Rightarrow a - b \in H_1 \cup H_2$ ($\because H_1 \cup H_2$ is a subgroup)

$\Rightarrow a-b \in H_1$ or $a-b \in H_2$

Suppose $a-b \in H_1$, then $a-b \in H_1$ and $a \in H_1$

$\Rightarrow a-(a-b) \in H_1$

$\Rightarrow b \in H_1$, a contradiction.

Suppose $a-b \in H_2$, then $a-b \in H_2$ and $b \in H_2$

$\Rightarrow (a-b)+b \in H_1$

$\Rightarrow a \in H_2$, a contradiction.

$\therefore H_1 \cup H_2$ is a subgroup iff $H_1 \subseteq H_2$ or $H_2 \subseteq H_1$

1.6 Summary

A binary operation is a rule that assigns to each ordered pair of elements of a set, a unique element of it.

A non-empty set together with an associative binary operation is called a semigroup.

A monoid is a semigroup with identity.

A non empty set together with a binary operation is called a group if it satisfies associative, identity and inverse axioms.

Properties of a group: If $(G, *)$ is a group, then

(1) The identity of G is unique.

(2) For each $a \in G$, a^{-1} is unique.

(3) $(a^{-1})^{-1} = a$, for $a \in G$.

(4) $(a*b)^{-1} = b^{-1} * a^{-1}$.

Theorem: If a, b, c are elements of a group G, then

i) $ab = ac$ implies $b = c$ (left cancellation law)

ii) $ba = ca$ implies $b = c$ (right cancellation law)

Theorem: If a and b are elements of a group G, then

i) The equation $ax = b$ has a unique solution in G

ii) The equation $ya = b$ has a unique solution in G.

A non-empty set H of a group G is called a subgroup if H itself is a group under the operation defined in G.

Theorem :

A non-empty subset H of a group G is a subgroup of G if and only if

3) $a, b \in H$ implies that $ab \in H$.

4) $a \in H$ implies that $a^{-1} \in H$.

1.7 Keywords

Semigroup, monoid, group, subgroup.

1.8 Supplementary problems

1.8.1. Define $x \times y = x - y$ on the set of all +ve integers. Is * a binary operation?

Solution: No

1.8.2. Show that the set N of natural numbers is a semigroup under the operation $x*y = max\{x, y\}$. Is it a monoid?

1.8.3. Let S be a finite set and $P(S)$ be the power set of S. Determine whether $(P(S), \cap)$ is a semigroup or a monoid.

1.8.4. Determine whether $(Z^+, *)$ where $x*y = x+y-xy$ is a semigroup or a monoid.

Solution: Monoid

1.8.5. Determine whether the set of even integers with the binary operations $x*y = \frac{xy}{2}$ forms a semigroup or a monoid.

Solution: Monoid

1.8.6. Determine whether the set Z with the binary operation *, ordinary multiplication is a group.

Solution: Not a group.

1.8.7. Is the set of +ve rationals a subgroup of the group of numbers under the operation of addition?

Solution: Not a subgroup

1.8.8. Let G be the non zero integers under the operation of multiplication and let $H = \{3^n | n \in R\}$. Is H a subgroup of G?

Solution: Yes

1.9 References:

1. Discrete mathematics for computer science, by Kolman, Busby and Ross, PHI publications.
2. Discrete Mathematical Structures, by N. G. Goudru, Himalaya Publishing House.
3. Discrete mathematical structures with applications to computer science, by Tremblay and Manohar (McGraw-Hill publications).
4. Topics in Algebra, by I. N. Herstein, Wiley eastern Ltd.

Unit 2: Cosets, Lagrange's Theorem and Normal Subgroups

Structure

2.0 Objectives

After going through this lesson you will be able to

- Explain the cosets;
- Analyse some theorems on cosets;
- Give an account of the Lagrange's theorem;
- Explain the normal subgroup;
- Analyse some theorems on normal subgroups;

2.1 Introduction

From the definition of a subgroup it is clear that not every subset of a group is a subgroup. To find those subsets which can qualify to become subgroups is an interesting problem. An important relationship exists between the subgroups and the group itself. This relationship is explained by a theorem known as Lagrange's theorem. This theorem has important application in the development of efficient group codes required in the transmission of information.

2.2 Cosets

Definition: Let H be a subgroup of G and $a \in G$. Then the set, $Ha = \{ha \mid h \in H\}$ is called the right coset of H generated by a. Similarly the set $aH = \{ah \mid h \in H\}$ is called the left coset of H generated by a.

Note: Since $eH=H=He$, we see that H itself is a right as well as a left coset. If the group operation is "addition" we define the right coset of H in G by $H+a=\{h+a \mid h \in H\}$. Similarly for the left coset of H in G by $a+H=\{a+h \mid h \in H\}$. Observe that cosets are not necessarily subgroups of G.

Example 1: Let $G=(Z,+)$ be the additive group of integers and $H=(2Z, +)$ the subgroup of even integers. Find the left cosets of H in G.

Solution:

The left cosets of H in G are,

$0+H=\{\ldots, 0+(-2), 0+0, 0+2, \ldots\}=\{\ldots, -2, 0, 2, \ldots\}=H.$

$1+H = \{\ldots, 1+(-2), 1+0, 1+2, \ldots\} = \{\ldots, -3, -1, 1, 3, \ldots\}$.

$2+H = \{\ldots, -4, -2, 0, 2, 4, 6 \ldots\} = H$.

$3+H = \{\ldots, -3, -1, 1, 3, \ldots\}$ and so on.

Notice that $2+H$ coincides with H, $3+H$ coincides with $1+H$, $4+H$ coincides with H, $5+H$ coincides with $1+H$ and so on. Hence there are only two distinct left cosets namely H and $1+H$.

Example 2: Let $G = \{1, -1, i, -i\}$ be a multiplicative group and $H=\{1, -1\}$ be a subgroup of G. Find the right cosets of H in G.

Solution: The right cosets are

$$H1 = \{1(1), -1(1)\} = \{1, -1\} = H.$$

$$H(-1) = \{1(-1), -1(-1)\} = \{-1, 1\} = H.$$

$$Hi = \{i, -i\}.$$

$$H(-i) = \{-i, i\} = Hi.$$

Note:

1) If G is abelian then right and left cosets of G coincide.
2) If $a \in H$, then $Ha = H$. If $a \in G$ such that $a \notin H$, then $Ha \neq H$.
3) H itself is a right coset and the number of elements in each right coset is the same as the number of elements in H.

2.2.1 Theorem: There is one-to-one correspondence between any two right cosets of a subgroup H of a group G.

Proof: Let $a, b \in G$. Let Ha and Hb be any two right cosets of H in G.

Define f: $Ha \rightarrow Hb$ by $f(ha) = hb \; \forall \; ha \in Ha$.

***f* is one-one:**

Let $h_1, h_2 \in H$

Then, $h_1a, h_2a \in Ha$

Now, $f(h_1a) = h_1b$, $f(h_2a) = h_2b$.

Suppose $f(h_1a) = f(h_2a)$

$\Rightarrow h_1 b = h_2 b$

$\Rightarrow h_1 = h_2$

$\Rightarrow h_1 a = h_2 a.$

$\therefore f$ is one-one .

f is onto:

Let $hb \in Hb$ be arbitary.

$\Rightarrow h \in H$, then there exists $ha \in Ha$.

$\therefore f(ha) = hb$ by the definition of f.

$\therefore f$ is onto

This proves the theorem.

2.2.2 Theorem: Let H be a subgroup of G and $a,b \in G$. Then $Ha=Hb$ if and only if $ab^{-1} \in H$

Proof: Let $Ha=Hb$.

Then $\exists$ elements h_1 and h_2 in H such that

$h_1 a = h_2 b$

$\Rightarrow h_1^{-1}(h_1 a) = h_1^{-1}(h_2 b)$ [By multiplying on the right by h_1^{-1}]

$\Rightarrow (h_1^{-1} h_1)a = (h_1^{-1} h_2)b$

$\Rightarrow ea = (h_1^{-1} h_2)b$

$\Rightarrow a = (h_1^{-1} h_2)b$

$\Rightarrow ab^{-1} = [(h_1^{-1} h_2)b]\, b^{-1}$

$\Rightarrow ab^{-1} = (h_1^{-1} h_2)(bb^{-1})$

$\Rightarrow ab^{-1} = (h_1^{-1} h_2)e$

$\Rightarrow ab^{-1} = h_1^{-1} h_2$

Since $h_1 \in H$, $h_1^{-1} \in H$. Also $h_2 \in H$

$\therefore h_1^{-1} h_2 \in H \Rightarrow ab^{-1} \in H.$

Conversely, suppose $ab^{-1} \in H$ for $a,b \in G$

Then $H=H(ab^{-1})$ [$\because H=Hh$ when $h \in H$]

$\therefore Hb=Hab^{-1}b$

$\Longrightarrow Ha=Hb.$

2.2.3 Theorem: Any two left (right) cosets of a subgroup are either disjoint or identical.

Proof: Let H be any subgroup of G and let aH and bH be two left cosets of H in G.

Suppose that aH and bH are not disjoint.

Let $c\in aH\cap bH$. Then,

$c\in aH$ and $c\in bH$.

Then $c=ah$ for some $h\in H$ and $c= bh^1$ for some $h^1\in H$.

$\Rightarrow ah=bh^1$

$\Rightarrow a=bh^1h^{-1}$

Since H is a subgroup, $h^1h^{-1}\in H$.

$\therefore a=bh_1$ for $h_1=h^1h^{-1}$.

$\therefore aH=(bh_1)H.$

$=b(h_1H)$

But, $h_1H = H$ [because $h_1\in H$]

$\therefore aH=bH$

Thus, $aH\cap bH = \phi$ or $aH=bH$

2.2.4 Theorem: Let H be a subgroup of a group G. Then G is equal to the union of all right cosets of H in G i.e. $G=\bigcup_{a\in G} Ha$

Proof: Since G is a group and H is a subgroup, for $a\in G$, $Ha\subseteq G$.

$\therefore \bigcup_{a\in G} Ha\subseteq G$ ---------- (1)

Let $x\in G$ be arbitrary.

Then, $x.e = x\in Hx$

$\therefore x\in \bigcup_{a\in G} Ha$

$\therefore G\subseteq \bigcup_{a\in G} Ha$ ----------(2)

From (1) and (2)

$G=\bigcup_{a\in G} Ha$

Note: Similarly it can be proved that G is also equal to the union of left cosets of H in G.

2.2.5 Coset decomposition:

We have seen that any two left (right) cosets are either disjoint or identical. Also, the union of all left (right) cosets of a subgroup H of G is equal to G. Hence the set of all left (right) cosets of a subgroup H constitutes a decomposition of G into mutually disjoint classes. As a matter of fact, he partition of a group G into mutually disjoint classes known as "cosets" is accomplished by defining an equivalence relation in G known as **Congruence relation.**

Relation of congruence modulo a subgroup H in a group G:

Definition: If H is a subgroup of a group G and a, b are two elements of G such that $ab^{-1} \in H$. Then we say that a is congruent to b modulo H, and write as $a \equiv b \ mod(H)$.

2.2.6 Theorem: If H is a subgroup of G with $a, b \in G$ then $a \equiv b \ mod(H)$ if and only if $ab^{-1} \in H$ is an equivalence relation.

Proof: The identity element $e \in H$ ($\because H$ is a subgroup)

$\therefore aa^{-1} = e \in H, \forall\ a \in G,$

Thus $a \equiv a (mod\ H), \forall\ a \in G$

$\therefore \equiv$ is reflexive.

Suppose $a \equiv b (mod\ H)$, for $a, b \in G \Leftrightarrow ab^{-1} \in H.$

$\Leftrightarrow (ab^{-1})^{-1} \in H$ [$\because H$ is a subgroup]

$\Leftrightarrow (b^{-1})^{-1}a^{-1} \in H$

$\Leftrightarrow ba^{-1} \in H$

$\Leftrightarrow b \equiv a (mod\ H)$, for $a, b \in G$.

$\therefore \equiv$ is symmetric.

Suppose $a \equiv b (mod\ H)$, $b \equiv c (mod\ H)$, for $a, b, c \in G$

$\Leftrightarrow ab^{-1} \in H$ and $bc^{-1} \in H$

$\Leftrightarrow (ab^{-1})(bc^{-1}) \in H$ [$\because H$ is a subgroup]

$\Leftrightarrow a(b^{-1}b)c^{-1} \in H$

$\Leftrightarrow ac^{-1} \in H$

$\Leftrightarrow a \equiv c (mod\ H)$

$\therefore \equiv$ is transitive.

$\therefore \equiv$ is an equivalence relation.

Definition: The number of distinct left (right) cosets of H in G is called the index of H in G denoted by $[G:H]$

2.3 Lagrange's Theorem

Theorem: If G is a finite group, and H is any subgroup of G, then the order of H divides the order of G.

Proof: Let $°(G)=n$ and $°(H)=m$. We consider the left coset decomposition of G relative to H.

First we show that every left coset aH for $a \in G$ has exactly m elements.

Let $H=\{h_1,h_2,\ldots,h_m\}$, h_i's are distinct.

Consider $aH=\{ah_1,ah_2,\ldots,ah_m\}$

ah_i's are distinct , for if $ah_i=ah_j$ for $i \neq j$.

$\Rightarrow h_i = h_j$, $i \neq j$ which is a contradiction.

$\therefore$ Every left coset aH has exactly m elements.

Since G is finite, the number of left cosets will also be finite. Let k be the number of distinct left cosets.

Then, $G = a_1H \cup a_2H \cup \ldots \cup a_kH$.

$\therefore$ the number of elements in G is equal to the number of elements in the k cosets. Since each coset contains m elements and there are k cosets, we get

$n = km$

i.e. $m|n$.

i.e. $o(H)|o(G)$.

2.4 Normal subgroup

Definition: A subgroup H of a group G is called a normal subgroup of G iff for every $x \in G$ and $h \in H, xhx^{-1} \in H$.

2.4.1 Theorem: A subgroup H of a group G is normal iff $xHx^{-1}=H, \forall\, x \in G$.

Proof: Suppose that $xHx^{-1}=H \; \forall\, x \in G$.

$\Rightarrow xHx^{-1} \subseteq H, \forall x \in G$

Thus for all $h \in H$, $xhx^{-1} \in H, \forall x \in G$.

$\therefore H$ is normal.

Conversely, let H be normal,

Then, $xHx^{-1} \subseteq H, \forall\, x \in G$----------(1)

and

$x^{-1}H(x^{-1})^{-1} \subseteq H, \forall x \in G$

i.e $x^{-1}Hx \subseteq H \; \forall x \in G$---------- (2)

Hence, $x(x^{-1}Hx)x^{-1} \subseteq xHx^{-1}$

i.e. $H \subseteq xHx^{-1}$---------- (3)

From (1) and (3)

$xHx^{-1}=H, \forall x \in G$.

2.4.2 Theorem: A subgroup H of a group G is a normal subgroup iff each right coset of H in G is a left coset of H in G.

Proof: Let H be normal subgroup of G.

Then, $xHx^{-1}=H \;\; \forall\, x \in G$. [By Theorem 2.4.1]

$\therefore (xHx^{-1})\, x=Hx$

$\Rightarrow xH = Hx \; \forall x \in G$

$\therefore$ Every left coset is a right coset.

Conversely, let every left coset be a right coset,

i.e. $xH = Hx \; \forall\, x \in G$

i. e. $(xH)x^{-1} = (Hx)x^{-1}$

i. e. $xHx^{-1}=H, \forall\, x \in G$.

Therefore H is normal in G. [By Theorem 2.4.1]

2.4.3 Theorem: The intersection of any two normal subgroup of a group is a normal subgroup.

Proof: Let N_1 and N_2 be two normal subgroups of G.

For $x \in N_1 \cap N_2$ and $g \in G$.

$\Rightarrow x \in N_1$ & $x \in N_2$, $g \in G$.

$\Rightarrow g\,xg^{-1} \in N_1$ and $g\,xg^{-1} \in N_2$ ($\because N_1$ & N_2 are normal).

$\therefore$ $gx\,g^{-1} \in N_1 \cap N_2$

$\therefore N_1 \cap N_2$ is normal.

2.4.4 Theorem: If G is an abelian group, then every subgroup of G is a normal subgroup.

Proof: Let H be a subgroup of G. Let $a \in G$ be arbitrary and $h \in H$.Then $ha=ah$, so $Ha=aH$, for every $a \in G$, which implies that H is a normal subgroup of G.

2.5 Solved Problems

2.5.1. Find the left cosets of $H= \{0, 3\}$ in the group $(Z_6, +_6)$.

Solution: $Z_6 = \{0, 1, 2, 3, 4, 5\}$

Left cosets of H in G are

$0+H=H$

$1+H= \{1, 4\}$

$2+H= \{2, 5\}$

$3+H= \{3, 0\}$

$4+H= \{4, 1\}$

$5+H= \{5, 2\}$

$\therefore$ Distinct left cosets of H in G are H, $1+H$, $2+H$.

2.5.2. Find the left cosets of $\{P_1, P_5, P_6\}$ in the group $\langle S_3, \bullet \rangle$

Solution: $S_3 = \{ P_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, P_2 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, P_3 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix},$

$P_4 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \quad P_5 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, P_6 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}\}$

Let $H=\{P_1, P_5, P_6\}$

Now left cosets of H in S_3 are,

$P_1 \bullet H = \{ P_1 \bullet P_1, P_1 \bullet P_5, P_1 \bullet P_6\}$

$= \{P_1, P_5, P_6\} = H$

$P_2 \bullet H = \{P_2 \bullet P_1, P_2 \bullet P_5, P_2 \bullet P_6\}$

$= \{P_2, P_3, P_4\}$

$P_3 \bullet H = \{P_3 \bullet P_1, P_3 \bullet P_5, P_3 \bullet P_6\}$

$= \{P_3, P_4, P_2\}$

$P_4 \bullet H = \{P_4 \bullet P_1, P_4 \bullet P_5, P_4 \bullet P_6\}$

$= \{P_4, P_2, P_3\}$

$P_5 \bullet H = \{P_5 \bullet P_1, P_5 \bullet P_5, P_5 \bullet P_6\}$

$= \{P_1, P_6, P_1\}$

$P_6 \bullet H = \{P_6 \bullet P_1, P_6 \bullet P_5, P_6 \bullet P_6\}$

$= \{P_6, P_1, P_5\}$

Therefore, the distinct left cosets of H in S_3 are H, $P_2 \bullet H$

2.5.3 Let $(Z_6, +)$ be a group and $H=\{0, 3\}$ be a subgroup. Is H a normal subgroup?

Solution: W.k.t. $Z_6 = \{0, 1, 2, 3, 4, 5\}$

Left cosets of H in G are

$0+H=H$

$1+H= \{1,4\}$

$2+H= \{2,5\}$

$3+H= \{3,0\}$

$4+H= \{4,1\}$

$5+H= \{5,2\}$

$\therefore$ Distinct left cosets of H in G are H, $1+H$, $2+H$.

Since $(Z_6, +)$ is an abelian group,

$a+H = H+a$ i.e. left coset is equal to right coset.

Thus, H is a Normal subgroup.

2.6 Summary

Definition: Let H be a subgroup of G and $a \in G$. Then the set, $Ha = \{ha / h \in H\}$ is called the right coset of H generated by a. Similarly the set $aH = \{ah / h \in H\}$ is called the left coset of H generated by a.

Theorem: There is one-to-one correspondence between any two right cosets of a subgroup H of a group G.

Theorem: Any two left (right) cosets of a subgroup are either disjoint or identical.

Theorem: Let H be a subgroup of a group G. Then G is equal to the union of all right cosets of H in G i.e. $G=\bigcup_{a \in G} Ha$

Definition: If H is a subgroup of a group G and a, b are two elements of G such that $ab^{-1} \in H$. Then we say that a is congruent to b modulo H, and write as $a \equiv b \ mod(H)$.

Theorem: If H is a subgroup of G with $a, b \in G$ then $a \equiv b \ mod(H)$ if and only if $ab^{-1} \in H$ is an equivalence relation.

Definition: The number of distinct left (right) cosets of H in G is called the index of H in G denoted by $[G:H]$.

2.7 Keywords

Coset, subgroup, normal subgroup.

2.8 Supplementary Problems:

2.8.1. Let $G=Z_8$, for each of the following subgroups H of G, determine all the left cosets of H in G, a) $H = \{[0],[4]\}$ b) $H = \{[0],[2],[4],[6]\}$.

2.8.2. Let G be the group of all non zero real numbers under the operation of multiplication and consider the subgroup $H=\{3^n | \ n \in R\}$ of G. Determine all the left cosets of H in G.

2.8.3. Let N be a subgroup of group G, Prove that N is a normal subgroup of G if and only if $a^{-1}Na \subseteq N$ for all $a \in G$.

2.8.4. Find the right cosets of $H = \{0, 3\}$ in the group $(Z_6, +_6)$.

2.8.5. Let $(Z_6, +)$ be a group and $H=\{0, 3\}$ be a subgroup. Is H a normal subgroup?

2.9 References

1. Discrete mathematical structures with applications to computer science, by Tremblay and Manohar (McGraw-Hill publications).

2. Discrete mathematics for computer science, by Kolman, Busby and Ross, PHI publications.

3. Discrete Mathematical Structures, by N. G. Goudru, Himalaya Publishing House.

4.Topics in Algebra, by I. N. Herstein, Wiley eastern Ltd.

Unit 3: Homomorphism, Isomorphism & Algebraic System

Structure

3.0 Objectives

After going through this lesson you will be able to

- Explain homomorphism and isomorphism;
- Analyse the procedure to establish isomorphism;
- Explain an algebraic system with two binary operations;

3.1 Introduction

The concept of isomorphism shows that two algebraic systems which are isomorphic to one another are structurally indistinguishable and that the results of operations in one system can be obtained from those of the other by simply relabeling the names of the elements and symbols for operations. This concept has useful applications in the sense that the results of one system permit an identical interpretation in the other system.

The algebraic systems with one binary operation like semigroups, monoids, groups are not adequate to describe the system of real numbers. We shall therefore consider an abstract algebraic system called a ring, which is a special case of a group on which an additional binary operation satisfying certain properties could be defined. Other algebraic systems with two binary operations will be obtained by imposing further restrictions on rings.

3.2 Homomorphism and Isomorphism

An isomorphism between two mathematical structures of the same type should preserve the distinguishing features of the structures.

Definition: Let $(S, *)$ & $(T, \bullet)$ be two Semigroups. A mapping $\varphi : (S, *) \rightarrow (T, \bullet)$ is called a semigroup homomorphism if $\varphi(a * b) = \varphi(a) \bullet \varphi(b)$, $\forall\ a, b \in S$

Further φ is called an isomorphism if φ is one-one and onto.

3.2.1 Procedure to establish isomorphism:

To show that the semigroups $(S,*)$ and $(T, \bullet)$ are isomorphic

Step1: Define a mapping $\varphi : S \rightarrow T$ with $\text{Dom}(\varphi) = S$.

Step 2: Show that $\varphi(a*b) = \varphi(a) \bullet \varphi(b)$.

Step3: Show that φ is one to one.

Step 4: Show that φ is onto.

Example 1: Let Z be the set of integers and 2Z be the set of even integers. Show that the semigroups $(Z, +)$ and $(2Z, +)$ are isomorphic.

Solution: We follow the above procedure to show that (Z, +) and $(2Z, +)$ are isomorphic.

Step1: Define a function $\varphi : Z \rightarrow 2Z$ by $\varphi(a) = 2a$.

Step 2: We have, $\varphi(a + b) = 2(a + b)$

$$= 2a+2b$$

$$= \varphi(a) + \varphi(b).$$

Step 3: We show that φ is one-one, suppose that $\varphi(a_1) = \varphi(a_2)$. Then $2a_1 = 2a_2$, so $a_1 = a_2$.

Step4: We show that φ is onto

Suppose that b is any even integer

Then $a = b/2 \in z$ and

$\varphi(a) = \varphi(b/2) = 2b/2 = b$

So φ is onto

Hence $(Z, +)$ and $(2Z, +)$ are isomorphic semigroups.

Definition: Let $(S, *, e_S)$ and $(T, \bullet, e_T)$ be two monoids, where e_S and e_T are identity elements of S and T with respect to the corresponding binary operations * and • respectively. A mapping $\varphi : S \rightarrow T$ is called a monoid homomorphism if

$$\varphi(a * b) = \varphi(a) \bullet \varphi(b), \ \forall \ a, b \in S \text{ and}$$

$$\varphi(e_S) = e_T .$$

Further φ is called an isomorphism if φ is one-one and onto.

3.2.2 Theorem: If f is a homomorphism from a commutative semigroup $(S, *)$ onto a semigroup $(T, \bullet)$, then $(T, \bullet)$ is also commutative.

Proof: Let t_1 and t_2 be any elements of T. Then there exit s_1 and s_2 in S with

$$t_1 = f(s_1) \text{ and } t_2 = f(s_2)$$

Consider,

$t_1 \bullet t_2 = f(s_1) \bullet f(s_2)$

$= f(s_1 * s_2)$

$= f(s_2 * s_1)$

$= f(s_2) \bullet f(s_1)$

$= t_2 \bullet t_1$

Hence, $(T, \bullet)$ is also commutative.

3.2.3 Theorem: Let $(S, *)$, $(T, \bullet)$ and $(V, \oplus)$ be semigroups and $g: S \to T$ & $h: T \to V$ be semigroup homomorphisms. Then $(h \bullet g): S \to V$ is a semigroup homomorphism from $(S, *)$ to $(V, \oplus)$.

Proof: Let $a,b \in S$. Then

$(h \bullet .g)(a*b) = h[g(a*b)]$

$= h[g(a) \bullet g(b)]$

$= h(g(a)) \oplus h(g(b))$

$= (h \circ g)(a) \oplus (h \circ g)(b)$

So, $(h \bullet g): S \to V$ is a semigroup homomorphism

3.3 Group Homomorphism

Definition: Let $(G, *)$ and $(G', \bullet)$ be two groups. Then a mapping $\varphi : (G, *) \to (G', \bullet)$ is a group homomorphism if $\varphi(a * b) = \varphi(a) \bullet \varphi(b)$, $\forall a, b \in G$.

Further φ is called an isomorphism if φ is one-one and onto.

3.3.1 Theorem: If φ is a homomorphism from a group G into a group G^1, then

1. $\varphi(e) = e^1$, where e is the identity in G and e^1 is the identity in G^1.
2. $\varphi(a^{-1}) = [\varphi(a)]^{-1}$, $\forall\ a \in G$.

Proof: 1) Let $a \in G$ then $\varphi(a) \in G^1$.

Consider,

$\varphi(a) \bullet e^1 = \varphi(a)$ $\quad [\because e^1$ is identity in $G^1]$

$= \varphi(a * e)$ $\quad [\because e$ is the identity in $G]$

$= \varphi(a) \bullet \varphi(e)$ [$\because \varphi$ is homomorphism]

Hence, $\varphi(e) = e^1$. [by left cancellation law in G']

2) Let $a \in G$ be arbitrary. Since G is a group $a^{-1} \in G$ and

$$aa^{-1} = e .$$

$$\Rightarrow \varphi(aa^{-1}) = \varphi(e)$$

$$\Rightarrow \varphi(a) \bullet \varphi(a^{-1}) = e^1, \quad [\because \varphi \text{ is homomorphism}]$$

$\therefore \varphi(a^{-1})$ is the inverse of G^1.

$$\therefore [\varphi(a)]^{-1} = \varphi(a^{-1}).$$

Example 2: Let $\varphi : G \to G^1$ be mapping from group G into G^1, defined by $\varphi(a) = e^1, \forall a \in G$. Then φ is a homomorphism. (e^1 is the identity in G^1)

Now for $a,b \in G$

$$\varphi(a * b) = e^1 = e^1 \bullet e^1$$

$$= \varphi(a) \bullet \varphi(b)$$

Example 3: Mapping $\varphi : (Z, +) \to (2Z, +)$ defined by $\varphi(n) = 2n$ is a homomorphism.

Consider

$$\varphi(n+m) = 2(n+m) \qquad \text{for } n, m \in (Z, +)$$

$$= 2n+2m$$

$$= \varphi(n) + \varphi(m).$$

3.4 Kernel of a Homomorphism

If $\varphi : G \to G^1$ is a group homomorphism then the set of all elements of G which are mapped onto the identity of G^1 is called kernel of φ.

i.e $\ker \varphi = \{x \in G / \varphi(x) = e^1\}$

3.4.1 Theorem: The kernel K of a homomorphism φ of G into G^1 is a normal subgroup of G.

Proof: First we show that K is a subgroup of G.

For $a, b \in K$, we have $\varphi(a) = e^1 = \varphi(b)$

Consider,

$$\varphi(ab^{-1}) = \varphi(a) \bullet \varphi(b^{-1}) = \varphi(a) \bullet [\varphi(b)]^{-1}$$
$$= e^1 \bullet (e^1)^{-1} = e^1$$

$\therefore ab^{-1} \in K.$

$\therefore K$ is a subgroup of G.

For $a \in K$ and $x \in G$,

Consider, $\varphi(xax^{-1}) = \varphi(x) \bullet \varphi(a) \bullet \varphi(x^{-1})$ [$\because \varphi$ is a homomorphism]

$= \varphi(x) \bullet e^1 \bullet \varphi(x^{-1})$ [$\because a \in K$]

$= \varphi(x) \bullet \varphi(x^{-1})$

$= \varphi(xx^{-1})$ [$\because \varphi$ is homomorphism]

$= \varphi(e) = e^1$

$\therefore xax^{-1} \in K$

Hence, K is normal in G.

Example 4: The map $\pi: R^2 \rightarrow R$ defined by $\pi(x, y) = x$ is a homomorphism and ker $\pi = R$.

Solution: Consider,

$$\pi[(x_1, y_1) + (x_2, y_2)] = \pi(x_1+x_2, y_1+y_2)$$
$$= x_1+x_2$$
$$= \pi(x_1, y_1) + \pi(x_2, y_2).$$

$\therefore$ π is a homomorphism.

Ker $\pi = \{(x, y) \in R^2 / \pi(x, y)=0\}$.

$= \{(x, y) \in R^2 / x=0\}$.

$= \{(0, y) \in R^2\} = R.$

Example 5: Show that the mapping $\varphi : (Z, +) \rightarrow (2Z, +)$ defined by $\varphi(n)=2n$ is an isomorphism.

Solution: For $n, m \in Z \Rightarrow n+m \in Z.$

$\therefore$ $\varphi(n+m)=2(n+m)$

$=2n+2m$

$= \varphi(n) + \varphi(m)$

$\therefore \varphi$ is a homomorphism.

If $\varphi(x) = \varphi(y)$ for $x, y \in Z$.

$\Rightarrow 2x = 2y$

$\Rightarrow x = y$

$\therefore \varphi$ is one – one.

For each $y \in 2Z$, we can find an element $y/2 \in Z$ such that

$$\varphi(y/2) = y$$

$\therefore \varphi$ is onto

Hence φ is an isomorphism.

Example 6: Determine whether the mapping $\varphi: (R, +) \rightarrow (R^+, \times)$ defined by $\varphi(x) = e^x$ is an isomorphism.

Solution:

For $x, y \in R$

$$\varphi(x+y) = e^{x+y} = e^x . e^y$$

$$= \varphi(x).\ \varphi(y)$$

$\therefore \varphi$ is a homomorphism.

If $\varphi(x) = \varphi(y)$

$e^x = e^y$

$\Rightarrow e^x . e^{-y} = 1$

$\Rightarrow e^{x-y} = 1$

$\Rightarrow x - y = 0$

$\Rightarrow x = y$

$\therefore \varphi$ is one-one.

For any $y \in (R^+, \times)$, $\exists \log_e^y \in R$ such that $\varphi(\log_e^y) = e^{\log y} = y$.

$\therefore$ φ is onto.

$\therefore$ φ is an isomorphism.

3.5 Algebraic System with two Binary Operations

3.5.1 Algebraic System:

A set with one or more n-ary operations on the set is called an algebraic system. We denote an algebraic system by $(S, f_1, f_2, \ldots)$ where S is a non-empty set and $f_1, f_2, \ldots$ are n-ary operations on S.

Example: Any group $(G, *)$ is an algebraic system consisting of a set G and a binary operation $*$.

Definition: An algebraic system $(S, +, \cdot)$ is called a ring if,

(i) $(S, +)$ is an abelian group.

(ii) $(S, \cdot)$ is a semigroup.

(iii) The operator $\cdot$ is distributive over $+$, that is for any $a, b, c \in S$

$a \cdot (b + c) = a \cdot b + a \cdot c$ and

$(b + c) \cdot a = b \cdot a + c \cdot a$.

Example 7: The set of all integers under operation addition $+$, and multiplication $\cdot$, is a ring called ring of integers.

Solution:

(i) From example 6 (Unit 1), $(Z, +)$ is an abelian group.

(ii) Again, from example 4(Unit 1), $(Z, \cdot)$ is a semigroup.

(iii) Multiplication of integers is distributive with respect to addition of integers i. e. for any $a, b, c \in S$,

$a \cdot (b + c) = a \cdot b + a \cdot c$ and

$(b + c) \cdot a = b \cdot a + c \cdot a$.

Therefore $(Z, +, \cdot)$ is a ring.

3.5.2 Special types of rings

Commutative ring: A ring R is commutative if the multiplication operation in R is commutative, that is, for all $a, b \in R$, $ab=ba$

Example: Ring of integers.

Ring with unity element : A ring R is said to be a ring with unity element if R has a multiplicative identity, i.e. if there exist an element in R denoted by 1 such that

$$1 \cdot a = a \cdot 1 = a \quad \forall\, a \in R.$$

Example: (i) Set of all rational numbers

(ii) Ring of integers.

Ring with zero divisors: A ring R is called a ring with zero divisor if there exist elements $a \neq 0$, $b \neq 0$ in R with $ab = 0$. Then we say that a is zero divisor of b and vice versa

Example: Consider, $Z_6 = \{0, 1, 2, 3, 4, 5\}$

Clearly $2 \neq 0, 3 \neq 0 \in Z_6$, but $2 \otimes_6 3 = 0$.

So, 2 is a zero divisor of 3.

Hence Z_6 is a ring with zero divisor.

Ring without zero divisor: A ring R is called a ring without zero divisor if for any $a, b \in R$ with $ab=0$ then either $a=0$ or $b=0$.

Example: Ring of integers

Integral domain: An integral domain is a commutative ring with unity which has no zero divisors.

Example: $(Z, +, \cdot)$, $(Q, +, \cdot)$, $(R, +, \cdot)$.

Field : A commutative ring with unity in which every non-zero element has the multiplicative inverse is called a field.

Example: $(Q, +, \cdot)$, $(C, +, \cdot)$.

3.6 Summary

Let $(S, *)$ & $(T, \bullet)$ be two semigroups. A mapping $\varphi: (S, *) \to (T, \bullet)$ is called a semigroup homomorphism if $\varphi(a * b) = \varphi(a) \bullet \varphi(b)$, $\forall\ a, b \in S$

Further φ is called an isomorphism if φ is one-one and onto.

Theorem: Let $(S, *)$, $(T, \bullet)$ and $(V, \oplus)$ be semigroups and $g: S \to T$ & $h: T \to V$ be semigroup homomorphisms. Then $(h \bullet g): S \to V$ is a semigroup homomorphism from $(S, *)$ to $(V, \oplus)$.

Theorem: If φ is a homomorphism from a group G into a group G^1, then

1. $\varphi(e) = e^1$, where e is the identity in G and e^1 is the identity in G^1.
2. $\varphi(a^{-1}) = [\varphi(a)]^{-1}$, $\forall\ a \in G$.

Definition: If $\varphi: G \to G^1$ is a group homomorphism then the set of all elements of G which are mapped onto the identity of G^1 is called kernel of φ.

Definition: A set with one or more n-ary operations on the set is called an algebraic system.

Definition: An algebraic system $(S, +, \cdot)$ is called a ring if,

(iv) $(S, +)$ is an abelian group.

(v) $(S, \cdot)$ is a semigroup.

(vi) The operator $\cdot$ is distributive over $+$.

3.7 Keywords

Homomorphism, isomorphism, algebraic system.

3.8 Supplementary Problems

3.8.1. What are the steps to be followed to check whether 2 semigroups $(S,*)$ and $(T,*)$ are isomorphic. Show that $(Z, +)$ and $(T, *)$ are isomorphic , where Z is the set of all even integers.

3.8.2. Let G be a group and let a be a fixed element of G. Then show that the function $f: G \to G$ defined by $f(x)=axa^{-1}$, where $x \in G$, is an isomorphism.

3.8.3. Let $(S,*)$ and $(T,*)$ be monoids with identities e and e^1 respectively, Let $f: S \to T$ be an isomorphism. Then prove that $f(e)= e^1$.

3.8.4. Let G be a group under addition and G^1 be a group under multiplication. Let $f: G \to G^1$ be defined by $f(x)=e^x$. Show that f is an isomorphism.

3.8.5. Let $(S_1,*)$, $(S_2,*^1)$, and $(S_1,*^{11})$ be semigroups and let $f: S_1 \to S_2$ and $g: S_2 \to S_3$ be isomorphisms. Show that $gof: S_1 \to S_3$ is an isomorphism.

3.8.6. Prove that the set of all reals, rationals, complex numbers forms a ring under usual addition and multiplication.

3.8.7. Prove that set M_n of all $n \times n$ matrices is a ring with respect to the addition and multiplication of matrices when the elements in matrices are numbers which are members of any ring of numbers.

3.9 References

1. Discrete mathematics for computer science, by Kolman, Busby and Ross, PHI publications.
2. Discrete Mathematical Structures, by N. G. Goudru, Himalaya Publishing House.
3. Discrete mathematical structures with applications to computer science, by Tremblay and Manohar (McGraw-Hill publications).
4. Topics in Algebra, by I. N. Herstein, Wiley eastern Ltd.

Unit 4: Introduction to Coding Theory

Structure

4.0 Objectives:

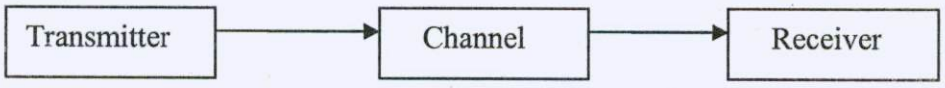
After going through this lesson you will be able to

- Explain the Group codes;
- Differentiate between encoding and decoding functions;
- Analyse the procedure of detecting errors in communication;
- Analyse the procedure of correcting errors in communication

4.1 Introduction:

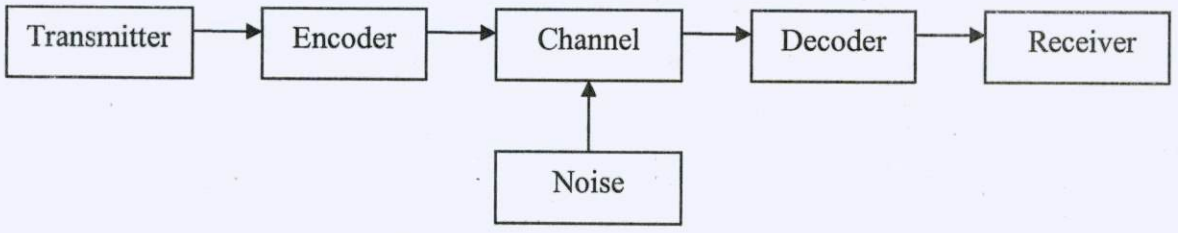
Error-detection and correction techniques have become increasingly important in the design of computer systems. Most systems contain telephone and communication lines which cause transmitted messages to be corrupted by the presence of noise. Peripheral equipment associated with such systems is by far the most unreliable component of these systems and both error detection and error correction are frequently performed. Algebraic structures have been the basis of the most important codes which have been designed.

Communication plays an important role. It takes place in a variety of ways. The three essential parts in an ideal communication system are transmitter, channel and receiver.



In practice, the transmission channel may suffer disturbances, which are called noise, due to weather interference, electrical problems and so on. The important task of a communication system is to minimize the errors in transmission.

A device used to improve the efficiency of communication channel is an encoder. Decoder is a device used to transform the encoded message into original form.



4.2 Encoding function

Message: Message is a basic unit of information. It is a finite sequence of characters from a finite alphabet.

Word: Let B= {0, 1} be the alphabet we choose. Every symbol, we want to transmit is represented as a sequence of m elements from B. Thus, word is a basic unit of information and is a sequence of m 0's &1's.

The set B is a group under the binary operation + mod2.

A group structure can be given to the set of all words, that is, binary strings of length m.

Let $B^m = B \times B \times B \ldots \times B$ (m factors) is a group under the operation $\oplus$ defined by

$$(x_1, x_2, \ldots, x_m) \oplus (y_1, y_2, \ldots y_m) = (x_1+y_1, x_2+y_2, \ldots, x_m+y_m).$$

i) For $(x_1, x_2, \ldots, x_m), (y_1, y_2, \ldots y_m) \in B^m$,

$$(x_1, x_2, \ldots, x_m) \oplus (y_1, y_2, \ldots y_m) = (x_1+y_1, x_2+y_2, \ldots, x_m+y_m) \in B^m$$

$(B^m, \oplus)$ satisfies closure axiom.

ii) $(B^m, \oplus)$ satisfies associative axiom.

iii) $0=(0,0,0\ldots,0)$ is the identity element.

iv) Every element is its own inverse.

Hence $(B^m, \oplus)$ is a group.

Note:

1) An element in B^m is written as $(b_1, b_2, \ldots, b_m)$ or simply as $b_1\ b_2 \ldots b_m$.

2) B^m has 2^m number of elements.

Definition: An (m, n) encoding function is a one to one function e: $B^m \rightarrow B^n$ with $n>m$. For every $b \in B^m$ there exists a distinct $e(b) \in B^n$ called the codeword representing b.

Definition: Let e be an encoding function. We say that the code word $x=e(b)$ has been transmitted with k or fewer errors if the received message x_t and x differ in at least one but no more than k positions.

Definition: Let e: $B^m \rightarrow B^n$ be an (m, n) encoding function. We say that e detects k or fewer errors if whenever $x=e(b)$ is transmitted with k or fewer errors, then x_t is not a code word.

Definition: If $x \in B^n$, then the number of 1's in x is called the weight of x and is denoted by $|x|$.

Example 1: Find the weight of each of the following words in B^7: a) x=010001 b) x= 1110000 c) x=0000000 d) x=1111111.

Solution: a) $|x|=2$ b)$|x|=3$ c) $|x|=0$ d)$|x|=7$

Example 2: Parity check code: The following encoding function e: $B^m \rightarrow B^{m+1}$ is called the parity $(m, m+1)$ check code:

If $b=b_1,b_2,\ldots b_m \in B^m$, define $e(b)= b_1,b_2,\ldots b_m\, b_{m+1}$

where $b_{m+1} = b_{m+1} = \begin{cases} 0 \text{ if } |b| \text{ is even} \\ 1 \text{ if } |b| \text{ is odd} \end{cases}$

To illustrate this encoding function, let m=2, Then,

$e(00)$=000, $e(01)$= 011, $e(10)$=101, $e(11)$=110

Let m=2, Then e:$B^2 \rightarrow B^3$

B^2={00,01,10,11}

To find the elements in B^3

B^3== 0 if $|b|$ is even

1 if $|b|$is odd

Weight of the word 00 = |00|=0, even, so $e(00)$=000

Weight of the word 01 = |01|=1, odd, so $e(01)$=011

Weight of the word 10 = |10|=1, odd, so $e(10)$=101

Weight of the word 11 = |11|=2, even, so $e(11)$=110

The code words in B^3 are {000,011,101,110}

Let m =3, Then e: $B^3 \rightarrow B^4$

B^3= {000,001,010,011,100,101,110,111}

For $b \in B^3$, B_4= 0 if $|b|$ is even

1 if $|b|$is odd

$e(000)$=0000, $e(001)$=0011, $e(010)$=0101

$e(011)=0110$, $e(100)=1001$, $e(101)=1010$

$e(110)=1100$, $e(111)=1111$

The code words in B^4 are {0000, 0011, 0101, 0110, 1001, 1010, 1100, 1111}

Example 3: Consider the (2, 3) parity check code. For each of the received words, determine whether an error will be detected a) 010 b) 110 c) 001 d) 110

Solution: Parity check code is e: $B^2 \rightarrow B^3$

a) Received word=010 $\in B^3$

Word= 01 $\in B^2$

Weight of 01 =|01|=1, odd. So, $e(01)$=011=code word

Since 010≠ 011, the received code word is not equal to the code word. Hence, an error detected.

b) Received word=110 $\in B^3$

Word= 11 $\in B^2$

Weight of 11 =|11|=2, even. So, $e(11)$=110=code word

Since 110=110, the received code word is equal to the code word. Hence, no error detected.

c) Received word=001 $\in B^3$

Word= 00 $\in B^2$

Weight of 00 =|00|=0, even .So, $e(00)$=000=code word

Since 001≠ 000, the received code word is not equal to the code word. Hence, an error detected.

d) Received word=100 $\in B^3$

Word= 10 $\in B^2$

Weight of 10 =|10|=1, odd. So, $e(10)$=101=code word

Since 100≠ 101, the received code word is not equal to the code word. Hence, an error detected.

4.2.1 Parity (m, $3m$) check code

Consider the encoding function $e:B^m \rightarrow B^{3m}$. If $b=(b_1, b_2,\ldots, b_m) \in B^m$. Define $e(b) = e(b_1,b_2,\ldots b_m) = b_1b_2\ldots b_m\ b_1b_2\ldots b_m\ b_1b_2\ldots b_m$.

Example 4: Determine the code words for the parity check code (m, $3m$) where m=2,

Solution: We know that, B^2={00,01,10,11}

Code words

$e(00)= 000000$

$e(10)=101010$

$e(01)=010101$

$e(11)=111111$

Example 5: Consider the $(m,3m)$ encoding function, where $m=2$, for each of the received words, determine whether an error will be detected a) 010100 b)1010101 c)111011 d)111111.

Solution: The encoding function is e: $B^2 \rightarrow B^6$

a) Received word =010100

$e(01)=010101 \neq 010100$

The received code word is not equal to the code word.

Hence error detected.

b) Received word =101010

$e(10)= 101010=101010$

The received code word is equal to the code word.

Hence error cannot be detected.

4.3 Hamming Distance

Definition: Let x and y be words in B^m. The hamming distance $H(x, y)$ between x and y is the weight, $|x \oplus y|$ of $x \oplus y$. Thus the distance between $x= x_1x_2\ldots x_m$ and $y=y_1y_2\ldots y_m$ is the number of positions in which x and y differ.

Example 6: Find the Hamming distance between x and y.

a) $x = 000101, y=010110.$

b) $x = 110110, y = 001100.$,

Solution: a) $x \oplus y = 010011$, so $|x \oplus y|=3$.

b) $x \oplus y = 111010$, so $|x \oplus y|=4$.

4.3.1 Theorem: (Properties of Distance Function)

Let x, y and z are the elements of B^m. Then,

a) $H(x, y) = H(y, x)$

b) $H(x, y) \geq 0$

c) $H(x, y) = 0$ if and only if $x = y$.

d) $H(x, y) \leq H(x, z) + H(z, y)$

Proof:

a) Let $x, y \in B^m$ so $x = (x_1, x_2, \ldots, x_m)$ and $y = (y_1, y_2, \ldots, y_m)$

$$H(x, y) = |x \oplus y|$$
$$= |(x_1, x_2, \ldots, x_m) \oplus (y_1, y_2, \ldots, y_m)|$$
$$= |x_1 \oplus y_1, x_2 \oplus y_2, \ldots, x_m \oplus y_m|$$
$$= |y_1 \oplus x_1, y_2 \oplus x_2, \ldots, y_m \oplus x_m|$$
$$= |y \oplus x|$$

b) $H(x, y) = |x \oplus y|$ is the distance between $x = x_1 x_2 \ldots x_m$ and $y = y_1 y_2 \ldots y_m$. Such that $x_i \neq y_i$. i.e., the number of positions in which x and y differ. Since $x_i, y_i \in \{0,1\}$,

$$H(x,y) = |x \oplus y| \geq 0$$
$$\Rightarrow H(x,y) \geq 0$$

c) If $x_i = y_i$

Let $x_i, y_i \in \{0\}$, then $|x \oplus y| = |(0,0,\ldots,0) \oplus (0,0, \ldots, 0)|$

$$= |(0,0,\ldots,0)|$$
$$= 0.$$
$$\Rightarrow \quad H(x, y) = 0.$$

Let $x_i, y_i \in \{1\}$ then

$|x \oplus y| = |(1, 1, \ldots,1) \oplus (1, 1, \ldots,1)|$ using mod 2 addition

$$= |(0,0,\ldots,0)|$$
$$= 0.$$
$$\Rightarrow \quad H(x, y) = 0.$$

d) For x and y in B^m,

$|x \oplus y| \leq |x| \oplus |y|$

If $z \in B^m$, then $z \oplus z = \bar{O}$, the identity element in B^m.

$H(x, y) = |x \oplus y|$

$= |x \oplus \bar{O} \oplus y|$

$= |x \oplus z \oplus z \oplus y|$

$\leq |x \oplus z| + |z \oplus y|$

$H(x, y) \leq H(x, z) + H(z, y)$

4.3.2 Minimum distance:

The minimum distance of an encoding function $e : B^m \to B^n$ is the minimum of the distances between all distinct pairs of code words; that is

$$\text{Min } \{ H(e(x), e(y)) / x, y \in B^m \}$$

4.3.3 Theorem: An (m, n) encoding function $e : B^m \to\to B^n$ can detect k or fewer errors if and only if its minimum distance is at least $k + 1$.

Proof: Suppose that the minimum distance between any two code words is at least $k+1$. Let $b \in B^m$, and let $x = e(b) \in B^n$ be the code word representing b. Then x is transmitted and is received as x_t. If x_t were a code word different from x, then $H(x, x_t) \geq k + 1$, so x would be transmitted with $k+1$ or more errors. Thus, if x is transmitted with k or fewer errors, then x_t cannot be a code word. This means that e can detect k or fewer errors.

Conversely, suppose that the minimum distance between code words is $r \leq k$, and let x and y be code words with $H(x, y) = r$. If $x_t = y$, that is, if x is transmitted and is mistakenly received as y, then $r \leq k$ errors have been committed and have not been detected. Thus it is not true, that e can detect k or fewer errors.

Example 7: (i) Find the minimum distances of the (2, 5) encoding function e: $B^2 \to B^5$ defined by $e(00)=00000$, $e(10)=00111$, $e(01)=01110$, $e(11)=11111$.

(ii) How many errors will e detect?

Solution: (i). $H(e(00), e(10))=|00000 \oplus 00111|=|00111|=3$.

$H(e(00), e(01))=|00000 \oplus 01110|=|01110|=3$.

$H(e(00), e(11))=|00000 \oplus 11111|=|11111|=5$.

$H(e(10), e(01))=|00111 \oplus 01110|=|01001|=2.$

$H(e(10), e(11))=|00111 \oplus 11111|=|11000|=2.$

$H(e(01), e(11))=|01110 \oplus 11111|=|10001|=2.$

Minimum distance= min{3, 3, 5, 2, 2, 2}=2.

(ii). The minimum distance of e is 2. By Theorem 4.3.3, we have $2 \geq k+1$ or $k \leq 1$. Thus the code can detect one error.

4.4 Group Codes

Definition: An (m, n) encoding function $e : B^m \rightarrow B^n$ is called a group code if

$e(B^m) = \{e(b)/b \in B^m\} = \text{Ran}(e)$ is a subgroup of B^n.

Example 8: Show that the (2, 5) encoding function $e : B^2 \rightarrow B^5$ defined by $e(00)=00000$, $e(01)=01110$, $e(10)=10101$, $e(11)=11011$ is a group code.

Solution:- Let $N = \{00000, 01110, 10101, 11011\}$.

Let $a = 00000$, $b = 01110$, $c = 10101$, $d = 11011$.

$\oplus$	a	b	c	d
a	a	b	c	c
b	b	a	d	c
c	c	d	a	b
d	d	c	b	a

The identity element $a = 00000$ of B^5 belongs to N.

From the table $(N, \oplus)$ is closed.

Every element is its own inverse.

So, the encoding function is a group code.

4.4.1 Theorem: Let e: $B^m \rightarrow B^n$ be a group code. The minimum distance of e is the minimum weight of a non-zero code word.

Proof: Let m be the minimum distance of the group code and let $m = H(x, y)$, where x and y are distinct code words. Let n be the minimum weight of a non-zero code word and suppose that $n = |z|$, for a code word z.

Since, e is a group code, $x \oplus y$ is a non-zero code word.

Then $m = H(x, y) = |x \oplus y| \geq n$ .....(1)

Since the identity element $\bar{O}$ and z are distinct code words.

$n = |z| = |z \oplus \bar{O}| = H(z, \bar{O}) \geq m$ .....(2)

From (1) and (2), $m = n$.

Hence, the minimum distance of e is the minimum weight of a non-zero code word.

4.5 Decoding and Error Correction

Consider an (m, n) encoding function $e: B^m \to B^n$. Once the encoded word $x = e(b) \in B^n$, for $b \in B^m$, is received as the word x_t, we are faced with the problem of identifying the word b that was the original message.

An onto function d: $B^n \to B^m$ is called an (n, m) decoding function associated with e if $d(x_t) = b' \in B^m$ is such that when the transmission channel has no noise, then $b' = b$, that is,

$d \circ e = 1_{B^m}$,

where 1_{B^m} is the identity function on B^m. The decoding function d is required to be onto so that every received word can be decoded to give a word in B^m. It decodes properly received words correctly, but the decoding of improperly received words may or may not be correct.

4.5.1 Decoding functions:

1) Let d: $B^{m+1} \to B^m$ be a $(m+1, m)$ decoding function. If $b = b_1b_2b_3....b_mb_{m+1} \in B^{m+1}$, then $d(b) = b_1b_2b_3.....b_m$. If there is no error, then

$(d \circ e)(b) = d(e(b))$

$= d(x)$

$= b.$

$d \circ e = 1_{B^m}$.

2) Let d: $B^{3m} \to B^m$ be a $(3m, m)$ decoding function.

Then, for $y = y_1, y_2 \ldots y_m, y_{m+1} \ldots y_{2m}, y_{2m+1} \ldots y_{3m}$.

$d(y) = z_1 z_2 \ldots z_m$ where

$$z_i = \begin{cases} 1 \text{ if } \{y_i, y_{i+m}, y_{i+2m}\} \text{ has at least two 1's} \\ 0 \text{ if } \{y_i, y_{i+m}, y_{i+2m}\} \text{ has less than two 1's.} \end{cases}$$

That is, the decoding function d examines the i^{th} digit in each of the three blocks transmitted. The digit that occurs at least twice in these three blocks is chosen as the decoded i^{th} digit.

Example 9: Let d be the (4, 3) decoding function determine $d(y)$ for the word $y \in B^4$

(a) $y = 0110$ (b) $y = 1011$

Solution: a) $y = 0110$

By definition of $(m+1, m)$ decoding function,

$d(b) = b_1, b_2, b_3 \ldots\ldots\ldots\ldots\ldots b_m, b_{m+1} \in B_{m+1}$ where $m=3$

$d(b) = b_1, b_2, b_3$ where $b = d(b) = b_1, b_2, b_3, b_4$.

So $d(y) = d(0110)$ where $y = 0110 \in B^4$

$d(y) = 011$

b) $d(y) = d(1011)$

$d(y) = 101$

Example10: Let d be the (6, 2) decoding function. Determine $d(y)$ for the word y=111011 in B^6.

Solution: y=111011

The received word y has 3 equal blocks like

y =11 10 11

B_1 B_2 B_3

To find z_1, compare the first digits of B_1, B_2 and B_3.

First digit of B_1 is 1

First digit of B_2 is 1

First digit of B_3 is 1

So, first digit of B_1, B_2 and B_3 has at least two 1's. Hence, first digit of z is 1.

To find z_2

Second digit of B_1 is 1

Second digit of B_2 is 0

Second digit of B_3 is 1

So, second digit of B_1, B_2 and B_3 has at least two 1's

Hence, second digit of z is 1

Since $z = z_1 z_2$, z=11.

$d(111011) = 11$

4.5.2 Maximum Likelihood Technique:

Given an (m, n) encoding function e: $B^m \rightarrow B^n$, we often need to determine an (n, m) Decoding function d : $B^n \rightarrow B^m$ associated with e. The maximum likelihood technique is to determine the decoding function d for a given e.

Since, B^m has 2^m elements, there are 2^m code words in B^m we first list that the code words in fixed order:

$x^{(1)}, x^{(2)}, \ldots\ldots\ldots x^{(2m)}$

If the received word is x_t, we compute $H(x^{(i)}, x_t)$ for $1 \leq i \leq 2^m$ and choose the first code word, say it is $x^{(s)}$, such that.

$$\underset{1 \leq i \leq 2^m}{Min}\{H(x^{(i)}, x_t)\} = H(x^{(s)}, x_t)$$

That is, $x^{(s)}$ is code word that is closest to x_t and the first in the list. If $x^{(s)} = e(b)$,we define the maximum likelihood decoding function d associated with e by $d(x_t)=b$. Observe that d depends on the particular order in which the code words in $e(B^m)$ are listed. If the code words are listed in a different order, we may obtain a different maximum likelihood decoding function d associated with e.

Example 11: Let $e: B^2 \rightarrow B^5$ be an encoding function defined by $e(00)$=00000, $e(01)$=01110, $e(10)$=10101, $e(11)$=11011. Decode the following words relative to a maximum likelihood decoding function. (a) 11110 (b) 10011

Solution: Let $x^{(1)}$=00000, $x^{(2)}$=01110, $x^{(3)}$=10101, $x^{(4)}$=11011

(a) Let x_t=11110

$H(x^{(1)}, x_r)$=|100000|+|111101| =|11110|=4

$H(x^{(2)}, x_r)$= |101110|+|111101|=|1100001|=1

$H(x^{(3)}, x_r)$=|110101|+|111101|=|1010111|=3

$H(x^{(4)}, x_r)$= |111011|+|111101|=|1001011|=2.

So, minimum $(H(x^{(i)}, x_r))$ = min{4,1,3,2}=1

Therefore,$(H(x^{(s)}, x_r))$ = min$(h(x^{(s)}, x_r))$

So, $x^{(s)} = x^{(2)} = e^{(b)}$

$x^{(s)} = 01110 = e^{(01)}$

Thus, maximum likelihood encoding word is b=01.

b)Let x_r=10011

$H(x^{(1)}, x_r)$=|00000+10011|=|10011|=3

$H(x^{(2)}, x_r)$=|01110+10011|=|11101|=4

$H(x^{(3)}, x_r)$=|10101+10011|=|00110|=2

$H(x^{(4)}, x_r)$=|11011+10011|=|01000|=1

so, minimum $(H(x^{(i)}, x_r))$=min{3,4,2,1}=1

Therefore,$(H(x^{(s)}, x_r))$=min$(H(x^{(4)}, x_r))$

So,$x^{(s)}=x^{(4)}=e^{(b)}$

$x(s)$=11011=$e^{(11)}$.

Thus, maximum likelihood decoded word is b=11.

4.5.3 Theorem: Suppose that e is an (m, n) encoding function and d is a maximum likelihood decoding function associated with e. Then (e, d) can correct k or fewer errors if and only if the minimum distance of e is at least $2k+1$

Proof: Assume that the minimum distance of e is at least $2k+1$.

Let $b \in B^m$ and $x = e(b) \in B^m$. Suppose that x is transmitted with k or fewer errors, and x_t is received. This means that $H(x, x_t) \leq k$. If z is any other code word, then

$$2k+1 \leq H(x, z) \leq H(x, x_t) + H(x_t, z) \leq k + H(x_t, z).$$

Thus $H(x_t, z) \geq 2k + 1 - k = k + 1$. This means that x is the unique code word that is closest to x_t, so $d(x_t)=b$. Hence (e, d) corrects k or fewer errors.

Conversely, assume that the minimum distance between code words is $r \leq 2k$, and let $x = e(b)$ and $x^1 = e(b^1)$ be code words with $H(x, x^1) = r$. Let $x = b_1\, b_2\, b_3 \ldots b_n$, $x^1 = b_1' b_2' \ldots b_n'$. Then $b_i \neq b_i'$ for exactly r integers, i between 1 and n. Assume, that $b_1 \neq b_1'$, $b_2 \neq b_2'$, …, $b_r \neq b_r'$, but $b_i = b_i'$, when $i > r$.

(a) Suppose that $r \leq k$. If x is transmitted as $x_t = x^1$, then $r \leq k$ errors have been committed but $d(x_t) = b^1$; so (e, d) has not corrected the r errors.

(b) Suppose that $k+1 \leq r \leq 2k$ and let

$y = b_1' b_2' \ldots b_k' b_{k+1} \ldots b_n$

If x is transmitted as $x_t = y$, then $H(x_t, x^1) = r - k \leq k$ and $H(x_t, x) \geq k$.

Thus, x^1 is at least as close to x_t as x is, and x^1 precedes x in the list of code words; so $d(x_t) \neq b$. Then we have committed k errors, which (e, d) has not corrected.

4.6 Summary

Definition: An (m, n) encoding function is a one to one function $e: B^m \rightarrow B^n$ with $n > m$. For every $b \in B^m$ there exists a distinct $e(b) \in B^m$, called the codeword representing b.

Definition: Let $e: B^m \rightarrow B^n$ be an (m, n) encoding function. We say that e detects k or fewer errors if whenever $x = e(b)$ is transmitted with k or fewer errors, then x_t is not a code word.

Definition: Let x and y be words in B^m. The hamming distance $H(x, y)$ between x and y is the weight, $|x \oplus y|$ of $x \oplus y$. Thus the distance between $x = x_1 x_2 \ldots x_m$ and $y = y_1 y_2 \ldots y_m$ is the number of positions in which x and y differ.

Definition: The minimum distance of an encoding function $e : B^m \rightarrow B^n$ is the minimum of the distances between all distinct pairs of code words; that is

$$\text{Min } \{ H(e(x), e(y)) / x, y \in B^m \}$$

Theorem: An (m, n) encoding function $e : B^m \rightarrow B^n$ can detect k or fewer errors if and only if its minimum distance is at least $k + 1$.

Theorem: Let $e: B^m \rightarrow B^n$ be a group code. The minimum distance of e is the minimum weight of a non-zero code word.

Theorem: Suppose that e is an (m, n) encoding function and d is a maximum likelihood decoding function associated with e. Then (e, d) can correct k or fewer errors if and only if the minimum distance of e is at least $2k+1$.

4.7 Keywords

Code word, group code, encoding function, decoding function.

4.8 Supplementary problems

4.8.1. Consider the (2, 3) parity check code. For each of the received words, determine whether an error will be detected a) 100 b) 101 c) 001 d) 110.

4.8.2. Determine the code words for the parity check code $(m, 3m)$ where $m=3$.

4.8.3. Find the minimum distances of the (2, 4) encoding function e: $B^2 \rightarrow B^4$ defined by $e(00)=0000$, $e(10)=0011$, $e(01)=0110$, $e(11)=1111$.

4.8.4. Determine whether the (2, 5) encoding function $e : B^2 \rightarrow B^5$ defined by $e(00)=00000$, $e(01)=01110$, $e(10)=10101$, $e(11)=11011$ is a group code.

4.8.5. Let d be the (6, 2) decoding function. Determine $d(y)$ for the word $y=101011$ in B^6.

4.8.6. Let $e:B^2 \rightarrow B^5$ be an encoding function defined by $e(00)=00000$, $e(01)=01110$, $e(10)=10101$, $e(11)=11011$. Decode the following words relative to a maximum likelihood decoding function. (a)11110 (b) 10011.

4.9 References

1. Discrete mathematics for computer science, by Kolman, Busby and Ross, PHI publications.
2. Discrete Mathematical Structures, by N. G. Goudru, Himalaya Publishing House.
3. Discrete mathematical structures with applications to computer science, by Tremblay and Manohar (McGraw-Hill publications).

Karnataka State Open University

Manasagangothri, Mysore-570006

M. Sc. (Computer Science)

MSC-501: DISCRETE MATHEMATICS

MODULE	UNITS
6	1 to 4

Unit 1: Introduction to Probability

Structure

1.0 Objectives:

After going through this lesson you will be able to

- Explain the classical and axiomatic definitions of probability;
- Explain the conditional probability;
- Give an account of the addition and multiplication rules;
- Analyse the Baye's theorem.

1.1 Introduction:

'Probable' or 'chance' is a word we often come across in our day-to-day life. We say that there is a high chance of raining today; that is we very much expect to have downpour today. This expectation comes from our knowledge about the conditions of the weather. In general the expectation is based on the present knowledge and belief about the system. But, we need a quantitative measure to quantify the expectations. For this, the theory of probability took birth in 17^{th} century in France.

The branch of mathematics which studies the influence of 'chance' is the theory of probability.

Before discussing the mathematical definition of probability, we need to define some of the terms.

1.2 Basic Terminology

Definition: An experiment is any physical action or process that is observed and the result noted.
Ex: Tossing a coin, firing a missile, getting up in the morning.

Definition: An experiment is called a random experiment, if, when repeated under the same conditions, it is such that the outcome cannot be predicted with certainty but all possible outcomes can be determined prior to the performance of the experiment.

Now onwards, an experiment means a random experiment.

Definition: An outcome or event is a result of the experiment.
Ex: In tossing a coin experiment, getting head or tail is an event.

Definition: The collection of all possible outcomes of a random experiment is called the sample space, denoted by S. The elements of a Sample space are called Sample points.
Ex: 1. In the toss of a single coin let the outcome 'head turning up' be denoted by H and the outcome 'tail turning up' be denoted by T. The coin is repeatedly tossed under the same

conditions. Then the toss of the coin results in the outcome H or T. Thus yielding the sample space: $S = \{H, T\}$

2. A die is numbered with 1, 2, 3, 4, 5, 6 on the faces. When this die is thrown the sample space is: $S = \{1, 2, 3, 4, 5, 6\}$

Definition: Any subset E of a sample space S is called an event.

Ex: A die is numbered with 1, 2, 3, 4, 5, 6 on the faces. When this die is thrown the sample space is $S = \{1, 2, 3, 4, 5, 6\}$, $E_1 = \{1, 3, 5\}$ is the event of getting an odd number, $E_2 = \{2, 4, 6\}$ is the event of getting an even number. Clearly E_1 is a subset of S and E_2 is subset of S.

Two events of S are of particular interest: S itself and the empty set ϕ.

Definition: The event S is called the sure event or certain event and the event ϕ is called an impossible event.

Definition: The events, which are favorable to a particular event of an experiment, are called favorable events.

Ex: When a die is rolled, getting 2, 4, or 6 are favorable events to the event 'getting an even number'.

Definition: Events of a random experiment are said to be mutually exclusive, if the occurrence of one event, prevents the occurrence of all the other events i.e., if no two or more of them can occur simultaneously in the same trial.

OR

Events $E_1, E_2, \ldots, E_r, \ldots$ are mutually exclusive if and only if $E_i \cap E_j = \phi$ for $i \neq j$.

Ex: When two teams E_1 and E_2 are playing a game, the events 'E_1 winning the game' and 'E_2 winning the game' are mutually exclusive.

1.3 Classical definition of Probability

If there are n mutually exclusive and equally likely events of a random experiment, out of which 's' events are favorable for a particular event E, then the probability of E is defined as

$$P(E) = s/n = \frac{\text{Number of favorable events with respect to } E}{\text{Total number of events of the experiment}}$$

1.4 Axiomatic definition of probability (Axioms of probability)

Probability is a number that is assigned to each member of a collection of events from a random experiment that satisfies the following properties.

If S is the sample space and E is any event in a random experiment,

1. $0 \leq P(E) \leq 1$ for each event E in S.
2. $P(S) = 1$.
3. If E_1 and E_2 are mutually exclusive events in S, then $P(E_1 \cup E_2) = P(E_1) + P(E_2)$.

The first axiom states that probability is a real number in the interval from 0 to1. The second axiom states that the sample space as a whole is assigned a probability of 1 and this express the idea that probability of a certain event is equal to1. The third axiom states that the probability of the sum of two mutually exclusive events is equal to the sum of their probabilities.

Example 1: A broker feels that the probability that a given stock will go up in value during the day's trading is 0.3 and the probability that it will go down in value is 0.1. What is the probability that it will go up or down?

Solution: Let E_1 be the event that the stock goes up in value, E_2 be the event that the stock goes down in value. The two events are mutually exclusive, since the closing price of the stock can't be both above and below its starting price simultaneously.

Then $P(E_1 \cup E_2) = P(E_1) + P(E_2) = 0.3 + 0.1 = 0.4$.

Example 2: When we roll a pair of balanced dice, what are the probabilities of getting 2, 3 or 12?

Solution: In throwing a pair of balanced dice simultaneously, there is 6×6 elementary events like (1, 1), (1, 2), ..., (1, 6), ..., (6, 6) in the sample space which are equally likely.

Let E_1 = event of getting 2 = {(1, 1)}

E_2 = event of getting 3 = {(1, 2), (2, 1)}; $E_1 \cap E_2 = \Phi$

E_3 = event of getting 12 = {(6,6)}; $E_1 \cap E_3 = \Phi, E_2 \cap E_3 = \Phi$

$P(E_1) = 1/36$, $P(E_2) = 2/36$, $P(E_3) = 1/36$

Clearly A= event of getting 2, (or) 3 (or) 12 = $E_1 \cup E_2 \cup E_3$

$$\therefore \quad P(A) = P(E_1) + P(E_2) + P(E_3)$$
$$= 1/36 + 2/36 + 1/36$$
$$= 4/36$$
$$= 1/9.$$

The third axiom of probability can be extended to include more than two mutually exclusive events.

1.4.1 Theorem: If $E_1, E_2, \ldots, E_n$ are mutually exclusive events, then $P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} P(E_i)$.

Now we are going to see the extension of axiom 3 that allows us to find the probability of the union of *any two events* in S regardless of whether they are mutually exclusive.

1.4.2 Theorem: (General addition rule):

If A and B are any events in S, then $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

When A and B are mutually exclusive, Then $P(A \cap B) = 0$, then the above theorem reduces to the third axiom of probability, that's why, axiom 3 is referred to as the special addition rule.

1.4.3 Theorem: The impossible event has Φ probability zero i.e., $P(\Phi)=0$

Proof: For any event E we have Then $E \cup \Phi = E$, $E \cap \Phi = \Phi$ i.e., E and Φ are mutually exclusive

Then $P(E \cup \Phi) = P(E)$

$\Rightarrow P(E)+P(\Phi)=P(E)$ (by axiom 3)

$\Rightarrow P(\Phi)=0$

Hence the theorem.

1.4.4 Theorem: (Prbability of complement event):

If A is any event in S, then $P(A^c)=1-P(A)$.

Proof: Let A be any event in S, then A and A^c are mutually exclusive events *i. e.*,

$A \cap A^c = \phi$, also $A \cup A^c = S$

$\Rightarrow P(A \cup A^c) = P(S) = 1$

$\Rightarrow P(A) + P(A^c) = 1$ (by axioms 2 and 3)

So that $P(A^c) = 1 - P(A)$.

Hence the theorem.

Example 3: A bag contains 12 balls numbered 1 to 12. If a ball is drawn at random, what is the probability of having a ball with a number which is a multiple of either 2 or 3?

Solution: Let A be an event that the ball number is a multiple of 2 and B be an event that the ball number is a multiple of 3. Then,

$A = \{2, 4, 6, 8, 10, 12\}$, $B = \{3, 6, 9, 12\}$, $A \cap B = \{6, 12\}$

$P(A) = 6/12 = 1/2$, $P(B) = 4/12 = 1/3$, $P(A \cap B) = 2/12$.

Now $P(A \text{ or } B) = P(\text{a ball with a number which is a multiple of either 2 or 3})$

$=P(A)+P(B)-P(A\cap B)$

$= 6/12 + 4/12 - 2/12$

$= 8/12$

$= 2/3.$

Example 4: A fair coin is tossed 5 times. What is the probability of having at least one head?

Solution: In tossing a coin 5 times simultaneously, sample space contains 32 elementary events (therefore $2^5=32$) and S= {TTTTT....., HHHHH}

Let A= the event of getting no head in 5 tosses of a fair coin.

Then, the probability of getting no head in 5 tosses of a coin is given by: $P(A) = 1/32$.

Now, P(getting at least one head)= 1-P(getting no head) $= 1 - P(A)$

$= 1- 1/32$

$=31/32$

1.5 Conditional probability

In many cases, the probabilities of two or more events depend on one another. That means, the occurrence of one event depends on the occurrence of another event.

Example: For a merchant of umbrellas, the probability to get profit on a rainy day is more than the probability of getting profit on any other day. Clearly the event of 'getting profit' depends on the 'event of raining'.

Definition: If A, E are any two events of a sample space S, then the event of "happening of E, after the happening of A" is called conditional event and is denoted by $P(E|A)$.

Conditional probability:

If E and A are any events in S, $P(A)>0$, the **conditional probability** of E given A is

$$P(E \mid A) = \frac{P(E \cap A)}{P(A)}$$

Example 5: A die is rolled. If the outcome is an odd number, what is the probability that it is prime?

Solution: when a die is rolled, the sample space is S = {1, 2, 3, 4, 5, 6}

Let A = event of getting an odd number = {1, 3, 5}

Let E = event of getting a prime number = {2, 3, 5}

Then $E \cap A$ = {3, 5}

Therefore, $P(A) = 3/6 = ½$ and $P(E \cap A)=2/6=1/3$.

P(getting a prime, already which is an odd number)

$= P$(getting a prime | getting an odd number)

$= P(E|A)$

$= \frac{P(E \cap A)}{P(A)}$

$= 1/3 \times 2/1$

$= 2/3$.

1.5.1 Theorem: (General multiplication Rule)

If A and B are any events in S, then

$P(A \cap B) = P(A) \,.\, P(B|A)$ if $P(A)>0$

$= P(B).\, P(A|B)$ if $P(B)>0$

Proof: The above statement can be obtained directly from the definition of conditional probability.

Definition: Two events are said to be independent if the occurrence or non occurrence of one event has no influence on the occurrence or nonoccurrence of the other.

OR

If A and B are any two events in a sample space S. Events A and B such that $P(A|B) = P(A)$ or $P(B|A) = P(B)$ are said to be independent events.

1.5.2 Theorem: (Special multiplication Rule for independent events)

If A and B are independent events, then

$$P(A \text{ and } B)=P(A \cap B)= P(A) \,.P(B)$$

Example 6: What is the probability of getting two heads in two tosses of a balanced coin?

Solution: Since the probability of getting head is ½ for each toss, the two tosses are independent; the probability is ½ x ½ =1/4.

The above rule is sometimes used as definition of independent events.

Example 7: In India approximately 46% of the population has type 'o' Blood. Approximately 39% have a negative *Rh* factor. If a person is selected at random, what is the probability that he or she will have type 'o'-negative blood? (Individuals blood type is independent of *Rh* factor)

Solution: Let A: the blood group is 'o'.

B: the *Rh* factor is negative.

Then,

P (type 'o' and negative Rh factor) $=P(A \text{ and } B)$

$=P(A).P(B)$

$=0.46\times0.39$

$= 0.179.$

That is, approximately 17.9% of all individuals in India have type 'o'-negative blood.

Example 8: We often assume that successful flips of a coin are independent and that the probability of a head in a flip is 0.5. Therefore, the probability of five tosses resulting in the sequence

[head ,head, head, tail, tail] is ½ * ½* ½ * ½* ½ = 1/32

1.6 Baye's Theorem

In some problems, the outcome of an experiment is due to a particular one of the possible 'causes' (causing events) of the outcome. Even though the general multiplication rule is useful in solving these types of problems, we need a systematic developed Formula:

For the purpose of obtaining a formula, let the sample space be divided into 'n' disjoint sets whose union is the sample space S. These events are denoted by $E_1,E_2,\ldots, E_n$. They represent the n possible causes (causing event) of an experimental outcome. Next, let A be an event that occurred when the experiment was performed, and consider the problem of calculating the probability that E_i is the cause of the occurrence of A. In the words of probability, the problem is to calculate the conditional probability $P(E_i|A)$.

1.6.1 Theorem of total probability (or) Rule of elimination

If the event A can occur only along with the event E. Suppose also E can occur only in n mutually exclusive ways $E_1, E_2,\ldots, E_n$. Then

$$P(A) = \sum_{i=1}^{n} P(E_i)P(A \mid E_i)\text{, provided } P(E_i) > 0 \text{ for all } i.$$

Example 9: A certain product is manufactured at two plants I and II. Plant I makes 70% of the requirement and plant II makes 30%. From plant I, 90% meets a particular standard and plant II only 80% meet such standard. Evaluate, out of every 100 items purchasing by a customer, how many will be up to the standard?

Solution: Let E denotes the event that the product is up to standard, $F1$ and $F2$ respectively, denote the events that the item is manufactured by plant I and plant II. Thus, $E = (E \cap F1) \cup (E \cap F2)$

Therefore $P(E) = P(E \cap F1) + P(E \cap F2) = P(F1)\, P(E|F1) + P(F2)\, P(E|F2)$

$$= 0.70 \times 0.90 + 0.30 \times 0.80$$

$$= 0.87.$$

And therefore out of every 100 items purchased, 87 will be up to standard.

Now consider the following formula called 'Baye's Rule', to calculate $P(Ei|A)$

1.6.2 Theorem: (Baye's theorem):

If $E_1, E_2, \ldots, E_n$ are mutually disjoint events with $P(E_i) \neq 0$ for each i then for any arbitrary event A which is a subset of $\bigcup_{i=1}^{n} E_i$ with $P(A) > 0$, we have

$$P(E_i \mid A) = \frac{P(E_i)P(A \mid E_i)}{\sum_{i=1}^{n} P(E_i)P(A \mid E_i)}, \; i = 1 \text{ to } n$$

The probability $P(E_1), P(E_2), \ldots, P(E_n)$ are termed as the 'a priori probabilities'.

The probabilities $P(A|Ei)$, $i = 1$ to n are called 'likelihoods'.

The probability $P(Ei|A)$, $i = 1$ to n are called 'a posteriori probabilities'.

Example 10: Two similar Urns, A, B contain 2 white and 3 red balls, 4 white and 5 red balls respectively. If a ball is selected at random from one of the Urns, then find the probability that the Urn is B, when the ball is red.

Solution:

Urn	White balls	Red balls
A	2	3
B	4	5

The Events of selecting the Urns are equally likely.

Let the event of selecting the first *Urn* A be E_1 and the second *Urn* be E_2.

Therefore, $P(E_1) = P(E_2) = ½$

$P(R|E_1)$ = probability of drawing a red ball from the first *Urn* = 3/5

$P(R|E_2)$ = probability of drawing a red ball from the second *Urn* B = 5/9

From Baye's theorem, $P(E_2 \mid R) = \dfrac{P(E_2)P(R \mid E_2)}{P(E_i)P(R \mid E_1) + P(E_2)P(R \mid E_2)}$

$= 25/52$

Example 11: A factory has three production lines I, II and III contributing 20%, 30% and 50% respectively, to its total output. The percentage of substandard items produced by lines I, II and III are, respectively, 15, 10 and 2. If an item chosen at random from the total output is found to be substandard, what is the probability that the item is from line I?

Solution: Let E_1, E_2, E_3 denote the events of productions from lines I, II and III respectively.

Let A denote the event that an item is substandard.

Then,

$P(E_1)= 0.2$, $P(E_2)= 0.3$, $P(E_3) = 0.5$

$P(A|E_1)= 0.15$, $P(A|E_2)= 0.1$, $P(A|E_3)= 0.02$

By Baye's theorem,

$$P(E_1 \mid A) = \frac{P(E_i)P(A \mid E_1)}{P(E_i)P(A \mid E_1) + P(E_2)P(A \mid E_2) + P(E_3)P(A \mid E_3)}$$

$$= \frac{0.2 * 0.15}{0.2*0.15+0.3*0.10+0.5*0.02}$$

$= 0.069$

1.7 Results:

1.7.1. If A and B are events of a sample space S such that $A \subseteq B$ then $P(A) \leq P(B)$.

Proof: Let $A \subseteq B$. Then $B = A \cup (B-A)$ and A and $(B-A)$ are disjoint events of S.

Hence $P(B) = P(A) + P(B-A)$. (1)

Hence $P(B) \geq P(A)$ (because, $P(B-A) \geq 0$)

Corollary: If $A \subseteq B$, then $P(B-A) = P(B) - P(A)$.

Corollary: Since $(A \cap B) \subset A$ and $(A \cap B) \subset B$, $P(A \cap B) \leq P(A)$, $P(A \cap B) \leq P(B)$

1.7.2. If A_1 and A_2 are two events, then $P(A_1 \cup A_2) \leq P(A_1) + P(A_2)$.

Proof: We have $P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$

$\leq P(A_1) + P(A_2)$ [because $P(A_1 \cap A_2) \geq 0$]

1.7.3. If A, B and C are any three events, then

$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) - P(C \cap A) + P(A \cap B \cap C)$

Proof: Consider $P(A \cup B \cup C) = P[(A \cup B) \cup C)]$

$= P(A \cup B) + P(C) - P[(A \cup B) \cap C)]$

$= [P(A)+P(B)-P(A \cap B)] + P(C) - P[(A \cap C) \cup (B \cap C)]$ [By distributive Law]

$= P(A)+P(B)+P(C)-P(A \cap B) - [P(A \cap C) + P(B \cap C) - P\{(A \cap C) \cap (B \cap C)\}]$

$= P(A)+P(B)+P(C)-P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$,

$[\because (A \cap C) \cap (B \cap C) = A \cap B \cap C]$

$= P(A)+P(B)+P(C)-P(A \cap B) - P(B \cap C) - P(C \cap A) + P(A \cap B \cap C)$

1.8 Solved Problems

1.8.1. From a pack of 52 cards two are drawn, the first being replaced before the second is drawn. Find the probability that the first one is a diamond and second is a king.

Solution: Let A_1 be the event of drawing a diamond.

$P(A_1) = \frac{13}{52} = \frac{1}{4}$

Let A_2 be the event of drawing a king, $P(A_2) = \frac{4}{52} = \frac{1}{13}$

These two events are independent and hence

$P(A_1 \cap A_2) = P(A_1)P(A_2) = \frac{1}{4} * \frac{1}{13} = \frac{1}{52}$

1.8.2. Find the probability that among two-digit numbers formed by 1,…,5 there is no repetition.

Solution: Let (x, y) stand for the number formed by the above digits. Total number of possible cases is 25, since each of x and y can be any of 1,…,5. All these cases are mutually exclusive and equally likely. To find $n(A)$ we note that A occurs if any of the following occurs: (x, y), $x \neq y = 1,\ldots,5$. Suppose x is chosen first and then y. x may be anything between 1 and 5 and thus x may be chosen in 5 ways. Since with each choice of x there are 4 ways of choosing y, the total numbers of cases (x, y), $x \neq y = 1,\ldots, 5$ is $5 \times 4 = 20$. Thus $n(A) = 20$ and hence $P(A) = 20/25 = 4/5$.

1.8.3. If a card is drawn at random from a pack of cards. Find the probability that the card is either a spade or a king.

Solution. Let the event of getting a spade be denoted by A and the event of getting a king be denoted by B

Total number of Spades = 13 $P(A) = 13/52$

Total number of Kings = 4 $P(B) = 4/52$

Number of King Spades = 1 $P(A \cap B) = 1/52$

From the general addition rule,

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$$= \frac{13}{52} + \frac{4}{52} - \frac{1}{52} = \frac{16}{52} = \frac{4}{13}$$

1.8.4. The probabilities that a husband and wife will be alive 30 years from now are given by 0.8 and 0.9 respectively. Find the probability that in 30 years (a) both (b) neither (c) at least one, will be alive.

Solution: Let H, W be the events that the husband and wife respectively will be alive in 30 years. Then $P(H) = 0.8$, $P(W) = 0.9$. We assume that H and W are independent events.

(a) P(both will be alive) $= P(H \cap W) = P(H)P(W) = 0.8 \times 0.9 = 0.72$

(b) P(neither will be alive) $= P(H^C \cap W^C) = P(H^C)\, P(W^C) = 0.2 \times 0.1 = 0.02$

(c) P(atleast one will be alive) = 1- P(neither will be alive)

$= 1 - 0.02 = 0.98.$

1.8.5. Among the workers in a factory only 30% receive bonus. Among those receiving the bonus only 20% are skilled. What is the probability of a randomly selected worker who is skilled and receiving bonus?

Solution: Let A denote the event of receiving bonus and let B denote the event of considering skilled workers

Given $P(A) = 0.3$ and $P(B \setminus A) = 0.2$

We have to find the probability of the event $A \cap B$

$$P(A \cap B) = P(A)\, P(B \setminus A) = 0.3 \times 0.2 = 0.06.$$

1.8.6. The probability that India wins a cricket test match against West Indies is known to be $2/5$. If India and West Indies play 3 test matches what is the probability that (i) India will lose all the three matches (ii) India will win at least one test match (iii) India will win all the tests (iv) India will win at most one match.

Solution: Let A, B, C denote the events that India wins the first, second, third test match against West Indies respectively.

Then

$P(A) = P(B) = P(C) = 2/5$, hence $P(A) = P(B) = P(C) = 3/5$

We note that the three events are independent events.

(i) P(India losing all the three matches) $= P(\overline{A} \cap \overline{B} \cap \overline{C})$

$$= P(\overline{A})\, P(\overline{B})\, P(\overline{C}) = \frac{27}{125}$$

(ii) P(India winning at least one match)

$$= 1 - P(\text{Loosing all the 3 matches})$$

$$= 1 - P(\overline{A} \cap \overline{B} \cap \overline{C})$$

$$= 1 - \frac{27}{125} = \frac{98}{125}$$

(iii) P(India winning all the 3 matches) $= P(A \cap B \cap C) = P(A)\, P(B)\, P(C)$

$$= \frac{2}{5} \times \frac{2}{5} \times \frac{2}{5} = \frac{8}{125}$$

(iv) P(India winning at most one test match) $= P(\text{wins} \leq 1)$

$$= P(\overline{A} \cap \overline{B} \cap \overline{C}) + P(A \cap \overline{B} \cap \overline{C}) + P(\overline{A} \cap B \cap \overline{C}) + P(\overline{A} \cap \overline{B} \cap \overline{C})$$

$$= \frac{27}{125} + P(A)\, P(\overline{B})\, P(\overline{C}) + P(\overline{A})\, P(B)\, P(\overline{C}) + P(\overline{A})\, P(\overline{B})\, P(C)$$

$$= \frac{81}{125}$$

1.8.7. Urn I contains three green and five red balls. Urn II contains two green, one red and two yellow balls. We select an Urn at random and then draw one ball at random from the Urn. What is the probability that we obtain 'green ball'?

Solution: The event 'green ball selected' can occur in one of these two mutually exclusive ways:

(i) Select I Urn and draw a green ball or

(ii) Select II Urn and draw a green ball

P(green) $= P$(I Urn and green) + P(II Urn and green)

$= P$(Urn I) P(green |Urn I) + P(Urn II) P(green|Urn II)

$$= \frac{1}{2} \times \frac{3}{8} + \frac{1}{2} \times \frac{2}{5}$$

$$= \frac{31}{80}.$$

1.8.8. The contents of 3 urns are

Urn I: 1 white, 3 red, 2 black balls

Urn II: 3 white, 1 red, 1 black balls

Urn III: 3 white, 3 red, 3 black balls

Two balls are chosen from a randomly selected Urn. If the balls are 1 white and 1 red ball, what is the probability that they come from Urn II?

Solution: Let A_1 denote that Urn I is chosen

Let A_2 denote that Urn II is chosen

Let A_3 denote that Urn III is chosen

$$P(A_1) = P(A_2) = P(A_3) = \frac{1}{3}$$

Let B denote the event that the balls drawn are 1 white and 1 red ball.

We need to find $P(A_2|B)$.

By Baye's theorem,

$$P(A_2 \mid B) = \frac{P(A_2)P(B \mid A_2)}{P(A_i)P(B \mid A_1) + P(A_2)P(B \mid A_2) + P(A_3)P(B \mid A_3)}$$

We now find $P(B \mid A_i)$ for i=1, 2, 3.

$P(B \mid A_1)$ = Probability of getting 1W and 1R balls in Urn I

$$= \frac{1 \times 3}{6c_2} = \frac{3}{15} = \frac{1}{5}$$

Similarly $P(B \mid A_2) = \frac{1 \times 3}{5c_2} = \frac{3}{10}$ and $P(B \mid A_3) = \frac{3 \times 3}{9c_2} = \frac{9}{36} = \frac{1}{4}$

$\therefore$ $P(A_2|B) = 2/5$.

1.8.9. A box contains 2000 components of which 5% are defective. A second box contains 500 components of which 40% are defective. Two other boxes contain 1000 components each with 10% defective components. We select at random one of the above boxes and draw from it at random a single component.

(i) What is the probability that this component is defective?

(ii) Finding that the selected component is defective, what is the probability that it was drawn from box 2?

Solution: Let A_i = event consisting of all components in the i^{th} box and B = event consisting of all defective components, we have $P(A_1) = P(A_2) = P(A_3) = P(A_4) = \frac{1}{4}$ as the boxes are selected at random

Given $P(B \mid A_1) = 0.05$

$P(B \mid A_2) = 0.4$

$P(B \mid A_3) = 0.1$

$$P(B \mid A_4) = 0.1$$

(i) P(defective component) $= P(B)$

$$= P(A_1)\,P(B \mid A_1) + P(A_2)\,P(B \mid A_2) + P(A_3)\,P(B \mid A_3) + P(A_4)\,P(B \mid A_4)$$

$$= 0.05\left(\frac{1}{4}\right) + 0.4\left(\frac{1}{4}\right) + 0.1\left(\frac{1}{4}\right) + 0.1\left(\frac{1}{4}\right)$$

$$= 0.1625.$$

(ii) Finding that the selected component is defective, the probability that it was drawn from box 2 is to find $P(B \mid A_2)$

Since $P(B) = 0.1625$; $P(B \mid A_2) = 0.4$; $P(A_2) = 0.25$

Gives $P(B \mid A_2) = \frac{P(B \mid A_2)\,P(A_2)}{P(B)}$

$$= \frac{0.4 \times 0.25}{0.1625}$$

$$= 0.615.$$

1.9 Summary:

Classical definition of Probability: If there are n mutually exclusive and equally likely events of a random experiment, out of which 's' events are favorable for a particular event E, then the probability of E is defined as

$$P(E) = s/n = \frac{\text{number of favorable events with respect to } E}{\text{total number of events of the experiment}}$$

Axiomatic definition of probability: If S is the sample space and E is any event in a random experiment,

1. $0 \le P(E) \le 1$ for each event E in S.
2. $P(S) = 1$.
3. If E_1 and E_2 are mutually exclusive events in S, then $P(E_1 \cup E_2) = P(E_1) + P(E_2)$.

General addition rule: If A and B are any events in S, then $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Conditional probability: If E and A are any events in S, $P(A)>0$, the conditional probability of E given A is $P(E \mid A) = \frac{P(E \cap A)}{P(A)}$

General multiplication Rule: If A and B are any events in S, then

$$P(A \cap B) = P(A)\,.\,P(B|A) \text{ if } P(A)>0$$

$$= P(B).\,P(A|B) \text{ if } P(B)>0$$

Baye's theorem: If E_1, E_2,..., E_n are mutually disjoint events with $P(Ei)\neq 0$ for each i then for any arbitrary event A which is a subset of $\bigcup_{i=1}^{n} E_i$ with $P(A) > 0$, we have

$$P(E_i \mid A) = \frac{P(E_i)P(A \mid E_i)}{\sum_{i=1}^{n} P(E_i)P(A \mid E_i)}, \; i = 1 \text{ to } n$$

1.10 Keywords

Event, sample space, conditional event, conditional probability.

1.11 Supplementary problems

1.11.1. If 2 page is randomly selected from 2 book of 100 pages, then find the probability that the sum of the digits of the page is 10.

Solution: $P(S) = \frac{9}{100}$

1.11.2. When 2 dice are rolled simultaneously, find the probability for the sum on the 2 faces will be 10?

Solution: $P(S) = \frac{1}{12}$

1.11.3. When 3 coins are tossed simultaneously, the probability to get at least one head?

Solution: $P(S) = \frac{7}{8}$

1.11.4. If 6 cards are drawn at random from pack of cards, then find the probability to 3 red and 3 black cards.

Solution: P(3R and 3B) = 0.33204

1.11.5. A, B, C are 3 news papers published from a city. 20% of the population read A, 16% read B, 14% read C, 8% both A and B, 5% read both A and C, 4% read B and C, and 2% read all three. Find the percentage of the population who read at least one paper.

Solution: $P(A \cup B \cup C) = 0.35 = 35\%$

1.11.6. The probability for a contractor to get a road contract $\frac{2}{3}$ and to get a building contract is ' $\frac{5}{9}$. The probability to get at least one contract is $\frac{4}{5}$. Find the probability to get both the contract?

Solution: $P(R \cap B) = \frac{19}{45}$

1.11.7. The probability of 3 students to solve 2 problems in mathematics are $\frac{1}{2}$, $\frac{1}{3}$, $\frac{1}{4}$ respectively. Find the probability that the problem is solved.

Solution: $\frac{3}{4}$

1.11.8. Six boys and six girls sit in 2 rows. Find the probability of

<i> all the girls sit together

<ii> all the girls sit together and also boys sit together

Solution: <i> $\frac{6!7!}{12!}$ <ii> $\frac{2(6!)(7!)}{12!}$

1.11.9. The probabilities of A and B to pass an examination are $\frac{2}{10}$, $\frac{3}{10}$. Find the probability that only one of them pass the examination.

Solution: 19/50.

1.11.10. A, B, C are three routes from the house to the office. On any day the route selected by the officer is independent of climate. On a rainy day the probabilities of reaching the office late through these routes are $\frac{1}{25}$, $\frac{1}{10}$, $\frac{1}{4}$ respectively. If in a rainy day the officer is late to the office then find the probability that the route to be B.

Solution: $\frac{10}{39}$

1.11.11. If $P(A \cup B) = 0.65$, $P(A \cap B) = 0.15$, then find the value of $P(\overline{A}) + P(\overline{B})$

Solution: 1.2

1.12 References

1. Probability and statistics for Engineers, by G.S.S. Bhishma Rao. (Scitech publications).
2. Probability and Statistics with Reliability, Queuing and Computer Science Applications, by K.S.Trivedi, PHI publications.

Unit 2: Discrete Probability Distribution

Structure

2.0 Objectives

After going through this lesson you will be able to

- Explain the meaning of a random variable;
- Differentiate pmf and CDF of discrete distribution;
- Explain the mean, and the variance of discrete distribution;
- Explain the Binomial distribution;
- Give an account of the Poisson distribution;
- Analyse the Geometric distribution;
- Analyse the Uniform distribution;

2.1. Introduction

We have already observed that, in a chance experiment, it is often not the actual outcome that concerns us but some quantity that depends upon the outcome. In a random experiment, we may be interested quite often in the numerical measure of the different outcomes.

Through the notion of random variable, we can develop methods for the study of experiments whose outcomes may be described numerically. Besides this convenience, Random variables also provide a more compact description of an experiment. The notion of random variables provides us the power of abstraction and thus allows us to discard unimportant details in the outcome of an experiment. All serious probabilistic computations are performed in terms of random variable. The description of a sample space of an experiment, which gives the information about the events along with their associated probabilities by random variable, can be termed as "Probability distribution" of the concerned experiment (or) of the random variable.

2.2. Basic Terminology

Definition: A random variable is a function that assigns a real number to each sample point in the sample space of a random experiment.

A random variable is denoted by an uppercase letter, such as X and a corresponding lowercase letter; 'x' is used to denote a possible value of X.

In other words, a real valued function X, defined on a sample space S, of a random experiment i.e., $X:S \to R$ is called a random variable.

We refer to the set of possible values of a random variable X as the values of (range of X).

Example 1: In the experiment of tossing two coins, we have the sample space S = {HH, TH, HT, TT}. We assign uniform probability ¼ to each element of S. consider a random variable X

which assigns to each element of S "the number of heads" in that element. Thus X: $S \rightarrow R$ is given by $X(HH) = 2$; $X(HT) = X(TH) = 1$; $X(TT) = 0$ i.e., Number of heads (X):
$\{HH, HT, TH, TT\} \rightarrow \{0, 1, 2\}$

Range of X=$\{0,1,2\}$

Now $X^{-1}(0) = X^{-1}(\text{no head}) = \{ s \in S; X(s) = 0 \} = \{TT\}$

$X^{-1}(1) = \{HT, TH\}$, $X^{-1}(2) = \{HH\}$

Therefore the probabilities of the events are: $P(X = 0) = 1/4$

$P(X = 1) = 1/2$, $P(X = 2) = 1/4$

Result: If X and Y are two random variables defined on the same sample space S and a and b are two real numbers

(i) $aX + bY$ is a random variable,. In particular $X - Y$ is a random variable

(ii) XY is a random variable

(iii) If $X(s) \neq 0$ for all $s \in S$ then $1/X$ is also a random variable

Definition: The event consisting of all outcomes for which $X=x$ is denoted as $\{X = x\}$ and the probability of this event is denoted as $P(X=x)$.

Random variables are usually classified according to the number of values they can take. In general, there are two types.

(i) Discrete random variable

(ii) Continuous random variable

Definition: A discrete random variable is a random variable with a finite (or countably infinite) range.

In other words, a discrete random variable is a random variable that can assume at most a finite or a countably infinite number of possible values.

Example 2: The random variable X = sum of the spots on two dice is discrete, since X can assume only the values 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 and 12. The number of values X can assume is 11 i.e. finite.

Definition: If the range of a random variable X is an interval of real numbers, then X is a continuous random variable.

Example 3: The random variable L= length of time it takes for a computer program to run is continuous. Let us assume that it is reasonable to expect that the value of L is less than four minutes. That is, the value of L lies in the interval (0, 4).

According to the type of random variable, in general we have two types of probability distributions, viz.,

1. Discrete probability distribution
2. Continuous probability distribution.

2.3 Discrete Probability Distribution

Definition: Let X be a discrete random variable or probability mass function (pmf) $f(x)$ for X is given by $f(x) = P(X = x)$ for real x.

The probability function $f(x)$ in the discrete case is a table or an equation that gives the possible values for X, together with the probability that X assumes those values.

Example 4: In tossing three coins, if X = Number of heads, we have the pmf as:

$X = x$	0	1	2	3
$f(x) = P(X = x)$	1/8	3/8	3/8	1/8

2.3.1 Properties that define a pmf

1. $f(x) \geq 0$ for each real number x
2. $\sum_{all\ x} f(x) = 1$

Example 5: Check whether the following can serve as (discrete) probability function

(a) $f(x) = \frac{x-2}{2}$ for $x = 1,2,3,4$

(b) $h(x) = \frac{x^2}{25}$ for $x = 0,1,2,3,4$

Solution:

(a) The function cannot serve as pmf, because $f(1)$ is negative.

(b) The function cannot serve as a pmf, because the sum of the five probabilities is 6/5 >1.

Now coming to the cumulative distribution function $F(x)$ for discrete random variable

Definition: The Cumulative Distribution Function $F(x)$ of a discrete random variable X is defined by $F(x) = P(X \leq x) = \sum_{xi \leq x} f(xi)$, where $f(x)$ is the discrete probability function.

Clearly, by the above definition, $F(x)$ can be obtained from $f(x)$ as follows:

$$F(x) = \begin{cases} 0, & -\infty < x < x_1 \\ f(x_1), & x_1 \leq x < x_2 \\ f(x_1) + f(x_2), & x_2 \leq x < x_3 \\ \quad\quad .... & \\ f(x_1) + ... + f(x_n), & x \geq x_n \end{cases}$$

where $x_1, x_2 \ldots\ldots, x_n$ are the values of the discrete random variable X.

Example 6: The discrete probability function $f(x)$ of X of the random experiment consists of 3 independent tosses of a fair coin

$X = xi$	0	1	2	3
$f(x)=P(X=xi)$	1/8	3/8	3/8	1/8

Then the Cumulative distributive function of X is given by

$$F(x) = \begin{cases} 0, & -\infty < x < 0 \\ 1/8, & 0 \leq x < 1 \\ 4/8, & 1 \leq x < 2 \\ 7/8, & 2 \leq x < 3 \\ 1, & x \geq 3 \end{cases}$$

2.4 General Properties of a discrete probability Distribution:

Now, we discuss some general characteristics of discrete probability of distributions.

Definition: If X is a discrete random variable which takes the values $x_1, x_2, \ldots, x_k$, with the probabilities $f(x_1), f(x_2) \ldots . f(x_k)$, then its mean (expected value) is $\sum_{i=1}^{k} xi\, f(x_i)$ or simply $\mu = \sum_{all\ x} x.f(x)$ where the mean is denoted by μ.

Some authors, use $E(x)$ for expected value. The mean of a probability distribution measures its center in the sense of an average.

Example 7: Let X be the discrete random variable taking the values 1, 2, .. . , 6 with probabilities $f(x_i) = 1/6$ for i = 1 to 6.

Then $\mu = \sum_{i=1}^{6} x_i f(x_i) = \frac{1}{6}(1+2+\ldots+6) = \frac{7}{2}$

Definition: Let X be a discrete random variable. The Variance of X is defined as

$$\sigma^2 = \sum_x x^2 f(x) - \mu^2$$

Example 8:

1) A random variable has the pmf

x	-2	-1	0	1	2	3
f(x)	0.1	k	0.2	$2k$	0.3	k

Find (i)Value of k (ii) Variance (iii)P(X≥2) (iv) P(X<2) (v) P(-1< X <3).

Solution:

(i) $\sum_x f(x)$ =0.1+k+0.2+2k+0.3+k=1

0.6+4k=1

4k=1;

k=1/4.

(ii) Mean

$$\mu = \sum_x xf(x)$$

= (-2)(0.1)+(-1)(0.1)+0(0.2)+1(0.2)+2(0.3)+3(0.1)

= 0.8.

Variance $\sigma^2 = \left[\sum_x x^2 f(x)\right] - \mu^2$

$= (-2)^2(0.1) + (-1)^2(0.1) + (1)^2(0.2) + (2)^2(0.3) + (3)^2(0.1) - (0.8)^2$

= 0.4 +0.1 + 0.2+ 1. 2+ 0.9 - 0.64

= 2.1

(iii) $P(X \geq 2)$

$= P(X=2) + P(X=3)$

$= f(2) + f(3)$

$= 0.3 + 0.1 = 0.4$

(iv) By complementation rule

$P(X<2) = 1 - PX \geq 2)$

$= 1 - 0.4 = 0.6$

(v) $P(-1 < X < 3)$

$= P(X=0) + P(X=1) + P(X=2)$

$= 0.2 + 0.2 + 0.3$

$= 0.7.$

2.5 Binomial Distribution

Binomial Distribution is applicable whenever we have a sequence of trials such that

1. The trials are independent
2. Each trial results in only two possible outcomes, i. e. "success" and "failure".
3. P(success)=p in each trial is a constant.

The random variable X that equals the number of trials that result in success has a binomial distribution with parameters p(probability) and n(number of trials).

Binomial pmf:

$b(x\,;\, n,\, p) = {}^{n}C_x\, p^x (1-p)^{n-x}$ for $x=0, 1, 2, \ldots, n$ with $p+q=1$.

Binomial CDF:

$$F(x) = P(X \leq x) = \sum_{k \leq x} b(k; n, p) = \sum_{0 \leq k \leq x} b(k; n, p) = \sum_{k=0}^{x} b(k; n, p)$$

Example: Four coins are tossed simultaneously. What is the probability of getting (i) 2 heads (ii) At least two heads and (iii) At least one head.

Solution: Let us call the occurrence of heads as success then

p=P(head with single coin)=½ , q=½, n=4.

Since the value of p is constant, and the trials are independent, it is binomial distribution with n=4, p=0.5, q=0.5

i. $P(X=2)={}^4C_2\,(1/2)^2(1/2)^{4-2} = 3/8.$

ii. P(at least two heads)=$P(X\geq 2)$=P(X=2)+P(X=3)+P(X=4)

$$= {}^4C_2\,(1/2)^2(1/2)^{4-2}+{}^4C_3\,(1/2)^3(1/2)^{4-3}+{}^4C_4\,(1/2)^4(1/2)^{4-4}$$

$$=11/16.$$

iii. P(at least 1 head)=1-P(no head)=$1-{}^4C_0\,(1/2)^0(1/2)^{4-0} = 15/16.$

2.5.1 Binomial mean and variance:

$$Mean\ \mu = \sum_x xf(x) = \sum_{x=0}^{n} xb(x;n,p)$$

$$= \sum_{x=0}^{n} x\ {}^nC_x p^x (1-p)^{n-x}$$

$$= \sum_{x=0}^{n} x \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$

$$= \sum_{x=1}^{n} \frac{n!}{(x-1)!(n-x)!} p^x (1-p)^{n-x}$$

$$= np \sum_{x=1}^{n} \frac{(n-1)!}{(x-1)!(n-x)!} p^{x-1} (1-p)^{n-x}$$

Now put $y=x-1$, $\therefore x = y+1$

When $x=1, y=0$; $x=n$, $y=n-1$.

$$\therefore \mu = np \sum_{y=0}^{n-1} \frac{(n-1)!}{y!(n-1-y)!} p^y (1-p)^{n-1-y}$$

$$= np \sum_{y=0}^{n-1} {}^{n-1}C_y\ p^y (1-p)^{n-1-y}$$

$$= np[p+(1-p)]^{n-1}$$

$$\therefore \mu = np.$$

$$Variance\ \sigma^2 = \sum_x x^2 f(x) - \mu^2 = \sum_{x=0}^{n}[x(x-1)f(x) + xf(x)] - \mu^2$$

$$= \sum_{x=0}^{n}[x(x-1)f(x)] + \mu - \mu^2.$$

$$Consider,\ \sum_{x=0}^{n} x(x-1)f(x)$$

$$= \sum_{x=0}^{n} x(x-1)\ {}^nC_x p^x (1-p)^{n-x}$$

$$= \sum_{x=0}^{n} x(x-1)\frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$

$$= \sum_{x=2}^{n} \frac{n!}{(x-2)!(n-x)!} p^x (1-p)^{n-x}$$

$$= n(n-1)p^2 \sum_{x=2}^{n} \frac{(n-2)!}{(x-2)!(n-x)!} p^{x-2} (1-p)^{n-x}$$

$$= n(n-1)p^2 \sum_{y=0}^{n-2} \frac{(n-2)!}{y!(n-2-y)!} p^y (1-p)^{n-2-y}, \quad [Put\ y = x-2]$$

$$= n(n-1)p^2 \sum_{y=0}^{n-2} {}^{n-2}C_y\ p^y (1-p)^{n-2-y},$$

$$= n(n-1)p^2[p + (1-p)]^{n-2}$$

$$= n(n-1)p^2.$$

$$\therefore \sigma^2 = n(n-1)p^2 + \mu - \mu^2$$

$$= np - np^2$$

$$= np(1-p).$$

$$\therefore \sigma^2 = npq.$$

The standard deviation is $\sigma = \sqrt{npq}$

Example 9: When a coin is tossed 200 times, determine the mean and standard deviation?

Solution:

Take n=200, p=½ , q=½.

∴ Mean(μ) = np = 200 × ½ = 100

Variance(σ^2) = npq = 50

∴ Standard deviation (σ) = $\sqrt{50}$ = 7.07

2.6 Poisson Distribution

Poisson distribution is a limiting case of the binomial distribution.

Poisson pmf:

$$f(x;\lambda)=\frac{e^{-\lambda}\lambda^{x}}{x!},\ \ for\ x=0,1,2,\ldots \text{ where } \lambda=np.$$

Poisson CDF:

$$F(x)=P(X\le x)$$

$$=\sum_{k=0}^{x} f(k;\lambda)$$

$$=\sum_{k=0}^{x}\frac{e^{-\lambda}\lambda^{k}}{k!}$$

2.6.1 Poisson mean and variance :

Mean $\mu=\sum_x xf(x)$

$$\mu=\sum_{all\,x}\frac{xe^{-\lambda}\lambda^{x}}{x!}$$

$$=\sum_{x=0}^{\infty}\frac{e^{-\lambda}\lambda^{x}}{(x-1)!}$$

$$=\lambda e^{-\lambda}\sum_{x=1}^{\infty}\frac{\lambda^{x-1}}{(x-1)!}$$

$$=\lambda e^{-\lambda}e^{\lambda}$$

$$=\lambda$$

$\therefore \mu=\lambda.$

Variance:

$$\sigma^2=\sum_x x^2 f(x;\lambda)-\mu^2$$

$$=\sum[x(x-1)f(x;\lambda)+xf(x;\lambda)]-\mu^2$$

$$=\sum x(x-1)f(x;\lambda)+\sum xf(x;\lambda)-\mu^2$$

$$= \sum_{x=0}^{\infty} x(x-1)\frac{e^{-\lambda}\lambda^x}{x!} + \mu - \mu^2$$

$$= \sum_{x=2}^{\infty} \frac{e^{-\lambda}\lambda^x}{(x-2)!} + \mu - \mu^2$$

$$= \lambda^2 e^{-\lambda} \sum_{x=2}^{\infty} \frac{\lambda^{x-2}}{(x-2)!} + \lambda - \lambda^2$$

$$= \lambda^2 e^{-\lambda} e^{\lambda} + \lambda - \lambda^2$$

$$= \lambda^2 + \lambda - l^2$$

$$= \lambda.$$

$$\therefore \sigma^2 = \lambda.$$

Example 10: Passengers arrive at an airport check out counter at an average rate of 1.5 per minute. Find the probabilities that

i. at most 4 will arrive at a given time
ii. at least 3 will arrive during an interval of 2 minutes

Solution:

The arrival of passengers at the check out counter is Poisson distributed with $\lambda=1.5$.

i. $P(X \leq 4) = \sum_{x=0}^{4} \frac{\lambda^x e^{-\lambda}}{x!} = 0.981$

ii. $P(X \geq 3)=1-P(X<3)=1-0.423=0.577$ since $\lambda=1.5 \times 2=3$.

Example 11: If a bank receives on the average $\alpha=6$ bad checks per day. What is the probability that it will receive four bad checks on the given day?

Solution: Substituting $x=4$ and $\lambda=\alpha \times T = 6 \times 1 = 6$

$$P(x=4) = f(4; 6) = \frac{6^4 x e^{-6}}{4!} = 0.134$$

2.7 Geometric Distribution:

Consider a fixed number n of trials, count the number of trials until the first 'success' occurs. If we let 0 denote a failure and let 1 denote a success, then the sample space of this experiment consists of the set of all binary strings with an arbitrary number of 0's followed by a single 1 i. e. $S = \{0^{i-1}\ 1 / i = 1, 2, 3, \ldots\}$.

Let Z be a random variable that represents the number of trials up to and including the first success.

To find the pmf of Z, the event $[Z=i]$ occurs if and only if we have a sequence of i-1 failures followed by one success.

Geometric pmf:

$f(i) = P(Z=i) = q^{i-1}p$, where $q=1-p$.

$= p(1-p)^{i-1}$ for $i = 1, 2, 3\ldots$

Geometric CDF:

$F_2(t) = \sum_{i=1}^{t} p(1-p)^{i-1} = 1-(1-p)^{[t]}$ for $t \geq 0$

2.7.1 Geometric Mean:

$$\mu = \sum_x x f(x)$$

$$\mu = \sum_{i=1}^{\infty} ipq^{i-1} = p\sum_{i=1}^{\infty} iq^{i-1} = p\sum_{i=0}^{\infty} \frac{d}{dq}(q^i) = p\frac{d}{dq}\left(\sum_{i=0}^{\infty} q^i\right) = p\frac{d}{dq}\left[\frac{1}{1-q}\right]$$

$$\mu = \frac{p}{(1-q)^2} = \frac{1}{p}$$

Variance:

$$\sigma^2 = \sum_{i=1}^{\infty} i^2 f(i) - \mu^2$$

$$= \sum_{i=1}^{\infty} i(i-1)f(i) + \sum_{i=1}^{\infty} if(i) - \mu^2$$

$$= \sum_{i=1}^{\infty} i(i-1)q^{i-1}p + \sum_{i=1}^{\infty} if(i) - \mu^2$$

$$= pq\sum_{i=2}^{\infty} i(i-1)q^{i-2} + \mu - \mu^2$$

$$= \frac{2pq}{(1-q)^3} + \frac{1}{p} - \frac{1}{p^2}$$

$$= \frac{2q}{p^2} + \frac{1}{p} - \frac{1}{p^2}$$

$$= \frac{2q+p-1}{p^2}$$

$$= \frac{1-p}{p^2}.$$

2.8 Uniform Distribution:

Let X be a discrete random variable with the range $\{x_1, x_2, x_3, \ldots, x_N\}$

Uniform pmf :

$$f(x) = \begin{cases} 1/N, & \text{if } x \text{ is a value of } X \\ 0, & \text{otherwise} \end{cases}$$

Uniform CDF:

If we let X take on the values $\{1, 2, 3, \ldots, N\}$ with $f(i)=1/N$, $1 \le i \le N$, then its distribution function is given by

$$F(t) = \sum_{i=1}^{t} f(i) = t/N, \qquad 1 \le t \le N$$

2. 8. 1 Uniform mean:

If X takes on the values $\{1, 2, 3, \ldots, n\}$ then

$$\mu = \sum_{i=1}^{n} if(i)$$

$$= \sum_{i=1}^{n} \frac{i}{n} \quad = \frac{1}{n}\sum_{i=1}^{n} i \quad = \frac{1}{n}\frac{n(n+1)}{2}$$

$$\therefore \mu = \frac{n+1}{2}.$$

Uniform Variance:

$$\sigma^2 = \sum_{i=1}^{n} i^2 f(i) - \mu^2$$

$$= \sum_{i=1}^{n} i^2 \frac{1}{n} - \mu^2$$

$$= \frac{1}{n}\sum_{i=1}^{n} i^2 - \mu^2$$

$$= \frac{1}{n}\frac{n(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4}$$

$$= \frac{(n+1)}{2}\left\{\frac{2n+1}{3} - \frac{n+1}{2}\right\}$$

$$= \frac{(n+1)}{2}\left\{\frac{n-1}{6}\right\} \quad = \frac{(n^2-1)}{12}.$$

2.9 Solved Problems

2.9.1. The distribution of a random variable X is given as follows

X = x	1	2	3	4
P(X=x)	1/10	2/10	3/10	4/10

Find the mean and variance of X.

Solution: W.K.T., Mean $\mu = \sum f(x).x$

$$\mu = 1.\frac{1}{10} + 2.\frac{2}{10} + 3.\frac{3}{10} + 4.\frac{4}{10}$$

$$\mu = \frac{1+4+9+16}{10}$$

$$\mu = 3.$$

W.K.T. Variance $\sigma^2 = \sum x^2 f(x) - \mu^2$

$$\sigma^2 = 1.\frac{1}{10} + 4.\frac{2}{10} + 9.\frac{3}{10} + 16.\frac{4}{10} - 9$$

$$\sigma^2 = \frac{1+8+27+64}{10} - 9$$

$$\sigma^2 = \frac{100}{10} - 9$$

$$\sigma^2 = 1$$

2.9.2. The range of a random variable X= {1, 2, 3,...} and the probabilities of X are such that $P(X = k) = \frac{c^k}{k!}$ { k= 1, 2, 3,....}. Find the value of c.

Solution:

$P(X = 1) = \frac{c^1}{1!}$; $P(X = 2) = \frac{c^2}{2!}$; $P(X = 3) = \frac{c^3}{3!}$, ..., $P(X = n) = \frac{c^n}{n!}$.

W.K.T. $\sum P(X=k) = 1$, for k= 1, 2, 3,....

$$\therefore \quad \frac{c^1}{1!} + \frac{c^2}{2!} + \frac{c^3}{3!} + \ldots = 1$$

$$\therefore \quad 1 + \frac{c^1}{1!} + \frac{c^2}{2!} + \frac{c^3}{3!} + \ldots = 2$$

w.k.t. $e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \infty$

$\therefore \quad e^c = 2$. Hence c = log 2.

2.9.3. It has been found that only 60% of the riders of two wheelers put on crash- helmets. Find the probability that

(i) 4 out of 5 will be using their helmets.

(ii) at least 4 out of 5 riders will use their helmets and

(ii) fewer than 4 out of 5 riders will be wearing their helmets.

Solution:

The above problem can be identified with binomial distribution b(x; n, p), where n=5, p=0.60, and if X denotes the number of riders using crash helmet, we have

(i) $P(X= 4) = b(4; 5, 0.6) = {}^5C_4\,(0.6)^4(1-0.6)^{5-4} = 0.26$

(ii) $p(X \geq 4) = P(X= 4)+P(X=5) = b(4;5,0.6)+b(5;5,0.6)$

$=0.26+0.08=0.34$

(ii) $P(X < 4) = 1-P(X\geq 4) = 1-0.34=0.66$

2.9.4. A product is claimed to be 90% free of defects. What is the expected value and standard deviation of the number of defects in a sample 0f 4?

Solution: Here n=4, p(defect)=0.1, q(no defect)=0.9

Mean =E(defect) = $4 \times 0.1=0.4$

s.d = $\sqrt{}$ variance = $\sqrt{(4\times 0.1\times 0.9)} = 0.6$

2.9.5. A car hire firm has two cars which it hires out day by day. The number of demands for a car on each day is distributed as poisson variate with mean 1.5. Calculate the proportion of days on which (i) neither car is used and (ii) some demand is refused.

Solution: Let X be the number of demands.

The proportion of days on which there are x demands for a car

P(x demands in a day) $= \frac{e^{-1.5}1.5^x}{x!}$

Since the number of demands for a car on any day is a poison variate with mean 1.5.

Thus $P(X=x) = \frac{e^{-1.5}1.5^x}{x!}$

(i) Proportion of days on which neither car is used is given by

$$P(X=0) = e^{-1.5} = 0.2231$$

(ii) Proportion of days on which some demand is refused is

$$P(X>2) = 1- P(X\leq 2)$$

$$= 1- [P(X=0) + P(X=1) + P(X=2)]$$
$$= 1- e^{-1.5}(1+1.5 +\frac{(1.5)^2}{21})$$
$$= 1- 0.2231 \times 3.625$$
$$= 0.79126$$

2.9.6. A manufacturer of cotton pins knows that 5% of his product is defective. If he sells cotton pins in boxes of 100 and guarantees that not more than 10 pins will be defective, what is the probability that a box will fail to meet the guaranteed quality?

Solution: We are given n=100.

Let X be the number of defective pins.

Let p= probability of defective pin = 0.05

λ = mean number of defective pins in a box of 100

$= np = 100\times 0.05 = 5$

Since p is small, we use Poisson distribution.

Probability that a box will fail to meet the guaranteed quality is

$$P(X> 10) = 1-P(X\leq 10) = 1- \sum_{x=0}^{10}\frac{e^{-5}5^x}{x!} = 1 - e^{-5}\sum_{x=0}^{10}\frac{5^x}{x!} = 0.0137.$$

2.9.7. The probability of getting no misprint in a page of a book is e^{-4}. Determine the probability that a page of a book contains more than 2 misprints.

Solution: Let X be the number of misprints.

Given $P(X=0) = e^{-4}$.

$$e^{-\lambda} = e^{-4} \Rightarrow \lambda = 4$$

$$P(X>2) = 1-P(X\leq 2)$$
$$= 1-[P(X=0) + P(X=0) + P(X=2)]$$
$$= 1- e^{-4(1+4+8)}$$
$$= 0.762.$$

2.9.8. A telephone switch board receives 20 calls on an average during an hour. Find the probability that during a period of 5 minutes

(a) No call is received

(b) Exactly 3 calls are received.

(c) More than 5 calls are received.

Assume that the time is measured in minutes

Solution: We assume that the number of incoming calls during any time period obeys a Poisson process. Let X be the number of calls received.

20 calls per hour is equivalent to 0.33 calls per minute, which is the mean rate of occurrence. Hence the number of calls in a 5 minute period follows a Poisson distribution with parameter λ=1.65.

(a) P(no call in 5 minute period) = $e^{-1.65} = 0.192$

(b) P(3 calls in a 5 minute period) = $\frac{e^{-1.65}(1.65)^3}{3!} = 0.144$

(c) P(more than 5 calls in a 5 minute period) = $\sum_{x=6}^{\infty} e^{-1.65}\frac{(1.65)^x}{x!} = 0.007$.

2.10 Summary

Definition: A random variable is a function that assigns a real number to each sample point in the sample space of a random experiment.

Definition: A discrete random variable is a random variable with a finite (or countably infinite) range

Definition: Let X be a discrete random variable or probability mass function (pmf). The $f(x)$ for X is given by $f(x) = P(X = x)$ for real x

Properties that define a pmf

1) $f(x) \geq 0$ for each real number x *(2).* $\sum_{all\ x} f(x) = 1$

Definition: The Cumulative Distribution Function $F(x)$ of a discrete random variable X is defined by $F(x) = P(X \leq x) = \sum_{xi \leq x} f(xi)$

Binomial distribution is a discrete probability distribution and its probability distribution function is given by $b(x:n,p)={}^{n}_{x}c\, p^x q^{n-x}$ for x=0,1,2,.....n.

μ=np and σ^2= np(1-p) for binomial distribution.

Poisson distribution is a discrete probability distribution and its probability distribution function is given by

$f(x, \lambda)=\frac{e^{-\lambda}\lambda^x}{x!}$ for x= 0, 1, 2, … and $\mu = \sigma^{2=} \lambda$, the parameter of P.D

Geometric distribution: The pmf is given by $f(i) = P(Z=i) = q^{i-1}p$, *where* $q=1-p$

$$\mu=\frac{1}{p} \quad \text{and} \quad \sigma^2 = \frac{1-p}{p^2}$$

Uniform distribution: The pmf is given by $f(x)= 1/N$ if x is a value of X

$\mu=\frac{n+1}{2}$ and $\sigma^2=\frac{n^2-1}{12}$

2.11 Keywords

Discrete random variable, pmf, CDF, mean, variance.

2.12 Supplementary problems

2.12.1. The range of random variable x is {0, 1, 2}. The probabilities are given by $p(x=0) = 3c^3$, (a) value of C (b) Probabilities $p(x < 1)$, $p(1 < x \leq 2)$, $p(0 < x \leq 3)$.

Solution: $C = {}^1/_3$ or 2.

2.12.2. The distribution of a random variable X is given as follows

$X = x_i$	-2	-1	0	1	2	3
$P(X= x_i)$	0.1	K	0.2	2K	0.3	K

Find the value of K, mean and variance of X.

Solution: K=0.1, Mean=0.8, Variance=2.16

2.12.3. If the probability of a binomial variate X is such that $p(x = 4) = p(x = 2)$ and n=6, then find the parameter p.

Solution: $p = {}^1/_4$

2.12.4. Assuming that the births of the male and female children are equally likely, find the probability for a family of the three children, to have at least one male child.

Solution: ${}^5/_8$

2.12.5. If the difference between the mean and the variance of twp binomial variates is ${}^5/_9$ then find the probability for the event of 2 successes, when the experiment is conducted 5 times.

Solution: $p = {}^1/_3$

2.12.6. An unbiased coin is tossed 8 times. Find the probability of obtaining (a) exactly 2 heads (b) more than 2 heads (c) all heads

Solution: (a) $p(x=2) = {}^7/_{64}$ (b) $p(x>2) = {}^{219}/_{256}$ (c) $p(x = 8) = {}^1/_{256}$

2.12.7. If the probability of a defective bolt is 0.1, Find (a) mean and (b) the standard deviation for the number of defective bolts in a total of 400 bolts.

Solution: (a) Mean =40 (b) S.D = 6

2.12.8. If X is a Poisson variate such that $p(X = 0) = P(X = 1) = K$ then show that $K = {}^1/_e$

2.12.9. A company knows on the basis of past experience, that 2% of its blades are defective. Find the probability of having 3 defective blades, in a sample of 100 blades.

Solution: $P(X = 3) = \frac{e^{-2}(2)^3}{3!} = 0.180447044.$

2.12.10. 20% of the bolts produced in a factory found to be defective. Find the probability that in a sample space of 10 bolts chosen at random exactly two will be defective by using (a) BD (b) Poisson approximation to BD.

Solution: (a) $p(X = 2) = 0.30198$. (b) $p(X = 2)=0.2707$.

2.13 References

1. Probability and statistics for Engineers, by G.S.S. Bhishma Rao. (Scitech publications).

2.Probability and Statistics with Reliability, Queuing and Computer Science Applications, by K.S.Trivedi, PHI publications.

Unit 3: Continuous Probability Distribution

Structure

3.0 Objectives

After going through this lesson you will be able to

- Differentiate pmf and CDF of continuous probability distribution;
- Explain the mean and variance of continuous probability distribution;
- Analyse the Exponential distribution;
- Explain the Normal distribution;
- Give an account of the Uniform distribution;

3.1 Introduction

So far, we have discussed discrete random variables and their distributions. In applications, such random variables denote the number of objects of certain type. Many situations, both applied and theoretical, require the use of random variables that are "continuous" rather than discrete.

The distribution function of a discrete random variable grows only by jumps whereas the distribution function of a continuous random variable has no jumps but grows continuously. Thus, a continuous random variable X is characterized by a distribution function $F_x(x)$ that is a continuous function of x for all x, $-\infty < x < \infty$.

3.2 Density and distribution functions

For a continuous random variable, X, $f(x) = \frac{d\,F(x)}{dx}$ is called the probability density function (pdf) of x, $-\infty<x<\infty$, where $F(x)$ is CDF.

The Cumulative Distribution function (CDF) is given by

$$F_x(x) = P(X=x) = \int_{-\infty}^{x} f(t)\,dt \quad -\infty < x < \infty.$$

The p.d.f, $f(x)$ satisfies the following properties

i) $f(x) \geq 0$, for all x

ii) $\int_{-\infty}^{\infty} f(x)dx = 1$

Properties of CDF:

As in the case for the CDF of a discrete random variable, the CDF of a continuous random variable F(x) satisfies the following properties.

i) $0 \leq F(x) \leq 1$, $-\infty < x < \infty$

ii) F(x) is a monotone non decreasing function of x.

iii) $\lim_{x\to-\infty} F(x) = 0$ and $\lim_{x\to+\infty} F(x) = 1$

iv) $P(X=C)=0$, where C is a constant

v) $P(a \le X \le b) = P(a< X\le b) = P(a\le X < b)$

$= P(a< X< b)$

$= \int_a^b f_x(x)dx$

$= F_x(b) - F_x(a)$

Example 1: Is the function defined by

$$f(x) = \begin{cases} 0, & x < 2 \\ \frac{3+2x}{18}, & 2 \le x \le 4 \\ 0, & x > 4 \end{cases}$$ a probability density function?

Solution:

i) Clearly $f(x) \ge 0$ for every x

ii) $\int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^{2} 0.dx + \int_2^4 \frac{3+2x}{18}\, dx + \int_4^{\infty} 0.dx$

$= 1$

Hence the given function is a density function.

Example 2) Find the cumulative function $F(x)$ corresponding to the density function $f(x) = \frac{1}{x^2+1}$, where $-\infty < x < \infty$

Solution: $F(x) = \int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{x} \frac{1}{t^2+1}\, dt$

$= \tan^{-1} t]_{-\infty}^{x}$

$= \tan^{-1} x + \frac{\pi}{2}.$

3.3 Mean and variance of a continuous probability distribution

The mean and variance of a continuous random variable are defined in a similar manner as that of a discrete random variable. Integration replaces the summation in the definition.

Mean:

Let X be a continuous random variable with probability density function f(x), $-\infty < x < \infty$.

The mean of X is defined as $\mu = \int_{-\infty}^{\infty} x\, f(x)dx$.

Variance:

The variance of X is defined as $\sigma^2 = \int_{-\infty}^{\infty}(x-\mu)^2\, f(x)dx$

Standard deviation:

$\sigma = \sqrt{\left(\int_{-\infty}^{\infty}(x-\mu)^2\, f(x)dx\right)}$

Example 3: If a continuous random variable has the probability density

$$f(x) = \begin{cases} 2e^{-2x} & for\ x > 0 \\ 0 & for\ x \leq 0 \end{cases}$$

Find the mean and the variance of the given probability density.

Solution: Mean $(\mu) = \int_{-\infty}^{\infty} x\, f(x)dx = \int_0^{\infty} x.2e^{-2x}\, dx$

$$= \frac{1}{2}$$

Variance $(\sigma^2) = \int_{-\infty}^{\infty}(x-\mu)^2 f(x)dx$

$$= \int_0^{\infty}(x-\frac{1}{2})^2 e^{2x}\, dx = \frac{1}{4}$$

3.4 Exponential distribution:

This distribution, sometimes called the negative exponential distribution occurs in application such as reliability theory and queuing theory.

The exponential cumulative distribution function is given by:

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & if\ o \leq x < \infty \\ 0, & otherwise \end{cases}$$

If a random variable X has CDF as defined above , we write X ~ EXP(λ).

The p.d.f of X is given by

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & if\ x > 0 \\ 0, & otherwise \end{cases}$$

3.4.1 Markov property of the exponential distribution

Now let us investigate the memoryless or markov property of exponential distribution.
Suppose we know that X exceeds some given value t; that is $x > t$

Ex: let X be the lifetime of a component and suppose we have observed that this component has already been operating for t hours. We may then be interested in the distribution of $Y = X - t$ the remaining lifetime.

Let the conditional probability of $Y \leq y$ given that $x > t$, be denoted by $G_t(y)$. Thus for $y \geq 0$ we have

$$G_t(y) = P(Y \leq y / X > t)$$
$$= P(X - t \leq y / X > t)$$
$$= P(X \leq y + t / X > t)$$
$$= \frac{P(X \leq y+t \text{ and } X>t)}{P(X>t)}, \quad \text{[By the definition of conditional probability]}$$
$$= \frac{P(t \leq X \leq y+t)}{P(X>t)}$$

Thus $G_t(y) = \dfrac{\int_t^{y+t} f(x)dx}{\int_t^{\infty} f(x)\,dx}$

$$= \frac{\int_t^{y+t} \lambda e^{-\lambda x}\,dx}{\int_t^{\infty} \lambda e^{-\lambda x}\,dx}$$
$$= \frac{e^{-\lambda t}(1-e^{-\lambda y})}{e^{-\lambda t}}$$
$$= 1-e^{-\lambda y}$$

Thus $G_t(y)$ is independent of t and is identical to the original exponential distribution of X

3.5 Normal Distribution

Normal distribution is one of the most widely used continuous probability distribution in applications of statistical methods. The normal distribution is of tremendous importance in the analysis and evaluation of every aspect of experimental data in science and medicine. Normal distribution is also referred to as the Gaussian distribution.

Definition: A continuous random variable X is said to have normal distribution with parameter μ and σ^2 if its density function is given by

$$f(x;\mu,\sigma)=\frac{1}{\sigma\sqrt{(2\pi)}}\ e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}\ ,\ -\infty<x<\infty,\ \text{where}\ -\infty<\mu<\infty,\ \sigma>0.$$

If a random variable X has the *p.d.f* as above, then we write $X \sim N(\mu, \sigma^2)$

Definition: A normal distribution with parameters $\mu=0$ and $\sigma = 1$ is called the standard normal distribution denoted by $Z\sim N(0, 1)$.

Standard normal density:

$$f(z;0,1)=\frac{1}{\sqrt{2\pi}}\exp\left(\frac{-z^2}{2}\right),$$

Standard normal CDF:

$$F_Z(z)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z}\exp\left(\frac{-t^2}{2}\right)dt\,.$$

Note:

1) $F_Z(-z)=1-F_Z(z)$.

2) For a value x of a normal random variable X, the corresponding value of the standardized variable Z is $z=(x-\mu)/\sigma$.

3) $F_X(x)=F_Z((x-\mu)/\sigma)$.

3.5.1 Mean and the variance of normal distribution

$$\text{Mean}=\frac{1}{\sigma\sqrt{(2\pi)}}\int_{-\infty}^{\infty}xe^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}\,dx$$

$$\text{Put } Z=\frac{x-\mu}{\sigma}$$

$$dx=\sigma\,dz$$

$$\text{Mean}=\frac{1}{\sqrt{(2\pi)}}\int_{-\infty}^{\infty}(\sigma z+\mu)e^{\frac{-z^2}{2}}dz$$

$$= \frac{\sigma}{\sqrt{(2\pi)}} \int_{-\infty}^{\infty} z e^{\frac{-z^2}{2}}\, dz + \frac{\mu}{\sqrt{(2\pi)}} \int_{-\infty}^{\infty} e^{\frac{-z^2}{2}}\, dz$$

$$= 0 + \frac{\mu}{\sqrt{(2\pi)}} (\sqrt{(2\pi)} \qquad [\because \int_{-\infty}^{\infty} z e^{\frac{z^2}{2}}\, dz = 0]$$

$$[\because \int_{-\infty}^{\infty} e^{\frac{-z^2}{2}}\, dz = \sqrt{(2\pi)}]$$

$$= \mu$$

Variance:

Variance = $\int_{-\infty}^{\infty} x^2 f(x;\, \mu, o^2) dx - \mu^2$

$$= \frac{1}{\sigma\sqrt{(2\pi)}} \int_{-\infty}^{\infty} x^2 e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}\, dx - \mu^2$$

$$= \frac{1}{\sqrt{(2\pi)}} \int_{-\infty}^{\infty} (\sigma z + \mu)^2\, e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}\, dz - \mu^2$$

$$= \frac{\sigma^2}{\sqrt{(2\pi)}} \int_{-\infty}^{\infty} z^2 e^{\frac{-z^2}{2}}\, dz + \frac{2\mu\sigma}{\sqrt{(2\pi)}} \int_{-\infty}^{\infty} z e^{\frac{-z^2}{2}}\, dz + \frac{\mu^2}{\sqrt{(2\pi)}} \int_{-\infty}^{\infty} e^{\frac{-z^2}{2}}\, dz$$

$$= \sigma^2 + 2\,\mu\sigma \times 0 + \mu^2 - \mu^2$$

$$= \sigma^2$$

3.6 Uniform or Rectangular distribution

A continuous random variable X is said to have a uniform distribution over the interval (a, b) if its density is given by:

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & a < x < b \\ 0, & otherwise \end{cases}$$

The distribution function is given by :

$$F(x)=\begin{cases}0, & x<a\\ \dfrac{x-a}{b-a}, & a\le x<b,\\ 1, & x\ge b\end{cases}$$

Uniform Mean:

$\mu = \int_{-\infty}^{\infty} xf(x)dx$

$= \int_{-\infty}^{a} x(0)\, ix + \int_{a}^{b} x\frac{x}{b-a}dx + \int_{b}^{\infty} x(0)dx$

$= 0 + \frac{1}{b-a}\left[\frac{x^2}{2}\right]_a^b + 0$

$= \frac{1}{b-a}\left[\frac{b^2-a^2}{2}\right]$

$= \frac{a+b}{2}$

Uniform variance:

$\sigma^2 = \int_{-\infty}^{\infty} xf(x)dx - \mu^2$

$= \int_{-\infty}^{a} x(0)dx + \int_{a}^{b} x^2\frac{x}{b-a}dx + \int_{b}^{\infty} x^2(0)dx - \mu^2$

$= 0 + \frac{1}{(b-a)}\left[\frac{x^3}{3}\right]_a^b + 0 - \mu^2$

$= \frac{1}{3(b-a)}(b^3 - a^3) - \frac{(a+b)^2}{4}$

$= \frac{1}{3(b-a)}(b-a)(b^2 + ba + a^2) - \frac{(a^2-b^2)+2ab}{4}$

$= \frac{1}{12}\left[\, 4b^2 + 4ab + 4a^2 - 3a^2 - 3b^2 - 6ab \,\right]$

$= \frac{1}{12}[a^2 + b^2 - 2ba]$

$= \frac{(a-b)^2}{12}$

3.7 Solved Problems

3.7.1. Find (i) the constant k such that the function

$$f(x)=\begin{cases}kx^2 & if\ 0<x<3\\ 0 & otherwise\end{cases}$$ is a probability density function

(i)Find the distribution Function $F(x)$

(ii) $P(1< X\leq 2) = ?$

Solution: $f(x)$ is the p.d.f of a continuous random variable X provided

$\int_{-\infty}^{\infty} f(x)dx = 1$. Hence $\int_0^3 kx^2 dx = 1$

$\therefore k\left(\frac{x^3}{3}\right) = 1$. Hence $k = \frac{1}{9}$

(i) $F(x) = \int_{-\infty}^{x} f(t)dt$

Since $f(x) = 0$ if $x \leq 0$, we have $F(x) = 0$ if $x \leq 0$.

Let $0 < x < 3$

Then $F(x) = \int_0^x \frac{t^2}{9}\,dt = \frac{1}{9}\left(\frac{t^3}{3}\right)]_0^x = \frac{x^3}{27}$

Also, $F(x) = 1$, if $x \geq 3$

Hence,

$$F(x) = \begin{cases} 0 & if\ x \leq 0 \\ \frac{x^3}{27} & if\ 0 < x < 3 \\ 1 & if\ x \geq 3 \end{cases}$$

(ii) $P(1< X\leq 2) = \int_1^2 f(x)dx = \int_1^2 \frac{x^2}{9}dx = \frac{x^3}{27}]_1^2 = \frac{8-1}{27} = \frac{7}{27}$

3.7.2. A continuous random variable has the distribution function

$$F(x) = \begin{cases} 0 & if\ x \leq 1 \\ k(x-1)^4 & if\ 1 < x \leq 3 \\ 1 & if\ x > 3 \end{cases}$$

Find (i) k and (ii) the p.d.f $f(x)$

Solution: We know that $f(x) = \frac{d}{dx}F(x)$

$$f(x) = \begin{cases} 0 & if\ x \leq 1 \\ 4k(x-1)^3 & if\ 1 < x \leq 3 \\ 0 & if\ x > 3 \end{cases}$$

We know that $\int_{-\infty}^{\infty} f(x)dx = 1$

$\therefore \int_1^3 f\,dx = 1$. Hence $4k\int_1^3 (x-1)^3 dx = 1$

$\therefore 4k\left[\frac{(x-1)^4}{4}\right]_1^3 = 1$

$\therefore 16k = 1$, Hence $k = \frac{1}{16}$

$$\therefore f(x) = \frac{(x-1)^3}{4} \text{ if } 1 \le x \le 3$$

3.7.3 If X is a normal variate with mean 30 and standard deviation 5, find the probabilities that

(i) $26 \le X \le 40$ (ii) $X \ge 45$

Solution: $\mu=30$, $\sigma = 5$, $Z = \sigma \frac{X-\mu}{\sigma}$

When $x=26$, $z=-0.8$

$x=40$, $z=2$

$x=45$, $z=3$

$P(26 \le X \le 40) = P(-0.8 \le Z \le 2) = F(2) - F(-0.8)$

$= F(2) - 1 + F(0.8)$

$= 0.9772 - 1 + 0.7881$

$= 0.7653$

$P(X \ge 45) = P(Z \ge 3) = 1 - P(Z < 3)$

$= 1 - F(3)$

$= 1 - 0.9987$

$= 0.0013.$

3.7.4 Assume the mean height of soldiers to be 68.22 inches with variance of 10.8 inches. How many soldiers in a regiment of 2000 soldiers would you expect to be atleast six feet tall? Assume heights to be normally distributed.

Solution: Let the variable X denote the height in inches of the soldiers

Mean $(\mu) = 68.22$, $\sigma^2 = 10.8$ so that $\sigma = 3.286$

$X \sim N(68.22, 3.286^2)$

$\therefore P(X > 6 \text{ ft}) = P(X > 72)$ (When $x = 72$, $z = 1.15$)

$= P(Z > 1.15)$

$= 1 - P(Z \le 1.15)$

$= 0.8749$

$\therefore$ The number of solders in the regiment of 2000 soldiers over 6 feet tall is $2000 \times 0.1251 = 250$.

3.8 Summary

For a continuous random variable, X, $f(x) = \frac{d\,F(x)}{dx}$ is called the probability density function of x, where, $-\infty < x < \infty$, $F(x)$ is CDF.

The Cumulative Distribution function (CDF) is given by

$$F_x(x) = P(X=x) = \int_{-\infty}^{\infty} f_x(t)dt, \; -\infty < x < \infty.$$

The p.d.f, $f(x)$ satisfies the following properties:

i) $f(x) \geq 0$, for all x

ii) $\int_{-\infty}^{\infty} f(x)dx = 1$

The exponential cumulative distribution function is given by:

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & if\ o \leq x < \infty \\ 0, & otherwise \end{cases}$$

A continuous random variable X is said to have normal distribution with parameter μ and σ^2 if its density function is given by

$$f(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{(2\pi)}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}, \ -\infty < x < \infty, \text{ where } -\infty < \mu < \infty, \sigma > 0.$$

Definition: A normal distribution with parameters μ=0 and σ = 1 is called the standard normal distribution denoted by Z~N(0, 1).

Standard normal density: $f(z;0,1) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-z^2}{2}\right)$,

Standard normal CDF: $F_Z(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} \exp\left(\frac{-t^2}{2}\right) dt$.

A continuous random variable X is said to have a uniform distribution over the interval (a, b) if its density is given by:

$$f(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & otherwise \end{cases}$$

3.9 Keywords

Continuous random variable, pdf, CDF, mean, variance.

3.10 Supplementary Problems:

3.10.1. Let the continuous random variable X have the p.d.f

$$f(x) = \begin{cases} \frac{2}{x^3} \text{ if } 1 < x < \infty \\ 0 \text{ otherwise} \end{cases}.$$

Find (i) $F(x)$ (ii) P(2<X<10).

3.10.2. The distribution function of a random variable X is $F(x) = \begin{cases} 1-e^{-2x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$.

Find (i) the density function (ii) P(X>3) (iii) P(-2<X<5).

3.10.3. If X is a normal variate with mean 12 and standard deviation 4. Find the probabilities that (i) $0 \leq X \leq 12$ (ii) $X \geq 20$ (iii) $X<20$.

3.10.4. A machine produces bolts which are 10% defective. Find the probability that in a random sample of 400 bolts produced (i) atmost 30 (ii) between 35 and 45 (iii) 55 or more bolts will be defective.

3.10.5 Assume the mean height of soldiers to be 68.22 inches with variance of 10.8 inches. How many soldiers in a regiment of 2000 soldiers would you expect to be six feet tall? Assume heights to be normally distributed.

3.11 References

1. Probability and statistics for Engineers, by G.S.S. Bhishma Rao. (Scitech publications).
2.Probability and Statistics with Reliability, Queuing and Computer Science Applications, by K.S.Trivedi, PHI publications.

Unit 4: Joint Distribution and Correlation

Structure

4. 0 Objectives

After going through this lesson you will be able to

- Explain the functions of random variables;
- Give an account of discrete random vector;
- Analyse the joint distribution of the random variables;
- Explain the correlation;
- Evaluate the correlation coefficient;

4.1 Introduction

So far, we have been concerned with the properties of a single random variable. In many practical problems, however, it is important to consider two or more random variables defined on the same sample space or probability space. So, we discuss discrete random vector and jointly distributed random variables. Also, we are interested in studying the relationship existing among the jointly distributed random variables. To have some understanding of exactly what this relation measures and how to estimate it from observed data, we consider the covariance between the random variables.

4.2 Functions of Random variables

Given the random variable X and its density $f(x)$, we can define $Y= \Phi(X)$ i.e, Y is a function of the random variable X. Further Y is a random variable if Φ is continuous or piecewise continuous.

Example 1: Let $Y= \Phi(X) =X^2$. Here X could denote the measurement error in a certain physical experiment and Y would then be the square of the error.

Y is a random variable, because X^2 is a continuous function. Let the CDF of Y be denoted by $F_Y(y)$, where y is a value of Y. Let us find $FY_y(y)$. Note that $F_Y(y) =0$ for $y\leq 0$, because always $X^2\geq 0$.

For $y>0$, $F_y(y) =P(Y \leq y)$

$$=P(X^2\leq y)$$

$$= P(-\sqrt{y} \leq X \leq \sqrt{y})$$

$$= F_X(\sqrt{y}) - F_X(-\sqrt{y})$$

Differentiating on both the sides

$$\frac{d}{dy}F_Y(y) = \frac{d}{dy}F_X(\sqrt{y}) - \frac{d}{dy}F_X(-\sqrt{y})$$

$$f_Y(y) = \frac{1}{2\sqrt{y}}f_X(\sqrt{y}) + \frac{1}{2\sqrt{y}}f_X(-\sqrt{y})$$

There fore, the density of Y is

$$f_Y(y)=\begin{cases}\frac{1}{2\sqrt{y}}f_X(\sqrt{y})+\frac{1}{2\sqrt{y}}f_X(-\sqrt{y}), & for\ y>0\\ 0, & otherwise\end{cases}$$

Example 2: Let X be uniformly distributed on (0, 1), we show that $Y=-\lambda^{-1}\log(1-X)$ has an exponential distribution with parameter $\lambda>0$.

Solution: We know that log is a non-negative function, so Y is a non-negative random variable, so that $F_Y(y)=0$, for $y\le 0$. For $y>0$, we have

$$\begin{aligned}F_y(y)&=P(Y\le y)\\ &=P(-\lambda^{-1}\log(1-X)\le y)\\ &=P(\log(1-X)\ge -\lambda y), && [by\ multiplying\ by -\lambda]\\ &=P((1-X)\ge e^{-\lambda y}), && [since\ e^x\ is\ an\ increasing\ function\ of\ x]\\ &=P(X\le 1-e^{-\lambda y})\\ &=F(1-e^{-\lambda y}).\end{aligned}$$

Since X is uniform over (0, 1),

$F_X(x)=x, for\ 0\le x\le 1.$

$$\begin{aligned}Thus,\ F_Y(y)&=F_X(1-e^{-\lambda y})\\ &=1-e^{-\lambda y}.\end{aligned}$$

Hence, $F_Y(y)=\begin{cases}1-e^{-\lambda y}, & for\ y>0\\ 0, & Otherwise\end{cases}$

$\therefore Y\sim EXP(y)$

The distribution of y is exponential with parameter λ.

Example 3: Let X be uniformly distributed on (0, 1), find the distribution of $Y=aX+b$

$$\begin{aligned}F_Y(y)=P(Y\le y)&=P(aX+b\le y)\\ &=P\left(X\le y-\frac{b}{a}\right)\\ &=F_X\left(y-\frac{b}{a}\right)\\ &=y-\frac{b}{a} && \therefore X\text{ is uniformly distributed in (0, 1) i.e.., } F_X(x)=x\\ &=y-\frac{b}{a}+b-b.\end{aligned}$$

$\therefore Y$ is uniformly distributed in $(b, a+b)$.

4.3 Discrete Random Vector

Definition: Let $X_1, X_2, \ldots, X_r$ be r discrete random variables defined on a sample space S. for each element s in S, let $X_1(s) = X_1, X_2(s) = X_2, X_r(s) = X_r$. Then $X = (X_1, X_2, \ldots, X_r)$ then X is an r- dimensional function from S to R^r with $X(s) = x = (x_1, x_2, \ldots, x_r)$. X is called a discrete random vector.

Definition: The compound (or joint) pmf for a discrete random vector X is defined to be

$$p_x(x) = P(X=x)$$
$$= P(x_1 = X1, X_2, \ldots\ldots\ldots x_r = x_r).$$

The compound pmf has the following four properties

(1) $P_X(x) \geq 0, x \in R^r$

(2) $\{x \mid p_x(x) \neq 0\}$ is a finite or countable infinite subset of R^r which is denoted by $\{x_1, x_2, \ldots\}$

(3) $P(X \in A) = \sum_{x \in A} p_x(x)$

(4) $\sum_j p_x(x_j) = 1$

Example 4: Consider a program with two modules and module execution times X & Y respectively. Let the images of X & Y be $\{1, 2\}$ & $\{1, 2, 3, 4\}$. The compound pmf is described by the table.

	$y=1$	$y=2$	$y=3$	$y=4$
$x=1$	$\frac{1}{4}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{8}$
$x=2$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{16}$

4.3.1 Marginal pmf:

In situations where we are concerned with more than one random variable, the pmf of a single variable, such as $p_x(x)$ is referred to as a marginal pmf . The marginal pmf is computed as $p_x(x) = \sum_j p_{x,y}(x, y_j)$.

Similarly $p_y(y) = \sum_i p_{x,y}(x_i, y)$

In the above example,

$$p_x(1) = p(1,1) + p(1,2) + p(1,3) + p(1,4)$$
$$= 1/4 + 1/16 + 1/16 + 1/8$$
$$= 1/2$$
$$p_x(2) = p(2,1) + p(2,2) + p(2,3) + p(2,4)$$
$$= 1/16 + 1/8 + 1/4 + 1/16$$

$= 1/2$

$P_y(1) = p(1,1) + p(2,1) = 1/4 + 1/16 = 5/16$

$P_y(2) = p(1,2) + p(2,2) = 1/16 + 1/8 = 3/16$

$P_y(3) = p(1,3) + p(2,3) = 1/16 + 1/4 = 5/16$

$P_y(4) = p(1,4) + p(2,4) = 1/8 + 1/16 = 3/16$

Example 5: Two discrete random variables X & Y have joint pmf given by

	Y			
X		1	2	3
	1	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{12}$
	2	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{12}$
	3	$\frac{1}{12}$	$\frac{1}{12}$	0

Compute the probability of each of the following events a) $X \leq 1$ ½ b) X is odd c) XY is even d) Y is odd given that is odd

Solution: a) $P(X \leq 1\ ½) = P(X = 1)$

$= p_x(1)$

$= p(1,1) + p(1,2) + p(1,3)$

$= \frac{1}{12} + \frac{1}{6} + \frac{1}{12}$

$= \frac{1}{3}$

b) P (X is odd) $= P(X = 1) + P(X = 3)$

$= p_x(1) + p_x(3)$

$= \frac{1}{3} + \frac{1}{12} + \frac{1}{12} + 0$

$= \frac{4}{6} = \frac{2}{3}$

c) P (XY is even) $= p(1,2) + p(2,1) + p(2,2) + p(2,3) + p(3,2)$

$= \frac{1}{6} + \frac{1}{6} + \frac{1}{4} + \frac{1}{12} + \frac{1}{12}$

$= \frac{3}{4}$

d) $P(Y)$ is odd $|$ X is odd) $= \frac{P(Y \text{ is odd } \& X \text{ is odd})}{P(X \text{ is odd})}$

$$= \frac{p(1,1)+p(1,3)+p(3,1)+p(3,3)}{\frac{2}{3}}$$

$$= \frac{1}{12}+\frac{1}{12}+\frac{1}{12} \div \frac{2}{3}$$

$$= \frac{3}{12} \times \frac{3}{2} = \frac{3}{8}$$

Example 6: Let X_1 and X_2 have the joint probability distribution given below:

	X_1			
X_2		0	1	2
	0	.1	.4	.1
	1	.2	.2	0

Find (i) $P(X_2 + X_1>1)$ (ii) Marginal distribution of X_1.

(iii) Find the conditional distribution X_1 given $X_2 =1$.

(iv)Are X_1, X_2 independent?

Solution:

(i) $P(X_1 + X_2>1) = P(X_1=1 \text{ and } X_2=1) + P(X_1=2 \text{ and } X_2=0) + P(X_1=2 \text{ and } X_2=1)$

$= 0.2+0.1+0$

$=0.3.$

(ii) $P(X_1=0) = 0.1+0.2=0.3$

$P(X_1=1)= 0.4+0.2=0.6$

$P(X_1=2)= 0.1+0=0.1$

(iii)

$$P(X_1 \mid X_2 = 1) = \frac{P(X_1 \text{ and } X_2 = 1)}{P(X_2 = 1)}$$

$$= \frac{0.2+0.2+0}{0.2+0.2+0}$$

$$= 1$$

(iv) X_1 and X_2 are independent if their joint pmf is the product of their marginal pmfs i.e.,

$P_{X,Y}(x,y) = P_X(x)\, P_Y(y)$ for all x and y.

$P(0,0) = 0.1$

$P_{X1}(0) = 0.1+0.2=0.3$

$P_{X2}(0) = 0.1+0.4+0.1=0.6$

But $P(0,0) \neq P_{X1}(0)\, P_{X2}(0)$

Therefore X_1 and X_2 are not independent.

4.4 Jointly distributed random variables

Let X & Y be two random variables defined on the same sample space S. The event $[X \leq x, Y \leq y] = [X \leq x] \cap [Y \leq y]$ consists of all sample points s € S such that X(s) ≤ x & $Y(s) \leq y$.

Definition: The joint (or compound) distribution function of random variables X&Y is defined by $F_{X,Y}(x,y)=P(X \leq x, Y \leq y)$, $-\infty < x < \infty$, $-\infty < y < \infty$.

The subscripts will be dropped whenever the two random variables under consideration are clear from the context; that is $F_{xy}(x, y)$ will be written as $F(x, y)$ such a function satisfies the following properties.

1) $0 \leq F(x, y) \leq 1$, $-\infty < x < \infty$, $-\infty < y < \infty$.
2) If either x or y approaches $-\infty$ then $F(x, y)$ approaches 0 and if both x & y approach $+\infty$ then $F(x, y)$ approaches 1.
3) $P(a < X \leq b \ \& \ c < Y \leq d) = F(b, d) - F(a, d) - F(b, c) + F(a, c)$.

4.4.1 Marginal Distribution functions

Given the joint distribution function the marginal distribution functions of X & Y are given by $F_X(x)_{xy}(x,y)$ and $F_Y(y) = \lim_{x \to \infty} F_{xy}(x,y)$.

4.4.2 Joint or Compound Probability density function

If X & Y are continuous random variables then we can find a function $f(x, y)$ such that

$$F(x, y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f(u, v)\, du\, dv$$

Such a function $f(x, y)$ is called the joint or the compound probability density function of X & Y

Note: - 1) $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(u,v)\,du\,dv = 1$

2) $p(a < X \leq b, c < Y \leq d) = \int_{a}^{b} \int_{c}^{d} f(x,y)\,dy\,dx$

4.4.3 Marginal Density function

If f(x, y) is the joint density of the c random variables X & Y then the marginal densities $f_x(x) = \int_{-\infty}^{\infty} f(x,y)dy$ and $f_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx$

4.4.4 Independent random variables:

We define two random variables X & Y to be independent if

$F(x, y) = F_x(x)\, F_y(y)$, $-\infty < x < \infty$, $-\infty < y < \infty$

Thus the independence of random variables X & Y implies that their joint CDF factors into the product of the marginal CDF's.

This definition is applicable regardless of the types of the random variables involved.

In the case that X & Y are discrete the above definition of independence is equivalent to

$$p(x, y) = p_x(x)\, p_y(y).$$

In the case that X & Y are continuous random variables having a joint PDF the above definition of independence is equivalent to

$$f(x, y) = f_x(x)\, f_y(y), -\infty < x < \infty, -\infty < y < \infty.$$

Example 7: The Lifetime X and the brightness Y of a light bulb are modeled as continuous random variables, let the joint pdf be given by $f(x, y) = \lambda_1 \lambda_2 e^{-(\lambda_1 x + \lambda_2 y)}$, $0 < x < \infty$, $0 < y < \infty$. Show that X & Y are independent random variables. Further, compute the joint distribution function.

Solution: The marginal density of X is

$f_x(x) = \int_{-\infty}^{\infty} f(x, y)dy = \int_{-\infty}^{0} 0\, dy + \int_0^{\infty} \lambda_1 \lambda_2 e^{-(\lambda_1 x + \lambda_2 y)}\, dy$

$= \lambda_1 \lambda_2 e^{-\lambda_1 x} \int_0^{\infty} e^{-\lambda_2 y}\, dy$

$= = -\lambda_1 e^{-\lambda_1 x}[e^{-\lambda_2 y}]$

$= -\lambda_1 e^{-\lambda_1 x}[0-1]$

$= \lambda_1 e^{-\lambda_1 x}$, $0 < x < \infty$

Similarly $f_y(y) = \lambda_2 e^{-\lambda_2 y}$, $0 < y < \infty$

$f_x(x) f_y(y) = \lambda_1 e^{-\lambda_1 x} \lambda_2 e^{-\lambda_2 y}$

$= \lambda_1 \lambda_2 e^{-(\lambda_1 x + \lambda_2 y)}$

$= f(x, y)$

$\therefore$ Therefore X & Y are independent random variables.

To find the joint distribution function

$F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u, v)\, du\, dv$

$= \int_0^x \int_0^y \lambda_1 \lambda_2 e^{-(\lambda_1 u + \lambda_2 v)}\, du\, dv$

$= \int_0^x \lambda_1 e^{-\lambda_1 u}\, du \int_0^y \lambda_2 e^{-\lambda_2 v}\, dv$

$= [-e^{-\lambda_1 u}]_0^x [-e^{-\lambda_2 v}]_0^y$

$$= (1- e^{-\lambda_1 x})(1- e^{-\lambda_2 y}),\ 0< x < \infty, 0 < y < \infty$$

Example 8: The joint density of two random variables X and Y is

$$f(x) = \begin{cases} 0.04e^{-0.2x=0.3y}, & for\ x, y < 0 \\ 0, & elsewhere \end{cases}$$

Find the marginal distribution of X.

Solution: The marginal distribution of X is

$$F_X(x) = \int_{-\infty}^{x}\int_{-\infty}^{\infty} f(u,y)dy\,du$$

$$= \int_{0}^{x}\int_{0}^{\infty} f(u,y)dy\,du$$

$$= \int_{0}^{x}\int_{0}^{\infty} 0.04e^{-0.2u=0.3y}dy\,du$$

$$= \int_{0}^{x} 0.04e^{-0.2u}\,dy \int_{0}^{\infty} e^{-0.3y}\,du$$

$$= \frac{0.04}{-0.2}\left[e^{-0.2u}\right]_0^x \frac{1}{-0.3}\left[e^{-0.3y}\right]_0^\infty$$

$$= \frac{2}{3}\left[e^{-0.2x} - 1\right](1-0)$$

$$= \frac{2}{3}\left[e^{-0.2x} - 1\right]$$

4.5 Covariance and correlation

Definition: Let X and Y be random variables with mean μ_x and μ_y respectively. The Covariance between X and Y, denoted by Cov (X, Y) is Cov $(X, Y) = \frac{\Sigma(X-\overline{X})(Y-\overline{Y})}{n}$ where $\overline{X}$ and $\overline{Y}$ are means of X and Y respectively.

4.5.1 Population correlation coefficient ρ (Karl Pearson Coefficient):

Definition: Let X and Y be random variables with means μ_X and μ_Y and variances σ^2_X and σ^2_Y respectively. The Population correlation coefficient (ρ) between X and Y is $\rho = \frac{\text{cov}(X,Y)}{\sqrt{(VarX)(VarY)}}$

Another convenient form of the formulas for computational purpose is as follows:

$$\text{Cov}(X, Y) = \frac{\sum xy}{n} - \bar{x}\,\bar{y}$$

$$\text{Var}(X) = \frac{\sum x^2}{n} - (\bar{x})^2$$

$$\text{Var}(Y) = \frac{\sum y^2}{n} - (\bar{y})^2$$

where (x_i, y_i); i=1 to n is the bivariate distribution related to two random variables X and Y; Also $\bar{x}$ and $\bar{y}$ are the values of means of two random variables X and Y respectively.

Example 9: Calculate the correlation coefficient for the following heights (in inches) of fathers (X) and their sons(Y)

x	65	66	67	67	68	69	70	72
y	67	68	65	68	72	72	69	71

Solution:

X	y	x^2	y^2	xy
65	67	4225	4489	4355
66	68	4356	4624	4488
67	65	4489	4225	4355
67	68	4489	4624	4556
68	72	4624	5184	4896
69	72	4761	5184	4986
70	69	4900	4761	4830
72	71	5184	5041	5112

$\sum x = 544$, $\sum y = 552$, $\sum x^2 = 37028$, $\sum y^2 = 38132$, $\sum xy = 37560$

$$\bar{x} = \frac{\sum x}{n} = \frac{544}{8} = 68, \quad \bar{y} = \frac{\sum y}{n} = \frac{552}{8} = 69$$

$$\rho = \frac{\text{Cov}(X,Y)}{\sqrt{((Var(X)(Var\,Y))}}$$

$$= \frac{\frac{\sum xy}{n} - \bar{x}\,\bar{y}}{\sqrt{\left(\left[\frac{\sum x^2}{n} - (\bar{x})^2\right]\left[\frac{\sum y^2}{n} - (\bar{y})^2\right]\right)}}$$

$$= \frac{\frac{37560}{8} - 68 \times 69}{\sqrt{\left(\left[\frac{37028}{8} - (68)^2\right]\left[\frac{38132}{8} - (69)^2\right]\right)}}$$

$\rho = 0.603$.

Note:

1) $-1 \leq \rho \leq 1$.

2) If $\rho = 1$, then the data points lie perfectly along a straight line with positive slope. This is called perfect positive correlation.

3) If $\rho = -1$ then the data points lie perfectly along a straight line with negative slope. This is called perfect negative correlation.

4) If $\rho = 0$ then the variables are uncorrelated. This indicates the absence of the linear relation.

Example 10:

x	-3	-2	-1	0	1	2	3	$\sum x = 0$
y	9	4	1	0	1	4	9	$\sum y = 28$
xy	-27	-8	-1	0	1	8	27	$\sum xy = 0$

$$\text{Cov}(X, Y) = \frac{\sum xy}{n} - \bar{x}\,\bar{y} = 0$$

$$\therefore \quad \rho = \frac{\text{Cov}(X,Y)}{\sqrt{((Var(X)(Var\,Y))}} = 0.$$

Therefore the variables X and Y are uncorrelated. But on careful examination we find that X and Y are connected by the relation $Y = X^2$. $\therefore \rho(X, Y) = 0$ implies the absence of any linear relationship between the variables X and Y. There may, however, exist some other form of relationship between them eg: Quadratic, Cubic or Trigonometric.

4.5.2 Rank Correlation Coefficient (Spearman's Coefficient)

The population correlation coefficient gives a numerical measure of linear dependence between X and Y (two characteristics), under the assumption that the distribution of X and Y is normal.

There are two kinds of situations in which this approach can't be used. In the first place, perhaps the variables cannot be assumed to have normal distributions. In the second, it might be impossible to measure the characteristics that are of interest.

Definition: Let (x_i, y_i), $i = 1$ to n be the ranks of the i^{th} individual in two characteristics A and B respectively. Then spearman rank-correlation coefficient is given by

$$r_s = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)}, \text{ where } d_i = \text{rank of } x_i - \text{rank of } y_i.$$

Note:

1) $-1 \leq r_s \leq 1$.

2) $r_s = 1$, indicates perfect agreement between the ranks of x and y.

3) $r_s = -1$ indicates that the ranks of y are in exactly the opposite order as the ranks of x.

4) If $r_s = 0$ indicates that x and y are independent.

Important Note: If measurements rather than ranks are given originally the measurement must be changed into ranks in order to find the rank-correlation coefficient.

Example 11: The observations in the following table are the scores that 8 salesmen made on a test that measures their aggressiveness (X), and their sales in thousands of rupees for their sales year with a certain company (Y).

Salesman	1	2	3	4	5	6	7	8
X	30	17	35	28	42	25	19	34
Y	35	31	40	46	50	32	33	42

Solution:

Salesman	x_i	y_i	$D_i^2 = (x_i - y_i)^2$
1	4	5	1
2	8	8	0
3	2	4	4
4	5	2	9
5	1	1	0
6	6	7	1
7	7	6	1
8	3	3	0

$\sum d_i^2 = 16$

The X- values are ranked from the largest to smallest (it does not matter whether the observations are ranked from largest to smallest or from smallest to largest, as long as both x and y values are ranked in the same way): 42, 35, 34, 30, 28, 25, 19, 17. The observation 42 is assigned rank 1, observation 35 is assigned rank 2 and so on. The Y values are ranked from largest to smallest also as 50, 46, 42, 40, 35, 33, 32, 31. The observation 50 is assigned the rank 1, the observation 46 is assigned the rank 2, and so on. Thus, salesman 1 ranks fourth on the test score and fifth in sales: $x_1 = 4$ and $y_1 = 5$. The ranks of the test scores (x_i) of the eight salesman and the ranks of their sales (y_i) are shown in table:

The squared differences between ranks are shown in the last column of the table. Substituting the values $n = 8$ and $\Sigma d_i^2 = 16$ into the following formula

$$r_s = 1 - \frac{6\Sigma d_i^2}{n(n^2-1)} = 1 - \frac{6(16)}{8(63)} = 0.8095 \cong 0.81.$$

Repeated Ranks: If any two or more variables (individuals) are bracketed equal in any classification with respect to characteristic A or B, or if there is more than one item with the same value in the series, then the spearman's formula for calculating the rank correlation coefficient is not suitable. In this case, common ranks are given to the repeated items. This common rank is the average of the ranks which these items would have assumed if they were slightly different from each other and the next item will get the rank next to the ranks already assumed. As a result of this, following adjustment or correction is made in the rank correlation formula.

In the formula, we add the factor $\frac{m(m^2-1)}{12}$ to Σd^2, where m is the number of times an item is repeated. This correction factor is to be added for each repeated value.

Example 12: Obtain the rank correlation coefficient for the following data:

X	68	64	75	50	64	80	75	40	55	64
Y	62	58	68	45	81	60	68	48	50	70

Solution: In the X-series we see that the value 75 occurs 2 times. The common rank given to these values is 2.5 which is the average of 2 and 3, the ranks which these values would have taken if they were different. The next value 68 then gets the next rank which is 4. Again we see that value 64 occurs thrice. The common rank given to it is 6 which is the average of 5, 6 and 7. Similarly in the Y-series value 68 occurs twice and its common rank is 3.5 which is the average of 3 and 4. As a result of these common rankings, the formula for 'r' has to be corrected. To Σd^2

we add $\frac{m(m^2-1)}{12}$ for each value repeated, where m is the number of times a value occurs. In the X-series the correlation is to be applied twice, once for the value 75 which occurs twice ($m = 2$) and then for the value 64 which occurs thrice ($m = 3$). The total correction for the X-series is $\frac{2(4-1)}{12}+\frac{3(9-1)}{12}=\frac{5}{2}$.

Similarly, this correction for the Y-series is $\frac{2(4-1)}{12}=\frac{1}{2}$ as the value 68 occurs twice.

X	Y	Rank of X (x_i)	Rank of Y (y_i)	$d_i = x_i - y_i$	d_i^2
68	62	4	5	-1	1
64	58	6	7	-1	1
75	68	2.5	3.5	-1	1
50	45	9	10	-1	1
64	81	6	1	5	25
80	60	1	6	-5	25
75	68	2.5	3.5	-1	1
40	48	10	9	+1	1
55	50	8	8	0	0
64	70	6	2	4	16
					$\sum d^2 = 72$

$$\text{Thus } r_s = 1-\frac{6[\sum d^2+\frac{5}{2}+\frac{1}{2}]}{n(n^2-1)}$$

$$\simeq 1-\frac{6(72+3)}{10*99}$$

$$= 0.545.$$

4.6 Summary

Given the random variable X and its density $f(x)$, we can define $Y = \Phi(X)$ i.e, Y is a function of the random variable X. Further Y is a random variable if Φ is continuous or piecewise continuous.

Definition: Let $X_1, X_2, \ldots, X_r$ be r discrete random variables defined on a sample space S. for each element s in S, let $X_1(s) = X_1, X_2(s) = X_2, X_r(s) = X_r$. Then $X = (X_1, X_2, \ldots, X_r)$ is an r- dimensional function from S to R^r with $X(s) = x = (x_1, x_2, \ldots, x_r)$. X is called a discrete random vector.

Definition: The joint (or compound) distribution function of random variables X and Y is defined by $F_{X,Y}(x, y)=P(X \leq x, Y \leq y)$, $-\infty < x < \infty$, $-\infty < y < \infty$.

Definition: Let X and Y be random variables with means μ_X and μ_Y and variances σ^2_X and σ^2_Y respectively. The Population correlation coefficient (ρ) between X and Y is $\rho = \dfrac{\text{cov}(X,Y)}{\sqrt{(VarX)(VarY)}}$

Definition: Let (x_i, y_i), $i = 1$ to n be the ranks of the i^{th} individual in two characteristics A and B respectively. Then spearman rank-correlation coefficient is given by

$$r_s = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)}, \text{ where } d_i = \text{rank of } x_i - \text{rank of } y_i.$$

4.7 Keywords

Random variable, covariance, correlation coefficient.

4.8 Supplementary problems

4.8.1. If X is uniform in (0, 1) show that $15X+20$ is uniform in (20, 35).

4.8.2. If X is uniform in (0,1), Verify that $-\log(X)/2$ is exponential with parameter 2.

4.8.3. Let X and Y have the joint probability distribution given below:

	X			
Y		0	1	2
	0	.1	.4	.1
	1	.2	.2	0

Find (i) $P(X + Y>1)$ (ii) Marginal distribution of Y.

(iii) Find the conditional distribution of Y given $X=2$.

(iv) Are X and Y independent?

4.8.4. If two random variables have the joint density

$$f(x_1, x_2) = \begin{cases} x_1 x_2, & 0 < x_1 < 1, \ 0 < x_2 < 1 \\ 0, & \text{elsewhere} \end{cases}$$

Find the probability that i) both are < 1 ii) sum <1 .

4.8.5. Two random variables X and Y have the following joint density

$$f(x,y)=\begin{cases}2(x^2+y^2), & 0<x,y<1\\ 0, & elsewhere\end{cases}$$. Find

(i) P(0.2<x<0.5 and 0.4<y<0.6). (ii)Marginal densities

(iii) Distribution function (iv) $P(X>0.8)$, $P(Y<0.5)$

4.8.6. Calculate the coefficient of correlation for the following data

X	78	36	98	25	75	82	90	62
Y	84	51	91	60	68	62	86	58

X	105	109	102	101	100	99	98	96
Y	101	103	100	98	95	96	104	92

4.8.7. Obtain the rank correlation coefficient for the following data

X	68	64	75	50	64	80	75	40	55	64
Y	62	58	68	45	81	60	68	48	50	70

4.8.8. The rankings of 10 students in 2 subjects A and B are as follows. Find the rank correlation coefficient.

A	3	5	8	4	7	10	2	1	6	9
B	6	4	9	8	1	2	3	10	5	7

4.9 References

1. Probability and statistics for Engineers, by G.S.S. Bhishma Rao. (Scitech publications).
2. Probability and Statistics with Reliability, Queuing and Computer Science Applications, by K.S.Trivedi, PHI publications.

NOTES

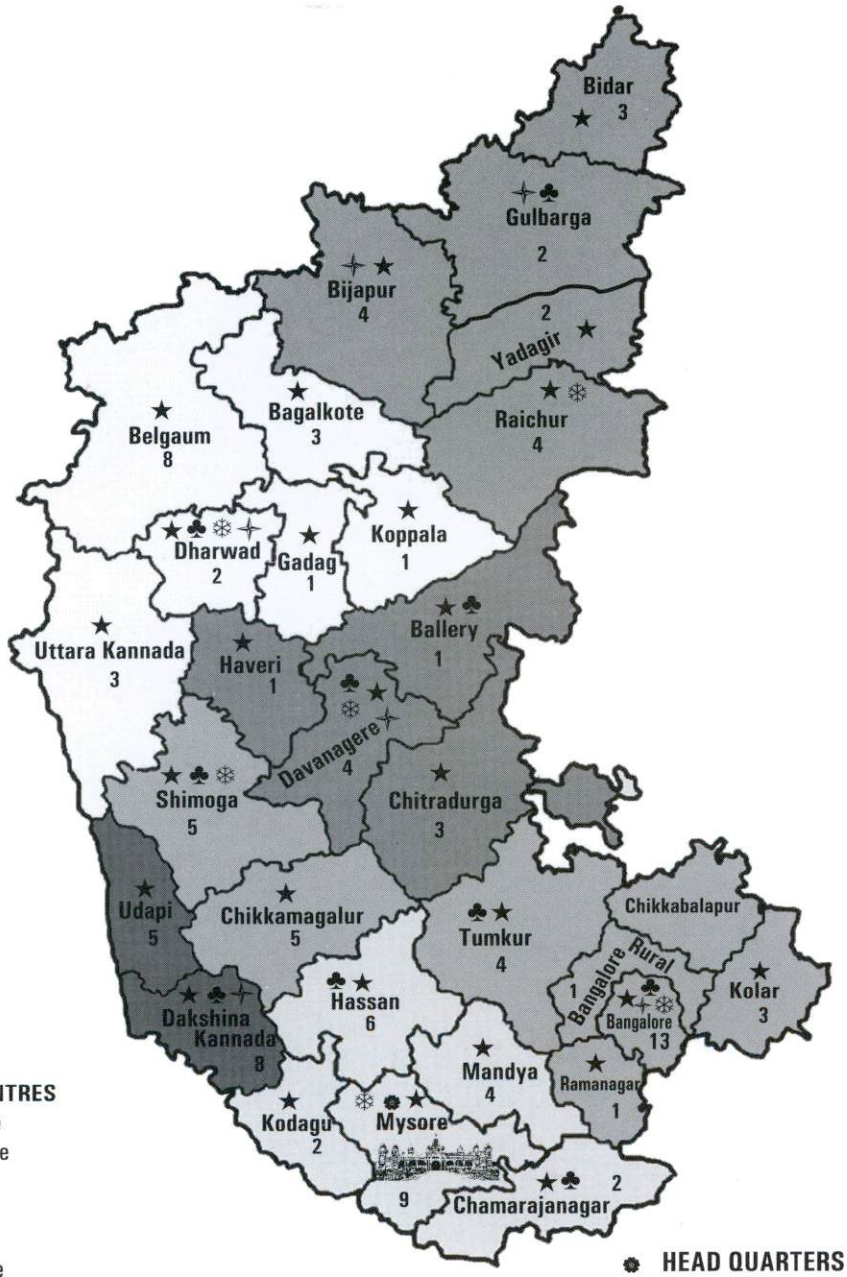
NOTES

NOTES

ಆದೇಶ ಸಂಖ್ಯೆ : ಕರಾಮುವಿ/ಸಿಪಾವಿ/2-1311/2010-2011 ದಿನಾಂಕ : 16/11/2010
ಒಳಪುಟ : 60 GSM ಮ್ಯಾಪ್ಲಿತೊ ಮತ್ತು ರಕ್ಷಾಪುಟ : 220 GSM ಆರ್ಟ್ ಕಾರ್ಡ್
ಮುದ್ರಕರು : ವಿನಾಯಕ ಆಫ್‌ಸೆಟ್ ಪ್ರಿಂಟರ್ಸ್, ಬೆಂಗಳೂರು – 560 070 ಪ್ರತಿಗಳು : 500

Karnataka State Open University

Manasagangotri Mysore - 570 006



♣ **REGIONAL CENTRES**

Bangalore
Davanagere
Gulbarga
Dharwad
Shimoga
Mangalore
Tumkur
Hassan
Chamarajanagar
Bellary

HEAD QUARTERS

★ **Total Study Centres : 111**

♣ **Regional Centres : 10**

❄ **B.Ed Study Centres : 10**

✢ **M.Ed Study Centres : 08**

VOP-02-356-500