JIS GROUP

ONLINE COURSE WARE

SUBJECT NAME: MATHEMATICS III

SUBJECT CODE: M (CSE) 301

TOTAL NO. OF LECTURES: 44

CONTACT HOURS: 44 HOURS.

CREDIT: 4

LESSION PLAN

MODULE NO	LECTURE NO	TOPIC	APPLICATION	REFERENCE BOOK
Module Number I: **Basic Probability Theory**	**Lecture 1**	INTRODUCTION TO PROBABILITY.	• Theory of Probability deals with law of governing the chances of occurrence of phenomenon which are unpredictable in nature. • Useful in almost all disciplines. • The outcome in terms of number of favorable outcomes and total outcomes in random experiment. • Laws of union, intersection etc. of events. • A formula for evaluating variable at beginning of table. • Probability on the Limiting value of trials. • Typical Problems. • Theorem of conditional and unconditional probability. • Versatile problems on probability. • Concept of presenting probability in terms of variable and its distribution. • Probability connected to Random Variable when it is Discrete and Continuous. • Advanced probability counts on the concept. • Probability of continuous variable and very practical problem. • Application in specific case (Chi-square and t). • Behavior of addition of	**1**.RusselMerris, Combinatorics, Wiley-Interscience series in Discrete Mathematics and Optimization **2**. N. Chandrasekaran and M. Umaparvati, Discrete Mathematics **3**. Gary Haggard, John Schlipf and Sue Whresides, Discrete Mathematics for Computer Science, CENGAGE **4**.Lipschutz S: Thory and Problems of Probability and Statistics (Schaum's Outline Series) – McGrew Hill Book Co. **5**.Spigel M R: Theory and Problems of Probabiliy and Statistics (Schaum's Outline Series)– McGrew Hill Book Co. **6**. Banerjee A.,
	Lecture 2	ADDITION THEOREM OF PROBABILITY		
	Lecture 3	CONDITIONAL PROBABILITY		
	Lecture 4	INDEPENDENT EVENTS & TOTAL PROBABILITY THEOREM		
	Lecture 5	BAYES' THEOREM.		
	Lecture 6	DISCRETE RANDOM VARIABLE AND IT'S PROBABILITY DISTRIBUTION.		
	Lecture 7	CONTINUOUS RANDOM VARIABLE AND IT'S PROBABILITY DISTRIBUTION.		
	Lecture 8	BINOMIAL DISTRIBUTION		
	Lecture 9	POISSON DISTRIBUTION		
	Lecture 10	NORMAL DISTRIBUTION		

	number of Random Variables. • Others are inequality in specific cases. • Law of Large number deals with the sequence of Random Variables. • Random Variables are transformed in special applications.	De S. K. and Sen S.: Mathematical Probability, U N Dhur& Sons. **7**. Deo N: Graph Theory with Application to Engg. & Computer Sc., Prentice Hal. **8**. Grewal B S, Higher Engg. Maths. (35 ed.), Khanna Pub. **9**. Kreyzig E., Advanced Engg. Maths., John Wiley & Sons. **10**. J K Sharma, Discrete Mathematics, Macmillan Publication. **11**. Winfried Karl Grassmann& Jean-Paul Tremblay, Logic & Discrete Maths., PEARSON. **12**. S K Chakraborty& B K Sarkar, Discrete Maths. OXFORD University Press **13.**

				Lakshsminarayan, Engg. Maths. 1.2.3 **14**. Spigel M R, Schiller J J&Srinivasan R A: Probability and Statistics (Schaum's Outline Series), TMH. **15**. Wilson: Introduction to graph theory, Pearson Education.
Module Number II: **PROPOSITIONAL LOGIC**	**Lecture 11**	Introduction to Propositional Calculus, Propositions, Logical Connectives	• Applications of Mathematical Logic to Formal Verification and program analysis • Logic is generally based on deduction which is a method of exact inference. • It is a study of correct reasoning that consist of language and reasoning. • Reasoning practically in our daily lives involves deciding what to do and when successful, issuing in an intention. • This is essential in General Problem Solving: It helps us to analyze concepts, definitions, arguments and problems, and contributes to our capacity to organize ideas and issues to deal with questions of value	**1.** Russell Merris, Combinatorics, Wiley-Interscience series in Discrete Mathematics and Optimisation **2.** N. Chandrasekaran and M. Umaparvathi, Discrete Mathematics, PHI **3.**Grewal B S: Higher Engineering Mathematics (thirtyfifthedn) - Khanna Pub.
	Lecture 12	Conjunction, Disjunction, Negation and their truth table		
	Lecture 13	Converse, Contrapositive, Inverse		
	Lecture 14	Logical Equivalence, Normal forms-CNF, DNF		
	Lecture 15	Predicate logic		
	Lecture 16	Logical Quantifications of propositions		

				4. Kreyzig E: Advanced Engineering Mathematics - John Wiley and Sons. **5.** J.K. Sharma, Discrete Mathematics, Macmillan **6.** Winfried Karl Grassmann and Jean-Paul Tremblay, Logic and Discrete Mathematics, PEARSON. **7.** S. K. Chakraborty and B. K. Sarkar, Discrete Mathematics, OXFORD University Press. 11. Douglas B. West, Introduction to graph Theory, PHI **8.** Discrete Mathematics and its Application: Kenneth H. Rosen, McGraw Hill Education.

Module Number III: **NUMBER THEORY & PARTIAL ORDER RELATION AND LATTICES**	**Lecture 17**	Introduction	• To enable the students to be aware of a class of functions which transform a finite set into another finite set which relates to input output functions in computer science. • Elementary mathematics tends to focus lopsidedly on computational structures. • Algebra is presented as a collection of problem-solving techniques and as a generalized symbolic arithmetic. • Quantitative structure-activity relationships are often based on standard multidimensional statistical analyses and sophisticated local and global molecular descriptors. • Here, the aim is to develop a tool helpful to define a molecule or a class of molecules which fulfills pre-described properties, i.e., an Inverse QSAR approach. If highly sophisticated descriptors are used in QSAR, the structure and then the synthesis recipe may be hard to derive. Thus, descriptors, from which the synthesis recipe can be easily derived, seem appropriate to be included within this study. However, if descriptors simple enough to be useful for	**1.** Russell Merris, Combinatorics, Wiley-Interscience series in Discrete Mathematics and Optimisation **2.** N. Chandrasekaran and M. Umaparvathi, Discrete Mathematics, PHI **3.**Grewal B S: Higher Engineering Mathematics (thirtyfifthedn) - Khanna Pub. **4.** Kreyzig E: Advanced Engineering Mathematics - John Wiley and Sons. **5.** J.K. Sharma, Discrete Mathematics, Macmillan **6.** Winfried Karl Grassmann and Jean-Paul Tremblay, Logic and Discrete
	Lecture 18	Fundamental Theorem of Arithmetic		
	Lecture 19	Greatest Common Divisor		
	Lecture 20	CONGRUENCE		
	Lecture 21	Residue classes of integer modulo n		
	Lecture 22	Relation, Partial Order Relation		
	Lecture 23	Hasse Diagram		
	Lecture 24	Lattice		

			defining syntheses recipes of chemicals were used, the accuracy of a numeric expression may fail. • This paper suggests a method, based on very simple elements of the theory of partially ordered sets, to find a qualitative basis for the relationship between such fairly simple descriptors on the one side and a series of ecotoxicological properties, • on the other side. The partial order ranking method assumes neither linearity nor certain statistical distribution properties. Therefore the method may be more general compared to many standard statistical techniques. • A series of chlorinated aliphatic compounds has been used as an illustrative example and a comparison with more sophisticated descriptors derived from quantum chemistry and graph theory is given. Among the results, it was disclosed that only for algae lethal concentration, as one of the four ecotoxicological properties, the synthesis specific predictors seem to be good estimators.	Mathematics, PEARSON. **7.** S. K. Chakraborty and B. K. Sarkar, Discrete Mathematics, OXFORD University Press. **8**. Douglas B. West, Introduction to graph Theory, PHI **9.** Discrete Mathematics and its Application: Kenneth H. Rosen, McGraw Hill Education.

Module Number IV: **PRINCIPLES OF COUNTING TECHNIQUES**	**Lecture 25**	INTRODUCTION TO COUNTING TECHNIQUES	• Counting is useful in computer science for several reasons: Determining the time and storage required to solve a computational problem • A central objective in computer science often comes down to solving a counting problem. Counting is the basis of probability theory, which plays a central role in all sciences, including computer science. • Two remarkable proof techniques, the “pigeon-hole principle” and “combinatorial proof,” rely on counting. • The number of different ways to select a dozen doughnuts when there are five varieties available. The number of 16-bit numbers with exactly 4 ones. • Our objective is to teach you simple counting as a practical skill, like integration. But counting can be tricky, and people make counting mistakes all the time, so a crucial part of counting skill is being able to verify a counting argument. • Sometimes this can be done simply by finding an alternative way to count and then comparing answers they better agree.	**1.** Russell Merris, Combinatorics, Wiley-Interscience series in Discrete Mathematics and Optimisation **2.** N. Chandrasekaran and M. Umaparvathi, Discrete Mathematics, PHI **3.**Grewal B S: Higher Engineering Mathematics (thirtyfifthedn) - Khanna Pub. **4.** Kreyzig E: Advanced Engineering Mathematics - John Wiley and Sons. **5.** J.K. Sharma, Discrete Mathematics, Macmillan **6.** Winfried Karl Grassmann and Jean-Paul
	Lecture 26	COMBINATIONS AND BINOMIAL COEFFICIENTS		
	Lecture 27	ADVANCED COUNTING TECHNIQUES		
	Lecture 28	RECURRENCE RELATION		
	Lecture 29	SOLUTION OF LINEAR RECURRENCE RELATION WITH CONSTANT COEFFICIENTS BY CHARACTERISTIC ROOT METHOD**.**		
	Lecture 30	SOLUTION OF LINEAR RECURRENCE RELATION WITH CONSTANT COEFFICIENTS BY GENERATING FUNCTION METHOD.		

			• Most elementary counting arguments reduce to finding a bijection between objects to be counted and easy-to-count sequences. • The chapter shows how explicitly defining these bijections —and verifying that they are bijections is another useful way to verify counting arguments.	Tremblay, Logic and Discrete Mathematics, PEARSON. **7.** S. K. Chakraborty and B. K. Sarkar, Discrete Mathematics, OXFORD University Press. **8**. Douglas B. West, Introduction to graph Theory, PHI **9.** Discrete Mathematics and its Application: Kenneth H. Rosen, McGraw Hill Education.
Module V: Algebraic Structure	**Lecture 31**	Introduction	• An algebraic structure associated with mapping of different compositions non-empty set. • A non-empty set is said to form a group w.r.t the binary composition '*' (say) if it is closed under the binary composition. • Different types of groups and the concerned properties is utilized in system.	1.Abstract algebra by Dummit and Foote 2.Abstract Algebra by S.k. Mapa. **3**. First Course in Abstract Algebra, A, 7th Edition John B. Fraleigh, University of Rhode Island
	Lecture 32	Groups		
	Lecture 33	Subgroups		
	Lecture 34	Cyclic groups		
	Lecture 35	Cosets, Quotient Group, Normal subgroup		
	Lecture 36	Ring, Field, Integral Domain		

Module V: Advanced Graph Theory	**Lecture 37**	Definition of Advanced Graph Theory	• The concept of color graph with no two vertices has been colored is Advanced graph.	**1.** Russell Merris, Combinatorics, Wiley-Interscience series in Discrete Mathematics and Optimisation **2.** N. Chandrasekaran and M. Umaparvathi, Discrete Mathematics, PHI **3.**Grewal B S: Higher Engineering Mathematics (thirtyfifthedn) - Khanna Pub. **4.** Kreyzig E: Advanced Engineering Mathematics - John Wiley and Sons. **5.** J.K. Sharma, Discrete Mathematics, Macmillan **6.** Winfried Karl Grassmann and Jean-Paul Tremblay, Logic and
	Lecture 38	Theorems of Advanced Graph Theory	• Properties related to chromatic graph and its bounds, concept of chromatic polynomial,	
	Lecture 39	Kuratowski's Graph, Generalized Euler's formula	• The concept of planer graph and connected theory.	
	Lecture 40	Planarity, Graph Coloring	• Concept of planarity and coloring graph.	
	Lecture 41	Dual and properties	• Dual graph and its properties.	
	Lecture 42	Graph Coloring	• Few examples of graph coloring.	
	Lecture 43	Theorems on Graph Coloring	• Some essential properties of graph coloring.	
	Lecture 44	Application of Graph: Coloring-matching	• Useful applications of graph coloring.	

				Discrete Mathematics, PEARSON. **7.** S. K. Chakraborty and B. K. Sarkar, Discrete Mathematics, OXFORD University Press. **8**. Douglas B. West, Introduction to graph Theory, PHI **9.** Discrete Mathematics and its Application: Kenneth H. Rosen, McGraw Hill Education.

MODULE I: THEORY OF PROBABILITY

(NUMBER OF LECTURES: 10)

LECTURE 1: INTRODUCTION TO PROBABILITY.

1.1. INTRODUCTION

In our daily life, we often used phrases such as 'It may rain today', or 'India may win the match' or ' I may be selected for this post.' These phrases involve an element of uncertainty. How can we measure this uncertainty? A measure of this uncertainty is provided by a branch of Mathematics, called the theory of probability. Probability Theory is designed to measure the degree of uncertainty regarding the happening of a given event. The dictionary meaning of probability is likely though not certain to occur. Thus, when a coin is tossed, a head is likely to occur but may not occur. Similarly, when a die is thrown, it may or may not show the number 6.

1.2. BRIEF HISTORY

Concepts of probability have been around for thousands of years, but probability theory did not arise as a branch of mathematics until the mid-seventeenth century. In 1654 the famous mathematician Blaise Pascal had a friend, Chevalier de Mere, a member of the French nobility and a gambler, who wanted to adjust gambling stakes so that he would be assured of winning if he played long enough. This gambler raised questions with Pascal such as the following:

"In eight throws of a die a player attempts to throw a one, but after three unsuccessful trials the game is interrupted. How should he be compensated?"

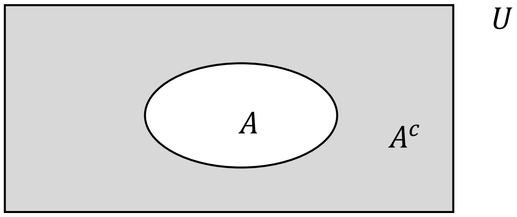
Pascal wrote to a leading mathematician of that day, Pierre de Fermat (1601–1665), about these problems, and their resulting correspondence represents the beginnings of the modern theory of mathematical probability.

The topic of probability is seen in many facets of the modern world. The theory of probability is not just taught in mathematics courses, but can be seen in practical fields, such as *insurance*, *industrial quality control*, *study of genetics*, *quantum mechanics*, and *the kinetic theory of gases*.

1.3. PREREQUISITE: BASIC TERMS AND NOTAIONS IN SET THEORY

- A countably infinite (or countable) set is a set with infinitely many elements which can be enumerated in a list, e.g., the set of all integers $\{0,-1,1,-2,2,\dots\}$. An example of an uncountable set is the set of all real numbers between 0 and 1, denoted $[0,1]$.
- ϕ denotes the empty set, i.e., the set that has no element.
- $S \subseteq T$ means that "S is a subset of T", i.e., every element of S is also an elements of T.
- $S \subset T$ means that "S is a proper subset of T", i.e., every element of S is also an elements of T and T has at least one element which is not in S.
- $S = T$ means that "S and T are two equal sets", i.e., $S \subseteq T$ as well as $T \subseteq S$.
- The set of all x that have a certain property P is denoted by $\{x : x\ satisfies\ P\}$, e.g., the interval $[0,1]$ can alternatively be written as $\{x : 0 \le x \le 1\}$

- The *complement* of a set A defined in the universal set U is the set $\{x: x \in U \; but \; x \notin A\}$ and is denoted by A^c or $\bar{A}$ or A'. It is to be noted here that $U^c = \phi$ and $\phi^c = U$.



1.4. TERMINOLOGY

Before we learn how find probability of an event associated with some random experiment, we will learn about some of terminologies which are frequently used in theory of probability.

Definition 1.1. (*Random Experiments***)**

The basic notion in probability is that of a random experiment: an experiment whose outcome cannot be determined in advance, but is nevertheless still subject to analysis.
Examples of random experiments are:

- tossing a die,
- measuring the amount of rainfall in Brisbane in January,
- counting the number of calls arriving at a telephone exchange during a fixed time period,
- selecting a random sample of fifty people and observing the number of left-handers,
- choosing at random ten people and measuring their height.

Definition 1.2. (*Sample Space***)**

The set of all possible outcomes of a random experiment is known as *sample space* that random experiment and is denoted by Ω or S. It is to be worthy to mention here that sample space, in the context of probability, play exactly the same role as that of universal set in the context of set theory.

Definition 1.3. (*Event***)**

Any subset of the sample space is known as *event* of the random experiment.
Note:

- Since, every set is subset of itself, sample space of a random experiment is considered as an event of that random experiment. Moreover, since the sample space contains all possible outcomes of the random experiment; every time the random experiment will be performed, sample space will occur as an event and hence it is known as *certain event*.

- Since, empty set, i.e., ϕ is a subset of every set, ϕ is considered as an event of that random experiment. Moreover, since ϕ contains no element, none of the results of the random experiment will favour the event ϕ to occur and hence it is known as *impossible event*.

Example 1.2. Using the setup in Example 2 we would describe the event that you get exactly two heads in words by E = 'exactly 2 heads'. Written as a subset this becomes E = {HHT, HT H, T HH}. You should get comfortable moving between describing events in words and as subsets of the sample space.

Definition 1.4. (*Complementary Event*)

In random experiment, non-occurrence of any event is also an event of that random experiment which is known as complementary event of the first event.

If A be any event of some random experiment, then the complementary event of the event A is denoted by A^C or $\bar{A}$ or A'.

Note:

- Since, complement of an event A is non-occurrence of the event A; the outcomes which are favourable for the event A, are not favourable for the complement of event A, i.e., A^C. Thus, $A^C = S - A \Longrightarrow n(A^C) = n(S) - n(A)$.
- Complement of impossible event ϕ is the certain event S, i.e., $\phi^c = S$.
- Complement of certain event S is the impossible event ϕ, i.e., $S^c = \phi$.
- Complement of complement of any event A in a random experiment is the event itself, i.e., $(A^c)^c = A$ [since, $(A^c)^c = S - A^c = S - (S - A) = A$].

Definition 1.4. (*Mutually Exclusive Events*)

In random experiment, two events are said to be mutually exclusive events if occurrence of one prevents the occurrence of the other event.

If A and B are any two events of some random experiment, then A and B are said to be mutually exclusive events if occurrence of the event A prevents the occurrence of the event B and vice versa. Thus, A and B can not occur simultaneously, i.e., $A \cap B = \phi$.

Definition 1.5. (*Mutually Exhaustive Events*)

In random experiment, two events are said to be mutually exhaustive events if at least one of them always occur whenever the random experiment is performed.

If A and B are any two events of some random experiment, then A and B are said to be mutually exhaustive events if at least one of the events A and B occur whenever the random experiment is performed. This is possible only when all the possible outcomes of that random experiment are favourable for at least one of the events A and B, i.e., $A \cup B = S$.

Definition 1.6. (*Equally Likely Events*)

In random experiment, two events are said to be equally likely events if all of them have equal chance to occure, i.e., no one will get any preference.

1.5. DEFINITIONS OF PROBABILITY

The probability of an event has been defined in several ways. Two of the most popular definitions are: the *relative frequency definition*, and the *classical definition*.

Definition 1.7. (*Relative Frequency Definition*)

Suppose that a random experiment is repeated n times. If the event A occurs n_A times, then the probability of A, denoted by $P(A)$, is defined as

$$P(A) = \lim_{n\to\infty} \left(\frac{n_A}{n}\right)$$

where $\frac{\mathrm{n}_A}{\mathrm{n}}$ represents the fraction of occurrence of A in n trials.

For small values of n , it is likely that $\frac{n_A}{n}$ will fluctuate quite badly. But as n becomes larger and larger, we expect, $\frac{n_A}{n}$ to tend to a definite limiting value. For example, let the experiment be that of tossing a coin and A the event 'outcome of a toss is Head'. If n is the order of 100, $\frac{n_A}{n}$ may not deviate from $\frac{1}{2}$ by more than, say ten percent and as n becomes larger and larger, we expect $\frac{n_A}{n}$ to converge to $\frac{1}{2}$.

The relative frequency definition given above has empirical flavor. In the classical approach, the probability of the event A is found without experimentation.

Definition 1.8. (*Classical Definition***)**

If a random experiment has n number of mutually exclusive, mutually exhaustive and equally likely possible outcomes out of which m number of outcomes are favorable to the occurrence of some event A of that experiment, then probability of the event A, denoted by $P(A)$, is defined as

$$P(A) = \frac{m}{n}$$

1.6. SOME IMPORTANT RESULTS

Result 1: $P(\phi) = 0$ and $P(S) = 1$.

Result 2: $0 \leq P(A) \leq 1$, for any event A of a random experiment.

Result 3: $P(A^c) = 1 - P(A)$, for any event A of a random experiment.

1.7. AXIOMS OF PROBABILITY

In random experiment, a probability space is the triple $(S, \mathcal{F}, P)$, wehere:

- the first object S is an arbitrary set of outcomes, sometimes called a sample space;
- the second object $\mathcal{F}$ is the collection of all events, that is a set of subsets of S;
- the third object P is defined a function from $\mathcal{F}$ to $[0,1]$, i.e., $P: \mathcal{F} \to [0,1]$ and known as *probability function*, s the collection of all events, that is a set

Finally, the probability P is a number attached to every event A and satisfies the following three axioms:

Axiom 1. For every event A, $P(A) \geq 0$.

Axiom 2. $P(S) = 1$.

Axiom 3. If $A_1, A_2, \cdots$ is a sequence of pair-wise disjoint events, then

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

Worked out Problems:

Problem 1.1. A die is rolled once. Find the probability of getting a '5'.

Solution: There are six possible ways in which a die can fall, viz., 1, 2, 3, 4, 5, 6 which are mutually exclusive, exhaustive and equally likely. Now, among these '5' is the only one result which is favourable to the event. Thus, probability of getting a 5 = $P(5) = \frac{1}{6}$.

Problem 1.2. A coin is tossed once. What is the probability of the coin coming up with head?

Solution: The coin can come up either 'head' (H) or a tail (T). Thus, the total number of possible outcomes

is two and one is favourable to the event. So, the probability of the coin coming up with head = $P(H) = \frac{1}{2}$.

Problem 1.3. A die is rolled once. What is the probability of getting a prime number?

Solution: There are six possible outcomes in a single throw of a die, viz., 1, 2, 3, 4, 5, 6 which are mutually exclusive, exhaustive and equally likely. Out of these, 2, 3 and 5 are prime numbers. So, the number of favourable cases for the event is 3. Thus, probability of getting a prime number = $P(prime\ number) = 3/6 = 1/2$.

Problem 1.4. A die is rolled once. What is the probability of the number '7' coming up? What is the probability of a number 'less than 7' coming up?

Solution: There are six possible outcomes in a single throw of a die, viz., 1, 2, 3, 4, 5, 6 which are mutually exclusive, exhaustive and equally likely. But, there is no face of the die with mark 7 and hence no favourable case. Thus, probability of getting a 7 = $P(7) = \frac{0}{6} = 0$.

[**Note:** That the probability of impossible event is zero].

Again, every face of a die is marked with a number less than 7 and hence the number of favourable cases for the event is 6. Thus, probability of getting a number 'less than 7' = $\frac{6}{6} = 1$.

[**Note:** That the probability of an event that is certain to happen is 1]

Problem 1.5. In a simultaneous toss of two coins, find the probability of (i) getting 2 heads (ii) exactly 1 head.

Solution: Here, the possible outcomes are HH, HT, TH, TT, i.e., the total number of possible outcomes = 4.

(i) Number of outcomes favourable to the event 'getting 2 heads' = 1 (i.e., HH). Thus, $P\ (getting\ 2\ heads)\ =\ \frac{1}{4}$.
(ii) Now, the event consisting of exactly one head has two favourable cases, viz., HT and TH. Thus, $P\ (getting\ exactly\ 1)\ =\ \frac{2}{4} = \frac{1}{2}$.

Problem 1.6. In a single throw of two dice, what is the probability that the sum is 9?

Solution: Here, the number of possible outcomes is $6^2 = 36$. We write them as given below:

$$\begin{gathered}(1,1),\ (1,2),\ (1,3),\ (1,4),\ (1,5),\ (1,6)\\(2,1),\ (2,2),\ (2,3),\ (1,4),\ (2,5),\ (2,6)\\(3,1),\ (3,2),\ (3,3),\ (3,4),\ (3,5),\ (3,6)\\(4,1),\ (4,2),\ (4,3),\ (4,4),\ (4,5),\ (4,6)\\(5,1),\ (5,2),\ (5,3),\ (5,4),\ (5,5),\ (5,6)\\(6,1),\ (6,2),\ (6,3),\ (6,4),\ (6,5),\ (6,6)\end{gathered}$$

Now, how do we get a total of 9. We have:

$$\begin{gathered}3\ +\ 6\ =\ 9\\4\ +\ 5\ =\ 9\\5\ +\ 4\ =\ 9\\6\ +\ 3\ =\ 9\end{gathered}$$

In other words, the outcomes (3, 6), (4, 5), (5, 4) and (6, 3) are favourable to the said event, i.e., the number of favourable outcomes is 4. Thus, $P\ (getting\ a\ total\ of\ 9\ )\ =\ \frac{4}{36}=\frac{1}{9}$.

Problem 1.7. From a bag containing 10 red, 4 blue and 6 black balls, a ball is drawn at random. Find the probability of drawing (i) a red ball, (ii) a blue ball, (iii) not a black ball.

Solution: There are 20 balls in total. So, the total number of ways by which one ball can be drawn out of 20 balls is ${}^{20}C_1 = 20$. Thus, the total number of possible outcomes is 20 (Random drawing of balls ensure equally likely outcomes).

(i) Now, for the occurrence of the event ‘drawing a red ball’, the selected ball must be from the 10 red balls of bag which can happen in ${}^{10}C_1$ ways, i.e., 10 ways. So, the number of outcomes favourable to the said event is 10. Thus, $P(drawing\ a\ red\ ball) = \frac{10}{20} = \frac{1}{2}$.

(ii) Again, for the occurrence of the event ‘drawing a blue ball’, the ball is to be selected from the 4 blue balls of bag which can be done in ${}^{4}C_1$ ways, i.e., 4 ways. So, the number of outcomes favourable to the said event is 4. Thus, $P(drawing\ a\ blue\ ball) = \frac{4}{20} = \frac{1}{5}$.

(iii) Now, the number of balls which are not black is $(10 + 4)$, i.e., 14. So, for the occurrence of the event ‘drawing a ball which is not black’, the ball is to be selected from the 14 balls which are not black of bag which can be done in ${}^{14}C_1$ ways, i.e., 14 ways. So, the number of outcomes favourable to the said event is 14. Thus, $P(drawing\ a\ ball\ which\ is\ not\ black) = \frac{14}{20} = \frac{7}{10}$.

Problem 1.8. A card is drawn at random from a well shuffled deck of 52 cards. If A is the event of getting a queen and B is the event of getting a card bearing a number greater than 4 but less than 10, find $P(A)$ and $P\ (B)$.

Solution: A card can be drawn from a well shuffled deck of 52 cards in ${}^{52}C_1$ ways, i.e., 52 ways. Since, well shuffled pack of cards ensures equally likely outcomes, the total number of possible outcomes is 52.

(i) There are 4 queens in a pack of cards. Thus, $P(A) = \frac{4}{52} = \frac{1}{13}$.

(ii) The cards bearing a number greater than 4 but less than 10 are 5, 6, 7, 8 and 9. Each card bearing any of the above number is of 4 suits diamond, spade, club or heart. Thus, the number of favourable outcomes is $(5\ \times\ 4)\ =\ 20$. Thus, $P(B) = \frac{20}{52} = \frac{5}{13}$.

LECTURE 2: ADDITION THEOREM OF PROBABILITY

2.1. INTRODUCTION

In the last lesson we have studied different terminologies related to the theory of probability as well as the definitions (classical, frequency and axiomatic) of probability. Moreover, the classical definition of probability is extensively exercised to find the probability of an event in a random experiment.

However, in practical problems, writing down the elements of sample space and counting the number of cases favourable to a given event often become very tedious. In such situations, the computation of probabilities can be facilitated to a great extent by fundamental theorem of addition. In this lesson we will learn Addition Theorem of Probability to find probability of occurrence for simultaneous trials under two conditions when events are mutually exclusive and when they are not mutually exclusive. But, before that we need to know about the following notations as we will use them frequently.

List of Symbols

$\boldsymbol{A} \cup \boldsymbol{B}$**:** An event which represents the happening of at least one of the events A and B i.e. either A occurs or B occurs or both A and B occur. This is also denoted as A or B (The region shaded by the patterns ▤, ▥ and ⁙ in the figure)

$\boldsymbol{A} \cap \boldsymbol{B}$**:** An event which represents the simultaneous happening of both A and B i.e. A and B (The region shaded by the patterns ⁙ in the figure).

$\overline{\boldsymbol{A}} \cap \overline{\boldsymbol{B}}$**:** Neither A nor B happens, i.e., none of A and B happens (The region shaded by the patterns ▒ in the figure).

$\boldsymbol{A} \cap \overline{\boldsymbol{B}}$: A happens but B does not happen (The region shaded by the patterns ▤ in the figure).

$\overline{\boldsymbol{A}} \cap \boldsymbol{B}$: A does not happen but B happens (The region shaded by the patterns ▥ in the figure).

$(\overline{\boldsymbol{A}} \cap \boldsymbol{B}) \cup (\boldsymbol{A} \cap \overline{\boldsymbol{B}})$**:** Exactly one of the two events A and B happens (The region shaded by the patterns ▥ and ▤ in the figure).

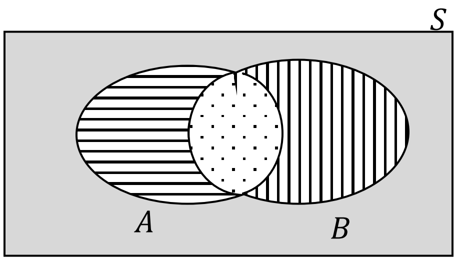


Figure 2.1. Venn diagram representation of two events A and B of a random experiment.

2.2. ADDITION THEOREM OF PROBABILITY (*For Mutually Exclusive Events)*

Statement: If A and B are any two mutually exclusive events of a random experiment, then $P(A \cup B) = P(A) + P(B)$.

Proof: Let, A and B are any two events of a random experiment and S is the sample space of that random experiment. Thus, according to the definition of event, $A, B \subseteq S$.

Let, the total number of possible outcomes of that random experiment is $n(S)$ and the numbers of favourable cases for the events A and B are respectively $n(A)$ and $n(B)$.

Since, A and B are two mutually exclusive events, A and B never occurs simultaneously, i.e., $A \cap B = \phi$. Thus, A and B have no common favourable cases, i.e., $n(A \cap B) = 0$.

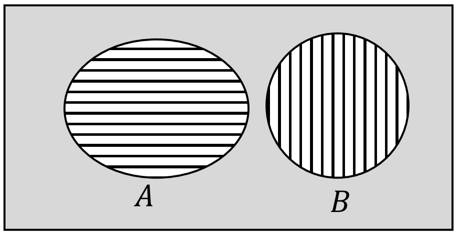


Figure 2.2. Venn diagram representation of two mutually exclusive events A and B of a random experiment.

Now, if the number of cases favourable to the event 'either A or B occur' (i.e., at least one occur) is $n(A \cup B)$, then

$$\therefore n(A \cup B) = n(A) + n(B)$$
$$\Rightarrow \frac{n(A \cup B)}{n(S)} = \frac{n(A)}{n(S)} + \frac{n(B)}{n(S)}$$
$$\Rightarrow P(A \cup B) = P(A) + P(B)$$

Generalization: This theorem can be extended to three or more mutually exclusive events of a random experiment. The probability of occurrence of any one of the several mutually exclusive events A, B and C is equal to the sum of their individual probabilities given by

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

In general, if $A_1, A_2 \cdots, A_n$ are mutually exclusive events of a random experiment, then

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = P(A_1) + P(A_2) + \cdots + P(A_n)$$

i.e., the probability of occurrence of any one of the n mutually disjoint events $A_1, A_2 \cdots, A_n$ is equal to the sum of their individual probabilities.

Note:

If n mutually exclusive events $A_1, A_2 \cdots, A_n$ are exhaustive also, so that probability of at least one of the n events to materialize is a certainty then the probability of the constituent events.

$$\therefore P(A_1 \cup A_2 \cup \cdots \cup A_n) = 1$$
$$\Rightarrow P(A_1) + P(A_2) + \cdots + P(A_n) = 1$$

Worked out Problems:

Problem 2.1. A card is drawn at random from a pack of 52 cards. Find the probability that the drawn card is either a club or an ace of diamond.

Solution: Let A : Event of drawing a card of club and B: Event of drawing an ace of diamond

A card can be drawn from a well shuffled deck of 52 cards in ${}^{52}\mathrm{C}_1$ ways, i.e., in 52 ways. Since, well shuffled pack of cards ensures equally likely outcomes; the total number of possible outcomes is 52.

Now, a well shuffled pack of 52 cards contains 13 cards of club. So, to ensure the occurrence of event A, the card must be selected from those 13 cards of club which can be done in ${}^{13}\mathrm{C}_1$ ways, i.e., in 13 ways. Thus, the number of outcomes favourable to the event A is 13.

$\therefore$ The probability of drawing a card of club $= P(A) = \frac{13}{52} = \frac{1}{4}$.

Again, a well shuffled pack of 52 cards contains 13 cards of diamond out of which only one card is ace. So, to ensure the occurrence of event B, that ace of diamond must be selected which can be done in ${}^{1}\mathrm{C}_1$ ways, i.e., in 1 ways. Thus, the number of outcomes favourable to the event B is 1.

$\therefore$ The probability of drawing a card of diamond $= P(B) = \frac{1}{52}$.

Since, the selected card can not be simultaneously a card of club and an ace of diamond; the two events A and B are mutually exclusive.

$\therefore$ The probability of the drawn card being a club or an ace of diamond $= P(A \cup B)$

$= P(A) + P(B)$

[By addition theorem of probability for mutually exclusive events]

$$= \frac{1}{4} + \frac{1}{52}$$

$$= \frac{13 + 1}{52}$$

$$= \frac{14}{52}$$

$$= \frac{7}{26}$$

Problem 2.2. A herd contains 30 cows numbered from 1 to 30. One cow is selected at random. Find the probability that number of the selected cow is a multiple of 5 or 8.

Solution: Let A be the event of number being a multiple of 5 within 30 and B be the event of number being a multiple of 8 within 30.

A cow can be selected from the 30 cows numbered from 1 to 30 in ${}^{30}\mathrm{C}_1$ ways, i.e., in 30 ways. Thus, the total number of possible outcomes is 30.

Now, there are exactly 6 numbers within 30 which are multiple of 5 (since, 6 is the quotient when 30 is divided by 5). So, to ensure the occurrence of event A, the cow must be selected only from those 6 cows numbered by multiple of 5 which can be done in ${}^{6}\mathrm{C}_1$ ways, i.e., in 6 ways. Thus, the number of outcomes favourable to the event A is 6.

$\therefore$ The probability of drawing a cow numbered by a multiple of 5 $= P(A) = \frac{6}{30} = \frac{1}{5}$.

Again, there are exactly 3 numbers within 30 which are multiple of 8 (since, 3 is the quotient when 30 is divided by 8). So, to ensure the occurrence of event B, the cow must be selected only from those 3 cows numbered by multiple of 8 which can be done in ${}^{3}\mathrm{C}_1$ ways, i.e., in 3 ways. Thus, the number of outcomes favourable to the event B is 3.

$\therefore$ The probability of drawing a cow numbered by a multiple of 8 $= P(B) = \frac{3}{30} = \frac{1}{10}$.

Since, the least common multiple of 5 and 8, i.e., $l.c.m.(5,8)$ is 40; there is no number within 30 which multiple of both 5 and 8. Thus, the number of the selected cow can not be simultaneously a multiple of 5 and multiple of 8, i.e., the two events A and B are mutually exclusive.

Thus, by addition theorem of probability,

$\therefore$ The probability of drawing a cow numbered by a multiple of 5 or 8 $= P(A \cup B)$

$= P(A) + P(B)$ [By addition theorem of probability for mutually exclusive events]

$= \frac{1}{5} + \frac{1}{10}$

$= \frac{2+1}{10}$

$= \frac{3}{10}$

2.3. ADDITION THEOREM OF PROBABILITY (*For Non Mutually Exclusive Events)*

Statement: If A and B are any two mutually exclusive events of a random experiment, then $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Proof: Let, A and B are any two events of a random experiment and S is the sample space of that random experiment.
According to the definition of event, $A, B \subseteq S$.
Let, the total number of possible outcomes of that random experiment is $n(S)$ and the numbers of favourable cases for the events A and B are respectively $n(A)$ and $n(B)$.
Since, A and B are two non-mutually exclusive events, A and B may occur simultaneously, i.e., $A \cap B \neq \phi$. Thus, A and B have some common favourable cases, i.e., $n(A \cap B) \neq 0$.

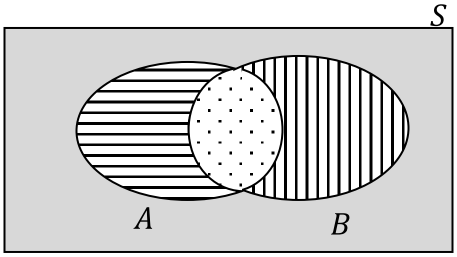


Figure 2.3. Venn diagram representation of two non-mutually exclusive events A and B of a random experiment.

Now, if the number of cases favourable to the event 'either A or B occur' (i.e., at least one occur) is $n(A \cup B)$, then

$$\therefore n(A \cup B) = n(A) + n(B) - n(A \cap B)$$
$$\Rightarrow \frac{n(A \cup B)}{n(S)} = \frac{n(A)}{n(S)} + \frac{n(B)}{n(S)} - \frac{n(A \cap B)}{n(S)}$$
$$\Rightarrow P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Generalization: The above theorem can be extended to three or more events. If A, B and C are not mutually exclusive events then the probability of the occurrence of at least one of them is given by

$P(A \cap B \cap C) = P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) - P(A \cap C) + P(A \cap B \cap C)$.

In general, if $A_1, A_2 \cdots, A_n$ are n non-mutually exclusive events of a random experiment, then

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) \\ = \sum_{i=1}^{n} P(A_i) - \sum_{\substack{0\le i,j \le n \\ 0<i<j}} P(A_i \cap A_j) + \sum_{\substack{0\le i,j,k \le n \\ 0<i<j<k}} P(A_i \cap A_j \cap A_k) - \cdots \\ + (-1)^{n-1} P(A_1 \cap A_2 \cap A_3 \cap \cdots \cap A_n)$$

i.e., the probability of occurrence of any one of the n mutually disjoint events $A_1, A_2 \cdots, A_n$ is equal to the sum of their individual probabilities.

The following worked out problems illustrate the application of this theorem:

Worked out Problems:

Problem 2.3. A card is drawn at random from a pack of 52 cards. Find the probability that the drawn card is either a spade or a king.

Solution: Let A: Event of drawing a card of spade and B: Event of drawing a king card.

∴ The probability of drawing a card of spade $\mathrm{P(A)} = \frac{13}{52}$

∴ The probability of drawing a king card $\mathrm{P(B)} = \frac{4}{52}$

Because one of the kings is a spade card also therefore, these events are not mutually exclusive. ∴ The probability of drawing a king $\mathrm{P(A \cap B)} = \frac{1}{52}$ of spade is
So, the probability of the drawing a spade or king card =
$$\mathrm{P(A \cup B) = P(A) + P(B) - P(A \cap B)} = \frac{13}{52} + \frac{4}{52} - \frac{1}{52} = \frac{16}{52} = \frac{4}{13}$$

Problem 2.4. A herd contains 30 cows numbered from 1 to 30. One cow is selected at random. Find the probability that the number of the selected cow is a multiple of 5 or 6.

Solution: Let A be the event of number being a multiple of 5 within 30 and B be the event of number being a multiple of 6 within 30.

Favourable cases for event A are {5, 10, 15, 20, 25, 30}
Similarly favourable cases for event B are {6, 12, 18, 24, 30}
The probability of the number being a multiple of 5 within 30 is $P(A) = \frac{6}{30}$
The probability of the number being a multiple of 6 within 30 is $P(B) = \frac{5}{30}$

Since 30 is a multiple of 5 as well as 6, therefore the events are not mutually exclusive
$$\mathrm{P(A \cap B) = P(A\ and\ B)} = \frac{1}{30}$$
The probability that the number of the selected cow is a multiple of 5 or 6 is :
$$\mathrm{P(A \cup B) = P(A) + P(B) - P(A \cap B)} = \frac{6}{30} + \frac{5}{30} - \frac{1}{30} = \frac{10}{30} = \frac{1}{10}$$

Problem 2.5. A number was drawn at random from the number 1 to 50. What is the probability that it will be a multiple of 2 or 3 or 10?

Solution: Probability of getting a multiple of 2: $P(A) = \frac{25}{50}$

Probability of getting a multiple of 3: $P(B) = \frac{16}{50}$

Probability of getting a multiple of 10: $P(C) = \frac{5}{50}$

Common Probability of getting a multiple of 2 and 3: $P(A \cap B) = \frac{8}{50}$

Common Probability of getting a multiple of 3 and 10: $P(B \cap C) = \frac{1}{50}$

Common Probability of getting a multiple of 2 and 10: $P(A \cap C) = \frac{5}{50}$

Common Probability of getting a multiple of 2, 3 and 10: $P(A \cap B \cap C) = \frac{1}{50}$

Probability that it is a multiple of 2 or 3 or 10:

$$\therefore P(A \cap B \cap C) = P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) - P(A \cap C) + P(A \cap B \cap C)$$

$$= \frac{25}{50} + \frac{16}{50} + \frac{5}{50} - \frac{8}{50} - \frac{1}{50} - \frac{5}{50} + \frac{1}{50}$$

$$= \frac{33}{50}$$

LECTURE 3: CONDITIONAL PROBABILITY

3.1. INTRODUCTION

In the previous lessons, we learned how to find probability by classical definition as well as by addition theorem of probability. Conditional probability answers the question ‘how does the probability of an event change if we have extra information’. We’ll illustrate with an example.

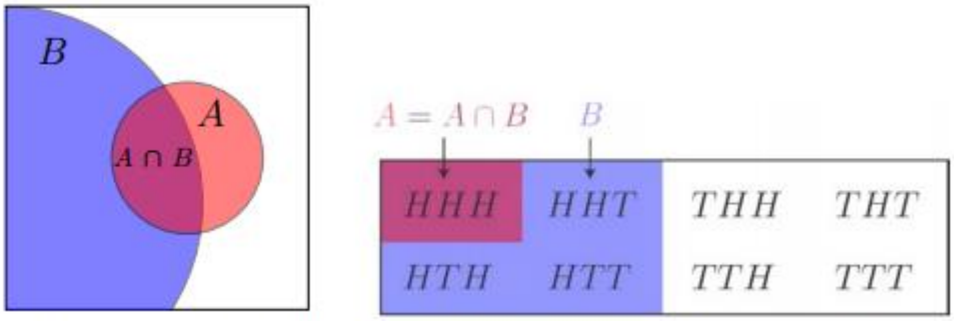
Example 3.1. Toss a fair coin 3 times.
(a) What is the probability of 3 heads?
answer: Sample space $\Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$. All outcomes are equally likely, so $P(3\ heads) = 1/8$.
(b) Suppose we are told that the first toss was heads. Given this information how should we compute the probability of 3 heads?
answer: We have a new (reduced) sample space: $\Omega' = \{HHH, HHT, HTH, HTT\}$.
All outcomes are equally likely, so

$$P(3\ heads\ given\ that\ the\ first\ toss\ is\ heads) = 1/4.$$

This is called conditional probability, since it takes into account additional conditions. To develop the notation, we rephrase (b) in terms of events.
Rephrased (b) Let A be the event ‘$all\ three\ tosses\ are\ heads$’ $= \{HHH\}$.
Let B be the event ‘$the\ first\ toss\ is\ heads$’ $= \{HHH, HHT, HTH, HTT\}$.
The conditional probability of A knowing that B occurred is written as $P(A|B)$.
This is read as

‘the conditional probability of A given B’
or
‘the probability of A conditioned on B’
or simply
‘the probability of A given B’.

We can visualize conditional probability as follows. Think of $P(A)$ as the proportion of the area of the whole sample space taken up by A. For $P(A|B)$ we restrict our attention to B. That is, $P(A|B)$ is the proportion of area of B taken up by A, i.e. $P(A \cap B)/P(B)$.



Conditional probability: Abstract visualization and coin example

Note, $A \subset B$ in the right-hand figure, so there are only two colors shown. The formal definition of conditional probability catches the gist of the above example and visualization.

3.2. FORMAL DEFINITION OF CONDITIONAL PROBABILITY

Definition 3.1.

Let A and B be events. We define the conditional probability of A given B as

$$P(A|B) = \frac{P(A \cap B)}{P(\,B)}\text{, provided } P(B) \neq 0. \qquad (1)$$

Let's redo the coin tossing example using the definition in Equation (1). Recall $A =$ '3 $heads$' and $B =$ '$first\ toss\ is\ heads$'. We have $P(A) = 1/8$ and $P(B) = 1/2$. Since $A \cap B = A$, we also have $P(A \cap B) = 1/8$. Now according to (1), $P(A|B) = \frac{1/8}{1/2} = 1/4$, which agrees with our answer in Example 1b.

Worked out Problems:

Problem 3.1. Toss two fair coins, blindfolded. Somebody tells you that you tossed at least one Heads. What is the probability that both your tosses are Heads?
Solution: Let, A = {both H}, B = {at least one H}

$$P(A|B) = \frac{P(A \cap B)}{P(\,B)} = \frac{P(both\,H)}{P(\,at\ least\ one\ H)} = \frac{\frac{1}{4}}{\frac{3}{4}} = \frac{1}{3}$$

3.3. MULTIPLICATION RULE OF PROBABILITY

The probability that two events *A* and *B* both occur is given by the *multiplication rule of probability* as:

$$P(A \cap B) = P(A\,|\,B) \times P(B)$$

or by:

$$P(A \cap B) = P(B\,|\,A) \times P(A)$$

Extension of Multiplication Rule of Probability
The multiplication rule can be extended to three or more events. In the case of three events, the rule looks like this:

$$P(A \cap B \cap C) = P\big((A \cap B) \cap C\big) = P(C\,|\,A \cap B) \times P(A \cap B)$$

But, since $P(A \cap B) = P(B\,|\,A) \times P(A)$, we have

$$P(A \cap B \cap C) = P(C\,|\,A \cap B) \times P(B\,|\,A) \times P(A)$$

Worked out Problem:

Problem 3.1. A bag contains 3 pink candies and 7 green candies. Two candies are taken out from the bag with replacement. Find the probability that both candies are pink.

Solution:

Let A = event that first candy is pink and B = event that second candy is pink.

$\therefore P\,(A) = \frac{3}{10}$ …(i)

Since, after the occurrence of the event A, the selected candy is replaced in the bag; the number of candies in the bag will remains unchanged.

$\therefore P\ (\mathrm{B}|\mathrm{A}) = P\ (B)\ =\ \frac{3}{10}$ …(ii)

Hence by the multiplication law we get,

$$P(A\ \cap\ B)\ =\ P(B\ |\ A)\ \times\ P(A)$$
$$\Rightarrow P(A\ \cap\ B)\ =\ \frac{3}{10} \times \frac{3}{10} = \frac{9}{100} = 0.09$$

Problem 3.2. A bag has 4 white cards and 5 blue cards. We draw two cards from the bag one by one without replacement. Find the probability of getting both cards white.

Solution: Let A = event that first card is white and B= event that second card is white.

$\therefore P\ (A)\ =\ \frac{4}{9}$ …(i)

Now, since the events given are dependent on each other, we have

$$P\ (B) \neq P\ (B|A)$$

Once the event A has occurred, the bag contains 3 white cards and 5 blue cards.

$$\therefore P\ (B|A) = \frac{3}{8}$$

Hence by the multiplication law we get,

$$P(A\ \cap\ B)\ =\ P(B\ |\ A) \times\ P(A) = \frac{3}{8} \times \frac{4}{9} = \frac{1}{6}$$

LECTURE 4: INDEPENDENT EVENTS & TOTAL PROBABILITY THEOREM

Definition 4.1a. (*Independent Events*)
Events A and B are *independent events* if the occurrence of one of them does not affect the probability of the occurrence of the other. That is, two events are independent if either:

$$P(B|A) = P(B), \text{ (provided that } P(A) > 0)$$

or:

$$P(A|B) = P(A), \text{ (provided that } P(B) > 0).$$

This substitution leads us to an alternative definition of independence.

Definition 4.1b. (*Independent Events*)
Events *A* and *B* are *independent events* if and only if:

$$P(A \cap B) = P(A) \times P(B)$$

Otherwise, *A* and *B* are called *dependent events*.

Recall that the "if and only if" (often written as "iff") in that definition means that the if-then statement works in both directions. That is, the definition tells us two things:

(i) If events A and B are independent, then $P(A \cap B) = P(A) \times P(B)$.
(ii) If $P(A \cap B) = P(A) \times P(B)$, then events A and B are independent.

The next example illustrates the first of these two directions, while the second example illustrates the second direction.

<u>Worked out Problem:</u>

Problem 4.1. A recent survey of students suggested that 10% of Penn State students commute by bike, while 40% of them have a significant other. Based on this survey, what percentage of Penn State students commute by bike *and* have a significant other?

Solution: Let's let *B* be the event that a randomly selected Penn State student commutes by bike, and *S* be the event that a randomly selected Penn State student has a significant other. If *B* and *S* are independent events (okay??), then the definition tells us that:

$$P(B \cap S) = P(B) \times P(S) = 0.10 \times 0.40 = 0.04$$

That is, 4% of Penn State students commute by bike *and* have a significant other.

Theorem 4.1.

Statement: If A and B are two independent events, then the following pairs of events are independent:

(i) $\bar{A}$ and $\bar{B}$
(ii) A and $\bar{B}$
(iii) $\bar{A}$ and B

Proof: Since A and B are independent, so

$$P(A \cap B) = P(A)P(B) \text{ .... (1)}$$

(i) Now, $P(\bar{A} \cap \bar{B}) = P(\overline{A \cup B})$ [by D'Morgan's law]

$$= 1 - P(A \cup B) = 1 - [P(A) + P(B) - P(A \cap B)]$$
$$= 1 - P(A) - P(B) + P(A \cap B)$$

$$= 1 - P(A) - P(B) + P(A)P(B) \quad \text{[by (i)]}$$
$$= \big(1 - P(A)\big)(1 - P(B))$$
$$= P(\bar{A})P(\bar{B}).$$

$\therefore \bar{A}$ and $\bar{B}$ are independent.

(ii) Again, $A \cap B$ and $A \cap \bar{B}$ are mutually exclusive and A can be expressed as follows:

$$A = (A \cap B) \cup (A \cap \bar{B})$$
$$\therefore P(A) = P(A \cap B) + P(A \cap \bar{B})$$
$$\Rightarrow P(A \cap \bar{B}) = P(A) - P(A \cap B)$$
$$= P(A) - P(A)P(B) \quad \text{[by (1)]}$$
$$= P(A)(1 - P(B))$$
$$= P(A)P(\bar{B}).$$

$\therefore A$ and $\bar{B}$ are independent.

(iii) Taking, $B = (A \cap B) \cup (\bar{A} \cap B)$, we can prove the result as (ii).

Worked out Problem:

Problem 4.2. A problem in statistics is given to three students $\mathrm{A, B,}$ and C whose chances of solving it are ½, ¾, and ¼ respectively. What is the probability that the problem will be solved if all of them try independently?

Solution: Let $\mathrm{A, B, C}$ denote the events that the problem is solved by the students $\mathrm{A, B, C}$ respectively. Then,

$\mathrm{P(A)} = \frac{1}{2}$, $\mathrm{P(B)} = \frac{3}{4}$, $\mathrm{P(C)} = \frac{1}{4}$

The problem will be solved if at least one of them solves the problem. Thus we have to calculate the probability of occurrence of at least one of the three events $\mathrm{A, B, C, i.e.\, P(A \cup B \cup C)}$.

$$\mathrm{P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)}$$

$$\mathrm{= P(A) + P(B) + P(C) - P(A)P(B) - P(A)P(C) - P(B)P(C) + P(A)P(B)P(C)}$$

($\therefore \mathrm{A, B, C}$ are mutually independent events.)

$$= \frac{1}{2} + \frac{3}{4} + \frac{1}{4} - \frac{1}{2}.\frac{3}{4} - \frac{1}{2}.\frac{1}{4} - \frac{3}{4}.\frac{1}{4} + \frac{1}{2}.\frac{3}{4}.\frac{1}{4} = \frac{29}{32}.$$

Theorem 4.2. (*Total Probability Theorem*)

Statement: If $A_1, A_1, \cdots, A_n$ are mutually exclusive and mutually exhaustive events of some random experiment and $B(\neq \phi)$ be any other event of the same random experiment, then the probability of the events B, denoted by $P(B)$, is as follows

$$P(B) = \sum_{i=1}^{n} P(B|A_i)P(A_i)$$

Worked out Problem:

Problem 4.3. Three bags contain 100 marbles each:

- Bag 1 has 75 red and 25 blue marbles;
- Bag 2 has 60 red and 40 blue marbles;
- Bag 3 has 45 red and 55 blue marbles

One of the bags chosen at random and then pick a marble from the chosen bag, also at random. What is the probability that the chosen marble is red?

Solution

Let R be the event that the chosen marble is red. Let B_i be the event that i^{th} Bag is chosen.

$\therefore P(B_i) = \frac{1}{3}, i = 1,2,3.$

We already know that

$$P(R|B_1) = 0.75,$$
$$P(R|B_2) = 0.60,$$
$$P(R|B_3) = 0.45$$

Using the law of total probability, we can write

$$\begin{aligned} P(R) &= P(R|B_1)P(B_1) + P(R|B_2)P(B_2) + P(R|B_3)P(B_3) \\ &= 0.75 \times \frac{1}{3} + 0.60 \times \frac{1}{3} + 0.45 \times \frac{1}{3} \\ &= 0.60 \end{aligned}$$

LECTURE 5 BAYES' THEOREM.

5.1.INTRODUCTION

In this lesson we extend the discussion of conditional probability to include applications of *Bayes' theorem* (or *Bayes' rule*), which we use for revising a probability value based on additional information that is later obtained. One key to understanding the essence of Bayes' theorem is to recognize that we are dealing with *sequential* events, whereby new additional information is obtained for a subsequent event, and that new information is used to revise the probability of the initial event. In this context, the terms *prior probability* and *posterior probability* are commonly used.

Definitions 5.1. (*Prior Probability*)

A *prior probability* is an initial probability value originally obtained before any additional information is obtained.

Definition 5.2. (*Posterior Probability*)

A *posterior probability* is a probability value that has been revised by using additional information that is later obtained.

5.2. BAYES' THEOREM

Statement: If $A_1, A_1, \cdots, A_n$ are mutually exclusive and mutually exhaustive events of some random experiment and $B(\neq \phi)$ be any other event of the same random experiment, then the probability of any one of the events A_i, for some i, given that the event B has already occurred, denoted by $P(A_i|B)$, is as follows

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{P(B)} - \qquad -$$

where $P(B) = \sum_{i=1}^{n} P(B|A_i)P(A_i)$

That's a formidable expression, but we will simplify its calculation. See the following example, which illustrates use of the above expression, but also see the alternative method based on a more intuitive application of Bayes' theorem.

<u>Worked out Problems:</u>

Problem 10.1. In Orange County, 51% of the adults are males. (It doesn't take too much advanced mathematics to deduce that the other 49% are females.) One adult is randomly selected for a survey involving credit card usage.

a. Find the prior probability that the selected person is a male.

b. It is later learned that the selected survey subject was smoking a cigar. Also, 9.5% of males smoke cigars, whereas 1.7% of females smoke cigars (based on data from the Substance Abuse and Mental Health Services Administration). Use this additional information to find the probability that the selected subject is a male.

Solution: Let's use the following notation:

M: the event that the adult is male
$\bar{M}$: the event that the adult is female
C: the event that the adult is cigar smoker
$\bar{C}$: the event that the adult is not cigar smoker

a. Before using the information given in part b, we know only that 51% of the adults in Orange County are males, so the probability of randomly selecting an adult and getting a male is given by P(M) = 0.51.

b. Based on the additional given information, we have the following:

$P(M) = 0.51$ [Since, 51% of the adults are males]

$P(\bar{M}) = 0.49$ [Since, 49% of the adults are females (not males)]

$P(\mathrm{C|M}) = 0.095$ [Since, 9.5% of the males smoke cigars (That is, the probability of getting someone who smokes]

$P(\mathrm{C|M}) = 0.017$ [Since, 1.7% of the females smoke cigars (That is, the probability of getting someone who smokes cigars, given that the person is a female, is 0.017.))]

Let's now apply Bayes' theorem by using the preceding formula with M in place of A, and C in place of B. We get the following result:

$$P(M \mid C) = \frac{P(C|\, M)P(\, M)\,)}{P(\, C)}$$

$$= \frac{P(C|\, M)P(\, M)}{P(C|\, M)P(\, M) + P(C|\, \bar{M})P(\, \bar{M})}$$

$$= \frac{0.095 \times 0.541}{0.095 \times 0.541 + 0.017 \times 0.49}$$

$$= 0.853$$

Before we knew that the survey subject smoked a cigar, there is a 0.51 probability that the survey subject is male (because 51% of the adults in Orange County are males). However, after learning that the subject smoked a cigar, we revised the probability to 0.853. There is a 0.853 probability that the cigar−smoking respondent is a male. This makes sense, because the likelihood of a male increases dramatically with the additional information that the subject smokes cigars (because so many more males smoke cigars than females).

Problem 10.2. An aircraft emergency locator transmitter (ELT) is a device designed to transmit a signal in the case of a crash. The Altigauge Manufacturing Company makes 80% of the ELTs, the Bryant Company makes 15% of them, and the Chartair Company makes the other 5%. The ELTs made by Altigauge have a 4% rate of defects, the Bryant ELTs have a 6% rate of defects, and the Chartair ELTs have a 9% rate of defects (which helps to explain why Chartair has the lowest market share).

a. If an ELT is randomly selected from the general population of all ELTs, find the probability that it was made by the Altigauge Manufacturing Company.

b. If a randomly selected ELT is then tested and is found to be defective, find the probability that it was made by the Altigauge Manufacturing Company.

Solution:

We use the following notation:

A = ELT manufactured by Altigauge
B = ELT manufactured by Bryant
C = ELT manufactured by Chartair

D = ELT is defective
$\overline{\text{D}}$ = ELT is not defective (or it is good)

a. If an ELT is randomly selected from the general population of all ELTs, the probability that it was made by Altigauge is 0.8 (because Altigauge manufactures 80% of them).

b. If we now have the additional information that the ELT was tested and was found to be defective, we want to revise the probability from part (a) so that the new information can be used. We want to find the value of P(A|D), which is the probability that the ELT was made by the Altigauge company given that it is defective. Based on the given information, we know these probabilities:

$P(A) = 0.80$ [Since, Altigauge makes 80% of the ELTs]
$P(B) = 0.15$ [Since, Bryant makes 15% of the ELTs]
$P(C) = 0.05$ [Since, Chartair makes 5% of the ELTs]

$P(\mathrm{D}|A) = 0.04$ [Since, 4% of the Altigauge ELTs are defective]
$P(\mathrm{D}|B) = 0.06$ [Since, 6% of the Bryant ELTs are defective]
$P(\mathrm{D}|C) = 0.09$ [Since, 9% of the Chartair ELTs are defective]

Here is Bayes' theorem extended to include three events corresponding to the selection of ELTs from the three manufacturers (A, B, C):

$$P(A|D) = \frac{P(D|\ A) \times P(\ A)}{P(\ D)}$$

$$= \frac{P(D|\ A) \times P(\ A)}{P(D|\ A) \times P(\ A) + P(D|\ B) \times P(\ B) + P(D|\ C) \times P(\ C)}$$

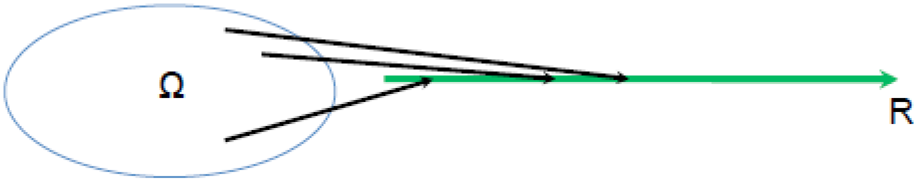
$$= \frac{0.04 \times 0.80}{0.04 \times 0.80 + 0.06 \times 0.15 + 0.09 \times 0.05}$$

$$= 0.703$$

LECTURE 6: DISCRETE RANDOM VARIABLE AND IT'S PROBABILITY DISTRIBUTION.

6.1. RANDOM VARIABLE

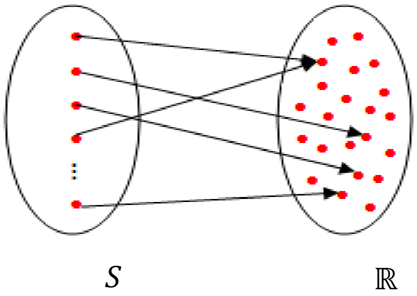
We are often interested not in the exact outcome of a random experiment but only some consequence of it (for example I toss 3 coins but I only care about how many heads occur not the exact outcome). In informal way of thinking, a random variable is an assignment of a value (number) – real or complex - to every outcome in the sample space. This number is called the numerical value or the experimental outcome of the random variable. But, we will restrict our discussion only on real valued random variable, i.e., the random variable which is a function from the sample space to the set of real numbers.



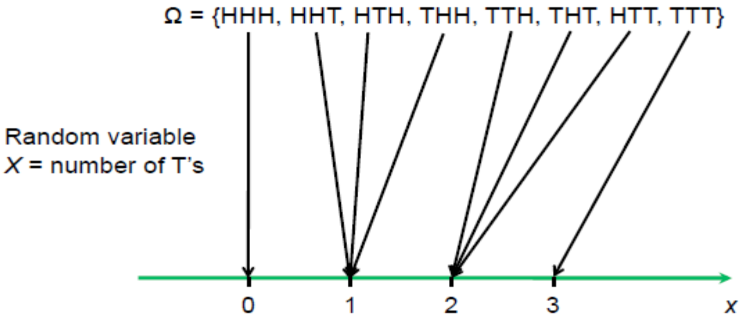
Definition 6.1. (*Random Variable*)

A random variable is a real-valued function defined on the sample space of a random experiment, i.e., a function from S to $\mathbb{R}$.

Notation: Generally, the R.V.s are denoted by uppercase letters of English alphabets, viz., *X*, *Y*, *Z*, *etc*. and their corresponding values are denoted by *x*, *y*, *z*, etc., respectively. Thus, if X is a R.V., then $X: S \to \mathbb{R}$. So, $\forall s \in S, X(s) = x \in \mathbb{R}$.



Example 6.1: Let, a coin is tossed three times and *X* is the number of T's in the outcome.



Remarks:

- An (informal) way of thinking is to regard a random variable as a measure about the outcome which takes a real number (e.g. how many heads occur in a coin
- tossing experiment) or a measurement made on the outcome.

- For a given a random variable X on S, statements like "$X = 3$" or "$X \leq 3$" are events. Specifically, "$X = 3$" is the event $\{s \in S : X(s) = 3\}$, i.e. the set of all outcomes in S for which X takes the value 3. Similarly, "$X \leq 3$" is the event $\{s \in S: X(s) \leq 3\}$ i.e. that is the set of all outcomes in S for which X takes a value at most 3.
- For a random variable X, the range of X is the set of all values taken by X. We denote it by Range (X) and is commonly known as *spectrum* of the random variable.
- Functions of random variables: Any function you are likely to run across of a random variable or random variables is a random variable. So if X and Y are random variables, then $X \pm Y$, XY , and $logX$ are all random variables.

6.2. TYPES OF RANDOM VARIABLES

Depending on the nature of possible values of the random variable, random variables are categorized into following two type:

- Discrete random variables – random variables that take on either finite or countably infinite number of values. For example, Number of phone messages awaiting me.
- Continuous random variables – random variables that take on a continuously infinite number of values. For example: Weight of a people selected randomly

6.3. DISTRIBUTION OF A RANDOM VARIABLE

The distribution of a random variable describes the probability that it takes on various values. In informal way of speaking, the probability distribution of a random variable is a representation which tells us which are the possible values and how the total probability is distributed over the several possible values.

Based on type of random variable, probability distribution of random variables are categorized as *discrete probability distribution* & *continuous probability distribution*.

6.4. DISCRETE PROBABILITY DISTRIBUTION

A discrete probability distribution is characterized by it's *probability mass function* (p.m.f).

Definition 6.2. (*Probability Mass Function*)

The probability mass function of a discrete random variable X is a function $f: \mathbb{R} \rightarrow [0,1]$, defined as follows:

$$f(x) = \begin{cases} P(X = x), \forall x \in spectrum \\ 0, \forall x \in \mathbb{R} - spectrum \end{cases}$$

<u>Properties of p.m.f.</u>

If f is the p.m.f. of a discrete random variable X, then f satisfies the following properties:

(i) $f(x) \geq 0, \forall x \in \mathbb{R}$

(ii) $\sum_{\forall x \in \mathbb{R}} f(x) = 1$

Definition 6.3. (*Cumulative Distribution Function***)**

The *cumulative distribution function* (c.d.f.) or simply *Distribution Function* (d.f) of a discrete random variable X is a function $F: \mathbb{R} \to [0,1]$, defined as follows:

$$F(x) = P(X \leq x), \forall x \in \mathbb{R} = \sum_{\forall t \epsilon \mathbb{R}: t \leq x} f(t)\,, \forall x \in \mathbb{R}$$

where f is the p.m.f. of the distribution.

Properties of c.d.f.

F satisfies the following properties:

(i) $F(x) \geq 0, -\infty < x < \infty$

(ii) $F(-\infty) = 0$ and $F(\infty) = 1$

(iii) $F(x_1) \leq F(x_1), if\ x_1 < x_2, \forall x_1, x_2 \in \mathbb{R}$

6.5. EXPECTATION, VARIANCE & S.D. OF DISCRETE RANDOM VARIABLES

Definition 6.4 (*Expectation*)

Let X be a discrete random variable and $f(x)$ be the probability mass function. Then expected value of X, denoted by $E(X)$, is defined by

$$E(X) \ = \sum_{x \epsilon \mathbb{R}} x f(x)$$

provided this sum converges absolutely. We often refer to the expected value as the mean, and denote $E(X)$ by μ for short. If the above sum does not converge absolutely, then we say that X does not have an expected value.

Worked out Problem:

Problem 6.1. Let an experiment consist of tossing a fair coin three times. Find the expected number of heads.

Solution: Let X denote the number of heads which appear. Then the possible values of X are 0, 1, 2 and 3. The corresponding probabilities are 1/8, 3/8, 3/8, and 1/8. Thus, the expected value of X is as follows:

$$E(X) = 0.\frac{1}{8} + 1.\frac{3}{8} + 2.\frac{3}{8} + 6.\frac{1}{8} = \frac{3}{2}$$

Definition 6.5 (*Variance*)

Let X be a discrete random variable and $f(x)$ be the probability mass function. Then variance of X, denoted by $V(X)$, is defined by

$$V(X) = E(X^2) - \{E(X)\}^2 = \sum_{x \epsilon \mathbb{R}} x^2 f(x) - \left\{ \sum_{x \epsilon \mathbb{R}} x f(x) \right\}^2$$

provided this sum converges absolutely.

Worked out Problem:

Problem 6.3. Let an experiment consist of tossing a fair coin three times. Find the variance of the number of heads appears.

Solution: Let X denote the number of heads which appear. Then the possible values of X are 0, 1, 2 and 3. The corresponding probabilities are 1/8, 3/8, 3/8, and 1/8. Thus, the expected value of X is as follows:

$$\therefore E(X) = 0.\frac{1}{8} + 1.\frac{3}{8} + 2.\frac{3}{8} + 6.\frac{1}{8} = \frac{3}{2}$$

$$E(X^2) = 0^2.\frac{1}{8} + 1^2.\frac{3}{8} + 2^2.\frac{3}{8} + 6^2.\frac{1}{8} = \frac{51}{8}$$

$$\therefore V(X) = E(X^2) - \{E(X)\}^2 = \frac{51}{8} - \left\{\frac{3}{2}\right\}^2 = \frac{51}{8} - \frac{9}{4} = \frac{51-18}{8} = \frac{33}{8}$$

Definition 6.6 (*Standard Deviation*)

Standard deviation of a discrete random variable X is the positive square root of variance, i.e., $s.d.of\ X = \sqrt{V(X)}$.

LECTURE 7: CONTINUOUS RANDOM VARIABLE AND IT'S PROBABILITY DISTRIBUTION.

A continuous probability distribution is characterized by it's *probability density function* (p.d.f.).

Definition 7.1. (*Probability Density Function***)**

The probability density function of a continuous random variable X is a function $f: \mathbb{R} \rightarrow [0,1]$, defined as follows:

$$f(x)\varepsilon \cong P\left[x - \frac{\varepsilon}{2} < X < x + \frac{\varepsilon}{2}\right], where\ \varepsilon > 0\ is\ small.$$

Note:

- $f(x)$ does not give the probability that the continuous random variable X takes on the value x.
- $f(x)\varepsilon$ is approximately equal to the probability that X takes on a value in an interval of length ε about x.
- For a continuous random variable X, $P(X = x) = 0, \forall x \in \mathbb{R}$

Properties of p.d.f.

If f is the p.d.f. of a continuous random variable X, then f satisfies the following properties:

(i) $f(x) \geq 0, \forall x \in \mathbb{R}$

(ii) $\int_{\forall x \in \mathbb{R}} f(x)dx = 1$

Definition 7.2. (*Cumulative Distribution Function***)**

The *cumulative distribution function* (c.d.f.) or simply *Distribution Function* (d.f) of a continuous random variable X is a function $F: \mathbb{R} \rightarrow [0,1]$, defined as follows:

$$F(x) = P(X \leq x), \forall x \in \mathbb{R} = \int_{\forall t \epsilon \mathbb{R} : t \leq x} f(t)dt\,, \forall t \leq x \in \mathbb{R}$$

where f is the p.d.f. of the distribution.

Properties of c.d.f.

F satisfies the following properties:

(i) $F(x) \geq 0, -\infty < x < \infty$

(ii) $F(-\infty) = 0$ and $F(\infty) = 1$

(iii) $F(x_1) \leq F(x_1), if\ x_1 < x_2, \forall x_1, x_2 \in \mathbb{R}$

EXPECTATION, VARIANCE & S.D. OF CONTINUOUS RANDOM VARIABLES

Definition 7.4 (*Expectation*)
Let X be a continuous random variable and $f(x)$ be the probability density function. Then expected value of X, denoted by $E(X)$, is defined by

$$E(X) = \int_{\forall x \in \mathbb{R}} x f(x) dx$$

Definition 7.5 (*Variance*)
Let X be a continuous random variable and $f(x)$ be the probability density function. Then variance of X, denoted by $V(X)$, is defined by

$$V(X) = E(X^2) - \{E(X)\}^2 = \sum_{x \epsilon \mathbb{R}} x^2 f(x) - \left\{ \sum_{x \epsilon \mathbb{R}} x f(x) \right\}^2$$

provided this sum converges absolutely.

Definition 7.6 (*Standard Deviation*)
Standard deviation of a continuous random variable X is the positive square root of variance, i.e., $s.d.of\ X = \sqrt{V(X)}$.

LECTURE 8: BINOMIAL DISTRIBUTION

Definition: A random variable X is said to follow **binomial distribution**, if it can assume only finite number of non-negative integral values and it's probability mass function (p.m.f.) is given by

$$f(x) = P(X = x) = \begin{cases} {}^{n}C_{x}p^{x}(1-p)^{n-x}\,, & x = 0, 1, 2, \dots., n;\ 0 \le p \le 1 \\ 0 \,, & otherwise \end{cases}$$

where the two independent constants n and p in the distribution are known as the parameters of the distribution.

Illustration: Let a random experiment be performed repeatedly, each repetition being called a trail and let the occurrence of an event in a trail be called a 'success' and it's non-occurrence be 'failure'. Now, in a series of $'n'$ independent trails, if the probability of 'success' in each trail is a constant $'p'$ and the probability of 'failure' in each trail is $'q'$ (where, $q = 1 - p$), then the probability of $'x'$ successes (and obviously $(n - x)$ failures) is given by the binomial distribution. So, it is clear that i)

Physical Condition for Binomial Distribution: We get the binomial distribution under the following experimental conditions:

i. Each trail results in two exhaustive and mutually disjoint outcomes, termed as success and failure.
ii. The number of trails $'n'$ is finite.
iii. The trails are independent of each other. Thus, the probability of success $'p'$ is same in each trail.

NOTE:

1. Binomial distribution is a discrete distribution as X can assume only finite number of isolated values.
2. Any random variable X which follows binomial distribution is known as binomial variate and is denoted by $X \sim B(n, p)$, where n and p are the parameters of the distribution. $\boldsymbol{np}$ **is the mean,** $\boldsymbol{npq}$ **is the variance** and $\sqrt{\boldsymbol{npq}}$ **is the standard deviation (s.d.)** of the normal distribution.
3. Let a random experiment be performed repeatedly, each repetition being called a trail and let the occurrence of an event in a trail be called a 'success' and it's non-occurrence be 'failure'. Now, in a series of $'n'$ independent trails, if the probability of 'success' in each trail is a constant $'p'$ and the probability of 'failure' in each trail is $'q'$ (where, $q = 1 - p$), then the probability of $'x'$ successes (and obviously $(n - x)$ failures) is given by the binomial distribution. So, it is clear that i)

4. **Physical Condition For Binomial Distribution:** We get the binomial distribution under the following experimental conditions:
   i. Each trail results in two exhaustive and mutually disjoint outcomes, termed as success and failure.
   ii. The number of trails $'n'$ is finite.
   iii. The trails are independent of each other. Thus, the probability of success $'p'$ is same in each trail.

Worked out Problem:

Problem 8.1. A biased coin is tossed 6 times. The probability of heads on any toss is 0.3. Let X denote the number of heads that come up. Calculate: (i) P(X = 2) (ii) P(X = 3) (iii) P(1 < X ≤ 5).

Solution: If we call heads a success then this X has a binomial distribution with parameters n = 6 and p = 0.3.

(i) $P(X = 2) = {}^{6}C_{2}\,(0.3)^{2}\,(0.7)^{4} = 0.324135.$

(ii) $P(X = 3) = {}^{6}C_{3}\,(0.3)^{3}\,(0.7)^{3} = 0.18522.$

(iii) $P(1 < X \leq 5) = P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5)$

$= 0.324 + 0.185 + 0.059 + 0.01$

$= 0.578$

Problem 8.2. A quality control engineer is in charge of testing whether or not 90% of the DVD players produced by his company conform to specifications. To do this, the engineer randomly selects a batch of 12 DVD players from each day's production. The day's production is acceptable provided no more than 1 DVD player fails to meet specifications. Otherwise, the entire day's production has to be tested.

(i) What is the probability that the engineer incorrectly passes a day's production as acceptable if only 80% of the day's DVD players actually conform to specifications?

(ii) What is the probability that the engineer unnecessarily requires the entire day's production to be tested if in fact 90% of the DVD players conform to specifications?

Solution: Let X denote the number of DVD players in the sample that fail to meet specifications.

X has a binomial distribution with parameters $n = 6$ and $p = 0.3$.

(i) X has a binomial distribution with parameters $n = 6$ and $p = 0.3$.

$P(X \leq 1) = P(X = 0) + P(X = 1)$

$= {}^{12}C_{0}\,(0.2)^{0}\,(0.8)^{12} + {}^{12}C_{1}\,(0.2)^{1}\,(0.8)^{11}$

$= 0.069 + 0.206$

$= 0.275$

(ii) X has a binomial distribution with parameters $n = 12$ and $p = 0.1$.

$P(X > 1) = 1 - P(X \leq 1)$

$= 1 - [\,P(X = 0) + P(X = 1)]$

$= 1 - \left[{}^{12}C_{0}\,(0.1)^{0}\,(0.9)^{12} + {}^{12}C_{1}\,(0.1)^{1}\,(0.9)^{11}\right]$

$= 1 - 0.659$

$= 0.341$

LECTURE 9: POISSON DISTRIBUTION

Definition: A random variable X is said to follow **Poisson distribution**, if it can assume only infinite number of non-negative integral values and it's probability mass function (p.m.f.) is given by

$$f(x) = P(X = x) = \begin{cases} \dfrac{e^{-\lambda}\lambda^x}{x!} , & x = 0, 1, 2, \ldots.; \lambda > 0 \\ 0 , & otherwise \end{cases}$$

where the independent constant λ in the distribution are known as the parameters of the distribution.

NOTE:

1. Poisson distribution is a discrete distribution as X can assume infinite number of isolated values.
2. Any random variable X which follows Poisson distribution is known as Poisson variate and is denoted by $X \sim P(\lambda)$, where λ is the parameters of the distribution.
3. Poisson distribution is such a discrete distribution where $mean = variance = \lambda$
4. **Physical Condition for Poisson Distribution:** Poisson distribution occurs when there are events which are do not occur as outcomes of definite number of trails (unlike that binomial distribution) of an experiment but which occur at random points of time and space. For examples,
   (i) the number of hits to a web site in a day;
   (ii) the number of calls that arrive in any one day on your mobile phone;
   (iii) the number of jobs arriving in any one minute in a busy computer centre;
   (iv) the number of messages arriving to a computer server in any one hour.

Worked out Problem:

Problem 9.1. The number of calls coming per minute into a hotels reservation center is Poisson random variable with mean 3. Find the probability that no calls come in a given 1 minute period.

Solution:

(a) Let X denote the number of calls coming in that given 1 minute period.

$$\therefore X \sim P(3)$$

Thus, $P(X = 0) = \frac{e^{-3}3^0}{0!} = e^{-3}$

Problem 9.2. Consider a computer system with Poisson job-arrival stream at an average of 2 per minute. Determine the probability that in any one-minute interval there will be

(i) 0 jobs;
(ii) exactly 2 jobs;
(iii) at most 3 arrivals.

Solution:

Let X denote the number of jobs -arrival in any one-minute.

$$\therefore X \sim P(2)$$

(i) No job arrivals:

$$P(X = 0) = e^{-2} = .135$$

(ii) Exactly 3 job arrivals:

$$P(X = 3) = \frac{e^{-2}2^{3}}{3!} = 0.18$$

(iii)At most 3 arrivals:

$$\begin{aligned} P(X = 3) &= P(0) + P(1) + P(2) + P(3) \\ &= e^{-2} + \frac{e^{-2}2^{1}}{1!} + \frac{e^{-2}2^{2}}{2!} + \frac{e^{-2}2^{3}}{3!} \\ &= 0.1353 + 0.2707 + 0.2707 + 0.1805 \\ &= 0.8571 \end{aligned}$$

LECTURE 10: NORMAL DISTRIBUTION

Definition: A random variable X is said to follow **normal distribution**, if it can assume any real real number and it's probability density function (p.d.f.) is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, -\infty < x < \infty; -\infty < \mu < \infty, \sigma > 0$$

where the two independent constants μ and σ in the distribution are known as the parameters of the distribution.

NOTE:

1. Normal distribution is a continuous distribution as X can assume any value within the interval$(-\infty, \infty)$.
2. Any random variable X which follows normal distribution is known as normal variate and is denoted by $X \sim N(\mu, \sigma)$, where μ and σ are the parameters of the distribution. **μ is the mean** and **σ is the standard deviation (s.d.)** of the normal distribution.
3. **Normal Curve:** The graph corresponding to the p.d.f. of a normal distribution with parameters μ and σ is known as normal curve. The normal curve is a bell-shaped curve which is symmetric about the ordinate $X = \mu$. The top of the bell is directly above $X = \mu$. For large values of σ, the curve tends to flatten out and for small values of σ, it has a sharp peak.
4. **Distribution Function:** The cumulative distribution function (c.d.f.) or distribution function (d.f.) of a normal distribution with parameters μ and σ, is denoted by F, is defined as follows:
$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(t)dt = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt$$

STANDARD NORMAL DISTRIBUTION

Definition: A normal distribution with parameters 0 and 1 is called **standard normal distribution**.

NOTE:

1. If X be a continuous random variable which follows standard normal distribution, then X is known as standard normal variate and is denoted by $X \sim N(0,1)$. Thus, the mean and standard deviation of a standard normal distribution are respectively 0 and 1.
2. Generally, the p.d.f. of any continuous probability distribution is denoted by f, but the p.d.f. of a standard normal distribution is always denoted by ϕ and is defined as follows (because here $\mu = 0$ and $\sigma = 1$)

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}, -\infty < x < \infty$$

3. **Standard Normal Curve:** The graph corresponding to the p.d.f. of a standard normal distribution known as standard normal curve. The standard normal curve is a bell-shaped curve which is symmetric about the ordinate $X = 0$ (because here $\mu = 0$). The top of the bell is directly above $X = 0$.
4. $Total\ area\ under\ the\ standard\ normal\ curve = \int_{-\infty}^{\infty} \phi(t)dt = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{1}{2}t^2} dt = 1\ square\ unit.$

   Since, standard normal curve is symmetric about the ordinate $X = 0$,
$$\int_{-\infty}^{0} \phi(t)dt = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{0} e^{-\frac{1}{2}t^2} dt = 0.5\ square\ unit$$
$$\int_{0}^{\infty} \phi(t)dt = \frac{1}{\sqrt{2\pi}}\int_{0}^{\infty} e^{-\frac{1}{2}t^2} dt = 0.5\ square\ unit$$

5. Generally, the c.d.f. of any continuous probability distribution is denoted by F, but the c.d.f. of a standard normal distribution is always denoted by Φ and is defined as follows (because here $\mu = 0$ and $\sigma = 1$)
$$\Phi(x) = P(X \leq x) = \int_{-\infty}^{x} \phi(t)dt = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{1}{2}t^2} dt\,, -\infty < x < \infty$$
   Thus, for $a > 0$,
$$\Phi(a) = P(X \leq a)$$
$$= \int_{-\infty}^{a} \phi(x)dx$$
$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{a} e^{-\frac{1}{2}x^2} dx$$
$$= Area\ under\ the\ standard\ normal\ curve\ on\ the\ left\ of\ the\ ordinate\ x = a$$

$$\Phi(-a) = P(X \leq -a)$$
$$= \int_{-\infty}^{-a} \phi(x)dx$$
$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{-a} e^{-\frac{1}{2}x^2} dx$$
$$= Area\ under\ the\ standard\ normal\ curve\ on\ the\ left\ of\ the\ ordinate\ x = -a$$
$$= Area\ under\ the\ standard\ normal\ curve\ on\ the\ right\ of\ the\ ordinate\ x = a$$
$$= 1 - Area\ under\ the\ standard\ normal\ curve\ on\ the\ left\ of\ the\ ordinate\ x = a$$
$$= 1 - \Phi(a)$$

Transformation of a Normal Distribution into Standard Normal Distribution

If $X \sim N(\mu, \sigma)$ and $Z = \frac{X-\mu}{\sigma}$, then $Z \sim N(0,1)$. Thus, the p.d.f. corresponding to Z is as follows

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}, -\infty < z < \infty$$

and the c.d.f. corresponding to Z is as follows

$$\Phi(z) = P(Z \le z) = \int_{-\infty}^{z} \phi(t)dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{1}{2}t^2} dt, -\infty < z < \infty$$

Technique to solve problems related to normal distribution

Problem related to normal distribution is solved by transforming the problem into a problem of standard normal distribution as follows:

Suppose $\boldsymbol{X \sim N(\mu, \sigma)}$ and we have to find the following probabilities

I $\boldsymbol{P(X \le a)}$

II $\boldsymbol{P(X \ge a)}$

III $\boldsymbol{P(a \le X \le b)}$

Let, $Z = \frac{X-\mu}{\sigma}$. Then $Z \sim N(0,1)$.

Therefore, the p.d.f. corresponding to Z is as follows

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}, -\infty < z < \infty$$

I $P(X \le a) = P\left(\frac{X-\mu}{\sigma} \le \frac{a-\mu}{\sigma}\right)$

$$= P\left(Z \le \frac{a-\mu}{\sigma}\right)$$

$$= \int_{-\infty}^{\frac{a-\mu}{\sigma}} \phi(z)dz$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{a-\mu}{\sigma}} e^{-\frac{1}{2}t^2} dt$$

$$= \Phi\left(\frac{a-\mu}{\sigma}\right)$$

$$= Area\ under\ the\ standard\ normal\ curve\ on\ the\ left\ of\ the\ ordinate\ z = \frac{a-\mu}{\sigma}$$

II $P(X \ge a) = P\left(\frac{X-\mu}{\sigma} \ge \frac{a-\mu}{\sigma}\right)$

$$= P\left(Z \ge \frac{a-\mu}{\sigma}\right)$$

$$= 1 - P\left(Z \le \frac{a-\mu}{\sigma}\right)$$

$$= 1 - \int_{-\infty}^{\frac{a-\mu}{\sigma}} \phi(z)dz$$

$$= 1 - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{a-\mu}{\sigma}} e^{-\frac{1}{2}t^2} dt$$

$$= 1 - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

$$= 1 - Area\ under\ the\ standard\ normal\ curve\ on\ the\ left\ of\ the\ ordinate\ z = \frac{a-\mu}{\sigma}$$

III $P(a \leq X \leq b) = P\left(\frac{a-\mu}{\sigma} \leq \frac{X-\mu}{\sigma} \leq \frac{b-\mu}{\sigma}\right)$

$$= P\left(\frac{a-\mu}{\sigma} \leq Z \leq \frac{b-\mu}{\sigma}\right)$$

$$= P\left(Z \leq \frac{b-\mu}{\sigma}\right) - P\left(Z \leq \frac{a-\mu}{\sigma}\right)$$

$$= \int_{-\infty}^{\frac{b-\mu}{\sigma}} \phi(z)dz - \int_{-\infty}^{\frac{a-\mu}{\sigma}} \phi(z)dz$$

$$= \frac{1}{\sqrt{2\pi}} \int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}} e^{-\frac{1}{2}t^2} dt$$

$$= \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

$$= Area\ under\ the\ standard\ normal\ curve\ between\ the\ ordinates\ z = \frac{a-\mu}{\sigma}\ and\ z = \frac{b-\mu}{\sigma}$$

IV $P(X \geq a) = P\left(\frac{X-\mu}{\sigma} \geq \frac{a-\mu}{\sigma}\right)$

$$= P\left(Z \geq \frac{a-\mu}{\sigma}\right)$$

$$= 1 - P\left(Z \leq \frac{a-\mu}{\sigma}\right)$$

$$= 1 - \int_{-\infty}^{\frac{a-\mu}{\sigma}} \phi(z)dz$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{a-\mu}{\sigma}} e^{-\frac{1}{2}t^2} dt$$

$$= \Phi\left(\frac{a-\mu}{\sigma}\right)$$

$$= Area\ under\ the\ standard\ normal\ curve\ on\ the\ left\ of\ the\ ordinate\ z = \frac{a-\mu}{\sigma}$$

Worked out Problem:

Problem 10.1. Suppose the diameter of a certain car component follows the normal distribution with $X \sim N(10, 3)$. Find, if we randomly select one of these components, the probability that its diameter will be larger than 13.4 mm.

Solution:

Here, $X \sim N(10, 3)$.

$$\therefore Z = \frac{X - 10}{3} \sim N(0,1)$$

$$\therefore P(X > 13.4) = P(X - 10 > 13.4 - 10)$$

$$= P\left(\frac{X - 10}{3} > \frac{13.4 - 10}{3}\right)$$

$$= P(Z > 1.13)$$

$$= 1 - P(Z \leq 1.13)$$

$$= 1 - 0.8708$$
$$= 0.1292.$$

MODULE II: MATHEMATICAL LOGIC

(NUMBER OF LECTURES: 6)

Lecture 11: Introduction to Propositional Calculus, Propositions, Logical Connectives

Mathematical logic is the discipline that mathematicians invented in the late nineteenth and early twentieth century so they could stop talking nonsense. It’s the most powerful tool we have for reasoning about things that we can’t really comprehend, which makes it a perfect tool for Computer Science.

11.1 The basic picture

Reality	Model	Theory
herds of sheep		
piles of rocks →	$N = \{0, 1, 2, \ldots\}$ →	$\forall x : \exists y : y = x + 1$
tally marks		

We want to model something we see in reality with something we can fit in our heads. Ideally we drop most of the features of the real thing that we don’t care about and keep the parts that we do care about. But there is a second problem: if our model is very big (and the natural numbers are very *very* big), how do we know what we can say about them?

11.2 Axioms, models and inference rules

One approach is to come up with a list of axioms that are true statements about the model and a list of inference rules that let us derive new true statements from the axioms. The axioms and inference rules together generate a theory that consists of all statements that can be constructed from the axioms by applying the inference rules. The rules of the game are that we can’t claim that some statement is true unless it’s a theorem: something we can derive as part of the theory.

Simple example: All fish are green (axiom). George Washington is a fish (axiom). From “all X are Y” and “Z is X”, we can derive “Z is Y” (inference rule). Thus George Washington is green (theorem). Since we can’t do anything else with our two axioms and one inference rule, these three statements together form our entire theory about George Washington, fish, greenness, etc.

Theories are attempts to describe models. A model is typically a collection of objects and relations between them. For a given theory, there may be many models that are consistent with it: for example, a model that includes both green fishy George Washington and MC 900-foot Abraham Lincoln is consistent with the theory above, because the theory doesn’t say anything about Abraham Lincoln.

11.3 Consistency

A theory is consistent if it can't prove both P and not$-P$ for any P. Consistency is incredibly important, since all the logics people actually use can prove anything starting from P and not$-P$.

11.4 The language of logic

The basis of mathematical logic is propositional logic, which was essentially invented by Aristotle. Here the model is a collection of statements that are either true or false. There is no ability to refer to actual things; though we might include the statement "George Washington is a fish", from the point of view of propositional logic that is an indivisible atomic chunk of truth or falsehood that says nothing in particular about George Washington or fish. If we treat it as an axiom we can prove the truth of more complicated statements like "George Washington is a fish or $2 + 2 = 5$" (true since the first part is true), but we can't really deduce much else. Still, this is a starting point.

If we want to talk about things and their properties, we must upgrade to predicate logic. Predicate logic adds both constants (stand-ins for objects in the model like "George Washington") and predicates (stand-ins for properties like "is a fish"). It also lets use quantify over variables and make universal statements like "For all x, if x is a fish then x is green." As a bonus, we usually get functions ("f (x) = the number of books George Washington owns about x") and equality ("George Washington = 12" implies "George Washington $+ 5 = 17$"). This is enough machinery to define and do pretty much all of modern mathematics.

We will discuss both of these logics in more detail below.

11.5 Propositional logic

Propositional logic is the simplest form of logic. Here the only statements that are considered are propositions, which contain no variables. Because propositions contain no variables, they are either always true or always false.

Examples of propositions:

- $2 + 2 = 4$. (Always true).

- $2 + 2 = 5$. (Always false).

Examples of non-propositions:

- $\mathrm{x} + 2 = 4$. (May be true, may not be true; it depends on the value of x.
- $x \cdot 0 = 0$. (Always true, but it's still not a proposition because of the variable.)
- $x \cdot 0 = 1$. (Always false, but not a proposition of the variable.)

As the last two examples show, it is not enough for a statement to be always true or always false—whether a statement is a proposition or not is a structural property. But if a statement doesn't contain any variables (or other undefined terms), it is a proposition, and as a side-effect of being a proposition it's always true or always false.

Lecture 12: Conjunction, Disjunction, Negation and their truth table

12.1 Operations on propositions

Propositions by themselves are pretty boring. So boring, in fact, that logicians quickly stop talking about specific propositions and instead haul out placeholder names like $\mathrm{p}, \mathrm{q},$ or r. But we can build slightly more interesting propositions by combining propositions together using various logical connectives, such as:

Negation The **negation** of p is written as $\neg\mathrm{p}$, or sometimes $\sim \mathrm{p}, -\mathrm{p}$. It has the property that it is false when p is true, and true when p is false.

Or The **or** of two propositions p and q is written as $p \lor q$, and is true as long as at least one, or possibly both, of p and q is true. This is not always the same as what "or" means in English; in English, "or" often is used for exclusive or which is not true if both p and q are true. For example, if someone says "You will give me all your money or I will stab you with this table knife", you would be justifiably upset if you turn over all your money and still get stabbed. But a logician would not be at all surprised, because the standard "or" in propositional logic is an **inclusive or** that allows for both outcomes.

Exclusive or If you want to exclude the possibility that both p and q are true, you can use exclusive or instead. This is written as $\mathrm{p} \oplus \mathrm{q}$, and is true precisely when exactly one of p or q is true.

And The **and** of p and q is written as $p \land q$, and is true only when both p and q are true. This is pretty much the same as in English, where "I like to eat ice cream and I own a private Caribbean island" is not a true statement when made by most people even though most people like to eat ice cream. The only complication in translating English expressions into logical and s is that logicians can't tell the difference between "and" and "but": the statement "$2 + 2 = 4$ but $3 + 3 = 6$" becomes simply "$(2 + 2 = 4) \land (3 + 3 = 6)$."

Implication This is the most important connective for proofs. An implication represents an "if. . . then" claim. If p implies q, then we write $p \to q$ or $p \Rightarrow q$, depending on our typographic convention and the availability of arrow symbols in our favourite font. In English, $p \to q$ is usually rendered as "If p, then q," as in "If you step on your own head, it will hurt." The meaning of $p \to q$ is that q is true whenever p is true, and the proposition $p \to q$ is true provided (a) p is false (in which case all bets are off), or (b) q is true.

In fact, the only way for $p \to q$ to be false is for p to be true but q to be false. Because of this, $p \to q$ can be rewritten as $\neg p \lor q$. So, for example, the statements "If $2 + 2 = 5$, then I'm the Pope", "If I'm the Pope, then $2 + 2 = 4$", and "If $2 + 2 = 4$, then $3 + 3 = 6$", are all true, provided the if/then is interpreted as implication. Normal English usage does not always match this pattern; instead, if/then in normal speech is often interpreted as the much stronger biconditional (see below).

Biconditional Suppose that $p \to q$ and $q \to p$, so that either both p and q are true or both p and q are false. In this case, we write $p \leftrightarrow q$ or $p \Leftrightarrow q$, and say that p holds if and only if q holds. The truth of $p \leftrightarrow q$ is still just a function of the truth or falsehood of p and q; though there doesn't need to be any connection between the two sides of the statement, "$2 + 2 = 5$ **if and only if** I am the Pope" is a true statement (provided it is not uttered by the Pope). The only way for $p \leftrightarrow q$ to be false is for one side to be true and one side to be false.

The result of applying any of these operations is called a **compound proposition.**

$NOT\ p$	$\neg p$	$p, \sim p$
$p\ AND\ q$	$p \wedge q$	
$p\ XOR\ q$	$p \oplus q$	
$p\ OR\ q$	$p \vee q$	
$p\ implies\ q$	$p \rightarrow q$	$p \Rightarrow q, p \supset q$
$p\ if\ and\ only\ if\ q$	$p \leftrightarrow q$	$p \Leftrightarrow q$

12.2 Truth tables

To define logical operations formally, we give a **truth table**. This gives, for any combination of truth values (true or false, which as computer scientists we often write as 1 or 0) of the inputs, the truth value of the output. In this usage, truth tables are to logic what addition and multiplication tables are to arithmetic.

Here is a truth table for negation:

p	$\neg p$
0	1
1	0

And here is a truth table for the rest of the logical operators:

p	q	$p \vee q$	$p \oplus q$	$p \wedge q$	$p \rightarrow q$	$p \leftrightarrow q$
0	0	0	0	0	1	1
0	1	1	1	0	1	0
1	0	1	1	0	0	0
1	1	1	0	1	1	1

12.3 Tautologies and logical equivalence

A compound proposition that is true no matter what the truth-values of the propositions it contains is called a tautology. For example, $p \rightarrow p, p \vee \neg p,$ and $\neg(p \wedge \neg p)$ are all tautologies, as can be verified by constructing truth tables. If a compound proposition is always false, it's a contradiction. The negation of a tautology is a contradiction and vice versa.

The most useful class of tautologies are logical equivalences. This is a tautology of the form $X \leftrightarrow Y$, where X and Y are compound propositions.

In this case, X and Y are said to be logically equivalent and we can substitute one for the other in more complex propositions. We write $X \equiv Y$ if X and Y are logically equivalent.

The nice thing about logical equivalence is that is does the same thing for Boolean formulas that equality does for algebraic formulas: if we know (for example), that $p \vee \neg p$ is equivalent to 1, and $q \vee 1$ is equivalent to 1, we can grind $q \vee p \vee \neg p \equiv q \vee 1 \equiv 1$ without having to do anything particularly clever.

To prove a logical equivalence, one either constructs a truth table to show that $X \leftrightarrow Y$ is a tautology, or transforms X to Y using previously-known logical equivalences.

Some examples:

- $p \wedge \neg p \equiv 0$: Construct a truth table

p	$\neg p$	$p \wedge \neg p$	0
0	1	0	0
1	0	0	0

and observe that the last two columns are always equal.

- $p \vee p \equiv p$: Use the truth table

p	p ∨ p
0	0
1	1

- $p \to q \equiv \neg p \vee q$: Again construct a truth table

p	q	$p \to q$	$\neg p \vee q$
0	0	1	1
0	1	1	1
1	0	0	0
1	1	1	1

- $\neg(p \vee q) \equiv \neg p \wedge \neg q$: (one of De Morgan's laws; the other is $\neg(p \wedge q) \equiv \neg p \vee \neg q$).

p	q	$p \vee q$	$\neg(p \vee q)$	$\neg p$	$\neg q$	$\neg p \wedge \neg q$
0	0	0	**1**	1	1	**1**
0	1	1	**0**	1	0	**0**
1	0	1	**0**	0	1	**0**
1	1	1	**0**	0	0	**0**

- $p \vee (q \wedge r) \equiv (p \vee q) \wedge (p \vee r)$ (one of the distributive laws; the other is $p \wedge (q \vee r) \equiv (p \wedge q) \vee (p \wedge r)$).

p	q	r	$q \wedge r$	$p \vee (q \wedge r)$	$p \vee q$	$p \vee r$	$(p \vee q) \wedge (p \vee r)$
0	0	0	0	**0**	0	0	**0**
0	0	1	0	**0**	0	1	**0**
0	1	0	0	**0**	1	0	**0**
0	1	1	1	**1**	1	1	**1**
1	0	0	0	**1**	1	1	**1**
1	0	1	0	**1**	1	1	**1**
1	1	0	0	**1**	1	1	**1**
1	1	1	1	**1**	1	1	**1**

- $(p \to r) \vee (q \to r) \equiv (p \wedge q) \to r$. Now things are getting messy, so building a full truth table may take awhile. But we have take a shortcut by using logical equivalences that we've already proved (plus associativity of $\vee$):

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

$\equiv (\neg p \vee r) \vee (\neg q \vee r)$ [Using $p \to q \equiv \neg p \vee q$ twice]

$\equiv \neg p \vee \neg q \vee r \vee r$ [Associativity and commutativity of $\vee$]

$\equiv \neg p \vee \neg q \vee r$ [$p \equiv p \vee p$]

$\equiv \neg(p \wedge q) \vee r$ [De Morgan's law]

$\equiv (p \wedge q) \to r.$ [$p \to q \equiv \neg p \vee q$]

Lecture 13: Converse, Contrapositive, Inverse

13.1 Inverses, converses, and contrapositives

The **contrapositive** of $p \to q$ is $\neg q \to \neg p$; it is logically equivalent to the original implication. For example, the contrapositive of "If I am Barack Obama then I am a Democrat" is "If I am not a Democrat then I am not Barack Obama". A proof by contraposition demonstrates that p implies q by assuming $\neg q$ and then proving $\neg p$; it is similar but not identical to an indirect proof, which assumes $\neg p$ and derives a contradiction.

The **inverse** of $p \to q$ is $\neg p \to \neg q$. So the inverse of "If you take CPSC 202, you will surely die" is "If you do not take CPSC 202, you will not surely die." There is often no connection between the truth of an implication and the truth of its inverse: "If I am Barack Obama then I am a Democrat" does not have the same truth-value as "If I am not Barack Obama then I am not a Democrat", at least according to current polling numbers.

The **converse** of $p \to q$ is $q \to p$. E.g. the converse of "If I am Barack Obama then I am a Democrat" is "If I am a Democrat then I am Barack Obama." The converse of a statement is always logically equivalent to the inverse. Often in proving a biconditional (e.g., "I am Barack Obama if and only if I am a Democrat"), one proceeds by proving first the implication in one direction and then either the inverse or the converse (which are logically equivalent).

$\neg\neg p \equiv p$	Double negation
$\neg(p \wedge q) \equiv \neg p \vee \neg q$	De Morgan's law
$\neg(p \vee q) \equiv \neg p \wedge \neg q$	De Morgan's law
$p \wedge q \equiv q \wedge p$	Commutativity of AND
$p \vee q \equiv q \vee p$	Commutativity of OR
$p \wedge (q \wedge r) \equiv p \wedge (q \wedge r)$	Associativity of AND
$p \vee (q \vee r) \equiv p \vee (q \vee r)$	Associativity of OR
$p \wedge (q \vee r) \equiv (p \wedge q) \vee (p \wedge r)$	AND distributes over OR
$p \vee (q \wedge r) \equiv (p \vee q) \wedge (p \vee r)$	OR distributes over AND
$p \to q \equiv \neg p \vee q$	Equivalence of implication and OR
$p \to q \equiv \neg q \to \neg p$	Contraposition
$p \leftrightarrow q \equiv (p \to q) \wedge (q \to p)$	Expansion of if and only if
$p \leftrightarrow q \equiv \neg p \leftrightarrow \neg q$	Inverse of if and only f

$p \leftrightarrow q \equiv q \leftrightarrow p$	Commutativity of if and only if

Lecture 14: Logical Equivalence, Normal forms-CNF, DNF

14.1 Equivalences involving true and false

Any tautology is equivalent to true; any contradiction is equivalent to false. Two important cases of this are the **law of the excluded middle** $P \vee \neg P \equiv 1$ and its dual, the **law of non-contradiction** $P \wedge \neg P \equiv 0$.

The law of the excluded middle is what allows us to do **case analysis**, where we prove that some proposition Q holds by showing first that P implies Q and then that ¬P also implies Q.

One strategy for simplifying logical expressions is to try to apply known equivalences to generate sub-expressions that reduce to true or false via

$$
\begin{array}{ll}
P \wedge 0 \equiv 0 & P \vee 0 \equiv P \\
P \wedge 1 \equiv P & P \vee 1 \equiv 1 \\
P \leftrightarrow 0 \equiv \neg P & P \oplus 0 \equiv P \\
P \leftrightarrow 1 \equiv P & P \oplus 1 \equiv \neg P \\
P \rightarrow 0 \equiv \neg P & 0 \rightarrow P \equiv 1 \\
P \rightarrow 1 \equiv 1 & 1 \rightarrow P \equiv P
\end{array}
$$

Example Let's show that $(P \wedge (P \rightarrow Q)) \rightarrow Q$ is a tautology. (This justifies the inference rule modus ponens, defined below.) Working from the inside out, we can compute

$$
\begin{array}{ll}
(P \wedge (P \rightarrow Q)) \rightarrow Q & \\
\equiv (P \wedge (\neg P \vee Q)) \rightarrow Q & \text{expand } \rightarrow \\
\equiv ((P \wedge \neg P) \vee (P \wedge Q)) \rightarrow Q & \text{distribute } \vee \text{ over } \wedge \\
\equiv (0 \vee (P \wedge Q)) \rightarrow Q & \text{non} - \text{contradiction} \\
\equiv (P \wedge Q) \rightarrow Q & \text{absorption} \\
\equiv \neg(P \wedge Q) \vee Q & \text{expand } \rightarrow \\
\equiv (\neg P \vee \neg Q) \vee Q & \text{De Morgan's law} \\
\equiv \neg P \vee (\neg Q \vee Q) & \text{associativity} \\
\equiv \neg P \vee 1 & \text{excluded middle} \\
\equiv 1 & \text{absorption}
\end{array}
$$

In this derivation, we've labeled each step with the equivalence we used. Most of the time we would not be this verbose.

14.2 Normal forms

A compound proposition is in **conjuctive normal form** (CNF for short) if it is obtained by AND-ing together ORs of one or more variables or their negations (an OR of one variable is just the variable itself). So for example

P, $(P \lor Q) \land R$, $(P \lor Q) \land (Q \lor R) \land (\neg P)$, and $(P \lor Q) \land (P \lor \neg R) \land (\neg P \lor Q \lor S \lor T \lor \neg U)$ are in CNF, but $(P \lor Q) \land (P \lor \neg R) \land (\neg P \land Q)$, $(P \lor Q) \land (P \to R) \land (\neg P \lor Q)$, and $(P \lor (Q \land R)) \land (P \lor \neg R) \land (\neg P \lor Q)$ are not. Using the equivalence $P \to Q \equiv \neg P \lor Q$, De Morgan's laws, and the distributive law, it is possible to rewrite any compound proposition in CNF.

CNF formulas are particularly useful because they support resolution Using the tautology $(P \lor Q) \land (\neg P \lor R) \to Q \lor R$, we can construct proofs from CNF formulas by looking for occurrences of some simple proposition and its negation and resolving them, which generates a new clause we can add to the list. For example, we can compute

$$\vdash (P \lor Q) \land (P \lor \neg R) \land (\neg P \lor Q) \land (\neg Q \lor R)$$

$$\vdash (P \lor Q) \land (P \lor \neg R) \land (\neg P \lor Q) \land (\neg Q \lor R) \land Q$$

$$\vdash (P \lor Q) \land (P \lor \neg R) \land (\neg P \lor Q) \land (\neg Q \lor R) \land Q \land R$$

$$\vdash (P \lor Q) \land (P \lor \neg R) \land (\neg P \lor Q) \land (\neg Q \lor R) \land Q \land R \land P$$

$$\vdash P.$$

Similarly, a compound proposition is in **disjunctive normal form** (DNF) if it consists of an OR of ANDs, e.g. $(P \land Q) \lor (P \land \neg R) \lor (\neg P \land Q)$. Just as any compound proposition can be transformed into CNF, it can similarly be transformed into DNF.

Note that conjunctive and disjunctive normal forms are not unique; for example, $P \land Q$ and $(P \lor \neg Q) \land (P \lor Q) \land (\neg P \lor Q)$ are both in conjunctive normal form and are logically equivalent to each other. So while CNF can be handy as a way of reducing the hairiness of a formula (by eliminating nested parentheses or negation of non-variables, for example), it doesn't necessarily let us see immediately if two formulas are really the same.

Lecture 15: Predicate logic

15.1 Predicate logic

Using only propositional logic, we can express a simple version of a famous argument:

- Socrates is a man.

- If Socrates is a man, then Socrates is mortal.

- Therefore, Socrates is mortal.

This is an application of the inference rule called modus ponens, which says that from p and $p \rightarrow q$ you can deduce q. The first two statements are axioms (meaning we are given them as true without proof), and the last is the conclusion of the argument.

What if we encounter Socrates's infinitely more logical cousin Spocrates? We'd like to argue

- Spocrates is a man.

- If Spocrates is a man, then Spocrates is mortal.

- Therefore, Spocrates is mortal.

Unfortunately, the second step depends on knowing that humanity implies mortality for everybody, not just Socrates. If we are unlucky in our choice of axioms, we may not know this. What we would like is a general way to say that humanity implies mortality for everybody, but with just propositional logic, we can't write this fact down.

15.2 Variables and predicates

The solution is to extend our language to allow formulas that involve variables. So we might let $x, y, z,$ etc. stand for any element of our universe of discourse or domain—essentially whatever things we happen to be talking about at the moment. We can now write statements like:

- "x is human."

- "x is the parent of y."

- "$x + 2 = x^2$."

These are not propositions because they have variables in them. Instead, they are predicates; statements whose truth-value depends on what concrete object takes the place of the variable. Predicates are often abbreviated by single capital letters followed by a list of arguments, the variables that appear in the predicate, e.g.:

- $H(x) =$ "x is human."

- $P(x, y) =$ "x is the parent of y."

- $Q(x) =$ "$x + 2 = x^2$."

We can also fill in specific values for the variables, e.g. H (Spocrates) = “Spocrates is human.” If we fill in specific values for all the variables, we have a proposition again, and can talk about that proposition being true

(e.g. $H\ (2)$ and $H\ (-1)$ are true) or false ($H\ (0)$ is false).

In **first-order logic**, which is what we will be using in this course, variables always refer to things and never to predicates: any predicate symbol is effectively a constant. There are higher-order logics that allow variables to refer to predicates, but most mathematics accomplishes the same thing by representing predicates with sets .

Lecture 16: Logical Quantifications of propositions

16.1 Quantifiers

What we really want is to be able to say when H or P or Q is true for many different values of their arguments. This means we have to be able to talk about the truth or falsehood of statements that include variables. To do this, we **bind** the variables using **quantifiers**, which state whether the claim we are making applies to all values of the variable (**universal quantification**), or whether it may only apply to some (**existential quantification**).

16.2 Universal quantifier

The universal quantifier ∀ (pronounced "for all") says that a statement must be true for all values of a variable within some universe of allowed values (which is often implicit). For example, "all humans are mortal" could be written $\forall x$: Human(x) → Mortal(x) and "if x is positive then $x + 1$ is positive" could be written $\forall x : x > 0 \rightarrow x + 1 > 0.$

If you want to make the universe explicit, use set membership notation, e.g. $\forall x \in Z : x > 0 \rightarrow x + 1 > 0$. This is logically equivalent to writing $\forall x : x \in Z \rightarrow (x > 0 \rightarrow x + 1 > 0)$ or to writing $\forall x : (x \in Z \wedge x > 0) \rightarrow x + 1 > 0$, but the short form makes it more clear that the intent of $x \in Z$ is to restrict the range of x^4.

The statement $\forall x : P\,(x)$ is equivalent to a very large AND; for example, $\forall x \in N : P\,(x)$ could be rewritten (if you had an infinite amount of paper) as $P\,(0) \wedge P\,(1) \wedge P\,(2) \wedge P\,(3) \wedge ....$ Normal first-order logic doesn't allow infinite expressions like this, but it may help in visualizing what $\forall x : P\,(x)$ actually means. Another way of thinking about it is to imagine that x is supplied by some adversary and you are responsible for showing that $P\,(x)$ is true; in this sense, the universal quantifier chooses the worst case value of x.

16.3 Existential quantifier

The **existential quantifier** ∃ (pronounced "there exists") says that a statement must be true for at least one value of the variable. So "some human is mortal" becomes $\exists x$: Human(x) ∧ Mortal(x). Note that we use AND rather than implication here; the statement $\exists x$: Human(x) → Mortal(x) makes the much weaker claim that "there is some thing x, such that if x is human, then x is mortal," which is true in any universe that contains an immortal purple penguin—since it isn't human, Human(penguin) → Mortal(penguin) is true.

As with $\forall, \exists$ can be limited to an explicit universe with set membership notation, e.g., $\exists x \in Z : x = x^2$. This is equivalent to writing $\exists x : x \in Z \wedge x = x2$.

The formula $\exists x : P\,(x)$ is equivalent to a very large OR, so that $\exists x \in N : P\,(x)$ could be rewritten as $P\,(0) \vee P\,(1) \vee P\,(2) \vee P\,(3) \vee ....$ Again, you can't generally write an expression like this if there are infinitely many terms, but it gets the idea across.

16.4 Negation and quantifiers

The following equivalences hold:

$$\neg\forall x : P(x) \equiv \exists x : \neg P(x).$$

$$\neg\exists x : P(x) \equiv \forall x : \neg P(x).$$

These are essentially the quantifier version of De Morgan’s laws: the first says that if you want to show that not all humans are mortal, it’s equivalent to finding some human that is not mortal. The second says that to show that no human is mortal, you have to show that all humans are not mortal.

16.6 Examples

Here we give some more examples of translating English into statements in predicate logic.

All crows are black. $\forall x : Crow(x) \rightarrow Black(x)$

The formula is logically equivalent to either of

$$\neg\exists x Crow(x) \wedge \neg Black(x)$$

$$or$$

$$\forall x : \neg Black(x) \rightarrow \neg Crow(x).$$

The latter is the core of a classic “paradox of induction” in philosophy: if seeing a black crow makes me think it’s more likely that all crows are black, shouldn’t seeing a logically equivalent non-black non-crow (e.g., a banana yellow AMC Gremlin) also make me think all non-black objects are non- crows, i.e., that all crows are black? The paradox suggests that logical equivalence works best for true/false and not so well for probabilities.

Some cows are brown. $\exists x$: Cow(=) $\wedge$ Brown(x)

No cows are blue. $\neg\exists x$: Cow(x) $\wedge$ Blue(x)

Some other equivalent versions:

$$\forall x : \neg(\text{Cow}(x) \wedge \text{Blue}(x)$$

$$\forall x : (\neg\text{Cow}(x) \vee \neg\text{Blue}(x))$$

$$\forall x : \text{Cow}(x) \rightarrow \neg\text{Blue}(x)$$

$$\forall x : \text{Blue}(x) \rightarrow \neg\text{Cow}(x).$$

All that glitters is not gold. $\neg\forall x$: Glitters(x) $\rightarrow$ Gold(x)

Or $\exists x$: Glitters(x) $\wedge \neg$Gold(x). Note that the English syntax is a bit ambiguous: a literal translation might look like $\forall x$: Glitters(x) $\rightarrow \neg$Gold(x), which is not logically equivalent. This is an example of how predicate logic is often more precise than natural language.

No shirt, no service. $\forall x : \neg$Shirt(x) $\rightarrow \neg$Served(x)

Every event has a cause. $\forall x \exists y : \text{Causes}(y, x)$

And a more complicated statement: Every even number greater than 2 can be expressed as the sum of two primes.

$\forall x : (\text{Even}(x) \wedge x > 2) \rightarrow (\exists p \exists q : \text{Prime}(p) \wedge \text{Prime}(q) \wedge (x = p + q))$

The last one is Goldbach's conjecture. The truth value of this statement is currently unknown.

16.7 Functions

A function symbol looks like a predicate but instead of computing a truth value it returns an object. So for example the successor function S in the Peano axioms for the natural numbers returns $x + 1$ when applied as $S(x)$. Sometimes when there is only a single argument we omit the parentheses, e.g., $Sx = S(x), SSSx = S(S(S(x)))$.

16.8 Equality

Often we include a special equality predicate $=$, written $x = y$. The interpretation of $x = y$ is that x and y are the same element of the domain. It satisfies the reflexivity axiom $\forall x : x = x$ and the substitution axiom schema:

$$\forall x \forall y : (x = y \rightarrow (Px \leftrightarrow Py))$$

where P is any predicate. This immediately gives a substitution rule that says $x = y, P(x) \vdash P(y)$. It's likely that almost every proof you ever wrote down in high school algebra consisted only of many applications of the substitution rule.

Example: We'll prove $\forall x \forall y : (x = y \rightarrow y = x)$ from the above axioms (this property is known as symmetry). Apply substitution to the predicate $P\,x \equiv y = x$ to get $\forall x \forall y : (x = y \rightarrow (y = x \leftrightarrow x = x))$. Use reflexivity to rewrite this as $\forall x \forall y : (x = y \rightarrow (y = x \leftrightarrow T\,))$ or $\forall x \forall y : (x = y \rightarrow y = x)$ as claimed.

Exercise: Prove $\forall x \forall y \forall z : (x = y \wedge y = z \rightarrow x = z)$. (This property is known as transitivity.)

16.9 Uniqueness

An occasionally useful abbreviation is $\exists ! x P(x)$, which stands for "there exists a unique x such that $P(x)$." This is short for

$$(\exists x P(x)) \wedge (\forall x \forall y : P(x) \wedge P(y) \rightarrow x = y).$$

An example is $\exists ! x : x + 1 = 12$. To prove this we'd have to show not only that there is some x for which $x + 1 = 12$ (11 comes to mind), but that if we have any two values

x and y such that $x + 1 = 12$ and $y + 1 = 12$, then $x = y$ (this is not hard to do). So the exclamation point encodes quite a bit of extra work, which is why we usually hope that $\exists x : x + 1 = 12$ is good enough.

16.10 Proofs

A proof is a way to derive statements from other statements. It starts with **axioms** (statements that are assumed in the current context always to be true), **theorems** or **lemmas** (statements that were proved already; the difference between a theorem and a lemma is whether it is intended as a final result or an intermediate tool), and **premises P** (assumptions we are making for the purpose of seeing what consequences they have), and uses **inference rules** to derive Q. The axioms, theorems, and premises are in a sense the starting position of a game whose rules are given by the inference rules. The goal of the game is to apply the inference rules until Q pops out. We refer to anything that isn't proved in the proof itself (i.e., an axiom, theorem, lemma, or premise) as a **hypothesis;** the result Q is the **conclusion.**

When a proof exists of Q from some premises $P_1, P_2, \ldots$, we say that Q is deducible or provable from $P_1, P_2, \ldots$, which is written as$P_1, P_2, \ldots, \vdash Q$.

If we can prove Q directly from our inference rules without making any assumptions, we may write ⊢Q

The **turnstile** symbol ⊢ has the specific meaning that we can derive the conclusion Q by applying inference rules to the premises. This is not quite the same thing as saying $P \rightarrow Q$. If our inference rules are particularly weak, it may be that $P \rightarrow Q$ is true but we can't prove Q starting with P. Conversely, if our inference rules are too strong (maybe they can prove anything, even things that aren't true) we might have $P \vdash Q$ but $P \rightarrow Q$ is false.

For propositions, most of the time we will use inference rules that are just right, meaning that $P \vdash Q$ implies $P \rightarrow Q$ **(soundness)** and $P \rightarrow Q$ implies $P \vdash Q$ **(completeness).** Here the distinction between ⊢ and → is then whether we want to talk about the existence of a proof (the first case) or about the logical relation between two statements (the second). Things get a little more complicated with statements involving predicates; in this case there are **incompleteness theorems** that say that sufficiently powerful sets of axioms have consequences that can't be proven unless the theory is inconsistent.

16.11 Inference Rules

Inference rules let us construct **valid** arguments, which have the useful property that if their premises are true, their conclusions are also true. The main source of inference rules is tautologies of the form $P_1, P_2, \ldots, \rightarrow Q$;given such a tautology, there is a corresponding inference rule that allows us to assert Q once we have $P_1, P_2, \ldots$ (either because each P_i is an axiom/theorem/premise or because we proved it already while doing the proof). The most important inference rule is modus ponens, based on the tautology $(p \wedge (p \rightarrow q)) \rightarrow q$; this lets us, for example, write the following famous argument:

1. If it doesn't fit, you must acquit. [Axiom]

2. It doesn't fit. [Premise]

3. You must acquit. [Modus ponens applied to $1 + 2$]

There are many named inference rules in classical propositional logic. We'll list some of them below. You don't need to remember the names of anything except modus ponens, and most of the rules are pretty much straightforward applications of modus ponens plus some convenient tautology that can be proved by truth tables or stock logical equivalences. (For example, the "addition" rule below is just the result of applying modus ponens to p and the tautology $p \to (p \vee q).)$

Inference rules are often written by putting the premises above a horizontal line and the conclusion below. In text, the horizontal line is often replaced by the symbol $\vdash$, which means exactly the same thing. Premises are listed on the left-hand side separated by commas, and the conclusion is placed on the right. We can then write

$p \vdash p \vee q.$	*Addition*
$p \wedge q \vdash p.$	*Simplification*
$p, q \vdash p \wedge q.$	*Conjunction*
$p, p \to q \vdash q.$	*Modus ponens*
$\neg q, p \to q \vdash \neg p.$	*Modus tollens*
$p \to q, q \to r \vdash` p \to r.$	*Hypothetical syllogism*
$p \vee q, \neg p \vdash q.$	*Disjunctive syllogism*
$p \vee q, \neg p \vee r \vdash q \vee r.$	*Resolution*

Of these rules, addition, simplification, and conjunction are mostly used to pack and unpack pieces of arguments. Modus ponens "the method of affirming" (and its reversed cousin modus tollens "the method of denying") let us apply implications. You don't need to remember modus tollens if you can remember the contraposition rule $(p \to q) \equiv (\neg q \to \neg p)$. Hypothetical syllogism just says that implication is transitive; it lets you paste together implications if the conclusion of one matches the premise of the other. Disjunctive syllogism is again a disguised version of modus ponens (via the logical equivalence $(p \vee q) \equiv (\neg p \to q)$); you don't need to remember it if you can remember this equivalence. Resolution is almost never used by humans but is very popular with computer theorem provers.

An argument is valid if the conclusion is true whenever the hypotheses are true. Any proof constructed using the inference rules is valid. It does not necessarily follow that the conclusion is true; it could be that one or more of the hypotheses is false:

1. If you give a mouse a cookie, he's going to ask for a glass of milk. [Axiom]

2. If he asks for a glass of milk, he will want a straw. [Axiom]

3. You gave a mouse a cookie. [Premise]

4. He asks for a glass of milk. [Modus ponens applied to 1 and 3.]

5. He will want a straw. [Modus ponens applied to 2 and 4.]

Will the mouse want a straw? No: Mice can't ask for glasses of milk, so Axiom 1 is false.

16.12 Proofs, implication, and natural deduction

Recall that $P \vdash Q$ means there is a proof of Q by applying inference rules to P, while $P \rightarrow Q$ says that Q holds whenever P does. These are not the same thing: provability ($\vdash$) is outside the theory (it's a statement about whether a proof exists or not) while implication ($\rightarrow$) is inside (it's a logical connective for making compound propositions). But most of the time they mean almost the same thing.

For example, suppose that $P \rightarrow Q$ is provable without any assumptions:

$$\vdash P \rightarrow Q.$$

Since we can always ignore extra premises, we get

$$P \vdash P \rightarrow Q$$

and thus

$$P \vdash P, P \to Q,$$

which gives

$$P \vdash Q$$

by applying modus ponens to the right-hand side.

So we can go from $\vdash P \to Q\ to\ P \vdash Q$.

This means that provability is in a sense weaker than implication: it holds (assuming modus ponens) whenever implication does. But we usually don't use this fact much, since $P \to Q$ is a much more useful statement than $P \vdash Q$. Can we go the other way?

Thanks to the James Aspnes, Yale University for his kind cooperation.

MODULE III: NUMBER THEORY & PARTIAL ORDER RELATION AND LATTICES

(NUMBER OF LECTURES: 8)

Lecture 17:

Introduction:

Number theory**,** branch of mathematics concerned with properties of the positive integers (1, 2, 3, …)also called natural numbers. However, the theory is not strictly confined to just the natural numbers or even to the set of all integers : 0, ±1, ±2,..... Sometimes called "higher arithmetic," it is among the oldest and most natural of mathematical pursuits.

Number theory has always fascinated amateurs as well as professional mathematicians. In contrast to other branches of mathematics, many of the problems and theorems of number theory can be understood by laypersons, although solutions to the problems and proofs of the theorems often require a sophisticated mathematical background.

Until the mid-20th century, number theory was considered the purest branch of mathematics, with no direct applications to the real world. The advent of digital computers and digital communications revealed that number theory could provide unexpected answers to real-world problems. At the same time, improvements in computer technology enabled number theorists to make remarkable advances in factoring large numbers, determining primes, testing conjectures, and solving numerical problems once considered out of reach.

Modern number theory is a broad subject that is classified into subheadings such as elementary number theory, algebraic number theory, analytic number theory, geometric number theory, and probabilistic number theory. These categories reflect the methods used to address problems concerning the integers.

Theorem [The well-ordering principle]: Every non-empty subset A of the set $\mathbb{N}$ of natural numbers has a least element.

Remark: By the "least element of a set A" it is meant that there exists an element $x \in A$ such that $x \leq y, \forall y \in A$.

Thus if S be a non-empty subset of natural numbers then there exists $p \in S$ such that $p \leq q, \forall q \in S$.

A set containing just one element has a smallest member, namely the element itself. Hence the well ordering principle is true for sets of size 1.

Now, let us assume that the principle is true for sets of size n, i.e. any set of n natural numbers has a smallest number.

Let us now consider a set S of $(n+1)$ numbers from which one element 'p' is removed. The remaining n numbers have a smallest element say q (by the induction hypothesis). The smaller of p and q is the smallest element of S.

Hence, by the principle of mathematical induction, it follows that any non-empty finite set of natural numbers has a smallest element.

Divisibility theory:

Definition. Given two integers a and b, we say $a(\neq 0)$ divides b if there is an integer c such that

$b = ac$. If a divides b, we write $a|b$. If a does not divide b, we write $a \nmid b$.

Note. (i) When a divides b, a is called a divisor or factor of b, and b is called multiple of a.

(ii) If a divides b then, $-a$ also divides b because $b = ac \Longrightarrow b = (-a)(-c)$, $-c$ is an integer.

i.e. $a|b \Rightarrow -a|b$

Illustrations.

(i) 38 is divisible by 19 because $38 = 2 \times 19$.

(ii) 11 is a divisor of 143 since $143 = 11 \times 13$.

(iii) 0 is divisible by every integer because $0 = x \times 0$, for every value of x.

Properties of divisibility:

For any integers a, b, c and d the following statements hold :

(i) $a|0$, $a|1$ and $a|a$

(ii) $x|y \Rightarrow -ax|ay, \forall a \in \mathbb{Z}$

(iii) $a|b$ and c$|d \Rightarrow ac|bd$.

(iv) $a|b$ and $b|c \Rightarrow a|c$ (Transitivity)

(v) $a|b$ and $b|a \Rightarrow a = \pm b$.

(vi) $a|b$ and $a|c \Rightarrow a|(bx + cy)$, for arbitrary integers x, y.

Remark: we use the symbol ‘ ∤ ’ to mean ‘does not divide’. Thus a∤b means ‘a does not divide b’ or ‘b is not divisible by a’.

Prime numbers: A positive integer $p > 1$ is called prime if the only positive factors of p are 1 and p. If a positive integer $n > 1$ is not prime, then n is called composite.

Note. (1) The positive integer 1 is neither prime nor composite.

(2) The only even prime integer is 2; rest of all primes are odd.

(3) If n is a composite, then there exists positive integers a and b such that $n = ab$ where $1 < a, b < n$

Ex. The integers 2,3,5,7 etc are prime where as 4,6,9,10 etc are composite since

$4 = 2.2, 6 = 3.2, 9 = 3.3, 10 = 5.2$

Lecture 18:

Fundamental Theorem of Arithmetic:

If a is an integer larger than $1(a > 1)$, then a can be written as a product of primes. Furthermore, this factorization is unique except for the order of the factors.

Note. (1) The unique expression for the integer $n > 1$ as a product of primes is called the **prime decomposition** or **prime factorisation** of n.

(2) If there be n_i prime factors of n, each equal to r_i , where $1 \leq i \leq p$, then n can be written as $n = n_1^{r_1} n_2^{r_2} \dots n_p^{r_p}$

Illustration. The prime factorisation of 100, 216 are given by

$$100 = 2 \times 2 \times 5 \times 5 = 2^2 \times 5^2$$

$$216 = 2 \times 2 \times 2 \times 3 \times 3 \times 3 = 2^3 \times 3^3$$

Theorem (Euclid): The number of prime numbers is infinite.

Proof: If possible, let the number of primes be finite and be equal to n. Let them be arranged as $p_1, p_2, \dots, p_n$. Then form the number

$$n = 1 + p_1 . p_2 \dots . . p_n$$

Now no one of $p's$ is a divisor of n.

∴ n is either a prime $> p_n$ or has a prime $> p_n$ as a factor.

But this contradicts our assumption that p_n is the greatest prime.

Hence the number of prime is infinite.

Theorem: If $n > 1$ is a composite integer then there exists a prime p such that $p|n$ and $p \leq \sqrt{n}$.

Proof: Since $n > 1$ is a composite integer, n can be expressed as $n = ab$, where $1 < a \leq b < n$, then

$a \leq \sqrt{n}$, since if it is not true, $a > \sqrt{n}$

$\therefore a . b > \sqrt{n} . \sqrt{n} \Longrightarrow ab > n$ which is a contradiction.

Now by fundamental theorem, either a is a prime or has a prime divisor p

$\therefore p|n$ and $p \leq \sqrt{n}$.

Note: To check a given integer n is prime it is sufficient to see that it is not divisible by any prime less than or equal to its square root.

Illustration: Let $n = 19$

$$\therefore 4 < \sqrt{19} < 5$$

Here 2 and 3 are the prime less than or equal to 4. But 19 is not divisible by 2 and 3.

∴ 19 must be a prime number.

Division algorithm:

Let a and b be two integers with $b > 0$. Then there exist unique integers q, r such that $a = qb + r$, where $0 \leq r < b$. The integer q is called the quotient and r, the remainder.

Ex.1. Use division algorithm to prove that the square of an odd integer is of the form $8k + 1$, where k is an integer.

Solution:

By division algorithm every integer, upon division by 4, leaves one of the remainders 0,1,2,3. Therefore any integer is one of the forms $4q, 4q + 1, 4q + 2, 4q + 3$, where q is an integer.

Odd integers are of the forms $4q + 1, 4q + 3$.

Now $(4q + 1)^2 = 8(2q^2 + q) + 1$ is of the form $8k + 1$

$(4q + 3)^2 = 8(2q^2 + 3q + 1) + 1$ is of the form $8k + 1$

Hence the square of an odd integer is of the form $8k + 1$.

Ex.2. Prove that the product of any m consecutive integers is divisible by m.

Solution:

Let the consecutive integers be $c, c + 1, c + 2, \ldots, c + (m - 1)$

Let q be the quotient and r be the remainder when c is divided by m.

Then $c = mq + r, \qquad 0 \leq r < m$

When $r = 0, c = mq$ and therefore $m|c$;

When $r = 1, c + (m - 1) = m(q + 1)$ and therefore $m|c + (m - 1)$;

When $r = 2, c + (m - 2) = m(q + 1)$ and therefore $m|c + (m - 2)$;

… … … … …

When $r = m - 1, c + 1 = m(q + 1)$ and therefore $m|c + 1$.

Therefore whatever integer r may be, m divides one of the integers $c, c + 1, \dots, c + (m - 1)$ and it follows that the product $c(c + 1)(c + 2)\dots(c + m - 1)$ is always divisible by m.

Lecture 19:

Greatest Common Divisor: A positive integer d is called a common divisor of the integers a and b, if d divides a and b. The greatest possible such d is called the greatest common divisor of a and b, denoted $gcd(a,b)$. If $gcd(a,b) = 1$ then a,b are called relatively prime.

Example: The set of positive divisors of 12 and 30 is {1,2,3,6}.

The greatest common divisor of 12 and 30 is $gcd(12,30) = 6$.

A few properties of divisors are the following.

Let m,n,d be integers. Then:

1. If $d|m$ and $d|n$ then $d|(m + n)$.

2. If $d|m$ and $d|n$ then $d|(m - n)$.

3. If $d|m$ then $d|mn$.

Some another properties:

1. If $c|ab$ and b,c are coprime, then$c|a$.
2. If a and b are coprime and a and c are coprime, then a and bc are coprime.
3. If $gcd(a,b) = 1$, then for any integer x, $gcd(ax,b) = gcd(x,b)$.

Note: Two numbers a and b are said to be coprime if $gcd(a,b) = 1$.

Another important result is the following:

Given integers a,b,c, the equation $ax + by = c$ has integer solutions if and only if $gcd(a,b)$ divides c. That is an example of a Diophantine equation. In general a Diophantine equation is an equation whose solutions must be integers.

Example: Consider the Diophantine equation $2x + 3y = 4$.

We have $gcd(2,3) = 1$, and 4 is divisible by 1. So, the given equation has integral solution. Now $1 = 2.(-1) + 3.1$

$\therefore 4 = 2.(-4) + 3.4$

Thus one integral solution is $x_0 = -4, y_o = 4$.

Lemma: For any integers a and b, we have

$gcd(a,b) = gcd(b,a) = gcd(\pm a, \pm b) = gcd(a, b-a) = gcd(a, b + a)$.

Theorem (Euclidean Algorithm): Let a and b be nonzero integers. Divide b into a and carry out further divisions according to the following method, where the old remainder becomes the new divisor:

$$a = bq_1 + r_1,\ 0 \leq r_1 < b$$

$$b = r_1 q_2 + r_2, 0 \leq r_2 < r_1$$

$$r_1 = r_2 q_3 + r_3, 0 \leq r_3 < r_2$$

$$\dots\dots\dots\dots$$

The non-negative remainders $r_1, r_2, \dots$. are strictly decreasing, and thus must eventually become 0. The last nonzero remainder is the greatest common divisor.

This algorithm of Euclid for finding $gcd(a,b)$ can be carried out very rapidly on a computer, even for very large integers which are not easy to factor into primes.

Example. let's see how it looks for the pair given below.

Consider $a = 19088597$ and $b = 39083$

$$19088597 = 39083 \cdot 488 + 16093$$

$$39083 = 16093 \cdot 2 + 6897$$

$$16093 = 6897 \cdot 2 + 2299$$

$$6897 = 2299 \cdot 3 + 0.$$

The last nonzero remainder is 2299, and we said gcd(19088597,39083) $= 2299$

Example. We compute $(322345, 21419)$:

$$322345 = 21419 \cdot 15 + 1060,$$

$$21419 = 1060 \cdot 20 + 219,$$

$$1060 = 219 \cdot 4 + 184,$$

$$219 = 184 \cdot 1 + 35,$$

$$184 = 35 \cdot 5 + 9,$$

$$35 = 9 \cdot 3 + 8,$$

$$9 \;=\; 8 \cdot 1 + 1,$$

$$8 \;=\; 1 \cdot 8 + 0.$$

Therefore $gcd(322345{,}21419) \;=\; 1$. The last equation was superfluous: if we ever reach a remainder of 1, then the next remainder is ≥ 0 and less than 1 and therefore must be 0, so 1 is the last nonzero remainder.

Not only is the last equation superfluous, but we could have stopped already in the fourth equation: here we meet a remainder of 35, which is small enough that we can factor it in our heads as $5 \cdot 7$. Therefore

$$(322345, 21419) \;=\; (184{,}35),$$

and we can easily check 5 and 7 are not factors of 184, so this greatest common divisor must be 1.

Theorem: Let a and b $(a > b)$ be any two integers. Then $gcd(a, b)$ can be expressed as an integral linear combination of a and b

i.e. $gcd(a, b) = ma + nb$, where m and n are integers.

Illustration: First we consider the steps used to find $gcd(1120{,}128)$ as given below:

$1120 = 8.128 + 96$ ....(1)

$128 = 1.96 + 32$ ....(2)

∴ From (2), we have

$$32 = 128 - 1.96$$

$$= 128 - 1.(1120 - 8.128) \text{ , by (1)}$$

$$=9.128 - 1.1120 \quad \text{...........(3)}$$

Thus $32 = gcd(1120{,}128) = (-1).1120 + 9.128$

$$\therefore m = -1, n = 9$$

Note: $gcd(1120{,}128) = \; (9 + 1120).128 - (1 + 128).1120$

$$= 1129.128 - 129.1120 \;\; = (-129).1120 + 1129.128$$

$$\therefore m = -129, n = 1129$$

Thus m and n are not unique.

Hence the expression of $gcd(a, b)$ in the form of $ma + nb$ is not unique.

Ex. Show that $gcd(a, a + 2) = 1$ or 2 for every integer a.

Ans. Let $d = gcd(a, a + 2)$ then $d|a$ and $d|a + 2$

Therefore $d|ax + (a + 2)y$ for all integers x, y.

Taking $x = -1$ and $y = 1$, it follows that $d|2$. i.e. d is either 1 or 2.

Theorem: If k be a positive integer, $gcd(ka, kb) = k. gcd(a, b)$.

Proof. Let $d = gcd(a, b)$. Then there exist integers u and v such that $d = au + bv$.

Since $d = gcd(a, b)$, $d|a$ and $d|b$.

$$d|a \Longrightarrow kd|ka, \qquad d|b \Longrightarrow kd|kb$$

Therefore kd is a common divisor of ka and kb.

Let c be a common divisor of ka and kb.

$c|ka \Longrightarrow ka = pc$ for some integer p, and $c|kb \Longrightarrow kb = qc$ for some integer q.

Now $kd = k(au + bv) = pcu + qcv = (pu + qv)c$.

As $(pu + qv)$ is an integer, it follows that $c|kd$.

Consequently, $kd = gcd(ka, kb)$, i.e. $gcd(ka, kb) = k. gcd(a, b)$.

Note: (1) Two integers a and b, not both zero, are said to be prime to each other (or relatively prime) if $gcd(a, b) = 1$.

(1) Let a and b be integers not both zero. Then a and b are prime to each other iff there exists integers u and v such that $1 = au + bv$.

Lecture 20:

CONGRUENCE:

If a and b are integers and $n > 0$, we write $a \equiv b \ (mod\ n)$ to mean $n|(b - a)$. We read this as "a is congruent to b modulo (or mod) n."

For example, $29 \equiv 8\ (mod\ 7)$, and $60 \equiv 0\ (mod\ 15)$.

Theorem: For any integers a and b, and positive integer n, we have:

1. $a \equiv a\ (mod\ n)$.

2. If $a \equiv b\ (mod\ n)$ then $b \equiv a\ (mod\ n)$.

3. If $a \equiv b\ (mod\ n)$ and $b \equiv c\ (mod\ n)$ then $a \equiv c\ (mod\ n)$.

These results are classically called:

1. Reflexivity; 2. Symmetry; and 3. Transitivity.

The proof is as follows:

1. $n|(a - a)$ since 0 is divisible by any integer. Therefore $a \equiv a\ (mod\ n)$.

2. If $a \equiv b\ (mod\ n)$ then $n|(b - a)$. Therefore, $n|(-1)(b - a)$ or $n|(a - b)$. Therefore, $b \equiv a\ (mod\ n)$.

3. If $a \equiv b\ (mod\ n)$ and $b \equiv c\ (mod\ n)$ then $n|(b - a)$ and $n|(c - b)$. Using the linear combination theorem, we have $n|(b - a + c - b)$ or $n|(c - a)$. Thus, $a \equiv c\ (mod\ n)$.

Theorem: If $a \equiv b\ (mod\ n)$ then $b = a + nq$ for some integer q, and conversely.

Proof: If $a \equiv b\ (mod\ n)$ then by definition $n|(b - a)$. Therefore, $b - a = nq$ for some q. Thus $b = a + nq$. Conversely if $b = a + nq$, then $b - a = nq$ and so $n|(b - a)$ and hence $a \equiv b\ (mod\ n)$.Then $b = a + nq$.

Theorem: If $a \equiv b\ (mod\ n)$ then a and b leave the same remainder when divided by n. Conversely if a and b leave the same remainder when divided by n, then $a \equiv b\ (mod\ n)$.

Proof: Suppose $a \equiv b\ (mod\ n)$. Then $b = a + nq$. If a leaves the remainder r when divided by n, we have $a = nQ + r$ with $0 \leq r < n$.

Therefore, $b = a + nq = nQ + r + nq = n(Q + r) + r$, and so b leaves the same remainder when divided by n. The converse is straightforward and we omit the proof.

Theorem: If $a \equiv b\ (mod\ n)$ and $c \equiv d\ (mod\ n)$ then

1. $a + c \equiv b + d\ (mod\ n)$.

2. $ac \equiv bd\ (mod\ n)$.

Proof: Write $b = a + nq_1$ and $d = c + nq_2$. Then adding equalities, we get $b + d = a + c + nq_1 + nq_2 = a + c + n(q_1 + q_2)$.

This shows that $a + c \equiv b + d\ (mod\ n)$.

Similarly, multiplying, we get

$bd = (a + nq_1)(c + nq_2) = ac + naq_2 + ncq_1 + n^2q_1q_2$.

Thus, $bd = ac + n(aq_2 + cq_1 + nq_1q_2)$, and so $ac \equiv bd\ (mod\ n)$.

Theorem: If $a \equiv b\ (mod\ n)$, then $ac \equiv bc\ (mod\ n)$.

Proof: Since $a \equiv b\ (mod\ n)$, $a - b$ is divisible by n.

$\therefore (a - b)c = ac - bc$ is also divisible by n.

$\therefore\ ac \equiv bc\ (mod\ n)$.

Theorem: $ca \equiv cb\ (\ mod\ m\ )$ implies $a \equiv b\ (\ mod\ m\ )$ if and only if $gcd(c, m) = 1$.

Proof: Note that we already know that $a \equiv b\ (\ mod\ m\ )$ implies $ca \equiv cb\ (\ mod\ m\ )$, We will prove the other direction, which is what is new, and that allows us to divide.

That is, if $gcd(c, m) = 1$, then $ca \equiv cb\ (\ mod\ m\ )$ implies $a \equiv b\ (\ mod\ m\ )$.

$ca \equiv cb\ (\ mod\ m\ )$ implies $m \mid (\ ca - cb\ )$. That is, $m \mid c(a - b)$.

Since $gcd(c, m) = 1$, a theorem from divisibility tells us that $m \mid (\ a - b)$.

Hence $a \equiv b\ (\ mod\ m\ )$.

Example1: Find the remainder when $25^{100} + 11^{500}$ is divided by 3.

We observe that $25 \equiv 1\ (\ mod\ 3\ )$ and $11 \equiv -1\ (\ mod\ 3\ )$. Raising these to the appropriate powers, $25^{100} \equiv 1^{100}\ \ (\ mod\ 3\ )$ and $11^{500} \equiv (-1)^{500}\ (\ mod\ 3\ )$.

That is, $25^{100} \equiv 1(\ mod\ 3\ )$ and $11^{500} \equiv\ \ (\ mod\ 3\ )$. Adding these congruencies, we get $25^{100} + 11^{500} \equiv 2(\ mod\ 3\ )$

Thus the remainder is 2.

Example 2: Prove that $3.4^{n+1} \equiv 3(mod\ 9)$ for all positive integer n.

Ans. $3.4^{n+1} = 12.4^{n} = 9.4^{n} + 3.4^{n}$
$3.4^{n} = 12.4^{n-1} = 9.4^{n-1} + 3.4^{n-1}$
… … … … … …
$3.4^{2} = 12.4 = 9.4 + 3.4$

$3.4 = 12 = 9 + 3$
Therefore, $3.4^{n+1} = 9(1 + 4 + 4^{2} + \cdots + 4^{n}) + 3$

Hence $3.4^{n+1} \equiv 3(mod\ 9)$

<u>Lecture 21:</u>

Residue classes of integer modulo n:

The set of all integers b which are congruent to a modulo n is called the congruence class of integer modulo n or residue classes of integer modulo n and is denoted by $[a]_n$ or $[a]$ or, $\bar{a}$.

Thus, $[a] = \{b \in \mathbb{Z} : b \equiv a \ (mod\ n)\}$.

For example, all congruence classes of integer modulo 4 are

$[0] = \{b \in \mathbb{Z} : b \equiv 0 \ (mod\ 4)\}$

$= \{b \in \mathbb{Z} : b\ is\ divisible\ by\ 4\ or\ b = 4k\ for\ some\ k)\}$

$= \{\ldots - 8, -4, 0, 4, 8, \ldots\}$

$[1] = \{b \in \mathbb{Z} : b \equiv 1 \ (mod\ 4)\}$

$= \{b \in \mathbb{Z} : b - 1 = 4k\ i.e. b = 1 + 4k\ for\ some\ k\}$

$= \{\ldots - 7, -3, 1, 5, 9, \ldots\}$

Similarly,

$[2] = \{\ldots - 6, -2, 2, 6, 10, \ldots\}$

$[3] = \{\ldots - 5, -1, 3, 7, 11, \ldots\}$

$[4] = \{\ldots - 8, -4, 0, 4, 8, \ldots\} = [0]$

Thus, there are only 4 distinct congruence classes of integer modulo 4, namely [0], [1], [2], [3].

In general [0], [1], [2],…, $[n - 1]$ are n distinct residue classes of integer modulo n.

It is evident that

$[a] = [a + n] = [a + 2n] = \cdots$

$[0] = [n] = [2n] = \cdots$

For any positive integer n, $\mathbb{Z}_n$ denote the set of all congruence classes of integer modulo n. Thus

$\mathbb{Z}_4 = \{[0], [1], [2], [3]\}$

In general,

$\mathbb{Z}_n = \{[0], [1], [2], \ldots [n - 1]\}$.

Theorem: The number of elements of $\mathbb{Z}_n$ is finite and this number is n.

Proof: Left as exercise.

Arithmetic of Residue classes: Addition and multiplication for residue classes of integer modulo n are defined as given below:

$[a] + [b] = [a + b]$ and $[a].[b] = [ab]$

As an illustration, consider the residue classes of integer modulo 3 namely $[0]$, $[1]$, $[2]$.

Then, $[0] + [1] = [1]$, $[1] + [2] = [3] = [0]$

$[1].[2] = [2]$, $[2].[2] = [4] = [1]$

Following composite tables show the addition and multiplication tables for the residue classes of integer modulo 3:

+	[0]	[1]	[2]
[0]	[0]	[1]	[2]
[1]	[1]	[2]	[0]
[2]	[2]	[0]	[1]

.	[0]	[1]	[2]
[0]	[0]	[0]	[0]
[1]	[0]	[1]	[2]
[2]	[0]	[2]	[1]

Definition: An element $[b] \in \mathbb{Z}_n$ is called an inverse of an element $[a] \in \mathbb{Z}_n$ if $[a][a] = [1]$ in $\mathbb{Z}_n$.

Definition: An element $[a] \in \mathbb{Z}_n$ is said to be a unit element in $\mathbb{Z}_n$ if $[a]$ has inverse in $\mathbb{Z}_n$.

Theorem: Let a and n be integers with $n \geq 2$ relatively prime. Then $[a]$ has an inverse in $\mathbb{Z}_n$ iff a and n are relatively prime.

Illustration: (i) The unit elements of $\mathbb{Z}_6$ are $[1]$and $[5]$, as $gcd(1,6) = gcd(5,6) = 1$

(ii) The inverse of $[5]$ in $\mathbb{Z}_6$ in $[3]$ as $[5][3] = [1]$

Ex. 1. Find all units of $\mathbb{Z}_{18}$.

Solution. We have

$$\mathbb{Z}_{18} = \{[0], [1], [2], \dots, [16], [17]\}$$

Now $gcd(1,18) = \text{gcd}(5,18) = \text{gcd}(7,18)$

$$= \text{gcd}(11,18) = \text{gcd}(13,18) = \text{gcd}(17,18) = 1.$$

Hence $[1], [5], [7], [11], [13], [17]$ are the only unit elements of $\mathbb{Z}_{18}$.

Ex.2. In $\mathbb{Z}_{16}$ find the inverse of $[9]$ and use it to solve $[9]x = [12]$

Solution: Since $gcd(9,16) = 1$, the inverse $[9]$ exists in $\mathbb{Z}_{16}$. From the Euclidean algorithm, we have

$$16 = 1.9 + 7$$

$$9 = 1.7 + 2$$

$$7 = 3.2 + 1$$

$$\therefore 1 = 7 - 3.2 = 7 - 3.(9 - 1.7)$$

$$= 4.7 - 3.9$$

$$= 4.(16 - 1.9) - 3.9$$

$$= 4.16 - 7.9$$

$$= 16.4 + 9(-7)$$

Hence,

$[1] = [16][4] + [9][-7]$

$=[0][4] + [9][9]$ $\quad\quad\left[\because 9 \equiv -7\ (mod\ 16) = [9][9]\right]$

Thus the inverse of $[9]$ is $[9]$

Now $[9]x = [12]$

$\Rightarrow [9][9]x = [12][9]$

$\Longrightarrow x = [-4][-7]$

$= [28]$

$= [12]$.

Linear congruence:

Let $f(x) = a_0x^n + a_1x^{n-1} + \cdots + a_n, (n \geq 1)$ be a polynomial with integer coefficients $a_0, a_1, \ldots, a_n$ with $a_0 \not\equiv 0\ (mod\ m)$. Then $f(x) \equiv 0\ (mod\ m)$ is said to be a polynomial congruence $(mod\ m)$ of degree n.

If there exists an integer x_0 such that $f(x_0) \equiv 0 \ (mod\, m)$, then x_0 is said to be a solution of the congruence.

Definition: A polynomial congruence of degree 1 is said to be a linear congruence. The general form of a linear congruence modulo a positive integer $m > 1$ is $ax \equiv b \ (mod\, m)$, where $a \not\equiv 0 \ (mod\, m)$.

An integer c is said to be a solution of the linear congruence $ax \equiv b \ (mod\, m)$ if

$ac \equiv b \ (mod\, m)$.

Theorem: If x_1 be a solution of the linear congruence $ax \equiv b \ (mod\, m)$ and if $x_2 \equiv x_1 \ (mod\, m)$,then x_2 is also a solution of the congruence.

Theorem: If $gcd(a, m) = 1$, then the linear congruence $ax \equiv b \ (mod\, m)$ has a unique solution.

Theorem: If $gcd(a, m) = d$, then the linear congruence $ax \equiv b \ (mod\, m)$ has no solution if d is not a divisor of b.

Ex.1. Solve the linear congruence $5x \equiv 3 \ (mod\ 11)$.

Ans: $gcd(5,11) = 1$. Hence the congruence has a unique solution.

Since $gcd(5,11) = 1$, there exists integers u and v such that $5u + 11v = 1$.

Here $u = -2, v = 1$. Therefore $5.(-2) + 11.1 = 1$ and this implies $5.(-2) \equiv 1 \ (mod\ 11)$.

Therefore $5.(-6) \equiv 3 \ (mod\ 11)$.

Hence $x = -6$ is a solution.

All solutions are $x \equiv -6 \ (mod\ 11), i.e. x \equiv 5 \ (mod\ 11)$.

All the solutions are congruent to 5 (mod 11) and therefore the given congruence has a unique solution.

Exercise: Solve the linear congruence $15x \equiv 9 \ (mod\ 18)$.

Lecture 22:

- **Cartesian product (review):**

Let A={a_1, a_2, ..a_k} and B={b_1,b_2,..b_m}.

The Cartesian product A x B is defined by a set of pairs {(a_1 b_1), (a_1, b_2), … (a_1, b_m), …, (a_k,b_m)}.

Cartesian product defines a product set, or a set of all ordered arrangements of elements in sets in the Cartesian product.

- **Definition of Relation:**

A relation is something that relates one set of values to another set of values. Sometimes the relationship that is specified between sets is meaningful, other times it is not.

In general, a relations are defined in the following manner:

A relation R defined over sets A and B is a subset of A x B. Thus, we have R $\subseteq$ A x B. This is known as a binary relation, because it relates elements between two sets.

- **Illustration:**

- Let R⊆ A x B means R is a set of ordered pairs of the form (a,b) where a ∈A and b∈ B.
- We use the notation **a R b** to denote (a , b)∈ R . If **a R b**, we say **a** is related to **b** by **R**.

Example: Let A={a,b,c} and B={1,2,3}.

- Is R={(a,1),(b,2),(c,2)} a relation from A to B? Yes.
- Is Q={(1,a),(2,b)} a relation from A to B? No.
- Is P={(a,a),(b,c),(b,a)} a relation from A to A? Yes

- **Domain and Range of a relation:**

In domain and range of a relation, if R be a relation from set A to set B, then

• The set of all first components of the ordered pairs belonging to R is called the domain of R. Thus, Dom(R) = {a ∈ A: (a, b) ∈ R for some b ∈ B}.

• The set of all second components of the ordered pairs belonging to R is called the range of R.

Thus, range of R = {b ∈ B: (a, b) ∈R for some a ∈ A}.

Therefore, Domain (R) = {a : (a, b) ∈R} and Range (R) = {b : (a, b) ∈R}

Note:

The domain of a relation from A to B is a subset of A.

The range of a relation from A to B is a subset of B.

- **Some Useful Definition:**

Let R be a binary relation on A.

- R is **reflexive** if for all $x \in$ A, $(x,x) \in$ R. (Equivalently, for all $x \in$ A, x R x.)
- R is **symmetric** if for all $x,y \in$ A, $(x,y) \in$ R implies $(y,x) \in$ R. (Equivalently, for all $x,y \in$ A, x R y implies that y R x.
- R is **transitive** if for all $x,y,z \in$ A, $(x,y) \in$ R and $(y,z) \in$ R implies $(x,z) \in$ R. (Equivalently, for all $x,y,z \in$ A, x R y and y R z implies x R z.)

Examples:

- Reflexive: The relation R on {1,2,3} given by R = {(1,1), (2,2), (2,3), (3,3)} is reflexive. (*All loops are present.*)

- Symmetric: The relation R on {1,2,3} given by R = {(1,1), (1,2), (2,1), (1,3), (3,1)} is symmetric. (*All paths are 2-way.*)

- Transitive: The relation R on {1,2,3} given by R = {(1,1), (1,2), (2,1), (2,2), (2,3), (1,3)} is transitive. (*If I can get from one point to another in 2 steps, then I can get there in 1 step.*)

- **The Anti-symmetry Property**
- **Definition:**

A relation R on a set A is called *anti-symmetric* if $(x,y) \in$ R and $(y,x) \in$ R implies $x = y$.

- This is equivalent to requiring that if $x \neq y$ and $(x,y) \in$ R, then $(y,x) \notin$ R. (*All streets are one-way.*)

- Example: R = {(1,1), (1,2), (3,2), (3,3)} is anti-symmetric.

- Is every relation symmetric or anti-symmetric?

- No! Consider R = {(1,2), (2,1), (1,3)}.

- **<u>Partial Order Relation</u>**:

  – A relation R on a set S is called a partial order if it is

    - Reflexive
    - Antisymmetric
    - Transitive

  – A set S together with a partial ordering R is called a partially ordered set (poset, for short) and is denote (S,R)

- Partial orderings are used to give an order to sets that may not have a natural one.
- In our renovation example, we could define an ordering such that (a,b) ∈ R if ‘a must be done before b can be done’.

Example:

- **Show that “greater than or equal” relation is a partial ordering on the set of integers?**

  a≥a for every integer a (reflexive)

  a≥b, b≥a, then a=b (anti-symmetric)

  a≥b, b≥c, then a≥c (transitive)

- Thus ≥ is a partial ordering on the set of integers
- (Z, ≥) is a poset.
- Similarly, the division symbol ‘|’ is a partial ordering on the set of positive integers.
- The inclusion relation is a partial ordering on the set of P(S)

- **Partial Orderings: Notation:**

- We use the notation:

  a $\preceq$ b, when (a,b)∈R

- The notation $\preceq$ is used to denote <u>any</u> partial ordering.

- **Comparable and Incomparable:**

- The elements a and b of a poset (S, $\preceq$) are called comparable, if either a $\preceq$ b or b $\preceq$ a. When a and b are elements of S such that neither a $\preceq$ b or b $\preceq$ a, they are called incomparable.
- In the Poset (Z^+,|), are the integers 3 and 9 comparable? Yes, as 3|9 => 3 $\preceq$ 9.
- But 5 and 7 are incomparable.

Properties of posets:

1. **ORDERED SETS:**

- **Definitions. (Partially and totally ordered sets.) :**

(1) The ordered pair (X,R) is called a partially ordered set if X is a set and R is a partial order relation in X.

(2) The ordered pair (X,R) is called a totally ordered set (or linear ordered set) if X is a set and R is a total (linear) order relation in X .

Let (X,R) be a (partially or totally) ordered set and $S \subset X$.

- **Special elements of a Po set:**

- **Definitions. (Upper and lower bounds, bounded set.) :**

(1) $u \in X$ is called an upper bound for S if $\forall x \in S\ (x,u) \in R$.

(2) $v \in X$ is called a lower bound for S if $\forall x \in S\ (v,x) \in R$.

(3) S is called bounded above if $\exists u \in X$ upper bound for S.

(4) S is called bounded below if $\exists v \in X$ lower bound for S.

(5) S is called bounded if it is bounded both above and below.

- **Definitions. (Maximal and minimal elements.) :**

(1) $M \in S$ is called a maximal element of S if there is no $x \in S$ such that $x \neq M$ and $(M,x) \in R$.

(2) $m \in S$ is called a minimal element of S if there is no $x \in S$ such that $x \neq m$ and $(x,m) \in R$.

- **Definitions. (Greatest and least elements.):**

(1) $M \in S$ is called the greatest element of S if $\forall x \in S\ (x,M) \in R$.
(2) $m \in S$ is called the least element of S if $\forall x \in S\ (m,x) \in R$.

- **Definitions. (Supremum and infimum.):**

(1) If S is bounded above and the set of all upper bounds for S has a least element, we call it the least upper bound or the supremum for S, and denote it by $\sup S$. (I.e., $\sup S \in U := \{u \in X : \forall x \in S\ (x,u) \in R\}$ and $\forall u \in U\ (\sup S, u) \in R$.)

(2) If S is bounded below and the set of all lower bounds for S has a greatest element, we call it the greateast lower bound or the infimum for S, and denote it by $\inf S$. (I.e., $\inf S \in V := \{v \in X : \forall x \in S\ (v,x) \in R\}$ and $\forall v \in V\ (v, \inf V) \in R$.)

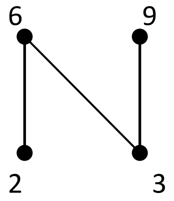
Example.

Let ρ be the divisibility in $N+$. Then $(N+,\rho)$ is a partially ordered set. 1 is the only minimal (and least) element of $N+$, 1 is also the infimum for $N+$. The primes are minimal elements of $N+\setminus\{1\}$, there is no least element of $N+\setminus\{1\}$, 1 is the infimum for $N+\setminus\{1\}$.

Lecture 23:

- **Hasse Diagram:**

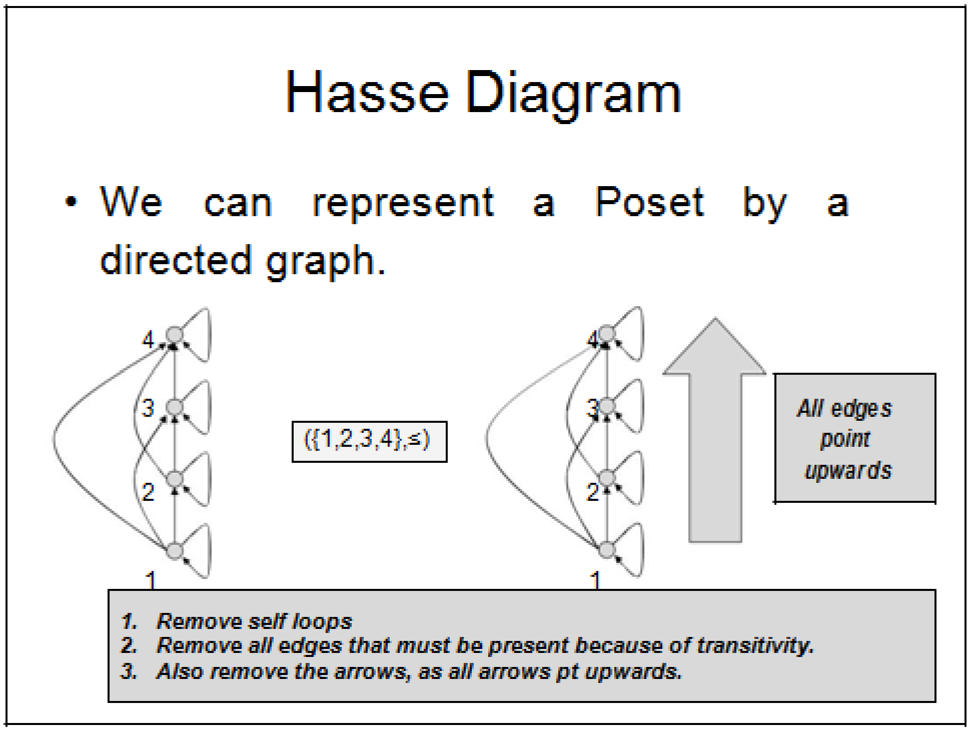
Posets are depicted with **Hasse diagrams**. The Hasse diagram of a poset shows the poset's covering relations. The Hasse diagram of P is formed by representing each element of P with a dot. These dots are arranged such that if y covers x, x is drawn below y and the two dots are connected with a line. For example, consider the set P={2,3,6,9}. Division gives a partial ordering for this set: x ≤ y if x divides y. With the dots labeled for clarity, the Hasse diagram for this poset is:



A **connected poset** is a poset whose Hasse diagram is a connected graph. As in graph theory, if y covers x in P, we call x a **child** of y and refer to y as a **parent** of x. Often we consider Hasse diagrams whose dots are labeled with integers. A **labeled poset on [n]** is a partial ordering on n: The elements in the Hasse diagram are labeled with the integers 1 to n. The poset is **naturally labeled** if i < j in P implies that i < j as integers. Visually, the label of a dot must be larger than the labels of its children.

Within a poset, there are maximal and minimal elements. A **maximal element** in P is an element that is not covered by any other elements. Similarly, a **minimal element** is one that does not cover any other elements. Hasse diagrams make maximal and minimal elements particularly easy to identify. Minimal elements have no lines below them and maximal elements have no lines drawn above them in the Hasse diagram. In the above example, 2 and 3 are minimal elements while 6 and 9 are maximal. At times we will consider posets with a unique maximal element. Observe that if a poset is not connected, each connected component of its Hasse diagram will have at least one maximal element. So posets with a unique maximal element are necessarily connected.

- **How to Draw Hasse diagram:**
  - There is a dot for each $a \in A$
  - If $a \prec b$, then the dot for b is positioned higher than the dot for a
  - If $a \prec b$ and there is no c such that $a \prec c \prec b$, then a line is drawn from a to b (say "b covers a").



Hasse Diagram

- We can represent a Poset by a directed graph.

1. *Remove self loops*
2. *Remove all edges that must be present because of transitivity.*
3. *Also remove the arrows, as all arrows pt upwards.*

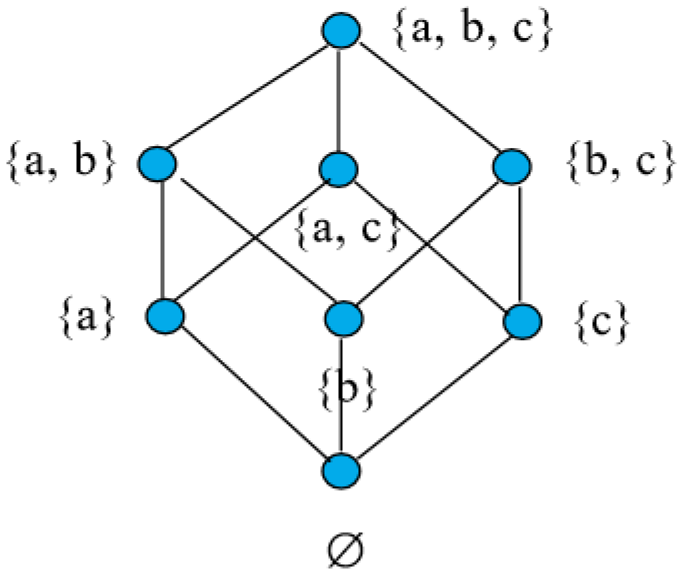
Example:

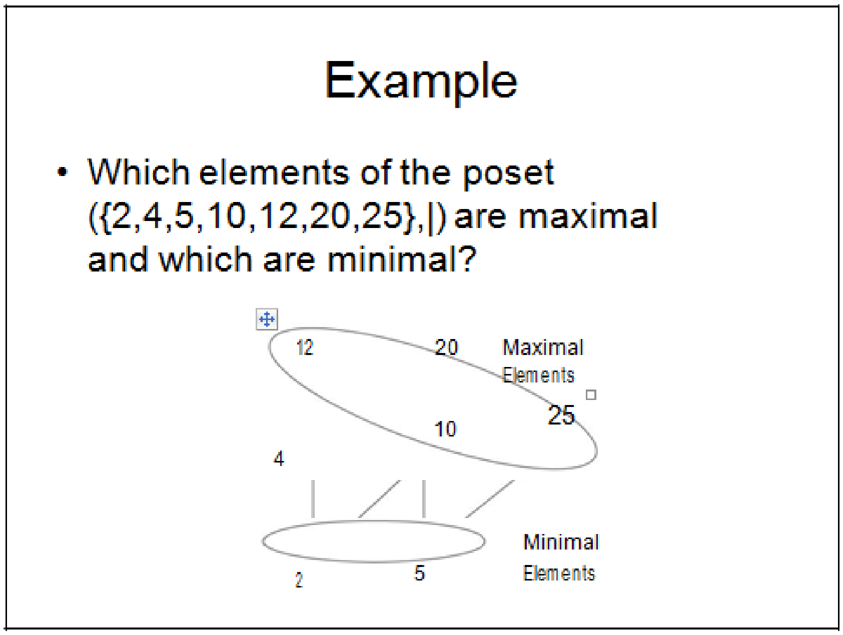
Construct the Hasse diagram of $(P(\{a, b, c\}), \subseteq)$

The elements of $P(\{a, b, c\})$ are

∅ ,{a}, {b}, {c} ,{a, b}, {a, c}, {b, c}{a, b, c}

The digraph is



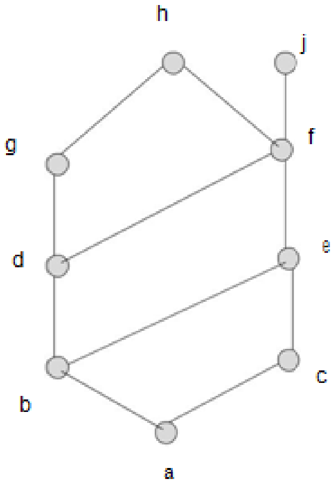


Example

- Which elements of the poset ({2,4,5,10,12,20,25},|) are maximal and which are minimal?

12
20
Maximal
Elements
25
10
4
Minimal
Elements
2
5

Example



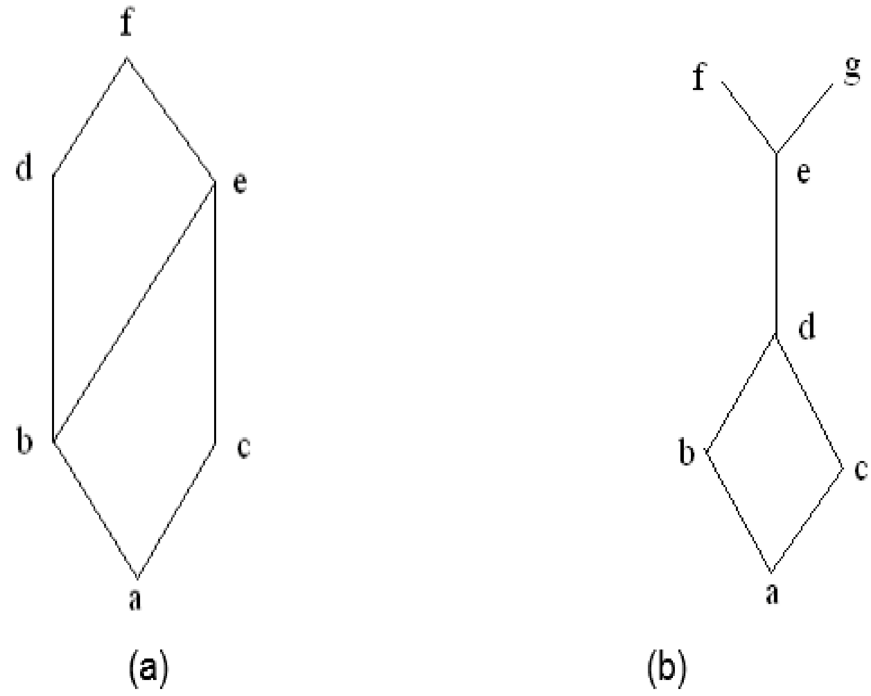
- UB({a,b,c})={e,f,j,h}
- LB({a,b,c})=a
- UB({j,h})={ }
- LB({j,h})={a,b,c,d,e,f}
- UB({a,c,d,f})={h,f,j}
- LB ({a,c,d,f})={a}
- glb({b,d,g})=max({a,b})=b
- lub({b,d,g})=min({g,h})=g

Lecture 24:

- **Lattice:**

A lattice is a poset in (L,≤) in which every subset {a,b} consisiting of two elements has a least upper bound and a greatest lower bound.

- LUB({a,b}) is denoted by a v b and is called the join of a and b.
- GLB({a,b}) is denoted by a Λ b and is called the meet of a and b.



a) is a lattice.
b) Is not a lattice because f v g does not exist.

- **Theorem*:***

If (L1, ≤) and (L2, ≤) are lattices then (L, ≤) is a lattice where L = L1xL2 and the partial order ≤ of L is the product partial order.

- **Properties of Lattices:**

- Idempotent Properties
  $a \vee a = a$
  $a \Lambda a = a$
- Commutative Properties
  $a \vee b = b \vee a$
  $a \Lambda b = b \Lambda a$
- Associative Properties
  $a \vee (b \vee c)= (a \vee b) \vee c$
  $a \Lambda(b \Lambda c)= (a \Lambda b) \Lambda c$
- Absorption Properties
  $a \vee (a \Lambda b) = a$
  $a \Lambda (a \vee b) = a$

- **Theorem:**
  - $a \vee b = b$ iff $a \leq b$
  - $a \Lambda b =a$ iff $a \leq b$
  - $a \Lambda b =a$ iff $a \vee b = b$

- **Types of Lattices:**

1. Isomorphic Lattices

If f: L1 -> L2 is an isomorphism from the poset (L1, ≤1) to the poset (L2, ≤2) then L1 is a lattice iff L2 is a lattice.

If a and b are elements of L1 then $f(a \Lambda b) = f(a) \Lambda f(b)$ and

$f(a \vee b) = f(a) \vee f(b)$

If two lattices are isomorphic as posets we say they are isomorphic lattices.

2. Bounded Lattice

A lattice L is said to be bounded if it has a greatest element I and a least element 0.

3. Complemented Lattice

A lattice L is said to be complemented if it is bounded and if every element in L has a complement.

- **Theorem:**

Let L = {a1,a2,a3,a4…..an} be a finite lattice. Then L is bounded.

- **Theorem:**

Let L be a bounded lattice with greatest element I and least element 0 and let a belong to L. an element a’ belong to L is a complement of a if

$$a \vee a' = I \text{ and } a \wedge a' =0$$

- **Theorem:**

Let L be a bounded distributive lattice. If complement exists it is unique.

MODULE IV: PRINCIPLES OF COUNTING TECHNIQUES

(NUMBER OF LECTURES: 6)

LECTURE 25: INTRODUCTION TO COUNTING TECHNIQUES

25.1. INTRODUCTION

In our daily life, we often need to find the number of ways that a particular event can occur. Moreover, we must count objects to solve many different types of problems. For instances:

- 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.
- Recently, it has played a key role in mathematical biology, especially in sequencing DNA.
- Furthermore, counting techniques are used extensively when probabilities of events are computed.

The problem of counting and enumeration of specified objects, patterns and designs, arise in every area of mathematics, might not be addressed without the knowledge of logical counting techniques. The branch of Mathematics which deals with such logical counting techniques is known as *Combinatorics*.

Today, the interest in combinatorial analysis is fueled by important problems in science, including chemistry, biology, physics and computer science. This subject was studied as long ago as the seventeenth century, when combinatorial questions arose in the study of gambling games.

We will here study mainly the following counting techniques:

- The *basic rules of counting*, which can solve a tremendous variety of problems. For instance, we can use these rules to enumerate the different telephone numbers possible in the United States, the allowable passwords on a computer system, and the different orders in which the runners in a race can finish.
- The *permutations* and *combinations*, which can phrase many counting problems in terms of ordered or unordered arrangements of the objects of a set with or without repetitions. For instance, suppose the 100 top finishers on a competitive exam taken by 2000 students are invited to a banquet. We can count the possible sets of 100 students that will be invited, as well as the ways in which the top 10 prizes can be awarded.
- The *pigeonhole principle*, which states that when objects are placed in boxes and there are more objects than boxes, then there is a box containing at least two objects. For instance, we can use this principle to show that among a set of 15 or more students, at least 3 were born on the same day of the week.
- The recurrence relation and generating functions, which involves generating all the arrangements of a specified kind. This is often important in computer simulations. We will devise algorithms to generate arrangements of various types.

25.2. BASIC COUNTING PRINCIPLES

There are two basic counting rules or principles which we use frequently in solving the problems of counting.

Rule 1: Addition Rule (Principle of Disjunctive Counting)

If an event E can occur in m ways and an another event F can occur in n ways, but the events cannot occur simultaneously, then the number of ways by which one event (E or F) can occur is $m + n$.

NOTE: More generally, if the events E_1, E_2, ……, E_n can occur in n_1, n_2, …, n_n ways respectively and no two of them can occur simultaneously, then the number of ways by which one event (E_1, E_2, …… or E_n) can occur is $n_1 + n_2 + \cdots + n_n$ ways.

Rule 2: Multiplication Rule (Principle of Sequential Counting)

If an event E can occur in m ways and, independent of this event, an another event F can occur in n ways, then the number of ways by which the events (E and F) can occur in the given order is mn.

NOTE: More generally, if the events E_1, E_2, ……, E_n can occur in n_1, n_2, …, n_n ways respectively and then the number of ways by which the events (E_1, E_2, …… and E_n) can occur in the order indecated is $n_1 n_2 \cdots n_n$ ways.

Worked out Problems:

Problem 25.1. Suppose you can travel from a place A to a place B by 3 buses, from place B to place C by 4 buses, from place C to place D by 2 buses and from place D to place E by 3 buses. In how many ways can you travel from A to E?

Solution:

The bus from A to B can be selected in 3 ways.
The bus from B to C can be selected in 4 ways.
The bus from C to D can be selected in 2 ways.
The bus from D to E can be selected in 3 ways.
So, by the multiplication rule of counting (i,e., Principle of sequential counting), one can travel from A to E in $3 \times 4 \times 2 \times 3 = \mathbf{72}$ **ways.**

Problem 25.2. How many 3-digit numbers can be formed with the digits 1,4,7,8 and 9 if the digits are not repeated?

Solution:

Three digit numbers will have unit's, ten's and hundred's place.
Out of 5 given digits any one can take the unit's place.
This can be done in 5 ways.
After filling the unit's place, any of the four remaining digits can take the ten's place.
This can be done in 4 ways.
After filling in ten's place, hundred's place can be filled from any of the three remaining digits.
This can be done in 3 ways.
So, by the multiplication rule of counting (i,e., Principle of sequential counting), the number of 3 digit numbers $= 5 \times 4 \times 3 = \mathbf{60}$

Problem 25.3. Suppose you have five story books and you want to distribute one each to Asha, Akhtar and Jasvinder. In how many ways can you do it?

Solution:

Any one of the five books can be given to Asha and after that any one of the remaining four books can be given to Akhtar. Thereafter, any one of the remaining four books can be given to Jasvinder. So, by the multiplication rule of counting (i,e., Principle of sequential counting), you can distribute the books in $5 \times 4 \times 3$ i.e. 60 ways.

Problem 25.4. Each user on a computer system has a password, which is six to eight characters long, where each character is an uppercase letter or a digit. Each password must contain at least one digit. How many possible passwords are there?

Solution: Let P be the total number of possible passwords, and let P_6, P_7, and P_8 denote the number of possible passwords of length 6, 7, and 8, respectively. By the sum rule, $P = P_6 + P_7 + P_8$.We will now find P_6, P_7, and P_8. Finding P_6 directly is difficult. To find P_6 it is easier to find the number of strings of uppercase letters and digits that are six characters long, including those with no digits, and subtract from this the number of strings with no digits. By the product rule, the number of strings of six characters is 36^6, and the number of strings with no digits is 26^6. Hence,

$$P_6 = 36^6 - 26^6 = 2{,}176{,}782{,}336 - 308{,}915{,}776 = 1{,}867{,}866{,}560.$$

Similarly, we have

$$P_7 = 36^7 - 26^7 = 78{,}364{,}164{,}096 - 8{,}031{,}810{,}176 = 70{,}332{,}353{,}920$$

and

$$P_8 = 36^8 - 26^8 = 2{,}821{,}109{,}907{,}456 - 208{,}827{,}064{,}576 = 2{,}612{,}282{,}842{,}880.$$

Consequently,

$$P = P_6 + P_7 + P_8 = 2{,}684{,}483{,}063{,}360.$$

Problem 25.5. In a version of the computer language BASIC, the name of a variable is a string of one or two alphanumeric characters, where uppercase and lowercase letters are not distinguished. (An alphanumeric character is either one of the 26 English letters or one of the 10 digits.) Moreover, a variable name must begin with a letter and must be different from the five strings of two characters that are reserved for programming use. How many different variable names are there in this version of BASIC?

Solution: Let V equal the number of different variable names in this version of BASIC. Let V_1 be the number of these that are one character long and V_2 be the number of these that are two characters long. Then by the sum rule, $V = V_1 + V_2$. Note that $V_1 = 26$, because a one-character variable name must be a letter. Furthermore, by the product rule there are $26\,.\,36$ strings of length two that begin with a letter and end with an alphanumeric character. However, five of these are excluded, so $V_2 = 26.36 - 5 = 931$. Hence, there are $V = V_1 + V_2 = 26 + 931 = 957$ different names for variables in this version of BASIC.

Problem 25.6. A new company with just two employees, Sanchez and Patel, rents a floor of a building with 12 offices. How many ways are there to assign different offices to these two employees?

Solution: The procedure of assigning offices to these two employees consists of assigning an office to Sanchez, which can be done in 12 ways, then assigning an office to Patel different from the office assigned to Sanchez, which can be done in 11 ways. By the product rule, there are $12 \cdot 11 = 132$ ways to assign offices to these two employees.

Problem 25.7. The chairs of an auditorium are to be labeled with an uppercase English letter followed by a positive integer not exceeding 100. What is the largest number of chairs that can be labeled differently?
Solution: The procedure of labeling a chair consists of two tasks, namely, assigning to the seat one of the 26 uppercase English letters, and then assigning to it one of the 100 possible integers.
The product rule shows that there are $26 \cdot 100 = 2600$ different ways that a chair can be labeled.
Therefore, the largest number of chairs that can be labeled differently is 2600.

Problem 25.8. There are 32 microcomputers in a computer center. Each microcomputer has 24 ports. How many different ports to a microcomputer in the center are there?
Solution: The procedure of choosing a port consists of two tasks, first picking a microcomputer and then picking a port on this microcomputer. Because there are 32 ways to choose the microcomputer and 24 ways to choose the port no matter which microcomputer has been selected, the product rule shows that there are $32 \cdot 24 = 768$ ports.

Problem 25.9. How many different bit strings of length seven are there?
Solution: Each of the seven bits can be chosen in two ways, because each bit is either 0 or 1.
Therefore, the product rule shows there are a total of $2^7 = 128$ different bit strings of length seven.
Problem 25.10. How many different license plates can be made if each plate contains a sequence of three uppercase English letters followed by three digits (and no sequences of letters are prohibited, even if they are obscene)?
Solution: There are 26 choices for each of the three uppercase English letters and ten choices for each of the three digits. Hence, by the product rule there are a total of 26 · 26 · 26 · 10 · 10 · 10 =17,576,000 possible license plates.

25.3. PERMUTATIONS

Definition 25.1.

Any arrangement of a given set of objects taking some or all at a time is called a permutation of the objects.

NOTE: An arrangement of any r objects taken from a set of n objects $(r \leq n)$ in a definite order is called r-permutation or a permutation of the n objects taken r at a time.

Examples 25.1

Suppose five letters, viz., a, b, c, d, e are given. Then the arrangements such as $abcde, bcdea, dcabe,$ .... etc. are examples of permutations of 5 objects taken all at a time. On the other hand, $acde, bcea, dcbe,$ .... etc. are examples of permutation of the 5 objects taken 4 at a time (4-permutation); $cde, bea, cbe,$ .... etc. are examples of permutation of the 5 objects taken 3 at a time (3-permutation); $de, ba, ae,$ .... etc. are examples of permutation of the 5 objects taken 2 at a time (2-permutation); e, a, d, c, b are examples of permutation of the 5 objects taken 1 at a time (1-permutation).

25.3.1. Rules for Finding the Number of Different Types of Permutations

As per our earlier discussion there exist different permutations of objects. People may be interested in counting the number of r-permutation. The rules for finding the number of different types of r-permutations can be derived using basic counting principles as follows.

Rule 1: The number of r-permutations of n distinct objects ($r \leq n$), in which repetition is not allowed, is ${}^{n}P_{r} = P(n,r) = \frac{n!}{(n-r)!}$.

NOTE: The number of permutations of n distinct objects taking all at a time, in which repetition is not allowed, is ${}^{n}P_{n} = P(n,n) = \frac{n!}{(n-n)!} = \frac{n!}{0!} = n!$.

Rule 2: The number of r-permutations of n distinct objects ($r \leq n$), in which repetition is allowed, is n^{r}.

NOTE: The number of permutations of n distinct objects taking all at a time, in which repetition is allowed, is n^{n}.

Rule 3: The number of r-permutations of n objects ($r \leq n$), when objects are not all distinct (*i.e.* repetition is obvious), say, $n_1$are of first kind, n_2 are of second kind, ….., n_k are of k^{th} kind, where $0 < n_i < n, i = 1,2,\ldots,k; \sum_{i=1}^{k} n_i = n,$ is $\frac{{}^{n}P_{r}}{n_1!\ n_2!\ldots.n_k!}$.

Rule 4: The number of r-permutations of n distinct objects ($r \leq n$), in which repetition is not allowed, but k particular objects ($k \leq r \leq n$) are

(i) always included, is ${}^{r}P_{k} \times {}^{n-k}P_{r-k}$

(ii) never included, is ${}^{n-k}P_{r}$.

Rule 5: The number of permutations of n distinct objects taking all at a time, in which repetition is not allowed, but k particular objects ($k \leq n$)

(i) always occur together, is $\{k! \times (n-k+1)!\}$

(ii) never occur together, is $[n! - \{k! \times (n-k+1)!\}]$.

Rule 6: The number of permutations of n distinct objects taking not more than r at a time, in which repetition is not allowed, is $\frac{n(n^{r}-1)}{(n-1)}$.

Remarks:

Arrangements of objects around a circle are known as ***Circular Permutations***. The fundamental difference between linear and that of circular permutation is that in the former, there are always two separate ends but in circular permutations we cannot distinguish the two ends. For this, in linear permutations, arrangements depend on the **absolute position** while in the case of circular permutations we shall be concerned with **relative positions** of the things. Thus, in case circular permutation, one object from n distinct objects is fixed in particular place of the circle and then the remaining objects are arranged around the circle which includes both clockwise and anti-clockwise arrangements in the same way of linear arrangements. Hence, the number of r-permutation (circular) is same as the number of $(r-1)$-permutation (linear).

∴ The number of circular permutations of n different things taken all at a time, when clockwise and anti-clockwise arrangements are distinguishable, is $(n-1)!$; whereas when clockwise and anti-clockwise arrangements are indistinguishable, the number of such permutations is $\frac{(n-1)!}{2}$.

25.3.2. Some Important Results Related to Permutations

The important results those can be derived using the concept of permutation are as follows.

Result 1: The number of integers between 1 and n which are divisible by $k(0 < k \leq n)$ is $\left[\frac{\mathrm{n}}{\mathrm{k}}\right]$, where [x] denotes the greatest integer not exceeding x.

Result 2: The number of distinct divisors (or factors) of a natural number $n(n > 1)$ is $(k_1 + 1)(k_2 + 1)(k_3 + 1) \ldots.(k_r + 1)$ if the given natural number n can be expressed as follows

$$n = {p_1}^{k_1} \times {p_2}^{k_2} \times {p_3}^{k_3} \times \ldots. \times {p_r}^{k_r}$$

where $p_1, p_2, p_3 \ldots, p_r$ are distinct prime factors of n and $k_1, k_2, k_3 \ldots, k_r$ are positive integers.

Result 3: ${}^{n}P_r = {}^{n-1}P_r + r.{}^{n-1}P_{r-1} = n.{}^{n-1}P_{r-1}$

<u>Worked out Problems:</u>

Problem 25.11. If you have 6 New Year greeting cards and you want to send them to 4 of your friends, in how many ways can this be done?

Solution:

We have to find number of permutations of 4 objects out of 6 objects.

This number is ${}^{6}P_4 = \frac{6!}{(6-4)!} = \frac{6!}{2!} = 3 \times 4 \times 5 \times 6 = 360.$

Therefore, cards can be sent in 360 ways.

Problem 25.12. Suppose you want to arrange your English, Hindi, Mathematics, History, Geography and Science books on a shelf. In how many ways can you do it?

Solution:

Here, 6 books are to be arranged.

∴ The number of possible arrangements = The number of permutations of 6 objects = 6! = 6.5.4.3.2.1 = $\mathbf{720.}$

Problem 25.13. In how many ways can 6 people be seated at a round table?

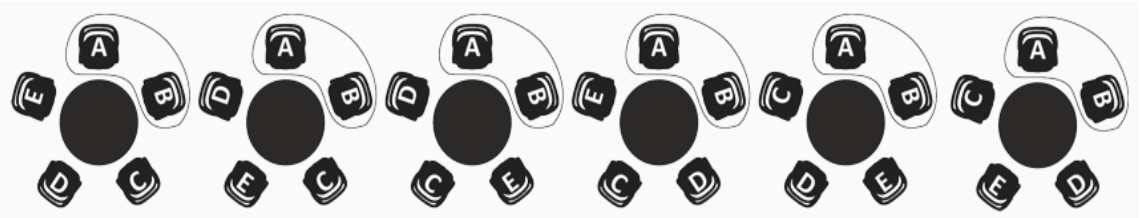
Solution: The number of ways will be (6 – 1)!, i.e., 120.

Problem 25.14. Find the number of ways in which 5 people A,B,C,D,E can be seated at a round table, such that

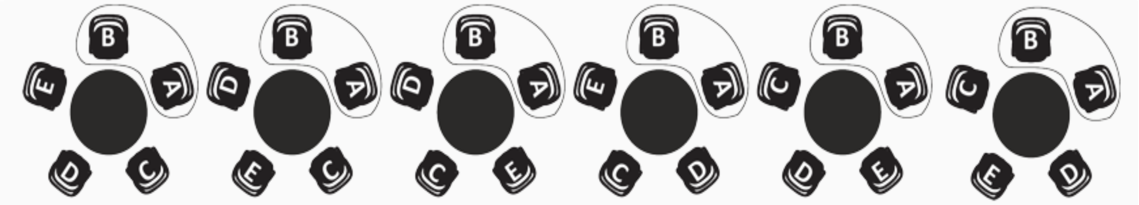
(i) A and B must always sit together.
(ii) C and D must not sit together.

Solution:

(i) If we wish to seat A and B together in all arrangements, we can consider these two as one unit, along with 3 others. So effectively we've to arrange 4 people in a circle, the number of ways being (4 – 1)! or 6. Let me show you the arrangements:



But in each of these arrangements, A and B can themselves interchange places in 2 ways. Here's what I'm talking about:

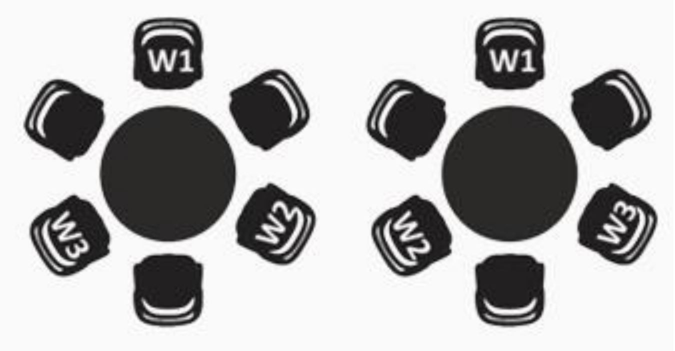


Therefore, the total number of ways will be 6 x 2 = 12.

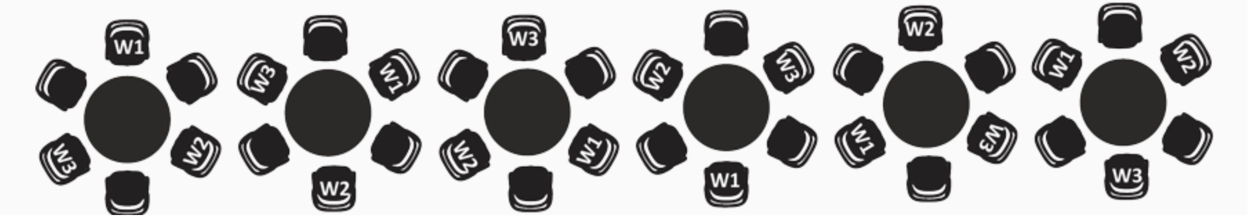
(ii) The number of ways in this case would be obtained by removing all those cases (from the total possible) in which C & D are together. The total number of ways will be (5 – 1)! or 24. Similar to (i) above, the number of cases in which C & D are seated together, will be 12. Therefore the required number of ways will be 24 – 12 = 12.

Problem 25.15. In how many ways can 3 men and 3 ladies be seated at around table such that no two men are seated together?

Solution: Since we don't want the men to be seated together, the only way to do this is to make the men and women sit alternately. We'll first seat the 3 women, on alternate seats, which can be done in (3 – 1)! or 2 ways, as shown below. (We're ignoring the other 3 seats for now)



Note that the following 6 arrangements are equivalent:



That is, if each of the women is shifted by a seat in any direction, the seating arrangement remains exactly the same. That is why we have only 2 arrangements, as shown in the previous figure.

Now that we've done this, the 3 men can be seated in the remaining seats in 3! or 6 ways. Note that we haven't used the formula for circular arrangements now. This is so because, after the women are seated, shifting the each of the men by 2 seats, will give a different arrangement. After fixing the position of the women (same as 'numbering' the seats), the arrangement on the remaining seats is equivalent to a linear arrangement.

Therefore the total number of ways in this case will be 2! X 3! = 12.

<u>Exercise:</u>

E 25.1. How many functions are there from a set with m elements to a set with n elements?

E 25.2. How many one-to-one functions are there from a set with m elements to one with n elements?

E 25.3. The *North American numbering plan* (NANP) specifies the format of telephone numbers in the U.S., Canada, and many other parts of North America. A telephone number in this plan consists of 10 digits, which are split into a three-digit area code, a three-digit office 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 denote a digit that can take any of the values 0 through 9, let N denote a digit that can take any of the values 2 through 9, and let Y denote a digit that must be a 0 or a 1. Two numbering plans, which will be called the old plan, and the new plan, will be discussed. (The old plan, in use in the 1960s, has been replaced by the new plan, but the recent rapid growth in demand for new numbers for mobile phones and devices will eventually 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 formats of the area code, office code, and station code are NYX, NNX, and XXXX, respectively, so that telephone numbers had the form NYX-NNX-XXXX. In the new plan, the formats of these codes are NXX, NXX, and XXXX, respectively, so that telephone numbers have the form NXX-NXX-XXXX. How many different North American telephone numbers are possible under the old plan and under the new plan?

E 25.4. What is the value of k after the following code, where $n_1, n_2, \ldots, n_m$ are positive integers, has been executed?

```
k := 0
for $i_1 := 1$ to $n_1$
    for $i_2 := 1$ to $n_2$
        .
        .
        .
        for $i_m := 1$ to $n_m$
            $k := k + 1$
```

E 25.5. Use the product rule to show that the number of different subsets of a finite set S is $2^{|S|}$.

LECTURE 26: COMBINATIONS AND BINOMIAL COEFFICIENTS

26.1. COMBINATIONS

Definition 26.1.

Any unordered selection of a given set of objects taking some or all at a time is called a combination of the objects.

Thus, by the combination of n different objects taken r $(r \leq n)$ at a time, we mean all possible grouping of r different objects taken from the n objects without giving any importance to the order of arrangement of objects forming the group. The number of combinations of n objects taken r at a time is denoted by ${}^{n}C_{r}$ or $C(n,r)$ or $\binom{n}{r}$

NOTE: An unordered selection of any r objects taken from a set of n objects $(r \leq n)$ is called r-combination or a combination of the n objects taken r at a time.

Example 26.1.

Suppose five letters, viz., a, b, c, d, e are given. Then the selection of all objects a, b, c, d, e (in any order) is a combination of 5 objects taken all at a time. On the other hand, the selection of four objects a, c, d, e (in any order), the selection of four objects b, c, d, e (in any order), the selection of four objects a, b, c, e (in any order), ….. etc. are examples of combinations of the 5 objects taken 4 at a time (4-combination); the selection of three objects a, c, d (in any order), the selection of three objects b, c, e (in any order), the selection of three objects a, b, e (in any order), ….. etc. are examples of combinations of the 5 objects taken 3 at a time (3-combination); the selection of two objects a, c (in any order), the selection of two objects b, c (in any order), the selection of two objects a, e (in any order), ….. etc. are examples of combinations of the 5 objects taken 2 at a time (2-combination); the selection of one object a, the selection of one object b, the selection of one object e, ….. etc. are examples of combinations of the 5 objects taken 1 at a time (1-combination).

26.1.1. Rules for Finding the Number of Different Types of Combinations

As per our earlier discussion (see *1.4.2*), there exists different combinations of objects. People may be interested in counting the number of r-combination. The rules for finding the number of different types of r-combinations can be derived using basic counting principles as follows.

Rule 1: The number of r-combinations of n distinct objects ($r \leq n$), in which repetition is not allowed, is ${}^{n}C_{r} = C(n,r) = \binom{n}{r} = \frac{n!}{r!(n-r)!}$.

NOTE: The number of permutations of n distinct objects taking all at a time, in which repetition is not allowed, is ${}^{n}C_{n} = C(n,n) = \binom{n}{n} = \frac{n!}{n!(n-n)!} = \frac{n!}{n!0!} = 1$.

Rule 2: The number of combinations of n distinct objects, taking at least one at a time is $2^n - 1$.

Rule 3: The number of combination of n objects, when objects are not all distinct, say, $n_1$are of first kind, n_2 are of second kind, ….., n_k are of k^{th} kind, where $0 < n_i < n, i = 1,2, \dots, k$; $\sum_{i=1}^{k} n_i = n$, is $[(n_1 + 1)(n_2 + 1) \dots (n_k + 1) - 1]$.

Rule 4: The number of combination of n objects taking some or all, when objects are not all distinct, say, $n_1$are of first kind, n_2 are of second kind, ….., n_k are of k^{th} kind and m are all distinct, where $0 < n_i, m < n, i = 1,2, \dots, k$; $\sum_{i=1}^{k} n_i + m = n$, is $[(n_1 + 1)(n_2 + 1) \dots (n_k + 1)2^m - 1]$.

Rule 5: The number of r-combination of n distinct objects ($r \leq n$.), in each of which k particular objects ($k \leq r \leq n$) are

(i) always included, is ${}^{n-k}C_{r-k}$
(ii) never included, is ${}^{n-k}C_r$.

Rule 6: The number of r-combination of n distinct objects, in which repetition is allowed, is ${}^{n+r-1}C_r$.

<u>Worked out Problems:</u>

Problem 26.1. Out of 7 consonants and 4 vowels, how many words of 3 consonants and 2 vowels can be formed?

Solution:

Number of ways of selecting 3 consonants from 7 = ${}^{7}C_3$.

Number of ways of selecting 2 vowels from 4 = ${}^{4}C_2$.

∴ Number of ways of selecting 3 consonants from 7 and 2 vowels from 4 = ${}^{7}C_3 \times {}^{4}C_2$

$$= \left(\frac{7\times6\times5}{3\times2\times1}\right) \times \left(\frac{4\times3}{2\times1}\right)$$

$$= 210$$

So, we can have 210 groups where each group contains total 5 letters (3 consonants and 2 vowels).

Now, the number of ways of arranging 5 letters among themselves $= 5! = 120$.

Hence, required number of ways $= 210 \times 120 = 25200$.

Problem 26.2. In a group of 6 boys and 4 girls, four children are to be selected. In how many different ways can they be selected such that at least one boy should be there?

Solution:

In a group of 6 boys and 4 girls, four children are to be selected such that at least one boy should be there.

Hence we have 4 options as given below:

<u>Option 1: All 4 members are boys</u>

The number of ways by which 4 boys can be selected out of 6 boys is ${}^6C_4$.

Option 2: 3 members are boys and 1 member is girl

The number of ways by which 3 boys can be selected out of 6 boys is ${}^6C_3$.
The number of ways by which 1 girl can be selected out of 4 girls is ${}^4C_1$.
So, by the multiplication rule of counting (i,e., Principle of sequential counting), the number of ways by which 3 boys and 1 girl can be selected for the group is ${}^6C_3 \times {}^4C_1$.

Option 3: 2 members are boys and 2 members are girls

The number of ways by which 2 boys can be selected out of 6 boys is ${}^6C_2$.
The number of ways by which 2 girls can be selected out of 4 girls is ${}^4C_2$.
So, by the multiplication rule of counting (i,e., Principle of sequential counting), the number of ways by which 2 boys and 2 girls can be selected for the group is ${}^6C_2 \times {}^4C_2$.

Option 4: 1 member is boy and 3 members are girls

The number of ways by which 1 boy can be selected out of 6 boys is ${}^6C_1$.
The number of ways by which 3 girls can be selected out of 4 girls is ${}^4C_3$.
So, by the multiplication rule of counting (i,e., Principle of sequential counting), the number of ways by which 1 boy and 3 girls can be selected for the group is ${}^6C_1 \times {}^4C_3$.

Since, the group of 4 members can be selected by considering only one of the 4 options mentioned above, by the addition rule of counting (i,e., Principle of disjunctive counting), the number of ways = ${}^6C_4 + {}^6C_3 \times {}^4C_1 + {}^6C_2 \times {}^4C_2 + {}^6C_1 \times {}^4C_3$

$$= \left(\frac{6 \times 5}{2 \times 1}\right) + \left(\frac{6 \times 5 \times 4}{3 \times 2 \times 1}\right) \times 4 + \left(\frac{6 \times 5}{2 \times 1}\right) \times \left(\frac{4 \times 3}{2 \times 1}\right) + (6 \times 4)$$
$$= 15 + 80 + 90 + 24$$
$$= 209$$

Problem 26.3. Example 6.1.21. In how many ways can you allocate 3 identical passes to 10 students so that each student receives at most one?

Solution: Here, we have only 3 passes to distribute among 10 students. Since, passes are identical and a student can receive at most one, we can classify the 10 students in 2 categories: one who will receive only one pass and other who will receive no pass. The first group consists of three students and the second group consists of 7 students. Thus, the required number of ways by which the passes can be allocated is same as the number of ways by which a group of 3 students can be formed out of 10 students (or same as the number of ways by which a group of 7 students can be formed out of 10 students).
∴ the required number of ways $= {}^{10}C_3$ (or ${}^{10}C_7$). [NOTE: ${}^nC_r = {}^nC_{n-r}$]

26.2. BINOMIAL THEOREM

Statement: For any real numbers x, y and any integer $n \geq 0$,

$$(x+y)^n = \sum_{r=0}^{n} {}^nC_r x^r y^{n-r}$$

NOTE: Here, ${}^{n}C_{r}$ is known as binomial coefficient. The binomial coefficient is also written as $C(n,r)$ or $\binom{n}{r}$. The symbol ${}^{n}C_{r}$ has two meaning

(i) ***Combinatorial meaning*** where it represents the number of ways of choosing r objects from given n distinct objects

(ii) ***Algebraic meaning*** where it is expressed as ${}^{n}C_{r} = \frac{n!}{r!(n-r)!}$

Hence, all theorems and identities about Binomial coefficients and factorials can be given two kinds of proofs a combinatorial proof and an algebraic proof.

Corollary: When n is a positive integer,

(i) $(1+x)^{n} = \sum_{r=0}^{n} {}^{n}C_{r}x^{r}$

(ii) $(1+x)^{-n} = \sum_{r=0}^{n}(-1)^{r\,n+r-1}C_{r}x^{r}$

(iii) $(1-x)^{-n} = \sum_{r=0}^{n} {}^{n+r-1}C_{r}x^{r}$

Some Important Identities:

(i) ${}^{n}C_{0} = {}^{n}C_{n}$=1

(ii) ${}^{n}C_{1} = {}^{n}C_{n-1} = n$

(iii) ${}^{n}C_{r} = \frac{n}{r}{}^{n-1}C_{r-1}$

(iv) ${}^{n}C_{0} + {}^{n}C_{1} + {}^{n}C_{2} + \cdots + {}^{n}C_{n} = 2^{n}$

(v) ${}^{n}C_{r} = {}^{n}C_{n-r}, for\ r \le n$

(vi) ${}^{n}C_{r} = {}^{n}C_{p} \Rightarrow r = p\ or\ r + p = n$

(vii) ${}^{n}C_{r} + {}^{n}C_{r-1} = {}^{n+1}C_{r}$

(viii) Newton's Identity: ${}^{n}C_{r}\,{}^{r}C_{k} = {}^{n}C_{k}\,{}^{n-k}C_{r-k}\ , for\ intgers\ 0 \le k \le r \le n$

(ix) Pascal Identity:

(x) Pascal's identity was probably first derived by Blaise Pascal, a 19th century French mathematician, whom the theorem is named after.

<u>*Statement:*</u> For any positive integers r and n, ${}^{n+1}C_{r} = {}^{n}C_{r} + {}^{n}C_{r-1}$

<u>*NOTE:*</u> (*Combinatorial Meaning*)

(xi) Vandermonde's Identity: ${}^{n+m}C_{r} = \sum_{k=0}^{r} {}^{m}C_{r-k}\,{}^{n}C_{k}\ , for\ integers\ 0 \le r \le m, n$

26.3. MULTINOMIAL THEOREM

The *multinomial theorem* describes how to expand the power of a sum of more than two terms. It is a generalization of the binomial theorem to polynomials with any number of terms. It expresses a power $(x_1 + x_2 + \cdots + x_k)^n$ as a weighted sum of monomials of the form ${x_1}^{b_1}{x_2}^{b_2} \ldots {x_k}^{b_k}$ where the weights are given by generalizations of binomial coefficients called *multinomial coefficients*.

Definition 26.2. (*Multinomial Coefficient*)

For non-negative integers $b_1, b_2, \ldots, b_k$ such that $\sum_{i=1}^{k} b_i = n$, the *multinomial coefficient* is

$$\binom{n}{b_1, b_2, \ldots, b_k} = \frac{n!}{b_1!\,b_2!\ldots b_k!}$$

NOTE: For $k = 2$, the *multinomial coefficient* reduces to binomial coefficient:

$$\binom{n}{b_1, b_2} = \frac{n!}{b_1!b_2!}, \text{ where } b_1 + b_2 = n$$

$$= \frac{n!}{b_1!\,(n - b_1)!}$$

$$= \binom{n}{b_1}$$

For a positive integer k and a non-negative integer n,

$$(x_1 + x_2 + \cdots + x_k)^n = \sum_{\substack{0 \leq\, b_1, b_2, \ldots, b_k \,\leq\, n \\ and\ b_1 + b_2 + \cdots + b_{k=n}}} \binom{n}{b_1, b_2, \ldots, b_k} \prod_{j=1}^{k} {x_j}^{b_j}$$

Theorem 26.1. (*Multinomial Theorem*)
Statement: For a positive integer k and a non-negative integer n,

$$(x_1 + x_2 + \cdots + x_k)^n = \sum_{\substack{0 \leq\, b_1, b_2, \ldots, b_k \,\leq\, n \\ and\ b_1 + b_2 + \cdots + b_{k=n}}} \binom{n}{b_1, b_2, \ldots, b_k} \prod_{j=1}^{k} {x_j}^{b_j}$$

LECTURE 27: ADVANCED COUNTING TECHNIQUES

27.1. PRINCIPLE OF INCLUSION AND EXCLUSION

The *Principle of Inclusion and Exclusion* (PIE) is a counting technique that computes the number of elements satisfying at least one of several properties while guaranteeing that elements satisfying more than one property are not counted twice.

An underlying idea behind PIE is that summing the number of elements that satisfy at least one of two categories and subtracting the overlap prevents double counting. For instance, the number of people that have at least one cat or at least one dog can be found by taking the number of people who own a cat, adding the number of people that have a dog, then subtracting the number of people who have both.

PIE is particularly useful in combinatorics and probability problem solving when it is necessary to devise a counting method that ensures an object is not counted twice.

Statement (for Two Sets):

In the case of objects being separated into two (possibly disjoint) sets A and B, the principle of inclusion and exclusion states

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

where $|S|$ denotes the cardinality, or number of elements, of set S in set notation.

Statement (for Three Sets):

In the case of objects being separated into three (possibly disjoint) sets A, B and C, the principle of inclusion and exclusion states

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

where $|S|$ denotes the cardinality, or number of elements, of set S in set notation.

Worked out Problems:

Problem 27.1. How many integers from 1 to 100 are multiples of 2 or 3?

Solution: Let, A and B are the sets of integers from 1 to 100 that are respectively multiples of 2 and 3.

$\therefore |A| = \left\lceil \frac{100}{2} \right\rceil = 50$ and $|B| = \left\lceil \frac{100}{3} \right\rceil = 33$, where $\lceil . \rceil$ denotes the ceiling function.

Now, $A \cap B$ be the set of integers from 1 to 100 that are multiples of both 2 and 3, and hence are multiples of 6, implying $|A \cap B| = \left\lceil \frac{100}{6} \right\rceil = 16$.

Since, $A \cap B$ is the set of integers from 1 to 100 that are multiples of 2 or 3, by principle of inclusion and exclusion,

$$|A \cup B| = |A| + |B| - |A \cap B| = 50 + 33 - 16 = 67$$

So, the total number of integers from 1 to 100 which are multiples of 2 or 3 is 67.

Problem 27.2. There are exactly three types of students in a school: the geeks, the wannabees, and the athletes. Each student is classified into at least one of these categories. Out of 1000 school students, 310 are geeks, 650 are wannabees, 440 are athletes, 170 are both geeks and wannabees, 150 are both geeks and athletes and 180 are both wannabees and athletes. What is the total number of students who fit into all 3 categories?

Solution: Let, A, B and C denote the set for geeks, wannabees, and athletes, respectively. Thus, $\therefore |A| = 310$, $|B| = 650$ and $|C| = 440$.
Since, the total number of students in the school is 1000 and each student is classified into at least one of these categories, we have

$$|A \cup B \cup C| = 1000$$

Now, $A \cap B$, $B \cap C$ and $C \cap A$ are the sets of school students who are both geeks and wannabees, both geeks and athletes, and both wannabees and athletes, respectively.
$\therefore |A \cap B| = 170$, $|B \cap C| = 150$ and $|C \cap A| = 180$.
Since, $A \cap B \cap C$ denotes the set of school students who fit into all 3 categories; by the principle of inclusion and exclusion, we have

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |B \cap C| - |C \cap A| + |A \cap B \cap C|$$
$$\Rightarrow 1000 = 310 + 650 + 440 - 170 - 150 - 180 + |A \cap B \cap C|$$
$$\Rightarrow |A \cap B \cap C| = 1000 - 310 - 650 - 440 + 170 + 150 + 180$$
$$\Rightarrow |A \cap B \cap C| = 100$$

So, the total number of students who fit into all 3 categories is 100.

27.2. DERANGEMENT

A *derangement* is an arrangement (i.e., a permutation) of some number of objects into positions such that no object goes to its specified position.

Theorem27.1.

Statement: Let $D(n)$ or D_n be the number of derangements for n different objects, then

$$D_n = n! \sum_{r=0}^{n} \frac{(-1)^r}{r!}$$

Proof: Let there be n distinct objects with their n distinct respective positions. So, The number of all possible arrangements is $n!$.

Now, let N is the number of ways of arranging the n objects in such a way that at least one object goes to its right position.

$\therefore$ The number of derangements $(D_n) = n! - N$.

Let, A_i be the set of permutations in which the i^{th} object goes into its right position. Then, $A_i \cap A_j$ is the set of permutations in which the i^{th} object and j^{th} object go into their right positions; $A_i \cap A_j \cap A_k$ is the set of permutations in which the i^{th} object, j^{th} object and k^{th} object go into their right positions and so on.

Thus, $|A_i| = (n-1)!$, $|A_i \cap A_j| = (n-2)!$, $|A_i \cap A_j \cap A_k| = (n-3)!$ and so on.

Now, $(\bigcup_{i=1}^{n} A_i)$ is the set of permutations in which at least one object goes into its right position.

$$\therefore N = \left|\bigcup_{i=1}^{n} A_i\right|$$

By the principle of inclusion and exclusion

$$N = \sum_{i=1}^{n}|A_i| - \sum_{\substack{1\leq i,j\leq n\\ i<j}}|A_i \cap A_j| + \sum_{\substack{1\leq i,j,k\leq n\\ i<j<k}}|A_i \cap A_j \cap A_k| - \cdots + (-1)^{n+1}|A_1 \cap A_2 \cap \dots \cap A_n|$$

$$= {}^{n}C_1(n-1)! - {}^{n}C_2(n-2)! + {}^{n}C_3(n-3)! - \cdots + (-1)^{n+1}\,{}^{n}C_n(n-n)!$$

$$= \sum_{r=1}^{n}(-1)^{r+1}\,{}^{n}C_r(n-r)!$$

$$= \sum_{r=1}^{n}(-1)^{r+1}\frac{n!}{r!}$$

Therefore, the number of derangements $(D_n) = n! - N$

$$= n! - \sum_{r=1}^{n}(-1)^{r+1}\frac{n!}{r!}$$

$$= n!\sum_{r=0}^{n}\frac{(-1)^r}{r!}$$

NOTE: The above result can also be rewritten as follows:

$$D_n = n!\sum_{r=0}^{n}\frac{(-1)^r}{r!}$$

$$= n!\left[\sum_{r=0}^{n-1}\frac{(-1)^r}{r!} + \frac{(-1)^n}{n!}\right]$$

$$= n(n-1)!\sum_{r=0}^{n-1}\frac{(-1)^r}{r!} + (-1)^n$$

$$= nD_{n-1} + (-1)^n$$

Problem 27.3. Mickey the mailman is very lazy. He has received 10 parcels to 10 different people. However, because he is lazy, he doesn't bother reading the address and delivers them off randomly. In how many ways can Mickey deliver the parcels such that no one gets the right parcel?

Solution: This is just a derangement problem (sending the set $\{1,2,3,\dots,10\}$ to another set such that none of the original elements are in the same place).

∴ The required number of ways $= the\ number\ of\ derangements\ of\ 10\ objects\ (D_{10})$

$$= 10! \sum_{r=0}^{10} \frac{(-1)^r}{r!}$$

$$= 10! \left(\frac{1}{0!} - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \cdots + \frac{1}{10!}\right)$$

$$= 1334961$$

Therefore, there are a total of 1334961 ways to deliver the parcels such that no one gets the right parcel.

Problem 27.4. There are 4 men: A, B, C and D. Each has a son. The four sons are asked to enter a dark room. Then A, B, C and D enter the dark room, and each of them walks out with just one child. If none of them comes out with his own son, in how many ways can this happen?

Solution: This is just a derangement problem (sending the set $\{1,2,3,4\}$ to another set such that none of the original elements are in the same place).

∴ The required number of ways $= the\ number\ of\ derangements\ of\ 4\ objects\ (D_4)$

$$= 4! \sum_{r=0}^{4} \frac{(-1)^r}{r!}$$

$$= 4! \left(\frac{1}{0!} - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!}\right)$$

$$= 24 \left(1 - 1 + \frac{1}{2} - \frac{1}{6} + \frac{1}{24}\right)$$

$$= 24 \left(\frac{1}{2} - \frac{1}{6} + \frac{1}{24}\right)$$

$$= 12 - 4 + 1$$

$$= 9$$

Therefore, there are a total of 9 ways that none of them comes out with his own son.

27.2. PIGEONHOLE PRINCIPLE

The *Pigeonhole Principle* (also known as the *Dirichlet box principle*, *Dirichlet principle* or *box principle*) is one of the simplest but most useful concept in mathematics, discovered all the way back in the 1800s.

Suppose there are fifteen pigeonholes and sixteen pigeons. A storm comes along, and all of the pigeons take shelter inside the pigeonholes. They could be arranged any number of ways. For all we know, all sixteen pigeons could be inside one hole, and the rest of the holes could be empty. What we know for sure, no matter what, is that there is at least one hole that contains more than one pigeon. The principle works no matter what the particular number of pigeons and pigeonholes, of course. As long as there are $(N-1)$ number of pigeonholes, and N number of pigeons, we know there will always be at least two pigeons in one hole.

Theorem 27.1 (*Pigeonhole Principle*)

Statement: If $n+1$ or more pigeons are placed in n holes, then one hole must contain two or more pigeons.
Proof: We 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.

Corollary 27.1.1. 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 contains $k+1$ or more elements and the codomain contains only k elements, the pigeonhole principle tells us that one of these boxes contains two or more elements x of the domain. This means that f cannot be one-to-one.

Theorem 27.2 (*Extension of Pigeonhole Principle*)
Statement: If k objects are placed in n boxes ($n < k$), then at least one box must contain at least $\left\lceil \frac{n}{k} \right\rceil$ objects, where $\lceil . \rceil$ denotes the ceiling function.

There are many examples which use pigeonhole principle. Few of the examples are given below:

Example 27.1 (*Golf*)

Let us suppose that there are 8 balls and 7 holes. If balls are to be put in different holes, then at least one hole must has more than one ball.

Example 27.2 (*Handshake*)

If a number of people does handshake with one another, then according to pigeonhole principle, there must exist two people who shake hands with same people.

Example 27.3 (*Birthday*)

Let us consider that n people are chosen at random from a group of people. Then, in order to find the probability of having same birthday, pigeonhole principle is applied. It says that at least two people will have same birthday.

Example 27.4 (*Marble picking*)

Consider that we have a mixture of different color marbles in a jar. In order to find at least how many marbles will be picked before two same color marbles are guaranteed. It can be calculated using pigeonhole principle assuming one pigeonhole per color will be assumed.

Worked out Problems:

Problem 27.5. How many students must be in a class to guarantee that at least two students receive the same score on the final exam, if the exam is graded on a scale from 0 to 100 points?

Solution: There are 101 possible scores on the final. The pigeonhole principle shows that among any 102 students there must be at least 2 students with the same score.

Problem 27.6. Show that for every integer n there is a multiple of n that has only 0s and 1s in its decimal expansion.

Solution: Let n be a positive integer. Consider the $n + 1$ integers $1, 11, 111, \ldots, 11 \ldots 1$ (where the last integer in this list is the integer with $n + 1$ 1s in its decimal expansion). Note that there are n possible remainders when an integer is divided by n. Because there are $n + 1$ integers in this list, by the pigeonhole principle there must be two with the same remainder when divided by n. The larger of these integers less the smaller one is a multiple of n, which has a decimal expansion consisting entirely of 0s and 1s.

Problem 27.7. A bag contains 10 red marbles, 10 white marbles, and 10 blue marbles. What is the minimum no. of marbles you have to choose randomly from the bag to ensure that we get 4 marbles of same color?

Solution: Apply pigeonhole principle.

No. of colors (pigeonholes) $n = 3$

No. of marbles (pigeons) $K + 1 = 4$

Therefore the minimum no. of marbles required $= Kn + 1$

By simplifying we get Kn+1 = 10.

Verification: ceil[Average] is [Kn+1/n] = 4

[Kn+1/3] = 4

Kn+1 = 10

i.e., 3 red + 3 white + 3 blue + 1(red or white or blue) = 10

Problem 27.8. If a Martian has an infinite number of red, blue, yellow, and black socks in a drawer, how many socks must the Martian pull out of the drawer to guarantee he has a pair?

Solution: The Martian must pull 5 socks out of the drawer to guarantee he has a pair. In this case the pigeons are the socks he pulls out and the holes are the colors. Thus, if he pulls out 5 socks, the Pigeonhole Principle states that some two of them have the same color. Also, note that it is possible to pull out 4 socks without obtaining a pair.

Problem 27.9. Suppose S is a set of $(\mathrm{n} + 1)$ integers. Prove that there exist distinct $\mathrm{a}, \mathrm{b} \in \mathrm{S}$ such that $\mathrm{a} - \mathrm{b}$ is a multiple of n.

Solution: Consider the residues of the elements of S, modulo n. By the Pigeonhole Principle, there exist distinct $a, b \in S$ such that $a \equiv b(mod\, n)$, as desired.

Problem 27.10. Show that in any group of n people, there are two who have an identical number of friends within the group.

Solution: The maximum number of friends that one person in the group can have is $n - 1$, and the minimum is 0. If all of the members have at least one friend, then each individual can have somewhere between 1 to $n - 1$ friends; as there are n individuals, by pigeonhole there must be at least two with the same number of friends. If one individual has no friends, then the remaining friends must have from 1 to $n - 2$ friends for the remaining friends not to also have no friends. By pigeonhole again, this leaves at least 1 other person with 0 friends.

Problem 27.11. Six distinct positive integers are randomly chosen between 1 and 2006, inclusive. What is the probability that some pair of these integers has a difference that is a multiple of 5?

Solution:

Let, a and b are any two integer such that $|a-b|$ is multiple of 5, i.e., $|a-b|=5k$, for some integer $k, 0 \leq k$. For the difference to be a multiple of 5, the two integers must have the same remainder when divided by 5. Since there are 5 possible remainders, viz., 0,1,2,3,4, by the pigeonhole principle, at least two of the integers must share the same remainder. Thus, the answer is 1 (E).

Exercise:

E.27.1. In a higher secondary examination 80% of the examinees have passed in English and 85% in Mathematics, while 75% passed in both English and Mathematics. If 45 candidates failed in both the subjects, find the total number of candidates.

E.27.2. A local grocery store in the outback newly opened. They were offering 1 free bottle Marmite to every 11^{th} customer, and 1 free pound of kangaroo meat for every 13^{th} customer. If there were 1000 customers that visited them on opening day, how many customers walked away with free goodies?

E.27.3. How many positive integers less than or equal to 60 are divisible by 3, 4 or 5?

E.27.4. There were 5 family members, who each brought their own gift, and then the 5 gifts were exchanged within the family. How many ways could they exchange such that none of them got their own present?

E.27.5. At the Winter Sochi Olympics Press Conference, there are 200 foreign journalists. Out of them, 175 people can speak German, 150 people can speak French, 180 people can speak English, 160 people can speak Japanese. What is the minimum number of foreigners that can speak all the four languages?

E.27.6. Seven line segments, with lengths no greater than 10 inches, and no shorter than 1 inch, are given. Show that one can choose three of them to represent the sides of a triangle.

(*Manhattan Mathematical Olympiad 2004*)

E.27.7. Prove that having 100 whole numbers, one can choose 15 of them so that the difference of any two is divisible by 7.

(*Manhattan Mathematical Olympiad 2005*)

E.27.8. Prove that from any set of one hundred whole numbers, one can choose either one number which is divisible by 100, or several numbers whose sum is divisible by 100.

(*Manhattan Mathematical Olympiad 2003*)

E.27.9. Prove that among any ten points located on a circle with diameter 5, there exist at least two at a distance less than 2 from each other.

(*Japan 1997*)

E.27.10. Every point in a plane is either red, green, or blue. Prove that there exists a rectangle in the plane such that all of its vertices are the same color.

(*USAMTS Year 18 - Round 1 - Problem 4*)

E.27.11. There are 51 senators in a senate. The senate needs to be divided into n committees such that each senator is on exactly one committee. Each senator hates exactly three other senators. (If senator A hates senator B, then senator B does 'not' necessarily hate senator A.) Find the smallest n such that it is always possible to arrange the committees so that no senator hates another senator on his or her committee.

LECTURE 28: RECURRENCE RELATION.

28.1. INTRODUCTION

An equation involving several terms of a sequence is called a recurrence relation. We shall think of the integer n as the independent variable, and restrict our attention to real sequence, so that the sequence a_n is considered as a function of the type

$$f: \mathbb{N} \cup \{0\} \longrightarrow \mathbb{R}: n \to a_n$$

A Recurrence relation is then an equation of the type

$$F(n, a_n, a_{n+1}, \dots, a_{n+k}) = 0\,, \quad \text{where } k \in \mathbb{N} \text{ is fixed.}$$

Recurrence relation is useful in counting problems. It defines a sequence by giving the nth value in terms of (n-1) th or (n-2) th value or in terms of other predecessors.

EXAMPLES

(i) $a_{n+1} = 5a_n$ is a Recurrence relation of order 1.

(ii) ${a_{n+1}}^4 + {a_n}^5 = n$ is a Recurrence relation of order 1.

(iii) $a_{n+3} + 5a_{n+2} + 4a_{n+1} + a_n = \cos n$ is a Recurrence relation of order 3.

(iv) $a_{n+2} + 5({a_{n+1}}^2 + a_n)^{\frac{1}{3}} = 0$ is a Recurrence relation of order 2.

28.2. ORDER OF A RECURRENCE RELATION

The order of a recurrence relation is the difference between the greatest and lowest subscripts of the terms of the sequence in the equation.

28.3. LINEAR AND NON-LINEAR RECURRENCE RELATION

A recurrence relation of order k is said to be linear if it is linear in $a_n, a_{n+1}, \dots, a_{n+k}$. Otherwise, the recurrence relation is said to be non-linear.

The recurrence relations in Examples (i) and (iii) are linear, while those in Examples (ii) and (iv) are non-linear.

28.4. FORMULATION OF DIFFERENT COUNTING PROBLEMS IN TERMS OF RECURRENCE RELATION

It is typical to want to derive a Recurrence relation with initial conditions for the number of objects satisfying certain conditions. The main technique involves giving counting argument that gives the number of objects of “size” n in terms of the number of objects of smaller size. This typically involves an analysis of several cases.

Suggestion: When attempting to derive a Recurrence relation with initial conditions, start by working out the first few cases directly. You’ll need these for the initial conditions anyway, and doing this might help you see how to proceed. If you do enough cases, then you can use them later to check your recurrence.

Example 28.1 Fibonacci numbers: Assume you start with one pair of new-born rabbits (one of each gender), and in each subsequent month each pair of rabbits which are more than 1 month old gives birth to a new pair of rabbits, one of each gender. Determine a Recurrence relation with initial conditions for f_n, the number of pairs of rabbits present at the end of n months.

Solution:
The statement tells us that $f_0 = 1$. Also, $f_1 = 1$ because the original pair of rabbits is not yet old enough to breed. At the end of two months, we have our pair from before, plus one new pair. At the end of 3 months, we have the f_2 pairs from before, and f_1 of them are old enough to breed, so we have f_3 = f_2 + f_1 = 3 pairs. Consider what happens at the end of n months. We still have the f_{n-1} pairs from the month before. The number of pairs old enough to breed is the number alive two months ago, or f_{n-2}, so we get f_{n-2} new pairs. Thus, f_n = f_{n-1}+ f_{n-2}, n ≥ 2, and f_0= f_1 = 1. Using the Recurrence relation with initial conditions yields the sequence 1, 1, 2, 3, 5, 8, ... which agrees with our initial counting.

Example 28.2 There are guests in a gathering. Each person shakes hand with everyone else exactly once. Formulate the Recurrence relation representing the number of handshake occurred in the gathering.

Solution:
Let there be n number of guests present in the gathering. $H_n =$ number of handshakes occurred in this gathering. Obviously $H_1 = 0$.
Now, $H_n =$ number of handshakes occurred among the gathering of $(n-1)$ guests $+$ number of handshakes made by the nth guests with the $(n-1)$ number of guests

$$= H_{n-1} + (n-1)$$

Thus the required Recurrence relation is

$$H_n = H_{n-1} + (n-1)$$

Exercise 28.1: Vinod deposits Rs. 15,000 in a savings account at a bank .The bank gives 8.5% interest per annum. Formulate the recurrence relation to compute the amount Vinod will have in his account at the end of nth year.

28.5. LINEAR RECURRENCE RELATION WITH CONSTANT COEFFICIENTS

Recall that a linear recurrence relation with constant coefficients $c_1, c_2 ,\cdots, c_k$ ($c_k \neq 0$) of degree k has the form
a_n= $c_1 a_{n-1}$ + $c_2 a_{n-2}$ +⋯+ $c_k a_{n-k}$ + F(n) (n ≥ k).
It follows from the general recursion theorem that for every string of initial values $a_0, a_1,\cdots, a_{n-k}$ there is exactly one sequence $\{a_n\}$ that satisfies the above recurrence relation and matches the given initial conditions. Consequently, if no initial conditions are imposed, there will always be an infinite set of solutions.
If F(n)=0 then it is called linear homogeneous recurrence relation with constant coefficients.

EXAMPLES

(i) $a_n = 2a_{n-1}$ is a 1st order linear homogeneous recurrence relation with constant coefficients.

(ii) $\mathrm{H_n} = 2\mathrm{H_{n-1}} + 1$ is a 1st order linear non homogeneous recurrence relation with constant coefficients.

(iii) $a_n = 2a_{n-1}a_{n-2} + n^2$ is a 2nd order non linear non homogeneous recurrence relation with constant coefficients.

28.5.1. SOLUTION OF RECURRENCE RELATION

A sequence is called a solution of a recurrence relation if its terms satisfy the recurrence relation. The relation $a_n = a_{n-1} + 2$ expresses a_n in terms of its preceding terms a_{n-1}.So it is a recurrence relation for the sequence $\{a_n\}$.Here we see that the sequence $\{2n - 1\}$ satisfies the recurrence relation. So it is a solution of the recurrence relation.

28.6. ITERATIVE METHOD

Let $\{a_n\}$ be the sequence defined by: a_k = a_{k-1}+ 2 with a_0 = 1.
• Plugging values of k into the relation, we get:
a_1 = a_0 + 2 = 1 + 2
a_2= a_1 + 2 = 1 + 2 + 2 = 1 + 2(2)
a_3 = a_2 + 2 = 1 + 2 + 2 + 2 = 1 + 3(2)
a_4 = a_3 + 2 = 1 + 2 + 2 + 2 + 2 = 1 + 4(2)
• Continuing in this fashion reinforces the apparent pattern that a_n = 1 + n(2) = 1 + 2n.
• This brute force technique is the Method of Iteration.

Example 28.3. Using Iterative method solve the recurrence relation

$$u_n = u_{n-1} + n \text{ for } n \geq 1;\ u_0 = 1 \quad \ldots\ldots\ldots\ldots\ldots\ldots(1)$$

Solution: Replacing n by $n-1, n-2, n-3, \ldots, 2,1$ in $u_n = u_{n-1} + n$ we get

$$u_{n-1} = u_{n-2} + n - 1$$
$$u_{n-2} = u_{n-3} + n - 2$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$

$$u_2 = u_1 + 2$$
$$u_1 = u_0 + 1$$

Then from above we get

$$\begin{aligned} u_n &= (u_{n-2} + n - 1) + n \\ &= u_{n-3} + (n-2) + (n-1) + n \\ &= u_{n-4} + (n-3) + (n-2) + (n-1) + n \\ &\qquad \cdots\cdots\cdots\cdots\cdots\cdots \\ &\qquad \cdots\cdots\cdots\cdots\cdots\cdots \\ &= u_2 + 3 + 4 + \cdots + (n-2) + (n-1) + n \\ &= u_0 + 1 + 2 + 3 + 4 + \cdots + (n-2) + (n-1) + n \\ &= 1 + 2 + 3 + 4 + \cdots + (n-2) + (n-1) + n \\ &= 1 + \frac{n(n+1)}{2} \end{aligned}$$

Therefore the required solution is $u_n = 1 + \frac{n(n+1)}{2},\ n \geq 1$

Exercise 28.2: Using Iterative method solve the recurrence relation

$$S_k = aS_{k-1} \text{ with } S_0 = b$$

Exercise 28.3: Using Iterative method solve the recurrence relation

$$t_n = 2t_{n-1} + 5^n, n \geq 1 \text{ and } t_0 = 1$$

LECTURE 29: SOLUTION OF LINEAR RECURRENCE RELATION WITH CONSTANT COEFFICIENTS BY CHARACTERISTIC ROOT METHOD.

29.1. METHOD OF CHARACTERISTIC ROOT

In this method the general solution of the recurrence relation $\sum_{i=0}^{k} c_i a_{n-i} = f(n) \ldots \ldots (1)$ consists of two parts (i.e two numeric function).The first one is called homogeneous solution which satisfies the homogeneous recurrence relation $\sum_{i=0}^{k} c_i a_{n-i} = 0$...................(2)
and the other is called particular solution which satisfies (1)

29.2. HOMOGENEOUS SOLUTION

Here the recurrence relation is $\sum_{i=0}^{k} c_i a_{n-i} = 0$...(3)
We seek for a solution of the form $a_n = r^n$ where r is a real number.
Putting in (3) we get
$c_0 r^n + c_1 r^{n-1} + c_2 r^{n-2} + \cdots + c_k r^{n-k} = 0$
Dividing by r^{n-k} we get
$c_0 r^k + c_1 r^{k-1} + c_2 r^{k-2} + \cdots + c_k = 0$
This is called the **characteristic equation** of the recurrence relation (3)

Case 1:
If the characteristic equation has k distinct roots $r_1, r_2, \ldots \ldots, r_k$
then a sequence $\{a_n\}$ is a solution of the recurrence relation iff
$a_n = \alpha_1 r_1{}^n + \alpha_2 r_2{}^n + \cdots + \alpha_k r_k{}^n$ for integers $n \geq 0$
Where $\alpha_1, \alpha_2 \ldots \ldots, \alpha_k$ are constants.

Case 2:
If the characteristic equation has t distinct roots $r_1, r_2, \ldots \ldots, r_t$ with multiplicities $m_1, m_2, m_3, \ldots \ldots, m_t$ respectively so that all $m_i' s$ are positive and $\sum_{i=1}^{t} m_i = k$
then a sequence $\{a_n\}$ is a solution of the recurrence relation iff
$a_n = (\alpha_{1,0} + \alpha_{1,1} n + \alpha_{1,2} n^2 + \cdots + \alpha_{1,m_1-1} n^{m_1-1}) r_1{}^n + (\alpha_{2,0} + \alpha_{2,1} n + \alpha_{2,2} n^2 + \cdots + \alpha_{2,m_2-1} n^{m_2-1}) r_2{}^n + \cdots + (\alpha_{t,0} + \alpha_{t,1} n + \alpha_{t,2} n^2 + \cdots + \alpha_{t,m_t-1} n^{m_t-1}) r_t{}^n$
for integers $n \geq 0$
where $\alpha_{i,j}$ are constants for $1 \leq i \leq t$; $0 \leq j \leq m_i - 1$

29.3. PARTICULAR SOLUTION

Let us consider the recurrence relation (1)
For particular solution there is no general method for every function f(n).However for certain types of functions , the solution can be obtained by the method of inspection which is sometimes called **trial sequence method.**

The following table shows a list of trial functions T(n) against specified f(n).The coefficients $A_0 . A_1, A_2, \ldots$ are unknown constants to be determined.

	f(n)	T(n)
1	b^n where b is not a root of the characteristic equation	Ab^n

2	b^n where b is a root of the characteristic equation with multiplicity m	A$n^m b^n$
3	Polynomial P(n) of degree s	$A_0 + A_1 n + A_2 n^2 + \cdots + A_s n^s$
4	c^n P(n) where Polynomial P(n) of degree s and c is not a root of the characteristic equation	$c^n(A_0 + A_1 n + A_2 n^2 + \cdots + A_s n^s\ )$
5	c^n P(n) where Polynomial P(n) of degree s and c is a root of the characteristic equation with multiplicity m	$n^m c^n(A_0 + A_1 n + A_2 n^2 + \cdots + A_s n^s\ )$

Example 29.1 Solve the recurrence relation by characteristic root method

$$a_n - 6a_{n-1} + 8a_{n-2} = 3^n$$

Solution: The characteristic equation of the given recurrence relation is

$x^2 - 6x + 8 = 0$ which gives $x = 2\ ,4$

Therefore the homogeneous solution(H.S) = A. $2^n + B.4^n$

Since right hand side of the given recurrence relation is 3^n and 3 is not a root of the characteristic equation so particular solution(P.S) is of the form $\beta.3^n$

Putting $a_n = \beta.3^n$ in the given recurrence relation we get

$\beta.3^n - 6\beta.3^{n-1} + \beta.3^{n-2} = 3^n$

$or\ \left(\beta - 2\beta + \frac{8\beta}{9}\right)3^n = 3^n$

$or - \frac{\beta}{9}3^n = 3^n$

Since this is an identity so coefficient of 3^n on both sides are equal.

i.e $-\frac{\beta}{9} = 1\ \ or\ \beta = -9$

therefore P.S = -9.3^n

therefore the general solution is $a_n =$ A. $2^n + B.4^n - 9.3^n$

Exercise 29.1: Solve the recurrence relation by characteristic root method

$$a_n = 2\ a_{n-1} +\ 3\ a_{n-2}\ ; a_0 = 1\ ,\ a_1 = 2$$

Exercise 29.2: Solve the recurrence relation by characteristic root method

$$a_n = 4a_{n-1} - 3a_{n-2} + 2^n + n + 3\ ;\ a_0 = 1, a_1 = 4$$

LECTURE 30: SOLUTION OF LINEAR RECURRENCE RELATION WITH CONSTANT COEFFICIENTS BY GENERATING FUNCTION METHOD.

30.1. GENERATING FUNCTION

For a sequence $\{a_n\}$ we define an infinite series
$a_0 + a_1x + a_2x^2 + \cdots + a_nx^n + \cdots + \infty = \sum_{n=0}^{\infty} a_nx^n$
Which is called the generating function of the sequence $\{a_n\}$ and is denoted by g(x).

For example, the generating function of the sequence $\{2^n\}$ is
g(x)= $2^0 + 2^1x + 2^2x^2 + \cdots + 2^nx^n + \cdots + \infty = \sum_{n=0}^{\infty} 2^nx^n$
which can be written in closed form g(x) = $(1-2x)^{-1} = \frac{1}{(1-2x)}$

30.2. CLOSED FORM OF SOME GENERATING FUNCTION

I. $(a+x)^n = \sum_{r=0}^{n} {}^nC_r a^{n-r}x^r$ where n is a positive integer
II. $(a-x)^n = \sum_{r=0}^{n}(-1)^r \, {}^nC_r a^{n-r}x^r$ where n is a positive integer
III. $(1-x)^{-1} = \sum_{r=0}^{\infty} x^r \; ; -1 < x < 1$
IV. $(1+x)^{-1} = \sum_{r=0}^{\infty}(-1)^r x^r \; ; -1 < x < 1$
V. $e^x = \sum_{r=0}^{\infty} \frac{x^r}{r!}$
VI. $log(1+x) = \sum_{r=1}^{\infty}(-1)^{r+1}\frac{x^r}{r}, -1 < x \leq 1$

30.3. THEOREM

If $g_1(x) and \ g_2(x)$ are generating functions of $\{a_n\}$ and $\{b_n\}$ respectively then $\alpha g_1(x) + \beta g_2(x)$ is generating function of the sequence $\{\alpha a_n + \beta b_n\}$ where α, β are real numbers independent of n.

Exercise 30.1: Find the generating function for the following sequence.

I. $\{1,1,1,1, \dots \dots\}$
II. $\{1,-2,3,4, \dots \dots\}$
III. $\{2.3^r\}$,$r \geq 0$

30.4. SOLUTION OF RECURRENCE RELATION

We solve the recurrence relation by using generating function method and we illustrate this by the following examples.

Example 30.1: Using generating function solve the recurrence relation,

$a_n - 7a_{n-1} + 10a_{n-2} = 0$ for all n> 1 and $a_0 = 3, a_1 = 3$

Solution: Let g(x) be the generating function of $\{a_n\}$

i.e, g(x) $= \sum_{n=0}^{\infty} a_n x^n$

multiplying the recurrence relation by x^n and summing from $n = 2\ to\ \infty$, we get

$\sum_{n=2}^{\infty} a_n x^n - 7\sum_{n=2}^{\infty} a_{n-1}x^n + 10\sum_{n=2}^{\infty} a_{n-2}x^n = 0$

$or\ (\sum_{n=0}^{\infty} a_n x^n - a_0 - a_1 x) - 7x\sum_{n=2}^{\infty} a_{n-1}x^{n-1} + 10x^2\sum_{n=2}^{\infty} a_{n-2}x^{n-2} = 0$

$or\ (g(x) - 3 - 3x) - 7x(\sum_{n=1}^{\infty} a_{n-1}x^{n-1} - a_0) + 10x^2\sum_{n=0}^{\infty} a_n x^n = 0$

$or\ (g(x) - 3 - 3x) - 7x(g(x) - 3) + 10x^2 g(x) = 0$

$or\ g(x) = \frac{3-18x}{(5x-1)(2x-1)}$

$or\ g(x) = \frac{4}{1-2x} - \frac{1}{1-5x}$

$or\ g(x) = 4(1-2x)^{-1} - (1-5x)^{-1}$

$or\ g(x) = 4\sum_{n=0}^{\infty} 2^n x^n - \sum_{n=0}^{\infty} 5^n x^n = \sum_{n=0}^{\infty}(2^{n+2} - 5^n)x^n$

Hence $a_n = 2^{n+2} - 5^n$ for all n≥ 0

Exercise 30.1: Using generating function solve the recurrence relation,

$$a_n - 8a_{n-1} = 10^{n-1}, a_0 = 1$$

MODULE V: ALGEBRAIC STRUCTURE

(NUMBER OF LECTURES: 6L)

Lecture 31: Introduction

31.1 What is this course about?

Take a look at the following questions.

- Give a number n which leaves a remainder of 20 when divided by 23 and 62 when divided by 83.
- How many different necklaces can you form with 2 black beads and 8 white beads? How many necklaces can you form with blue, green and black beads?
- What are the last two digits of $a^{40}$when a is not divisible by 2 or 5?
- When does the equations of the form $x - y = z$ make sense? If x is a natural number or an integer or a matrix or an apple or a permutation?

When we look at these questions, they seem unrelated and seem to have no common thread.. Mathematicians realized long time back that problems in algebra, number theory and even geometry can be solved using very similar techniques. They were interested in finding out the common element among these proofs and were interested in searching for more domains where such techniques are applicable. It turns out that there is a single mathematical theory which can help us understand these questions in a single framework and give us answers to these seemingly non-related topics.

The mathematical framework which ties these questions together is called abstract algebra. Not surprisingly, given the name, the course is going to be about abstract algebra.

Exercise What does abstract mean?

Note. The exercises given in the course notes are practice problems with the exception of this particular introduction. The exercises given in this particular document are to motivate the study of abstract algebra. You should try to think about them but remember that there are no clear answers.

We will precisely study the mathematical structures which can represent numbers, matrices, permutations, geometric objects under different parameters. The first step would be to define these mathematical (algebraic) structures like groups, rings and fields. The next step is to find properties of these algebraic structures. Finally we will also see how these properties give so many beautiful results in different areas of mathematics.

Let’s start with a more basic question,

Exercise What does algebra mean?

31.2 Arithmetic and algebra

Most of the people when asked the above question, think about numbers, equations and operations between them. So let's make the previous question more precise. What is the difference between arithmetic and algebra? Arithmetic is the study of numbers and the operations (like addition, subtraction, multiplication) between them. Algebra, intuitively, talks about equations, variables, symbols and relations between them.

The primary difference is the use of variables, which can stand for an unknown or a group of numbers. These variables are somewhat abstract but really help us in manipulating equations and solving them. It would be too cumbersome to write things in words instead of using equations and variables.

Exercise Give an example where using a variable helps you to write a statement concisely.

Now we know what algebra is, let's talk about abstract part of it.

31.3 Abstraction

All of us like numbers (or at least understand the importance of it). One of the reasons is that numbers are very well-behaved. In other words, there are so many nice properties that it is easy to manipulate and work with numbers. Let's look at one of the most fundamental properties,

Theorem *Fundamental theorem of arithmetic: Every integer greater than 1 can be uniquely expressed as the product of primes up to different orderings.*

Since this property is so useful, we should ask, are there other objects which satisfy similar theorems.

Exercise Do we have unique factorization theorem for matrices or permutations.

There is a very important methodology to generalize given proofs. You look at the proof and figure out the crucial step and properties which make the proof work. So one way to approach this question would be, carefully look at the proof of the theorem and figure out the properties of integers we have used at different step. Then check if another mathematical object satisfies the same properties.

In other words, any mathematical object which satisfies these properties will also have a unique factorization theorem. The abstract object which has all these properties can be given an appropriate name. This is similar to variables. As variables can take different values, this abstract object can be assigned different mathematical objects.

We will turn this method upside down. We will consider some basic properties and give a name to the abstract structure which satisfies these "basic properties".

Exercise Who decides these basic properties?

Using these "basic properties" we will come up with multiple theorems like the unique factorization theorem above. By the above discussion any mathematical object (from arithmetic, algebra, geometry or anywhere else) which has these "basic properties" will satisfy all the theorems too. Hence in one shot we will get theorems in diverse areas.

You are already familiar with one such abstract structure, set. A collection of objects is called set and it needs no other property to be satisfied.

Exercise What kind of theorems can you prove for sets?

In the course we will look at the collection of objects (sets) with certain composition properties. These will give rise to groups, rings etc.. The first such abstraction we will study is group.

Exercise Should we choose as many basic properties as possible or as less basic properties as possible?

Lecture 32: Groups

These notes are about the first abstract mathematical structure we are going to study, groups. You are already familiar with set, which is just a collection of objects. Most of the sets we encounter in mathematics are useful because of the operations we can perform on them. We can do addition, multiplication, AND, OR, take power etc..

Sets, by definition, need not have such operations. For example, S = {Apple; Oranges; M201; Monitor} is a set. But, if we look at more interesting sets like integers, matrices, permutations etc., we generally have operations which can be done on them. For example, you can add matrices, multiply permutations, add and multiply integers and so on.

Our next task is to define an abstract object (say a special set) with operation to compose elements inside the object. But first let's ask a basic question. What are the nice properties of addition of two natural numbers? What about integers?

To begin with, it is great that we can add two numbers, that is, the addition of any two numbers is a number. Another property not present in natural numbers is that we can always solve $a + x = b$ (a, b are given, x is unknown). Notice that we have to assume the existence of Zero.

32.1 Groups

A group G is a set with binary operation $*$, s.t.,

- Closure: For any two elements $a, b \in G$; their composition under the binary operation $a * b \in G$.
- Associativity: For all $a, b, c \in G$;we have $a * (b * c) = (a * b) * c$. This property basically means that any bracketing of $a_1 * a_2 * \ldots\ldots a_k$ is same (exercise).
- Identity: There is an element identity (e) in G, s.t., $a * e = e * a = a$ for all $a \in G$.
- Inverse: For all $a \in G$, there exist $a^{-1} \in G$, s.t., $a * a^{-1} = a^{-1} * a = e$.

Note Some texts define binary operation as something which has closure property. In that case, the first property is redundant. For the sake of brevity, it is sometimes easier to write xy instead of $x * y$.

Sometime we denote a group by its set and the operation, e.g., $(Z; +)$ is the group of integers under addition.

Exercise Show that integers form a group under addition (In other words, Integers have a group structure with respect to addition). Do they form a group under multiplication?

You can think of groups as being inspired by integers. In other words, we wanted to abstract out some of the fundamental properties of integers. We will later see that all groups share some properties with integers, but more interestingly, there are a lot of other groups which do not look like integers. That means there are some properties of integers which are not captured by the definition of groups. So what properties of integers do you think is not captured by groups?

To start with, we haven't specified *commutativity* as one of the basic properties. The properties are chosen so that we have many examples of groups and simultaneously we can prove a lot of theorems (properties) of this group structure. Later we will see that some important groups do not have commutativity property.

Definition A group is called commutative or abelian if, $\forall a, b \in G; a * b = b * a$.

32.2 Examples of groups

Exercise Can you think of any other group except integers under addition? Is it commutative?

The whole exercise of abstraction will be a waste if integers (addition) is the only set which follow group property. Indeed, there are many examples of groups around you, or at least in the mathematics books around you ☺

- Integers, Rationals, Reals, Complex numbers under addition. Clearly for all these 0 is the identity element. The inverse of an element is the negative of that element.
- Rationals, Reals, Complex numbers (without zero) under multiplication. Identity for these groups is the element 1. Why did we exclude integers?
- Positive rationals, positive reals under multiplication.
- The group Z_n, set of all remainders modulo n under addition modulo n. Will it be a group under multiplication? How can you make it a group under multiplication?

Till now all the examples taken are from numbers. They are all subsets of complex numbers. Let's look at a few diverse ones.

- The symmetries of a regular polygon under composition. In other words, the operations which keep the polygon fixed. The symmetries are either obtained through rotation or reflection or combination of both. This group is called Dihedral group.
- The set of all permutations of {1, 2, ... , n} under composition. What is the inverse element?
- The set of all $n \times n$ matrices under addition. The identity in this case is the all 0 matrix,

$$\begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$$

- The set of all $n \times n$ invertible matrices of real numbers. What is the identity element?

We have seen so many examples of groups. Are they all similar (we will define the word similar later). Can we represent a group in a succinct way. One of the trivial representation is the multiplication table of the group. It is a matrix with rows and columns both indexed by group elements. The $(i;\ j)^{th}$ entry denotes the sum of i^{th} and j^{th} group element. For example, let's look at the multiplication table of Z_5^+under multiplication. Here Z_5^+ denotes all the remainders modulo 5 Co-prime to 5 (gcd with 5 is 1).

	1	2	3	4
1	1	2	3	4
2	2	4	1	3
3	3	1	4	2
4	4	3	2	1

Exercise Notice that every element occurs exactly once in every row and every column. Do you think this property is true for any group or just Z_5^+ ?

Multiplication table gives us all the information about the group but is a pretty long description. Specifically it is quadratic in the size of the group. It turns out that groups have lot of properties which can help us in giving a more succinct representation. We already showed one property, that the identity is unique. What other theorems can be shown for groups?

32.3 Properties of groups

To start with, we need to define few quantities. Suppose we are given an element $x \neq e$ of group G. What other elements can be constructed with x. The composition with identity will not give anything new, so let's compose it with itself. Since G is a group, $x^2 := x * x$; $x^3 := x^2 * x$ (notice the new notation) and so on will be elements of group G. In this way we can create new elements in G except if these elements start repeating.

Suppose G is finite, then sooner or later there will exist i and j, s.t., $x^i = x^j$.

Exercise Show that the first element which will repeat is e.

The least positive j for which $x^j = \mathrm{e}$ is called the order of x and is denoted by $|x|$. Clearly the only element with order 1 is e and everything else will have a bigger order.

We will now go on to prove more properties of groups, but before that there is a warning. Groups are inspired by numbers and the notations are very similar. It is not surprising that sometimes you can get carried away and use properties of integers which are not really true for groups (e.g., commutativity).

For all the proofs for the theorems given below, notice that we will use the already known properties like closure, associativity, inverse, existence of identity. Then using those theorems we can prove other results. Now check your proofs for the exercises given in this section above.

This distinction can be made more clear by an analogy which we will use later too. Working with groups is like playing football☺. In general, for any activity you use your hands, feet or any other tool. But in case of football there is a restriction that you only use your feet. Using your feet you develop other skills which can be used to score a goal.

Our goal would be to prove theorems. Our feet will be the defining properties of groups (closure, associativity, inverse, identity). And the intermediate theorems would be like dribbling or kicking. You should not foul (use properties of integers) to prove a theorem (score a goal). So let's play football. We will use G to denote a group.

- The inverse of an element is unique.

Proof: Suppose a has two inverses b and c. Then $c = (ba)c = b(ac) = b$. What properties of groups did we use in this proof?

- Cancellation laws: Given $a, b, x \in G$, we know $ax = bx \Longrightarrow a = b$, and also $xa = xb \Longrightarrow a = b$. These are called respectively the right and the left cancellation law.

Exercise Prove the assertion. What does it say about the rows (or columns) of multiplication table?

- $x \in G$ and x^{-1} have the same order.

Proof: We will show that order of x^{-1} is at most the order of x, by symmetry this will prove the assertion. Suppose $x^n = e$. Multiply this equality by x^{-n} and we get $x^{-n} = e$ and hence the order of x^{-1} is less than n.

Exercise We did not define x^n. What do you think it should be?

For a finite group we have shown that its order is less than the cardinality (also called the order) of the group. Actually order of an element can be restricted to just the divisors of the order of the group. Look carefully at the following theorem and proof.

Theorem Suppose G is a finite group with n elements (n is the order of the group). If d is the order of an element $x \in G$ then n is a multiple of d $(d \mid n)$.

Proof: We will prove the theorem in two steps. First, we will show that $x^n = e \, \forall \, x \in G$. Second, if there is any m, s.t., $x^m = e$ then d divides m. From these two steps the conclusion can be easily inferred.

From the cancellation laws, it is clear that $S_x = \{xg : g \in G\} = G$ is a set. All elements of S_x are distinct, in G and hence they are just a permutation of elements of G. Taking the product over all elements of S_x ,

$$\Pi_{s \in S_x s} = \Pi_{g \in G xg} = x^n \Pi_{g \in G g} = x^n \Pi_{s \in S_x s}.$$

Using the first and the last step,

$$e = x^n.$$

So for every element $x \in G$, we know $x^n = e$.

For the second part, suppose $m = kd + r$ by division. Here k is the quotient and $r < d$ is the remainder. Then looking at x^m,

$$e = x^m = x^{kd + r} = x^r.$$

So there exists $r < n$, s.t. $x^r = e$. By the definition of order, $r = 0$. Hence d divides m.

Actually the proof given above is not correct.

Exercise Where is the mistake in the proof? Hint: It is in the first part.

If you look at the proof of fact that $x^n = e$,then it was proved using commutativity. So we have only proved that for a commutative or abelian group the theorem. 2.10 is true. It turns out that it is true for non-commutative groups too. We will prove the full generalization later with a different technique.

32.4 Isomorphism and homomorphism of a group

As discusses above we want to find out what kind of groups are there. Are they all *similar*. Let us formalize the notion of similarity now. Clearly if two sets are equal if and only if there is a bijection between them. But the bijection need not respect the composition. That means the composition properties of two groups might be completely different even if they have a bijection between them.

Exercise Would you say that groups $(Z_4, +)$ and $(Z_8^+, \times)$ similar (both have four elements). The second group is the set of all remainders modulo 8 which are Co-prime to 8.

Hint: Look at the orders of different elements in these groups.

Hence for group similarity, we need to take care of composition too. Two groups are considered same if they are isomorphic to each other. In other words there exist an isomorphism between the two. To define, a group $\boldsymbol{G_1}$ is isomorphic to group $\boldsymbol{G_2}$ if there exist a bijection: $\phi: \boldsymbol{G_1} \rightarrow \boldsymbol{G_2}$, s.t.,

$$\forall g, h \in G_1 : \phi(g)\phi(h) = \phi(gh).$$

The second property takes care of the composition. A related notion is called homomorphism where we drop the bijection criteria. So $\boldsymbol{G_1}$ is homomorphic to $\boldsymbol{G_2}$ if there exist a map: $\phi: \boldsymbol{G_1} \rightarrow \boldsymbol{G_2}$, s.t.,

$$\forall g, h \in G_1 : \phi(g)\phi(h) = \phi(gh).$$

Exercise Give a homomorphism which is not an isomorphism from a group G to itself.

Assignment

Exercise For any $a_1, a_2, \ldots\ldots a_k \in G$, show that expression $a_1 * a_2 * \ldots\ldots a_k$ is independent of bracketing.

Hint: Show it using induction that all expression are same as $a_1 * (a_2 * (\ldots\ldots * a_k) \ldots)$.

Exercise Prove that the identity is unique for a group.

Exercise Which Groups are commutative from the list of groups given in the section 2.1.1?

Exercise Prove that $G = \{a + b\sqrt{2} \mid a, b \in Q\}$ is a group under addition.

Exercise Which of them are groups under addition?

- The set of all rational numbers with absolute value < 1.
- The set of all rational number with absolute value ≥1.
- The set of all rational numbers with denominator either 1 or 2 in the reduced form.

Exercise Find the order of following,

- 3 in $(Z_5, +)$.
- 5 in $(Z_7, \times)$.

Exercise Give an example of a finite group where order of an element is different from order of the group.

Exercise If all elements have order 2 for a group G, prove that it is abelian.

Exercise Show that if $\boldsymbol{G_1}$ is isomorphic to $\boldsymbol{G_2}$ then $\boldsymbol{G_2}$ is isomorphic to $\boldsymbol{G_1}$.

Exercise Show an isomorphism from real numbers with addition to positive real numbers with multiplication.

Lecture 33: Subgroups

We are interested in studying the properties and structure of the group. By properties, we mean the theorems which can be proven about groups in general. Then any mathematical construct having the group structure (satisfy closure, associativity etc.) will satisfy those theorems.

Another important task is to understand the structure of group itself. It is deeply related to the properties of group. It ultimately helps us in figuring out which groups are similar (with respect to isomorphism) and can we list out all possible kind of groups (not isomorphic to each other).

One of the natural questions is that if groups can exist inside a group.

Exercise Can we have a subset of group which itself is a group under the group operation? Try to construct such a set in Z.

33.1 Definition

As the intuition would suggest,

Definition A subset H of a group G is called a subgroup if it is not empty, closed under group operation and has inverses. The notation $H \leq G$ denotes that H is a subgroup of G.

Note The subgroup has the same operation as the original group itself

Exercise Why did we not consider associativity, existence of inverse?

Every group G has two trivial subgroups $\{e\}$ and the group G itself. Let's look at few examples of non-trivial subgroups. Try to prove that each of them is a subgroup.

- nZ, the set of all multiples of n is a subgroup of Integers.
- Under addition, integers (Z) are a subgroup of Rationals (Q) which are a subgroup of Reals (R). Reals are a subgroup of Complex numbers, C.
- Z^+, the set of all positive integers is not a subgroup of Z. Why?
- The set $S = \{a + b\sqrt{2} \mid a, b \in Z\}$is a subgroup of R under addition.
- *Center of a group*: The center of a group G is the set of elements which commute with every element of G.

$$C(G) = \{h \in G : hg = gh \ \forall g \in G\}.$$

We will show that center is the subgroup. Associativity follows from G and existence of identity is clear. Suppose $h, k \in C(G)$, then for any $g \in G$,

$$g(hk) = hgk = (hk)g.$$

Hence $C(G)$ is closed. For the inverse, note that $gh = hg$ is equivalent to $h^{-1}gh = g$ and $g = hgh^{-1}$. Hence existence of inverse follows (Why?).

Lecture 34: Cyclic groups

We noticed that $\{e\}$ is a subgroup of every group. Let's try to construct more subgroups. Suppose x is some element which is not the identity of the group G. If k is the order of x then $S_x = \{e, x, x^2, \dots, x^{k-1}\}$ is a set with all distinct entries. It is clear from previous discussion of groups that S_x is a subgroup.

Exercise Prove that S_x is a subgroup.

While proving the previous exercise, we need to use the fact that k is finite. What happens when k is infinite? Can we construct a group then? The answer is yes, if we include the inverses too. All these kind of groups, generated from a single element, are called *cyclic groups.*

Definition A group is called cyclic if it can be generated by a single element. In other words, there exist an element $x \in G$, s.t., all elements of G come from the set,

$$< x >= \{\dots, x^{-2}, x^{-1}, e, x, x^2, \dots\}$$

There are many things to note here:

- For an infinite group, we need to consider inverses explicitly. For a finite group, inverses occur in the positive powers.
- The group generated by the set S is the group containing all possible elements obtained from S through composition (assuming associativity, inverses etc.).
- The notation for the group generated by S is $< S >$.

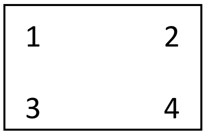
The structure of cyclic groups seems very simple. You take an element and keep composing. What different kind of cyclic groups can be there? Look at different examples of cyclic groups of order 4 in the following figure. The next theorem shows that all these are isomorphic.

$$Z_4 = \{\bar{0}, \bar{1}, \bar{2}, \bar{3}\}$$

$$Z_5^+ = \{\bar{1}, \bar{2}, \bar{3}, \bar{4}\}$$

$$Permutations \subset S_4$$

(1 2 2 4) (4 1 2 3) (3 4 1 2) (2 3 4 1)

Rotational Symmetryies of 

Theorem Every finite cyclic group G of order n is isomorphic to Z_n.

Proof. Suppose x is a generator for G. It exists by the definition of G. Then since the order is finite, group G is,

$$G = \{e, x, x^2, \dots, x^{n-1}\}$$

Lets look at the obvious bijection from Z_n to G. The element k is mapped to x^k. It is a bijection because, the inverse maps x^k to k. For the above bijection,

$$\phi(j+k) = x^{j+k \mod n} = x^j * x^k = \phi(j) * \phi(k).$$

Where first inequality follows from the definition of Z_n and second from the fact that $x^n = 1$. This shows that ϕ is an isomorphism. Hence proved.

Using the previous theorem and exercise (assignment), we have given complete characterization of cyclic groups. This loosely means that we can get all the properties of any cyclic group of order n from Z_n and an in finite cyclic group with integers.

This is called a classification of cyclic groups. We would ideally like to give classification of groups and finding out more properties of groups. These two questions are not independent. We will explore both simultaneously and progress in one question helps in finding the answer for other.

Exercise What are the subgroups of a cyclic group?

Lecture 35: Cosets

The next step in understanding the structure of a group is to partition it using a subgroup. Suppose we are given a group G and its subgroup H. We will show that G can be partitioned into disjoint sets of equal size ($|H|$)This will imply that$|G|$ is always divisible by$|H|$. Lets define these parts first and then we can prove the fact given above.

Definition *Cosets*: The left coset (gH) of H with respect to an element g in G is the set of all elements which can be obtained by multiplying g with an element of H,

$$gH \ = \{gh : h \in H\}.$$

This is called the left coset because g is multiplied on the left. We can similarly define the right cosets Hg.

Exercise How are left and right coset related for commutative groups?

Let us show some properties of these cosets. Remember not to use any illegal property while proving these. Without loss of generality we will assume that cosets are left. Same properties hold true for right ones too.

Every element of G is in at least one coset. H is one of the cosets too.

Proof: Exercise.

The cardinality of all cosets is equal and hence their cardinality is $|H|$.

Proof: Consider a coset gH and a subgroup $H \ = \{h_1, h_2, \dots, h_k\}$. The elements of the left coset gH are $\{gh_1, gh_2, \dots, gh_k\}$.. It is easy to show that any two elements in this set are distinct (why?). Hence all cosets have cardinality $k \ = |H|$.

For any two elements $g_1, g_2$of G either g_1H, g_2H are completely distinct (disjoint) or completely same $g_1H = \ g_2H$.

Proof: Suppose there is one element common in $g_1H \ and \ g_2H$ (otherwise they are completely distinct). Say it is $g_1H = \ g_2H$, then,

$$g_1 = \ g_2h_2h_1^{-1} \rightarrow \exists h \in H: g_1 = \ g_2h.$$

Now you can prove a simple exercise.

Exercise If $\exists h \in H: g_1 = \ g_2h$ then show that$g_1H \subseteq \ g_2H$.

But if $g_1 = \ g_2h$ then$g_2 = \ g_1h^{-1}$. This will show from the previous exercise that $g_2H \subseteq \ g_1H$. Hence both the sets $g_1H \ and \ g_2H$ are the same.

Using the properties we have shown that the two columns of the following table are completely the same or completely distinct.

${}^{G}/_{H}$	e	g_2	$\cdots$	g_n
e	e	g_2	$\cdots$	g_n
h_2	h_2	g_2h_2	$\cdots$	g_nh_2
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
h_k	h_k	g_2h_k	$\cdots$	g_nh_k

This conclusion is beautifully summarized in Lagrange’s theorem.

35.1 Lagrange's theorem

Using the previous list of properties it is clear that if we look at the distinct cosets of H then they partition the group G into disjoint parts of equal size.

Exercise What is the size of these parts?

Theorem Lagrange: Given a group G and a subgroup H of this group, the order of H divided the order of G.

Proof: The proof is left as an exercise. You should try to do it without looking at the hint given in the next line.

Hint: From the previous discussion, the $\left|\frac{G}{H}\right|$is just the number of distinct cosets of H.

Note If the set of left and right cosets coincide the subgroup is called normal. In this case, the set of cosets actually forms a group, called the quotient group $\frac{G}{H}$ (What is the composition rule?).

This is a great discovery. The statement of Lagrange's theorem does not do justice to the implications. We started with an abstract structure with some basic properties like associativity, inverses etc. (group). The proof of Lagrange's theorem implies that if we can find a subgroup of the group then the whole group can be seen as a disjoint partition with all parts related to the subgroup. Notice that it is easy to construct a cyclic subgroup of a group.

Exercise Prove that the order of an element always divides the order of a group. We had proved this for commutative groups in an earlier lecture.

Exercise What does Lagrange's theorem say about groups with prime order?

Let’s look at one application of Lagrange's theorem in the case of $Z_m^{\times}$. We know that this group contains all the remainders $mod\ m$ which are coprime (gcd is 1) to m. If m is a prime p then $Z_p^{\times}$ contains $p - 1$ elements. This proves the well known *Fermat's little theorem*.

Exercise Fermat's little theorem: For a prime p and any number a, $a^{p-1} = 1 \mod p$.

Prove this theorem.

Assignment

Exercise List all possible subgroups of Z_6 under addition.

Exercise The kernel of a homomorphism $\phi: G \rightarrow L$ is the subset of G which maps to identity of L. Hence, $Ker(\phi) = \{g \in G : \phi\&(g) = e_L\}$.

Similarly, the image of ϕ are the elements of L which have some element mapped to them through ϕ. $Img(G) = \{h \in L : \exists g \in G \, for \, which \, (g) = h.\}$

Show that $Img(G)$ and $Ker(G)$ are subgroups.

Exercise Show that a subset H is a subgroup of G if it is non-empty and $\forall x, y \in H: xy^{-1} \in H.$

Note. Because H is a subset, the set of properties we need to check are much less.

Exercise Show that Z_n is cyclic under addition. Give some examples of cyclic subgroups and some examples of non-cyclic subgroups in Z_n^+under multiplication.

Exercise Show that all cyclic groups are commutative (abelian).

Hint: Look for the obvious bijection between the group and Z. Show that it is an isomorphism.

Exercise Find the order of every element of group Z_n where p is a prime.

Exercise *Euler's theorem:* For a number m, say (m) is the number of positive elements coprime to m and less than m. For any a which is co-prime to m, $a^{\phi(m)} = 1$ mod m.

Prove this theorem.

Exercise Show that there always exist a cyclic subgroup of any finite group G.

Exercise Show that the subgroup of a cyclic group is cyclic.

Lecture 35 A: Quotient Group, Normal subgroup

We have seen that the cosets of a subgroup partition the entire group into disjoint parts. Every part has the same size and hence Lagrange's theorem follows. If you are not comfortable with cosets or Lagrange's theorem, please refer to earlier notes and refresh these concepts.

So we have information about the size of the cosets and the number of them. But we lack the understanding of their structure and relations between them. In this lecture, the concept of normal subgroups will be introduced and we will form a group of cosets themselves!!

35A.1 Normal subgroup

Suppose we are given two elements g, n from a group G. The conjugate of n by g is the group element gng^{-1}.

Exercise When is the conjugate of n equal to itself ?

Clearly the conjugate of n by g is n itself iff n and g commute.

We can similarly define the conjugate of a set $N \subseteq G$ by g,

$gNg^{-1} \coloneqq \{gng^{-1}: n \in N\}$.

Definition *Normal subgroup*: A subgroup N of G is normal if for every element g in G, the conjugate of N is N itself.

$$gNg^{-1} = N \ \forall g \in G.$$

We noticed that $gng^{-1} = n$ iff g; n commute with each other.

Exercise When is $gNg^{-1} = N$?

In this case the left and right cosets are the same for any element g with respect to subgroup N. Hence, a subgroup is normal if its left and right cosets coincide.

Exercise Show that following are equivalent. So you need to show that each of them applies any other.

1. N is a normal subgroup.
2. The set $S = \{g : gN = Ng\}$ is G itself.
3. For all elements $g \in G, gNg^{-1} \subseteq N$.

Hint: Instead of showing all $2 \times \binom{3}{2}$ implications, you can show $1) \implies 2) \implies 3) \implies 1$.

35A.2 Quotient group

We have introduced the concept of normal subgroups without really emphasizing why it is defined. Let's move to our original question. What can be said about the set of cosets, do they form a group?

Suppose G is a group and H is a subgroup. Denote by S, the set of cosets of G with respect to H. For S to be a group it needs a law of composition. The most natural composition rule which comes to mind is,

$$(gH)(kH) = (gk)H.$$

Here gH and kH represent two different cosets. The problem with this definition is that it might not be well-defined. It might happen that $g' \in gH$and $k' \in kH$ when multiplied give a totally different coset $(g'k')H$ then $(gk)H$.

Exercise Show that this operation is well-defined for commutative (abelian) groups.

What about the general groups? Here comes the normal subgroup to the rescue.

Theorem *Suppose G is a group and H is its subgroup, the operation,*

$$(gH)(kH) = (gk)H,$$

is well defined if and only if H is a normal subgroup.

Note Every subgroup of a commutative group is normal.

Proof: We need to show that if the operation is well defined then $ghg^{-1} \in H$ for every $g \in G, h \in H$. Consider the multiplication of H with $g^{-1}H$. Since $e, h \in H$, we know $eH = hH$. Since the multiplication is well defined,

$$(eg^{-1})H = (eH)(g^{-1}H) = (hH)(g^{-1}H) = (hg^{-1})H \Longrightarrow g^{-1}H = (hg^{-1})H.$$

Again using the fact that $e \in H, hg^{-1} \in g^{-1}H$. This implies $hg^{-1} = g^{-1}h' \Longrightarrow ghg^{-1} = h'$ for some $h' \in H$.

Suppose N is a normal subgroup. Given $g' = gn$ and $k' = kn'$, where $g, g', k, k' \in G$ and $n, n' \in N$ we need to show that $(gk)N = (g'k')N$.

$$(g'k')N = (gnkn')N.$$

Definition Given a group G and a normal subgroup N, the group of cosets formed is known as the quotient group and is denoted by $\frac{G}{N}$.

Using Lagrange's theorem,

Theorem Given a group G and a normal subgroup N,

$$|G| = |N|\left|\frac{G}{N}\right|$$

35A.3 Relationship between quotient group and homomorphisms

Let us revisit the concept of homomorphisms between groups. The homomorphism between two groups G and H is a mapping $\phi: G \rightarrow H$ that preserves composition.

$$\phi(gg') = \phi(g)\phi(g').$$

For every homomorphism we can define two important sets.

Image: The set of all elements h of H, s.t., there exists $g \in G$ for which $\phi(g) = h$.

$$Img(\phi) = \{h \in H : \exists g \in G \ \phi(g) = h\}$$

Generally, you can restrict your attention to $Img(\phi)$ instead of the entire H. *Kernel:* The set of all elements of G which are mapped to identity in H.

$$Ker(\phi) = \{g \in G : \phi(g) \ = \ e_H\}.$$

Notice how we have used the subscript to differentiate between the identity of G and H.

Note $Img(\phi)$ is a subset of H and $Ker(\phi)$ is a subset of G.

Exercise Prove that $Img(\phi)$ and $Ker(\phi)$form a group under composition with respect to $H$and G respectively.

Exercise Show that $Ker(\phi)$ is a normal subgroup.

There is a beautiful relation between the quotient groups and homomorphisms. We know that $Ker(\phi)$is the set of elements of G which map to identity. What do the cosets of $Ker(\phi)$ represent? Lets take two elements g, h of a coset $gKer(\phi)$. Hence $h \ = \ gk$ where $\phi(k) \ = \ e_H$. Then by the composition rule of homomorphism $\phi(g) \ = \ \phi(h)$.

Exercise Prove that $\phi(g) \ = \ \phi(h)$ if and only if g and h belong to the same coset with respect to $Ker(\phi)$.

The set of elements of G which map to the same element in H are called the *fibers* of ϕ. The previous exercise tell us that fibers are essentially the cosets with respect to $Ker(\phi)$ (the quotient group).

The fibers are mapped to some element in $Img(\phi)$by ϕ. Hence there is a one to one relationship between the quotient group $\frac{G}{Ker(\phi)}$ and $Img(\phi)$. Actually the relation is much stronger.

It is an easy exercise to show that the mapping between quotient group $\frac{G}{Ker(\phi)}$ and $Img(\phi)$. is an isomorphism.

We have shown that $Ker(\phi)$is normal. It can also be shown that any normal subgroup N is a kernel of some homomorphism (exercise).

Assignment

Exercise Given a group G and a normal subgroup N. Say the set of cosets is called S and has composition operation $(gH)(kH) \ = \ (gk)H$. Show that,

- Identity exists in this set.
- Inverses exist in this set.
- Associativity is satisfied.

Since Closure is obvious we get that S is a group with respect to the above mentioned composition rule.

Exercise Suppose G is an abelian group and H is a subgroup. Show that $\frac{G}{H}$ is abelian.

Exercise Given N is a normal subgroup, prove that $g^k(N) \ = \ (gN)^k$.

Exercise Suppose N is normal in G, show that for a subgroup $H, H \cap N$ is a normal subgroup in H.

Exercise Show that a subgroup N is normal in G iff it is the kernel of a homomorphism from G to some group H.

Lecture 36: Ring, Field, Integral Domain

We have shown that Z_n is a group under addition and Z_n^+ is a group under multiplication (set of all numbers co-prime to n in Z_n). Till now, the two operations + and have been treated differently. But from our experience with integers and even matrices, these operations satisfy properties like "distribution" $a(b + c) = (ab + ac)$.

Hence, after success in defining an abstract structure with one operation (group), now we define another abstract structure with 2 operations. The first question is, what should be the defining properties of this new abstract structure. We will be inspired by integers again and define the concept of *Rings.*

36.1 Rings

Consider two operations $+$ and $\times$ in a set R.

Definition The set R with the two operations $+$ and $\times$ is a ring, if,

- R is a commutative group under $+$.
- R is associative, closed and has an identity with respect to the operation $\times$.
- The two operations $+$ and $\times$ follow the distributive law, i.e.,

$$a \times (b + c) = a \times b + a \times c \; and \; (a + b) \times c = a \times c + b \times c.$$

Note. We will always assume that the multiplicative identity is different from additive identity. The additive identity will be denoted by 0 and multiplicative identity by 1. For brevity, we will denote $a \times b$ as ab.

Exer**cise** Are the two conditions under the distributive law same?

Ex**ercise** Why did we assume commutativity under addition for a ring?

There are many examples of rings, many of these sets we have encountered before.

- The sets Z, Q, R, C are rings with addition and multiplication.
- The set of integers modulo m, Z_m, is a ring with addition and multiplication.
- The set of 2×2 matrices with integer entries is a ring. Actually if R is a ring then set of 2×2 matrices with entries in R is also a ring.

Another ring which will be of our particular interest is the ring of polynomials. The set $R[x]$ is the set of all polynomials with coefficients from ring R. If the multiplication in R is commutative then $R[x]$ is also a commutative ring.

Note The addition and multiplication of polynomials is defined in the same way as in regular polynomials.

Exercise Check that you can define these operations on polynomials with entries from a ring R. Why do we need that multiplication is commutative in the original ring?

Hence we have polynomial rings $Z[x], Q[x], R[x], C[x]$ having commutative multiplication.

36.2 Units of a ring

The ring is not a group with respect to multiplication. That is because inverses need not exist in a ring (e.g., integers). The elements of rings which have inverses inside the ring with respect to multiplication are called units or invertible elements.

The set of units for Z are just ± 1.

Exercise Prove that the set of units form a group under multiplication.

36.3 Characteristic of a ring

Rings have two identities $e_\times$ and e_+ (we will denote them by 1 and 0 respectively). For a ring an important criteria is the additive group generated by 1. The elements of that group are $1, 1 + 1, 1 + 1 + 1$ and so on. The smallest number of times we need to sum 1 to get 0 is called the characteristic of the ring.

For some cases, like reals, the sum never reaches the additive identity 0. In these cases we say that the characteristic is zero.

Exercise Prove that 1×0 = 0 in a ring.

36.4 Homomorphism for a ring

We have already defined the homomorphism for a group. How should we define the homomorphism for a ring?

Exercise Try to come up with a definition of ring homomorphism. Remember that the mapping should be well behaved with respect to both the operators.

When not clear from the context, we specify if it is a group homomorphism or a ring isomorphism.

We can define the kernel of a homomorphism $\phi: R \to S$ from a ring R to ring S as the set of elements of R which map to the additive identity 0 of S. A bijective homomorphism is called an isomorphism.

We showed in previous lectures that the kernel of a group homomorphism is a normal subgroup. What about the kernel of a ring homomorphism? For this, the concept of ideals will be defined.

36.5 Ideal

The ring R is a group under addition. A subgroup I of R under addition is called an *ideal* if

$$\forall x \in I, r \in R: xr, rx \in I$$

For example, the set of all elements divisible by n is an ideal in Z.

Exercise Show that nZ is an ideal of Z.

Ideal is similar to the normal subgroup, but belongs to a ring. Suppose I is an ideal. Then we can define the set of cosets of I with respect to R as $\frac{R}{I}$. We denote the elements of the set by $r + I$.

We know that $\frac{R}{I}$ is a group (why?), but it can be shown that it is a ring under the following operations too.

$$(r + I) + (S + I) = (r + s) + I \qquad (r + I)(s + I) = (rs) + I$$

Exercise Show that the kernel of a ring homomorphism is an ideal.

Kernel of any ring homomorphism is an ideal and every ideal can be viewed this way. We can define quotient ring using ideals as we defined quotient group using normal subgroup. It turns out,

Theorem Given a homomorphism ϕ: $R \to S$,

$$\frac{R}{Ker(\phi)} \cong Img(\phi)$$

Given a set $S \subseteq I$, we can always come up with the ideal generated by the set. Suppose the multiplication is commutative, then

$$I = \{ r_1x_1 + r_2x_2 + \cdots .. + r_nx_{n1}: \forall i \, r_i \in R, x_i \in S\},$$

is the ideal generated by S.

Exercise Prove that it is an ideal.

36.6 Integral domain

Our main motivation was to study integers. We know that integers are rings but they are not fields. We also saw (through exercise) that integers are more special than rings. The next abstract structure is very close to integers and is called *integral domain*.

An integral domain is a commutative ring (multiplication is commutative) where product of two non-zero elements is also non-zero. In other words, if $ab = 0$ then either $a = 0$ or $b = 0$ or both.

Exercise Give some examples of an integral domain. Give some examples of rings which are not integral domains.

We said that integral domain is closer to integers than rings. The first thing to notice is that integral domains have cancellation property.

Exercise If $ab = ac$ in an integral domain, then either $a = 0$ or $b = c$.

Now we will see that the properties of divisibility, primes etc. can be defined for integral domains. Given two elements $a, b \in R$, we say that a divides b (b is a multiple of a) if there exist an $x \in R$, s.t., $ax = b$.

Exercise If a divides b and b divides a then they are called *associates*. Show,

- Being associates is an equivalence relation.
- a and b are associates iff $a = ub$ where u is a unit.

You can guess (from the example of integers), the numbers 0 and units (± 1) are not relevant for divisibility. A non-zero non-unit x is *irreducible* if it can't be expressed as a product of two non-zero non-units. A non-zero non-unit x is prime if whenever x divides ab, it divides either a or b.

Notice that for integers the definition of irreducible and prime is the same. But this need not be true in general for integral domain. For examples, look at any standard text.

Exercise What is the problem with defining divisibility in ring?

36.7 Fields

If you look at the definition of rings, it seems we were a bit unfair towards multiplication. R was a commutative group under addition but for multiplication the properties were very relaxed (no inverses, no commutativity). Field is the abstract structure where the set is almost a commutative group under multiplication.

Definition The set F with the two operations $+$ and $\times$ is a field, if,

- F is a commutative group under $+$.
- $F - \{0\}$ is a commutative group under $\times$ (it has inverses).
- The two operations $+$ and $\times$ follow the distributive law, i.e.,

$$a \times (b + c) = a \times b + a \times c \; and \; (a + b) \times c = a \times c + b \times c.$$

Exercise Why are we excluding the identity of addition when the multiplicative group is defined?

As you can see Field has the strongest structure (most properties) among the things (groups, rings etc..) we have studied. Hence many theorems can be proven using Fields. Fields is one of the most important abstract structure for computer scientists.

Note The notion of divisibility etc. are trivial in fields.

Let us look at some of the examples of fields.

- Z is NOT a field.
- Q, R and C are fields.
- Z_m is a field iff m is a ________. (Fill in the blank)

The last example is of fields which have finite size. These fields are called finite fields and will be of great interest to us.

Assignment

Exercise Give a rule that is satisfied by Integers but need not be satisfied by rings in general.

Exercise Find the set of units in the ring Z_8.

Exercise If all the ideals in the ring can be generated by a single element then it is called a principal ideal domain. Show that Z is a principal ideal domain.

Exercise Show that if $ab = 0$ for a, b in a field F then show that either $a = 0$ or $b = 0$.

Exercise What are the units of a field?

Exercise Show that a finite integral domain is a field.

Exercise Find a number n which leaves remainder 23 with $31, 2$ with 37 and 61 with 73.

Exercise Find a number n which leaves remainder 3 when divided by 33 and 62 when divided by 81.

Thanks to the Rajat Mittal, IIT kanpur for his kind cooperation.

MODULE VI: ADVANCED GRAPH THEORY

(NUMBER OF LECTURES: 8)

LECTURE 37:**Advance Graph Theory**

37.1. PLANAR GRAPHS

Basic definitions:

Isomorphic graphs

Two graphs G1(V1,E1) and G2(V2,E2) are isomorphic if there is a one-to-one correspondence F of their vertices such that the following holds:

For all $u,v\in V1, uv \in E1$ implies $F(u)F(v) \in E2$

For all $x,y \in V1, x,y \notin E1$ implies $F(x)F(y) \notin E2$

Plane graph (or embedded graph)

A graph that is drawn on the plane without edge crossing, is called a Plane graph

Planar graph

A graph is called Planar, if it is isomorphic with a Plane graph

Phases

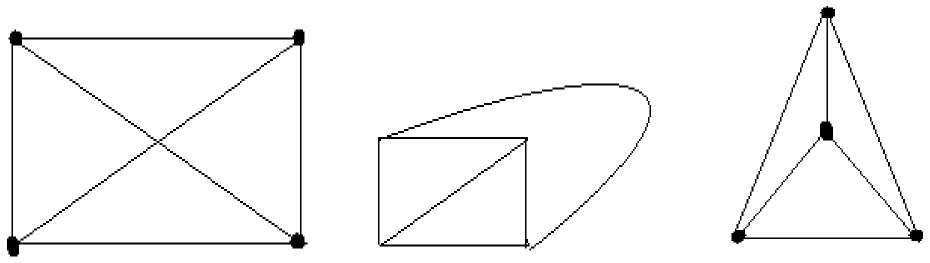
A planar representation of a graph divides the plane in to a number of connected regions, called faces, each bounded by edges of the graph.

For every graph G, we denote n(G) the number of vertices , e(G) the number of edges, f(G) the number of faces

Degree

We define the degree of a face d(f), to be the number of edges bounding the face f.

Examples: The following graphs are isomorphic to 4 (the complete graph with 4 vertices)



2nd and3rd graph is planer but 1st is not.1st graph is planar as it is isomorphic to 2nd and 3rd graph.

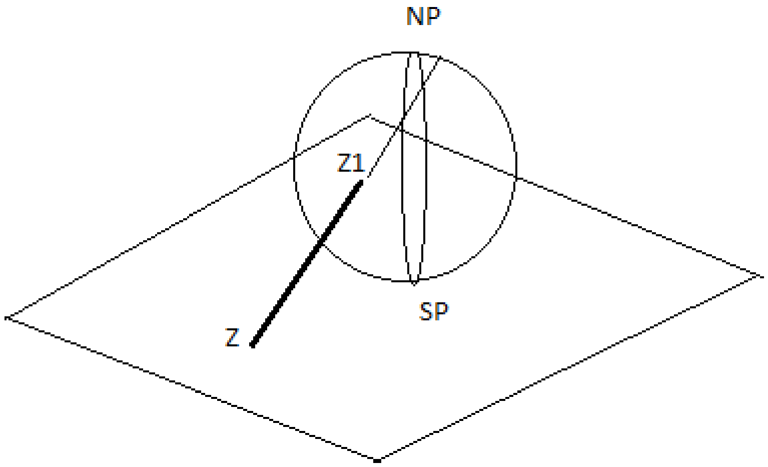
LECTURE 38:

38.1 Theorem 1

A graph is embeddable in the sphere if and only if it is embeddable in the plane.

Proof

We show this by using a mapping known as stereographic projection. Consider a spherical surface S touching a plane P at the point SP (called south pole). The point NP (called the point of projection or north pole) is on S and diametrically opposite SP. Any point z on P can be projected uniquely onto S at z' by making NP, z and z' collinear. In this way any graph embedded in P can be projected onto S. Conversely, we can project any graph embedded in s onto P , choosing NP so as not to lie an any vertex or edge of the graph.



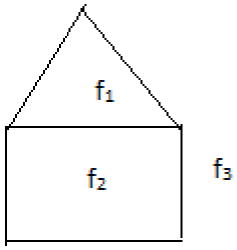
38.2 Theorem 2

A planar embedding G' of a graph G can be transformed in to another embedding such that any specified face becomes the exterior face.

Proof

Any face of G' is defined by the path which forms its boundary. Any such path, T, identified in a particular planar representation P of G, may be made to define the exterior face of a different planar representation P' as follows. We form a spherical embedding P'' of P. P' is then formed by projecting P'' onto the plane in such a way that the point of projection lies in the face defined by the image of T on the sphere.

Example



38.3 Theorem 3 (Euler's formula)

If G is a connected planar graph, for any embedding G' the following formula holds:

n(G)+f(g)=e(G)+2

Proof.

By induction on f

· For f(G) = 1, G is a tree. For every tree, e(G) = n(G)-1, so n(G)+1= e(G) + 2 implies n(G)+f(G) =e(G) +2 and the formula holds.

· Suppose it holds for all planar graphs with less than f faces and suppose that G' has f ≥2 faces.

· Let (u,v) be an edge of G which is not a cut-edge. Such an edge must exists because G' has more than one face. The removal of (u.v) will cause the two faces separated by (u,v) to combine, forming a single face.

Hence (G-(u,v))' is a planar embedding of a connected graph with one less face than G' , hence:

f(G - (u,v)) = f(G) - 1

n(G - (u,v) = n(G)

e(G = (u,v)) = e(G) – 1

But by the induction hypothesis:

n(G- (u,v)) + f(G - (u,v)) = e(G - (u,v)) + 2 and so, by substitution: n(G) + f(G) = e(G) + 2

Hence, by induction, Euler's formula holds for all connected planar graphs.

Lemma 1

For any embedding G' of any simple connected planar graph G,$\sum d(f)$ =2e(G).

Proof.

Each edge contributes 1 to each face it is a bound, so it contributes 2 to the total sum. So the e(G) edges contributes 2e(G) to the total sum.

38.4 Theorem 4(GenerelisedEuler"s formula)

A planar graph G with n vertices,e number of edges and knumber of connected components determines f=e-n+k+1.

Proof

Let the connected components of G be G_1,G_2,……,G_k.Let G_i hasn_i number of vertices,e_i number of edges and f_i number of regions(i=1,2,…n).Then by the previous theorem f_i=e_i-n_i+2 for i=1,2,,….k.Now the exterior region is same for all components.If the exterior regions are not considered then number of interior regions for each components is given by g_i=f_i-1.

Thus the total number of interior regions of G is

=(f_1-1)+(f_2-1)+……+(f_k-1)

$=\sum_{i=1}^{k} f\text{i} - \text{k}$

$=\sum_{i=1}^{k}(ei - ni + 2)$

$=\sum_{i=1}^{k} ei$-$\sum_{i=1}^{k} ni$+$\sum_{i=1}^{k} 2$ -k

=e-n+2k-k

=e-n+k

So the total number of interior regions are f=e-n+k+1.

38.5 Theorem 5

Let G be a simple connected planar graph with n vertices,e edges and f regions.Then

(a)e≥3f/2

(b)e≤3n-6.

Proof

If n=3 then G may have 2 or 3 edges.If G has 3 edges then G has 2 faces otherwise G has 1 region.Thus if e=3,f=2 and if e=2,then f=1.In any case(a) is true.

So we assume n=4 or n>4.If G is tree then e=n-1; and f=1 as there is no circuit to enclose a finite region.In that case e=n-1≥4-1=$\frac{3}{2}$2>3/2=3/2.f proving the result (a).If G is not a tree then it must contain a circuit and at least one circuit all of whose edges are boundary of infinite region of G.Now since G has no loop or parallel so number of boundary edges≥3,So sum of boundary edges≥3f.

Now the left side of the above inequality each edge is counted either once or twice.

Thus LHS of the inequality≤2e.

Therefore 2e≥3f

i.e. e≥$\frac{3f}{2}$

(b) Now from a previous theorem we have f=e-n+2.Then from result (a) we have

e≥$\frac{3(e-n+2)}{2}$

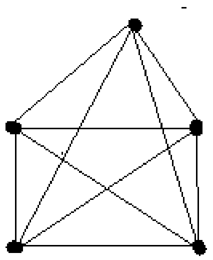
or,2e≥3e-3n+6

or,e≤ 3n-6

LECTURE 39:

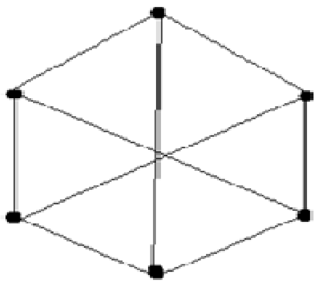
39.1 Kuratowski First Graph

A complete graph with five vertices is Kuratowski First graph. It is denoted by K_5.It is shown in the following graph.



39.2 Kuratowski Second Graph

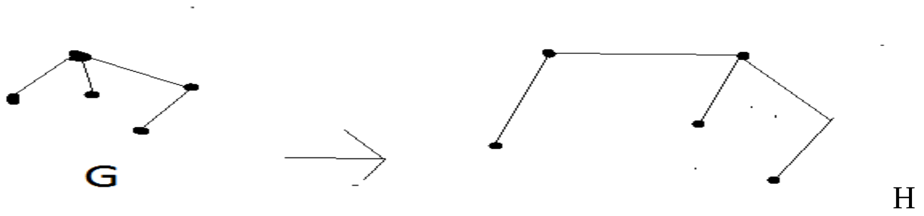
A regular connected graph with six vertices and nine edges is Kuratowski Second Graph.It is denoted by $K_{3,3}$..It is shown as follows



39.3 Homeomorphic Graph

Let e be an edge joining the two vertices u and v in a graph G.Let a new graph H be formed by deleting the edge e and introducing a new vertex w and two edges ,one joining u and w and the other joining v and w. This operation of replacement of an edge by two edges and a new vertex is called **edge subdivision.**A graph is obtained from a graph G by a sequence of subdivision is called a**Homeomorph.**Two graph are said to be**Homeomorphic**of each is a homeomorph of same graph.

EXAMPLE



Note1: Kurtatowski**first** graph is not planar.

Note 2: Kuratowski second graph is not planar.

Note-3: Removal of one edge of both Kuratowski graph make them planar.

Note-4:A graph G is planar if and only of G does not contain either of the Kuratowski two graph.

LECTURE 40:

40.1 Detection of Planarity

Given any graph G we reduce it to a simple form through the following simplifying steps.Then it becomes easy to detect whether the graph is planar or not.

Step 1: If G has several components,consider only one component at a time(since G is planar iff each component of its be planar).

Step 2: If any component is separable it would have several blocks.Consider only one block at a time(since component is planar iff each block of its be planar).

Step 3: Remove all self-loops from G addition or deletion of self-loops does not affect the planarity.

Step 4: Keep only one edge between any two vertices by removing all parallel edges between them, since this does not affect the planarity.

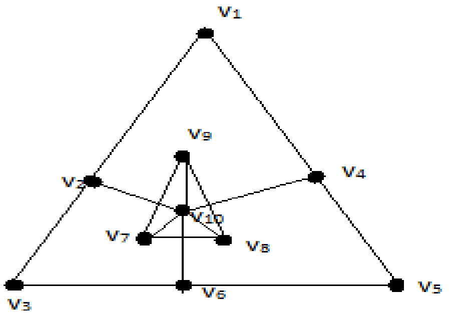
Step 5:If G has two edges having exactly one vertex in common and if this vertex is of degree 2 thenelimination of such vertex does not affect the planarity of G.So remove all such vertex from G.

Step 6: Repeat the above steps so long as we can.After going through the above steps a block or component of G would be look like (1) a single edge,(2)a complete graph with four vertices or (3) a graph with number of vertices ≥ 5 and number of edges ≥ 7.

If it looks like (1) or (2) then it would be planar and if it be look like (3) then it is non-planar.

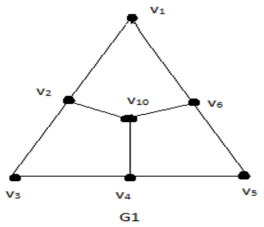
EXAMPLE

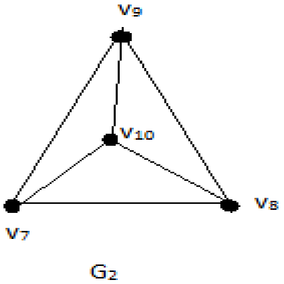
Check whether graph is planar or not



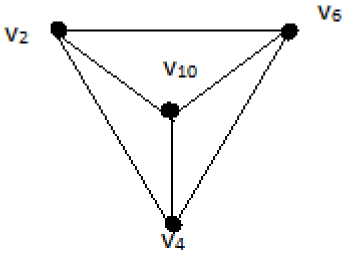
G

We see the vertex v_{10} in G is a cut vertex.So G is separable.so the two blocks are





In G_1 the vertices v_1,v_3 ,v_5 has degree 2 each.These are also common vertices of edges v_1v_3 and $v_1v_6$etc.So in G_1 we remove the vertices v_1,v_3,v_5 and G_1 is converted to



This graph is planar since every edge of this subgraph is drawn without any cross over.Similarly other block G_2 is also planar.

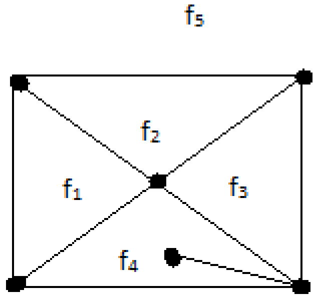
Thus the given graph is planar.

LECTURE 41:

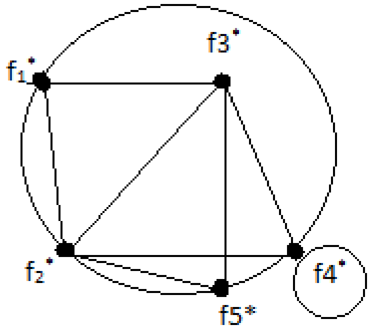
41.1 **Dual graphs**

Introdution

Let G be a plane graph. The dual of G is defined to be the graph G^* constructed as follows. To each region f of G there is a corresponding vertex f^* of G^* and to each edge e of G there is corresponding edge e^* in G^* such that if the edge e occurs on the boundary of the two regions f and g, then the edge e^* joins the corresponding vertices $f^*$and g^*in G^*. If the edge e is a bridge, i.e., the edge e lies entirely in one region f , then the corresponding edge e^* is a loop incident with the vertex f^* in G^*. For example, consider the graph shown in Figure



G



G^*

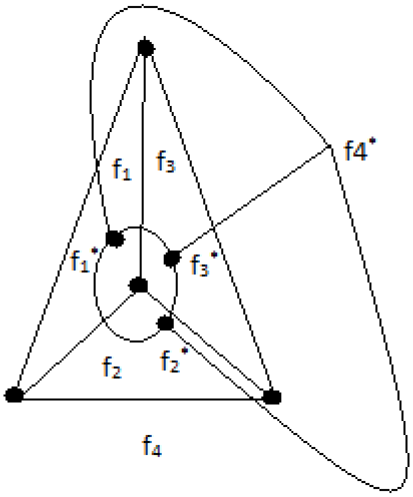
41.2 Theorem

The dual G^* of a plane graph is planar.

Proof

Let G be a plane graph and let G^* be the dual of G. The following construction of G^* shows that G^* is planar.

Place each vertex f^*_k of G^* inside its corresponding region f_i . If the edge ei lies on the boundary of two regions f j and fk of G, join the two vertices f^*_j and f^*_k by the edge e^*_i , drawing so that it crosses the edge e exactly once and crosses no other edge of G



Remarks

Clearly, there is one-one correspondence between the edges of plane graph G and its dual G * with one edge of G * intersecting one edge of G.

41.3 Relation between Planar and Dual graph

Let G be a graph and G* be its dual. The relation between them are as follows:

1. An edge forming a self-loop in G gives a pendant edge in G * (An edge incident on a pendant vertex is called a pendant edge).

2. A pendant edge in G gives a self-loop in G * .

3. Edges that are in series in G produce parallel edges in G *.

4. Parallel edges in G produce edges in series in G * .

5. The number of edges forming the boundary of a region f_i in G is equal to the degree of the corresponding vertex f $^*{}_i$ in G * .

6. Considering the process of drawing a dual G * from G, it is evident that G is a dual of G * . Therefore, instead of calling G *a dual of G, we usually say that G and G *are dual graphs.

7. a)No. of vertices in G* =No. of regions in G

   b) No. of regions in G* =No. of vertices in G

c) No. of regions in G^* =No. of regions in G.

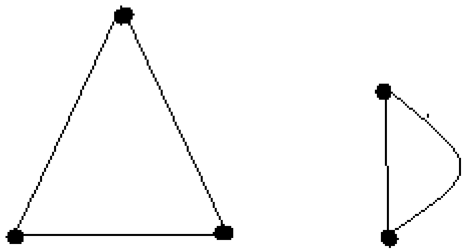
8. G^* is always connected,even when G is disconnected.

Example

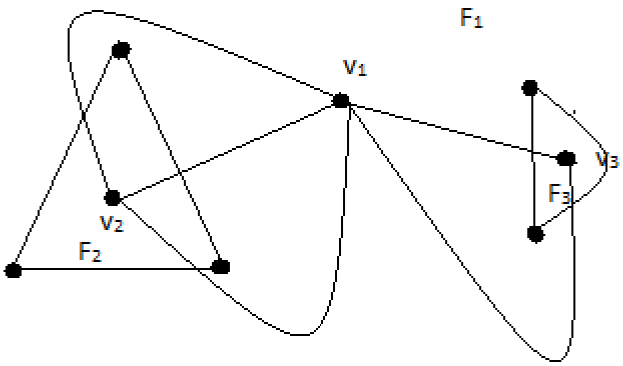
A disconnected graph G is given below.Construct its dual and show that the dual is connected.

Solution

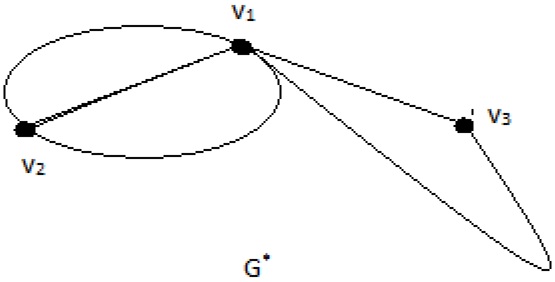
Let us consider the following graph



F_1,F_2,F_3 are three regions determined by G.v_1,v_2,v_3 are three points taken in F_1,F_2,F_3 respectively



.They are joined accordingly and dual of the given graph G^* is formed.


v_1
v_3
v_2
G^*

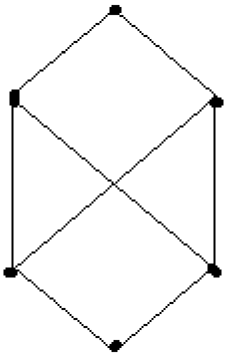
LECTURE 42:

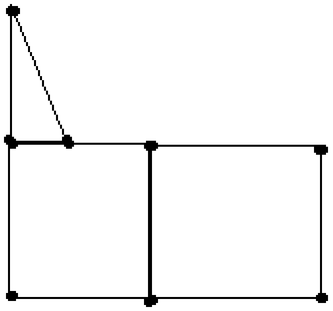
42.1 Graph colouring

Colouring is one of the important branches of graph theory and has attracted the attention of almost all graph theorists, mainly because of the four colour theorem, the details of which can be seen in this Chapter.

42.2 Vertex colouring

A vertex colouring (or simply colouring) of a graph G is a labelling $f : V(G) \rightarrow \{1, 2, ...\}$; the labels called colours, such that no two adjacent vertices get the same colour and each vertex gets one colour. A k-colouring of a graph G consists of k different colours and G is then called k-colourable. . A 2-colourable and a 3-colourable graph are shown in Figure. It follows from this definition that the k-colouring of a graph G(V, E) partitions the vertex set V into k independent sets $V_1, V_2, ..., V_k$ such that $V = V_1 \cup V_2 \cup ... \cup V_k$. The independent sets $V_1, V_2, ..., V_k$ are called the colour classes and the functionf : $V(G) \rightarrow \{1, 2, ..., k\}$ such that $f(v) = i$ for $v \in V_i$, $1 \leq i \leq k$, is called the colour function.





number (chromatic index) of G and is denoted by $\chi(G)$. If $\chi(G) = k$, the graph G is said to be k-chromatic. The minimum number k for which there is a k-colouring of the graph is called the chromatic

We observe that colouring any one of the components in a disconnected graph does not affect the colouring of its other components. Also, parallel edges can be replaced by single edges, since it does not affect the adjacencies of the vertices. Thus, for colouring considerations, we opt only for simple connected graphs.

The following observations are the immediate consequences of the definitions introduced above.

1. A graph is 1-chromatic if and only if it is totally disconnected.

2. A graph having at least one edge is at least 2-chromatic (bichromatic).

3. A graph G having n vertices has $\chi(G) \leq n$

. 4. If H is subgraph of a graph G, then $\chi(H) \leq \chi(G)$.

5. A complete graph with n vertices is n-chromatic, because all its vertices are adjacent. So, $\chi(K_n) = n$ and $\chi(K_n) = 1$. Therefore we see that a graph containing a complete graph of r vertices is at least r-chromatic. For example, every graph containing a triangle is at least 3-chromatic.

6. A cycle of length $n \geq 3$ is 2-chromatic if n is even and 3-chromatic if n is odd. To see this, let the vertices of the cycle be labelled 1, 2, ..., n, and assign one colour to odd vertices and another to even. If n is even, no adjacent vertices get the same colour, if n is odd, the nth vertex and the first vertex are adjacent and have the same colour, therefore need the third colour for colouring.

7. If G1,G2,...,Gr are the components of a disconnected graph G, then

$\chi(G) = \max_{1 \leq i \leq r} \chi(Gi)$.

We note that trees with greater or equal to two vertices are bichromatic as is seen in the following result.

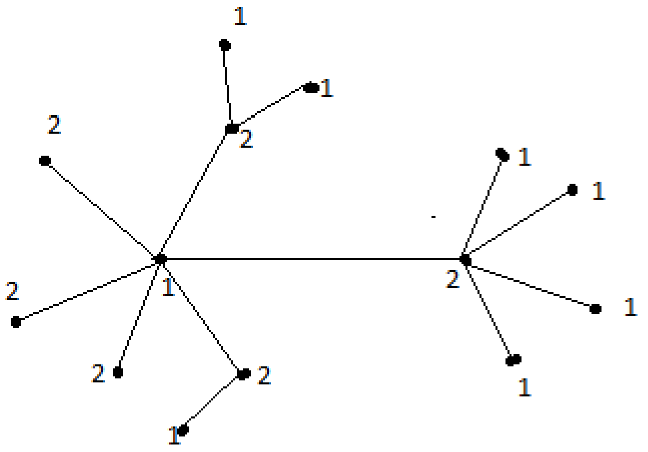
LECTURE 43:

43.1 **Theorem 1**

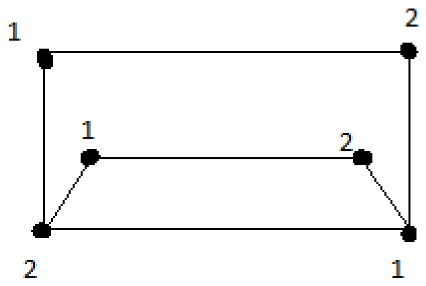
Every tree with $n \geq 2$ vertices is 2-chromatic.

Proof

Let T be a tree with $n \geq 2$ vertices. Consider any vertex v of T and assume T to be rooted at vertex v. Assign colour 1 to v. Then assign colour 2 to all vertices which are adjacent to v. Let $v_1, v_2, \ldots, v_r$ be the vertices which have been assigned colour 2. Now assign colour 1 to all the vertices which are adjacent to $v_1, v_2, \ldots, v_r$. Continue this process till every vertex in T has been assigned the colour. . We observe that in T all vertices at odd distances from v have colour 2, and v and vertices at even distances from v have colour 1. Therefore along any path in T , the vertices are of alternating colours. Since there is one and only one path between any two vertices in a tree, no two adjacent vertices have the same colour. Thus T is coloured with two colours. Hence T is 2-chromatic.



The converse of the above theorem is not true, i. e., every 2-chromatic graph need not be a tree. To see this, consider the graph shown in Figure 7. Clearly, G is 2-chromatic, but is not a tree.



43.2 Theorem 2(Konig)

A graph is bicolourable (2-chromatic) if and only if it has no odd cycles.

Proof

Let G be a connected graph with cycles of only even length and let T be a spanning tree in G. Then, by Theorem 1, T can be coloured with two colours. . Now add the chords to T one by one. As G contains cycles of even length only, the end vertices of every chord get different colours of T . Thus G is coloured with two colours and hence is 2-chromatic. Conversely, let G be bicolourable, that is, 2-chromatic. We prove G has even cycles only. Assume to the contrary that G has an odd cycle. Then by observation (6), G is 3-chromatic, a contradiction. Hence G has no odd cycles.

Corollary

For a graph G, $\chi(G) \geq 3$ if and only if G has an odd cycle. The following result is yet another characterization of 2-chromatic graphs.

43.3 Theorem 3

A nonempty graph G is bicolourable if and only if G is bipartite.

Proof

Let G be a bipartite graph. Then its vertex set V can be partitioned into two nonempty disjoint sets V1 and V2 such that $V = V_1 \cup V_2$. Now assigning colour 1 to all vertices in V1 and colour 2 to all vertices in V2 gives a 2-colouring of G. Since G is nonempty, $\chi(G) = 2$.

Conversely, let G be bicolourable, that is, G has a 2-colouring. Denote by V_1 the set of all those vertices coloured 1 and by V_2 the set of all those vertices coloured 2. Then no two vertices in V_1 are adjacent and no two vertices in V_2 are adjacent. Thus any edge in G joins a vertex in V1 and a vertex in V2. Hence G is bipartite with bipartition $V = V_1 \cup V_2$.

43.4 Theorem 4

For any graph G, $\chi(G) \leq 4(G) + 1$.

Proof

Let G be any graph with n vertices. To prove the result, we induct on n. For n = 1,G = K1 and $\chi(G) = 1$ and $\Delta(G) = 0$. Therefore the result is true for n = 1.

Assume that the result is true for all graphs with n−1 vertices and therefore by induction hypothesis, $\chi(G) \leq \Delta(G-v)+1$. This shows that G−v can be coloured by using $\Delta(G-v)+1$ colours. Since $\Delta(G)$ is the maximum degree of a vertex in G, vertex v has at most $\Delta(G)$ neighbours in G. Thus these neighbours use up at most $\Delta(G)$ colours in the colouring of G−v.

If $\Delta(G) = \Delta(G-v)$, then there is at least one colour not used by v's neighbours and that can be used to colour v giving a $\Delta(G) +1$ colouring for G.

In caseΔ (G) 6=Δ (G − v), thenΔ (G − v) <Δ (G). Therefore, using a new colour for v, we have aΔ (G−v) +2 colouring of G and clearly$\Delta(G-v) +2 \leq \Delta(G) +1$. Hence in both cases, it follows that $\chi(G) \leq \Delta(G) +1$.

43.5 Theorem 5-Brook's Theorem-

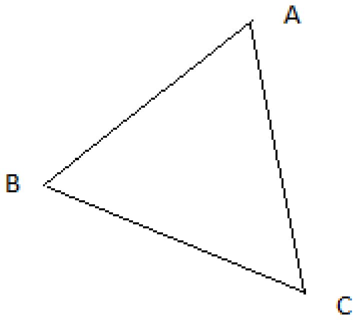
If G is a connected graph which is neither complete nor an odd cycle, then $\chi(G) \leq \Delta(G)$.

LECTURE 44:

44.1 Chromatic Polynomials(Chromials)

Let f(G,x) be the number of different colourings of a graph G with x or fewer colours.Then it can be shown that f(G,x) will be a polynomial of x.This f(G,x) is called chromatic polynomial of G.

Consider the following graph G



And the x number of colours C_1,C_2,C_3,…..C_x.Since A,B,C are pairwise adjacent to each to each other,each should them should be assigned with distinct colour.Now A can be assigned with x number of colour.Therefore G may be colored in x(x-1)(x-2) way or fewer colours.Thus f(G,x)= x(x-1)(x-2)=x^3-3x^2+2x is the chromatic polynomial of G.

44.2 Theorem-5

Chromatic polynomial for complete graph with n vertices (K_n) is f(K_n,x)=x(x-1)(x-2)….(x-n+1).

Proof

Let the vertices of K_n be v_1,v_2,v_3,…..v_n and we have x number of colours C_1,C_2,C_3,….C_x.Obviouslyx≥n.Since the vertices are pairwise adjacent to each other of them should be assigned with distinct colours.Now v_1 can be assigned with x number of colours;v_2 can be assigned with x-1 number of colours.Similarly,v_3 can be assigned x-2 number of colours.Continuing this way v_n can be assigned with x-n+1 number of colours.

So,by principle of counting ,K_n can be coloured in x(x-1)(x-2)….(x-n+1) ways or fewer.

44.3 Theorem-6

The chromatic polynomial of a tree with vertices is $x(x-1)^{n-1}$.

Proof

Let G be a tree wuth n vertices. We shall prove its chromatic polynomial is $x(x-1)^{n-1}$.We prove that by the method of induction.For n=1 the result is obvious.For n=2 G has two vertices v_1 and v_2.v_1 may assigned with x number of colours from a collection of x colours.Since,v_2 is adjacent to v_1,v_2 can be assigned with x-1 colours

Therefore, f(G,x)=x(x-1) proving the result for n=2.Let the result hold for n=m.We shall prove that it hold for n=m+1.Since G is a tree it must have at least once pendant vertex say v.Then the graph G-v is also a tree with m vertex.Then by hypothesis its $f(G-V,x)=x(x-1)^{m-1}$.Since v is pendant so its adjacent to only one vertex say v_1.Since v cannot be assigned with the colour of v_1 so v can be assigned with x-1 number colours.So by the principle of counting

$f(G,x) =x(x-1)^{m-}1(x-1)=x(x-1)^{m}$.

44.4 Theorem -7

The chromatic polynomial is a polynomial.

Proof

Let G be graph with n vertices.Let G be coloured with i number of ofcolours in c_i different ways.Since i colours can be choosen from x number of colours in ${}^xC_i$ ways ,there are ${}^xC_i$ways of colouring with i number of colours taken from x number of colours.Since G has n vertices so it is not possible to use more than x colours for colouring of G.Therefore i=1,2,….n.

Thus the chromatic polynomial is

$f(G,x)=c_1\,{}^xC_1 +c_2\,{}^xC_2 +…..+c_n{}^xC_n$

each c_i ha sto be evaluated individually for the given graph but it can be mentioned that each c_i is positive integer.Again${}^xC_i=x(x-1)(x-2)0….(x-i+1)/i!$ is a polynomial of x.So f(G,x) is a polynomial of degree at most n.

44.5 Decomposition Theorem

Let v_1 andv_2 be two non-empty adjacent vertices of a simple graph G.Let G'be a graph obtained by adding an edge between v_1 and v_2. Let G'' be a graph obtained from G merging v_1 and v_2 together.Then $f(G,x)= f(G',x)+ f(G'',x)$.

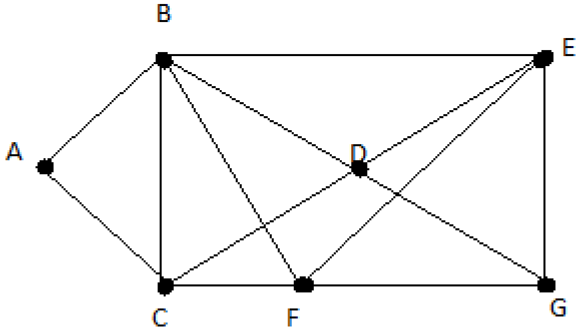
44.6 Four Colour Theorem

Every planar graph has a chromatic number four or less.

44.7 Five Colour Theorem

A planar graph can be coloured with five colours.

EXAMPLEFind the chromatic number of the following graph



SOLUTION

Since the vertices A,B,C form a triangle so three colours are must for them.Say Red to A,Blue to B,Green to C.

D is adjacent to B and C.Therefore D is assigned to D and C so F is asssigned with Blue.Similarly E is assigned with Green.Finally G is assigned with Red.

Thus the graph is 3 vertex colourable but not 2.So chromatic number of this graph is 3.