SEMESTER IV

MAT256	PROBABILITY AND STATISTICAL MODELLING	Category	L	T	P	Credit	Year of Introduction
		BSC	3	1	0	4	2019

Preamble: Study of this course provides the learners a clear understanding of fundamental concepts in probability and statistics. This course covers the modern theory of probability and statistics, important models of sampling, techniques of hypothesis testing and correlation & regression. The course helps the learners to find varied applications in engineering and science like disease modelling, climate prediction and computer networks.

Prerequisite: A sound knowledge in Calculus.

Mapping of course outcomes with program outcomes

CO1	Explain the concept, properties and important models of discrete random variables and use them to analyze suitable random phenomena**(Cognitive Knowledge Level: Apply)**
CO2	Summarize the properties and relevant models of continuous random variables and use them to analyze suitable random phenomena**(Cognitive Knowledge Level: Apply)**
CO3	Make use of concepts of sampling and theory of estimation to solve application level problems **(Cognitive Knowledge Level: Apply)**
CO4	Organize the basic concepts in hypothesis testing and develop decision procedures for the most frequently encountered testing problems**(Cognitive Knowledge Level: Apply)**
CO5	Build statistical methods like correlation and regression analysis to interpret experimental data **(Cognitive Knowledge Level: Apply)**

Mapping of course outcomes with program outcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	✓	✓	✓	✓								✓
CO2	✓	✓	✓	✓								✓
CO3	✓	✓	✓	✓								✓
CO4	✓	✓	✓	✓								✓
CO5	✓	✓	✓	✓								✓
CO6	✓	✓	✓	✓								✓

Abstract POs defined by National Board of Accreditation			
PO#	**Broad PO**	**PO#**	**Broad PO**
PO1	Engineering Knowledge	**PO7**	Environment and Sustainability
PO2	Problem Analysis	**PO8**	Ethics
PO3	Design/Development of solutions	**PO9**	Individual and team work
PO4	Conduct investigations of complex problems	**PO10**	Communication
PO5	Modern tool usage	**PO11**	Project Management and Finance
PO6	The Engineer and Society	**PO12**	Lifelong learning

Assessment Pattern

Bloom's Category	Continuous Assessment Tests		End Semester Examination Marks (%)
	Test 1 (%)	Test 2 (%)	
Remember	**30**	**30**	**30**
Understand	**30**	**30**	**30**
Apply	**40**	**40**	**40**
Analyze			
Evaluate			
Create			

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
150	**50**	**100**	**3**

Continuous Internal Evaluation Pattern:

Attendance	**10 marks**
Continuous Assessment Tests(Average of Internal Tests1&2)	**25 marks**
Continuous Assessment Assignment	**15 marks**

Internal Examination Pattern

Each of the two internal examinations has to be conducted out of 50 marks. First series test shall be preferably conducted after completing the first half of the syllabus and the second series test shall be preferably conducted after completing remaining part of the syllabus. There will be two parts: Part A and Part B. Part A contains 5 questions (preferably, 2 questions each from the completed modules and 1 question from the partly completed module), having 3 marks for each question adding up to 15 marks for part A. Students should answer all questions from Part A. Part B contains 7 questions (preferably, 3 questions each from the completed modules and 1 question from the partly completed module), each with 7 marks. Out of the 7 questions, a student should answer any5.

End Semester Examination Pattern:

There will be two parts; Part A and Part B. Part A contains 10 questions with 2 questions from each module, having 3 marks for each question. Students should answer all questions. Part B contains 2 full questions from each module of which student should answer any one. Each question can have maximum 2 sub-divisions and carries 14 marks.

Syllabus

Module-1 (Discrete probability distributions)

Discrete random variables and their probability distributions, Expectation, mean and variance, Binomial distribution, Poisson distribution, Poisson approximation to the binomial distribution, Discrete bivariate distributions, marginal distributions, Independent random variables, Expectation ,multiple random variables.

Module - 2(Continuous probability distributions)

Continuous random variables and their probability distributions, Expectation, mean and variance, Uniform, exponential and normal distributions, Continuous bivariate distributions, marginal distributions, Independent random variables. Expectation-multiple random variables, independent and identically distributed (i.i.d) random variables and Central limit theorem (Proof not required).

Module - 3(Sampling Techniques)

Need for Sampling, Some Fundamental Definitions, Important Sampling Distributions, Sampling Theory, Sandler's A-test, Concept of Standard Error, Estimation, Estimating the Population Mean(μ), Estimating Population Proportion, Sample Size and its Determination, Determination of

Sample Size through the Approach Based on Precision Rate and Confidence Level, Determination of Sample Size through the Approach Based on Bayesian Statistics

Module– 4**(Testing of Hypothesis)**

Hypothesis and Test Procedures, Tests about a population mean, Tests concerning a population proportion, p-values, Single factor ANOVA, F-test, Multiple comparisons in ANOVA, Two factor ANOVA

Module - 5 (Correlation and Regression Analysis)

Simple Linear Regression Model, Estimating model parameters, Correlation, Non-Linear and multiple regression, Assessing Model Adequacy, Regression with transformed values, Polynomial Regression, Multiple Regression Analysis

Text Books

1. Jay L. Devore, Probability and Statistics for Engineering and the Sciences, 8th edition, Cengage, 2012
2. Research Methodology: Methods and Techniques: C.R. Kothari, New Age International Publishers

Reference Books

1. HosseinPishro-Nik, Introduction to Probability, Statistics and Random Processes, Kappa Research, 2014 (Also available online at www.probabilitycourse.com)
2. Sheldon M. Ross, Introduction to probability and statistics for engineers and scientists, 4th edition, Elsevier, 2009.
3. T. VeeraRajan, Probability, Statistics and Random processes, Tata McGraw-Hill,2008
4. B.S. Grewal, Higher Engineering Mathematics, Khanna Publishers, 36 Edition, 2010
5. Levin R.I. and Rubin D.S., Statistics for Management, 7th edition, Prentice Hall of India Pvt. Ltd., New Delhi, 2001.
6. Srivastava TN, Shailaja Rego, Statistics for Management, Tata McGraw Hill, 2008.
7. Anand Sharma, Statistics for Management, Himalaya Publishing House, Second Revised edition, 2008.
8. Goon A.M., Gupta M.K. and Dasgupta B. (2002): Fundamentals of Statistics, Vol. I & II, 8th Edition. The World Press, Kolkata.
9. Miller, Irwin and Miller, Marylees (2006): John E. Freund's Mathematical Statistics with Applications, (7th Edition.), Pearson Education, Asia.
10. Sampling of Populations: Methods and Applications (2008): Paul S. Levy , Stanley Lemeshow (Fourth Edition), John Wiley &Sons

Course Level Assessment Questions

Course Outcome1 (CO1):

1. Organizers of a concert are limiting tickets sales to a maximum of 4 tickets per customer. Let T be the number of tickets purchased by a random customer. Here is the probability distribution of T:

T=#of tickets	1	2	3	4
P(T)	0.1	0.3	0.2	0.4

Calculate the expected value of T.

2. X is a binomial random variable B (n, p) with n = 100 and p= 0.1. How would you approximate it by a Poisson random variable?
3. Three balls are drawn at random without replacement from a box containing 2 white,3 red and 4 black balls. If X denotes the number of white balls drawn and Y denotes the number of red balls drawn, find the joint probability distribution of (X, Y).

Course Outcome 2(CO2):

1. What can you say about P(X = a) for any real number a when X is a (i) discrete random variable? (ii) continuous random variable?
2. Let X be a random variable with PDF given by

$$f_X(x)= \begin{cases} cx^2 & |x| \le 1 \\ 0 & \text{Otherwise} \end{cases}$$

   a. Find the constant c.
   b. Find E(X) and Var(X).
   c. Find P(X≥1/2).
3. A string, 1 meter long, is cut into two pieces at a random point between its ends. What is the probability that the length of one piece is at least twice the length of the other?

Course Outcome 3(CO3):

1. In a random selection of 64 of the 2400 intersections in a small city, the mean number of scooter accidents per year was 3.2 and the sample standard deviation was 0.8.
   (a) Make an estimate of the standard deviation of the population from the sample standard deviation.
   (b) Work out the standard error of mean for this finite population.
   (c) If the desired confidence level is 0.90, what will be the upper and lower limits of the confidence interval for the mean number of accidents per intersection per year?

2. Suppose a certain hotel management is interested in determining the percentage of the hotel's guests who stay for more than 3 days. The reservation manager wants to be 95 per cent confident that the percentage has been estimated to be within ± 3% of the true value. What is the most conservative sample size needed for this problem?
3. 500 articles were selected at random out of a batch containing 10000 articles and 30 were found defective. How many defective articles would you reasonably expect to find in the whole batch?

Course Outcome 4(CO4):

1. A manufacturer of sprinkler systems used for fire protection in office buildings claims that the true average system-activation temperature is 130°F. A sample of n=9 systems, when tested, yields a sample average activation temperature of 131.08°F. If the distribution of activation times is normal with standard deviation 1.5°F, does the data contradict the manufacturer's claim at significance level α=0.01?
2. Let m denote the true average radioactivity level (picocuries per liter). The value 5 pCi/L is considered the dividing line between safe and unsafe water. Would you recommend testing H_0: μ =5 versus H_a: μ >5 or H_0: μ =5 versus H_a: μ <5 ? Explain your reasoning.
3. Pairs of P-values and significance levels, a, are given. For each pair, state whether the observed *P*-value would lead to rejection of H_0 at the given significance level.
   a. P-value=0.084, α=0.05
   b. P-value=0.003, α=0.001

Course Outcome 5 (CO5):

1.Calculate and interpret the correlation coefficient of the two variables below.

Person	Hand	Height
A	17	150
B	15	154
C	19	169
D	17	172
E	21	175

2. You are told that a 95% CI for expected lead content when traffic flow is 15, based on a sample of n=10 observations is (462.1, 597.7). Calculate a CI with confidence level 99% for expected lead content when traffic flow is 15.
3. A trucking company considered a multiple regression model for relating the dependent variable y=total daily travel time for one of its drivers (hours) to the predictors x_1=distance travelled (miles) and x_2=the number of deliveries made. Suppose that the model equation is $Y = \mathbf{-0.800 + 0.060\, x_1 + 0.900 x_2 + \varepsilon}$. What is the mean value of travel time when distance traveled is 50 miles and three deliveries are made?

Model Question Paper

QP CODE:

Reg No: _____

Name: _____ **PAGES : 4**

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

FOURTH SEMESTER B.TECH DEGREE EXAMINATION, MONTH & YEAR

Course Code: MAT256

Course Name: Probability and Statistical Modelling

Max.Marks:100 **Duration: 3 Hours**

PART A

Answer All Questions. Each Question Carries 3 Marks

1. Let X denote the number that shows up when an unfair die is tossed. Faces 1 to 5 of the die are equally likely, while face 6 is twice as likely as any other. Find the probability distribution, mean and variance of X.

2. An equipment consists of 5 components each of which may fail independently with probability 0.15. If the equipment is able to function properly when at least 3 of the components are operational, what is the probability that it functions properly?

3. A random variable has a normal distribution with standard deviation 10. If the probability that it will take on a value less than 82.5 is 0.82, what is the probability that it will take on a value more than 58.3?

4. X and Y are independent random variables with X following an exponential distribution with parameter μ and Y following and exponential distribution with parameterλ. Find P (X+ Y$\leqslant$ 1).

5. Discuss the difference between F-distribution and Chi-square distribution.

6. From a random sample of 36 New Delhi civil service personnel, the mean age and the sample standard deviation were found to be 40 years and 4.5 years

respectively. Construct a 95 per cent confidence interval for the mean age of civil servants in New Delhi.

7. A sample of 50 lenses used in eyeglasses yields a sample mean thickness of 3.05 mm and a sample standard deviation of .34 mm. The desired true average thickness of such lenses is 3.20 mm. Does the data strongly suggest that the true average thickness of such lenses is something other than what is desired? Test using α=0.05.

8. A random sample of 110 lightning flashes in a certain region resulted in a sample average radar echo duration of 0.81 sec and a sample standard deviation of 0.34 sec. Calculate a 99% (two-sided) confidence interval for the true average echo duration m, and interpret the resulting interval.

9. Let the test statistic T have a t distribution when H_0 is true. Give the significance level for the following situation H_a: $\mu > \mu_0$, df=15, rejection region $t \geq 3.733$.

10. Calculate the regression coefficient and obtain the lines of regression for the following data **(10x3=30)**

X	1	2	3	4	5	6	7
Y	9	8	10	12	11	13	14

Part B

(Answer any one question from each module. Each question carries 14 Marks)

11. (a) The probability mass function of a discrete random variable is P(x) = kx ; x = 1,2,3 where k is positive constant. Find (i) the value of k (ii) P(X $\leq$2) (iii) E[X] (iv) var(1-X). **(7)**

(b) Find the mean and variance of a binomial random variable **(7)**

OR

12. (a) Accidents occur at an intersection at a Poisson rate of 2 per day. What is the probability that there would be no accidents on a given day? What is the probability that in January there are at least 3 days (not necessarily consecutive) without any accidents? **(7)**

(b) One fair die is rolled. Let X denote the number on the die and Y = 0 or 1, according as the die shows an even number or odd number. Find (i) the joint probability distribution of X and Y, (ii) the marginal distributions. (iii) Are X and Y independent? **(7)**

13. (a) The IQ of an individual randomly selected from a population is a normal distribution with mean 100 and standard deviation 15. Find the probability that an individual has IQ (i) above 140 (ii) between 120 and 130. **(7)**

(b) A continuous random variable X is uniformly distributed with mean 1 and variance 4/3. Find $P(X < 0)$? **(7)**

OR

14. (a) The joint density function of random variables X and Y is given by **(7)**

$$f(x, y) = \begin{cases} e^{-(x+y)}, & x>0, y>0 \\ 0 & \text{otherwise} \end{cases}$$

Find $P(X + Y \leq 1)$. Are X and Y independent? Justify

(b) The lifetime of a certain type of electric bulb may be considered as an exponential random variable with mean 50 hours. Using central limit theorem, find the approximate probability that 100 of these electric bulbs will provide a total of more than 6000 hours of burning time. **(7)**

15. (a) A market research survey in which 64 consumers were contacted and states that 64 percent of all consumers of a certain product were motivated by the product's advertising. Find the confidence limits for the proportion of consumers motivated by advertising in the population, given a confidence level equal to 0.95. **(7)**

(b) Determine the size of the sample for estimating the true weight of the cereal containers for the universe with N = 5000 on the basis of the following information: **(7)**

(i) the variance of weight = 4 ounces on the basis of past records.

(ii) estimate should be within 0.8 ounces of the true average weight with 99% probability.

OR

16. (a) The foreman of *ABC* mining company has estimated the average quantity of iron ore extracted to be 36.8 tons per shift and the sample standard deviation to be 2.8 tons per shift, based upon a random selection of 4 shifts. Construct a 90 percent confidence interval around this estimate. **(7)**

(b) What should be the size of the sample if a simple random sample from a population of 4000 items is to be drawn to estimate the percent defective within 2 per cent of the true value with 95.5 per cent probability? What would be the size of the sample if the population is assumed to be infinite in the given case? **(7)**

17. The calibration of a scale is to be checked by weighing a10-kg test specimen 25 times. Suppose that the results of different weighings are independent of one another and that the weight on each trial is normally distributed with σ =0.200kg. Let μ denote the true average weight reading on the scale.
(a)What hypotheses should be tested? **(7)**
(b)Suppose the scale is to be recalibrated if either $\bar{x} \geq 10.1032$ or $\bar{x} \leq 0.8968$. What is the probability that recalibration is carried out when it is actually unnecessary? **(7)**

OR

18. (a) Lightbulbs of a certain type are advertised as having an average lifetime of 750 hours. The price of these bulbs is very favorable, so a potential customer has decided to go ahead with a purchase arrangement unless it can be conclusively demonstrated that the true average lifetime is smaller than what is advertised. A random sample of 50 bulbs was selected, the lifetime of each bulb determined, and the appropriate hypotheses were tested using Minitab, resulting in the accompanying output.

Variable	N	Mean	StDev	SEMean	Z	P-Value
lifetime	50	738.44	38.20	5.40	-2.14	0

What conclusion would be appropriate for a significance level of 0.05? A significance level of 0.01? What significance level and conclusion would you recommend**?** **(7)**

(b) The recommended daily dietary allowance for zinc among males older than age 50 years is 15 mg/day. The article "Nutrient Intakes and Dietary Patterns of Older Americans: A National Study" reports the following summary data on intake for a sample of males age 65–74 years: n=115, $\bar{x}$ =11.3, and s=6.43. Does this data indicate that average daily zinc intake in the population of all males ages 65–74 falls below the recommended allowance? **(7)**

19. The flow rate y (m^3/min) in a device used for air-quality measurement depends on the pressure drop x (inches of water) across the device's filter. Suppose that for x values between 5 and 20, the two variables are related according to the simple linear regression model with true regression line
$y = -0.12 + 0.095x$

(a)What is the expected change in flow rate associated with a 1 inch increase in pressure drop? Explain. **(7)**

(b)What change in flow rate can be expected when pressure drop decreases by 5 inches? **(7)**

OR

20. Suppose that in a certain chemical process the reaction time y (hr) is related to the temperature (°F) in the chamber in which the reaction takes place according to the simple linear regression model with equation $y = 5.00 - 0.01x$ and $\sigma = 0.075$

(a)What is the expected change in reaction time for a 1°F increase and 10°F increase in temperature? **(7)**

(b)What is the expected reaction time when temperature is 200°F and 250°F? **(7)**

Teaching Plan

No	Contents	No. of Lecture Hours (45 hrs)
Module 1- (Discrete Probability distributions) (9 hours)		
1.1	Discrete random variables	1 hour
1.2	Probability Distributions	1 hour
1.3	Expectation, mean and variance	1 hour
1.4	Binomial distribution	1 hour
1.5	Poisson distribution	1 hour
1.6	Poisson approximation to binomial Distribution	1 hour
1.7	Discrete bivariate distributions	1 hour
1.8	Marginal distributions, Independent Random variables	1 hour
1.9	Expectation-multiple random variables	1 hour
Module-2 Continuous Probability distributions(9 hours)		
2.1	Continuous random variables and probability distributions	1 hour

2.2	Expectation, mean and variance	1 hour
2.3	Uniform distributions	1 hour
2.4	Exponential Distribution	1 hour
2.5	Normal distribution	1 hour
2.6	Continuous Bivariate distributions	1 hour
2.7	Marginal distributions, Independent random variables	1 hour
2.8	Expectation-multiple random variables, i.i.d random variables	1 hour
2.9	Central limit theorem.	1 hour
Module-3 (Sampling Techniques) (9 hours)		
3.1	Need for Sampling	1 hour
3.2	Some fundamental Definitions, Important Sampling Distributions	1 hour
3.3	Sampling Theory, Sandler's A-test	1 hour
3.4	Concept of Standard Error, Estimation , Estimating the Population Mean(μ)	1 hour
3.5	Estimating Population Proportion	1 hour
3.6	Sample Size and its Determination	1 hour
3.7	Determination of Sample Size through the Approach Based on Precision Rate and Confidence Level	1 hour
3.8	Determination of Sample Size through the Approach Based on Bayesian Statistics	1 hour
3.9	Determination of Sample Size through the Approach Based on Bayesian Statistics(continued)	1 hour
Module-4 (Testing of Hypothesis) (9 hours)		
4.1	Null and alternate Hypothesis	1 hour
4.2	Test Procedures	1 hour
4.3	Test Tests about a population mean	1 hour
4.4	Tests concerning a population proportion	1 hour
4.5	p-values	1 hour

4.6	Single factor ANOVA	1 hour
4.7	F-Test	1 hour
4.8	Multiple comparisons in ANOVA	1 hour
4.9	Two factor ANOVA	1 hour
Module-5 (Correlation and Regression Analysis) (9 hours)		
5.1	Simple Linear Regression Model(Lecture 1)	1 hour
5.2	Simple Linear Regression Model(Lecture 2)	1 hour
5.3	Estimating model parameters	1 hour
5.4	Correlation	1 hour
5.5	Non-Linear and multiple regression	1 hour
5.6	Assessing Model Adequacy	1 hour
5.7	Regression with transformed values	1 hour
5.8	Polynomial Regression	1 hour
5.9	Multiple Regression Analysis	1 hour

CST202	COMPUTER ORGANISATION AND ARCHITECTURE	CATEGORY	L	T	P	CREDIT	YEAR OF INTRODUCTION
		PCC	3	1	0	4	2019

Preamble:

The course is prepared with the view of enabling the learners capable of understanding the fundamental architecture of a digital computer. Study of Computer Organization and Architecture is essential to understand the hardware behind the code and its execution at physical level by interacting with existing memory and I/O structure. It helps the learners to understand the fundamentals about computer system design so that they can extend the features of computer organization to detect and solve problems occurring in computer architecture.

Prerequisite : Topics covered under the course Logic System Design (CST 203)

Course Outcomes: After the completion of the course the student will be able to

CO#	CO
CO1	Recognize and express the relevance of basic components, I/O organization and pipelining schemes in a digital computer **(Cognitive knowledge: Understand)**
CO2	Explain the types of memory systems and mapping functions used in memory systems **(Cognitive Knowledge Level: Understand)**
CO3	Demonstrate the control signals required for the execution of a given instruction **(Cognitive Knowledge Level: Apply))**
CO4	Illustrate the design of Arithmetic Logic Unit and explain the usage of registers in it (**Cognitive Knowledge Level: Apply**)
CO5	Explain the implementation aspects of arithmetic algorithms in a digital computer **(Cognitive Knowledge Level:Apply)**
CO6	Develop the control logic for a given arithmetic problem (**Cognitive Knowledge Level: Apply**)

Mapping of course outcomes with program outcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	✓	✓	✓	✓								✓
CO2	✓	✓	✓	✓						✓		✓
CO3	✓	✓	✓	✓						✓		✓
CO4	✓	✓	✓	✓						✓		✓
CO5	✓	✓	✓							✓		✓
CO6	✓	✓	✓	✓								✓

Abstract POs defined by National Board of Accreditation			
PO#	**Broad PO**	**PO#**	**Broad PO**
PO1	Engineering Knowledge	**PO7**	Environment and Sustainability
PO2	Problem Analysis	**PO8**	Ethics
PO3	Design/Development of solutions	**PO9**	Individual and team work
PO4	Conduct investigations of complex problems	**PO10**	Communication
PO5	Modern tool usage	**PO11**	Project Management and Finance
PO6	The Engineer and Society	**PO12**	Life long learning

Assessment Pattern

Bloom's Category	Continuous Assessment Tests		End Semester Examination Marks (%)
	Test1 (%)	Test2 (%)	
Remember	20	20	30
Understand	40	40	30
Apply	40	40	40
Analyze			

Evaluate			
Create			

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
150	**50**	**100**	**3 hours**

Continuous Internal Evaluation Pattern:

Attendance : 10 marks

Continuous Assessment Tests : 25 marks

Continuous Assessment Assignment : 15 marks

Internal Examination Pattern:

Each of the two internal examinations has to be conducted out of 50 marks

First Internal Examination shall be preferably conducted after completing the first half of the syllabus and the Second Internal Examination shall be preferably conducted after completing remaining part of the syllabus.

There will be two parts: Part A and Part B. Part A contains 5 questions (preferably, 2 questions each from the completed modules and 1 question from the partly covered module), having 3 marks for each question adding up to 15 marks for part A. Students should answer all questions from Part A. Part B contains 7 questions (preferably, 3 questions each from the completed modules and 1 question from the partly covered module), each with 7 marks. Out of the 7 questions in Part B, a student should answer any 5.

End Semester Examination Pattern:

There will be two parts; Part A and Part B. Part A contains 10 questions with 2 questions from each module, having 3 marks for each question. Students should answer all questions. Part B contains 2 questions from each module of which a student should answer any one. Each question can have maximum 2 sub-divisions and carries 14 marks.

Syllabus

Module 1

Basic Structure of computers – functional units - basic operational concepts - bus structures. Memory locations and addresses - memory operations,Instructions and instruction sequencing , addressing modes.

Basic processing unit – fundamental concepts – instruction cycle – execution of a complete instruction - single bus and multiple bus organization

Module 2

Register transfer logic: inter register transfer – arithmetic, logic and shift micro operations.

Processor logic design: - processor organization – Arithmetic logic unit - design of arithmetic circuit - design of logic circuit - Design of arithmetic logic unit - status register – design of shifter - processor unit – design of accumulator.

Module 3

Arithmetic algorithms: Algorithms for multiplication and division (restoring method) of binary numbers. Array multiplier , Booth's multiplication algorithm.

Pipelining: Basic principles**,** classification of pipeline processors, instruction and arithmetic pipelines (Design examples not required), hazard detection and resolution.

Module 4

Control Logic Design: Control organization – Hard_wired control-microprogram control – control of processor unit - Microprogram sequencer,micro programmed CPU organization - horizontal and vertical micro instructions.

Module 5

I/O organization: accessing of I/O devices – interrupts, interrupt hardware -Direct memory access.

Memory system: basic concepts – semiconductor RAMs. memory system considerations – ROMs, Content addressable memory, cache memories - mapping functions.

Text Books

1. Hamacher C., Z. Vranesic and S. Zaky, Computer Organization ,5/e, McGraw Hill, 2011
2. Mano M. M., Digital Logic & Computer Design, PHI, 2004
3. KaiHwang, Faye Alye Briggs, Computer architecture and parallel processing McGraw-Hill, 1984

Reference Books

1. Mano M. M., Digital Logic & Computer Design, 3/e, Pearson Education, 2013.
2. Patterson D.A. and J. L. Hennessy, Computer Organization and Design, 5/e, Morgan Kaufmann Publishers, 2013.
3. William Stallings, Computer Organization and Architecture: Designing for Performance, Pearson, 9/e, 2013.
4. Chaudhuri P., Computer Organization and Design, 2/e, Prentice Hall, 2008.
5. Rajaraman V. and T. Radhakrishnan, Computer Organization and Architecture, Prentice Hall, 2011

Sample Course Level Assessment Questions

Course Outcome1(CO1): Which are the registers involved in a memory access operation and how are they involved in it?

Course Outcome 2(CO2): Explain the steps taken by the system to handle a write miss condition inside the cache memory.

Course Outcome 3(CO3): Generate the sequence of control signals required for the execution of the instruction MOV [R1],R2 in a threebus organization.

Course Outcome 4(CO4): Design a 4-bit combinational logic shifter with 2 control signals H0 and H1 that perform the following operations :

H1	H0	Operation
0	0	Transfer 1's to all output line
0	1	No shift operation
1	0	Shift left
1	1	Shift right

Course Outcome 5(CO5): Explain the restoring algorithm for binary division. Also trace the algorithm to divide $(1001)_2$ by $(11)_2$

Course Outcome 6(CO6): Design a software control logic based on microprogramed control to perform the addition of 2 signed numbers represented in sign magnitude form.

Model Question Paper

QP CODE: PAGES:2

Reg No:_____

Name:_____

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

THIRD SEMESTER B.TECH DEGREE EXAMINATION, MONTH & YEAR

Course Code: CST 202

Course Name: Computer organisation and architecture

Max.Marks:100 **Duration: 3 Hours**

PART A

Answer all Questions. Each question carries 3 Marks

1. Give the significance of instruction cycle.
2. Distinguish between big endian and little endian notations. Also give the significance of these notations.
3. Compare I/O mapped I/O and memory mapped I/O.
4. Give the importance of interrupts in I/O interconnection.
5. Justify the significance of status register.
6. How does the arithmetic circuitry perform logical operations in an ALU.
7. Illustrate divide overflow with an example.
8. Write notes on arithmetic pipeline.
9. Briefly explain the role of micro program sequence.
10. Differentiate between horizontal and vertical micro instructions.

Part B

Answer any one Question from each module. Each question carries 14 Marks

11.

11.(a) What is the significance of addressing modes in computer architecture.

(4)

11.(b) Write the control sequence for the instruction DIV R1,[R2] in a three bus structure.

(10)

OR

12. Explain the concept of a single bus organization with help of a diagram. Write the control sequence for the instruction ADD [R1],[R2].

(14)

13. Explain various register transfer logics.

(14)

OR

14.

14.(a) Design a 4 bit combinational logic shifter with 2 control signals H1 and H2 that perform the following operations (bit values given in parenthesis are the values of control variable H1 and H2 respectively.) : Transfer of 0's to S (00), shift right (01), shift left (10), no shift (11).

(5)

14.(b) Design an ALU unit which will perform arithmetic and logic operation with a given binary adder.

(9)

15.

15.(a) Give the logic used behind Booth's multiplication algorithm.

(4)

15.(b) Identify the appropriate algorithm available inside the system to perform the multiplication between -14 and -9. Also trace the algorithm for the above input.

(10)

OR

16.

16.(a) List and explain the different pipeline hazards and their possible solutions

(10)

16.(b) Design a combinational circuit for 3x2 multiplication. (4)

17. Design a hardwared control unit used to perform addition/subtraction of 2 numbers represented in sign magnitude form. (14)

OR

18. Give the structure of the micro program sequencer and its role in sequencing the micro instructions. (14)

19.

19.(a) Explain the different ways in which interrupt priority schemes can be implemented (10)

19.(b) Give the structure of SRAM cell. (4)

OR

20.

20.(a) Explain the various mapping functions available in cache memory. (9)

20.(b) Briefly explain content addressable memory. (5)

TEACHING PLAN		
No	**Contents**	**No of Lecture Hrs**
Module 1 : (Basic Structure of computers) (9 hours)		
1.1	Functional units,basic operational concepts,bus structures (introduction)	1
1.2	Memory locations and addresses , memory operations	1
1.3	Instructions and instruction sequencing	1
1.4	Addressing modes	1
1.5	Fundamental concepts of instruction execution, instruction cycle	1
1.6	Execution of a complete instruction - single bus organization (Lecture 1)	1
1.7	Execution of a complete instruction - single bus organization (Lecture 2)	1
1.8	Execution of a complete instruction - multiple bus organization (Lecture 1)	1
1.9	Execution of a complete instruction - multiple bus organization (Lecture 2)	1
Module 2 :(Register transfer logic and Processor logic design) (10 hours)		
2.1	Inter register transfer – arithmetic micro operations	1
2.2	Inter register transfer – logic and shift micro operations	1
2.3	Processor organization	1
2.4	Design of arithmetic circuit	1
2.5	Design of logic circuit	1
2.6	Design of arithmetic logic unit	1
2.7	Design of status register	1
2.8	Design of shifter - processor unit	1

2.9	Design of accumulator (Lecture 1)	1
2.10	Design of accumulator (Lecture 2)	1
Module 3 : (Arithmetic algorithms and Pipelining) (9 hours)		
3.1	Algorithm for multiplication of binary numbers	1
3.2	Algorithm for division (restoring method) of binary numbers	1
3.3	Array multiplier	1
3.4	Booth's multiplication algorithm	1
3.5	Pipelining: Basic principles	1
3.6	Classification of pipeline processors (Lecture 1)	1
3.7	Classification of pipeline processors (Lecture 2)	1
3.8	Instruction and arithmetic pipelines (Design examples not required)	1
3.9	Hazard detection and resolution	1
Module 4 :(Control Logic Design) (9 hours)		
4.1	Control organization –design of hardwired control logic (Lecture 1)	1
4.2	Control organization –design of hardwired control logic (Lecture 2)	1
4.3	Control organization –design of hardwired control logic (Lecture 3)	1
4.4	Design of microprogram control logic–control of processor unit (Lecture1)	1
4.5	Design of microprogram control logic–control of processor unit (Lecture2)	1
4.6	Design of microprogram control logic–control of processor unit (Lecture3)	1
4.7	Microprogram sequencer	1
4.8	Micro programmed CPU organization	1
4.9	Microinstructions –horizontal and vertical micro instructions	1
Module 5 : (Basic processing units, I/O and memory) (8 hours)		
5.1	Accessing of I/O devices –interrupts	1
5.2	Interrupt hardware	1

5.3	Direct memory access	1
5.4	Memory system: basic concepts –semiconductor RAMs	1
5.5	Memory system considerations – ROMs	1
5.6	Content addressable memory	1
5.7	Cache memories -mapping functions (Lecture 1)	1
5.8	Cache memories -mapping functions (Lecture 2)	1

CST 204	DATABASE MANAGEMENT SYSTEMS	CATEGORY	L	T	P	CREDIT	YEAR OF INTRODUCTION
		PCC	3	1	0	4	2019

Preamble: This course provides a clear understanding of fundamental principles of Database Management Systems (DBMS) with special focus on relational databases to the learners. The topics covered in this course are basic concepts of DBMS, Entity Relationship (ER) model, Relational Database principles, Relational Algebra, Structured Query Language (SQL), Physical Data Organization, Normalization and Transaction Processing Concepts. The course also gives a glimpse of the alternative data management model, NoSQL. This course helps the learners to manage data efficiently by identifying suitable structures to maintain data assets of organizations and to develop applications that utilize database technologies.

Prerequisite: Topics covered under the course Data Structures (CST 201), Exposure to a High Level Language like C/python.

Course Outcomes: After the completion of the course the student will be able to

CO1	Summarize and exemplify fundamental nature and characteristics of database systems (**Cognitive Knowledge Level: Understand)**
CO2	Model real word scenarios given as informal descriptions, using Entity Relationship diagrams. **(Cognitive Knowledge Level: Apply**)
CO3	Model and design solutions for efficiently representing and querying data using relational model (**Cognitive Knowledge Level: Analyze**)
CO4	Demonstrate the features of indexing and hashing in database applications **(Cognitive Knowledge Level: Apply**)
CO5	Discuss and compare the aspects of Concurrency Control and Recovery in Database systems **(Cognitive Knowledge Level: Apply**)
CO6	Explain various types of NoSQL databases (**Cognitive Knowledge Level: Understand**)

Mapping of course outcomes with program outcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	✓	✓	✓									✓
CO2	✓	✓	✓	✓								✓
CO3	✓	✓	✓	✓								✓
CO4	✓	✓	✓							✓		✓
CO5	✓	✓	✓							✓		✓
CO6	✓	✓	✓		✓					✓		✓

Abstract POs defined by National Board of Accreditation			
PO#	**Broad PO**	**PO#**	**Broad PO**
PO1	Engineering Knowledge	**PO7**	Environment and Sustainability
PO2	Problem Analysis	**PO8**	Ethics
PO3	Design/Development of solutions	**PO9**	Individual and team work
PO4	Conduct investigations of complex problems	**PO10**	Communication
PO5	Modern tool usage	**PO11**	Project Management and Finance
PO6	The Engineer and Society	**PO12**	Life long learning

Assessment Pattern

Bloom's Category	Continuous Assessment Tests		End Semester Examination Marks (%)
	Test1 (%)	Test2 (%)	
Remember	30	30	30
Understand	40	40	40
Apply	30	30	30

Analyze			
Evaluate			
Create			

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
150	**50**	**100**	**3 hours**

Continuous Internal Evaluation Pattern:

Attendance : 10 marks

Continuous Assessment Tests : 25 marks

Continuous Assessment Assignment : 15 marks

Internal Examination Pattern:

Each of the two internal examinations has to be conducted out of 50 marks

First Internal Examination shall be preferably conducted after completing the first half of the syllabus and the Second Internal Examination shall be preferably conducted after completing remaining part of the syllabus.

There will be two parts: Part A and Part B. Part A contains 5 questions (preferably, 2 questions each from the completed modules and 1 question from the partly covered module), having 3 marks for each question adding up to 15 marks for part A. Students should answer all questions from Part A. Part B contains 7 questions (preferably, 3 questions each from the completed modules and 1 question from the partly covered module), each with 7 marks. Out of the 7 questions in Part B, a student should answer any 5.

End Semester Examination Pattern:

There will be two parts; Part A and Part B. Part A contains 10 questions with 2 questions from each module, having 3 marks for each question. Students should answer all questions. Part B contains 2 questions from each module of which a student should answer any one. Each question can have maximum 2 sub-divisions and carries 14 marks.

Syllabus

Module 1: Introduction & Entity Relationship (ER) Model

Concept & Overview of Database Management Systems (DBMS) - Characteristics of Database system, Database Users, structured, semi-structured and unstructured data. Data Models and Schema - Three Schema architecture. Database Languages, Database architectures and classification.

ER model - Basic concepts, entity set & attributes, notations, Relationships and constraints, cardinality, participation, notations, weak entities, relationships of degree 3.

Module 2: Relational Model

Structure of Relational Databases - Integrity Constraints, Synthesizing ER diagram to relational schema

Introduction to Relational Algebra - select, project, cartesian product operations, join - Equi-join, natural join. query examples, introduction to Structured Query Language (SQL), Data Definition Language (DDL), Table definitions and operations – CREATE, DROP, ALTER, INSERT, DELETE, UPDATE.

Module 3: SQL DML (Data Manipulation Language), Physical Data Organization

SQL DML (Data Manipulation Language) - SQL queries on single and multiple tables, Nested queries (correlated and non-correlated), Aggregation and grouping, Views, assertions, Triggers, SQL data types.

Physical Data Organization - Review of terms: physical and logical records, blocking factor, pinned and unpinned organization. Heap files, Indexing, Singe level indices, numerical examples, Multi-level-indices, numerical examples, B-Trees & B+-Trees (structure only, algorithms not required), Extendible Hashing, Indexing on multiple keys – grid files.

Module 4: Normalization

Different anomalies in designing a database, The idea of normalization, Functional dependency, Armstrong's Axioms (proofs not required), Closures and their computation, Equivalence of Functional Dependencies (FD), Minimal Cover (proofs not required). First Normal Form (1NF), Second Normal Form (2NF), Third Normal Form (3NF), Boyce Codd Normal Form (BCNF), Lossless join and dependency preserving decomposition, Algorithms for checking Lossless Join (LJ) and Dependency Preserving (DP) properties.

Module 5: Transactions, Concurrency and Recovery, Recent Topics

Transaction Processing Concepts - overview of concurrency control, Transaction Model, Significance of concurrency Control & Recovery, Transaction States, System Log, Desirable Properties of transactions.

Serial schedules, Concurrent and Serializable Schedules, Conflict equivalence and conflict serializability, Recoverable and cascade-less schedules, Locking, Two-phase locking and its variations. Log-based recovery, Deferred database modification, check-pointing.

Introduction to NoSQL Databases, Main characteristics of Key-value DB (examples from: Redis), Document DB (examples from: MongoDB)

Main characteristics of Column - Family DB (examples from: Cassandra) and Graph DB (examples from : ArangoDB)

Text Books

1. Elmasri R. and S. Navathe, Database Systems: Models, Languages, Design and Application Programming, Pearson Education, 2013.
2. Sliberschatz A., H. F. Korth and S. Sudarshan, Database System Concepts, 6/e, McGraw Hill, 2011.

Reference Books:

1. Adam Fowler, NoSQL for Dummies, John Wiley & Sons, 2015
2. NoSQL Data Models: Trends and Challenges (Computer Engineering: Databases and Big Data), Wiley, 2018
3. Web Resource: https://www.w3resource.com/redis/
4. web Resource: https://www.w3schools.in/category/mongodb/
5. Web Resource: https://www.tutorialspoint.com/cassandra/cassandra_introduction.htm
6. Web Resource : https://www.tutorialspoint.com/arangodb/index.htm

Sample Course Level Assessment Questions

Course Outcome1 (CO1):

1. List out any three salient features of database systems, which distinguish it from a file system.
2. Give one example each for logical and physical data independence.

Course Outcome 2(CO2):

1. What facts about the relationships between entities EMPLOYEE and PROJECT are conveyed by the following ER diagram?

1. Design an ER diagram for the following scenario:
   There is a set of teams, each team has an ID (unique identifier), name, main stadium, and to which city this team belongs. Each team has many players, and each player belongs to one team. Each player has a number (unique identifier), name, DoB, start year, and shirt number that he uses. Teams play matches, in each match there is a host team and a guest team.

Course Outcome 3(CO3):

1. For the SQL query, SELECT *A, B FROM R WHERE B='apple' AND C = 'orange'* on the table R(A, B, C, D), where A is a key, write any three equivalent relational algebra expressions.
2. Given the FDs P→Q, P→R, QR→S, Q→T, QR→U, PR→U, write the sequence of *Armstrong's Axioms* needed to arrive at the following FDs: (a) P → T (b) PR → S (c) QR → SU
3. Consider a relation PLAYER (PLAYER-NO, PLAYER-NAME, PLAYER-POSN, TEAM, TEAM-COLOR, COACH-NO, COACH-NAME, TEAM-CAPTAIN). Assume that PLAYER-NO is the *only* key of the relation and that the following dependencies hold:

   TEAM→{TEAM-COLOR, COACH-NO, TEAM-CAPTAIN}
   COACH-NO→COACH-NAME.

   i. Is the relation in 2NF? If not, decompose to 2NF.
   ii. Is the relation in 3NF? If not, decompose to 3NF.

4. In the following tables foreign keys have the same name as primary keys except DIRECTED-BY, which refers to the primary key ARTIST-ID. Consider only *single-director* movies.
   MOVIES(MOVIE-ID, MNAME, GENRE, LENGTH, DIRECTED-BY)
   ARTIST(ARTIST-ID, ANAME)
   ACTING(ARTIST-ID, MOVIE-ID)
   Write SQL expressions for the following queries:
   (a) *Name(s) and director name(s) of movie(s) acted by 'Jenny'.*
   (b) *Names of actors who have never acted with 'Rony'*
   (c) *Count of movies genre-wise.*
   (d) *Name(s) of movies with maximum length.*

Course Outcome 4(CO4):

1. Consider an EMPLOYEE file with 10000 records where each record is of size 80 bytes. The file is sorted on employee number (15 bytes long), which is the primary key. Assuming un-spanned organization, block size of 512 bytes and block pointer size of 5 bytes. Compute the number of block accesses needed for retrieving an employee record based on employee number if (i) No index is used (ii) Multi-level primary index is used.

Course Outcome 5(CO5):

1. Determine if the following schedule is *recoverable*. Is the schedule *cascade-less*? Justify your answer. *r1(X), r2(Z), r1(Z), r3(X), r3(Y), w1(X), c1, w3(Y), c3, r2(Y), w2(Z), w2(Y), c2*. (*Note: ri(X)/wi(X)* means transaction *Ti* issues read/write on item X; *ci* means transaction *Ti* commits.)
2. Two-phase locking protocol ensures serializability. Justify.

Course Outcome 6(CO6):

1. List out any three salient features of NoSQL databases. Give example of a document in MongoDB.

Model Question paper

QPCODE

Reg No:_____

Name:_____

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

FOURTH SEMESTER B.TECH DEGREE EXAMINATION, MONTH & YEAR

Course Code: CST 204

Course Name: **Database Management Systems**

Max.Marks:100 **Duration: 3 Hours**

PART A

Answer all Questions. Each question carries 3 Marks

1 List out any three salient features of a database systems.

2 When is multi-valued composite attribute used in ER modelling?

3 For the SQL query, SELECT *A, B FROM R WHERE B='apple' AND C = 'orange'* on the table R(A, B, C, D), where A is a key, write any two equivalent relational algebra expressions.

4 Outline the concept of *theta*-join.

5 How is the purpose of *where* clause is different from that of having clause?

6 What is the use of a trigger?

7 When do you say that a relation is not in 1NF?

8 Given the FDs P→Q, P→R, QR→S, Q→T, QR→U, PR→U, write the sequence of Armstrong's Axioms needed to arrive at a. P → T b. PR → S

9 What is meant by the lost update problem?

10 What is meant by check pointing?

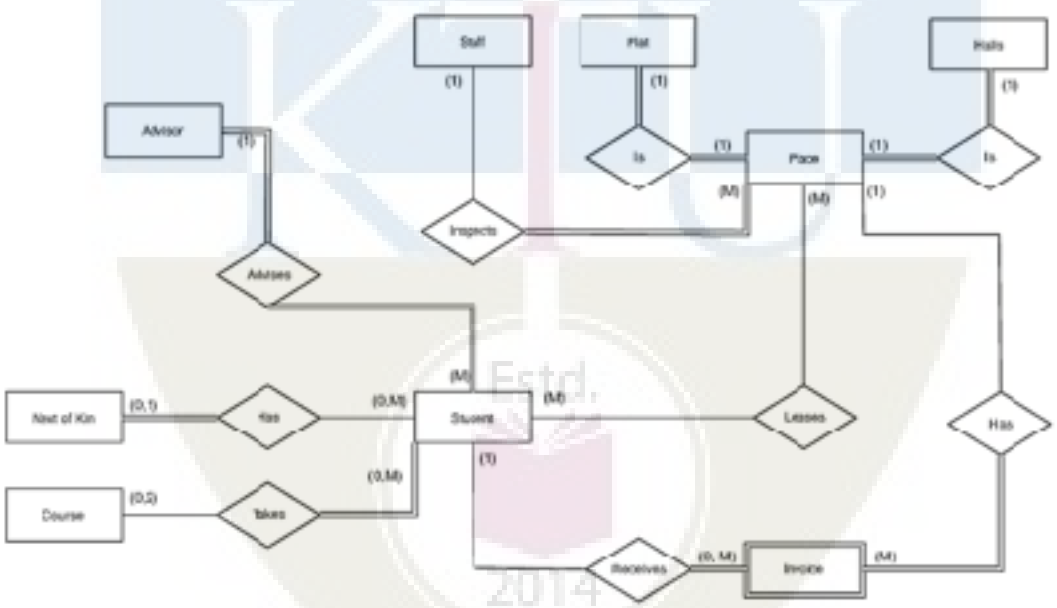
PART B

Answer any one Question from each module. Each question carries 14 Marks

11 a. Design an ER diagram for the following scenario: There is a set of teams, each team has an ID (unique identifier), name, main stadium, and to which city this team belongs. Each team has many players, and each player belongs to one team. Each player has a number (unique identifier), name, DoB, start year, and shirt number that he uses. Teams play matches, in each match there is a host team and a guest team. The match takes place in the stadium of the host team. For each match we need to keep track of the following: The date on which the game is played The final result of the match. The players participated in the match. For each player, how many goals he scored, whether or not he took yellow card, and whether or not he took red card. During the match, one player may substitute another player. We want to capture this substitution and the time at which it took place. Each match has exactly three referees. For each referee we have an ID (unique identifier), name, DoB, years of experience. One referee is the main referee and the other two are assistant referee. (14)

OR

12 a. Interpret the the following ER diagram. (8)



b. Distinguish between physical data independence and logical data independence with suitable examples. (6)

13 **EMPLOYEE(ENO, NAME, ADDRESS, DOB, AGE, GENDER, SALARY, DNUM, SUPERENO)**
DEPARTMENT(DNO, DNAME, DLOCATION, DPHONE, MGRENO)
PROJECT(PNO, PNAME, PLOCATION, PCOST, CDNO) (14)

DNUM is a foreign key that identifies the department to which an employee belongs. MGRENO is a foreign key identifying the employee who manages the department. CDNO is a foreign key identifying the department that controls the project. SUPERENO is a foreign key identifying the supervisor of each employee.

Write relational algebra expressions for the following queries:-

(a) Names of female employees whose salary is more than 20000.
(b) Salaries of employee from 'Accounts' department
(c) Names of employees along with his/her superviser's name
(d) For each employee return name of the employee along with his department name and the names of projects in which he/she works
(e) Names of employees working in all the departments

OR

14 a. Write SQL DDL statements for the the following (Assume suitable domain types): (10)

i. Create the tables STUDENT(ROLLNO, NAME, CLASS, SEM, ADVISER), FACULTY(FID, NAME, SALARY, DEPT). Assume that ADVISER is a foreign key referring FACUTY table.

ii. Delete department with name 'CS' and all employees of the department.

iii. Increment salary of every faculty by 10%.

b. Illustrate foreign key constraint with a typical example. (4)

15 For the relation schema below, give an expression in SQL for each of the queries that follows: (14)

employee(<u>employee-name</u>, street, city)
works(<u>employee-name, company-name</u>, salary)
company(<u>company-name</u>, city)
manages(<u>employee-name, manager-name</u>)

a) Find the names, street address, and cities of residence for all employees who work for the Company 'RIL Inc.' and earn more than $10,000.
b) Find the names of all employees who live in the same cities as the companies for which they work.
c) Find the names of all employees who do not work for 'KYS Inc.'. Assume that all people work for exactly one company.
d) Find the names of all employees who earn more than every employee of 'SB Corporation'. Assume that all people work for at most one company.
e) List out number of employees company-wise in the decreasing order of number of employees.

OR

16 a. Consider an EMPLOYEE file with 10000 records where each record is of size 80 bytes. The file is sorted on employee number (15 bytes long), which is the primary key. Assuming un-spanned organization and block size of 512 bytes compute the number of block accesses needed for selecting records based on employee number if, (9)

i. No index is used
ii. Single level primary index is used
iii. Multi-level primary index is used

Assume a block pointer size of 6 bytes.

b. Illustrate correlated and non-correlated nested queries with real examples. (5)

17 a. Illstrate3NF and BCNF with suitable real examples. (6)

b. Given a relation R(A1,A2,A3,A4,A5) with functional dependencies A1→A2A4 and A4→A5, check if the decomposition R1(A1,A2,A3), R2(A1,A4), R3(A2,A4,A5) is lossless. (8)

OR

18 a. Consider the un-normalized relation R(A, B, C, D, E, F, G) with the FDs A→B , AC→G, AD→EF, EF→G, CDE→AB. Trace the normalization process to reach 3NF relations. (7)

b. Illustrate Lossless Join Decomposition and Dependency Preserving Decomposition with typical examples. (7)

19 a. Discuss the four ACID properties and their importance. (7)

b. Determine if the following schedule is conflict serializable. Is the schedule recoverable? Is the schedule cascade-less? Justify your answers. (7)

r1(X), r2(Z), r1(Z), r3(X), r3(Y), w1(X), c1, w3(Y), c3, r2(Y), w2(Z), w2(Y), c2

(Note: ri(X)/wi(X) means transaction Ti issues read/write on item X; ci means transaction Ti commits.)

OR

20 a. Discuss the main characteristics of Key-value DB and Graph DB. (7)

b. Illustrate two-phase locking with a schedule containing three transactions. Argue that 2PL ensures serializability. Also argue that 2Pl can lead to deadlock. (7)

Teaching Plan

	Course Name	Hours (48)
	Module 1: Introduction & ER Model	**8**
1.1	Concept & Overview of DBMS, Characteristics of DB system, Database Users.	1
1.2	Structured, semi-structured and unstructured data. Data Models and Schema	1
1.3	Three-Schema-architecture. Database Languages	1
1.4	Database architectures and classification	1
1.5	ER model: basic concepts, entity set & attributes, notations	1
1.6	Relationships and constraints – cardinality, participation, notations	1
1.7	Weak entities, relationships of degree 3	1
1.8	ER diagram – exercises	1
	Module 2: Relational Model	**7**
2.1	Structure of relational Databases, Integrity Constraints	1
2.2	Synthesizing ER diagram to relational schema, Introduction to relational algebra.	1
2.3	Relational algebra: select, project, Cartesian product operations	1
2.4	Relational Algebra: join - Equi-join, Natural join	1
2.5	Query examples	1
2.6	Introduction to SQL, important data types	1
2.7	DDL, Table definitions and operations – CREATE, DROP, ALTER, INSERT, DELETE, UPDATE	1
	Module 3: SQL DML, Physical Data Organization	**11**
3.1	SQL DML, SQL queries on single and multiple tables	1
3.2	Nested queries (correlated and non-correlated)	1
3.3	Aggregation and grouping	1

	Course Name	Hours (48)
3.4	Views, assertions (with examples)	1
3.5	Triggers (with examples), SQL data types	1
3.6	Review of terms: physical and logical records, blocking factor, pinned and unpinned organization. Heap files, Indexing	1
3.7	Singe level indices, numerical examples	1
3.8	Multi-level-indices, numerical examples	1
3.9	B-Trees and B+Trees (structure only, algorithms not required)	1
3.10	Extendible Hashing	1
3.11	Indexing on multiple keys – grid files	1
	Module 4: Normalization	**8**
4.1	Different anomalies in designing a database, The idea of normalization	1
4.2	Functional dependency, Armstrong's Axioms (proofs not required)	1
4.3	Closures and their computation, Equivalence of FDs, minimal Cover (proofs not required).	1
4.4	1NF, 2NF	1
4.5	3NF, BCNF	1
4.6	Lossless join and dependency preserving decomposition	1
4.7	Algorithms for checking Lossless Join and Dependency preserving properties (Lecture 1)	1
4.8	Algorithms for checking Lossless Join and Dependency preserving properties (Lecture 2)	1
	Module 5: Transactions, Concurrency and Recovery, Recent Topics	**14**
5.1	Transaction Processing Concepts: Transaction Model	1
5.2	Overview of concurrency control, Significance of concurrency Control & Recovery	1
5.3	Transaction States, System Log	1

	Course Name	Hours (48)
5.4	Desirable Properties of transactions, Serial schedules	1
5.5	Concurrent and Serializable Schedules	1
5.6	Conflict equivalence and conflict serializability	1
5.7	Recoverable and cascade-less schedules	1
5.8	Locking, Two-phase locking, strict 2PL.	1
5.9	Log-based recovery	1
5.10	Deferred database modification (serial schedule), example	1
5.11	Deferred database modification (concurrent schedule) example, check-pointing	1
5.12	Introduction to NoSQL Databases	1
5.13	Main characteristics of Key-value DB (examples from: Redis), Document DB (examples from: MongoDB) [detailed study not expected]	1
5.14	Main characteristics of Column-Family DB (examples from: Cassandra) and Graph DB (examples from : ArangoDB) [detailed study not expected]	1

CST 206	OPERATING SYSTEMS	Category	L	T	P	Credit	Year of Introduction
		PCC	3	1	0	4	2019

Preamble: Study of operating system is an essential to understand the overall working of computer system, tradeoffs between performance and functionality and the division of jobs between hardware and software. This course introduces the concepts of memory management, device management, process management, file management and security & protection mechanisms available in an operating system. The course helps the learner to understand the fundamentals about any operating system design so that they can extend the features of operating system to detect and solve many problems occurring in operating system and to manage the computer resources appropriately.

Prerequisite: Topics covered in the courses are **Data Structures (CST 201)** and **Programming in C (EST 102)**

Course Outcomes: After the completion of the course the student will be able to

CO1	Explain the relevance, structure and functions of Operating Systems in computing devices. (**Cognitive knowledge: Understand**)
CO2	Illustrate the concepts of process management and process scheduling mechanisms employed in Operating Systems. (**Cognitive knowledge: Understand**)
CO3	Explain process synchronization in Operating Systems and illustrate process synchronization mechanisms using Mutex Locks, Semaphores and Monitors (**Cognitive knowledge: Understand**)
CO4	Explain any one method for detection, prevention, avoidance and recovery for managing deadlocks in Operating Systems. (**Cognitive knowledge: Understand**)
CO5	Explain the memory management algorithms in Operating Systems. (**Cognitive knowledge: Understand**)
CO6	Explain the security aspects and algorithms for file and storage management in Operating Systems. **(Cognitive knowledge: Understand)**

Mapping of course outcomes with program outcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	✓	✓	✓							✓		✓
CO2	✓	✓	✓	✓						✓		✓
CO3	✓	✓	✓	✓						✓		✓
CO4	✓	✓	✓	✓						✓		✓
CO5	✓	✓	✓	✓						✓		✓
CO6	✓	✓	✓	✓						✓		✓

Abstract POs defined by National Board of Accreditation			
PO#	**Broad PO**	PO#	**Broad PO**
PO1	Engineering Knowledge	PO7	Environment and Sustainability
PO2	Problem Analysis	PO8	Ethics
PO3	Design/Development of solutions	PO9	Individual and team work
PO4	Conduct investigations of complex problems	PO10	Communication
PO5	Modern tool usage	PO11	Project Management and Finance
PO6	The Engineer and Society	PO12	Life long learning

Assessment Pattern

Bloom's Category	Test 1 (Marks in percentage)	Test 2 (Marks in percentage)	End Semester Examination (Marks in percentage)
Remember	30	30	30
Understand	30	30	30
Apply	40	40	40
Analyse			
Evaluate			
Create			

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
150	50	100	3

Continuous Internal Evaluation Pattern:
Attendance : 10 marks
Continuous Assessment Test : 25 marks
Continuous Assessment Assignment : 15 marks

Internal Examination Pattern:

Each of the two internal examinations has to be conducted out of 50 marks. First series test shall be preferably conducted after completing the first half of the syllabus and the second series test shall be preferably conducted after completing remaining part of the syllabus. There will be two parts: Part A and Part B. Part A contains 5 questions (preferably, 2 questions each from the completed modules and 1 question from the partly completed module), having 3 marks for each question adding up to 15 marks for part A. Students should answer all questions from Part A. Part B contains 7 questions (preferably, 3 questions each from the completed modules and 1 question from the partly completed module), each with 7 marks. Out of the 7 questions, a student should answer any 5.

End Semester Examination Pattern:
There will be two parts; Part A and Part B. Part A contains 10 questions with 2 questions from each module, having 3 marks for each question. Students should answer all questions. Part B contains 2 questions from each module of which a student should answer any one. Each question can have maximum 2 sub-divisions and carries 14 marks.

Syllabus

Module I

Introduction: Operating system overview – Operations, Functions, Service – System calls, Types – Operating System structure - Simple structure, Layered approach, Microkernel, Modules – System boot process.

Module II

Processes - Process states, Process control block, threads, scheduling, Operations on processes - process creation and termination – Inter-process communication - shared memory systems, Message passing systems.

Process Scheduling – Basic concepts- Scheduling criteria -scheduling algorithms- First come First Served, Shortest Job Firs, Priority scheduling, Round robin scheduling

Module III

Process synchronization- Race conditions – Critical section problem – Peterson's solution, Synchronization hardware, Mutex Locks, Semaphores, Monitors – Synchronization problems - Producer Consumer, Dining Philosophers and Readers-Writers.

Deadlocks: Necessary conditions, Resource allocation graphs, Deadlock prevention, Deadlock avoidance – Banker's algorithms, Deadlock detection, Recovery from deadlock.

Module IV

Memory Management: Concept of address spaces, Swapping, Contiguous memory allocation, fixed and variable partitions, Segmentation, Paging. Virtual memory, Demand paging, Page replacement algorithms.

Module V

File System: File concept - Attributes, Operations, types, structure – Access methods, Protection. File-system implementation, Directory implementation. Allocation methods.
Storage Management: Magnetic disks, Solid-state disks, Disk Structure, Disk scheduling, Disk formatting.

Text Book

Abraham Silberschatz, Peter Baer Galvin, Greg Gagne, ' Operating System Concepts' 9th Edition, Wiley India 2015.

Reference Books:

1. Andrew S Tanenbaum, "Modern Operating Systems" , 4th Edition, Prentice Hall, 2015.
2. William Stallings, "Operating systems", 6th Edition, Pearson, Global Edition, 2015.
3. Garry Nutt, Nabendu Chaki, Sarmistha Neogy, "Operating Systems", 3rd Edition, Pearson Education.
4. D.M.Dhamdhere, "Operating Systems", 2nd Edition, Tata McGraw Hill, 2011.
5. Sibsankar Haldar, Alex A Aravind, "Operating Systems", Pearson Education.

Sample Course Level Assessment Questions

Course Outcome1 (CO1): What is the main advantage of the micro kernel approach to system design? How do user program and system program interact in a microkernel architecture?

Course Outcome 2 (CO2): Define process. With the help of a neat diagram explain different states of process.

Course Outcome 3 (CO3): What do you mean by binary semaphore and counting semaphore? With C, explain implementation of wait () and signal().

Course Outcome 4 (CO4): Describe resource allocation graph for the following. a) with a deadlock b) with a cycle but no deadlock.

Course Outcome 5 (CO5): Consider the following page reference string 1, 2, 3, 4, 2, 1, 5, 6, 2, 1, 2, 3, 7, 6, 3, 2, 1, 2, 3, 6. Find out the number of page faults if there are 4 page frames, using the following page replacement algorithms. i) LRU ii) FIFO iii) Optimal

Course Outcome 6 (CO6): Explain the different file allocation methods with advantages and disadvantages.

Model Question Paper

QP CODE: PAGES: ___

Reg No:_____
Name:_____

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY
FOURTH SEMESTER B.TECH DEGREE EXAMINATION, MONTH & YEAR

Course Code: CST 206

Course name : OPERATING SYSTEMS

Max Marks: 100 **Duration: 3 Hours**

PART-A

(Answer All Questions. Each question carries 3 marks)

1. How does hardware find the Operating System kernel after system switch-on?
2. What is the purpose of system call in operating system?
3. Why is context switching considered as an overhead to the system?

4. How is inter process communication implement using shared memory?

5. Describe resource allocation graph for the following.

   a) with a deadlock b)with a cycle but no deadlock.

6. What is critical section? What requirement should be satisfied by a solution to the critical section problem?

7. Consider the reference string 1, 2, 3, 4, 2, 1, 5, 6, 2, 1, 2, 3, 7, 6, 3, 2, 1, 2, 3, 6. How many page faults occur while using FCFS for the following cases.

   a) frame=2 b)frame=3

8. Differentiate between internal and external fragmentations.

9. Compare sequential access and direct access methods of storage devices.

10. Define the terms (i) Disk bandwidth (ii) Seek time.

PART-B(Answer any one question from each module)

11. a) Explain the following structures of operating system (i) Monolithic systems (ii) Layered Systems (iii) Micro Kernel (iv) Modular approach. **(12)**

    b) Under what circumstances would a user be better of using a time sharing system than a PC or a single user workstation? **(2)**

OR

12. a) What is the main advantage of the micro kernel approach to system design? How do user program and system program interact in a microkernel architecture? **(8)**

    b) Describe the differences between symmetric and asymmetric multiprocessing? What are the advantages and disadvantages of multiprocessor systems? **(6)**

13. a) Define process. With the help of a neat diagram explain different states of process. **(8)**

    b) Explain how a new process can be created in Unix using fork system call. **(6)**

OR

14 a) Find the average waiting time and average turnaround time for the processes given in the table below using:- i) SRT scheduling algorithm ii) Priority scheduling algorithm **(9)**

Process	Arrival Time (ms)	CPU Burst Time (ms)	Priority
P1	0	5	3
P2	2	4	1
P3	3	1	2
P4	5	2	4

b) What is a Process Control Block? Explain the fields used in a Process Control Block. **(5)**

15. Consider a system with five processes P_0 through P_4 and three resources of type A, B, C. Resource type A has 10 instances, B has 5 instances and C has 7 instances. Suppose at time t_0 following snapshot of the system has been taken:

Process	Allocation	Max	Available
	A B C	A B C	A B C
P_0	0 1 0	7 5 3	3 3 2
P_1	2 0 0	3 2 2	
P_2	3 0 2	9 0 2	
P_3	2 1 1	2 2 2	
P_4	0 0 2	4 3 3	

i) What will be the content of the Need matrix? Is the system in a safe state? If Yes, then what is the safe sequence? **(8)**

iii)What will happen if process P_1 requests one additional instance of resource type A and two instances of resource type C? **(6)**

OR

16. a) State dining philosopher's problem and give a solution using semaphores. **(7)**

b) What do you mean by binary semaphore and counting semaphore? With C struct, explain implementation of wait () and signal() **(7)**

17. a) Consider the following page reference string 1, 2, 3, 4, 2, 1, 5, 6, 2, 1, 2, 3, 7, 6, 3, 2, 1, 2, 3, 6. Find out the number of page faults if there are 4 page frames, using the following page replacement algorithms i) LRU ii) FIFO iii) Optimal **(9)**

b) Explain the steps involved in handling a page fault. **(5)**

OR

18. a) With a diagram, explain how paging is done with TLB. **(5)**

b) Memory partitions of sizes 100 kb, 500 kb, 200 kb, 300 kb, 600 kb are available, how would best ,worst and first fit algorithms place processes of size 212 kb, 417 kb, 112 kb, 426 kb in order. Rank the algorithms in terms of how efficiently they uses memory. **(9)**

19. a) Suppose that a disk drive has 5000 cylinders, numbered 0 to 4999. the drive currently services a request at cylinder 143, and the previous request was at cylinder 125. the queue of pending request in FIFO order is 86, 1470, 913, 1774, 948, 1509, 1022, 1750, 130. Starting from the current position, what is the total distance (in cylinders) that the disk arm moves to satisfy all pending requests for each of the following algorithms

i) FCFS ii) SSFT iii) SCAN iv) LOOK v) C-SCAN **(10)**

b) What is the use of access matrix in protection mechanism? **(4)**

OR

20. a) Explain the different file allocation operations with advantages and disadvantages. **(8)**

b) Explain the following i) file types ii) file operation iii) file attributes **(6)**

Teaching Plan

	Module 1 - Introduction	**5 Hours**
1.1	Introduction to Operating System	1
1.2	Operating System operations, functions, service	1
1.3	System calls, Types	1
1.4	Operating System Structure: Simple, Layered, Microkernel, Modules	1
1.5	System Boot Process	1
	Module 2 – Processes and Process Scheduling	**9 Hours**
2.1	Processes, Process states	1
2.2	Process Control Block, Threads	1

2.3	Scheduling	1
2.4	Operations on processes: process creation and termination	1
2.5	Inter-process communication: Shared memory systems, Message Passing	1
2.6	Process Scheduling – Basic concepts, Scheduling Criteria	1
2.7	Scheduling algorithms - Basics	1
2.8	First come First Served, Shortest Job First	1
2.9	Priority scheduling, Round Robin Scheduling	1
	Module 3 - Process synchronization and Dead locks	**13 Hours**
3.1	Process synchronization, Race conditions	1
3.2	Critical Section problem, Peterson's solution	1
3.3	Synchronization hardware, Mutex Locks	1
3.4	Semaphores	1
3.5	Monitors	1
3.6	Synchronization problem examples (Lecture 1)	1
3.7	Synchronization problem examples (Lecture 2)	1
3.8	Deadlocks: Necessary conditions, Resource Allocation Graphs	1
3.9	Deadlock prevention	1
3.10	Deadlock avoidance	1
3.11	Banker's algorithm	1
3.12	Deadlock detection	1
3.13	Deadlock recovery	1
	Module 4 - Memory Management	**9 Hours**
4.1	Memory Management: Concept of Address spaces	1
4.2	Swapping	1
4.3	Contiguous memory allocation, fixed and variable partitions	1
4.4	Segmentation.	1
4.5	Paging (Lecture 1)	1
4.6	Paging (Lecture 2)	1
4.7	Virtual memory, Demand Paging	1

4.8	Page replacement algorithms (Lecture 1)	1
4.9	Page replacement algorithms (Lecture 2)	1
	Module 5 - File and Disk management	**9 Hours**
5.1	File concept, Attributes, Operations, types, structure	1
5.2	Access methods	1
5.3	Protection	1
5.4	File-System implementation	1
5.5	Directory implementation	1
5.6	Allocation methods	1
5.7	Magnetic disks, Solid-state disks, Disk structure	1
5.8	Disk scheduling	1
5.9	Disk formatting	1

ADL202	PYTHON AND STATISTICAL MODELLING LAB	Category	L	T	P	Credits	Year of introduction
		PCC	0	0	3	2	2019

Preamble: The Python and Statistical modelling course is intended to impart the elementary concepts of Python and apply various statistical techniques to a variety of data. This course provides the learners with hands-on experience in Python and statistical processes like measures of central tendency, measures of dispersion, probability distributions, graphical analysis, correlation analysis and use of statistical analysis software. The course enables the students to get an exposure to Python programming and use proper methods to analyze and interpret data effectively.

Prerequisite: A basic knowledge of Probability and Statistical Modelling.

Course Outcomes: After the completion of the course the student will be able to

CO#	Course Outcomes
CO 1	Experiment with concepts of iteration, function, string and list **(Cognitive Knowledge Level: Apply)**
CO 2	Identify the importance of tuples, dictionary traversal, dictionary methods, files and operations **(Cognitive Knowledge Level: Apply)**
CO 3	Model graphical representation of data, measures of central tendency and measures of dispersion **(Cognitive Knowledge Level: Apply)**
CO 4	Solve problems based on Binomial distribution, Poisson distribution, sampling and regression analysis **(Cognitive Knowledge Level: Apply)**
CO 5	Make use of various correlation tests and utilize statistical analysis software **(Cognitive Knowledge Level: Apply)**

Mapping of course outcomes with program outcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	✓	✓	✓					✓				✓
CO2	✓	✓	✓	✓				✓				✓
CO3	✓	✓	✓	✓				✓				✓
CO4	✓	✓	✓	✓				✓				✓
CO5	✓	✓	✓	✓	✓			✓				✓

Abstract POs defined by National Board of Accreditation			
PO#	**Broad PO**	**PO#**	**Broad PO**
PO1	Engineering Knowledge	**PO7**	Environment and Sustainability
PO2	Problem Analysis	**PO8**	Ethics
PO3	Design/Development of solutions	**PO9**	Individual and team work
PO4	Conduct investigations of complex problems	**PO10**	Communication
PO5	Modern tool usage	**PO11**	Project Management and Finance
PO6	The Engineer and Society	**PO12**	Lifelong learning

Assessment Pattern

Bloom's Category	Continuous Assessment Tests	End Semester Examination (Percentage)
Remember	20	20
Understand	20	20
Apply	60	60
Analyze		
Evaluate		
Create		

Mark distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
150	75	75	3 hours

Continuous Internal Evaluation Pattern:

Attendance : 15 marks
Continuous Evaluation in Lab : 30 marks
Continuous Assessment Test : 15 marks
Viva Voce : 15 marks

Internal Examination Pattern:

The marks will be distributed as Design/Algorithm 30 marks, Implementation/Program 20 marks, Output 20 marks and Viva 30 marks. Total 100 marks which will be converted out of 15 while calculating Internal Evaluation marks.

End Semester Examination Pattern:

The marks will be distributed as Design/Algorithm 30 marks, Implementation/Program 20 marks, Output 20 marks and Viva 30 marks. Total 100 marks will be converted out of 75 for End Semester Examination.

Fair Lab Record:

All Students attending the Statistical Modelling Using Python Lab should have a Fair Record. The fair record should be produced in the University Lab Examination. Every experiment conducted in the lab should be noted in the fair record. For every experiment in the fair record, the right-hand page should contain Experiment Heading, Experiment Number, Date of experiment, Aim of the Experiment and the operations performed on them, Details of experiment including algorithm and result of Experiment. The left-hand page should contain a print out of the code used for experiment and sample output obtained for a set of input.

SYLLABUS

PYTHON AND STATISTICAL MODELLING LAB

1. Familiarization of expressions, conditional and iteration statements.
2. Problems on function and function calls. **
3. String traversal and other important string methods. **
4. List traversal and list operations. **
5. Tuples, dictionary traversal and dictionary methods. **
6. Problems based on files and operations. **
7. Problems on graphical representation of data. **
8. Problems based on measures of central tendency and measures of dispersion using raw data and grouped data. **
9. Application problems based on Binomial and Poisson distribution. **
10. Implement Chi-square test for goodness of fit. **
11. Perform t-test for difference of means. **
12. Implement Correlation tests. (Karl Pearson correlation coefficient and Spearman rank correlation coefficient).
13. Estimation of gain in precision due to stratification. **
14. Analysis of a one way/ two-way ANOVA.
15. Problems on Lines of regression, regression coefficients, angle between regression lines.
16. Familiarization with statistical analysis software. (SPSS or similar) **

**mandatory

PYTHON AND STATISTICAL MODELLING LAB – Practice Questions

1. Write a program to find the largest of three numbers.
2. Write a program to print the multiplication table of a number n.
3. Write a program to find Surface area and volume of a cylinder using function.
4. Write a program to replace a word by another word in a sentence.
5. Write a program to confirm the validity of an email id by verifying its format.
6. Write a program to remove every occurrence of a number from a list.
7. Write a program to add two matrices.
8. Write a program to read a tuple of numbers and print even tuple and odd tuple.
9. Create a dictionary with a set of book title and corresponding stock. Write a program to update the stock and to add or delete books.
10. A set of numbers are stored in a file. Write a program to print the prime numbers among them.
11. Write a program to count the number of words, sentences, upper case letters, lowercase letters and special symbols in a text stored in file.
12. Plot a graph y = f(x).
13. The areas of the various continents of the world (in millions of square miles) are as follows:11.7 for Africa; 10.4 for Asia; 1.9 for Europe; 9.4 for North America; 3.3 Oceania; 6.9 South America; 7.9 Soviet Union. Draw a bar chart representing the given data.
14. Draw the histogram of the following data:

Height of student(m)	135 - 140	140 - 145	145 - 150	150 - 155
No. of students	4	12	16	8

15. Table contains population and murder rates (in units of murders per 100,000 people per year) for different states. Compute the mean, median and variance for the population.

State	**Population**	**Murder**
Alabama	4,779,736	5.7
Alaska	710,231	5.6
Arizona	6,392,017	4.7
Arkansas	2,915,918	5.6
California	37,253,956	4.4
Colorado	5,029,196	2.8
Connecticut	3,574,097	2.4
Delaware	897,934	5.8

16. Calculate the S.D. and coefficient of variation (C.V.) for the following table:

Class:	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80
Frequency:	5	10	20	40	30	20	10	5

17. If X is binomially distributed with 6 trials and a probability of success equal to 0.25 at each attempt, what is the probability of:
    a) exactly 4 successes b) at least one success
18. If the random variable X follows a Poisson distribution with mean 3.4, find P(X=6).
19. A random sample of 395 people were surveyed and each person was asked to report the highest education level they obtained. The data that resulted from the survey is summarized in the following table. Are gender and education level dependent at 5% level of significance?

	High School	**Bachelors**	**Masters**	**Ph.D.**	**Total**
Female	60	54	46	41	201
Male	40	44	53	57	194
Total	100	98	99	98	395

20. Calculate the correlation coefficient of the two variables shown in the table below.

Person	**Hand**	**Height**
A	17	150
B	15	154
C	19	169
D	17	172
E	21	175

21. Suppose a sample of 16 light trucks is randomly selected off the assembly line. The trucks are driven 1000 miles and the fuel mileage (MPG) of each truck is recorded. It is found that the mean MPG is 22 with a SD equal to 3. The previous model of the light truck got 20 MPG. Conduct a t- test of the null hypothesis at p = 0.05

22. The mean productivity rating for all employees at a company was 3.8 on a five-point scale last year. This year you get ratings from a representative sample of fifteen employees from the Human Research Management. Do the data from this sample provide evidence that employee productivity in the department of Human Resource Management is significantly higher than in the company as a whole? Write the null and alternative hypotheses for this problem. Use statistical analysis software to test the null hypothesis stated above.

23. Obtain the regression equation for predicting systolic blood pressure from job satisfaction with reference to the given data using statistical analysis software. If one knows that a subject in the future has a score on job satisfaction of 15, what is their systolic blood pressure predicted to be? What is the standard error of estimate?

Job Satisfaction	**Systolic BP**
34	124
23	128
19	157
43	133
56	116
47	125
32	147
16	167
55	110
25	156

CSL204	OPERATING SYSTEMS LAB	CATEGORY	L	T	P	CREDIT	YEAR OF INTRODUCTION
		PCC	0	0	3	2	2019

Preamble: The course aims to offer students a hands-on experience on Operating System concepts using a constructivist approach and problem-oriented learning. Operating systems are the fundamental part of every computing device to run any type of software.

Prerequisite: Topics covered in the courses are **Data Structures (CST 201)** and **Programming in C (EST 102)**

Course Outcomes:

At the end of the course, the student should be able to

CO1	Illustrate the use of systems calls in Operating Systems. **(Cognitive knowledge: Understand)**
CO2	Implement Process Creation and Inter Process Communication in Operating Systems. **(Cognitive knowledge: Apply)**
CO3	Implement Fist Come First Served, Shortest Job First, Round Robin and Priority-based CPU Scheduling Algorithms. **(Cognitive knowledge: Apply)**
CO4	Illustrate the performance of First In First Out, Least Recently Used and Least Frequently Used Page Replacement Algorithms. **(Cognitive knowledge: Apply)**
CO5	Implement modules for Deadlock Detection and Deadlock Avoidance in Operating Systems. **(Cognitive knowledge: Apply)**
CO6	Implement modules for Storage Management and Disk Scheduling in Operating Systems. **(Cognitive knowledge: Apply)**

Mapping of course outcomes with program outcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	✓	✓	✓					✓		✓		✓
CO2	✓	✓	✓					✓		✓		✓
CO3	✓	✓	✓	✓				✓		✓		✓
CO4	✓	✓	✓	✓				✓		✓		✓
CO5	✓	✓	✓	✓				✓		✓		✓
CO6	✓	✓	✓	✓				✓		✓		✓

Abstract POs defined by National Board of Accreditation			
PO#	**Broad PO**	**PO#**	**Broad PO**
PO1	Engineering Knowledge	PO7	Environment and Sustainability
PO2	Problem Analysis	PO8	Ethics
PO3	Design/Development of solutions	PO9	Individual and team work
PO4	Conduct investigations of complex problems	PO10	Communication
PO5	Modern tool usage	PO11	Project Management and Finance
PO6	The Engineer and Society	PO12	Life long learning

Assessment Pattern:

Bloom's Category	Continuous Assessment Test (Internal Exam) Marks in percentage	End Semester Examination Marks in percentage
Remember	20	20
Understand	20	20
Apply	60	60
Analyse		
Evaluate		
Create		

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
150	**75**	**75**	**3 hours**

Continuous Internal Evaluation Pattern:

Attendance	:	15 marks
Continuous Evaluation in Lab	:	30 marks
Continuous Assessment Test	:	15 marks
Viva Voce	:	15 marks

Internal Examination Pattern: The marks will be distributed as Algorithm 30 marks, Program 20 marks, Output 20 marks and Viva 30 marks. Total 100 marks which will be converted out of 15 while calculating Internal Evaluation marks.

End Semester Examination Pattern: The percentage of marks will be distributed as Algorithm 30 marks, Program 20 marks, Output 20 marks and Viva 30 marks. Total 75 marks.

Operating System to Use in Lab : Linux

Compiler/Software to Use in Lab : gcc

Progamming Language to Use in Lab : Ansi C

Fair Lab Record:

All Students attending the Operating System Lab should have a Fair Record. The fair record should be produced in the University Lab Examination. Every experiment conducted in the lab should be noted in the fair record. For every experiment in the fair record, the right hand page should contain Experiment Heading, Experiment Number, Date of experiment, Aim of the Experiment and the operations performed on them, Details of experiment including algorithm and result of Experiment. The left hand page should contain a print out of the code used for experiment and sample output obtained for a set of input.

SYLLABUS

OPERATING SYSTEMS LAB

* mandatory

1. Basic Linux commands
2. Shell programming
   -Command syntax
   -Write simple functions with basic tests, loops, patterns
3. System calls of Linux operating system:*
   fork, exec, getpid, exit, wait, close, stat, opendir, readdir
4. Write programs using the I/O system calls of Linux operating system (open, read, write)
5. Implement programs for Inter Process Communication using Shared Memory *
6. Implement Semaphores*
7. Implementation of CPU scheduling algorithms. a) Round Robin b) SJF c) FCFS d) Priority *
8. Implementation of the Memory Allocation Methods for fixed partition*
   a) First Fit b) Worst Fit c) Best Fit
9. Implement l page replacement algorithms a) FIFO b) LRU c) LFU*
10. Implement the banker's algorithm for deadlock avoidance. *
11. Implementation of Deadlock detection algorithm
12. Simulate file allocation strategies.
    b) Sequential b) Indexed c) Linked
13. Simulate disk scheduling algorithms. *
    c) FCFS b)SCAN c) C-SCAN

OPERATING SYSTEMS LAB - PRACTICE QUESTIONS

1. Write a program to create a process in linux.
2. Write programs using the following system calls of Linux operating system:
   fork, exec, getpid, exit, wait, close, stat, opendir, readdir
3. Write programs using the I/O system calls of Linux operating system (open, read, write)

4. Given the list of processes, their CPU burst times and arrival times, display/print the Gantt chart for FCFS and SJF. For each of the scheduling policies, compute and print the average waiting time and average turnaround time
5. Write a C program to simulate following non-preemptive CPU scheduling algorithms to find turnaround time and waiting time.

   a)FCFS b) SJF c) Round Robin (pre-emptive) d) Priority
6. Write a C program to simulate following contiguous memory allocation techniques

   a) Worst-fit b) Best-fit c) First-fit
7. Write a C program to simulate paging technique of memory management.
8. Write a C program to simulate Bankers algorithm for the purpose of deadlock avoidance.
9. Write a C program to simulate disk scheduling algorithms a) FCFS b) SCAN c) C-SCAN
10. Write a C program to simulate page replacement algorithms a) FIFO b) LRU c) LFU
11. Write a C program to simulate producer-consumer problem using semaphores.
12. Write a program for file manipulation for display a file and directory in memory.
13. Write a program to simulate algorithm for deadlock prevention.
14. Write a C program to simulate following file allocation strategies.

    a)Sequential b) Indexed c) Linked

SEMESTER IV

HONOURS

CODE	COURSE NAME	CATEGORY	L	T	P	CREDIT	Year of Introduction
CST292	NUMBER THEORY	VAC	4	0	0	4	2019

Preamble: This is the foundational course for awarding B. Tech. Honours in Computer Science and Engineering with specialization in ***Security in Computing***. The purpose of this course is to create awareness among learners about the important areas of number theory used in computer science. This course covers Divisibility & Modular Arithmetic, Primes & Congruences, Euler's Function, Quadratic Residues and Arithmetic Functions, Sum of Squares and Continued fractions. Concepts in Number Theory help the learner to apply them eventually in practical applications in Computer organization & Security, Coding & Cryptography, Random number generation, Hash functions and Graphics.

Prerequisite: A sound background in Higher Secondary School Mathematics

Course Outcomes: After the completion of the course the student will be able to

CO1	Illustrate modular arithmetic operations, methods and techniques **(Cognitive Knowledge Level:Understand)**
CO2	Use the methods - Induction, Contraposition or Contradiction to verify the correctness of mathematical assertions (**Cognitive Knowledge Level: Apply)**
CO3	Utilize theorems and results about prime numbers, congruences, quadratic residues and integer factorization for ensuring security in computing systems (**Cognitive Knowledge Level: Analyse)**
CO4	Illustrate uses of Chinese Remainder Theorem & Euclidean algorithm in Cryptography and Security (**Cognitive Knowledge Level: Apply)**
CO5	Explain applications of arithmetic functions in Computer Science **(Cognitive Knowledge Level:Understand)**
CO6	Implement Number Theoretic Algorithms using a programming language (**Cognitive Knowledge Level: Apply)**

Mapping of course outcomes with program outcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	✓	✓	✓	✓						✓		✓
CO2	✓	✓	✓	✓								✓
CO3	✓	✓	✓	✓		✓						✓
CO4	✓	✓	✓	✓		✓						✓
CO5	✓	✓	✓	✓						✓		✓
CO6	✓	✓	✓	✓	✓			✓				✓

Abstract POs defined by National Board of Accreditation			
PO#	**Broad PO**	**PO#**	**Broad PO**
PO1	Engineering Knowledge	**PO7**	Environment and Sustainability
PO2	Problem Analysis	**PO8**	Ethics
PO3	Design/Development of solutions	**PO9**	Individual and team work
PO4	Conduct investigations of complex problems	**PO10**	Communication
PO5	Modern tool usage	**PO11**	Project Management and Finance
PO6	The Engineer and Society	**PO12**	Life long learning

Assessment Pattern

Bloom's Category	Continuous Assessment Tests		End Semester Examination Marks (Percentage)
	Test1 (Percentage)	Test2 (Percentage)	
Remember	30	30	30
Understand	30	30	30
Apply	40	40	40
Analyse			
Evaluate			
Create			

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
150	**50**	**100**	**3 hours**

Continuous Internal Evaluation Pattern:

Attendance : 10 marks

Continuous Assessment Tests : 25 marks

Continuous Assessment Assignment : 15 marks

Internal Examination Pattern:

Each of the two internal examinations has to be conducted out of 50 marks

First Internal Examination shall be preferably conducted after completing the first half of the syllabus and the Second Internal Examination shall be preferably conducted after completing remaining part of the syllabus.

There will be two parts: Part A and Part B. Part A contains 5 questions (preferably, 2 questions each from the completed modules and 1 question from the partly covered module), having 3 marks for each question adding up to 15 marks for part A. Students should answer all questions from Part A. Part B contains 7 questions (preferably, 3 questions each from the completed modules and 1 question from the partly covered module), each with 7 marks. Out of the 7 questions in Part B, a student should answer any 5.

End Semester Examination Pattern:

There will be two parts; Part A and Part B. Part A contains 10 questions with 2 questions from each module, having 3 marks for each question. Students should answer all questions. Part B contains 2 questions from each module of which a student should answer any one. Each question can have maximum 2 sub-divisions and carries 14 marks.

SYLLABUS

Module 1

Divisibility and Modular Arithmetic:

Finite Fields – Groups, Rings and Fields.

Divisibility - Divisibility and Division Algorithms, Well ordering Principle,Bezout's Identity.

Modular Arithmetic- Properties, Euclid's algorithm for the greatest common divisor, Extended Euclid's Algorithm, Least Common multiple, Solving Linear Diophantine Equations, Modular Division.

Module 2

Primes and Congruences:

Prime Numbers-Prime Numbers andprime-powerfactorization, Fermat and Mersenne primes., Primality testing and factorization.

Congruences-Linear congruences, Simultaneous linear congruences, Chinese Remainder Theorem, Fermat's little theorem, Wilson's theorem.

Module 3

Congruences with a Prime-Power Modulus&Euler's Function:

Congruences with a Prime-Power Modulus-Arithmetic modulo p, Pseudoprimes and Carmichael numbers, Solving congruences modulo prime powers.

Euler's Function-Euler's Totient function, Applications of Euler's Totient function, Traditional Cryptosystem, Limitations.

The Group of units- The group U_n,Primitive roots, Existence of primitive roots, Applications of primitive roots.

Module 4

Quadratic Residues & Arithmetic Functions :

Quadratic Residues- Quadratic Congruences, The group of Quadratic residues, Legendre symbol, Jacobi Symbol, Quadratic reciprocity.

Arithmetic Functions- Definition and examples, Perfect numbers, Mobius function and its properties, Mobius inversion formula, The Dirichlet Products.

Module 5

Sum of Squares and Continued Fractions:

Sum of Squares- Sum of two squares, The Gaussian Integers, Sum of three squares, Sum of four squares.

Continued Fractions -Finite continued fractions, Infinite continued fractions, Pell's Equation, Solution of Pell's equation by continued fractions.

Text Books

1. G.A. Jones & J.M. Jones, Elementary Number Theory, Springer UTM, 2007.

2. Joseph Silverman, A Friendly introduction to Number Theory, Pearson Ed. 2009.

Reference Books

1. William Stallings, Cryptography and Network Security Principles and Practice, Pearson Ed.

2. Tom M.Apostol, 'Introduction to Analytic Number Theory', Narosa Publishing House Pvt. Ltd, New Delhi, (1996).

3. Neal Koblitz, A course in Number Theory and Cryptography, 2nd Edition, Springer ,2004.

Sample Course Level Assessment Questions

Course Outcome 1 (CO1): Describe the properties of modular arithmetic and modulo operator.

Course Outcome 2 (CO2): Prove that the equation $y^2 = x^3 - 2$ has only the integer solution $(3, \pm 5)$.

Course Outcome 3 (CO3): State the law of reciprocity for Jacobi symbols and use it to determine whether 888 is a quadratic residue or non residue of the prime 1999.

Course Outcome 4 (CO4): Using Chinese remainder theorem, solve the system of congruence $x \equiv 2 \pmod 3$, $x \equiv 3 \pmod 5$, $x \equiv 2 \pmod 7$

Course Outcome 5(CO5): State and prove Dirichlet product.

Course Outcome 6 (CO6): Use extended Euclid's algorithm to solve Diophantine equations efficiently. Given three numbers $a>0$, $b>0$, and c, the algorithm should return some x and y such that $a x + b y = c$.

Model Question Paper

QP CODE: PAGES: 03

RegNo : …..... Name :……..……

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY
FOURTH SEMESTER BTECH (HONOURS) DEGREE EXAMINATION, MONTH &YEAR

Course Code:CST 292 Course
Name: Number Theory

Max.Marks:100 **Duration: 3 Hours**

PART A

Answer all Questions. Each question carries 3 Marks (10x3=30)

1. State and prove well ordering principle.
2. Find gcd d of x=525 and y=231 and express d as ax + by where a and b are integers.
3. Solve the congruence equation 103 x≡ 57 (mod 211).
4. Use Fermat's Little theorem to show that 91 is not a prime.
5. If m is relatively prime to n , show that $\Phi(mn) = \Phi(m)\,\Phi(n)$.
6. Explain how public key cryptography can be used for digital signatures.
7. Define Mobius function and prove Mobius function is a multiplicative.
8. State and prove Dirichlet product.
9. Show that every prime of the form 4k+1 canbe represented uniquely as the sum of two squares.
10. Find the continued fraction representation of the rational number 55/89.

Part B

Answer any one Question from each module.
Each question carries 14 Marks

11. (a) State the Euclidean algorithm and its extension with an example. (7)
    (b) Find all the solutions of 24x + 34 y = 6. (7)

OR

12. (a) Describe the properties of modular arithmetic and modulo operator. (7)
    (b) Explain Extended Euclidean algorithm. Using the algorithm find the

multiplicative inverse of 135 mod 61 (7)

13. (a) State and prove Wilson's theorem . (7)

(b) Explain Fermat's factorization method and use it to factor 809009 (7)

OR

14. (a) Using Chinese remainder theorem, solve the system of congruences, $x \equiv 2 (\text{mod } 3)$, $x \equiv 3 (\text{mod } 5)$, $x \equiv 2 (\text{mod } 7)$ (7)

(b) Define Fermat primes. Show that any two distinct Fermat numbers are Relatively prime. (7)

15. (a) Distinguish between public key and private key encryption techniques. Also point out the merits and demerits of both. (7)

(b) Define Carmichael number and show that a Carmichael number must be the product of at least three distinct primes. (7)

OR

16. (a)Define a pseudo prime to a base and find all non trivial bases for which 15 is a pseudo prime. (6)

(b) Find an element of

i) order 5 modulo 11 ii) order 4 modulo 13

iii) order 8 modulo 17 iv) order 6 modulo 19 (8)

17. (a) Determine the quadratic residues and non residues modulo 17. Also determine whether 219 is a quadratic residue or non residue of the prime 383. (8)

(b) State the law of quadratic reciprocity. Determine those odd primes p for which 3 is a quadratic residue and those for which it is a non residue. (6)

OR

18. (a) State and prove properties of Legendre's symbol. (7)

(b) State the law of reciprocity for Jacobi symbols and using it determine whether 888 is a quadratic residue or non residue of the prime 1999. (7)

19. (a) Prove that the equation $y^2 = x^3 - 2$ has only the integer solution $(3, \pm 5)$. (7)

(b) Define a Gaussian integer. Factorize the Gaussian integer 440 − 55i. (7)

OR

20. (a) If *m*, and *n* can be expressed as sum of four squares, then show that *mn* can also be expressed the sum of four squares. (7)

(b) Find all the solutions of the Diophantine equation $x^2 - 6y^2 = 1$. (7)

Teaching Plan

Module 1: Divisibility and Euclidean Algorithm		**9 hours**
1.1	Finite Fields – Groups and Rings.	1 hour
1.2	Finite Fields – Fields.	1 hour
1.3	Divisibility and Division Algorithms, Well ordering Principle.	1 hour
1.4	Decimal Expansion of a positive Integer, Greatest Common Divisor, Bezout's Theorem.	1 hour
1.5	Modular Arithmetic- Properties of congruences, Modular Arithmetic Operations, Properties of Modular Arithmetic.	1 hour
1.6	Euclid's algorithm for the greatest common divisor, Extended Euclid's Algorithm.	1 hour
1.7	Solving Linear Diophantine Equations.	1 hour
1.8	Least Common multiple and Modular Division.	1 hour
1.9	Implementation of Euclid's algorithm, Extended Euclid's Algorithm and solution of Linear Diophantine Equations.	1 hour
Module 2: Primes and Congruences		**9 hours**
2.1	Prime Numbersand prime-powerFactorization.	1 hour
2.2	Fermat and Mersenne primes.	1 hour
2.3	Primality testing and factorization, Miller -Rabin Test for Primality.	1 hour
2.4	Pollard's Rho Method for Factorization, Fermat's Factorization.	1 hour

2.5	Linear congruences, Simultaneous linear congruences.	1 hour
2.6	Chinese Remainder Theorem.	1 hour
2.7	Implementation of Chinese Remainder Theorem.	1 hour
2.8	Fermat's little theorem.	1 hour
2.9	Wilson's theorem.	1 hour
Module 3: Congruences with a Prime-Power Modulus &Euler's Function		**9 hours**
3.1	Congruences with a Prime-Power Modulus, Arithmetic modulo p.	1 hour
3.2	Pseudo-primes and Carmichael numbers.	1 hour
3.3	Solving congruences modulo prime powers.	1 hour
3.4	Definition of Euler Totient function, Examples and properties.	1 hour
3.5	Multiplicativity of Euler's Totient function.	1 hour
3.6	Applications of Euler’s function, Euler’s Theorem.	1 hour
3.7	Traditional Cryptosystem, Limitations, Public Key Cryptography.	1 hour
3.8	The Group of Units, Primitive Roots.	1 hour
3.9	Existence of primitive roots for Primes, Applications of primitive roots.	1 hour
Module 4: Quadratic Residues and Arithmetic Functions		**9 hours**
4.1	Quadratic congruences, The group of Quadratic Residues.	1 hour
4.2	Legendre symbol, Jacobi Symbol.	1 hour
4.3	Quadratic reciprocity.	1 hour
4.4	Quadratic residues for prime-power moduli.	1 hour
4.5	Arithmetic Functions: Definition and examples.	1 hour

4.6	Perfect numbers, Definition and proposition.	1 hour
4.7	Mobius inversion formula., application of the Mobius inversion formula.	1 hour
4.8	Mobius function and its properties.	1 hour
4.9	The Dirichlet Product, Definition and proof.	1 hour
Module 5: Sum of Squares and Continued Fractions		**9 hours**
5.1	Sum of Squares, Sum of two squares.	1 hour
5.2	The Gaussian Integers.	1 hour
5.3	Sum of three squares.	1 hour
5.4	Sum of four squares.	1 hour
5.5	Continued Fractions, Finite continued fractions.	1 hour
5.6	Continued Fractions, Finite continued fractions.	1 hour
5.7	Infinite continued fractions.	1 hour
5.8	Pell's Equation, Definition.	1 hour
5.9	Solution of Pell's equation by continued fractions.	1 hour

ADT294	COMPUTATIONAL FUNDAMENTALS FOR BIOINFORMATICS	Category	L	T	P	Credit	Year of Introduction
		VAC	3	1	0	4	2020

Preamble: Bioinformatics is an interdisciplinary area that combines Computer Science, Molecular Biology, and Mathematics and allied areas of Science. This course covers computational fundamentals of Bioinformatics and Computational Biology such as DNA, genes and proteins, transcription, translation, sequence alignment, representation and basic Python programming required for handling bioinformatics data. The learners will be able to solve basic bioinformatics problems using python programming.

Prerequisite: Basic understanding of programming languages.

Mapping of course outcomes with program outcomes

CO 1	Describe the basic concepts of Bioinformatics with an emphasis on biological macromolecules-DNA, RNA and Protein and synthesis of biomolecules (**Cognitive knowledge level : Understand**)
CO 2	Identify biological data formats and databases, retrieve bio-sequences, and align bio-sequences to identify similarity, dynamic programming (**Cognitive knowledge level : Apply)**
CO 3	Illustrate nucleotide attributes and transcription using programming tools (**Cognitive knowledge level : Apply)**
CO 4	Demonstrate the concepts of Parsing FASTA and Sequences Analysis (**Cognitive knowledge level : Apply)**
CO 5	Compute k-mers, translation of DNA subsequences and Open reading frame. (**Cognitive knowledge level : Apply**)

Mapping of course outcomes with program outcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	√	√	√									√
CO2	√	√	√	√	√	√						√
CO3	√	√	√	√	√							√
CO4	√	√	√	√	√							√
CO5	√	√	√	√								√

Abstract POs defined by National Board of Accreditation			
PO#	**Broad PO**	**PO#**	**Broad PO**
PO1	Engineering Knowledge	**PO7**	Environment and Sustainability
PO2	Problem Analysis	**PO8**	Ethics
PO3	Design/Development of solutions	**PO9**	Individual and team work
PO4	Conduct investigations of complex problems	**PO10**	Communication
PO5	Modern tool usage	**PO11**	Project Management and Finance
PO6	The Engineer and Society	**PO12**	Lifelong learning

Assessment Pattern

Bloom's Category	Continuous Assessment Tests		End Semester Examination
	Test1 (%)	Test2 (%)	
Remember	**10**	**10**	**10**
Understand	**30**	**30**	**70**
Apply	**10**	**10**	**20**
Analyse			
Evaluate			
Create			

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
150	50	100	3

Continuous Internal Evaluation Pattern:

Attendance 10 marks

Continuous Assessment Tests (Average of Internal Tests 1 & 2) 25 marks

Continuous Assessment Assignment 15 marks

Internal Examination Pattern

Each of the two internal examinations has to be conducted out of 50 marks. First series test shall be preferably conducted after completing the first half of the syllabus and the second series test shall be preferably conducted after completing remaining part of the syllabus. There will be two parts: Part A and Part B. Part A contains 5 questions (preferably, 2 questions each from the completed modules and 1 question from the partly completed module), having 3 marks for each question adding up to 15 marks for part A. Students should answer all questions from Part A. Part B contains 7 questions (preferably, 3 questions each from the completed modules and 1 question from the partly completed module), each with 7 marks. Out of the 7 questions, a student should answer any 5.

End Semester Examination Pattern:

There will be two parts; Part A and Part B. Part A contains 10 questions with 2 questions from each module, having 3 marks for each question. Students should answer all questions. Part B contains 2 full questions from each module of which student should answer any one. Each question can have maximum 2 sub-divisions and carries 14 marks.

SYLLABUS

Module-1 (Introduction to bioinformatics)
Introduction to bioinformatics, Nature & Scope of Bioinformatics, animal vs plants, Eukaryote vs prokaryote,Nucleus. Chromosome, gene DNA, RNA, amino acids, and Protein, The Central Dogma, Messenger RNA, tRNA, rRNA, Genetic code, Gene Structure, Transcription, translation.

Module-2 (Introduction to bio sequences and analysis)
Introduction to Biological Databases and data storage, NCBI, Genbank, Bio sequence formats-Database Similarity Searching, BLAST, Sequence alignment, Scoring Matrices, Multiple-Sequence Alignment, Dynamic programming

Module 3: (Introduction to Processing Nucleotides)
Tetranucleotide Frequency, Counting the Nucleotides, Writing and Verifying a Solution, Transcribing DNA into mRNA: Mutating Strings, Reading and Writing Files, Reverse Complement of DNA, String Manipulation, Iterating Over a Reversed String.

Module 4: (Processing Nucleotides GC Content and Hamming Distance)
Creating the Fibonacci Sequence, Writing, Testing, and Benchmarking Algorithms, retrieving FASTA Using Biopython, Iterating the Sequences Using a for Loop, Parsing FASTA and Analyzing Sequences, Computing GC Content, Finding the Hamming Distance, Counting Point Mutations

Module 5 (Translation of DNA and subsequence)

K-mers and Codons, Translating Codons, Translating mRNA into Protein, Finding Subsequences of DNA, Find a Motif in DNA, Finding Overlapping Patterns Using Regular Expressions, Sequence Similarity, Finding the Shortest Sequence in a FASTA File, Extracting K-mers from a Sequence, Counting Frequencies of K-mers, Finding Open Reading Frames

Text Books

1. Mount, D. W.. Bioinformatics: Sequence and Genome Analysis. India, CBS Publishers & Distributors, 2005.
2. Youens-Clark, Ken. *Mastering Python for Bioinformatics*. United States: O'Reilly Media, 2021.

References

1. Kelley, S.T. and Didulo, D, *Computational Biology: A Hypertextbook*. John Wiley & Sons, 2020
2. Baxevanis, Andreas D., Gary D. Bader, and David S. Wishart, eds. *Bioinformatics*. John Wiley & Sons, 2020.
3. Shaik, Noor Ahmad, et al. *Essentials of Bioinformatics, Volume I.* Springer, 2019
4. Selzer, Paul M., Richard J. Marhöfer, and Andreas Rohwer, *Applied bioinformatics. An introduction–Springer, Verlag,*, 2008.
5. S C Rastogi, N Mendiratta and P Rastogi, *Bioinformatics: Methods and Applications,* PHI Learning Private Limited, New Delhi, 2015.
6. D E Krane and M L Raymer, *Fundamental Concepts of Bioinformatics*, Pearson Education, 2006.
7. Bassi, Sebastian. *Python for Bioinformatics*. United Kingdom: CRC Press, 2017.
8. Model, Mitchell L. Bioinformatics Programming Using Python. United States: O'Reilly Media, 2010.
9. Antao, Tiago. *Bioinformatics with Python Cookbook*. United Kingdom: Packt Publishing, 2015.Antao, Tiago. Bioinformatics with Python Cookbook: Learn how to Use Modern Python Bioinformatics Libraries and Applications to Do Cutting-edge Research in Computational Biology, 2nd Edition. United Kingdom: Packt Publishing, 2018.

Course Level Assessment Questions

Course Outcome 1 (CO1)

1. Compare and contrast the DNA and RNA on the basis of structure and functions.
2. Demonstrate with the help of a flow diagram the generation of protein using the transcription and translation process.

Course Outcome 2 (CO2):

1. Identify the following qualifiers for GenBank and give their definitions: [ACCN], [ALL], [AUTH], [ECNO], [FKEY], [GENE], [JOUR], [KYWD]

2. Find the sequence alignment between the following two sequences, locally and Globally
   Sequence1: GATTCTATCTAACTA, Sequence2: GTTCTATTCTAAC
3. Retrieve sequence of Severe acute respiratory syndrome coronavirus 2 and use BLAST to find the similar sequences

Course Outcome 3 (CO3):

1. Write a Python program pseudocode to read the below given sequence as command line argument and print the counts for each of the bases A, C, G, and T.
   Sequence: ACTGCAACGGGCAATATGTCTC
2. Write a python pseudocode to transcribe the following DNA sequence to its mRNA sequence.
   Sequence: TGCAACGGGCAATATGTCTC

Course Outcome 4 (CO4)

1. Solve the problem of generating the Fibonacci sequence using Python**.**
2. Use a simple python program using a list to find the DNA string having the highest GC content, provided any 5 random DNA strings.

Course Outcome 5 (CO5)

1. Illustrate with the help of an example how an RNA string is getting converted to a protein string.
2. Write a python code to print the position and the number of times a subsequence is present in a given DNA string.

Model Question Paper

QP CODE:

Reg No: _____

Name: _____ **PAGES : 4**

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

FOURTH SEMESTER B.TECH DEGREE (HONOURS) EXAMINATION, MONTH & YEAR

Course Code: ADT294

Course Name: COMPUTATIONAL FUNDAMENTALS FOR BIOINFORMATICS

Max. Marks : 100 **Duration: 3 Hours**

PART A

Answer All Questions. Each Question Carries 3 Marks

1. Differentiate DNA, Gene, genome and chromosome.
2. What do you mean by Gene expression?
3. Specify the functions of mRNA, tRNA and rRNA?
4. Differentiate between local and global alignment.
5. Find the reverse complement of the following DNA given in 5'-3'order? AAAACCCGGT
6. List any 3 string manipulation construct used in processing nucleotides.
7. Illustrate how recursion is implemented using a Python pseudocode.
8. What is GC content? Give the GC content of the DNA string: "AGCTATAG".
9. Discuss the role of K-mers and codons in protein synthesis.
10. Define motif in DNA. Mention its importance in finding a conserved sequence.

(10x3=30)

Part B

(Answer any one question from each module. Each question carries 14 Marks)

11.	(a)	Discuss the central dogma of molecular biology.	**(7)**
	(b)	How is the primary transcript produced by a prokaryote different from that produced by a eukaryotic cell?	**(7)**
		OR	
12.	(a)	Differentiate between Prokaryote and Eukaryote Cell	**(7)**
	(b)	Describe with the help of a neat diagram, the structure of DNA.	**(7)**
13.	(a)	What is sequence alignment? Explain any five applications of sequence alignment in Bioinformatics?	**(7)**
	(b)	Discuss variants of BLAST with its input and output	**(7)**
		OR	
14.	(a)	Explain the working principles of the Nucleotide BLAST with an example	**(7)**
	(b)	Differentiate primary and secondary databases in Bioinformatics.	**(7)**
15.	(a)	How do you find the reverse complement of a DNA sequence? Write at least 2 different Python pseudocodes using different constructs to print the reverse complement of a given the 5'-3' end of a DNA sequence.	**(10)**
	(b)	Write a Python pseudocode to convert DNA sequence to RNA sequence by using the re.sub() regular expression construct.	**(4)**
		OR	
16.	(a)	What is the need for 'argparse' module in Python? How can we use this module in different ways to do a tetra nucleotide frequency count?	**(7)**
	(b)	Write a Python program pseudocode to read the below given sequence as command line argument and print the counts for each of the bases A, C, G, and T.	**(7)**
17.	(a)	Generate a random DNA sequence using python and find the transcribed	(7)

DNA sequence of its reverse complement

(b) Write a python code using regular expressions to find the DNA sequence having the highest GC content in a DNA sequence. **(7)**

OR

18. (a) Define Hamming distance. Using hamming distance, find the percentage of similarity between the sequence AAACCCGGGTTT and AACCCGGGTTTA with one sequence in line with other. **(7)**

(b) Write a Python code using zip() function to find the hamming distance between 2 sequence. Give comments on each construct used in the code. **(7)**

19. (a) Write a Python program using function and a list comprehension to translate RNA into protein. Illustrate working of the program with an example RNA string. **(10)**

(b) Illustrate with python pseudocode to show how the str.find( ) function can be used to find a substring and its position in an input sequence. **(4)**

OR

20. (a) Illustrate with the help of an example how an RNA string is getting converted to a protein string.. **(6)**

(b) Write notes on ORF. Write a python code to find the ORF using the str.find( ) and str.partition( ) functions. **(8)**

TEACHING PLAN

No	Contents	No of Lecture Hrs
Module-1 (Introduction to bioinformatics)(10 hrs)		
1.1	Introduction to bioinformatics	1
1.2	Nature & Scope of Bioinformatics	1
1.3	Animal vs plants, Eukaryote vs prokaryote	1
1.4	Nucleus. Chromosome, gene	1
1.5	DNA, RNA, and Protein	1
1.6	The Central Dogma introduction	1
1.7	Messenger RNA, tRNA, rRNA,	1
1.8	Genetic code	1
1.9	Gene Structure and Control	1
1.10	Transcription, Translation	1
Module-2 (Introduction to bio sequences and analysis) (10 hrs)		
2.1	Introduction to Biological Databases and data storage	1
2.2	NCBI, Genbank	1
2.3	NCBI, Genbank Sequence retrieval	1
2.4	Bio sequence formats	1
2.5	Database Similarity Searching, BLAST	1
2.6	BLAST Exercises	1
2.7	Sequence alignment	1
2.8	Scoring Matrices	1
2.9	Multiple-Sequence Alignment	1
2.10	Introduction to Dynamic programming in MSA	1
Module-3 (Introduction to Processing Nucleotides) (8 hrs)		
3.1	Counting the Nucleotides, Writing and Verifying a Solution	1

3.2	Transcribing DNA into mRNA	1
3.3	Iterating the Input Files	1
3.4	Mutating Strings	1
3.5	Writing and Reading Output Sequences	1
3.6	Reverse Complement of DNA	1
3.7	String Manipulation	1
3.8	Iterating Over a Reversed String	1

Module-4 (Processing Nucleotides GC Content and Hamming Distance) (8 hrs)		
4.1	Creating the Fibonacci Sequence	1
4.2	Writing, Testing, and Benchmarking Algorithms	1
4.3	Retrieving FASTA Using Biopython	1
4.4	Parsing FASTA and Analysing Sequences	1
4.5	Computing GC Content	1
4.6	Finding the Hamming Distance	1
4.7	Iterating the Characters of Two Strings	1
4.8	Counting Point Mutations	1
Module-5 (Translation of DNA and subsequence) (9 hrs)		
5.1	K-mers and Codons	1
5.2	Translating mRNA into Protein	1
5.3	Finding Subsequence of DNA	1
5.4	Find a Motif in DNA	1
5.5	Finding Overlapping Patterns Using Regular Expressions	1
5.6	Sequence Similarity	1
5.7	Finding the Shortest Sequence in a FASTA File , Extracting K-mers from a Sequence	1
5.8	Counting Frequencies of K-mers	1
5.9	Finding Open Reading Frames	1

ADT296	ADVANCED TOPICS IN COMPUTER GRAPHICS	CATEGORY	L	T	P	CREDITS
		VAC	3	1	0	4

Preamble: This course helps the learners to make awareness about strong theoretical concept in computer graphics. It covers the three-dimensional environment representation in a computer, transformation of 2D/3D objects, basic mathematical techniques and algorithms used to build useful applications. This course enables the learners to develop the ability to create image processing frameworks for different domains and develops algorithms for emerging display technologies.

Prerequisite: A sound knowledge of Mathematics and concepts of any programming language.

Course Outcomes: After the completion of the course the student will be able to

CO#	CO
CO1	Describe the working principles of graphics devices**(Cognitive Knowledge level: Understand)**
CO2	Illustrate line drawing, circle drawing and polygon filling algorithms**(Cognitive Knowledge level: Apply)**
CO3	Demonstrate geometric representations and transformations on 2D & 3D objects. **(Cognitive Knowledge level: Apply)**
CO4	Demonstrate the working of various clipping algorithms and projection algorithms. **(Cognitive Knowledge level: Apply)**
CO5	Summarize visible surface detection methods**(Cognitive Knowledge level: Understand)**
CO6	Explain the concept of realism in a scene and its performance preservation**(Cognitive Knowledge level: Understand)**

Mapping of course outcomes with program outcomes

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	✓											✓
CO2	✓	✓	✓	✓	✓							✓
CO3	✓	✓	✓	✓	✓							✓
CO4	✓	✓	✓	✓	✓							✓
CO5	✓											✓
CO6	✓	✓										✓

Abstract POs defined by National Board of Accreditation			
PO#	**Broad PO**	**PO#**	**Broad PO**
PO1	Engineering Knowledge	**PO7**	Environment and Sustainability
PO2	Problem Analysis	**PO8**	Ethics
PO3	Design/Development of solutions	**PO9**	Individual and team work
PO4	Conduct investigations of complex problems	**PO10**	Communication
PO5	Modern tool usage	**PO11**	Project Management and Finance
PO6	The Engineer and Society	**PO12**	Life long learning

Assessment Pattern

Bloom's Category	Continuous Assessment Tests		End Semester Examination Marks (%)
	Test 1 (%)	Test 2 (%)	
Remember	30	30	30
Understand	30	30	30
Apply	40	40	40
Analyze			
Evaluate			
Create			

Mark Distribution

Total Marks	CIE Marks	ESE Marks	ESE Duration
150	50	100	3

Continuous Internal Evaluation Pattern:

Attendance **10 marks**

Continuous Assessment Tests (Average of SeriesTests1& 2) **25 marks**

Continuous Assessment Assignment **15 marks**

Internal Examination Pattern:

Each of the two internal examinations has to be conducted out of 50 marks. The first series test shall be preferably conducted after completing the first half of the syllabus and the second series test shall be preferably conducted after completing the remaining part of the syllabus. There will be two parts: Part A and Part B. Part A contains 5 questions (preferably, 2 questions each from the completed modules and 1 question from the partly completed module), having 3 marks for each question adding up to 15 marks for part A. Students should answer all questions from Part A. Part B contains 7 questions (preferably, 3 questions each from the completed modules and 1 question from the partly completed module), each with 7 marks. Out of the 7 questions, a student should answer any5.

End Semester Examination Pattern:

There will be two parts; Part A and Part B. Part A contains 10 questions with 2 questions from each module, having 3 marks for each question. Students should answer all questions. Part B contains 2 full questions from each module of which student should answer any one full question. Each question can have maximum 2 sub-divisions and carries 14 marks.

SYLLABUS

Module – 1(Line and Circle drawing algorithms)

Basics of Computer Graphics and its applications. Video Display devices - Refresh Cathode Ray Tubes, Random Scan Displays and systems, Raster scan displays and systems, Color CRT displays, Flat panel display and its categories. Line drawing algorithms - DDA, Bresenham's algorithm. Circle drawing algorithms - Midpoint Circle generation algorithm, Bresenham's algorithm.

Module - 2(Filled Area Primitives and Two dimensional transformations)

Filled Area Primitives- Scan line polygon filling, Boundary filling and flood filling. Two dimensional transformations-Translation, Rotation, Scaling, Reflection and Shearing, Composite transformations, Matrix representations and homogeneous coordinates.

Module - 3 (Clipping and 3D transformations)

Window to viewport transformation. Cohen Sutherland and Midpoint subdivision line clipping algorithms, Sutherland Hodgeman and Weiler Atherton Polygon clipping algorithms. Three dimensional viewing pipeline. Basic 3D transformations.

Module - 4 (Projections and Visible Surface detection)

Projections- Parallel and Perspective projections. Visible surface detection algorithms- Back face detection, Depth buffer algorithm, Scan line algorithm, A buffer algorithm

Module - 5 (Realism and performance)

Realism - Illumination Shading, Shadows, Texture mapping, Bump mapping, Environment mapping, Transparency, Accumulation Buffer, Back face Culling, Visibility Culling.

Text Books

1. Donald Hearn and M. Pauline Baker, Computer Graphics, PHI, 2e, 1996
2. Aditi Majumder and M.Gopi , Introduction to VISUAL COMPUTING Core Concepts in Computer Vision, Graphics, and Image Processing, 2018

References

1) William M. Newman and Robert F. Sproull, Principles of Interactive Computer Graphics. McGraw Hill, 2001
2) Zhigang Xiang and Roy Plastock, Computer Graphics (Schaum's outline Series), McGraw Hill, 2019.
3) David F. Rogers , Procedural Elements for Computer Graphics, Tata McGraw Hill,2001.
4) Donald Hearn, M. Pauline Baker and Warren Carithers, Computer Graphics with OpenGL, PHI, 4e, 2013

Course Level Assessment Questions

Course Outcome 1 (CO1):

1. Compare the working principle of raster scan systems and random scan systems.
2. How much time is spent scanning across each row of pixels during screen refresh on a raster system with resolution of 1280*1024 and a refresh rate of 60 frames per second?

Course Outcome 2 (CO2):

1. Rasterize the line with end points accepted from the user(2,3) and (5,8) using Bresenham's line drawing algorithm and implement it using any appropriate programming language. (Assignment)
2. Illustrate how the 4-connected area filling approach differs from 8- connected area filling in boundary filling algorithm and implement it using any appropriate programming language.(Assignment)

Course Outcome 3 (CO3):

1. Rotate a triangle ABC 45 degree counter clockwise about the pivot point (10,3) , where the position vector of the coordinate ABC is given as A(4,1), B(5,2) and C(4,3).
2. Implement the above transformation using any appropriate programming language with user inputs. (Assignment)

Course Outcome 4 (CO4):

1. Given a clipping window A(20,20), B(60,20), C(60,40) and D(20,40). Using Cohen Sutherland algorithm, find the visible portion of the line segment joining the points P(40,80) and Q(120,30).
2. Implement Cohen Sutherland clipping algorithm using any appropriate programming language with user inputs. (Assignment)

Course Outcome 5 (CO5):

1. Explain scan line algorithm for detecting visible surfaces in an object.

Course Outcome 6 (CO6):

1. You are rendering a black and white checkered tiled floor using a single texture mapped polygon. The view is simulating a person standing on the floor and looking at a point far away from him on the floor. (1)Artifacts at the distant end of the floor can be seen. How would you remove these artifacts? (2) How can you explain why this method works using the sampling theorem?
2. You are seeing an object which is either texture mapped, bump mapped or displacement mapped but you don't know which one. However, you have the liberty to move the light and the viewpoint of an object and see it from different angles and for different positions of the light. How will you figure out which technique was used?

Model Question Paper

QP CODE:

Reg No: _____

Name: _____ **PAGES : 4**

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

FOURTH SEMESTER B.TECH DEGREE (HONOURS) EXAMINATION, MONTH & YEAR

Course Code:ADT296

Course Name: Advanced Topics in Computer Graphics

Max. Marks : 100 **Duration: 3 Hours**

PART A

Answer All Questions. Each Question Carries 3 Marks

1. Consider a raster system with a resolution of 1024*1024. Compute the size of the raster needed to store 4 bits per pixel? How much storage is needed if 8 bits per pixel are to be stored?
2. How 8-way symmetry of circle can be used for writing circle drawing algorithms? Write the symmetric points if (x, y) is a point on the circle with centre at origin.
3. Show that two successive reflections about either of the coordinate axes is equivalent to a single rotation about the coordinate origin.
4. Determine a sequence of basic transformations that are equivalent to the x-direction shearing matrix.

5. Find the window to viewport normalization transformation with window lower left corner at (1,1) and upper right corner at (2,6).

6. How does Cohen Sutherland algorithm determine whether a line is visible, invisible or a candidate for clipping based on the region codes assigned to the end points of the line?

7. Define the terms (i) Centre of projection (ii) Principal vanishing point

8. Differentiate between the object space and image space method for the hidden surface removal of an image.

9. Describe the steps used to convert the normal map to bump mapping.

10. One artifact of Gouraud shading is that it can miss specular highlights in the interior of the triangles. How can this be explained as an aliasing artifact? **(10x3=30)**

Part B

(Answer any one question from each module. Each question carries 14 Marks)

11. (a) Derive the initial decision parameter of Bresenham's line drawing algorithm and rasterize a line with endpoints (2,2) and (10,10). **(8)**

(b) Draw the architecture of raster scan display systems and explain its working principle **(6)**

OR

12. (a) Explain the working principle of a Refresh CRT monitor with suitable diagrams. **(7)**

(b) Write Midpoint circle drawing algorithm and plot a circle with radius=20 and center (50,30) using the algorithm. **(7)**

13. (a) Differentiate between boundary fill and flood fill algorithms. **(5)**

(b) Reflect a triangle ABC about the line 3x-4y+8=0, where the position vector of the coordinate ABC is given as A(4,1), B(5,2) and C(4,3). **(9)**

OR

14. (a) A diamond shaped polygon is located at P(-1,0), Q(0,-2), R(1,0) and S(0,2). Find the transformation matrix which would rotate the triangle by 90 degree counter clockwise about the point Q. Using the transformation matrix, find the coordinates of the rotated polygon. **(7)**

(b) Illustrate the working principle of scan line polygon filling algorithm **(7)**

15. (a) Illustrate Weiler – Atherton polygon clipping algorithm. **(6)**

(b) Explain Cohen-Sutherland line clipping algorithm. Use the algorithm to clip line P1 (70, 20) and P2(100,10) against a window lower left hand corner (50,10) and upper right hand corner (80,40). **(8)**

OR

16. (a) Describe the steps required for a general 3D rotation if the rotation axis is not parallel to any one of the principal axis. The rotation axis is defined by the points P1(x1,y1,z1) and P2(x2,y2,z2). Give its composite matrix representation **(6)**

(b) Describe Sutherland Hodgeman polygon clipping algorithm and list out its limitations **(8)**

17. (a) Explain how visible surfaces can be detected using depth buffer algorithm. **(7)**

(b) Define parallel projection. Describe orthographic and oblique parallel projection. **(7)**

OR

18. (a) Illustrate the scan line method used in visible surface detection. **(7)**

(b) Explain the steps involved in performing perspective projections **(7)**

19. (a) Specify any three shading algorithms used in interactive graphics. **(6)**

(b) Explain the procedure of texture to object space mapping. **(8)**

OR

20. (a) Explain the mapping scheme in which the effects of small bumps on the surface of an object can be simulate without changing the number of primitives **(8)**

(b) Describe about object to screen space mapping. **(6)**

TEACHING PLAN

No	Contents	No of Lecture Hrs
Module – 1 (Line and Circle drawing algorithms) (10 hrs)		
1.1	Basics of Computer Graphics and applications	1
1.2	Refresh Cathode Ray Tubes	1
1.3	Random and Raster Scan Displays and systems,	1
1.4	Color CRT displays	1
1.5	Flat panel display and its categories.	1
1.6	DDA Line drawing Algorithm	1
1.7	Bresenham's line drawing algorithm	1
1.8	Midpoint Circle generation algorithm	1
1.9	Bresenham's Circle generation algorithm	1
1.10	Illustration of line and circle drawing algorithms	1
Module - 2 (Filled Area Primitives and Two dimensional transformations) (9 hrs)		
2.1	Scan line polygon filling	1
2.2	Boundary filling and flood filling	1
2.3	Basic 2D transformations-Translation	1
2.4	Basic 2D transformations- Rotation	1
2.5	Basic 2D transformations- Scaling	1
2.6	Reflection and Shearing	1
2.7	Illustration of Basic 2D Transformations	1
2.8	Composite transformations	1
2.9	Matrix representations and homogeneous coordinates	1
Module - 3 (Clipping and 3D transformations) (8 hrs)		
3.1	Window to viewport transformation	1
3.2	Cohen Sutherland Line clipping algorithm	1
3.3	Midpoint subdivision Line clipping algorithm	1
3.4	Sutherland Hodgeman Polygon clipping algorithm	1
3.5	Weiler Atherton Polygon clipping algorithm	1
3.6	Three dimensional viewing pipeline	1

3.7	Basic 3D transformation-Translation and scaling	1
3.8	Basic 3D transformation-Rotation	1
Module - 4 (Projections and Visible Surface detection) (7 hrs)		
4.1	Projections-Parallel projections	1
4.2	Projections- Perspective projections	1
4.3	Illustration of projection methods	1
4.4	Visible surface detection algorithms- Back face detection	1
4.5	Depth buffer algorithm	1
4.6	Scan line visible surface detection algorithm	1
4.7	*A buffer* algorithm	1
Module - 5 (Realism and performance)(10 hrs)		
5.1	Illumination	1
5.2	Shading and Shadows	1
5.3	Texture mapping-Texture to object space mapping	1
5.4	Texture mapping-Object to screen space mapping and Mip Mapping	1
5.5	Bump mapping	1
5.6	Bump mapping-Illustration	1
5.7	Environment mapping and Transparency	1
5.8	Accumulation Buffer and Back face Culling	1
5.9	Visibility Culling	1
5.10	Visibility Culling	1