SOMAIYA VIDYAVIHAR UNIVERSITY

08·05·2024 (E)

Semester: Jan 2024 - May 2024		
Maximum Marks: 100 **Examination: ESE Examination**		**Duration:3 Hrs.**
Programme code: 01/04 **Programme:** B. Tech Computer/IT Engineering	**Class:** SY	**Semester: IV (SVU 2020)**
Name of the Constituent College: K. J. Somaiya College of Engineering	**Name of the department: Computer/IT**	
Course Code:116U01C401/116U04C401	**Name of the Course:** Probability, Statistics and Optimization Techniques	
Instructions: 1) All questions are compulsory 2) Assume suitable data wherever necessary		

Que. No.	Question	Max. Marks
Q1	Solve any **Four** of the following.	**20**
i)	Consider 3 candidates X, Y, Z, with chances of getting appointed are 4: 2: 3 respectively. The probability that Mr X is selected and company become profitable is 0.3. The probability of Y & Z for the same is 0.5 & 0.8 respectively. If the company become profitable then what is the probability that Mr Z is appointed? (Solve by Bayes' Theorem)	5
ii)	Obtain the Spearman's rank correlation coefficient (R) from the following data. X : 20, 22, 28, 28, 25, 50 Y : 32, 38, 45, 45, 70, 45	5
iii)	A random sample of 200 observations has mean 6.5 cm. Can it be a random sample From population whose mean is 7 cm and variance is 8.5 cm at 2% LOS? (Two tailed Z test with critical value 2.33)	5
iv)	Convert the given LPP into the standard form Minimise $z = -3x_1 + 2x_2 - x_3$ Subject to $x_1 - 3x_2 + 2x_3 \geq -6,\ 3x_1 + 4x_3 \leq 3, -3x_1 + 5x_2 \leq 4$ Where $x_1, x_2 \geq 0$ and x_3 is unrestricted in sign.	5
v)	Arrivals at telephone booth are considered to be Poisson with an average time of 10 min. between one arrival and the next .The length of phone calls is assumed to be distributed exponentially with a mean of 3 min. **(a)** What is the probability that a person arriving at the booth will have to wait? **(b)** What is the probability that it will take a customer more than 10 min. altogether to wait for the phone and complete his call? Assume The system is (M/M/1/∞)	5
vi)	Given $f_{xy}(x,y) = \begin{cases} cx(x-y), & 0 < x < 2, -x < y < x \\ 0, & elsewhere \end{cases}$ **(a)** Evaluate c **(b)** find $f_x(x)$ (Marginal distribution with respect to x)	5

Q2 A	Solve the following.	**10**
i)	If $\sigma_y = \sigma_y = \sigma$ and the angle between the lines of regression is $\tan^{-1} 3$, find the coefficient of correlation.	5
ii)	A die was thrown 132 times and the following frequencies were observed No obtained: 1 | 2 | 3 | 4 | 5 | 6 | Total Frequency: 15 | 20 | 25 | 15 | 29 | 28 | 132 Use χ^2 test to test the hypothesis that the die is unbiased at 5% LOS.	5
	OR	
Q2 A	Using Lagrange's Multiplier method solve the following NLPP $z = x_1^2 + x_2^2 + x_3^2 - 10x_1 - 6x_2 - 4x_3$ Subject to $x_1 + x_2 + x_3 = 7, x_1, x_2, x_3 \geq 0$.	10
Q 2 B	Solve any **One** of the following.	**10**
i)	A diesel pump has capacity to accommodate only 4 trucks including the one at the pump. On an average trucks arrive at the rate of 5 per hour and the service rate is 6 per hour. Assume that the arrival process is Poisson and the service time is an exponential random variable. **(a)** What is the average time for which a truck is at the pump? **(b)** What is the average waiting time for a truck in the queue? **(c)** What percentage of trucks will be turned away?	10
ii)	It is known that the probability of an item produced by a certain machine will be defective is 0.05. If the produced item are send to the market in packets of 20, find the number of packets containing **(a)** at least 2 **(b)** exactly 2 & **(c)** at most 2 defective items in a consignment of 1000 packets using Binomial distribution & Poisson distribution.	10
Q3	Solve any **Two** of the following.	**20**
i)	**(a)** A manufacturer knows from his experience that the resistance of resistor he produces is normal with $\mu = 100$ ohms & standard deviation $\sigma = 2$ ohms. What percentage of resistors will have resistance between 98 ohms & 102 ohms?	05
	(b) A sample of 200 fish of a particular kind taken as random from one end of a lake had mean weight of 20 lbs & standard deviation of 2 lbs. At the other end of the lake, a sample of 80 fish of the same kind had mean weight of 20.51lbs & standard deviation of 2 lbs. Is the difference between the mean weights significant at 1% level of significance?	05
ii)	Find the equations of two regression lines for the following data X: 78 | 36 | 98 | 25 | 45 | 82 | 90 | 62 | 65 | 39 Y: 84 | 51 | 91 | 60 | 68 | 62 | 86 | 58 | 53 | 47 Estimate Y when $X = 40$ and Estimate X when $Y = 80$.	10

iii)	Define the following terms: Solution, Basic solution, Basic Feasible solution and degenerate solution of LPP. Also find **(a)** All basic solutions **(b)** All feasible basic solutions **(c)** All degenerate solutions. Hence decide the optimal feasible basic for the following L.P.P. Maximise $z = 2x_1 - 2x_2 + 4x_3 - 5x_4$ Subject to $x_1 + 4x_2 - 2x_3 + 8x_4 = 2,\ -x_1 + 2x_2 + 3x_3 + 4x_4 = 1$ where $x_1, x_2, x_3, x_4 \geq 0$	10
Q4	Solve any **Two** of the following.	**20**
i)	For a normal variate X with mean 2.5 and standard deviation 3.5 find the probability that **(a)** $P(2 \leq X \leq 4.5)$ **(b)** $P(-1.5 \leq X \leq 5.5)$ **(c)** $P(\lvert X \rvert \leq 2)$ **(d)** Find c such that $P(X < c) = 0.2$	10
ii)	In a certain experiment to compare two types of pig foods A & B the following results of increasing weights were obtained Pig no: 1 2 3 4 5 6 7 8 Increase in weight by A: 49 53 51 52 47 50 52 53 Increase in weight by B: 52 55 52 53 50 54 54 53 Can we conclude that food B is better than food A if the same set of pigs were used in both cases? (Use one tailed test at 5% LOS)	10
iii)	Solve the given LPP by Penalty method (Big M) Minimize $z = 2x_1 + x_2$ Subject to $3x_1 + x_2 = 3,\ 4x_1 + 3x_2 \geq 6,\ x_1 + 2x_2 \leq 3$ where $x_1, x_2 \geq 0$	10
Q5	Solve any **Four** of the following.	**20**
i)	Suppose a random number N is taken from 690 to 850 in uniform distribution .Find the probability that number N is greater than 790. Find the Mean and the Variance of the distribution.	5
ii)	Calculate the two regression coefficients and the coefficient of correlation from the data: $N = 10, \sum x = 350, \sum y = 310,\ \sum(x-35)^2 = 162,$ $\sum(y-31)^2 = 222,\ \sum(x-35)(y-31) = 35$	5
iii)	The mean value of random sample of 20 items was found to be 145 with standard deviation of 40. Find 90% confidence limits for the population mean.	5
iv)	Find the relative maximum or minimum of the function $z =$ $Z = 9x_1 + 6x_3 + x_1x_2 - {x_1}^2 - {x_2}^2 - {x_3}^2$	5
v)	Find the probability that a customer has to wait in an M/M/1/∞ model if $\lambda = 8$, $\mu = 10$ per hour. Also find the probability that a customer has to wait in the queue more than 15 minutes.	5
vi)	Obtain the dual of given LPP Maximize $z = 2x_1 - x_2 + 3x_3$ Subject to $x_1 - 2x_2 + x_3 \geq 4,\ 2x_1 + x_3 \leq 10,\ x_1 + x_2 + 3x_3 = 20,\ x_1, x_2, x_3, \geq 0$	5

SOMAIYA

10.05.2024 (E)

Maximum Marks: 100	Semester: January 2024 –April 2024 Examination: ESE Examination	Duration:3 Hrs.
Programme code: 04 Programme: B. Tech	Class: SY	Semester: IV(SVU 2020)
Name of the Constituent College: K. J. Somaiya College of Engineering	Name of the department: IT	
Course Code: 116U04C402	Name of the Course: Information Theory and Coding	
Instructions: 1)Draw neat diagrams 2) All questions are compulsory 3) Assume suitable data wherever necessary		

Que. No.	Question	Max. Marks
Q1	Solve any **Four**	**20**
i)	Describe a symmetric channel. How do we find the capacity?	5
ii)	A source emits one of the four symbols S0, S1, S2 and S3 with probabilities **1/3, 1/6, 1/4** and **1/4** respectively. The successive symbols emitted by the source are statistically independent. Calculate the entropy of the source.	5
iii)	The binary symbol '0' and '1' are transmitted with probabilities 1/4 and 3/4 respectively. Find the corresponding self-information.	5
iv)	Define information, average information, information rate, symbol rate.	5
v)	What is Mutual Information? Mention its properties.	5
vi)	A Binary Channel has the following characteristics: $P(Y/X) = \begin{bmatrix} 2/3 & 1/3 \\ 1/3 & 2/3 \end{bmatrix}$ If the input symbols are transmitted ' ' with probabilities ¾ and ¼ respectively. Find entropies H(X), H(X,Y) and H(Y/X).	5

Que No.	Question	Max. Marks
Q2 A	Solve the following	**10**
i)	Explain briefly with relevant example: a) Hamming code b) Cyclic Code	5
ii)	What are the properties of Cyclic codes?	5
	OR	
Q2 A	Consider a (6,3) linear code whose generator matrix is $G = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}$ a) Find all code vectors b) Find all hamming weight and distances c) Find Parity Check Matrix d) Draw encoder circuit for the above codes	10

Q 2 B	Solve any **One**	10
i)	Describe Standard array. How it is used in syndrome decoding? Explain with an example	10
ii)	For a systematic (6,3) Linear Block Code, the parity matrix 'P' for the received code vector R=[110010] is given by: $P = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$ Detect and correct the single error that has occurred due to noise.	10

Que. No.	Question	Max. Marks
Q3	Solve any **Two**	20
i)	Using Huffman Encoding Algorithm, find the code words for the given symbols. Determine the code efficiency and redundancy of the code. S_0: 0.55, S_1: 0.15, S_2: 0.15, S_3: 0.1, S_4: 0.05	10
ii)	Explain Image compression and its types. Explain Run Length Coding with example	10
iii)	Given messages X_1, X_2, X_3, X_4, X_5 and X_6 with probabilities **0.4, 0.2, 0.2, 0.1, 0.07** and **0.03** respectively. Construct a binary code by applying Shannon Fano Encoding Procedure. Determine the code efficiency and redundancy of the code.	10

Que. No.	Question	Max. Marks
Q4	Solve any **Two**	20
i)	Find the value of x for the given set of congruent equations using Chinese Remainder theorem. $x \equiv 2 \pmod 3$ $x \equiv 3 \pmod 5$ $x \equiv 2 \pmod 7$	10
ii)	Use the Vigenere cipher with the keyword "VIG" to encrypt and decrypt the following message "THE BOY HAS THE BALL"	10
iii)	What is substitution cipher? Show Affine Cipher encryption and decryption process with an example.	10

Que. No.	Question	Max. Marks
Q5	(Write notes / Short question type) on any **four**	20
i)	Explain LZW coding with example.	5
ii)	Explain Trellis Diagram.	5
iii)	Define entropy. State its properties.	5
iv)	What is Joint Probability Matrix? Mention its properties	5
v)	Explain Linear Block code with diagram	5
vi)	Shannon's characteristic of good cipher	5

13/5/2024(E)

SOMAIYA

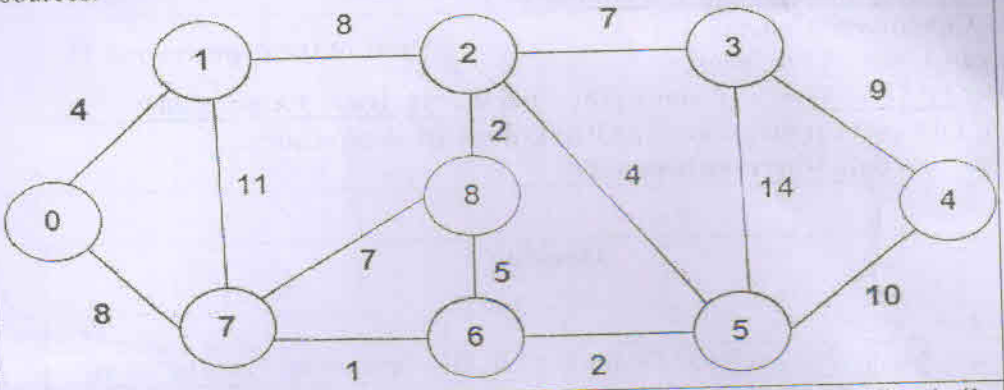
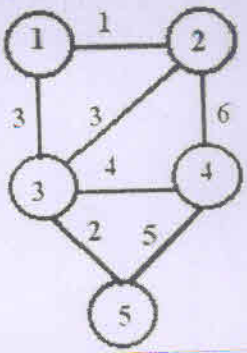
Semester: January 2024 – April 2024			
Maximum Marks: 100	Examination: ESE Examination		Duration:3 Hrs.
Programme code: 04			
Programme: B. Tech Information Technology	Class: SY	Semester: IV(SVU 2020)	
Name of the Constituent College: K. J. Somaiya College of Engineering		Name of the department: IT	
Course Code: 116U04C403	Name of the Course: Analysis of Algorithms		
Instructions: 1)Draw neat diagrams 2) All questions are compulsory 3) Assume suitable data wherever necessary			

Que. No.	Question	Max. Marks
Q1	Solve any Four	20
i)	What is Asymptotic analysis? Define Big Oh, Big Omega and Theta notation?	5
ii)	Derive the complexity of counting sort.	5
iii)	Solve the following using master method $T(n) = 8T(n/2) + n^2$	5
iv)	Write the difference between greedy method and dynamic programming.	5

v)	Determine the Time Complexity for all statement in the following algorithm Segment.
	
	I=1;
	While(I<=n)
	{
	X=X+I;
	I=I+1;
	}
		5

vi)	Define P class, NP class problems	5

Que. No.	Question	Max. Marks
Q2 A	Solve the following	10
i)	Write an algorithm to find all pairs shortest path using dynamic programming.	5
ii)	Write short note on optimal binary search tree.	5
	OR	
Q2 A	Find longest common subsequence for following strings: using LCS algorithm. X= ababcde Y= bacadb	10
Q 2 B	Solve any One	10
i)	Apply the quick sort algorithm to sort the list (E, X, A, M, P, L, E) in alphabetical order. Find Time complexities of quick sort in worst case using Recursion tree.	10
ii)	Sort the list of the elements 10,5,7,6,1,4,8,3,2,9 using merge sort algorithm and show its computing time is O(n logn).	10

Que. No.	Question	Max. Marks
Q3	Solve any **Two**	20
i)	Obtain the solution to 0/1 knapsack problem by Greedy method n=7, m=15 (p1,P2,...P7)=(10,5,15,7,6,18,3), (W1,W2,...W7)=(2,3,5,7,1,4,1).	10
ii)	Apply Dijkstra's algorithm on the following graph. Consider vertex 0 as sources. 	10
iii)	Find the cost of minimal spanning tree of the given graph by using Kruskal's Algorithm and also write its algorithm. 	10

Que. No.	Question	Max. Marks
Q4	Solve any **Two**	20
i)	What do you understand by NP Complete? Explain is subset sum problem is NP complete or not with justification?	10
ii)	Describe the travelling sales person problem and discuss how to solve it using dynamic programming with example.	10
iii)	Write an algorithm to solve N Queens problem. Show its working for N = 4.	10

Que. No.	Question	Max. Marks
Q5	Write notes on any **four**	20
i)	Job sequencing with deadlines	5
ii)	Branch and Bound strategy	5
iii)	Dynamic Programming	5
iv)	15 Puzzle Problem	5
v)	Strassen's matrix multiplication	5
vi)	Hamiltonian cycles	5

15/5/2024 (E)

SOMAIYA VIDYAVIHAR UNIVERSITY

Semester: January 2024 –April 2024		
Maximum Marks: 100 **Examination: ESE Examination**		**Duration:3 Hrs.**
Programme code: 04 **Programme:BTech in IT**	**Class:SYBTech**	**Semester: IV (SVU 2020)**
Name of the Constituent College: **K. J. Somaiya College of Engineering**	**Name of the department: IT**	
Course Code: 116U04C404	**Name of the Course: Advanced** Databases	
Instructions: 1)Draw neat diagrams 2) All questions are compulsory **3) Assume suitable data wherever necessary**		

Que. No.	Question	Max. Marks
Q1	Solve any **Four**	**20**
i)	Explain how skew in partitioning is handled using virtual processor partitioning.	5
ii)	Explain concept of distributed query optimization.	5
iii)	Explain the concept of array and multisets in ORDBMS.	5
iv)	Write short note on spatial SQL queries.	5
v)	TYPE 2 changes in dimensions of datawarehouse. Explain.	5
vi)	Explain OLAP models: MOLAP, ROLAP, HOLAP.	5

Que. No.	Question	Max. Marks
Q2 A	Solve the following	**10**
i)	Explain Heterogeneous distributed database systems with diagram.	5
ii)	Explain range partitioning parallel sort in brief.	5
	OR	
Q2 A	Consider the distributed database given below. Site 1: **EMPLOYEE** Fname | Minit | Lname | Ssn | Bdate | Address | Sex | Salary | Super_ssn | Dno 10,000 records each record is 100 bytes long Ssn field is 9 bytes long — Fname field is 15 bytes long Dno field is 4 bytes long — Lname field is 15 bytes long Site 2: **DEPARTMENT** Dname | Dnumber | Mgr_ssn | Mgr_start_date 100 records each record is 35 bytes long Dnumber field is 4 bytes long — Dname field is 10 bytes long Mgr_ssn field is 9 bytes long Consider the query "For each employee, retrieve employee name and department name,Where the employee works" submitted from site 2 then by applying the optimization criteria as minimizing data transfer find out the best strategy and cost (in terms of number of bytes transferred) for given distributed query.	10
Q 2 B	Solve any **One**	**10**
i)	Explain different techniques used for partitioning data across multiple disks for parallel I/O with examples.	10

ii)	Write advantages of distributed databases system and parallel DB systems.	10

Que. No.	Question	Max. Marks
Q3	Solve any **Two**	**20**
i)	Explain creation of structured type using nested types. Create table using this nested type and insert two records in this table.	10
ii)	Explain Vector and Raster data models in detail.	10
iii)	How No-SQL databases are different from the SQL databases? Explain the concept of Wide column family NoSQL database with proper diagram and example.	10

Que. No.	Question	Max. Marks
Q4	Solve any **Two**	**20**
i)	Differentiate a. OLAP and OLTP b. ETL and ELT	10
ii)	Explain significance of metadata in Data Warehousing.	10
iii)	Design a snowflake schema for university database. Consider at least 4 dimensions and one fact table with 2 measures.	10

Que. No.	Question	Max. Marks
Q5	(Write notes / Short question type) on any **four**	**20**
i)	Spatial relations.	5
ii)	Document database.	5
iii)	Data Transformation for data standardization.	5
iv)	Types of data extraction for ETL.	5
v)	Types of Data Marts.	5
vi)	Data Lake .	5