INDIAN STATISTICAL INSTITUTE

Mid-Semestral Examination: 2012-2013

Course: M.Tech. (CS) I

Subject: Introduction to programming

Date: 03/09/2012 Maximum Marks: 20 Duration: 2 hours

Answer all

1. Write a program to convert any decimal number in binary format. 6

2. Is there any difference between following declarations? Explain if any.
   a. extern int myfunction();
   b. int myfunction ();

   3

3. Which bitwise operator is suitable for turning off a particular bit in a number? 1

4. Is there any maximum combined length of the command-line arguments including the spaces between adjacent arguments? 1

5. List the FILE structure members. 5

6. Is always *switch* statement better than multiple *if* statements? Explain your answer. 4

INDIAN STATISTICAL INSTITUTE

PERIODICAL EXAMINATION
M.TECH.(CS) I YEAR

ELEMENTS OF ALGEBRAIC STRUCTURES

03|09|1

Date: 03.09.2012 Maximum marks: 70 Duration: 2 hrs

The paper contains 85 marks. Answer as much as you can, the maximum you can score is 70.

1. (a) Let $\mathbb{N} = \{0, 1, 2, \ldots\}$. Show that there is an injection from $\mathbb{N} \times \mathbb{N}$ to $\mathbb{N}$.
   (b) Find the multiplicative inverse of 32 modulo 75.
   (c) Compute the value of $(p-1)^{p^2+p-2} \bmod p$ where p is a prime.

(7 + 7 + 6 = 20)

2. Suppose that N and M are two normal subgroups of a group G and $N \cap M = \{e\}$. Show that for any $n \in N$ and $m \in M$, $nm = mn$.

(10)

3. Let G be a group. If H is a subgroup of G and N is a normal subgroup of G, then show that $H \cap N$ is a normal subgroup of H.

(10)

4. Let

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 4 & 7 & 8 & 2 & 5 & 3 & 1 & 6 \end{pmatrix} \text{ and } \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 4 & 1 & 8 & 2 & 6 & 3 & 7 & 5 \end{pmatrix}$$

Are these two conjugate permutations? If so, exhibit a conjugating permutation.

(10)

5. Let G be a cyclic group of order 10. What are the automorphisms of G? Justify your answer.

(10)

6. Let $\mathbb{R}$ be the set of real numbers and for $a, b \in \mathbb{R}$, $a \neq 0$, let $\tau_{a,b} : \mathbb{R} \to \mathbb{R}$ be defined by $\tau_{a,b}(x) = ax + b$. Let

$$G = \{\tau_{a,b} : a, b \in \mathbb{R}, a \neq 0\} \text{ and } N = \{\tau_{1,b} \in G\}.$$

Assuming the operation on G to be composition of functions, show that N is a normal subgroup of G and that G/N is isomorphic to the group of non-zero reals under multiplication.

(10)

7. Let A and B be two subgroups of a finite group G. Define a relation $\sim$ on G as follows: for $x, y \in G$, $y \sim x$ if $y = axb$ for some $a \in A$ and $b \in B$.

(a) Show that $\sim$ is an equivalence relation on G and the equivalence classes of $\sim$ are AxB for $x \in G$.

(b) Show that for any $x \in G$, xBx^{-1} is a subgroup of G and $o(xBx^{-1}) = o(B)$.

(c) Show that for any $x \in G$,

$$\#AxB = \#AxBx^{-1} = \frac{o(A)o(xBx^{-1})}{o(A \cap xBx^{-1})}.$$

(d) If A and B are of size p^α such that $p^\alpha || o(G)$, show that $A = gBg^{-1}$ for some $g \in G$.

$(2 + 3 + 3 + 7 = 15)$

INDIAN STATISTICAL INSTITUTE
Mid - Semester Examination : 2012-13

M. Tech. (CS) I Year
Discrete Mathematics

Date: 05.09.2012 **Full Marks : 100** **Time: 3 hours**

N O T E : There are two groups of questions.
Use separate answer sheet for each group.

GROUP - A (Total marks 38)

Answer as many as you can.

1 Let R be a binary relation on the set of all positive integers such that

$R = \{ \text{a, b} \mid \text{a - b is an positive integer} \}$

Is R
i) reflexive ? ii) symmetric ? iii) antisymmetric ? iv) transitive ?

[8]

2 Show that all natural numbers greater than 6 can be expressed as the sum of distinct primes.

[5]

3 Prove that $\sum_{i=0}^{n} F_i = F_{n+2} - 1$, where F_i is the i^{th} Fibonacci number.

[5]

4 Prove by using the law of contrapositive "If the product of two integers is odd, then both must be odd integers".

[6]

5 Translate the following in propositional logic:

i) Unless Sita comes to the party, Mita will not be happy.

ii) If Usha jumps and Anju does not make a leap, Saina will have to do a gigantic step.

[2 x 3 = 6]

6 Prove or disprove the following by tree method:

Let the universe of discourse be $\{0, 1\}$. Then $\exists \text{y} \ \left[(\text{y} - 1)^2 \neq \text{y}^2 - 1 \right]$.

[8]

*****

Group – B (total marks = 62)

Note : There are alternatives in several questions. Carefully identify those alternatives.

1. Give a combinatorial proof of the following identity :

$$\binom{m}{0}\binom{n}{0}+\binom{m}{1}\binom{n}{1}+ \dots + \binom{m}{n}\binom{n}{n} = \binom{m+n}{n}$$

[6]

OR

Show that the number of different pairings possible for the first round of a tennis tournament with $2n$ participants is equal to $\frac{(2n)!}{2^n n!}$

[6]

2. a) Suppose 3 identical dice are rolled. What is the total number of different possible outcomes ?

b) In how many ways a given integer n can be partitioned into exactly r parts, where order counts ?

[5+5 = 10]

3. a) Prove that

$$1(1!)+2\ (2!)\ +3(3!)+\dots+n(n!)=(n\ +\ 1)\ !-1$$

Hence, show that for any given positive integer n, there is a unique set of positive integers $a_1, a_2, a_3, \dots$, such that $n = a_1\ (1!)\ +\ a_2\ (2!) + a_3\ (3!) + \dots$, where $a_i < i$ for all $i \geq 1$.

[4+6=10]

b) Let $p_o\ (n)$ and $p_d\ (n)$ be the total number of ways of partitioning n in odd parts and distinct parts, respectively. Show that $p_o\ (n) = p_d\ (n)$.

[5]

OR

3. Let $p(n)$ = total number of ways of partitioning n where order does not count. Show that $p(n) < e^{3\sqrt{n}}$.

[15]

4. In how many different ways can a man climb up a ladder having 20 rungs if at each step, he can climb either one rung or two rungs ?

[5]

5. There are 10 people standing in a queue at a box office. Admission fee is Rs. 50 per person, and only 5 of the people have exactly this amount each. The other 5 each have exactly a 100 rupee bill. Unfortunately, the box office starts off with no change. A sequence of 10 people is workable if at each point, the box office has the correct change to pay each person, who needs it. Calculate the total number of such workable sequences.

[8]

6. Using the principle of inclusion and exclusion, calculate the total number of numbers less than n which are co-prime to n.

[8]

7. In how many ways, 6 apples, 1 orange, 1 pear, 1 peach, 1 plum, 1 strawberry and 1 grape can be divided among 3 people ? There is no restriction on the distribution; a person may get all of these items or none of these.

[10]

OR

a) Let $\left\{ n \atop r \right\}$ denote the Stirling number of the second kind which gives the number of ways of distributing n distinct balls in r unlabelled boxes so that no box remains empty. Show that $\left\{ n \atop r \right\} = r \left\{ n-1 \atop r \right\} + \left\{ n-1 \atop r-1 \right\}$.

b) Hence, find the value of $\left\{ 5 \atop 2 \right\}$.

[6+4=10]

*****

INDIAN STATISTICAL INSTITUTE

First Mid-Semestral Examination: 2012-13

Subject Name : **Probability and Stochastic Process** DATE- 07.09.12
Course Name : M.Tech. (CS) I yr. Maximum Score: 30 Duration: 3 Hours

Note: Attempt all questions. Marks are given in brackets. State the results clearly you use. Use separate page for each question.

Problem 1 (3). We pick at random two letters without replacement from the letters of "PROBABILITYTHEORY". Find the probability that one of the letter is 'B'.

Problem 2 (2). Suppose $P(X = x, Y = y) = \frac{c}{xy}$ for $1 \leq x, y \leq n$ and $c \in \mathbb{R}$. Express the probability mass function (p.m.f.) of X without using c.

Problem 3 (4). Prove or disprove the following statement.

For a probability function $P(\cdot)$ and events A_1, A_2 and A_3, if $P(A_1 \cap A_2 \cap A_3) = P(A_1) \times P(A_2) \times P(A_3)$ then at least one pair of the events A_1, A_2 and A_3 is independent.

Problem 4 (3+3). Suppose we draw a ball, say numbered X_1, "at random" from a box containing n balls numbered from 1 to n. We then throw away all balls with larger numbers (i.e. $X_1 + 1, X_1 + 2, \ldots, n$) and return the ball numbered X_1 to the box. We again draw a ball, say X_2, "at random" from the box. Given that the second ball has number 1, what is the probability that the first ball has number n? How do you guess the number for the first ball, given that you have observed that the second ball number is i, for some $1 \leq i \leq n$.?

Problem 5 (4). Let $X \sim Poi(\theta)$. For any $x, n \geq 0$, compute $P(X = n + x | X \geq x)$.

Problem 6 (4). Let $X_1, \ldots, X_k \overset{iid}{\sim} unif(\{0, 1, \ldots, n\})$ and let $Y = X_1 + \ldots + X_k$. Prove that for all $0 \leq i \leq nk$, $P[Y = i] = P[Y = nk - i]$.

Problem 7 (3+3). Consider the following random experiment for a set $S = \{x_1, \ldots, x_n\}$: For each $1 \leq i \leq n$, Alice and Bob independently picks x_i "at random" with probability p and q respectively.

(i) Write down the sample space and corresponding p.m.f. of the combined experiment.
(ii) Compute the probability that there is no common element picked by both.

Problem 8 (4). A novel is viewed as a finite sequence of English letters $a, b, \ldots, z$ (ignoring all other symbols or replacing capital letters by corresponding smaller letters). Suppose a monkey is typing the letters "at random" indefinitely. What is the probability that the monkey at some point of time types the complete novel *Hamlet* continuously?

INDIAN STATISTICAL INSTITUTE

Periodical Examination

M. Tech (CS) - I Year (Semester - I)

Data and File Structures

Date : 10.9.2012 Maximum Marks : 100 Duration : 3.5 Hours

Note : You may answer any part of any question, but maximum you can score is 100.

1.(a) Consider the linked list representation of a polynomial of one variable where

(i) each term is stored in separate node,

(ii) the terms are arranged in increasing order of their exponents, and

(iii) each node has only one link field.

Write an efficient algorithm for create a separate polynomial $C = P \times Q$, such that P and Q remains as it is. You may use some extra scalar locations if necessary. Give an estimate of the number of nodes of P, Q and C visited by your algorithm.

(b) Given a doubly-linked list containing n integer information. Write an efficient algorithm to sort the elements of the list in increasing order of their values. Analyze the number of comparisons and exchanges made by your algorithm. [10+10 = 20]

2.(a) Suppose that you are given a binary three which consists of a linear chain of left children. Suppose further that the tree has exactly $n = 2k - 1$ nodes, for some integer $k \geq 1$. Derive an algorithm which using only AVL single left- and right-rotations, maps this tree into a perfectly balanced complete tree. Explain how your algorithm works and give an example.

(b) You are given a binary search tree, where each node contains a pointer *left* to its left child, a pointer *right* to its right child, a pointer *parent* to its parent, and a *key*, which is an integer that can be positive, negative or zero. Describe a procedure that performs the following operation, and then write pseudo-code for it: The operation is given a pointer to the root of the tree, and needs to output the number of nodes who have the property that the sum of the keys in the nodes sub-tree is positive. The procedure should take $O(n)$ time, where n is the number of nodes in the tree. [12+13=25]

3.(a) Consider an application where the key values are stored at the leaf level, and the intermediate nodes store the *discriminant values* to guide the search path. Define weight balanced (α, where $0 < \alpha \leq 1/2$) ($WB[\alpha]$) tree to maintain a balanced search structure in this environment. What is its advantage over height balanced tree.

(b) Show that for a $WB[\alpha]$ tree if $\frac{1}{3} < \alpha < \frac{1}{2}$, then $WB[\alpha] = WB[\frac{1}{2}]$.

(c) How do you dynamically maintain (perform insertion and deletion) such a tree? Describe a scheme such that the amortized cost of insertion and deletion of a node in a $WB[\alpha]$-tree is $O(\log n)$. [6+6+13=25]

4. Define a *B-tree* of order m.

Consider a B-tree of order m and containing n key values residing in a disk. Show that the maximum number of nodes (h) to be accessed from disk to retrieve a key value x satisfies the relation $h \leq \log_{\lceil \frac{m}{2} \rceil} \frac{n+1}{2}$.

Write an algorithm for reporting all the key values in an interval $[a, b]$. Mention the time complexity of your algorithm in terms of k, m and n, where k is the number of reported answers, m and n are as mentioned above. [4 + 8 + 8 = 20]

5.(a) Describe briefly the different techniques for address resolution if collision occurs during hashing. State the merits and demerits of each of these techniques.

(b) Show that, in the open addressing scheme of collission resolution, the expected number of probes required for a successful search is

$$\frac{1}{\lambda} \log_e \frac{1}{1-\lambda}$$

where λ is the load factor of the hash table ($\frac{n}{m}$, n and m are respectively the number of keys present in the hash table and the size of the hash table respectively).

(c) Give an idea of maintaining hash table in secondary storage, where the record size is fixed for all records. [9+9+7=25]

INDIAN STATISTICAL INSTITUTE

SEMESTRAL-I EXAMINATION (2012-13)
M.TECH.(CS) I YEAR

ELEMENTS OF ALGEBRAIC STRUCTURES

Date: 19.11.12 Maximum marks: 100 Duration: 3 hours

The paper contains 120 marks. Answer as much as you can, the maximum you can score is 100.

1. (a) Define the greatest common divisor (gcd) of two positive integers.
   (b) Describe an efficient algorithm to compute the gcd of two positive integers. Show the correctness of your algorithm and find its time complexity.
   (c) Show that any subgroup of a cyclic group is also cyclic.
   (d) If a cyclic subgroup T of G is normal in G, then show that every subgroup of T is normal in G.

   $(2 + 10 + 6 + 7 = 25)$

2. (a) Show that a commutative ring D is an integral domain if and only if for $a, b \in D$ with $a \neq 0$, the relation $ab = ac$ implies $b = c$.
   (b) If R is a ring with identity and ϕ is a homomorphism of R *onto* R' prove that $\phi(1)$ is the identity of R'.
   (c) Let R be a ring with identity, R not necessarily commutative, such that the only right-ideals of R are (0) and R. Prove that R is a division ring.
   (d) Let F be the field of real numbers. Prove that $F[x]/(x^2+1)$ is a field and it is isomorphic to the field of complex numbers.
   (e) If R is an integral domain with identity, prove that any unit in $R[x]$ must already be a unit in R.

   $(6 + 6 + 6 + 6 + 6 = 30)$

3. (a) Let F_5 be the field of 5 elements and consider the following vectors from F_5^4.

   $$(0,1,2,0), (1,3,2,0), (1,0,4,3), (2,4,0,3).$$

   Are these vectors linearly independent?
   (b) Let U and V be vector spaces over the same field and let T be a homomorphism of U onto V with kernel W. Prove that there is a one-to-one correspondence between the subspaces of V and the subspaces of U which contain W.
   (c) Suppose that V is a finite-dimensional vector space and $T : V \to V$ is a homomorphism which is not onto. Prove that there is some $v \neq 0$ in V such that $T(v) = 0$.

(d) Give an example of a 3×3 matrix whose minimal polynomial is not equal to its characteristic polynomial. Justify your answer.

(e) Suppose V is finite dimensional inner product space having dimension n. Suppose that for $m < n$, $\{w_1, \ldots, w_m\}$ is an orthonormal set in V. Prove that there exist vectors $w_{m+1}, \ldots, w_n$ such that $\{w_1, \ldots, w_m, w_{m+1}, \ldots, w_n\}$ is an orthonormal set.

$(6 + 6 + 6 + 6 + 6 = 30)$

4. (a) Find an irreducible polynomial of degree 4 in $F_3[x]$, where F_3 is the finite field of three elements.

(b) Let α be a primitive element in a field of cardinality p^d and let $m(x)$ be the minimal polynomial of α. Find the degree of $m(x)$.

(c) Let $F = GF(2)$ be the finite field of two elements and $\tau(x)$ be an irreducible polynomial of degree d in $F[x]$.

i. Let K be an extension of F such that there is an $\alpha \in K$ satisfying $\tau(\alpha) = 0$. Show that the zeros of $\tau(x)$ are

$$\alpha, \alpha^2, \alpha^{2^2}, \ldots, \alpha^{2^{d-1}}.$$

ii. Show that $\tau(x) = \tau(1 + x)$ if and only if $\tau(x) | (x^{2^i} + x + 1)$ for some i.

$(5 + 10 + 7 + 13 = 35)$

INDIAN STATISTICAL INSTITUTE

First Semester Examination: 2012-13

Course Name: M.Tech (CS) I

Subject Name: Introduction to Programming

Date: 21/11/2012 **Maximum Marks: 100** **Duration: 3hours**

The default data type is integer. Mention your assumptions, if any, clearly.

1. Write a C program to find the transpose of a given ***n x n*** matrix ***M***. The matrix ***M*** should be dynamically created after reading the value from the key board. Your program should store the resultant in ***M*** only. Do not use any additional matrix.
20
2. Write a C program to find the ***n***th smallest element of a given unsorted array ***A*** having ***e*** element.
15
3. Write a C program to count the number of vowels in a given text file.
15
4. Write a C program to print "ISI" ***n*** times without using any conditional operators.
10
5. Write a C program using recursion to find greatest common divisor of two numbers given as command line argument.
15
6. Are the following declarations different? Explain.
   a. ***char a[]*** and ***char *a***
5
7. Is the following statement a correct C syntax? Explain.
   a. ***iA>?= iB;***
5
8. Is it possible to rename any function in c? Explain.
5
9. Answer in short
   A. How would you insert pre-written code into a current C program?
   B. The first expression in a ***for*** loop is?
   C. What is the ***continue*** statement used for?
   D. Why we cannot initialize ***extern*** variables?
   E. Which storage class is not allowed with ***structure***?

2x5=10

INDIAN STATISTICAL INSTITUTE

Semestral Examination

M. Tech (CS) - I Year (Semester - I)

Data and File Structures

Date : 30.11.2012 Maximum Marks : 100 Duration : 4 Hours

Note : You may answer any part of any question, but maximum you can score is 100.

1. In the online version of the convex hull problem, the number of points in the set P is not known in advance, and they come one by one (but no point is deleted during the execution). At every instance, we may need to update the convex hull. State a data structure for maintaining the convex hull, where the size of the data structure is proportional to the number of hull vertices at the present instant of time, and $O(\log n + k)$ time is enough to update the data structure during the insertion of a point, where k is the total number of insertions and deletions of the hull vertices on the existing one to get the revised one. [8]

2. Consider the UNION-FIND problem, where each set is represented by a tree, and the set name is written at the root of that tree. The *union by rank* heuristic for the union operation of two sets works as follows. Let U and V be two sets represented by two trees rooted at u and v respectively. If $|U| < |V|$, then u is attached to v; otherwise, v is attached to u. The name of the new set is writen at the root (v or u) of the resulting tree, and its size is set to $|U| + |V|$.

   Show that if the *union by rank* heuristic is followed, then the height of the tree of a set containing k nodes can be at most $\lfloor \log k \rfloor$.

   State the method of path compression when a sequence of union and find operations are executed. What is the advantage of this method with respect to the execution time of the algorithm ? Justify your answer.

   Also describe the data structure to be used for storing the sets in the implementation of the above two methods for the UNION-FIND problem. [5+12+5=22]

3.(a) Define skip list. Give an example of a skip list with a set S of distinct integer key values where $|S| = 20$.

(b) Show that the expected search time for a key in a skip list of size n is $O(\log n)$. Mention the assumptions clearly which are made for constructing the skip list.

(c) Also show that the expected space complexity of your method of organizing the skip list is $O(n)$.

(d) How can insertions and deletions be made in the skip list maintaining the aforesaid search time and space complexities ? [6+4+4+6=20]

4.(a) An orthogonal range tree for a set of n points in the plane is given. State an algorithm for the axis-parallel rectangular range query which reports the points inside the query rectangle in $O(k + \log^2 n)$ time, where k is the number of reported answers. Explain the details of the data structure and the algorithms for (i) preprocessing and (ii) query, along with the complexity analysis for each of these items.

(b) If each point is attached with a weight, and the objective is to get the sum of weights of all the points inside the query rectangle, then what are the modifications needed in the data structure so that the query can be answered in $O(\log^2 n)$ time ?

(c) Formulate a scheme of storing the orthogonal range tree in a disk using records of same size such $k = O(n)$, and an axis-parallel rectangular range query can be performed with $O(\log \log n)$ number of disk accesses. [(6+5+5)+5+9=30]

5.(a) Information about a set of employees and their dates of joining an institute is available. The institute administration frequently needs to know the names of all the employees whose date of joining is on or after some year y. Construct a data structure for this type of problem such that the above query can be answered efficiently. State the preprocessing time, space, and query time complexities for your proposed data structure with proper justification.

(b) The same institute authority is now interested to find all the employees whose date of joining is after some year y and whose monthly salary lies in between x_1 and x_2 for some x_1 and x_2. Design a linear space data structure which can answer this type of query efficiently. State the preprocessing and the query algorithms clearly, with proper justification of preprocessing time, query time, and space complexities. State whether your proposed data structure can answer the previous type of query (stated in 5(a)) as well. If yes, then state the query time complexity.

(c) Movie time and duration for different television channels are different. Viewers frequently ask which movies are going on at a particular time. Design an efficient data structure for this type of query. State preprocessing time, space and the query time complexities of your proposed data structure with proper justification. [5+(8+5)+7=25]

6.(a) A bit-array $A[]$ of size n consisting of 0s and 1s is given. RANK(i) for $i \in \{1, 2, \ldots n\}$ is defined as the number of 1's in the array $A[]$ upto $A[i]$. On the other hand, SELECT(i) for $1 \leq i < n$ is the position of ith 1 inside the array $A[]$. Design a space-efficient data structure which can answer both RANK and SELECT queries efficiently.

(b) How many bits are needed to store a traditional binary tree with n nodes? How to represent that tree such that it will need $O(n)$ bits ? State how to access the parent and children for any node of a tree that is thus represented. [10+5=15]

INDIAN STATISTICAL INSTITUTE

First Semester Examination: 2012-13

Subject Name : **Probability and Stochastic Processes** DATE – 23.11.12
Course Name : M.Tech. (CS) I yr. Maximum Score: 60 Duration: 3 Hours

Note: Attempt all questions carrying 68 marks, maximum you can score 60. Marks are given in brackets. State the results clearly you use. Use separate page for each question.

Problem 1 (5 + 5 = 10). If $P[X = x, Y = y, Z = z] = f(x, z)g(y, z)$ then show that X and Y are conditionally independent given Z. State and prove the converse.

Problem 2 (8). Let X and Y be independent random variables, each having the geometric distribution with parameter p. Find $E(Y|X + Y = n)$, $n \geq 2$.

Problem 3 (10). Let $Y_0, Y_1, \ldots$ be a sequence of independent, identically distributed random variables on natural numbers. Show that $(X_n)_{n\geq 0}$ is a Markov chain with stationary transition probability where $X_0 = Y_0$ and $n \geq 1$,

$$X_n = \begin{cases} X_{n-1} - Y_n, & \text{if } X_{n-1} > 0; \\ X_{n-1} + Y_n, & \text{if } X_{n-1} \leq 0. \end{cases}$$

Problem 4 (6 + 6 = 12). Consider a Markov chain with $N = \{0, 1, 2 \ldots\}$ as state space and transition probabilities $p_{i,j}$, $i, j \in N$ given by

$$p_{i,j} = \begin{cases} 1, & \text{if } i = j = 0; \\ p, & \text{if } i = j > 0; \\ q, & \text{if } i - j = 1; \\ 0, & \text{otherwise }; \end{cases}$$

where $p + q = 1, p, q > 0$. Find $P(T_0 = n|X_0 = j)$ for $j \in N$, and show that $E(T_0|X_0 = j) = j/q$, where T_0 is the hitting time of state 0

Problem 5 (6 + 4 =10). (i) Let $X := (X_1, \ldots, X_n)$ where $X_i \overset{i.i.d.}{\sim} ber(p)$ and $Z = \sum_i X_i$ then compute the conditional entropy $H(X|Z)$.

(ii) Prove that for any real number a there is a continuous random variable X such that the differential entropy $H(X) = a$.

Problem 6 (8). Let M_n denote the number of comparisons in randomized quick sort for n elements and let $a_n = E(M_n)$. Find a recurrence relation on a_n. (Recall that quick sort is a recursive algorithm in which we choose a number at random, say x_i, and then make three groups (L, x_i, R) where L and R are the set of all numbers less and greater than x_i respectively. Then we recursively sort L and R.)

Problem 7 (10). Let $X \sim N(0, 1)$. Find the p.d.f. of $\sqrt{|X|}$.

P.T.O

1 Syllabus

Probability, conditional probability and independence; Random variables and their distributions (discrete and continuous), bivariate and multivariate distributions; Laws of large numbers, central limit theorem (statement and use only).

Stochastic process: Definition and examples of stochastic processes, Markov chains with finite and countable state spaces classification of states, birth and death processes, branching processes, queuing processes.

Entropy, Differential entropy, maximization of entropy.

Indian Statistical Institute

First Semester Examination (2012-2013)

M.Tech. (CS) First Year

Discrete Mathematics

Date: November 26, 2012 — Maximum Marks: 100 — Time: 3 hours

Answer as many questions as you can, but the maximum you may score is 100.

Marks alloted to each question are indicated within square brackets near the right margin.

1. Prove that the graph G_1 shown below does not have a Hamiltonian cycle. [4]

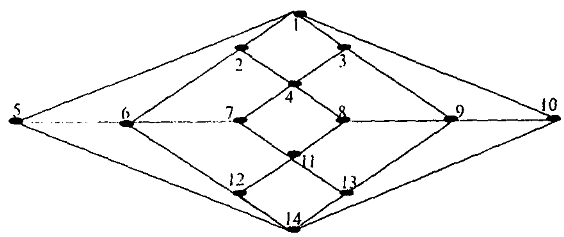

Figure 1: G_1

2. (a) Consider the following graph G_2:

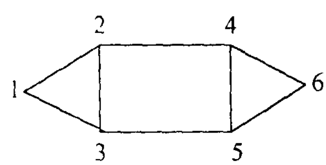

Figure 2: G_2

   (i) How many fundamental cut-sets does G_2 have with respect to a spanning tree?

   (ii) How many fundamental circuits does G_2 have with respect to a spanning tree?

   (b) Show that every circuit in a graph has an even number of edges in common with any cut-set.

   (c) Show that every cut-set in a connected graph G must contain at least one edge of every spanning tree of G. [(2+2)+3+3=10]

3. (a) Let G be a maximal planar graph, and n_i be the number of vertices of G having degree i. Prove that $\sum_i (6-i)n_i = 12$.

(b) Show that all wheel graphs are self-dual.

(c) What are the values of genus γ and crossing number ν for Petersen graph shown below?

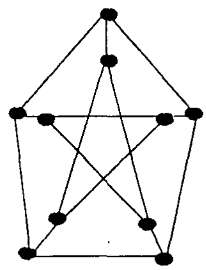

Figure 3: Petersen graph

(d) Write a predicate logic formula for the following statement:
Every outerplanar graph with at least seven vertices has a non-outerplanar complement.

[7+7+(2+2)+7=25]

4. (a) Consider G, the join of two graphs G_1 and G_2. Prove that the chromatic number $\chi(G) = \chi(G_1) + \chi(G_2)$.

(b) Argue whether $x^4 - 4x^3 + 3x^2$ is a chromatic polynomial of a graph.

(c) Show that for any graph G, $\beta_1(G) \leq \alpha_0(G) \leq 2\beta_1(G)$, where $\beta_1(G)$ and $\alpha_0(G)$ denote respectively the cardinalities of a maximum matching and a minimum vertex cover of G. For any $k \in \mathbb{N}$, construct a graph with $\alpha_0(G) = 2k$, and $\beta_1(G) = k$.

[5+5+(6+4)=20]

5. (a) Prove or disprove: For each $n > 1$, every simple digraph with n vertices has two vertices with the same outdegree or two vertices with the same indegree.

(b) (i) Find the values of κ and λ of the graph G_1 in Figure ~~2~~ ~~below~~ above.
(ii) Show that if G is a regular graph of degree r and $\kappa = 1$, then $\lambda \leq \lceil r/2 \rceil$.

[6+(4+5)=15]

6. (a) Find the total solution for the recurrence $a_r = 7a_{r-1} + 18(2^{2r-2})$, $r \geq 1$. Assume $a_0 = 1$.

(b) Consider the following list of statements about a book:
(i) There are three statements in this list.
(ii) Two of them are not true.
(iii) The average increase in IQ scores of those who read this book is more than 20 points.
Is statement (iii) **true**? Justify your answer.

[10+6=16]

7. Give a combinatorial proof of the following identity:

$$\binom{m+n}{r} = \binom{m}{0}\binom{n}{r} + \binom{m}{1}\binom{n}{r-1} + \cdots + \binom{m}{r}\binom{n}{0}.$$

[5]

8. (a) For the Fibonacci numbers, show that $F_{n+m} = F_m F_{n+1} + F_{m-1} F_n$. Hence, show that F_{5k}, $k > 0$ is a multiple of 5.

   (b) Show that $F_{n-1}^2 + F_{n+1}^2 = F_{2n-1}$.

[6+4=10]

9. (a) Show that the number of partitions of an integer n into k or fewer parts is equal to the number of partitions of n into parts of size at most k.

   (b) Compute the total number of valid ways of parenthesizing a sequence of n distinct integers $< a_1, a_2, \ldots, a_n >$ to find their product, assuming that multiplication of two integers is the primitive operation. For example, all possible ways of parenthesizing the sequence $< a_1, a_2, a_3 >$ are: $(a_1(a_2a_3))$ and $((a_1a_2)a_3)$.

[4+6=10]

10. Suppose 4-digit numbers are printed on a set of identical rectangular tags with one 4-digit number per tag. Find the smallest number of distinctly printed tags needed to cover all the 10^4 numbers? For example, the tag with 0066 may also be read as 9900. [6]

INDIAN STATISTICAL INSTITUTE

Back-paper Examination

M. Tech (CS) - I Year (Semester - I)

Data and File Structures

Date : 26.12.2012 | Maximum Marks : 100 | Duration : 3.5 Hours

Note : You may answer any part of any question, but maximum you can score is 100.

1.(a) Give an $O(n\log k)$ time algorithm for merging k sorted lists into one sorted list, where n is the total number of elements in all the input lists.

(b) Suppose you have two heaps each containing $(2^k - 1)$ elements. Design an efficient algorithm for merging these two heaps into a single heap. Mention (with justification) the time complexity of your algorithm.

(c) Consider a binary heap with n nodes. How long does it take to find the third smallest key in the heap? (You do not have to delete this key, just report its value.) Explain briefly. [8+8+4=20]

2.(a) Suppose you are given a *height balanced binary search tree* T, and you have to print the key values of that tree in sorted order. Write an algorithm for this problem assuming that recursion is not supported in your environment.

(b) Define $WB[\alpha]$ (weight balanced) binary tree.

What are its distinct advantages and disadvantages over the height balanced binary tree ? [10+(5+5)=20]

3. Define a B-tree of order m.

Consider a B-tree of order m stored in a disk that contains n key values. Show that the maximum number of nodes (h) to be accessed from disk to retrieve a key value x satisfies the relation $h \leq \log_{\lceil \frac{m}{2} \rceil} \frac{n+1}{2}$.

Write an algorithm for inserting a key value in a B-tree. Mention the number of disk accesses made by your algorithm in terms of m and n. [4 + 8 + (9+3) = 24]

4. Consider the problem of inserting the keys 10, 22, 31, 4, 15, 28, 17, 88, 59, 63 into a hash table of length $m = 11$ using open addressing with the primary hash function $h_1(k) = k \; mod \; m$. Illustrate the result of inserting these keys using (i) linear probing, (ii) quadratic probing with $c_1 = 1$ and $c_2 = 3$, and (iii) double hashing with $h_2(k) = 1 + (k \; mod \; (m-1))$. [5 × 3 = 15]

5. Define *range tree* for 2 dimensional range searching. For a given set of n tuples, what will be the time and space complexities for constructing the range tree ?

Can you achieve an $O(\log n + k)$ time algorithm for one-dimensional range searching in a two-dimensional range tree, where the query range is an interval $[\alpha_1, \alpha_2]$ on the (i) x-axis and (ii) y-axis? [(5+8)+(5+7) = 25]

6. What is the difference between internal and external fragmentation?

Give pseudocode for a procedure dealloc(p) that deallocates a block of memory with starting address at p that was allocated using the dynamic storage allocation method described in the class.

You may assume that simple utilities have been provided for manipulating the doubly-linked available space list. Remember that after deallocation, there should be no two consecutive available blocks. You may assume that the block being deallocated is not the first or the last block in the heap. Your procedure should run in O(1) time.

As always, do not just give the code. Explain how your code works, and if helpful give an illustration of the various cases that arise. [12]

INDIAN STATISTICAL INSTITUTE
First Semestral Back Paper Examination: 2012-13

27/1

Subject Name : **Probability and Stochastic Process**
Course Name : M.Tech. (CS) I yr. Maximum Score: 100 Duration: 3 Hours

Note: Attempt all questions. Marks are given in brackets. State the results clearly you use. Use separate page for each question.

Problem 1 (10). Let $X_1, \ldots, X_n \overset{i.i.d.}{\sim} \text{unif}(\{0,1\}^m)$. For any subset $S \subseteq \{1, 2, \ldots, n\}$ define $Y_S := \oplus_{i \in S} X_i$. Prove that Y_S's are pairwise independent.

Problem 2 (10 + 5 = 15). Compute the moment generating function of negative binomial distribution and hence find its expectation.

Problem 3 (3 + 10). State and prove the weak law of large numbers.

Problem 4 (8 + 8 = 16). We choose a number repeatedly at random from $\{1, 2, \ldots, 1000\}$ until the number is divisible by 4. Let T be the number of trials required to stop and X denote the final number. Compute the probability mass function of X. Compute $E(T)$.

Problem 5 (8 + 8 = 16). Prove that the fair random walk is a markov process with stationary transition probabilities. Prove that all states are recurrent.

Problem 6 (5 + 10 = 15). Compute the differential entropy of $N(\mu, \sigma^2)$ and prove that the normal density has maximum entropy among all densities with a fixed variance σ^2.

Problem 7 (15). Let $M, N \overset{i.i.d.}{\sim} \text{Poi}(\theta)$. Compute $E(M^N)$.

INDIAN STATISTICAL INSTITUTE

SEMESTRAL-I BACK PAPER EXAMINATION (2012-13)
M.TECH.(CS) I YEAR

ELEMENTS OF ALGEBRAIC STRUCTURES

Date: 28/12/12 Maximum marks: 100 Duration: 3 hours

The paper contains 100 marks. Each question carries 10 marks. Answer all questions.

1. Let m and n be positive integers such that $\gcd(m, n) = 1$. Show that given any two integers a and b, there exists an integer x such that $x \equiv a \bmod m$ and $x \equiv b \bmod n$.

2. Let S_3 be the group of all permutations of the set $\{a, b, c\}$. In S_3 show that there are two elements x and y such that $(x \cdot y)^2 \neq x^2 \cdot y^2$.

3. Let G be any group and g a fixed element of G. Define $\phi : G \to G$ by $\phi(x) = gxg^{-1}$. Prove that g is an isomorphism of G onto G.

4. Give an example of an integral domain which is of positive characteristic, but, has an infinite number of elements.

5. If R is a ring with identity 1 and ϕ is a homomorphism of R *onto* R', prove that $\phi(1)$ is the identity of R'.

6. Find the greatest common divisor of the following polynomials over the field of rational numbers.
$$x^2 + 1 \text{ and } x^6 + x^3 + x + 1.$$

7. Let F be the field of real numbers and let V be the set of all sequences $(a_1, a_2, \ldots, a_n, \ldots)$, $a_i \in F$, where equality, addition and scalar multiplication are defined componentwise. Prove that V is a vector space over F.

8. If F is the field of real numbers, prove that the vectors $(1, 1, 0, 0)$, $(0, 1, -1, 0)$ and $(0, 0, 0, 3)$ in F^4 are linearly independent over F.

9. Let V be an inner product space and $\{w_1, \ldots, w_m\}$ is an orthonormal set in V. Show that
$$\sum_{i=1}^{m} |\langle w_i, v\rangle| \leq \|v\|^2 \quad \text{for any } v \in V.$$

10. Show that the polynomial $\tau(x) = x^3 + x + 1$ is irreducible over F_2, the field of two elements. Write down the elements of $F_2[x]/(\tau(x))$.

Indian Statistical Institute
Mid-semestral Examination: 2012-2013
(first year second semester)
Course Name: M. Tech in Computer Science
Subject Name: Operating System

Date: 18-02-2013 Maximum Marks: 30 Duration: 2 hours

Instructions:
You **may** attempt **all** questions which carry a total of **36** marks. However, the maximum marks you can score is only **30**.

1. (a) Consider the following set of processes, with the length of the CPU burst given in milliseconds:

Process	Burst Time	Priority
P_1	10	3
P_2	1	1
P_3	2	3
P_4	1	4
P_5	5	2

The processes are assumed to have arrived in the order P_1, P_2, P_3, P_4, and P_5 all at time 0.

i. What is the turnaround time of each process for each of the scheduling algorithms FCFS, SJF and a nonpreemptive priority (a smaller priority number implies a higher priority)? [2]

ii. What is the waiting time of each process for each of the scheduling algorithms in part (i)? [2]

(b) Discuss how the following pairs of scheduling criteria conflict in certain settings.

i. CPU utilization and response time [1]

ii. Average turnaround time and maximum waiting time [1]

iii. I/O device utilization and CPU utilization [1]

(c) Why is it important for the scheduler to distinguish I/O-bound programs from CPU-bound programs? [2]

(d) How do clustered systems differ from multiprocessor systems? What is required for two machines belonging to a cluster to cooperate to provide a highly available service? [1+2]

2. (a) What is a thread? Give some benefits of multithreaded programming. [1+2]

(b) Consider a multiprocessor system and a multithreaded program written using the many-to-many threading model. Let the number of user-level threads in the program be more than the number of processors in the system. Discuss the performance implications of the following scenarios.

i. The number of kernel threads allocated to the program is less than the number of processors. [1.5]

P.T.O

ii. The number of kernel threads allocated to the program is equal to the number of processors. [1.5]

iii. The number of kernel threads allocated to the program is greater than the number of processors but less than the number of user level threads. [1.5]

(c) Describe the differences between symmetric and asymmetric multiprocessing. What are the advantages of multiprocessor systems? [1+2]

(d) Using the following program, explain what the output will be at Line A. [1.5]

```
#include< sys/types.h >
#include< stdio.h >
#include< unistd.h >
int value = 5;
int main()
{
pid_t pid;
pid=fork();
if(pid == 0)     /* child process */
{
  value = value + 15;
  return(0);
}
else if (pid > 0)     /* parent process */
{
  wait(NULL);
  printf("value=%d",value);     /* Line A */
  return(0);
}
}
```

3. (a) Describe the actions taken by a kernel to contex-switch between processes. [2]

(b) List out some reasons for which a parent may require to terminate the execution of one of its children. [2]

(c) What is a pipe? List out the important issues that must be considered while implementing a pipe. Give an example of a situation in which ordinary pipes are more suitable than named pipes and an example of a situation in which named pipes are more suitable than ordinary pipes. [1+2+2]

(d) Briefly describe the mechanism for interprocess communication using remote procedure calls. [3]

Indian Statistical Institute

Semester-II 2012-2013

M.Tech.(CS) - First Year

Mid-semester Examination (20 February, 2013)

Subject: Automata, Languages and Computation

Maximum marks: 45 Total marks: 50 Duration 3 hrs.

To the extent possible, please be formal in your proofs, arguments, etc.

1. (a) Formally define
   (i) a non-deterministic finite automaton (NFA),
   (ii) the extended transition function δ^* for NFAs,
   (iii) acceptance of a string by an NFA.

   (b) When is a relation over strings said to be right invariant?

   (c) When is a context free grammar (CFG) said to be ambiguous?

   [(3+3+1) + 2 + 2 = 11]

2. Draw the state diagram of a deterministic finite automaton (DFA) that recognises the set of all binary strings where the difference between the number of 0s and the number of 1s is even. [6]

3. Let Σ be a finite alphabet, L a language over Σ, and $h : \Sigma \to \Sigma^*$ a homomorphism. Prove that (i) $h(h^{-1}(L)) \subseteq L$; (ii) $h^{-1}(h(L)) \supseteq L$. [5 + 5 = 10]

4. Let L be a regular language over an alphabet Σ. Prove that there exists a constant N such that if y is any string of length N and there are strings x and z such that $xyz \in L$, then y can be written as $y = uvw$ such that $|v| \geq 1$ and for each $i \geq 0$, $xuv^iwz \in L$. [7]

5. Let $\Sigma = \{0, 1, \approx, \boxplus\}$, and let

   $$L = \{a \approx b \boxplus c \mid a, b, c \in (0+1)^* \text{ and } a = b+c \text{ if } a, b, c \text{ are interpreted as unsigned binary integers}\}.$$

   Note that the string $10 \approx 01 \boxplus 01$ is in L since $2 = 1 + 1$, but the string $10 \approx 11 \boxplus 01$ is not since $2 \neq 3 + 1$. Show that L is not regular. [6]

6. Let L be the set of strings consisting of balanced pairs of square and round brackets. Thus, the strings ( ), [ ], ([ ]), and (( )) [ ] all belong to L, but ([ )] $\notin L$.

   (a) Give an inductive definition for L. Assume that $\varepsilon \notin L$.

   (b) Using (a), construct a context free grammar for L. Your grammar should have only one non-terminal.

   [7 + 3 = 10]

INDIAN STATISTICAL INSTITUTE
Mid-Semester Examination: 2012 – 13

Course Name: M. TECH. (CS) I YEAR

Subject: Computer Networks

Date: 22.02.13

Maximum Marks: 60 Duration: 2 hrs. 30 minutes

Answer as much as you can

1. a) Consider an application which transmits data at a steady rate (e.g., the sender generates a N bit unit of data every k time units, where k is small and fixed). Also, when such an application starts, it will stay on for relatively long period of time. Would a packet-switched network or a circuit-switched network be more appropriate for this application? Why?

b) Why are the network architecture layered? A system has n layer protocol hierarchy. User application generates messages with M bytes. At each of the layers, an h-byte header is added. What fraction of the network bandwidth is wasted with headers for such message transmission?

c) Suppose that x-bits of user data are to be transmitted over a k-hop path in a packet-switched network as a series of packets. Each packet containing p data bits and h header bits, with x being much greater than $p + h$. The bit rate of the transmission lines is b bps and the propagation delay is negligible. Find the total time for the packets to arrive at the destination.

3+(2+4)+6 = 15

2.a) Consider a slotted ALOHA system having four stations. If the respective offer loads for these stations are $G_1 = 0.1$, $G_2 = 0.5$, $G_3 = 0.2$, and $G_4 = 0.2$ packets per second, find the individual throughput rates for each station and the total throughput.

b) A small slotted ALOHA system has only k customers, each of whom has a probability $1/k$ of transmitting during any slot (original+re-transmissions combined). What if the channel throughput as a function of k? Evaluate its value for $k = 2$ and $k \rightarrow \infty$.

c) If the bit string 011110111110111110 is subjected to bit stuffing, what is the output string?

6+6+3 = 15

3. a) i) Given the address 132.6.17.85, find the network address.
ii) Find the class in the following address: 11110011 10011011 11111011 00001111.

b) A router outside the organization receives a packet with destination address 190.240.7.91 /16. Show how it finds the network address to route the packet.

c) A router inside the organization receives the same packet with destination address 190.240.33.91 /19. Show how it finds the subnetwork address to route the packet.

(3+3)+4+5 = 15

4. a) An organisation is granted the block 130.56.0.0/16. The administrator wants to create 1024 subnets:
i. Find the subnet mask required
ii. Find the number of addresses in each subnet

iii. Find the first and last allocatable addresses in subnet 1
iv. Find the first and last allocatable addresses in subnet 32

b) Frames of 1000 bits are sent over a 1 Mbps satellite channel. Assume that the propagation delay over a satellite channel is 270 msec. Acknowledgements are always piggybacked onto data frames. The headers are very short. Three bit sequence numbers are used. What is the maximum achievable channel utilisation for Go back-n protocol? 10+5 = 15

5. a) Consider building a CSMA/CD network running at 1 Gbps over a 1km cable with no repeaters. The signal speed (signal propagation time) in the cable is 200,000 km/sec. What is the minimum frame size?

b) A 3000 km long T1 trunk is used to transmit 64-byte frames using a Go-back-n protocol. If the propagation speed is 6 μsec/km, how many bits should the sequence number be?

c) A code scheme has a Hamming distance $d_{min} = 4$. What is the error detection and correction capability of this scheme? Explain briefly. 4+7+4 = 15

6. a) Ten signals, each requiring 4000 Hz, are multiplexed on to a single channel using FDM. How much minimum bandwidth is required for the multiplexed channel? Assume that the guard bands are 400 Hz wide.

b) Find the minimum hamming distance in the following code table:

Dataword	*Codeword*
00	00000
01	01011
10	10101
11	11110

Discuss whether this qualifies as a linear block code.

c) Assume the following parameters for a switching network:
N = number of hops between two given stations
L = message length, in bits
B = data rate, in bps, on all links
P = packet-size, in bits
H = overhead (header) bits per packet
S = call setup time (circuit switching or virtual circuit) in seconds
D = propagation delay per hop in seconds

Derive general expressions for end-to-end delay for the four techniques: circuit switching, virtual circuit packet switching and datagram packet switching. Assume that there are no acknowledgements. 3+4+8 = 15

7. Write short notes on (Any 2): $7\frac{1}{2}\times 2 = 15$

a) 1-persistent CSMA
b) Code Division Multiple Access
c) Token Ring Networks
d) Loopback with IPv4 Addresses.

INDIAN STATISTICAL INSTITUTE

Periodical Examination

M. Tech (CS) - I Year (Semester - II)

Design and Analysis of Algorithms

Date : February 26, 2013 Maximum Marks : 60 Duration : 3 Hours

Note : You may answer any part of any question, but maximum you can score is 60.

1. You are given an array of n integers. You need to find the median ($\lceil \frac{n}{2} \rceil$-th smallest element) of these integers. Write an efficient algorithm for this problem using minimum amount of extra space. Analyze the worst case time and space complexities of your algorithm. [10]

2. Consider the following algorithm for finding the largest element in an array $A[1\ldots, n]$.

```
RANDOMMAX(A[1...n]):
max = ∞;
for i = 1 to n in random order
    if A[i] > max
        max = A[i]          (*)
return max
```

   (a) In the worst case, how many times does RANDOMMAX execute line (*)?

   (b) What is the exact probability that line (*) is executed during the last iteration of the for loop?

   (c) What is the exact expected number of executions of line (*)?

   Justify your answers. [3+5+7=15]

3. You are given a set of points $P = \{p_1, p_2, \ldots, p_n\}$ on a 2D plane, and a real number ρ. Write an expected $O(n)$ time algorithm to test the existance of a pair of points in P whose distance is less than or equal to ρ. Mention explicitly the assumptions we are making to justify the time complexity of your algorithm. [10]

4. You are given a connected undirected graph $G = (V, E)$ in which the weight of each edge is either 1 or 2. Present an $O(V + E)$ time algorithm to compute a minimum spanning tree for G. Explain the correctness of your algorithm, and analyze its time complexity. [10]

5. Recall that a bridge is an edge in a connected, undirected graph $G = (V, E)$ whose removal causes the graph to become unconnected. Present an $O(|V| + |E|)$ algorithm, which reports all the bridges in a graph. [10]

6. Suppose that you wish to route flow through a network of pipes. We model the network as a connected, undirected graph $G = (V, E)$, in which each edge has a numeric value $c(u, v)$, which represents the capacity of the edge $(u, v) \in E$, that is, the amount of flow it can take. Given any path $P = u_1, u_2, \ldots, u_k$, its capacity is defined to be the minimum capacity of any edge on the path, that is $cap(P) = \min\{c(u_1, u_2), c(u_2, u_3), \ldots, c(u_{k1}, u_k)\}$. By convention, $c(u, u) = \infty$, for all $u \in V$. For every $u, v \in V$, define $cap(u, v)$ to be the maximum capacity over all paths from u to v.

   Given a source vertex $s \in V$, present an algorithm that computes $cap(s, u)$ for all $u \in V$. Your algorithm should run in $O(|E| \log |V|)$ time. Note that, it is sufficient just to compute the capacity of the paths. It is not required to compute the actual paths. [12]

INDIAN STATISTICAL INSTITUTE
Mid-Semester Examination : (2012-2013)
M.Tech. (CS) I Year
Database Management Systems

Date : 27.02.2013 Maximum Marks : 50 Duration : 2 Hrs.

1. A school database maintains the following relations for its students, teachers and subjects:
Student (st_name, st_address, class, section, roll_no, regn_no)
Teacher (t_name, t_address, tel_no)
Subject(s_name, t_name, text_book, class, section, year)
However, the above set of relations does not meet all the requirements. Many constraints have not been represented properly and many new attributes need to be included. The requirements are listed below. They include both new attributes and constraints on existing data.

   a) A student after admission to the school is assigned with a unique regn_no. However, a student also gets a roll_no that starts from 0 for each class and section. A class can have many sections and a student is placed in only one class and section as expected in a school.
   b) In the school a teacher's name (t_name) has been found to be unique. However, more than one teacher may stay at the same address and the tel_no is a land line connection where an address will have only one such telephone.
   c) There are two types of teachers: permanent and contractual. When a permanent teacher gets a monthly salary and allowances, a contractual teacher gets a consolidated amount at the end of a month.
   d) A subject name (s_name) is unique but the same subject may be taught in many classes and sections (for example, History may be taught in many classes with different contents but s_name remains the same). Every subject has a prescribed text_book for a class. In a particular year, one teacher is assigned to each class and each section for teaching a subject.
   e) The school conducts weekly tests and an annual examination for each subject. While in a weekly test a student will either pass or fail, in the annual examination a student will get marks within 100.
   f) Weekly tests and annual examination should relate a student, a teacher and a subject.

   Considering the above requirements draw an ER/EER diagram. Using the standard mapping rules create a set of new relations or modify the existing relations to arrive at the same result. Now to meet the specified constraints, identify the functional dependencies and normalize the relations so that they are free from partial and transitive dependencies. For each relation underline the primary key. If any relation has more than one candidate key, mention them.
(10+10+15=35)

2. Using the set of relations created in Question 1, identify the teachers (if any) who can teach all subjects. Use relational Algebra to answer the query. Don't use division operator.
(15)

※

Indian Statistical Institute
Semester-II 2012-2013
M.Tech.(CS) - First Year
End-semester Examination (25 April, 2013)
Subject: Automata, Languages and Computation

Maximum marks: 100 Total marks: 110 Duration 4 hrs.

To the extent possible, please be formal in your proofs, arguments, etc.

---

1. Let G be a context-free grammar (CFG) given by the following productions (S is a non-terminal; all other symbols ($a, +, -$, etc.) are terminals).

$$S \to SS+$$
$$S \to SS*$$
$$S \to a$$

(a) Prove that G is unambiguous. (HINT: Use induction.)

(b) Construct an equivalent grammar in Chomsky Normal Form.

(c) Use the CYK algorithm to fill in the following table of V_{ij}s (symbols have their usual significance) for the string $aaa+*$.

	$i \to$	a	a	a	$+$	$*$
		1	2	3	4	5
	1					
j	2					
$\downarrow$	3					
	4					
	5					

Hence determine whether the string is in $L(G)$.

[10 + 6 + 16 = 32]

2. (a) Let $\Sigma = \{0, 1\}$. Let $\overline{w}$ denote the Boolean complement of w, i.e., it is the string obtained from w by changing all 0s to 1s and 1s to 0s. Let $L = \{w \mid w^R = \overline{w}\}$ (w^R denotes the reverse of the string w). Give a CFG for L. Your CFG should have exactly one non-terminal S. Explain your answer.

(b) Using the Pumping Lemma for context-free languages (CFLs), show that $L = \{0^i 1^j \mid j = i^2\}$ is not context-free.

(c) Show that if L is a CFL, and R is a regular set, then $L \cap R$ is context-free.

[8 + 12 + 8 = 28]

P.T.O.

3. (a) Let $\Sigma = \{a_1, a_2, \ldots, a_n\}$ be an ordered alphabet, and let $f : \Sigma^* \to \Sigma^*$ be given by $f(w) =$ the string that immediately follows w when the strings in Σ^* are enumerated in lexicographic order. Draw the diagram for a machine schema to compute f.

   (b) Recall that $L_1/L_2 = \{x \in \Sigma^* \mid \text{ for some } y \in L_2, xy \in L_1\}$. Suppose that L_1 and L_2 are accepted by Turing Machines (TMs) M_1 and M_2 respectively. Clearly describe how you would construct a TM M to accept L_1/L_2.

[10 + 14 = 24]

4. (a) Let L be a language consisting of encodings of TMs that accept at most 25042013 strings. *Without using Rice's Theorem*, prove that L is not Turing decidable.

   (b) Is the following question decidable: given a TM M, are there infinitely many TMs equivalent to M? Justify your answer.

   (c) Prove that the following instance of Post's Correspondence Problem does not have a solution:

$$w_1 = 10,\ u_1 = 101; \qquad w_2 = 011,\ u_2 = 11; \qquad w_3 = 101,\ u_3 = 011.$$

[14 + 4 + 8 = 26]

INDIAN STATISTICAL INSTITUTE

Semestral Examination

M. Tech (CS) - I Year (Semester - II)

Design and Analysis of Algorithms

Date : April 30, 2013 | Maximum Marks : 100 | Duration : 3.5 Hours

Note : You may answer any part of any question, but maximum you can score is 100.

1.(a) Let $T = t_1, t_2, \ldots, t_n$ and $P = p_1, p_2, \ldots, p_k$ be two sequences of characters, and $k \leq n$. *P is said to be a* sub-sequence *of T if there exist indices* $1 \leq i_1 \leq i_2 \leq \ldots \leq i_k \leq n$ *such that for all* $j, 1 \leq j \leq k$, *we have* $t_{i_j} = p_j$.

Design an $O(n)$ time algorithm to determine whether P is a sub-sequence of T according to the above definition of a *sub-sequence.*

(b) Describe Huffmans coding of information. When is such a variable length coding scheme used ? What is an optimal Huffman code for the following set of frequencies?

a	b	c	d	e	f	g	h
1	1	2	3	5	8	13	21

[8+(3+3+7)=21]

2.(a) Consider two 2×2 matrices $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ and $B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$. The elements of the matrices A and B are real numbers. C is a 2×2 matrix obtained by multiplying A and B. Show that computation of C can be done using at most 7 multiplications. You may use as many additions/subtractions as required.

(b) Suppose that you want to multiply two polynomials $1 + x$ and $2 + 3x^2$ using FFT. Choose an appropriate power of 2, find the FFT of the two sequences, multiply the results componentwise, compute the inverse FFT, and output the final result.

[10+15=25]

3.(a) Decide whether the following statement is *true* or *false*. If it is true, give proper justification, and if it is false, give a counterexample.

> Let $G = (V, E)$ be an arbitrary flow network, with a source s and a sink t. Each edge $e \in E$ is attached with a positive integer capacity c_e. Let (A, B) be the minimum $s - t$ cut in graph G with respect to the edge costs $\{c_e, e \in E\}$. Now, if we add 1 to the capacity of every edge, then (A, B) still remains the minimum $s - t$ cut in the revised flow network.

(b) The edge contraction method for computing the mincut of a weighted undirected graph $G = (V, E)$ is as follows.

Let $(x, y) \in E$ be contracted to get the graph $G \setminus xy$. The vertices of the resulting graph are $V \setminus \{x, y\} \cup xy$, where xy is a single vertex. All the edges incident to x and y in G are incident to the vertex xy in the graph $G \setminus xy$. More than one edge incident to a vertex is replaced by a single edge with sum of weights of the original edges as the weight of the new edge. Self loop is not removed.

Write an algorithm for finding a cut of the graph G using the aforesaid contraction procedure.

Show that if G is unweighted and in each stage (when needed in your algorithm) the edge to be contracted is chosen randomly, then you can find a mincut with constant probability (You may choose your own constant value < 0.5, but have to justify your answer). State the time complexity of your algorithm.

[10+8+12=30]

4.(a) Let P_1 and P_2 be two problems such that P_1 is polynomial time reducible to P_2, the time complexity of this reduction is $O(n^2)$, and the time complexity of solving P_2 is $O(n^4)$. What can you say about the time complexity of the problem P_1?

(b) A **k-SAT** problem is defined as follows: *Given a CNF expression with n boolean variables such that each of its clauses contains at most k variables (in either complemented or uncomplemented form), does there exist any assignment of values to these n boolean variables such that the value of the expression is 1?*

Show that **2-SAT** decision problem can be solved in polynomial time.

(c) The 3D-matching problem is defined as follows:

Given a tripartite graph $G = (V, E)$, where $V = V_1 \cup V_2 \cup V_3$ and $E = E_{12} \cup E_{13} \cup E_{23}$; for each edge $(u, v) \in E_{ij}$ $u \in V_i$ and $v \in V_j$, $1 \leq i < j \leq 3$. Now given an integer k it asks whether there exist k disjoint cycles of size 3 in the graph G.

Assuming that **3-SAT** with each variable occuring in at most 3 clauses in a CNF expression is NP-complete, show that the 3D-matching problem is also NP-comlete,

[6+8+8=22]

5. Consider the optimization version of the (0-1)-knapsack problem as follows:

KNAPSACK: Given two sequences of integers $(w_1, w_2, \ldots, w_n)$ and $(c_1, c_2, \ldots, c_n)$ and an integer K, where w_i and c_i are volume and price respectively of item i, $i = 1, 2, \ldots, n$, and K is the size of the knapsack; the objective is to choose items from $\{1, 2, \ldots, n\}$ such that their total volume is less than or equal to K and the total price is maximum.

Part 1: Write an efficient pseudo-polynomial time algorithm for the *KNAPSACK* problem (time complexity may be polynomial in n and K).

Part 2: Let *KNAPSACK'* be another (0-1)-knapsack problem with parameters $(w_1, w_2, \ldots, w_n)$, $(c'_1, c'_2, \ldots, c'_n)$ and K, where $c'_i = 10 \times \lceil \frac{c_i}{10} \rceil$. Let OPT and OPT' be optimal solutions of *KNAPSACK* and *KNAPSACK'* respectively. Show that $|OPT - OPT'| \leq 10n$.

Part 3: Use this fact to suggest a PTAS for the (0-1)-knapsack problem.

[10+10+10=30]

INDIAN STATISTICAL INSTITUTE

Second Semester Examination : (2012-2013)

M.Tech.(CS) I Year

Database Management Systems

Date : 03.05.2013 **Maximum Marks : 50** **Duration : 2.5 Hours**

. Answer all questions.

1 The concurrent schedule shown below involves three transactions T1 to T3 using two data items A and B for read and write.

Time Slot	T1	T2	T3
1	read(A)		
2		read(B)	
3		write(A)	
4			write(B)
5			read(A)
6	read(B)		
7			write(A)
8	write(B)		

a) Drawing a precedence graph show whether the above schedule is conflict serializable.

b) To examine view serializability, draw the labeled precedence graph for two data items separately to show whether they are individually view serializable.
Also draw the composite labeled precedence graph to examine whether the schedule is view serializable considering the two data items together.

(4x4=16)

2 Let us consider that timestamp based protocol is used for the execution of the concurrent schedule given in Question 1. According to the schedule, each read or write operation needs one time slot. In case of any time conflict, a transaction is allowed to rollback and restart with a new timestamp higher than all the existing timestamps. No such conflict is ignored, i.e. Thomas' Write Rule is not considered. It is assumed that a rolled back transaction is rescheduled immediately. The unused time slots obtained due to the rollback of a transaction may be utilized by a rescheduled transaction. A transaction may have any number of rollback and restart. Find the total number of time slots required to execute all the three transactions without changing the given schedule. The timestamps of the three transactions are related as,

$$TS(T1) < TS(T2) < TS(T3)$$

For each instruction executed, show the status of the read and write timestamps of the data item involved.

(10)

3 Considering Wound-Wait scheme as deadlock detection mechanism, evaluate the final execution order of the three transactions if they follow the schedule given in Question 1. Find also the total number of rollbacks needed before all the transactions complete their executions. The three transactions follow the timestamp order given in Question 2. Consider that between two transactions, one having lower value of timestamp has a higher priority.

(10)

4 Considering the concurrent schedule given in Question 1, examine whether the schedule is executable under two-phase locking protocol with upgrade facility.

(7)

5 Two relations **R** and **S** are to be joined against a common attribute **a** where **a** is the primary key of **S**. **R** has 10000 tuples and **S** has 2000 tuples. If for both the relations 100 tuples form a page, what would be the estimated number of disk accesses for reading the relations. Consider that the smaller relation can be totally accommodated in the main memory and block-oriented nested loop join is used as the join algorithm. If each tuple in relation **R** and **S** are of length 100 bytes and 200 bytes respectively and common attribute **a** is 30 bytes long, find the estimated size of the joined relation in bytes.

(7)

-x-