Indian Statistical Institute

M.Tech (CS) I

Discrete Mathematics

Mid Semester Examination

Maximum Marks: 70

Date: September 8, 2014,
Time 2.5 hours

The question paper contains 7 questions. Total marks is 70. Maximum you can score is 60 Unless otherwise mentioned, all notations are the same as presented in class.

1. Without using recurrence relations and generating functions, show that the n-th Catalan number is given by $C_n = \frac{1}{n+1}\binom{2n}{n}$. (10)

2. Solve the following recurrence relation

$$x_n = 3x_{n-1} + 10x_{n-2} + 7.5^n,$$

where $x_0 = 4$ and $x_1 = 3$. (10)

3. Every positive integer n can be written uniquely as a product of powers of primes.

$$n = p_1^{e_1} p_2^{e_2} \cdots p_r^{e_r},$$

where $p_1, p_2, \cdots, p_r$ are distinct primes and $e_i \geq 1$, for all i. The *Mocbius function* $\mu(n)$, is defined by,

$$\mu(n) = \begin{cases} 1 & \text{if } n = 1 \\ 0 & \text{if } e_i > 1 \text{ any } i \\ (-1)^r & \text{if } e_1, e_2 \cdots e_r \text{ all equal 1} \end{cases}$$

(a) Prove that

$$\sum_{d|n} \mu(d) = \begin{cases} 1 & \text{if } n = 1 \\ 0 & \text{if } n > 1 \end{cases}$$

(b) Show that if $\phi(n)$ is the Euler function, then

$$n = \sum_{d|n} \phi(d).$$

(c) Prove that

$$\phi(n) = \sum_{d|n} \mu(d)\frac{n}{d}.$$

(4+3+3 = 10)

4. Prove that a simple graph with n vertices and k components can have atmost $(n-k)(n-k+1)/2$ edges. 10

5. Write an algorithm to test if a graph is bipartite? Prove the correctness of the algorithm and calculate the time complexity. 4 + 3 + 3 = 10

6. Prove that if G has more than $n^2/4$ edges, then G has a vertex whose deletion leaves a graph with more than $(n-1)^2/4$ edges. Use (a) to prove that G contains a triangle if $e(G) > n^2/4$. (5+5 = 10)

7. Prove that the following is a necessary condition for the existance of k pairwise edge-disjoint spanning trees in a graph G: For any partition of the vertices of G into r parts, there are at least $k(r-1)$ edges of G whose endpoints are in different parts of the partition. (10)

INDIAN STATISTICAL INSTITUTE

PERIODICAL EXAMINATION
M.TECH.(CS) I YEAR

ELEMENTS OF ALGEBRAIC STRUCTURES

Date: 09.09.2014 Maximum marks: 60 Duration: 3 hours

The paper contains questions of 82 marks. Answer as much as you can, the maximum you can score is 60. If you use any result, not proven in the class, you must prove it.

Notation: *fg, the product of two functions f and g is the result of first applying g followed by f.*

1. (a) Let p be an integer greater than one. Prove that p is a prime if and only if $(p-1)! \equiv -1 \bmod p$. [8]
   (b) Prove that any subgroup of a cyclic group is cyclic. [8]
   (c) Let G be a group. Show that the map $\varphi : G \to G$ defined by $\varphi(a) = a^{-1}$ is a group homomorphism if and only if G is abelian. [8]

2. **Definition 1** *Let G be a group. We say G is simple if it contains no proper normal subgroup.*

   Definition 2 *A group of order a power of prime p is called p-**group**. Let G be a finite group of order $n = p^k m$ where p is a prime and p does not divide m. A subgroup of order p^k is called **Sylow p-subgroup** of G.*

   Theorem 3 *Let G be a finite group of order $n = p^k m$ where p is a prime and p does not divide m.*

   (a) The number of Sylow p-subgroup is congruent to 1 modulo p, and divides n.

   (b) Any two Sylow p-subgroups are conjugate.

   Lemma 4 *Every simple p-group has prime order.*

   Prove the following propositions.

   Proposition 5 *Let G be a finite group and and let p be a prime dividing the order of G.*

   (a) G has at least one Sylow p-subgroup P. [3]

   (b) If P is the only Sylow p-subgroup, then P is normal in G. [7]

   Proposition 6 *Let G be a group of order pq where both p and q are primes and $p < q$. Then G has a normal subgroup of order q. [8]*

Proposition 7 *There is no simple group of order 12. [6]*

Proposition 8 *If G is a simple group of order less than 16, then G is of prime order. [4]*

3. (a) Prove that every permutation is a product of transpositions. [8]

   (b) Let $(x_1\ x_2\ \cdots\ x_k)$ be a k-cycle. Prove the following

   i. $(x_1\ x_2\ x_3\ \cdots\ x_k)(x_1\ x_2) = (x_1\ x_3\ \cdots\ x_k)$. [4]

   ii. For a y different from any of $x_1, x_2, \cdots, x_k$,

   $$(x_1\ x_2\ \cdots\ x_k)(x_1\ y) = (x_1\ y\ x_2\ \cdots\ x_k).[4]$$

   iii. $(x_1\ x_2\ x_3\ x_4)(x_2\ x_4) = (x_1\ x_2)(x_3\ x_4)$. [4]

   (c) As a part of a treasure hunt, a group of n people land up in an unknown island. They discover a mind switching device which allows two users to switch minds. Unfortunately, after a few mind switching took place, they realize it is a trap. A mind swap is not repeatable; no pair of bodies can swap their minds more than once. Luckily they found two mathematicians (with their mind in their respective bodies) who were ready to co-operate. They came up with a series of mind switches, without knowing the earlier operations, which return the minds to the original bodies! Can you find the method? (Hint: Use the two extra bodies. Prove that any permutation can be converted to identity permutation with a sequence of products of transpositions each containing only one of the extra bodies.)[10]

INDIAN STATISTICAL INSTITUTE
Mid-Semestral Examination, Semester I, 2014-15
M.Tech. (Computer Science)
Computer Organization

Full Marks : 100
Time : 3 Hours

Date : 10. 09. 2014

Note : *Answer all questions. Marks on each part of a question are indicated in the right margin within parentheses.*

1. a) Represent each of the three decimal integers *16, -1* and *-32* in *8*-bit binary, using i) sign-magnitude, ii) *1's* complement and iii) *2's* complement modes of representation. **(9)**

   b) How are the decimal integers *6* and *7* represented in *4*-bit Gray code and what will be the Hamming distance between these two codes? Represent these two numbers also in two-out-of-five code and calculate again the Hamming distance between these two codes. **(6)**

   c) While adding two binary numbers represented in *2's* complement form, how can we determine the conditions for overflow and underflow? Justify your answer with appropriate reasoning. **(4)**

   d) Represent the following numbers in *IEEE 754-1985* format : **(3)**

   i) 2.75 *ii)* 0.375×2^{-130} *iii)* $-\infty$

2. Consider the following truth table of a combinational circuit where *A, B* and *C* are its input Boolean variables with *f* as the output. Represent *f* in the sum of product form and also in the product of sum form. Prove that these two forms are equivalent.

A	*B*	*C*	*f*
0	*0*	*0*	*1*
0	*0*	*1*	*0*
0	*1*	*0*	*1*
0	*1*	*1*	*0*
1	*0*	*0*	*1*
1	*0*	*1*	*0*
1	*1*	*0*	*1*
1	*1*	*1*	*1*

Show how *f* can be implemented using only three *NAND* gates. **(4 + 4 + 3 = 11)**

3. a) Show, with appropriate logic, how an *EX-OR* operation between two Boolean variables *A* and *B* can be implemented using only *4 NAND* gates, assuming that the complements of the variables *A* and *B* are not available. **(5)**

b) Minimize the Boolean function $g = \sum(0,1, 2, 6, 7, 12, 14, 15)$. **(10)**

c) Given the function *f* corresponding to the truth-table of *Q. 2* above, show how *f* can be implemented using only *2* to *1* multiplexers. Justify your implementation scheme with proper logic. **(8)**

4. Derive a minimal form of a sequential machine with a single input line x and having the following state transition table with seven states *A, B, C, D, E, F* and *G* : **(12)**

Present State	*Next State and Output*	
	x = 0	*x = 1*
A	*E,0*	*C,0*
B	*C,0*	*A,0*
C	*B,0*	*G,0*
D	*G,0*	*A,0*
E	*F,1*	*B,0*
F	*E,0*	*D,0*
G	*D,0*	*G,0*

5. a) Draw the circuit diagram of a Master-Slave *J-K* flip-flop using *NAND* gates. **(8)**

b) Show how a synchronous modulo *5* counter can be designed using *J-K* flip-flops. **(8)**

6. a) Assume that you have to add two *4*-bit binary numbers by means of a parallel binary full adder circuit with carry look-ahead mechanism. Write down the logic functions to generate the carry at all the bit positions for this parallel adder circuit.

b) Consider a parallel binary adder which first computes the precarry vector to generate the carry at different bit positions. Assuming that the two binary numbers to be added are *10110101* and *11001010* respectively, compute the precarry vector and also the resulting carry at the different bit positions. **(8 + 8 = 16)**

INDIAN STATISTICAL INSTITUTE

Mid-Semestral Examination

M.Tech(CS)-I Year, 2014-2015 (Semester-I)

Data and File structures

Date: **September 12, 2014** Maximum Marks: **75** Duration: $2\frac{1}{2}$ **Hours**

Note: The question paper carries a total of 85 marks. You can answer as much as you can, but the maximum you can score is 75.

---

1. (a) Let D be a data structure that internally has two stacks (A and B) and supports only the following operations:
Insert (x): push the key x into stack A.
Remove (x): pop stack B and output the element.
Shift: If stack A is nonempty, pop A and push the output to B.

   (i) Show that all stack-permutable permutations can be generated by D.

   (ii) Is there any non-stack-permutable permutation that can be generated by D? Justify.

   (b) Suppose you are given k sorted lists, each containing m integers in ascending order. Assume that the lists are stored as singly-linked lists with one integer in each node and the head pointers of these lists are stored in an array A. Write an efficient algorithm that merges the above k sorted lists into a single sorted list using additional $\Theta(1)$ additional storage. Analyze the time complexity of your proposed algorithm.

((6+4)+(8+2)=20)

2. (a) Describe an implementation of a data-structure which supports the following operations:
insert(x): insert the key x into the data structure only if it is not already there.
Delete(x): delete the key x from the data structure if it is there.
$Find - Smallest(k)$: find the kth smallest key in the data structure.
All these operations should take $O(\log n)$ time in the worst case, where n is the number of elements in the data structure.

   (b) An $n \times n$ matrix A is called a $NICE$ matrix, if each row and each column of A has exactly one non-zero element equal to 1.

   (i) Suggest a method of storing a $NICE$ matrix in an $O(n)$ size array.

   (ii) Design an $O(n)$ time algorithm for computing $R = P \times Q$, where P and Q are both $NICE$ matrices stored as above. Justify the time complexity of your algorithm.

(9+(4+10)=23)

3. (a) Explain how you may implement a priority queue in an array such that insertion and deletion time of an element is $O(\log n)$ in the worst case.

   (b) Suppose H is a max-heap (implemented in an array) of n real numbers. Write an algorithm to find all elements in H, which are larger than a given value x. The algorithm should run in $O(k)$ time, where k is the number of elements greater than x in H.

(5+(10+2)=17)

4. (a) Suppose T_1 and T_2 are two arbitrary binary trees, each having n nodes. How many rotations are sufficient to convert T_1 into T_2? Justify your answer.

   (b) Show that in an AVL tree of height h, the depth of every leaf node is at least $\lceil h/2 \rceil$.

(c) Let T be an binary search tree whose nodes are of the following structure.

```
typedef struct node *nodeptr
typedef struct{
int Info;
int Size;
nodeptr left;
nodeptr right;
} node;
```

The *Info* field of the nodes of T contain distinct key values and *Size* field of the nodes contain the number of elements in the subtree rooted at the corresponding node. Write an algorithm to find the number of elements in T whose key value x satisfies $x_1 < x \leq x_2$ for two given key values x_1 and x_2 with $x_1 \leq x_2$. The algorithm should run in $O(h)$ time, where h is the height of the tree.

(7+7+(9+2)=25)

INDIAN STATISTICAL INSTITUTE

PERIODICAL EXAMINATION
M.TECH.(CS) I YEAR

PROBABILITY AND STOCHASTIC PROCESSES

Date: 15.09.2014 Maximum marks: 50 Duration: 2 hours

Each question carries 8 marks. Answer as much as you can, the maximum you can score is 50.

1. Let A, B and C be three events. Express $A \cup B \cup C$ as the union of some mutually exclusive events.

2. A man is given n keys of which only one fits his door. He tries them successively (sampling without replacement). This procedure may require $1, 2, \ldots, n$ trials. Show that each of these n outcomes has probability $1/n$.

3. Define σ-algebra and probability as a measure. Describe the Lebesgue measure on $[0, 1]$.

4. Let A be an event and $B_1, \ldots, B_n$ be a partition of the event space. Show

$$\max_{1 \le i \le n} \mathsf{Pr}[A \wedge B_i] \le \mathsf{Pr}[A] \le \max_{1 \le i \le n} \mathsf{Pr}[A \mid B_i].$$

5. Consider the set of all outcomes of n tosses of a fair coin. Let A be the event "there is at most one head" and B be the event "both head and tail occur". For what values of n are A and B independent?

6. A court is investigating the actual occurrence of an event T which occurs with probability 0.001. There are two witnesses A and B both of whom tell the truth with probability 0.9 What is the probability that T has occurred when both A and B testify to its occurrence.

7. Consider a succession of $\mathsf{Ber}(p)$ trials. For any $r > 0$, show that with probability 1, r successes will occur in a finite number of trials.

8. Let $A_1, A_2, \ldots$ be a sequence of events. Define $\limsup_{n \to \infty} A_n$. If $A_1, A_2, \ldots$ are mutually independent and $\sum_n \mathsf{Pr}[A_n] \to \infty$ show that

$$\mathsf{Pr}\left[\limsup_{n \to \infty} A_n\right] = 1.$$

INDIAN STATISTICAL INSTITUTE

Mid-Semestral Examination : 2014-15

Course : Master of Technology in Computer Science (First Year)

Subject : Introduction to Programming (A1) for M.Tech. (CS) I

Date : 15 September 2014 Maximum Marks : 60 Duration : 3 Hours

Note: You are expected to code in C to solve the following problem(s), and you are allowed to use any academic resource of your choice during the course of this examination. Write the theoretical part(s) of the solution(s) in the answer-sheet provided, and submit the relevant C source file(s) separately. In the theoretical part(s) of the solution(s), you must provide basic sketch of the algorithm(s) that you use, along with basic complexity analysis of your chosen algorithm(s), wherever appropriate. You must justify the C data-type(s) that you use while coding, and provide basic pictorial illustration(s) of the memory stack-frame(s) for clarification, wherever appropriate.

Problem 1

Maximum Marks: 40 = 10 + 10 + 15 + 5

Write by yourself all required string functions, without using the standard `string.h` library of C.

1A. Accept an English word from the user as a *command-line argument*. If the word is k-letters long, print the following – the original k-letter word, last $(k-1)$ letters of the word, last $(k-2)$ letters of the word, and so on, till the last two letters of the word, and the last letter of the word.

Example Input: `sourav`

Example Output: `sourav, ourav, urav, rav, av, v`

1B. In continuation to the previous problem, write a function `wordRev` to reverse the original word and store the output as a new word. Write another function `wordRevInp` to reverse this new word in place, and overwrite itself. Print the following – the original word, the new word formed by applying `wordRev` to the original word, and the overwritten new word after applying `wordRevInp`.

Example Output: `sourav, varuos, sourav`

1C. In continuation to the previous problem, accept a positive integer r from the user as a *run-time input*. Write a function `wordRot` to right-rotate the original input word by r letters, and store the output as a new word. Write another function `wordRotInp` to right-rotate this new word by r letters, in place, and overwrite itself. Print the following – the original word, the new word formed by applying `wordRot` to the original word, and the overwritten new word after applying `wordRotInp`.

Example Input: 2

Example Output: `sourav, avsour, uravso`

1D. In continuation to the previous problem, print all possible right-rotations of the original word.

Example Output: `sourav, vsoura, avsour, ravsou, uravso, ouravs`

Problem 2

Maximum Marks: 20 – 10 + 10

An *anagram* is a type of word play, where the result of rearranging the letters of a word, using all the original letters exactly once, produce a new word that belongs to some predefined dictionary.

2A. Accept two lowercase English words from the user as *command-line arguments*. Check if one word is an *anagram* of the other, and print `Anagrams` or `NOT Anagrams` accordingly.

Example Input: `stop post`

Example Output: `Anagrams`

Example Input: `stop sign`

Example Output: `NOT Anagrams`

2B. Accept an arbitrary list of lowercase English words from the user as *command-line arguments*, and store them as the predefined *dictionary*. Accept another lowercase English word from the user as a *run-time input*, and determine if any *anagram* of this word exists in the aforesaid dictionary. If the answer is affirmative, print all anagrams of the word from the dictionary.

Example Input: `{stop block sign post club fly run dance}` and `tops`

Example Output: `Anagrams found : stop post`

Example Input: `{stop block sign post club fly run dance}` and `sourav`

Example Output: `Anagrams found : None!`

Solving a problem correctly is necessary, but not sufficient, as it does not guarantee the maximum marks alloted for the problem. Sufficient credit is reserved in each case for *smart* algorithm and *good* coding practice. Good luck! ☺

INDIAN STATISTICAL INSTITUTE

SEMESTRAL EXAMINATION

M. Stat (1st Year) – B Stream, 2014 – 2015

Subject: Applied Stochastic processes

Date: 14/11/2014 **Full Marks: 60**

Time: 3 hours.

Attempt all questions

1. State and prove 1st Hardy – Weinberg law in connection to ~~unisex~~ the relevant population with two alleles A, a. In this context also show that proportions obtained in generations by random mating and by random pairing of gametes from the gametic pool are same.

[(2 + 6) + 4]

2. Write down Kermack-Mckendrick Equations. Describe how those equations are derived from the model of General Deterministic Epidemics. Using those equations show that limiting number of susceptible with time $t \to \infty$ is positive.

[4 + 6 + 6]

3. a) Consider Neyman – Scott model in the context of spread of Epidemic in R^2. Show that if there are no deserts and there is full mobility then the epidemic goes for extinction at u_0 implies that it goes for extinction at all other points, for all $u_0 \in R^2$.

b) Let the mobility function be given by

$$f_u(x) = 1/\pi \text{ if } d(u, x) < 1$$
$$= 0 \text{ otherwise,}$$

where $d(u, x)$ is distance between u and $x \in R^2$. Probability generating function of the number of infected persons from one infectious at $u = (u_1, u_2) \in R^2$ given by

$$g(t/u) = 1/3 + 1/3t + 1/3\,t^2 \quad \text{if } u_1 u_2 = 0$$
$$= 1/3t + 1/3t^2 + 1/3t^3 \quad \text{if } u_1 u_2 \neq 0$$

Discuss the probabilities of the process to become extinct at different points.

[12 + 6]

4. a) Let $\{X_n\}_{n \geq 0}$ be a branching process with $X_0 = 1$, $E(X_1) = m > 1$, $Var(X_1) = \sigma^2 > 0$. Let $W_n = X_n / m^n$. Show that there is a random variable W such that $W_n \to W$ almost surely.

b) Consider a branching chain with initial size N and probability generating function $g(s) = q + ps$, where $0 < p < 1$, $q = 1 - p$. If T = time when the process becomes extinct, show that $P(T = n) = (1 - p^n)^N - (1 - p^{n-1})^N$.

[10 + 4]

----------o----------

INDIAN STATISTICAL INSTITUTE

Back Paper Examination

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

Design and Analysis of Algorithms

17.11.2014

Date : ~~**.5.2014~~ Maximum Marks : 100 Duration : 3.5 Hours

1. Consider an array S of size n. Each element of the array is a member in the set $X = \{x_1, x_2, \ldots, x_k\}$. Here, $k << n$ (k is much less than n) and the elements in X are known to be integers, but their values are not known apriori. Write an efficient ($O(n \log k)$ time if possible) algorithm for sorting the array S. Justify the time and space complexities of your proposed algorithm. Is your proposed algorithm in-place? If the elements are known apriori, then can you reduce the time and/or space complexities of your algorithm? [10]

2. Write a method to check whether a given string of characters $T = a_1 a_2 \ldots a_n$ is the circular rotation of another string of characters $T' = b_1 b_2 \ldots b_n$. For example, *arc* and *car* are cyclic rotation of each other. Analyze the time complexity of your proposed method. [10]

3. Let $P = \{p_1, p_2, \ldots, p_n\}$ be a set of n points distributed on a 2D plane. The objective is to test whether the convex hull of P is a quadrilateral or not. Write a linear time implementable algorithm for solving this problem. Justify the time complexity of your algorithm. [10]

4. (a) When is a graph said to be bi-connected ? Prove/disprove the following statement:

   An undirected graph is bi-connected if for every pair of vertices v and w there exists two vertex disjoint paths.

   (b) Write an algorithm to identify the bi-connected components in a graph. Mention the time complexity of your algorithm.

   [(3+7) + 10 = 20]

5. Describe Ford-Fulkerson's algorithm for computing a maximum flow from the source vertex s to the sink vertex t in a directed graph where each edge has integer capacity.

   Show that Ford-Fulkerson's algorithm may take exponential time to terminate depending on the nature of its edge capacities.

   Show that if Ford-Fulkerson's algorithm terminates, it produces the maximum flow value from s to t. [10+7+8=25]

6. (a) When is a problem said to be (i) in NP, and (ii) NP-complete?

   (b) Show that the problem of computing the vertex-cover of a graph is NP-complete.

(c) Write a polynomial time algorithm for the vertex-cover problem which produces a solution of size at most $2 * opt$, where opt is the size of the optimum solution.

[5+10+10=25]

INDIAN STATISTICAL INSTITUTE

First Semestral Examination : 2014-15

Course : Master of Technology in Computer Science (First Year)

Subject : Introduction to Programming (A1) for M.Tech. (CS) I

Date : 19 November 2014 | Maximum Marks : 100 | Duration : 3 Hours

Note: You are expected to code in C to solve the following problem(s), and you are allowed to use any academic resource of your choice (except the Internet) during the course of this examination. Write the theoretical part(s) of the solution(s) in the answer-sheet provided, and submit the relevant C source file(s) separately. In the theoretical part(s) of the solution(s), you must provide basic sketch of the algorithm(s) that you use, along with basic complexity analysis of your chosen algorithm(s), wherever appropriate. You must justify the C data-type(s) that you use while coding, and provide basic pictorial illustration(s) of the memory stack-frame(s) for clarification, wherever appropriate.

Problem 1 : Short Questions

Maximum Marks: 50 = 5 + 5 + 5 + 10 + 10 + 15

1A. Is it possible to determine the byte-size of a structure in C without using the inbuilt `sizeof()` function? If so, write a sample C code snippet to demonstrate your strategy.

1B. Devise a one-pass algorithm to find/print the data located exactly at the middle of a singly linked list containing odd number of nodes. Write the pseudocode for your algorithm.

1C. Hamming weight of a binary string refers to the number of 1's in the string. Write a C program that takes a binary string as input from the user, and outputs the Hamming weight of the same.

1D. Palindromes are words or phrases (or any sequence of symbols) that read the same forward or reversed, ignoring the punctuations and blank spaces. Write a C program that takes an English string (may be a few words long) as input from the user, and checks if it is a palindrome.

1E. Write a C program that takes an English word (size within 10 characters) as input from the user, and writes all possible permutations of the same in a file. Make sure that the permutations are generated and written in a lexicographic (alphabetic) order.

1F. Write a C program that imitates the standard UNIX function `cat`. The program should take as command-line input one or more filenames, and print on the standard output all the files in a concatenated format. Make sure that the program exhibits plausible errors in a proper format.

Problem 2 : Integer Polynomials

Maximum Marks: 25

Polynomial arithmetic is analogous to Integer arithmetic, especially in terms of addition and multiplication. If we define the notion of size for polynomials as their degree, with constant polynomials defined to have degree zero, we get a natural notion for the division algorithm with polynomials $a(x)$ and $b(x)$, where $\deg(a) \geq \deg(b)$, as follows.

$$a(x) \quad q(x) \cdot b(x) + r(x) \qquad \text{with quotient } q(x) \text{ and remainder } r(x) \text{ with } 0 \leq \deg(r) < \deg(b).$$

This automatically lets us define divisibility of polynomials as '$b(x)$ divides $a(x)$' if and only if $r(x) = 0$, and therefore, extends the notion of GCD naturally to the polynomials. The GCD of two polynomials $a(x)$ and $b(x)$ is a polynomial $d(x)$, of the highest possible degree, that divides both $a(x)$ and $b(x)$.

Write a C program to represent polynomials with integer coefficients in a format suitable for basic arithmetic operations like addition and multiplication. Use the basic operations to implement a quotient-remainder division routine for integer polynomials, and then use this division to implement a complete GCD routine for integer polynomials. At the end of the day, your program should take two polynomials as input from the user, and output their GCD.

Example 1 : Input $x^4 + 2x^3 + 2x^2 + 2x + 1$ and $x^3 + 2x^2 + 2x + 1$. Output $x + 1$.
Example 2 : Input $x^4 + x^2 + x + 1$ and $x^3 + 2x^2 + 2x + 1$. Output 1.
Example 3 : Input $x^2 + 2x + 1$ and $x + 1$. Output $x + 1$.

Problem 3 : Regular Expressions

Maximum Marks: 25

Regular expressions are sequences of characters forming a search pattern, mainly for use in pattern matching with strings. We will consider a small class of regular expressions, consisting of the following search patterns and identifiers.

- `x` : Matches the exact character `x` in input string
- `.` : Matches any single character in input string
- `*` : Matches zero or more occurrences of previous character
- `^` : Matches only at the start of input string
- `$` : Matches only at the end of input string

Write a C program that matches regular expressions covering the above-mentioned classes of search patterns. Your program should accept as input a regular expression conforming to the classes defined above, and an input string. It should output 1 if the regular expression matches the input string (all lower case letters), and 0 otherwise.

Example 1 : Input `^so` and `sourav`. Output 1.
Example 2 : Input `^ra` and `sourav`. Output 0.
Example 3 : Input `^.ra` and `sourav`. Output 0.
Example 4 : Input `^.*ra` and `sourav`. Output 1.
Example 5 : Input `ra$` and `sourav`. Output 0.
Example 6 : Input `rav$` and `sourav`. Output 1.
Example 7 : Input `ra.$` and `sourav`. Output 1.

Solving a problem correctly is necessary, but not sufficient, as it does not guarantee the maximum marks alloted for the problem. Sufficient credit is reserved in each case for *smart* algorithm and *good* coding practice. Good luck! ☺

INDIAN STATISTICAL INSTITUTE

SEMESTRAL EXAMINATION
M.TECH.(CS) I YEAR

ELEMENTS OF ALGEBRAIC STRUCTURES

Date: 21.11.2014 Maximum marks: 100 Duration: 3 hours

The paper contains 110 marks. Answer as much as you can, the maximum you can score is 100.

1. (a) An element a of a ring R is called nilpotent if $a^n = 0$ for some positive integer n. Prove that the set of nilpotent elements in a commutative ring R forms a subring of R. [8]
   (b) Suppose R is a commutative ring with unity and that $x \in R$ is nilpotent. Show that $1 + x$ is a unit. [6]
   (c) Deduce that the sum of a unit and a nilpotent element is a unit. [4]
   (d) Let R be a commutative ring with unity and let $f := a_0 + a_1 x + \cdots + a_n x^n$ be an element of the ring $R[x]$. Prove that f is nilpotent if and only if $a_0, a_1, \cdots, a_n$ are nilpotent. [8]
2. (a) Prove that a homomorphism from a field F onto a ring R with more that one element must be an isomorphism. [8]
   (b) Let $\phi : R \to S$ be a ring homomorphism such that the image of ϕ is not $\{0\}$. If R has a unity 1 and S is an integral domain, show that ϕ maps the unity of R to the unity of S. [8]
   (c) Let

$$R = \left\{ \begin{pmatrix} m & 2n \\ n & m \end{pmatrix} | m, n \in \mathbb{Z} \right\}.$$

   and

$$S = \{m + n\sqrt{2} | m, n \in \mathbb{Z}\}$$

   be rings with respect to their usual operations. Prove that R and S are isomorphic rings. [8]
   (d) Let $\phi : R \to R'$ be a ring homomorphism. Prove that if $x \in R$ is nilpotent, then $\phi(x)$ is nilpotent in R'.[4]
3. (a) Let $V = \{(x, y, z) | x, y, z \in \mathbb{R}, x^2 - y^2 - z^2 = 0\}$. Is V a vector space over $\mathbb{R}$? justify your answer. [8]
   (b) Consider the vector space $\mathbb{R}^\kappa$ over $\mathbb{R}$. Let $\{e_i | i = 1, \cdots, n\}$ denote the standard basis. Prove that the set $\{e_1, e_2 - e_1, e_3 - e_2, \cdots, e_n - e_{n-1}\}$ form a basis too. [8]

(c) Let $\mathbf{V}$ be an n-dimesional vector space over F. Prove that any spanning set of $\mathbf{V}$ contains at least n elements. [8]

4. (a) Let $\mathbf{V}, \mathbf{W}$ and $\mathbf{Z}$ be finite dimesional vector spaces over a field F. If $T : \mathbf{V} \to \mathbf{W}$ and $U : \mathbf{W} \to \mathbf{Z}$ be linear transformations. Prove that,

$$rank(UT) \leq min\{rank(U), rank(T)\}.$$

[10]

(b) Let $\mathbf{V}$ be an n-dimensional vector space over F. Let $\mathbf{V}^*$ be the dual space. Find a basis of $\mathbf{V}^*$ over F. Justify your answer. [8]

5. (a) Let $F = \mathbf{F}_3[x]/(x^2 + 1)$. Find a generator of F^*. [6]

(b) Construct a finite field with $2^3 = 8$ elements. [8]

INDIAN STATISTICAL INSTITUTE

SEMESTRAL EXAMINATION
M.TECH.(CS) I YEAR

PROBABILITY AND STOCHASTIC PROCESSES

Date: 24.11.2014 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) Let $X_1, \ldots, X_n$ be independent Poisson trials with $\Pr[X_i = 1] = p_i$, $0 < p_i < 1$ for $i = 1, \ldots, n$; $X = X_1 + \cdots + X_n$ and $\mu = p_1 + \cdots + p_n$. Show that for any $\delta > 0$,

$$\Pr[X > (1+\delta)\mu] < \left(\frac{e^\delta}{(1+\delta)^{(1+\delta)}}\right)^\mu.$$

(b) In a sequence of independent throws of a symmetric die, let X_k be the number scored in the kth trial. Let $S_{200} = X_1 + \cdots + X_{200}$. Find a pair of real numbers δ_1 and δ_2 with $\delta_1 < \delta_2$ such that

$$\mathsf{P}\left[S_{200} \in [\delta_1, \delta_2]\right] \approx \mathsf{P}[S_{200} \notin [\delta_1, \delta_2]]\,.$$

(Assume that $\Phi(\alpha) - \Phi(-\alpha) = 0.5$ for $\alpha = 0.6744....$)

(c) A total of n persons, all of different heights are standing in a queue facing the ticket window in a cinema hall. What is the expected number of persons visible to the person at the window? (Assume that each permutation of n persons is equally likely.)

(d) If a random variable X takes only non-negative integer values show that

$$E[X] = \sum_{n \geq 0} \Pr[X > n].$$

(e) An urn contains b blue and r red balls. Balls are removed at random one by one until the first blue ball is drawn. Show that the expected number of balls drawn up to that stage is $(b + r + 1)/(b + 1)$.

$(8 + 7 + 8 + 6 + 6 = 35)$

2. (a) Describe how the one-dimensional, n-step, random walk problem can be transformed into the problem of considering paths from $(0, 0)$ to (n, a). Find an expression for $N_{n,a}$, the number of paths from $(0, 0)$ to (n, a).

(b) If $b > a > 0$, then show that there are $N_{n,a} - N_{n,2b-a}$ paths from $(0, 0)$ to (n, a) satisfying the condition that all partial sums are *less* than b.

(c) Suppose that in a ballot, candidate P scores p votes and candidate Q scores q votes where $p > q$. Show that throughout the counting, the probability that there are always more votes for P than for Q equals $(p-q)/(p+q)$.

$(5+7+8=20)$

3. (a) A die is rolled repeatedly. Which of the following are Markov chains? For those that are, provide the transition matrix and for the others provide justification.

   i. The largest number X_n shown up to the nth roll.

   ii. The number N_n of sixes in n rolls.

   iii. At time r, the time C_r since the most recent six.

   iv. At time r, the time B_r until the next six.

   (b) Let $\{Y_k\}$ be a sequence of mutually independent random variables, each assuming the values ± 1 with probability $1/2$. Let $X_n = (Y_n + Y_{n+1})/2$. Does $\{X_n\}$ form a Markov process? Justify your answer.

   (c) Age of lamps is measured in days. If a lamp is k days old, then with probability p_{k+1} it will survive to the $(k+1)$st day and with probability q_{k+1} it will need to be replaced. Model this as a Markov chain and find the transition matrix.

   (d) For a time invariant Markov chain, let $p_{j,k}^{(n)}$ be the probability that the state e_k is reached from state e_j in n steps. Show that

$$p_{j,k}^{(m+n)} = \sum_{\nu} p_{j,\nu}^{(m)} p_{\nu,k}^{(n)}.$$

   (e) Show that all states of an irreducible Markov chain are of the same type.

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

4. (a) Let X be a random variable having its moment generating function to be

$$M_X(t) = \frac{1}{5}\left[\frac{1+2e^{2t}}{e^{4t}} + \frac{e^{3t}+e^{6t}}{e^{5t}}\right].$$

   Find the mean and the variance for X.

   (b) Let X_1, X_2, X_3, X_4 be i.i.d. $\sim N(\mu, \sigma^2)$ with μ unknown but σ^2 known. Consider

$$\begin{aligned} T_1 &= X_1 + X_3, & T_2 &= X_1 + X_2 + X_3 + X_4, \\ T_3 &= (X_1 + X_2 + X_3 + X_4)/4, & T_4 &= (X_1 + X_2 + X_3 + X_4)/3. \end{aligned}$$

   State with justifications which of these are (i) sufficient for μ; (ii) UMVUE for μ; (iii) MLE for μ.

   (c) Define Fisher information. Find the information in X about μ when $X \sim N(\mu, \sigma^2)$. With reference to this example, discuss the suitability of this measure as a measure of information.

   (d) Consider a test of $H_0 : \theta = \theta_0$ versus $H_1 : \theta = \theta_1$. Design a most powerful level α test for this scenario. Justify your answer.

$(5+10+5+10=30)$

INDIAN STATISTICAL INSTITUTE
Semestral Examination, Semester I, 2014-15
M.Tech. (Computer Science)
Computer Organization

Full Marks : 100
Time : 3 Hours

Date : 26.11. 2014

Note : Answer all questions. Marks on each part of a question are indicated in the right margin within parentheses.

1. a) Compute all the prime implicants of the Boolean function $F = \sum(0,1, 2, 5, 7, 8, 9, 10, 13, 15)$ using Quine-McCluskey's method. Hence, derive an irredundant cover of these prime implicants to implement the function F with the minimum circuit complexity.

   b) Show how a synchronous BCD counter can be designed using *J-K* flip-flops. **((8+4) + 8 = 20)**

2. a) Show how two *n*-bit binary numbers can be added in *O(log n)* time using *O(n log n)* hardware with the help of a *precarry adder*. Clearly explain the basic principle of this adder along with the derivation of its time and circuit complexities.

   b) Explain how a carry-propagation-free adder can be designed for adding two *n*-bit binary numbers in *O(1)* time using only *O(n)* hardware. Show the truth table of the logic function to be implemented at each bit position for implementing this idea.

   c) To add two *n*-bit signed integers in 2's complement mode, how would you check for overflow and underflow situations? Justify your answer with proper reasoning. **((6+4) + 8 + 2 = 20)**

3. Explain the underlying principle of the Booth's algorithm for multiplying two binary numbers and show how this idea can be implemented in hardware. Estimate the time needed for multiplying two *n*-bit numbers by using this algorithm and then comment on its capability over the conventional *add and shift* algorithm for multiplication. **(10)**

4. a) Describe how two floating-point numbers represented in IEEE 754-1985 single precision format can be multiplied, taking the rounding off operation, overflow and underflow in consideration. (Clearly explain the roles of guard bit, sticky bit and rounding bit).

   b) Consider that the mantissa field of a floating point number is represented in only 4 bits after the binary point (instead of the 23 bits as in IEEE 754-1985 single precision format). Then show how the decimal numbers -6 and 11 can be

represented, and also show the results of multiplying these two numbers using the above technique after i) rounding to zero, ii) rounding to $+\infty$, iii) rounding to $-\infty$ and iv) rounding to nearest.

c) Show how division of two numbers represented by n-bits in floating-point format can be performed in $O(\log n)$ integer multiplication time.

(10+6+4 = 20)

5. Assuming the MIPS instruction set architecture with four different types of instruction format (R-format, I-format, conditional branch format and unconditional jump format) as an example and considering five different phases of instruction execution, identify the specific operations to be performed at each phase of an instruction execution cycle. Then specify the corresponding hardware blocks (registers, memory, ALU, multiplexers, etc.) and also specify the appropriate data path/control path interconnecting these blocks. Label each such hardware block and data path/control path with brief explanations of their functions (for labelling purposes, you may use the notations given on the attached sheet of paper). Hence, using the description of the required control signals as given on the attached sheet of paper, draw a state transition diagram of the multicycle implementation of the control unit of the processor.

(10 + 8 = 18)

6. a) Consider a computer system with 4 GB main memory, 8 KB cache memory divided in blocks each of size 32 bytes. How many bits are needed for the index field and the tag field for i) direct mapping, ii) 4-way set associative mapping and iii) fully associative mapping of the main memory on the cache memory blocks?

   b) What actions are taken on a read miss of the cache memory?

   c) Distinguish among the terms – compulsory miss, capacity miss and conflict miss of the cache memory.

(3+2+3 = 8)

7. Represent the numbers below in floating-point form, following IEEE 754-1985 single precision standard :

   i) 250 ii) 0 iii) $(-3)^{1/2}$ iv) -18.75×2^{-135}

(4 x 1 = 4)

Actions of the 1-bit control signals

Signal name	Effect when deasserted	Effect when asserted
RegDst	The register file destination number for the Write register comes from the rt field.	The register file destination number for the Write register comes from the rd field.
RegWrite	None	The general-purpose register selected by the Write register number is written with the value of the Write data input.
ALUSrcA	The first ALU operand is the PC.	The first ALU operand comes from the A register.
MemRead	None	Content of memory at the location specified by the Address input is put on Memory data output.
MemWrite	None	Memory contents at the location specified by the Address input is replaced by value on Write data input.
MemtoReg	The value fed to the register file Write data input comes from ALUOut.	The value fed to the register file Write data input comes from the MDR.
IorD	The PC is used to supply the address to the memory unit.	ALUOut is used to supply the address to the memory unit.
IRWrite	None	The output of the memory is written into the IR.
PCWrite	None	The PC is written; the source is controlled by PCSource.
PCWriteCond	None	The PC is written if the Zero output from the ALU is also active.

Actions of the 2-bit control signals

Signal name	Value	Effect
ALUOp	00	The ALU performs an add operation.
	01	The ALU performs a subtract operation.
	10	The funct field of the instruction determines the ALU operation.
ALUSrcB	00	The second input to the ALU comes from the B register.
	01	The second input to the ALU is the constant 4.
	10	The second input to the ALU is the sign-extended, lower 16 bits of the IR.
	11	The second input to the ALU is the sign-extended, lower 16 bits of the IR shifted left 2 bits.
PCSource	00	Output of the ALU (PC + 4) is sent to the PC for writing.
	01	The contents of ALUOut (the branch target address) are sent to the PC for writing.
	10	The jump target address (IR[25–0] shifted left 2 bits and concatenated with PC + 4[31–28]) is sent to the PC for writing.

INDIAN STATISTICAL INSTITUTE

Semestral Examination

M.Tech(CS)-I Year, 2014-2015 (Semester-I)

Data and File Structures

Date: **November 28, 2014** Maximum Marks: **100** Duration: 3.5 **Hours**

Note: The question paper carries a total of 120 marks. You can answer as much as you can, but the maximum you can score is 100.

---

1. Show how a queue may be implemented using two stacks. No other data structures like arrays, linked lists etc. are available to you. Derive the time complexities for insertion and deletion in your queue.

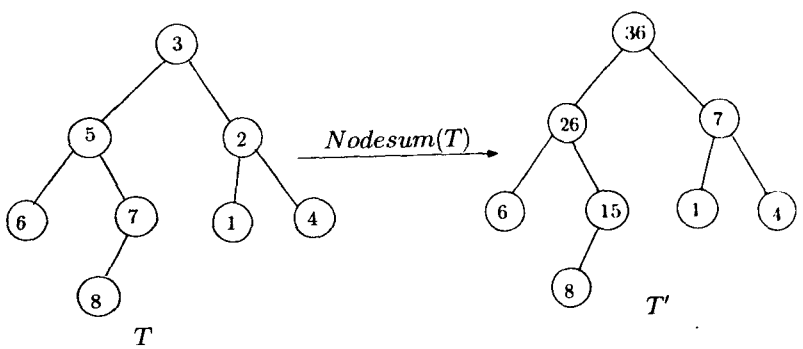
(7+3=10)

2. (a) What is the maximum number of rotations that might be needed to insert a new node into an AVL tree containing n keys? Justify your answer.

   (b) Let T be a binary tree whose nodes are of the following structure.

```
typedef struct node *nodeptr
typedef struct{
int data;
nodeptr left;
nodeptr right;
} node;
```

Let T' be the binary tree generated from T by an operation *Nodesum(T)* which is defined as follows: *Nodesum(T)* replaces the *data* field of each node in T with the sum of its data field and those of all its descendants. Following is an example:



Write an efficient algorithm to implement the operation *Nodesum(T)*.

(10+15=25)

3. (a) Define a B-tree of order m.

   (b) Write an algorithm for inserting a key value in a B-tree. Derive the time complexity of your algorithm.

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

(4+(7+3)+10=24)

4. (a) We call a vertex of a Red-Black tree R a *virtual leaf* if it is a leaf or has a single child. The virtual min-height of a vertex v of R is the length of a shortest path from v to a virtual leaf.

   (i) Prove that all the vertices on a shortest path from a black vertex to a virtual leaf are black.

   (ii) Prove that a non-leaf vertex v is red iff it has a black parent and virtual min-height of v is greater than that of its sibling.

   (b) Show that there exists a sequence of n insert operations that results in a splay tree of height $\Omega(n)$.

((10+6)+6=22)

5. Consider hashing by open addressing. Prove or disprove the following statements:

   (i) The average time to retrieve a key from a hash table can either increase or decrease when a key is deleted from the table (by marking its slot as 'empty').

   (ii) When there are no collisions, linear probing and double hashing are equally good collision resolution methods, in term of worst case number of probes required for an unsuccessful search.

   (iii) In a hash table of size $4N$, the expected number of probes to insert the second N keys is more than double the number of probes required to insert the first N keys, using uniform hashing.

   (iv) The expected number of probes to retrieve the i-th key inserted in a hash table is equal to the expected number of probes in an unsuccessful lookup in table containing $i - 1$ keys.

   (v) When the number of keys N equals the number of slots in the hash table, the expected number of probes for a successful search using linear probing exceeds $\log N$.

$(5 \times 3 = 15)$

6. Consider the *UNION-FIND* problem. Prove that if T is a tree with n nodes created as a result of *weighting rule for union*, no node in T has a level greater than $\lfloor \log n \rfloor + 1$.

(12)

7. (a) Briefly describe the actions taken by the Buffer Manager when a page arrives.

   (b) Briefly describe how variable-length records are organized in a file system.

(7+5=12)

INDIAN STATISTICAL INSTITUTE

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

PROBABILITY AND STOCHASTIC PROCESSES
(BACK PAPER)
Date: 13.12.15 Maximum marks: 100 Duration: 3 hours

Answer all questions. Each question carries 10 marks.

1. Let A, B, C be three arbitrary events. Find expressions for the following events.
   (a) At least one of A, B, C occurs.
   (b) None of the events A, B, C occur.
   (c) Not more than two of the events occur.

2. Let X_1 and X_2 be independent random variables with Poisson distributions $p(k; \lambda_1)$ and $p(k; \lambda_2)$. Show that $X_1 + X_2$ has the Poisson distribution $p(k; \lambda_1 + \lambda_2)$.

3. There are three persons and you want to pick one such that the probability of any one of them getting chosen is 1/3. You are given a fair coin. Describe a method for doing this.

4. State the weak law of large numbers for a sequence of i.i.d. random variables having a common distribution. Assuming that the variance exists, prove the law.

5. Suppose a Markov chain has two states 0 and 1. With probability p_i the chain moves from state i to state $(1-i)$ and with probability $(1-p_i)$ it remains in state i. Find the stationary distribution of this Markov chain.

6. Suppose that in an election, a candidate P get p votes and another candidate Q gets q votes, with $p > q$. Show that the probability that throughout the counting there are always more votes for P than for Q equals $(p-q)/(p+q)$.

7. Let $\mathcal{E}$ be a recurrent event and

$$\begin{aligned} u_n &= \mathsf{P}[\mathcal{E} \text{ occurs at the } n\text{th trial }]; \\ f_n &= \mathsf{P}[\mathcal{E} \text{ occurs for the first time at the } n\text{th trial }]. \end{aligned}$$

   Let $U(s)$ and $F(s)$ be the generating functions for $\{u_n\}$ and $\{f_n\}$. Show that

$$U(s) = 1/(1 - F(s)).$$

8. Define mean deviation from a measure of central tendency. Show that the mean deviation is minimum when taken about the median.

9. The frequency distribution of family-size for 250 families is given below.

family-size	frequency
1	4
2	22
3	25
4	45
5	52
6	41
7	36
8	15
9	7
10	3

Find the mean, median and the mode.

10. When is a statistic called sufficient for a parameter? Let $X_1, \ldots, X_n$ be i.i.d. Geometric(p) with common p.m.f. $f(x;p) = p(1-p)^x$, $x = 0, 1, 2, \ldots$. Show that $\sum_{i=1}^{n} X_i$ is sufficient for p.

Indian Statistical Institute

Semester-II 2014-2015

M.Tech.(CS) - First Year

Mid-term Examination (23 February, 2015)

Subject: Operating Systems

Total: 40 marks Maximum marks: 30 Duration: 2.5 hrs.

Please keep your answers brief and to the point.

1. For each of the standard C library functions listed below, explain in 1-2 lines whether the function may result in a system call.

   (a) `strcpy(char *dest, const char *src)`: copies the string pointed to by **src**, including the terminating null byte, to the buffer pointed to by **dest**;

   (b) `strdup(const char *s)`: returns a pointer to a new string which is a duplicate of **s** (memory for the new string is obtained with `malloc()`, and can be freed with `free()`).

   (c) `stat(const char *path, struct stat *buf)`: returns (in **buf**) information about a file (the size of the file, the time when it was last modified, etc.).

   [2×3=6]

2. (a) What is the difference between the types of information stored by the OS in the *u area* and the *proc* structure of a process?

   (b) Using your answer above, explain in 1-2 lines each whether the following items of information should be stored in the *proc* structure or the *u* area: (i) process state, (ii) disk quota and resource limits.

   (c) Using an example, explain why environment variables are a useful component of a process' context. Where are environment variables usually stored?

   [2+3+(2+1)=8]

3. Suppose that the (conventional) processes described in the table below are the *only* processes that exist on a single-processor system that is running version 2.4 of the Linux kernel. The table also provides the durations of successive CPU and I/O bursts (in ms) for these processes. Assume that all I/O requests are for different devices.

Process	nice	counter	CPU burst 1	I/O burst 1	CPU burst 2
P_1	-5	10	3	5	4
P_2	0	6	4	6	3
P_3	4	9	6	1	7

Construct a Gantt chart showing how the processes are scheduled **up to the end of the current epoch.** [12]

P.T.O.

4. Consider the following function to add a node p to a binary search tree whose root is specified via the root pointer.

```
void insert(NODE *p, NODE **root)
{
    NODE *r;
    if (root == NULL) return;
    if (*root == NULL) {
        *root = p; return;
    }
    r = *root;
    if (p->value == r->value) return;
    if (p->value < r->value) {
        insert(p, &(r->left)); return;
    }
    if (p->value > r->value) {
        insert(p, &(r->right)); return;
    }
}
```

Suppose two processes P_1 and P_2 share a binary search tree. Using an example, show how the tree may get corrupted if the two processes execute `insert` concurrently. You should show the initial state of the tree, the calls to `insert` by the processes, the intended final state of the tree, the interleaving of operations by the two processes, and the actual final state of the tree. [8]

5. Consider the following solution to the n-process Critical Section Problem (CSP) that uses an atomic `TestAndSet` function.

```
shared char lock = 0;
...
while (TestAndSet(&lock)); critical_section(); lock = 0;
```

For each of the three desirable properties that a solution to the CSP should possess, explain (i) the property; and (ii) whether the above solution possesses the property. [6]

INDIAN STATISTICAL INSTITUTE

Mid-Semestral Examination: 2014-15

Course Name: M. Tech. I Year

Subject Name: Database Management Systems

Date: 24/02/2015 **Maximum Marks: 50** **Duration: 2hours**

Answer all questions. Write your assumptions clearly (if any).

1. A socio-economy survey has been done in a set of remote tribal villages in Eastern India to study their agricultural practices. A village is identified by a name, which may be considered unique. Each village has certain number of houses indentified by a house number. Each house belongs to a certain tribe. In each house, the number of persons involved in agriculture may vary. Some house may also have some non-agricultural come.

   Villages have three types of lands – highly fertile, moderately fertile and un-fertile. Un-fertile lands are used for raising fruit trees. In each village, each plot of land is of specific land type and is identified by a plot number and cropping practice (mono-cropping and multi-cropping). Each plot is owned by any one of the houses in that village.

   Village authority, for each village, keeps an account of the crops cultivated in that village, in a calendar year and the average yield of each of them in tons/acre.

   From the above description draw an ER/EER diagram. From the diagram derive a set of relation applying the standard mapping rules.

   20+20=40

2. Consider the following relations:
   Parts(part_id, part_name, part_colour)
   Suppliers (supplier_id, supplier_name, supplier_address)
   Billing(Supplier_id, Part_id, Price)

   a. Find the parts each of which are supplied by the entire supplier at the same price.
   b. Find the supplier who sells any parts at the highest price.

   05+05=10

INDIAN STATISTICAL INSTITUTE

Mid-Semestral Examination

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

Design and Analysis of Algorithm

Date : 26.02.15 Maximum Marks : 60 Duration : 2.5 Hours

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

1. Consider an n-node complete binary tree T. Each node v of T is labeled with a real number x_v and label of all nodes are distinct. A node v of T is local minimum if the label x_v is less than the label of its parent as well as all the children of node v. Write an efficient algorithm to find a local minimum of T and compute the complexity of the algorithm. [15]

2. We are looking at the price of a given item over n consecutive days, numbered $1, 2, \ldots, n$. For each day i, we have a price $p(i)$ per unit for the said element on that day. How should we choose a day i on which to buy the element and a later day on which to sell it in order to maximized the profit. Writean efficient algorithm that finds the optimal members i and j and analyze the algorithm complexity. [15]

3. Suppose you are given a set of positive integers $A = \{a_1, a_2, \ldots, a_n\}$ and a positive integer B. A subset $S \subseteq A$ is called feasible if the sum of the numbers in S does not exceed B. The sum of the numbers in S is called the total sum of S.

   You would like to select a feasible subset S of A whose total sum is as large as possible.

   Below is an algorithm for this problem.

   Initialized S as null set and T as zero;

   for $i = 1, 2, \ldots, n$ do

   if $T + a_i \leq B$ then $S \leftarrow S \cup \{a_i\}$; $T \leftarrow T + a_i$ endif

   endfor

   Give an instance in which the total sum of the set S returned by this algorithm is less than half the total sum of some other feasible subset of A.

   Give a polynomial time approximation algorithm for this problem with the following guarantee: It returns a feasible set $S \subseteq A$ whose total sum is at least half as large as the maximum total sum of any feasible set $S' \subseteq A$. Your algorithm should have a running time of at most $O(n \log n)$. [10+10=20]

4. An array $A[1..n]$ is unimodal if it consists of an increasing sequence followed by a decreasing sequence, or more precisely, if there is an index $m \in \{1, 2, \ldots, n\}$ such that $A[i] < A[i+1]$

for all $1 \leq i < m$, and $A[i] > A[i+1]$ for all $m \leq i < n$. In particular, $A[m]$ is the maximum element, and it is the unique locally maximum element whose neighboring elements $A[m-1]$ and $A[m+1]$ have values less than it. Give an algorithm to compute the maximum element of a unimodal input array $A[1..n]$ in $O(logn)$ time. [10]

5. You are given a set of real numbers $A = \{a_1, a_2, \ldots, a_n\}$. Present an algorithm that computes the largest interval $[x, y]$, such that no elements of A lie within this interval. Your algorithm should run in $O(n)$ time. Justify your algorithms correctness, and derive its running time. [15]

Indian Statistical Institute
Mid-Semester Examination: 2015
Course Name: M. Tech in Computer Science
Subject Name: Computer Networks

Date: 27-02-2014 Maximum Marks: 60 Duration: 2 hours 30 minutes

Instructions:

You **may** attempt **all** questions which cary a total of **65** marks. However, the maximum marks you can score is only **60**.

1. (a) Define delta modulation. Define quantizing noise and slope overload noise. State the effects of quantization level on quantizing noise and slope overload noise. [2 + (2 + 2) + 2 = 8]

   (b) What is scrambling technique? What are its design goals? What is the result of scrambling the sequence 1 0 0 0 0 0 0 0 0 using $B8ZS$ scrambling technique? Assume that the last voltage pulse preceding the octet of eight zeros was positive. [2 + 2 + 2 = 6]

   (c) For the bit stream 1 1 1 1 1 1 1 0 1 1 1 1 1 1 1, sketch the waveforms for each of the following encoding scheme. Assume that the signal level for the most recent preceding 1 bit for Bipolar-AMI has a negative voltage, and the most recent preceding 0 bit for pseudoternary has a negative voltage.

      i. NRZ-L.
      ii. Bipolar-AMI.
      iii. Manchester.
      iv. Pseudoternary. [2 + 2 + 2 + 2 = 8]

   (d) Derive an expression for baud rate D as a function of bitrate R for $QPSK$ using NRZ-L digital encoding technique. [2]

   (e) Consider a periodic signal $x(t)$ whose period is 5 millisecond. It has been determined that frequencies from f to $8f$ are required to represent the signal. Calculate the effective bandwidth of $x(t)$ in Hz. [3]

   (f) Consider a communication link with bandwidth $B = 4000\ Hz$ and $S/N = 30\ dB$. Calculate its maximum data rate in bits per second according to Shannon's theorem. Nyquist's theorem cannot be applied here because a factor is unknown. Point out what factor is unknown. [3 + 2 = 5]

   (g) What does the sampling theorem tell us concerning the rate of sampling required for an analog signal? [2]

2. (a) Consider the use of Cyclic Redundancy Code (CRC) for error detection. If the data frame is 1 1 1 0 0 0 1 1 and the generator code is 1 1 0 0 1 1, calculate the CRC and produce the final transmitted frame. Show the steps of your calculation. [3 + 2 = 5]

   (b) Can the generator polynomial $x^8 + x^2 + x + 1$ in CRC detect all odd number of errors? Justify your answer. [3]

   (c) Suppose that the Hamming distance of a code is d. Show that if the receiver maps the received string into a codeword at smallest distance from the string, all combinations of fewer than $d/2$ errors will be corrected by this code. [4]

(d) Assuming that even parity is used in the Hamming code, show the bit pattern transmitted for the message 1 1 0 1 0 0 1 1. Suppose that the third bit from the left is inverted due to transmission errors. Show how Hamming code detects and corrects it. [5]

(e) The following character encoding is used in a data link protocol: A 0 1 0 0 0 1 1 1; B 1 1 1 0 0 0 1 1; $FLAG$ 0 1 1 1 1 1 1 0; ESC 1 1 1 0 0 0 0 0. Show the bit sequence transmitted (in binary) for the four-character frame: A B ESC $FLAG$ when each of the following framing methods are used:

- Character count
- Flag bytes with byte stuffing
- Starting and ending flag bytes with bit stuffing [2+2+2=6]

3. A channel has a data rate of R *bps* and a propagation delay of t *sec/km*. The distance between the sending and receiving nodes is L *km*. Nodes exchange fixed-size frames of B bits using a sliding window protocol with go-back-N ARQ. Assume that ACK frames are negligible in size and processing time at the nodes is also negligible.

(a) Find a formula that gives the required window size W (in terms of R, t, B and L) for the maximum channel utilization.

(b) Find a formula that gives the minimum value of k (in terms of R, t, B and L) corresponding to the window size W obtained above, where k is the number of bits used for the representation of sequence numbers.

[4+4=8]

INDIAN STATISTICAL INSTITUTE
Mid-Semester Examination
M. Tech. (CS) I year (2nd Sem): 2014–2015
Automata, Languages and Computation

Date: 03. 03. 2015 Marks : 80 Time : 3 Hours

Answer all questions.

Please try to write all the part answers of a question at the same place.

1. (a) Give examples of languages L such that
   i. both L and L^* are finite,
   ii. both L and L^* are infinite,
   iii. L is finite, but L^* is infinite,
   iv. L is infinite, but L^* is finite.

   If you think that some of the cases listed above are impossible, then provide just a single-line explanation for each such case.

   (b) If L is a language over $\{a, b\}$, and $|L| = n$, prove that $|L^2| \geq 2n - 1$.

   [(3 + 3 + 3 + 3) + 5]

2. (a) Construct a DFA to recognize all strings over $\{a, b\}$ that begin with a, but do not contain aab as a substring.

   (b) Prove that any DFA that recognizes the language represented by the regular expression $(ab)^* + a$ requires at least 4 states.

   [10 + 10]

3. (a) Derive regular expression for the complement of $L(a^*b(aa)^*)$.

   (b) What is the worst case time complexity of converting
   i. an n-state NFA into its equivalent DFA?
   ii. an n-state DFA into its equivalent RG?

   Justify your answers.

   (c) Are irregular languages closed under union and intersection?

   [8 + (3 + 3) + (3 + 3)]

4. Are the following languages regular or irregular? Prove your claims.

   (a) $L = \{w \in \{a, b\}^* : w = w^R\}$,

   (b) $L = \{a^n : n \text{ is a perfect square}\}$.

   (c) $L = \{a^n : n \text{ is divisible by 2 or 3 or 5}\}$.

   [6 + 8 + 6]

INDIAN STATISTICAL INSTITUTE
Back-Paper Examination, Semester I, 2014-15
M.Tech. (Computer Science)
Computer Organization

Full Marks : 100
Time : 3 Hours

Date : 18.03. 2015

Note : Answer all questions. Marks on each part of a question are indicated in the right margin within parentheses.

1 a) Show, with appropriate logic, how an *EX-OR* operation between two Boolean variables *A* and *B* can be implemented using only *4 NAND* gates, assuming that the complements of the variables *A* and *B* are not available.

b) Minimize the Boolean function $f = \sum(0,1, 2, 6, 7, 12, 14, 15)$.

c) Show how a synchronous BCD counter can be designed using *J-K* flip-flops. **(6+6+8 = 20)**

2. a) Assume that you have to add two 4-bit binary numbers by means of a parallel binary full adder circuit with carry look-ahead mechanism. Write down the logic functions to generate the carry at all the bit positions for this parallel adder circuit.

b) Explain the underlying principle of the Booth's algorithm for multiplying two binary numbers and show how this idea can be implemented in hardware. Estimate the time needed for multiplying two *n*-bit numbers by using this algorithm and then comment on its capability over the conventional *add and shift* algorithm for multiplication. **(10+10=20)**

3. a) Describe how two floating-point numbers represented in IEEE 754-1985 single precision format can be multiplied, taking the rounding off operation, overflow and underflow in consideration (Clearly explain the roles of guard bit, sticky bit and rounding bit).

b) Consider that the mantissa field of a floating point number is represented in only 4 bits after the binary point (instead of the 23 bits as in IEEE 754-1985 single precision format). Then show how the decimal numbers -5 and 10 can be represented, and also show the results of multiplying these two numbers using the above technique after i) rounding to zero, ii) rounding to $+\infty$, iii) rounding to $-\infty$ and iv) rounding to nearest.

(10+10 = 20)

4. Assuming the MIPS instruction set architecture with four different types of instruction format (R-format, I-format, conditional branch format and unconditional jump format) as an example and considering five different phases of instruction execution, identify the specific operations to be

performed at each phase of an instruction execution cycle. Then specify the corresponding hardware blocks (registers, memory, ALU, multiplexers, etc.) and also specify the appropriate data path/control path interconnecting these blocks. Label each such hardware block and data path/control path with brief explanations of their functions (for labelling purposes, you may use the notations given on the attached sheet of paper).

Hence, using the description of the required control signals as given on the attached sheet of paper, draw a state transition diagram of the multicycle implementation of the control unit of the processor. **(10 + 6 = 16)**

5. a) Consider a computer system with 1 GB main memory, 4 KB cache memory divided in blocks each of size 16 bytes. How many bits are needed for the index field and the tag field for i) direct mapping, ii) 2-way set associative mapping and iii) fully associative mapping of the main memory on the cache memory blocks?

   b) What actions are taken on a read hit of the cache memory?

   c) What actions are taken on a write miss of the cache memory?

   (6+4+6=16)

6. Represent the numbers below in floating-point form, following IEEE 754-1985 single precision standard :

   i) 16 ii) 0 iii) $(-5)^{3/4}$ iv) 6.5×2^{-140}

   (4 x 2 = 8)

Indian Statistical Institute

M.Tech. (CS) I

Discrete Mathematics

Back Paper Examination

Maximum Marks: 100

Date: March 19, 2015
Time 3 hours

The question paper contains 8 questions. Total marks is 100. Unless otherwise mentioned, all notations are the same as presented in class.

1. Let $\mathcal{A} = (A_1, A_2, \ldots A_m)$ be a collection of subsets of a set X. A system of distinct representatives (SDR) for $\mathcal{A}$ is a set of distinct elements $a_1, a_2, \ldots a_m$ in X such that $a_i \in A_i$. Prove that $\mathcal{A}$ has an SDR if and only if $|\cup_{i \in S} A_i| \geq |S|$, for every $S \subseteq \{1, 2, \ldots, m\}$ (10)

2. (a) Prove that every planar graph has a vertex whose degree is less than or equal to 5
   (b) Let G be a graph on n vertices. Prove that if $\omega(G) \leq r$, then $e(G) \leq (1 - 1/r)n^2/2$. Hence prove that, $\chi(G) \geq n^2/(n^2 - 2e(G))$. (5+(7+3) = 15)

3. Prove that a set of edges in a connected planar graph G forms a spanning tree of G if and only if the duals of the remaining edges form a spanning tree of G^*. (15)

4. Prove that every strongly connected tournament has a Hamiltonian circuit (15)

5. Let G be a connected graph with atmost two odd vertices. Write an algorithm which constructs an Eulerian trail. Prove the correctness of the algorithm and deduce the time complexity. (8 + 4 + 3 = 15)

6. Deduce the number of distinct colorings that can be obtained by coloring the vertices with colors blue and red. (10)

7. Prove that $\alpha \vee \beta \in \Delta$, if and only if $\alpha \in \Delta$ or $\beta \in \Delta$. (5+5=10)

8. If x does not occur free in α, then prove that,

$$\vdash (\alpha \rightarrow \forall x\, \beta) \leftrightarrow \forall x(\alpha \rightarrow \beta).$$

   Use this result to prove that,

$$\vdash (\exists x\, \beta \rightarrow \alpha) \leftrightarrow \forall x(\beta \rightarrow \alpha).$$

(6 + 4 = 10)

INDIAN STATISTICAL INSTITUTE

First Semestral Examination : 2014-15 : Back Paper

Course : Master of Technology in Computer Science (First Year)

Subject : Introduction to Programming (A1) for M.Tech. (CS) I

Date : 20·03·2015 [redacted] Maximum Marks : 100 Duration : 3 Hours

Note: You are expected to code in C to solve the following problem(s), and you are allowed to use any academic resource of your choice (except the Internet) during the course of this examination. Write the theoretical part(s) of the solution(s) in the answer-sheet provided, and submit the relevant C source file(s) separately. In the theoretical part(s) of the solution(s), you must provide basic sketch of the algorithm(s) that you use, along with basic complexity analysis of your chosen algorithm(s), wherever appropriate. You must justify the C data-type(s) that you use while coding, and provide basic pictorial illustration(s) of the memory stack-frame(s) for clarification, wherever appropriate.

Problem 1 : Short Questions

Maximum Marks: 50 = 5 + 5 + 10 + 10 + 10 + 10

1A. If you increment a pointer p, simply using p++, how much does the pointer value increase by? Write a C code snippet to demonstrate how you can increase the value of an integer pointer by 1.

1B. Write the pseudocode algorithm and/or the C code snippet (only for the relevant function) to insert a new node in a linked list. Take care of the appropriate extremal cases, if any.

1C. Write a C code to determine the length of an input string. You are not allowed to use the `string.h` library, and the input string should be accepted as a command-line input from the user.

1D. Hamming distance of two binary strings of equal length refers to the number of places where they differ. Write a C program that takes two binary strings as input from the user, and outputs the Hamming distance between them.

1E. Write a C program that takes two integers as input, and outputs their *greatest common divisor*. Your program should use the principles of recursive function calls to compute the GCD.

1F. Write a C program that takes the name of an existing text file as input from the user, and outputs the individual counts of alpha-numeric characters, words, and lines in the file.

Problem 2 : Square Matrices

Maximum Marks: 25

The determinant of a square matrix ($n \times n$, say) is generally computed using the method of 'determinant expansion by minors'. For every element x_{ij} in the $n \times n$ square matrix X, the corresponding minor M_{ij} is defined as the $n - 1) \cdot (n - 1)$ square matrix formed by eliminating the i-th row and j-th column of the original matrix X. The determinant of X is computed in terms of the elements and minors of a single row or column, as follows.

$$\det(X) \quad \sum(-1)^{i+j} \cdot x_{ij} \cdot \det(M_{ij}), \qquad \text{where either row } i \text{ or column } j \text{ is fixed.}$$

In the special case of a 2×2 matrix X, the expression simplifies to $\det(X) = x_{11} \cdot x_{22} - x_{12} \cdot x_{21}$.

Design an algorithm that uses recursion to compute the determinant of any square matrix. Write a C program to implement your algorithm. It should take as input from the user a square matrix X of arbitrary dimension $n \times n$, consisting only of integer elements, and compute its determinant using 'determinant expansion by minors'.

Problem 3 : Rational Numbers

Maximum Marks: 25

Rational numbers are expressed in the form $\frac{x}{y}$, where $x, y \in \mathbb{Z}$ and $y \neq 0$. Write a C program that creates an appropriate structure to represent rational numbers, and devise a strategy to numerically compare the magnitude (equal, less than or greater than) of any two rational numbers of this form.

Finally, your C program should take as input from the user an array of rational numbers (of arbitrary length n), and output the sorted version of the array. You may use any inbuilt sorting routine that you know, or implement any sorting algorithm of your choice.

Solving a problem correctly is necessary, but not sufficient, as it does not guarantee the maximum marks alloted for the problem. Sufficient credit is reserved in each case for *smart* algorithm and *good* coding practice. Good luck! ☺

Indian Statistical Institute

Second Semestral Examination: 2015

Course: M. Tech. in Computer Science

Subject: Computer Networks

Date: 20-04-2015 Total Marks: 110 Duration: 3 hours

Instructions: You **may** attempt **all** questions. But the maximum marks you can score is only **100**.

1 (a) Describe the leaky bucket mechanism for congestion control. What problems of the leaky bucket are resolved by using a token bucket mechanism? [3+3=6]

(b) Computer A has 19.5 *MBytes* to send over a network and transmits the data in a burst at the rate of 6 *Mbps*. The maximum transmission rate across routers in the network is 4 *Mbps*. If Computer A's transmission is shaped using a leaky bucket, how much capacity must the queue in the bucket hold not to discard any data? [5]

(c) Suppose that the TCP congestion window is set to 18 KB and a timeout occurs. How large should the window be if all the next four transmission bursts are successful? Assume that the maximum segment size is 1 KB. [5]

2 (a) How routing tables are updated using distance vector routing algorithm? Explain with an example how information about a node failure propagates using this algorithm. What problem is encountered in deciding whether a host has become unreachable? In what circumstances is it possible to resolve this problem? [5+5+3+2=15]

(b) Consider the subnet as shown in Figure 1 with communication delays shown as labels on the bidirectional links. Distance vector routing is used, and the following vectors have just come into router C: from B: $(5, 0, 8, 12, 6, 2)$; from D: $(16, 12, 6, 0, 9, 10)$; and from E: $(7, 6, 3, 9, 0, 4)$. The measured delays to B, D, and E, are 6, 3, and 5, respectively. What is C's new routing table? Mention both the outgoing link to be used and the expected delay. [10]

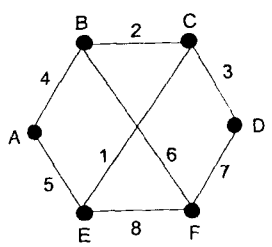


Figure 1: A subnet.

(c) Explain with an example how route discovery is performed in AODV routing. [8]

3 (a) Briefly describe UDP. [5]

(b) Write the purpose of the following two fields of the TCP header:

i. Window size

ii. Urgent pointer [3+3=6]

4 (a) Consider the wireless topology shown in Figure 2, comprised of five nodes. Nodes A, B, C, and D all have equal-sized transmission ranges, while E has a smaller range. Transmission range of a node is shown in the figure as a circle centered around that node. Assume that the transmissions of two nodes interfere if and only if they transmit at the same time and their transmission ranges overlap. In these problems, assume that packet losses occur only due to collisions. List the potential hidden terminals and exposed terminals when

i. Node A transmits to node B.

ii. Node B transmits to node C. [4+4=8]

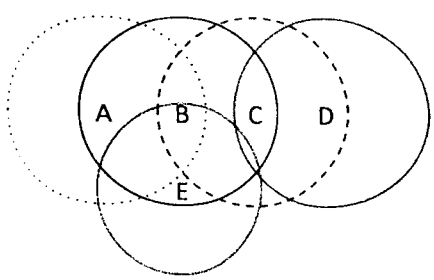


Figure 2: Wireless network with five nodes.

(b) Explain MACA and comment whether MACA solves the hidden terminal problems. [5]

6 (a) Explain adaptive tree walk protocol. A collection of 2^n stations uses the adaptive tree walk protocol to arbitrate access to a shared cable. At certain instant, two of them becomes ready. What is the 1) minimum, 2) maximum, and 3) mean number of slots to walk the tree if $2^n >> 1$? [2+2+4=8]

(b) Consider the Pure ALOHA, Slotted ALOHA, and Non-persistent CSMA protocols. Which one will you use at high load? Justify your answer. [4]

(c) Ten thousand airline stations are competing for the use of a single slotted ALOHA channel. On an average a station makes 18 $requests/hour$. A slot is 125 microseconds. What is the approximate total channel load? [5]

5 Consider the network as shown in Figure 3, with link capacities (in $Kbps$) shown as labels on the links. Links are bidirectional. The matrix shown in Figure 4 gives: 1) number of packets to be sent from node i to node j, and 2) the route from node i to node j to be

used. Packet size is exponentially distributed with a mean of $1/\mu = 800$ *bits*. Based on the given routing and traffic matrix:

(a) Compute the mean number of hops per packet.

(b) Compute the mean packet delay.

(c) Compute the mean packet delay when the network is almost empty.

(d) Compute the maximum amount of traffic the network can support under the given routing algorithm.

(e) Compute the theoretical maximum amount of traffic the network can support under any optimal routing algorithm. [3+5+3+5+4=20]

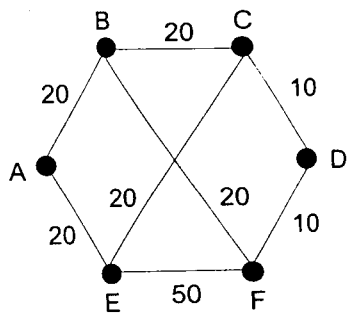


Figure 3: Subnet and delays.

	A	B	C	D	E	F
A		9 AB	4 ABC	1 ABFD	7 AE	4 AEF
B	9 BA		8 BC	3 BFD	2 BFE	4 BF
C	4 CBA	8 CB		3 CD	3 CE	2 CEF
D	1 DFBA	3 DFB	3 DC		3 DCE	4 DF
E	7 EA	2 EFB	3 EC	3 ECD		5 EF
F	4 FEA	4 FB	2 FEC	4 FD	5 FE	

Figure 4: Routing and traffic matrix.

INDIAN STATISTICAL INSTITUTE
Semester Examination
M. Tech. (CS) I year (2nd Sem): 2014–2015
Automata, Languages and Computation

Date: 23. 04. 2015 Marks : 60 Time : 3 Hours

Answer any 6 questions.
Please try to write all the part answers of a question at the same place.

1. (a) Given an alphabet $\Sigma = \{0, 1\}$, is the relation ρ on Σ^* right-invariant for each of the following cases? For two strings $w_1, w_2 \in \Sigma^*$, define $w_1 \rho w_2$ iff they contain
   i. the same number of zeros.
   ii. an odd number of zeros.

   (b) Let $L = \{ww : w \in \{0,1\}^*\}$. Find the value of $index(\equiv_L)$, where $w_1 \equiv_L w_2$ iff $\forall z \in \{0,1\}^*$, $w_1 z \in L \iff w_2 z \in L$. From this value, what can you say about the regularity or irregularity of L?

   [(2 + 2) + 6]

2. Suppose you use Myhill-Nerode theorem to minimize a DFA D to D_{min}. Suppose I use another method, say GKP method, completely different from the approach of Myhill-Nerode theorem, to obtain a reduced DFA D_{red} from D. Let n_{min} and n_{red} be the number of states of D_{min} and D_{red} respectively.

   (a) For arbitrary D, is it always true that $n_{min} = n_{red}$?

   (b) (i) If yes, formally prove your claim.
   (ii) If not, what can you say about their relationship? In this case, if for some D, it so happens that $n_{min} = n_{red}$, what can you say about the relationship between D_{min} and D_{red}? Formally prove your claim.

   [2 + 8]

3. Define acceptance by empty stack and by final state for a pushdown automaton. Prove that these two mechanisms of acceptance are equivalent. [2 + 8]

4. (a) In the proof of the Pumping Lemma for context-free languages, what will be the problem if you allow the first repetition along the longest path from a leaf to the root of the derivation tree to be the starting non-terminal of the grammar? How is this unwanted situation prevented in the proof?

   (b) State Ogden's Lemma.

   (c) Show that the language L defined in Question 1(b) is not context-free.

   [(2 + 1) + 3 + 4]

5. (a) Convert $\{\{S, A, B\}, \{a, b\}, \{S \to ASA \mid aB, A \to B \mid S, B \to b \mid \epsilon\}, S\}$ into Chomsky normal form.

(b) Why and how to remove left recursion during conversion of a context-free grammar into Greibach normal form?

[6 + (2 + 2)]

6. (a) Show that the grammar $G = \{\{S\}, \{a, b\}, \{S \rightarrow aSb \mid aS \mid \epsilon\}, S\}$ is ambiguous.

(b) Apply Cocke-Younger-Kasami algorithm to test whether $aabb \in L(G)$, where G in Chomsky normal form is given as $\{\{S, S_1, Z, A, B\}, \{a, b\}, P, S\}$, with $P = \{S \rightarrow AZ \mid AS_1 \mid a, S_1 \rightarrow AZ \mid AS_1 \mid a, Z \rightarrow S_1B, A \rightarrow a, B \rightarrow b\}$. **(No need to show details, just show the matrix of substring-generating non-terminals).**

[4 + 6]

7. (a) If the space-time geometry would allow you to replace the one-dimensional input tape of a Turing machine by a 5-dimensional tape, would that increase the power of the machine? Justify your answer.

(b) Let $L_{DFA} = \{\langle D, w\rangle$: the DFA D accepts the string $w\}$. Prove that L_{DFA} is decidable.

(c) What is the meaning of a function $f : \mathbb{N}^k \rightarrow \mathbb{N}$ to be computable, where $\mathbb{N}$ is the set of natural numbers, and k is a fixed positive integer?

[(2 + 2) + 4 + 2]

8. Construct a language that is not recursively enumerable. Give a detailed proof that your constructed language is indeed not recursively enumerable. [4 + 6]

9. (a) What is the Post Correspondence Problem (PCP)? If you have an oracle that solves the PCP, can you use that oracle to solve the halting problem?

(b) What is Strong Church-Turing Thesis? Can it be refuted by an alternative model of computation?

[(4 + 2) + (2 + 2)]

Indian Statistical Institute

Semester-II 2014-2015

M.Tech.(CS) - First Year

Semestral Examination (27 April, 2015)

Subject: Operating Systems

Total: 55 marks Maximum marks: 50 Duration: 3.5 hrs.

Please keep your answers brief and to the point.

1. (a) Recall that the integer value of a binary semaphore can only be 0 or 1. Write pseudo-code to show the functioning of the blocking version of the *wait* and *signal* operations on **binary** semaphores. (Do not write code for ordinary counting semaphores.)

   (b) Show how you would implement the above *wait* and *signal* system calls using an atomic `swap(int *x, int *y)` operation. [4 + 6 = 10]

2. (a) Define internal and external fragmentation.

   (b) For which of the following memory management schemes is internal / external fragmentation possible: (i) paging, (ii) segmentation? In each case, justify your answer in 2-3 lines.

   (c) Explain how page-fault frequency monitoring can be used to prevent thrashing.

   (d) Clearly explain what hardware support is provided in i386 processors' architecture to speed up the translation of 2-dimensional logical addresses to 32-bit linear addresses.

   [2 + (2 + 2) + 3 + 3 = 12]

3. Consider a variant of the SVR2 filesystem in which the inode table of contents consists of 15 entries, comprising 10 direct, 2 single indirect, 2 double indirect, and 1 triple indirect entry. Assume that

   - the block size for this filesystem is 4096 bytes,
   - each block number can be stored in 32 bits,
   - the format of a directory data block is the same as in SVR2 filesystems.

   (a) Calculate the maximum possible file size for this filesystem.

   (b) The `stat` command is used to display information about a file on UNIX-like filesystems. When the command is run on a particular computer, it shows that the inode number corresponding to both the root directory (/) and the `/home` directory is 2. Explain why / how this can happen.

   (c) Recall that the `mkdir` and `touch` commands can be used to create the specified directories and files respectively. Draw a diagram of the data block for the directory `myNewDirectory` after the following commands are executed to create some new files and directories.

```
$ mkdir myNewDirectory
$ mkdir myNewDirectory/a myNewDirectory/b myNewDirectory/c
$ touch myNewDirectory/d myNewDirectory/e
```

   What will change if the file `myNewDirectory/d` is now deleted?

(d) Recall that the *link count* for an inode (with inode number n, say) within a filesystem may be defined as the **total** number of directory data block entries (over the entire filesystem) in which the inode number field contains n. Using this definition, explain why the link count for the directory `myNewDirectory` would be 5 after the above operations. [3+2+(2+1)+5=13]

4. *ProZilla* is an application that is used to download files from HTTP or FTP servers. It makes multiple connections to a server and downloads a file in multiple parts simultaneously to reduce the time required for a download. Suppose that you implement ProZilla using the following outline.

```
...
partSize = totalDownloadSize / numParts;
fp = fopen(targetFile, "w");
for (count = 0; count < numParts; count++) {
    if (0 == (p = fork())) {
        socketfd = socket(...); /* open connection */
        getPartialData(socketfd, count, buffer, partSize); /* download my part into buffer */
        fseek(fp, count * partSize, SEEK_SET); /* go to my part of file */
        fwrite(buffer, sizeof(char), partSize, fp); /* dump my part */
        return;
    }
    else {
        count++;
        printf("Forked child (pid %d) to take care of fragment %d\n", p, count);
    }
}
```

(a) Draw a diagram to show the relationship between the inode cache, global file table (GFT), and process file descriptor tables for the processes involved. Show the relevant fields in the GFT entries. Hence explain why this is an incorrect solution.

(b) Suggest a simple modification that will make the scheme work correctly. [(4+3)+3=10]

5. Consider the following Resource Allocation State involving 5 processes and 5 resources.

$$Total = [\ 5\ 4\ 5\ 3\ 6\ ]$$

$$Max = \begin{bmatrix} 5 & 1 & 1 & 1 & 5 \\ 2 & 2 & 2 & 2 & 2 \\ 0 & 1 & 2 & 3 & 4 \\ 4 & 0 & 4 & 0 & 4 \\ 2 & 0 & 1 & 2 & 3 \end{bmatrix} \quad Alloc = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 \\ 2 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} \quad Req = \begin{bmatrix} 1 & 0 & 0 & 0 & 4 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 2 & 2 \\ 1 & 0 & 0 & 0 & 3 \\ 2 & 0 & 1 & 1 & 1 \end{bmatrix}$$

$Total[i]$ specifies the total number of instances of resource i that exist in the system (including both allocated and free instances). $Max[i,j]$, $Alloc[i,j]$, and $Req[i,j]$ denote, respectively:

- the maximum number of instances that process i may request of resource j,
- the number of instances of resource j currently allocated to process i, and
- the number of instances of resource j that process i is currently requesting.

Show that the system is neither in a safe state nor in a deadlocked state. [5+5=10]

INDIAN STATISTICAL INSTITUTE

Semestral Examination

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

Design and Analysis of Algorithms

Date 30·4·15 Maximum Marks : 100 Duration : 3 Hours

Note : Answer all questions from Group A and you may answer any part of any question from Group B. Maximum you can score is 100.

Group A

Answer all questions in this group. Each question caries 1 mark. If you ommit or answer wrongly, then you will get negetive marks and in that case 3 marks will be duducted for each such wrong answer or ommited question.

1. Let $f(n)$ and $g(n)$ be asymptotically positive functions. State whether the following statement is *true* or *false*.

   $f(n) + g(n)$ is $\Theta(\min(f(n), g(n)))$

2. Suppose $T(n)$ is the running time of an algorithm. We know that $T(1) = 1$, and $T(n) = T(\lfloor \frac{n}{2} \rfloor) + T(\lceil \frac{n}{2} \rceil) + O(1)$. Compute the complexity of the algorithm. (You can ommit the derivation.)

3. Consider the problem of sorting n real numbers. Choose the correct option among the following choices for the class this problem belongs to.

   (A) NP (B) NP-Complete

   (C) NP-hard (D) None of these above choices.

Group B

Answer all questions in this group. There is no negetive marks in this group.

4. You are given a connected, undirected graph $G(V, E)$, where $|V| = n$ and $|E| = m$. Each edge of G is labeled *red* or *blue*. You are also given an integer k, where $1 \leq k \leq n - 1$. You need to build a spanning tree for G, where at least k edges of your spanning tree are red edges. Write an efficient algorithm that reports such a spanning tree if it exists.

   Modify your algorithm or otherwise write a separate efficient algorithm that reports a spanning tree for G where exactly k edges are red (and hence exactly $n - k - 1$ edges are blue) if such spanning tree exists. 15+10=25

5. The *3-Coloring problem* is to recognize whether it is possible, to assign a color to each vertex using a set of 3 colors such that no two adjacent vertices are of same color.

Highly intelligent aliens from another world come to Earth and tell us that the 3-Coloring problem (which is NP-complete) is solvable in $O(n^9)$ time. They also tell us that there is no algorithm for 3-Coloring that runs faster than $\Omega(n^7)$ time in the worst case. (Here n denote the number of nodes in the graph.) For each of the following assertions, indicate whether it follows from the information the aliens have given us. Also, provide a short explanation in each case.

(i) All NP-complete problems are solvable in polynomial time.

(ii) All problems in NP, even those that are not NP-complete, are solvable in polynomial time.

(iii) All NP-hard problems are solvable in polynomial time.

(iv) All NP-complete problems are solvable in$O(n^9)$ time.

(v) No NP-complete problem can be solved faster than $\Omega(n^7)$ time in the worst case.

(vi) There is no algorithm for the 4-Coloring problem that runs in time faster than $\Omega(n^7)$ time.

[3*6=18]

6. Consider the problem GEOMBASE as follows.

GEOMBASE: Given a set of n points with integer coordinates on three horizontal lines $y = 0$, $y = 1$, and $y = 2$, determine whether there exists a non-horizontal line containing three of the points.

Prove that GEOMBASE is 3-SUM Hard problem. [18]

7. Prove that for any $\epsilon > 0$, there is no $(3/2 - \epsilon)$ approximation algorithm for Binpacking, unless P = NP. [15]

8. You are given a graph $G(V, E)$ with $|V| = n$, and a parameter k. Present an algorithm that determines whether there exists a vertex cover for graph G of size less than or equal to k in $O(2^k.kn)$ time. [12]

9. For the following problem, either give a short proof of the correctness of your claim (if true) or give a counterexample (if false).

Consider a weighted undirected graph G. Suppose you replace the weight of every edge with its negation (e.g. $w(u; v)$ becomes $-w(u; v)$), and compute the minimum spanning tree of the resulting graph using Kruskals algorithm. State and argue whether the statement *"the resulting tree is a maximum cost spanning tree for the original graph"* is true or false.

[10]

10. You are given a set of n real numbers $x_1, x_2, \ldots x_n$. Your task is to cover these points by intervals of unit length (that is, by intervals of the form $[a, a+1] = x \in \Re | a \leq x \leq a + 1$ for real numbers a). Your goal is to minimize the number of intervals in the cover.

(a) Consider the following greedy strategy. Choose an interval of unit length to cover the maximum number of points in the given collection. Output this interval, and remove the points covered by this interval from the collection. Repeat until no points are left. Give an example to demonstrate that this greedy algorithm may fail to provide an optimal solution.

(b)Although the greedy strategy of Part (a) fails, there exist other greedy strategies that efficiently compute optimal solutions. Describe such a strategy. The running time of your greedy algorithm must be bounded by a polynomial in n. [5+5=10]

INDIAN STATISTICAL INSTITUTE

Semestral Examination: 2014-15

Course Name: M. Tech. I Year

Subject Name: Database Management Systems

Date: 05/05/2015 **Maximum Marks: 100** **Duration: 3hours**

Note: Answer all questions

1. Define all the variations of the join operation. Express them in terms of cross product, selection and projection, if possible. Why is the join operation given special attention for query optimization?

10

2. a) Give a set of FDs for the relation schema R(A, B, C, D) with primary key {A, B} under which R is in 2NF but not in 3NF.
   b) Consider the following set of FDs:
   - $A \rightarrow BC$
   - $B \rightarrow C$
   - $A \rightarrow B$
   - $AB \rightarrow C$
   - $AC \rightarrow D$

   Find a irreducible set.

   c) Proof or disproof the following claim:

   "Any two-attribute relation is in BCNF"

20

3. Design a generalization – specialization hierarchy for a motor vehicle sales company. The company sells motorcycles, autos, passenger cars, buses and trucks. Justify your placement of attributes at each level of the hierarchy. Explain why they should not be placed higher or lower levels.

10

4. Let transactions T1, T2, and T3 be defined to perform the following operations on A which is some numeric item in the database:
   - T1: Add one to A
   - T2: Double A
   - T3: Display A on Screen and then set A to one

   a) Suppose transactions T1, T2, T3 are allowed to execute concurrently. If A has initial value zero, how many possible correct results are there? Enumerate them.
   b) Assume the transactions execute without any locking, how many possible schedules will be there? (Write the pseudocode for T1,T2 and T3).

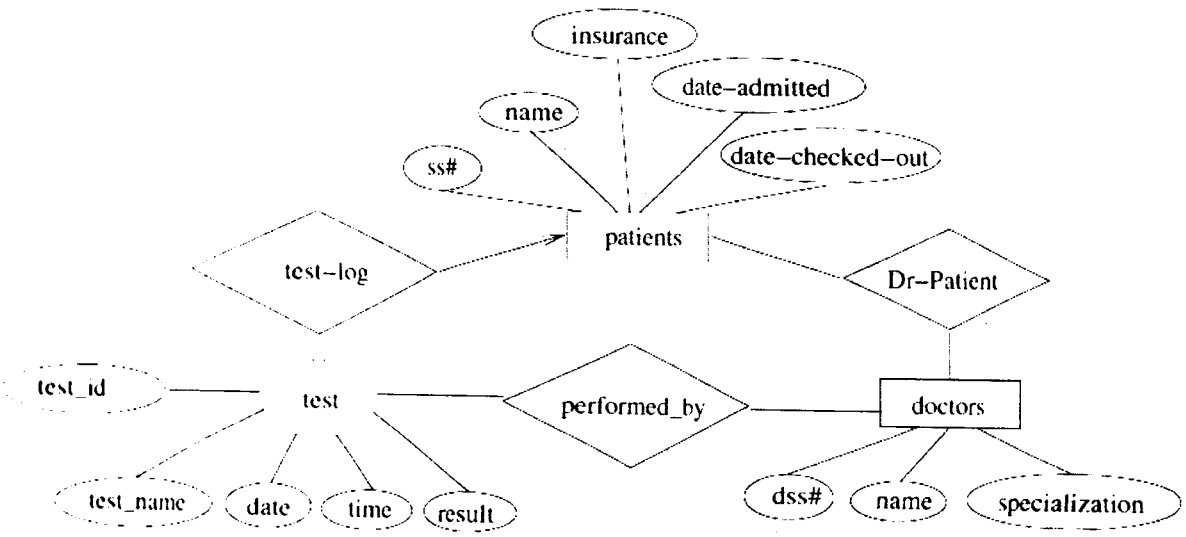
10

P.T.O

5. a) What will be the occupancy of each leaf node of a B^+ tree, if index entries are inserted in sorted order? Explain your answer.
b) Suppose you have a relation r with n_r tuples on which a secondary B^+ tree is to be constructed.
   a) Give a formula for the cost of building the B^+ tree index by inserting one record at a time. Assume each block will hold an average of f entries, and that all levels of the tree are in memory.
   b) Assuming a random access disk takes 10 milliseconds, what is the cost of index construction on a relation with 1,000,00 records.

20

6. Consider the following sample ER Diagram for a hospital:



   a) Write the problem statement that will lead the above diagram.
   b) Map the diagram to appropriate tables.

20

7. Consider the following relations:
Suppliers(supplier_no: int, supplier_name: char(50), status: int, city: char(40))
Parts(part_no: int, part_name: char(100), color: char(10), weight: real)
Projetes(project_no: int, project_name: char(20), city: char(40))
Shipment(supplier_no: int, part_no: int, project_no: int, quantity: int)
Give SQL formulation for the following updates:
   a) Delete all the projects for which there are no shipments.
   b) Change the colour of "red" coloured "nut" weighing 12.5 to "blue"
   c) Insert a new supplier "New Sup" from "Kolkata". His supplier_no will be 1 more than the existing maximum supplier_no, but status is yet known.

Give a relational algebra or tuple / domain relational calculus expression for the following queries:
   d) Find the maximum value of part_no
   e) All parts with same name, color and weight.
   f) All the projects for which there are no shipments.

10