INDIAN STATISTICAL INSTITUTE

Second Semestral Examination: 2012-2013 (Back paper)

M. Tech. (CS) II year

Data Mining and Knowledge Discovery

Date: 14.08.13 Maximum Marks: 80 Duration: 3 hours

1. (i) What are the limitations of DBSCAN?

   (ii) Describe algorithm OPTICS. [6+8=14]

2. (i) Explain the role of Kohonen's self-organization in data mining.

   (ii) How is genetic algorithm used in multilayer perceptron learning?

   (iii) Describe the backpropagation algorithm. [5+5+10=20]

3. (i) How does RAINFOREST work?

   (ii) Elaborate on some performance measures for evaluating classifiers.

   (iii) What is ROC? [6+8+4=18]

4. (i) Describe the apriori algorithm for association rule mining.

   (ii) How does FP tree help in rule mining? [8+4=12]

5. (i) Explain the difference between feature extraction and feature selection procedures. What are the basic steps of a feature selection method? How can principal component analysis (PCA) be used as a dimension reduction technique?
   (ii) What is the difference between filter and wrapper methods in the context of feature selection? Describe a filter method in detail.
   (iii) Describe a feature clustering algorithm. [7+5+4=16]

INDIAN STATISTICAL INSTITUTE

Mid-Semester Examination: 2013-2014

M. Tech. (CS) 2nd Year

Artificial Intelligence

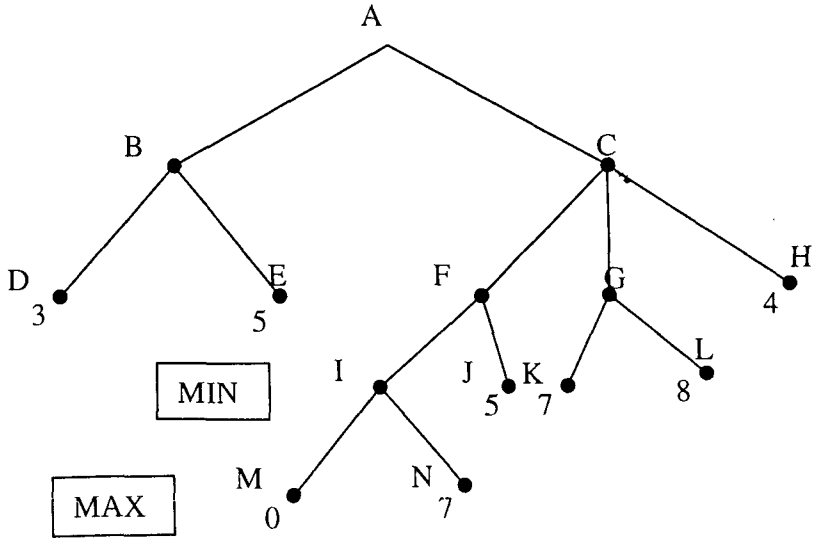
Date: 09.09.2013 Maximum Marks: 60 Duration: 2 hours

Answer all questions in brief.

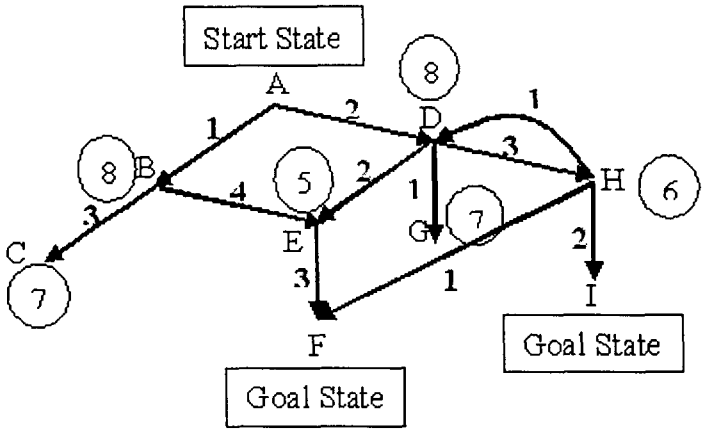
1. Describe the *depth-first iterative deepening* algorithm. Prove that it is asymptotically optimal among brute-force tree searches in terms of time, space, and length of solution. [4 + (3 + 2 + 1) = 10]

2. Prove that any monotonic heuristic is *admissible*. Describe the simulated annealing approach. [4 + 6 = 10]

3. Perform the *minimax* search procedure on the game tree shown below in which static scores are all from the first player's point of view and MAX is allowed to move first. Perform the left-to-right and right-to-left α-β pruning procedure on this tree and show how many nodes can be pruned. Discuss why a different pruning occurs. [3 + 3 + 3 + 1 = 10]



4. Execute the *uniform cost search* and *best first search* algorithms on the following search graph, and show the solution path, along with its cost and list the expanded nodes for each case (each node of the graph is represented by a letter and the encircled value is the heuristic evaluation of the corresponding node, while the bolded numerical value represents the actual length of the path between two nodes). [5 + 5 = 10]



5. Consider a sliding block puzzle with the following initial configuration:

W	W	W	B	B	B	E

There are three white tiles (W), three black tiles (B), and an empty cell (E). The puzzle has the following moves:

(i) A tile may move into an adjacent empty cell with unit cost.
(ii) A tile may hop over at most two other tiles into an empty cell with a cost equal to the number of tiles hopped over.

The goal of the puzzle is to have all the black tiles to the left of all the white tiles without regard for the position of the empty cell. Define the problem as a state space graph problem and find a sequence of moves that will transform the initial configuration to a goal configuration. What is the cost of the solution? [4 + 5 + 1 = 10]

6. Solve the following cryptarithmetic problem:

```
  S E N D
+ M O R E
-----------
M O N E Y
-----------
```

[10]

Indian Statistical Institute

Periodical Examination of First Semester (2013-2014)
M. TECH. (CS) 2 Year
Subject: Information and Coding Theory
Date: 09/09/2013 Time: 2 hours Maximum Marks: 40

1. For a pair of random variables (X, Y), prove that $H(X,Y) = H(X) + H(Y|X)$. [5]

2. Prove that for any instantaneous code over an alphabet of size D, the codeword lengths $l_1, l_2, \ldots, l_m$ must satisfy the inequality
$$\sum_i D^{-l_i} \leq 1.$$
Conversely, given a set of codeword lengths that satisfy this inequlity, there exists an instantaneous code with these word lengths. [10]

3. State the channel capacity theorem and give an outline of the proof. [10]

4. Let $\mathcal{X} = \{1, 2, 3\}$ and $X_1, X_2, X_3 \ldots$ be a time invariant Markov chain where X_is take values from the set $\mathcal{X}$. The transition matrix is
$$P = \begin{bmatrix} 2/3 & 1/6 & 1/6 \\ 1/2 & 0 & 1/2 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}.$$
Find the stationary distribution of the Markov chain. [5]

5. Let (X, Y) have the following joint distribution:

Y \ X	1	2	3	4
1	1/8	1/16	1/32	1/32
2	1/16	1/8	1/32	1/32
3	1/16	1/16	1/16	1/16
4	1/4	0	0	0

Find $H(X), H(Y), H(X|Y), H(Y|X)$ and $I(X;Y)$. [10]

Indian Statistical Institute

Mid-Semester Examination 2013-14
M. TECH. (CS) II Year
Subject: Document Processing and Retrieval
Full Marks: 50 Duration: 2 hrs.

(Answer all questions) 10·09·13

1. (a) For an English font define Baseline, x-height, Ascender, Kerning, Color, weight and typographic contrast. [7]
   (b) Find an expression for cubic Bezier Polynomial. Suppose two consecutive control points have the same (x,y) co-ordinate values. What will be the difference in the shape of the cubic Bezier curve? How to draw a loop by a cubic Bezier function? [8+2+2=12]

2. Generate a feature based tree classifier to recognize the following alphanumeric characters.
   (y, v, 6, 9, W, M, B, D, O, f, t). [10]

3. Discuss about different errors that may occur in an optical character recognition system. [7]

4. Describe a suitable technique to segment characters from arbitrarily oriented words of Devnagari/Bangla script. [7]

5. Describe a Hough Transform based method to detect the skew of a document image. [7]

Indian Statistical Institute
Mid-Semestral Examination
Course: Master of Technology (Computer Science)
Subject: Computer Graphics
September 10, 2013
Maximum Marks - 50
Answer all questions.

1. A 2D point (x,y) is transformed to a point (a,b) using the following equation:

$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} t_{31} \\ t_{32} \end{bmatrix}$$

   (a) What is this transformation known as? (b) What is the effect of the transformation due to the matrix $\begin{bmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{bmatrix}$? (c) What is the effect of the vector $\begin{bmatrix} t_{31} \\ t_{32} \end{bmatrix}$? (d) A square shaped object is transformed to an image of trapezium. How the above transformation should be changed to account for this square to trapezium transformation?

   [1+4+2+3=10]

2. Given a rectangular bounding box and a point in 2D, (a) write a pseudo-code to test whether the point is within the bounding box or not. Assume an arbitrary straight line is passing through the point. (b) Test whether the line is within the box and find out the portion of the line within the bounding box. (c) Given a point on a circle in a discrete grid, how can you decide the coordinates of the next point on the circle? Assume clock-wise direction of trace of points on the circle.

   [2+3+5=10]

3. Given coordinates of points **A**(0,2) and **B**(6,4) find out all the intermediate points lying on the digital line **AB**. Follow Bresenham's scan conversion principle to draw the digital line **AB**.

   [10]

4. (a) How is it possible to approximate an arbitrary surface using a polygon (planar) mesh? (b) Write the algorithmic steps for Scan-line based Polygon Filling algorithm.

   [5+5=10]

5. Write True or False

a) Voxels can have 6 neighbourhood connectivity in a digital 3D grid.
b) RGB image can be transformed to HSV format.
c) Aliasing is a special visual effect that improves quality of rendered image.
d) Midpoint subdivision algorithm can be used for line clipping.
e) Two-dimensional projective object to image transformation has 6 degrees of freedom.
f) Orthographic parallel projection foreshortens image of an object due to viewpoint.
g) Homogeneous transformation takes n-dimensional Cartesian coordinate to $(n+1)$-dimensional homogeneous space.
h) Complexity of seed fill algorithm is independent of the size of the polygon to be filled.
i) Distances between the viewer and 3D planar patches determine visibility of 2D planar patches with respect to the viewer.
j) Vanishing point is a 2D point where parallel lines meet under orthographic projection.

[10]

INDIAN STATISTICAL INSTITUTE
M.Tech. (Computer Science)
Second year, First Semester
Computer Architecture
Mid-Sem Examination, 2013-14

10.09.13

Full Marks : 60 Time : 3 Hours

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

1. Consider a user program which consists of 10 Billion machine language instructions. If this program is executed on a computer which has an average cycle per instruction (CPI) of 1.2 and a 1 GHz clock, then calculate the time needed to execute this program.

   Consider a user program which is compiled on a *Computer A*, generating an object code having 10 Billion instructions with a mix of 40% ALU instructions, 30% load-store instructions and 30% branch instructions. The same user program, when compiled on another *Computer B* (with a different architecture), was translated to an object code in which the number of ALU instructions was reduced by a factor of two, while the number of all other instructions remained the same as that in case of *Computer A* . Assume that on both these computers, the CPI values for ALU instructions, load-store instructions and branch instructions are 1, 3 and 2, respectively. Now compute the following :

   i) MIPS value for *Computer A*, ii) MIPS value for *Computer B*, iii) execution time of the program on *Computer A* and iv) execution time of the program on *Computer B*.

   What conclusion can you draw from these computed values? **(2 + 3 + 4 + 2 + 2 + 2 = 15)**

2. a) Consider an instruction pipeline with five phases as *instruction fetch, instruction decode/resister fetch, execute/address calculation, memory access* and *write back*, respectively. Referring to this pipeline structure , explain the terms 'Structural hazards' and 'Control hazards', if any, in executing the instructions on this architecture. **(3+2 = 5)**

   Is it possible to eliminate structural hazards altogether in the above pipeline structure? If so, then how? If not, then what are those cases. **(3)**

   b) Consider the execution of the following machine language instructions (meaning of each instruction being as specified in the code segment) on the above pipelined architecture :

   ```
   ADD   R1, R2, R3 ;    R1 ← R2 + R3
   SUB   R4, R1, R5 ;    R4 ← R1 – R5
   MUL   R6, R7, R1 ;    R6 ← R7 * R1
   OR    R8, R1, R9 ;    R8 ← R1 OR R9
   DIV   R10, R11, R1;   R10 ← R11 / R1
   ```

   Indicate the possible data hazards that may be encountered in executing the above code section. Briefly mention any possible means to eliminate such data hazards. **(4 + 3 = 7)**

3. a) Consider the following assembly language program segment for a loop executed on an instruction-pipelined computer as described in Q. 2 above, where LD stands for a 'load double word' (8 bytes) instruction, ADDD stands for addition of two double word floating point numbers, SD stands for a double word store instruction, SUBI stands for subtracting an immediate operand from a register and BNEZ stands for 'branch on not equal to zero'. Assume that each data element is stored in the memory as a double word.

```
LOOP :   LD      F0,0(R1)
         ADDD    F4,F0,F2
         SD      0(R1),F4
         SUBI    R1,R1,#8
         BNEZ    LOOP
```

Assume that the LD instruction needs 2 cycles and the ADDD instruction needs 3 cycles to complete. Also, a stall of one cycle is needed after the SUBI instruction as the BNEZ instruction computes the branch address and the condition for branching in the *instruction decode/resister fetch* phase of the pipeline. Now answer the following :

How many cycles per data element are needed for the above program segment and how can it be reduced by only rescheduling the instructions? Give at least one specific example for such a rescheduled code segment with the corresponding number of cycles needed per data element.

Can it be ever possible to reduce the number of cycles needed per data element to i) exactly 3, ii) a value more than 3 but less than 4? Justify your answer with proper illustrative examples and appropriate reasoning.

(5+1+1 = 7)

b) Show how control hazards can be eliminated by rescheduling of instructions from different parts of the user program. What are the limitations of this approach? **(5 + 3 = 8)**

4. 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? **(3)**

b) Consider a computer having 16 KB instruction cache with a miss rate of 0.60% and another 16 KB data cache with a miss rate of 6.40%. Assuming that 75% of the memory accesses are for fetching the instructions, calculate the effective cache miss rate for this system.
If, instead of such a split cache, one unified cache of size 32 KB were considered for both instructions and data, then the cache miss rate were observed to be only 2.00%. Assuming that one cycle is spent on resolving the structural hazard on this unified cache (with instruction being fetched in first cycle and data being accessed in the second cycle), calculate the overall effective memory access time in terms of machine cycles for both the split cache system and the unified cache system. **(2+4 = 6)**

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

i) 250 ii) 0 iii) $(-3)^{1/2}$ iv) ∞ v) 1.5×2^{-63} vi) -18.75×2^{-135}

(6 x 1 = 6)

INDIAN STATISTICAL INSTITUTE
Mid-Semester Examination: 2013
Course Name: M.Tech. In Computer Science
Subject Name: Software Design and Validation

Date: 11.09.2013 *Maximum Marks: 60* *Duration: 2 hours*

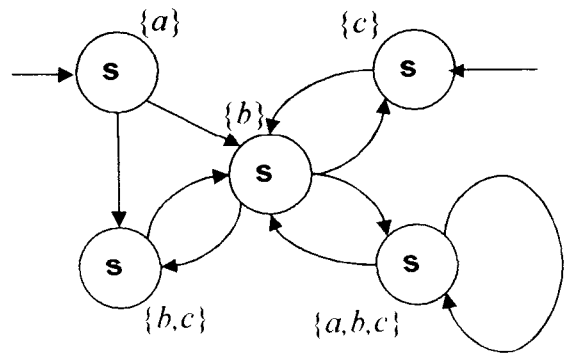
Answer all questions

---

1. For each of the following statements, indicate whether the statement is true or false. For the ones that are true, provide a brief justification. For the ones that are false, provide a small counter-example. [4 X 5 = 20]
   i) The LTL property FG **p** is equivalent to the property AFAG **p** in CTL, where **p** is an atomic proposition.
   ii) Any CNF formula in which no clause has more than two literals is solvable in polynomial time.
   iii) ω-Regular Languages are closed under intersection.
   iv) For 3 MSCs M1, M2 and M3, (M1 × M2) × M3 = M1 × (M2 × M3), where × is the synchronous concatenation operator.
   v) The forward slice with respect to a slicing criterion is always smaller than the backward slice with respect to the same slicing criterion.

2. [Temporal Logic and Model Checking] [2 X 3 +14 = 20]
   i) Indicate whether the following statements are true for LTL with brief justification:
      - If P1 and P2 are liveness properties, then P1 ∩ P2 is also a liveness property
      - If P1 and P2 are safety properties, then P1 ∩ P2 is also a safety property
      - The two properties GF ($\Psi_1 \wedge$ F Ψ_2) and GF ($\Psi_2 \wedge$ F Ψ_1) are equivalent.

   ii) Consider the following transition system, TS over the set of atomic propositions AP ={ a, b, c }



Decide for each of the LTL formulae φ_i below, whether TS $\models \varphi_i$ holds. If TS does not satisfy φ_i then provide a path $\pi \in$ Paths(TS) such that π does not satisfy φ_i

$\varphi_1 = \mathsf{F}\mathsf{G}\, c$ $\qquad$ $\varphi_2 = \mathsf{G}\mathsf{F}\, c$

$\varphi_3 = \mathsf{X}\neg c \Rightarrow \mathsf{X}\mathsf{X}\, c$ $\qquad$ $\varphi_4 = \mathsf{G}\, a$

$\varphi_5 = a\ \mathsf{U}\ \mathsf{G}(b \vee c)$ $\qquad$ $\varphi_6 = (\mathsf{X}\mathsf{X}\, b)\ \mathsf{U}\ (b \vee c)$

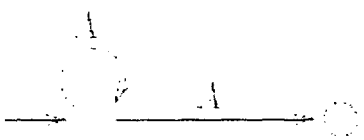
3. [Preliminaries]

i) [Satisfiability]

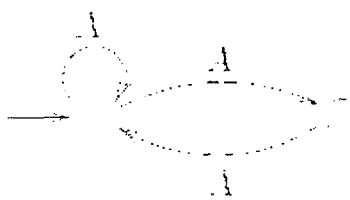
Show a formulation in propositional logic of the following problem: given a directed graph, does it contain a Hamiltonian cycle? [**Note**: A Hamiltonian cycle in a graph is a closed path that visits each node, other than the first, exactly once].

[10 marks]

ii) [Automata on infinite words]



A_1



A_2

Assuming that these are NFAs, indicate the languages accepted by them:

$\mathcal{L}(A_1)$ =
$\mathcal{L}(A_2)$ =

Assuming that these are Nondeterministic Buchi Automata (NBAs), indicate the languages accepted by them:

$\mathcal{L}_\omega(A_1)$=
$\mathcal{L}_\omega(A_2)$=

[4+6 = 10 marks]

INDIAN STATISTICAL INSTITUTE

Mid-Semester Examination: (2013-2014)
M.Tech. (CS) I I Year
Information Security & Assurance

Date : 12.09.2013 **Maximum Marks : 40** **Duration : 2 Hrs.**

1. Following the Bell-LaPadula's model a security lattice has been built with two security labels Graduate > Undergraduate.
   Indicate whether the following requests should be permitted or denied, where $\lambda(s)$ and $\lambda(o)$ represents the security label of the subject and object, respectively,
   a) s requests to write o $\lambda(s)=(UG), \lambda(o)=(UG)$
   b) s requests to read o1 and o2 $\lambda(s)=(G), \lambda(o1)=(UG), \lambda(o2)=(G)$
   c) s requests to write o $\lambda(s)=(G), \lambda(o)=(G)$
   d) s requests to read o1, o2, o3 $\lambda(s)=(UG), \lambda(o1)=(UG)$ $\lambda(o2)=(UG), \lambda(o3)=(UG)$
   e) s requests to write o1, o2 $\lambda(s)=(UG), \lambda(o1)=(G), \lambda(o2)=(UG)$

   (5x2=10)

2. Let two servers S1 and S2 has two users U1 and U2. Execution of an application A1 in S1 uses a dataset D1 in S2 and execution of an application A2 in S2 uses a dataset D2 in S1. However, execution of A1 updates D2 in S1 and execution of A2 updates D1 in S2. Both the users U1 and U2 have relevant access permissions to S1, S2 and necessary authorizations for executing applications A1 and A2. However, only user U1 has write permission to both D1 and D2. U2 can only read D1 and D2.
   From the above description, write down the low level policies in the form (subject, object, access right, sign). Use standard boolean operators to create composite policies to satisfy the constraints given in the problem. If needed, Inclusion, Connection, Assignment and Reachability operators may also be used.

   (20)

3. Describe when $P(o_x)$ signifying a policy related to object o_x will be considered vulnerable for a system.

   (10)

INDIAN STATISTICAL INSTITUTE

Periodical Examination: (2013)

M.Tech. (CS) II Year

Parallel Processing: Architectures and Algorithms

Date: Sept. 13, 2013 Maximum Marks: 100 Duration: 3 hours

NOTE: A student may answer all questions but maximum marks attainable is 100.

1. Indicate whether each of the following statements is true or false and justify your answer:
a) In Parallel Random Access Machine (PRAM), all processors must execute the same instruction synchronously.
b) An *n*-processor EREW PRAM can be no more than *O(log n)* times slower than a CRCW PRAM.
c) The AT^2 bound defined by the VLSI complexity model plays an important role in the design of parallel processing system.
d) In parallel processing, the computational granularity is independent of the interprocessor communication latency.

[4 × 5 =20]

2. Consider the following sequential program segment with five statements, S_1 through S_5:

```
S1 :  A = B + C
S2 :  C = B × D
S3 :  SUM = 0
S4 :  Do I= A, A+7
          SUM = SUM + X(I)
      End Do
S5 :  IF (SUM .GT. 1000)  C = C + SUM
```

(a) Restructure the program in order to achieve maximum parallelism.
(b) Assuming that addition and multiplication takes 20 cycles and 50 cycles respectively show a possible scheduling of the program. Use as many processors as you need. Calculate the speed-up and utilization, ignoring the overhead due to interprocessor communication.

[6+(7+3) =16]

3. (a) Define a Cayley graph. Prove that a Cayley graph is always vertex symmetric.
(b) Define a pancake graph. Show that the upper bound on the diameter of an n-pancake graph is (2n-3). In a 4-pancake graph find the routing path (sequence of generators) from the node (3214) to the node (4312).

[(3+4)+(3+5+3) =18]

4. Answer in brief:
(a) Assume that each node of a Cube-Connected-Cycle (CCC) of order *n* has one datum. Write an algorithm to find the *maximum*. How many steps of communication and computation are required?
(b) Prove that an *(n-1)*-node complete binary tree is not a subgraph of an *n*-node hypercube.
(c) In a regular graph with *N* nodes each of degree *d*, derive an expression for the lower bound on the diameter *k*.

[3 × 7 = 21]

P.T.O

4. (a) Show the diagrams of a 8×8 Baseline network and a 8×8 3-stage Clos network with minimum number of middle-stage switches. Compare these two networks in terms of number of conflict-free permutations and the cost. Assume that the cost of an $n \times n$ switch is $O(n^2)$.
(b) Show the routing of the identity permutation P on a 8×8 Baseline network. Is it realizable in single pass? If not, show the conflict graph, and find the minimum number of passes for routing. Using path matrix, check if P is conflict-free on a 8×8 Omega network.

[(3+3+6) +(5+1+ 4+3)=25]

5. Given the following CUDA kernel with some errors,

```
__global__ void  test( int *A, int *B, int *C )
{
        __shared__ int temp[THREADS_PER_BLOCK];
        temp[threadIdx.x] = B[blockIdx.x] + A[blockIdx.x];

                if( 0 == threadIdx.x ) {
                        int sum = 0;
                        for( int i = 0; i < THREADS_PER_BLOCK; i++ )
                                sum += temp[i];
                        atomicAdd( C, sum ); // to avoid race condition
                }
}
```

identify a function the kernel can compute with slight modifications. Rewrite the code with necessary changes to compute the function correctly.

[4 + 6 =10]

INDIAN STATISTICAL INSTITUTE

Periodical Examination

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

Advanced Algorithms for Graph and Combinatorial Optimization Problems

Date: 13.9.2013 Maximum Marks: 60 Duration: 2.5 Hours

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

1.(a) For any simple planar graph G with n vertices and e edges, prove the followings:

(i) There is a vertex of G that has degree at most 5.

(ii) If G does not contain a cycle of length 3 then $e \leq 2n - 4$.

(b) Prove that any undirected planar graph $G = (V, E)$ with non-negative edge weights can be transformed into an undirected planar graph $G' = (V', E')$ with maximum node-degree 3 such that,

- for any $u, v \in V$, $\Delta(u, v) = \Delta(f(u), f(v))$, where $f : V \rightarrow V'$ maps vertices between G and G' and $\Delta(u, v)$ denotes the length of the shortest path between nodes u and v in the respective graph; and
- $|V'| = O(|V|)$.

[(5+3)+10=18]

2.(a) Let $T = (V, E)$ be a connected undirected tree such that each vertex has degree at most 3. Let $n = |V|$. Show that T has an edge whose removal disconnects T into two disjoint subtrees with no more than $(2n + 1)/3$ vertices each. Give a linear time algorithm to find such an edge; prove its correctness.

(b) Use the above algorithm or otherwise provide an algorithm running in $O(n \log k)$ time to partition the binary tree with n vertices into $k (k \leq n)$ subtrees, so that each of the subtree is of size at most $(2/3)^k n$.

[10+8=18]

3. Let $G = (V, E)$ be a comparability graph, and a transitive orientation of G is given. Define layering of the comparability graph as follows:

The vertex v with indegree 0 is assigned in layer $\ell(v) = 1$. A vertex $u \neq v$ is assigned in layer $\ell(u) = 1 + max_{w \in adj(u)} \ell(w)$, where $adj(u)$ is the set of predecessors of u

(a) If m layers are needed for layering all the vertices of the graph, then what is the size of the maximum clique of the graph G ?

(b) What is the size of the maximum independent set of the graph G ?

(c) How will you compute all maximal cliques of the graph G ? [4+6+5=15]

4.(a) Define *simplicial vertex* and *perfect elimination order*.

(b) Show that every triangulated graph $G = (V, E)$ has a simplicial vertex. Also show that, if G is not complete, then it has at least two non-adjacent simplicial vertices.

(c) Show that the perfect elimination order of the vertices of a triangulated graph can be found in time linear in the number of edges of the graph. [4+7+5=17]

Indian Statistical Institute

Mid-Semestral Examination : 2013 – 2014

Master of Technology in Computer Science, Semester III

Functional Brain Signal Processing: EEG & fMRI

Date: 14 September 2013 Maximum Marks: 50 Duration: 2 hours

Attempt all the questions. Credit will be given for precise and brief answers.

1. Write short notes on each of the following EEG bands: (1) delta, (2) theta, (3) alpha, (4) beta and (5) gamma. 5 x 3 = 15

2. You are given 10 human scalp EEG signals recorded from different locations on the head of the same subject during a particular task execution pertaining to a single trial and you are asked to determine the ensemble cross-correlation across all the 10 signals. Describe an algorithm to accomplish this measure. You can describe the entire process in plain English (mathematical equations and derivations are not required) in numbered steps. 10

3. Define *Hilbert transformation*. Show that Hilbert transformation of $\sin(t)$ is $-\cos(t)$. Hint: $\int_{-\infty}^{\infty} \frac{\sin(x)}{x} dx = \pi$ and $\int_{-\infty}^{\infty} \frac{\cos(x)}{x} dx = 0$. Describe a wavelet based phase synchronization measure between a pair of EEG channels. Precisely, but concisely describe in numbered steps. 2 + 4 + 4 = 10

4. Consider the following configuration in a two dimensional space. If a neural network is designed to identify this configuration at the least how many hidden layers the neural network must have and why? 5

5. a) In Fisher's linear discriminant the expression $J(\mathbf{w}) = \frac{\mathbf{w}^T \mathbf{S}_B \mathbf{w}}{\mathbf{w}^T \mathbf{S}_w \mathbf{w}}$ has to be maximized, where $\mathbf{w}$ is a vector, $\mathbf{S}_B$ and $\mathbf{S}_w$ are nonsingular, square matrices of appropriate dimension. Show that when $J(\mathbf{w})$ is maximum, $\mathbf{S}_B \mathbf{w} = \lambda \mathbf{S}_w \mathbf{w}$ will have to hold for some

scalar λ. Also show that for the optimally discriminating hyperplane $\mathbf{w}^T\mathbf{x} = c$, $\mathbf{w}$ is given by $\mathbf{w} = \mathbf{S}_w^{-1}(\mathbf{m}_1 - \mathbf{m}_2)$ (assume that $\mathbf{S}_B$ is in the direction of $\mathbf{m}_1 - \mathbf{m}_2$, only the direction of $\mathbf{w}$ matters not the magnitude). $\mathbf{m}_1$ and $\mathbf{m}_2$ are mean of the two data sets respectively which will have to be optimally discriminated (separated) from each other by a hyperplane. 2.5 + 2.5 = 5

b) Write a short note on logistic regression and mention why logistic regression is more efficient than Fisher's discriminant. 5

INDIAN STATISTICAL INSTITUTE
M. Tech (CS) II year : 2013-2014
Cryptology
Periodical Examination

Date: 14. 09. 2013 Maximum Marks : 100 Time : 3 Hours

Answer any part of any question. Maximum marks you can obtain is 100.
Please answer all parts of a question at the same place.

1. Consider the Geffe generator where the output of three LFSRs are combined by a multiplexer to generate the key stream and then message bits are XORed with the key stream bits to generate the cipher bits. Clearly explain a ciphertext only attack on this scheme which requires lesser time complexity than the exhaustive search. 30

2. (a) Write a program in C language to calculate the GCD of two integers.
   (b) Write the extended Euclidean algorithm to find the inverse of an integer a modulo an integer n.
   (c) Explain each step of the above algorithm to find out the inverses of 23 and 35 modulo 119.
   (d) Prove that if one can factorize easily then the RSA public key algorithm is broken.

   5+10+5+10 = 30

3. (a) Using an efficient method, calculate the nonlinearity of the Boolean function $f(x_1, x_2, x_3, x_4) = x_1x_2 \oplus x_3x_4$.
   (b) Consider that four pseudo-random bit generators are available and no weakness about them is known yet. The output bits from each of these are fed to the inputs of the Boolean function f described above. Explain whether the output stream of the Boolean function can be used as a pseudo-random bit stream.

   10+10 = 20

4. (a) What is a Fiestel structure in a block cipher?
   (b) Briefly describe different modes of operations in a block cipher.
   (c) Consider a block cipher that accepts 128-bit plain-text block and 128-bit secret key to produce 128-bit cipher text block. Let $B_0, B_1, \ldots, B_{l-1}$ be the plain text blocks such that all of them are same except one. Discuss the problems in ECB mode compared to CBC mode of operation in such a situation.

   10+10+10 = 30

Indian Statistical Institute
Mid-Semestral Examination: 2013
Course Name: M. Tech. in Computer Science
Subject Name: Mobile Computing

Date: 11-09-2013 Maximum Marks: 60 Duration: 2 hours 30 minutes

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

1. Briefly explain the channel assignment problem (CAP) in cellular network. Represent the CAP in terms of classical vertex coloring problem. Describe a method by which the CAP with non-homogeneous demand can be partitioned into a sequence of smaller subproblems where each subproblem has a homogeneous demand from a subset of nodes of the network. Suppose *Algorithm A* is an algorithm that can solve the CAP with homogeneous demand of single channel from a subset of nodes of the network. Describe an algorithm to solve the CAP with non-homogeneous demand by using the *Algorithm A*. [3+3+5+5=16]

2. State the differences between soft handover and hard handover. What is fade margin? Discuss the basis for which there can be coverage gain due to fade margin in soft handover in UMTS network. Discuss how signals from multiple base stations are combined by the rake receiver using the maximal ratio combining in soft handover in UMTS network? [3+3+3+5=14]

3. Briefly explain the perturbation-minimizing frequency assignment problem. [5]

4. State the differences between horizontal handover and vertical handover. Briefly explain the following RSS based horizontal handover decision strategies: i) relative RSS, ii) relative RSS with threshold, iii) relative RSS with hysteresis, and iv) relative RSS with hysteresis and threshold. What are the main stages of the vertical handover process in $4G$ heterogeneous wireless networks. List out the most commonly used vertical handover decision (VHD) criteria and performance evaluation metrics for VHD algorithms. [3+8+3+8=22]

5. Consider the problem of optimal association of wireless stations (STAs) to access points (APs) in an IEEE 802.11 wireless local area network. Write the formulas for computing the throughput obtained by the STAs from their respective associating APs, for two MAC scheduling policies namely random polling access and proportional fair access. Compare these two MAC scheduling policies in terms of the individual throughput obtained by the STAs.

[3+3+2=8]

Indian Statistical Institute

Semester-I 2013-2014

M.Tech.(CS) - Second Year

Mid-semester Examination (16 September, 2013)

Subject: Compiler Construction

Maximum marks: 40 Total marks: 45 Duration: 2.5 hrs.

Please keep your answers brief and to the point.

1. Consider `findmail`, a hypothetical command-line application that can be used to search through email. It takes some keywords as arguments and produces output in XML format. The following example shows how `findmail` is run, and a portion of the output it produces.

```
$ findmail compiler end-sem
<messages>
    <message>
        <from>Mandar Mitra &lt;mandar@isical.ac.in&gt;</from>
        <to>Dean of Studies &lt;dean@isical.ac.in&gt;, rnm@isical.ac.in</to>
        <subject>marks for compiler construction</subject>
        <date>15 December, 2013</date>
        <size>18360</size>
        <msgid>abcd1234!@#$</msgid>
        <some other unimportant tag>   </some other unimportant tag>
        ...
     </message>
    <message>
    ...
    </message>
    ..
</messages>
```

You have to write a postprocessor using (f)lex that will take `findmail`'s output as input, and generate a summary with one line per message.

- The one-line summary will contain the date (10 characters wide), sender's name (25 characters wide), and the subject (35 characters wide).
- In case a field is shorter than the specified length, you should pad it with blanks; if it is longer, just truncate it to the specified length.
- The fields should be separated by exactly three blank spaces.

For the above example, your postprocessor should produce the following output. Note that < and > are encoded as `<` and `>` respectively in XML.

```
15 Decembe   Mandar Mitra <mandar@isic   marks for compiler construction
```

Note 1: You may make reasonable assumptions where necessary. **Clearly state any assumptions you make.**

Note 2: In Lex, . is a regular expression that matches any single character other than the newline.

Warning: Lex always tries to find the largest possible match for any given pattern.

2. Consider the following grammar (capital letters denote non-terminals, **tb**, **te**, **rb**, **re**, **cb**, **ce** and **d** are terminals).

$S \rightarrow \mathbf{tb}\ R\ \mathbf{te} \quad R \rightarrow RQ \quad Q \quad Q \rightarrow \mathbf{rb}\ C\ \mathbf{re} \quad C \rightarrow CB \quad B \quad B \rightarrow \mathbf{cb}\ D\ \mathbf{ce} \quad D \rightarrow \mathbf{d}$

(a) Eliminate left-recursion from the above grammar.

(b) Compute the *FIRST* set for the right hand side of each production of the modified (non-left-recursive) grammar.

(c) Compute the *FOLLOW* set for each non-terminal of the modified grammar. Show your rough work.

(d) Construct the $LL(1)$ parsing table for the modified grammar.

(e) Construct the canonical collection of $LR(0)$ (i.e. SLR) items for the **original** grammar.

You should get about 18 sets in your answer.

2+3+4+4+9=22

3. Recall that $LL(1)$ parsers read the input from left to right, and parse the input string in a way that corresponds to a left-most derivation. Suppose that we want to develop an $RR(1)$ parser in analogy with $LL(1)$ parsers. We will use the following grammar G as an example (S is a non-terminal; all other symbols (a, $+$, $*$) are terminals).

$$S \rightarrow SS+ \qquad S \rightarrow SS* \qquad S \rightarrow a$$

(a) The $LAST$ set of a grammar symbol is the set of all terminals that can occur in the last position of strings derived from that grammar symbol. Give a formal definition (in set-theoretic notation) for $LAST$.

(b) Give an algorithm for computing the $LAST$ set of a grammar symbol.

(c) Compute the $LAST$ set for S in the example above.

(d) We define the *PRECEDES* set for a non-terminal A as follows.

$$PRECEDES(A) = \{a \mid S \Rightarrow \alpha a A \beta\}$$

Give an algorithm for computing the *PRECEDES* set for non-terminal symbols. You may assume that the special symbol BOF stands for an imaginary token that lies at the beginning of a file.

(e) Compute the *PRECEDES* set for S in the example above.

2+3+1+3+2 = 11

INDIAN STATISTICAL INSTITUTE

Mid-semestral Examination:(2013-2014)

MTech C.S. 2nd Year

Digital Signal Processing

Date: 16.9.2013 Maximum Marks: 60 Duration: 2 hours

Note: The marks add up to 76. The maximum you can score is 60. The exam is open-book, open-notes. Use of calculators is permitted.

1. Consider a continuous-time signal

$$x_c(t) = 3\cos(32\pi t) - 2\cos(24\pi t) - 4\cos(120\pi t) + 8\cos(144\pi t)$$

It is sampled at 40 Hz.

(a) Determine the expression for the corresponding discrete-time signal $x[n]$.

(b) $x[n]$ is now passed through an ideal low-pass filter with cut-off 20 Hz. Determine the reconstructed signal.

[8+8]

2. A continuous-time signal $x_c(t)$ is uniformly sampled at the Nyquist rate for 6 s, yielding 12000 samples.

(a) What is the highest frequency that could be present in $x_c(t)$

(b) Design a system for changing the sampling rate to 3 kHz.

[4+6]

3. A sequence $x[n] = \{3 \quad 1 \quad -5 \quad 1 \quad 0 \quad 8\} \quad -2 \leq n \leq 3$ has a discrete time Fourier Transform (DTFT) $X(e^{j\omega})$ Without actually computing $X(e^{j\omega})$, determine

(a) $\int_{-\pi}^{\pi} |\frac{dX(e^{j\omega})}{d\omega}|^2 \, d\omega$

(b) $Im\{X(e^{j(\omega-\pi/3)})\}$

[6+6]

4. Two LTI systems with impulse responses $n\alpha^n u[n]$ and $u[n]$ are cascaded together. Determine the overall impulse response. [10]

5. Determine the inverse z-transform of: $X(z) = \frac{1}{1-2z^{-5}}$ $|z| > 1$ [5]

6. A causal LTI system has the difference equation:

$$y[n] = 0.4y[n-1] + 0.05y[n-2] + 3x[n]$$

   (a) Determine the transfer function and impulse response of the system.

   (b) Sketch the poles and zeros and state if and why the system can/cannot be stable.

   ~~(c) Determine the output for an input $x[n] = (0.5)^n u[n]$.~~

[10+5]

7. For the system described by $y[n] = x[5-n]$, determine if the system is linear, time-invariant, causal and stable. [10]

INDIAN STATISTICAL INSTITUTE

M. Tech (Computer Science) II year, 2013 – 14

Pattern Recognition and Image Processing

Periodical Examination

Date: 18.09.13 **Maximum marks: 60** **Duration: 2 hours**

Note: Answer all the questions

1. State the Bayes decision rule for three-class classification problem and show that it minimizes the probability of misclassification. [3+10=13]
2. Let there be two classes C_1 and C_2 with prior probabilities P and (1-P), and class conditional probability density functions p_1 and p_2 where

$$p_1(x) = e^{-x};\ 0 < x < \infty,$$
$$= 0 \text{ otherwise,}$$

and $$p_2(x) = 2e^{-2x};\ 0 < x < \infty$$
$$= 0 \text{ otherwise.}$$

(i) Find the Bayes decision rule for the above classification problem and find its probability of misclassification.

(ii) Find the probability of misclassification for the following decision rule and verify that it is not less than the misclassification probability of the Bayes decision rule.

Put x in class 1 if $x \geq 2$. Otherwise put it in class 2.

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

3. (i) State the minimum within cluster distance criterion.
   (ii) Describe the k-means algorithm for clustering.
   (iii) Give an example of a data set and two of its initial partitions for which the resultant clusterings would be different when k-means algorithm is applied.

[5+5+7=17]

4. Write short notes on the following.
   (i) Training and test sets
   (ii) Estimation of parameters of Multivariate normal distribution. [5+5=10]

INDIAN STATISTICAL INSTITUTE

Periodical Examination

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

Multidimensional Searching and Computational Geometry

Date : 20.09.2013 Maximum Marks : 60 Duration : 3 Hours

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

1. Let R be a set of n rectangles in $\mathbb{R}^2$, and $G(R) = (V, E)$ be a planar graph

   - whose vertices are the corner points of all the rectangles and the intersection points of the boundaries of pair of members in R,
   - a pair of vertices $u, v \in V$ have an edge $(u, v) \in E$ if (i) both u and v lie on the boundary of a rectangle in R and (ii) the (horizontal/vertical) line segment $[u, v]$ does not contain any other vertex of V, and
   - a face is the maximal portion of the plane that does not contain the boundary of any rectangle.

   For each face f, $\mu(f)$ denotes the number of rectangles containing the face f. A face f is said to be *maximal* if $\mu(f) > \mu(f')$ for any f' adjacent to f (f and f' share an edge).

   (a) Show that each maximal face is a rectangle.

   (b) Show that the number of maximal faces can be $O(n^2)$ in the worst case.

   (c) Describe an efficient algorithm for reporting all maximal faces of $G(R)$ (an $O(n^2 \log n)$ time algorithm is easy. Try to design an $O(n \log n + k)$ time algorithm using sweep line algorithm and interval tree, where k is the number of maximal faces in $G(R)$). [8+6+12=26]

2. Describe a randomized incremental algorithm to compute the convex hull of a set of n points in $\mathbb{R}^2$. Analyze the expected running time of your algorithm. [8+8=16]

3. Let P be a monotone polygon whose vertices are given in clockwise order in an array. Preprocess it so that given any arbitrary query point p, the testing of whether $p \in P$ or not can be performed in $O(\log n)$ time. [10]

4. A rectilinear polygon is a simple polygon whose edges are horizontal or vertical line segments,

   (a) Let P be a rectilinear polygon with n vertices. Show that $\lceil \frac{n}{4} \rceil$ point guards are sometimes necessary and always sufficient to guard the polygon P.

   (b) Propose an algorithm that computes the position of a set of guards that can cover the entire polygon P and the number of guards reported by your algorithm is not more that $\log n \times OPT$, where OPT is the minimum number of guards required to cover P. (Note that OPT for the polygon P is unknown to the designer of the algorithm). [8+10=18]

5. Let Q be a set of points in $\mathbb{R}^2$, where both the coordinates of each point are positive (points are in the first quadrant of the coordinate system). It is assumed that, no two points in Q lie on (i) the same horizontal line and (ii) the same vertical line. A point $q \in Q$ is said to be dominated by another point $q' \in Q$ if $q_x < q'_x$ and $q_y < q'_y$. Design an algorithm to report the largest subset $Q' \in Q$ such that each member in Q' is not dominated by any member of Q. [10]

INDIAN STATISTICAL INSTITUTE

Mid-Semestral Examination : (2013 - 2014)

Course Name : M.Tech (CS) **Year : 2nd year**

Subject Name : Neural Networks & Applications

Date : September 20, 2013 Maximum Marks : 50 Duration : 2 hrs

---

Answer all the questions.

1. Consider a two-input XOR gate. The inputs for which the outputs of the gate are 1 belong to class C_1, while the others are in class C_2. Show how C_1 and C_2 become linearly separable under a RBF network framework using a suitable radial basis function. [10]

2. Derive an expression for the updated weights in terms of the current weights under backpropagation learning. Why is it called backpropagation? [30]

3. Explain the various issues you need to consider while designing an artificial neural network model. [10]

INDIAN STATISTICAL INSTITUTE

Semestral Examination: (2013 – 2014)

M.Tech. (CS) II Year

Parallel Processing: Architectures and Algorithms

Date: Nov. 25, 2013 Total Marks: 116 Duration: 3 hours

NOTE: Answer as much as you can; maximum marks you can score is 100.

1. a) According to the multiplicity of instruction and data streams, classify the parallel machine architectures with a schematic diagram for each. Mention the basic differences between these architectures and the *CUDA GPU* architecture.
b) Given the choices: i) *2-D Torus*, ii) *binary hypercube*, and iii) *cube-connected cycle*, select the best interconnection network for interconnecting *64* nodes of a multicomputer based on the parameter $(B/(dk))$, B, d and k being the bisection width, the average node degree, and the network diameter respectively.

[(4+4)+8=16]

2. a) Consider the following program segment with seven instructions:

$P_1 : A = B \times C$

$P_2 : D = B + A$

$P_3 : G = B \times E$

$P_4 : E = G + C$

$P_5 : A = D \times A$

$P_6 : F = E \times G$

$P_7 : H = E \times D$

Draw the data dependence graph considering each statement as a process. Show a possible scheduling of the processes exploiting the maximum parallelism existing among the processes. Assuming that addition takes 10 time units, multiplication 100 time units, and interprocessor communication 150 time units, calculate the speed-up achieved by your scheduling.

[6+6+5= 17]

3. a) Consider an N-node hypercube, $(N= 2^n)$, where each node wants to broadcast a datum to every other node (all-to-all broadcast). Write an algorithm to complete the procedure exactly in $(N-1)$ steps. Assume that each link is bi-directional, i.e., can communicate in both directions simultaneously.
b) Given an array $\{a_1, a_2, \ldots\ldots, a_n\}$, write an algorithm for computing the prefix sums $S_i = (a_1 + a_2 + \ldots + a_i)$, for $1 \leq i \leq n$, in $O(\log n)$ time on an EREW SM SIMD machine.

[8+12= 20]

4. Consider an $n \times n$ mesh of processors where the boundary processors, i.e., the processors in row *1*, and column *1* only are capable of handling input-output operations. Design an $O(n)$ algorithm for multiplying two $n \times n$ matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ on this architecture. Justify that this is the fastest possible algorithm for the given architecture.

[12 + 5 = 17]

P.T.O

5. Explain how the Newton's method for solving non-linear equations can be implemented efficiently on a shared memory parallel processor to solve the equation $f(x) = 0$. Assume that the equation $f(x) = 0$ has one and only one root in an interval (a, b). Which of the models SIMD or MIMD will be better and why?

[12+3=15]

6. a) Draw a 5×5 Mesh-of-Tree (*MoT*) interconnection network.
b) Given an input sequence of n samples, describe an algorithm to compute the *Discrete Fourier Transform* (*DFT*) on SIMD computer where processors are interconnected by *MoT* interconnection network. Analyze the time complexity and the cost.

[3+(10+5)=18]

7. a) Draw *Batcher's odd-even merging* network for merging two sorted sequences of lengths 4 and 7 respectively.
b) Describe *Batcher's odd-even merging* algorithm for merging two sorted sequences of lengths m and n, where m and n are any two positive integers.
c) Prove the correctness of the algorithm using *0 -1 principle*.

[5+3+5=13]

INDIAN STATISTICAL INSTITUTE

Semester Examination: (2013-2014)
M.Tech. (CS) I I Year
Information Security & Assurance

Date: 25.11.2013 **Maximum Marks: 50** **Duration: 2 Hrs.**

1. An organization wishes to implement an access control mechanism distributes the documents created within the organization among different security classes (sc) arranged in a partial order. The employees are also given clearances at different security levels. One document is placed in only one security class and an employee is also given clearance to only one security class. So, each employee as well as document is marked with a unique sc-value. Any document created at a lower class is available for reading to any employee at the higher class but he/she cannot alter it. However, any document at the higher class is not available to any employee of the lower class either to read or alter. Any document created by an employee is given the security class of the concerned employee. Documents stored in the document database are categorized with document classes (dc) and each document marked with a unique dc-value. So, documents with the same dc-value may be placed in different security classes. A set of employees marked as DBA (not allowed to create or alter any document) are responsible for maintaining the document database. Each DBA is given clearance to read any document of a particular dc-value across all security classes. From the above description,
   a) Specify the rules of information flow among different security classes.
   b) Specify the rules for each DBA to allow reading documents among all security classes for the assigned dc-value and also to restrict any writing or alteration of documents.
   c) Specify how a DBA will store the documents in the document database so that documents having same dc-value distributed among different security classes can be uniquely identified.

   (10+15+5=30)

2. An organization producing cosmetic products wishes to give online advertisement of the products. Orders can only be placed online. Once an order is placed, it is kept in an Order Database. Against each such order, payment information is also maintained along with the mode of payment (credit/debit card or bank transfer). Each customer wishes to avail of the online order placement facility has to register with the company creating a personal profile that gives credit/debit card no. or bank account no. along with associated bank information. Customer related information are maintained in a Customer database. Following restrictions have been imposed.
   a) Employees of the purchase department can only read order information from the Order database but cannot alter it. They can, however, follow up the order till it is marked as complete by the accounts department.
   b) Employees of the accounts department can generate an invoice and a bill against an order and mark it complete when both supply and corresponding payment are done. However, they cannot change any existing content of an order.
   c) Employees in the IT department receive the online orders, update the Order database and maintain the Customer database. Customer database cannot be read by any employee other than those in the IT department and cannot be altered by anyone in the organization.
   d) A customer creates his/her own profile, can read and alter it. However, a customer can neither read nor alter the profile of any other customer.

   From the above description,
   i) Design a system with a Citizen Portal, two separate database systems for Order and Customer database. Place appropriate firewalls, Webservers etc.
   ii) Generate a set of rules for Purchase related, Accounts related and IT employees so that above restrictions can be imposed.

   (10+10=20)

-x-

INDIAN STATISTICAL INSTITUTE
First Semestral Examination: (2013)
M Tech (Computer Science) – II yr.
Computer Graphics

Date: 25. 11. 2013 Maximum Marks: 100 Duration: 3 Hours

The answers should be presented point-wise and **not** in descriptive style. Clearly specify the input and output in case of an algorithm.

1. Write True OR False (answer all questions) [10x1=10]:
   a) If 2D points are expressed in homogeneous coordinates, applications of translation, rotation and scaling, taken all together, to points can be treated as a multiplication by a matrix.
   b) The cubic curve is the lowest order polynomial that has C^1 and C^2 continuity.
   c) Z-buffer value cannot be linearly interpolated within a surface patch.
   d) Goraud shading uses surface normal interpolation scheme.
   e) Ambient intensity of a surface point depends on the external light source.
   f) Orthographic parallel projection foreshortens image of an object.
   g) Aliasing is a special visual effect that improves quality of rendered image.
   h) Intensity attenuation adds to realistic rendering effect.
   i) BMP is an image format.
   j) RGB image cannot be transformed to CMY image.

Answer any **eight** from the group of questions 2 – 11:

2. (a) Write equation of a plane. (b) Find the normal vector to the plane. (c) If $P(x,y,z)$ is a vertex of the plane, find the perpendicular distance of the point P from the plane. (d) Write parametric equations of a line and a curve. (1+1+1+1+1=5 marks)

3. 2D Cohen-Sutherland algorithm determines the clipped line visible within a 2D rectangular view-plane. Extend this algorithm for clipping a 3D line to be visible within a 3D unit cube view-volume. (5 marks)

4. (a) Suppose a circle is drawn in a discrete plane. The equation of the circle is: $(x-h)^2+(y-k)^2-R^2=0$. State the steps to test whether a point (u,v) lies within the circle. (b) Assume a polygon is drawn in a discrete plane. State the steps to test whether a point (u,v) lies within the polygon (point-in-polygon test). (3+2=5 marks)

5. Write algorithmic steps of ray tracing algorithm for visible surface determination. Assume that the 3D world contains only spherical objects. (5 marks)

6. Given a 3D viewpoint and three non-intersecting (not necessarily parallel) planar patches, Find the closest planar patch with respect to the viewpoint. (5 marks)

7. How a straight line can be clipped using mid-point sub-division algorithm? (5 marks)

8. Assume a Lambertian and Specular planar object is illuminated by a point light source. You need to assign a pixel value for each image point corresponding to each point of the object plane. Design the model to generate the pixel values. (5 marks)

P.T.O

9. Assume pixel values I_1, I_2, I_3 and I_4 are given. We would like to generate a set of 100 pixel values in between pixel values I_2 and I_3. However, the interpolated pixel values between I_2 and I_3 should follow a cubic B-spline function. State algorithmic steps to generate these 100 pixel values. State assumptions if any. (5 marks)

10. Information about (linear) boundaries of a polygon can be stored using an edge table. Design a suitable data structure for this edge table and explain how it can help filling a region. (2+3=5 marks)

11. How distance function can be used to represent a closed curve displayed in a discrete 2D grid? Given the closed curve, how you propose to smooth the contour of the curve in an iterative fashion? (3+2=5 marks)

Answer any **five** from the group of questions 12 – 17:

12. State Bresenham's algorithm for drawing a straight line with slope between 0 and 1. (10 marks)

13. Show that vanishing point under perspective projection is a direction and not a point. (10 marks)

14. We need to display a surface represented by $z=f(x,y)$. State the algorithmic steps to display the surface z after removing the hidden lines. (10 marks)

15. Given the equation of an ellipse as $f(x,y) \equiv b^2x^2 + a^2y^2 - a^2b^2 = 0$, how the ellipse can be drawn in a discrete plane? (10 marks)

16. Design a two-pass Z-buffer algorithm to find the shadow region in a rendered scene consists of planar shapes. (10 marks)

Fig. A

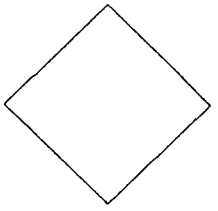


Fig. B

17. The image of Fig. A is to be transformed to fill the face of the cube shown in Fig. B. State steps, transformation and assumption (if any) to achieve this. (10 marks)

Indian Statistical Institute

Semester-1 2013-2014

M.Tech.(CS) - Second Year

End-semester Examination (27 November, 2013)

Subject: Compiler Construction

Total marks: 120 Maximum marks: 100 Duration 3.5 hrs.

Please keep your answers brief and to the point.

1. Construct an example to show that the introduction of marker non-terminals in an LR(1) grammar can introduce conflicts in the LR(1) parsing table. [14]

2. The aim of this problem is to write a translation scheme (TS) to print to standard output a version of an input program that uses whitespace and indentation in a "proper" way. Assume that the program is written in a hypothetical language C--, whose syntax is given by the following grammar (S is the start symbol, all symbols not occurring on the left hand side of a production are terminals).

$S \rightarrow P$ (*Plist*) { *Stmt* }	*Stmt* $\rightarrow$ { *Stmt* }
$P \rightarrow T$ **id**	*Stmt* $\rightarrow$ **id** $= E$;
Plist $\rightarrow$ *Plist*, P | P	*Stmt* $\rightarrow$ **if** E **then** *Stmt* **else** *Stmt*
$T \rightarrow$ **char** | **int** | **void**	*Stmt* $\rightarrow$ **while** E **do** *Stmt*
$E \rightarrow E + E \mid E * E \mid -E \mid (E) \mid$ **id** | **num**	*Stmt* $\rightarrow$ **return** E | **return**

Add semantic rules to your grammar to "pretty print" the input program in accordance with the following guidelines. Clearly state any additional conventions that you follow.

- A single blank should precede and follow any binary operator (+, *, =).
- There must be no whitespace between a unary operator and its argument.
- No whitespace must follow an opening parenthesis; no whitespace must precede a closing parenthesis.
- The bodies of compound statements, conditionals and loops must be indented deeper than the first token of the statement.
- Braces surrounding a compound statement should appear on separate lines by themselves.

[20]

3. Consider the following function in C. Assume that pointers, and variables of type `int` occupy 4 bytes each.

```
static void
cumulative_freq(int *A, int num)
{
    int i, j;
    for (i = num; i > 0; i--)
        for (j = 0; j < i-1; j++)
            A[i-1] = A[i-1] + A[j];
    return;
}
```

(a) Convert the body of the procedure into 3-address code. Each time a temporary variable is needed, use a new temporary. You may use the name of an array instead of the constant (base address) associated with that array. **Do not perform any optimization at this stage.**

required

(b) Compute the minimum number of storage locations to store the temporary variables in your intermediate code. Draw a diagram showing the locations of temporaries.

[If you need to make use of next-use information, you may calculate this information by inspection.]

(c) Write the machine code for the calling sequence and return sequence for a call to `cumulative_freq` from `main`. Assume that (i) the stack grows from low addresses to high addresses; (ii) the stack pointer points to the beginning (i.e. lowest address) of an AR; and (iii) the AR for `main` occupies 320 bytes.

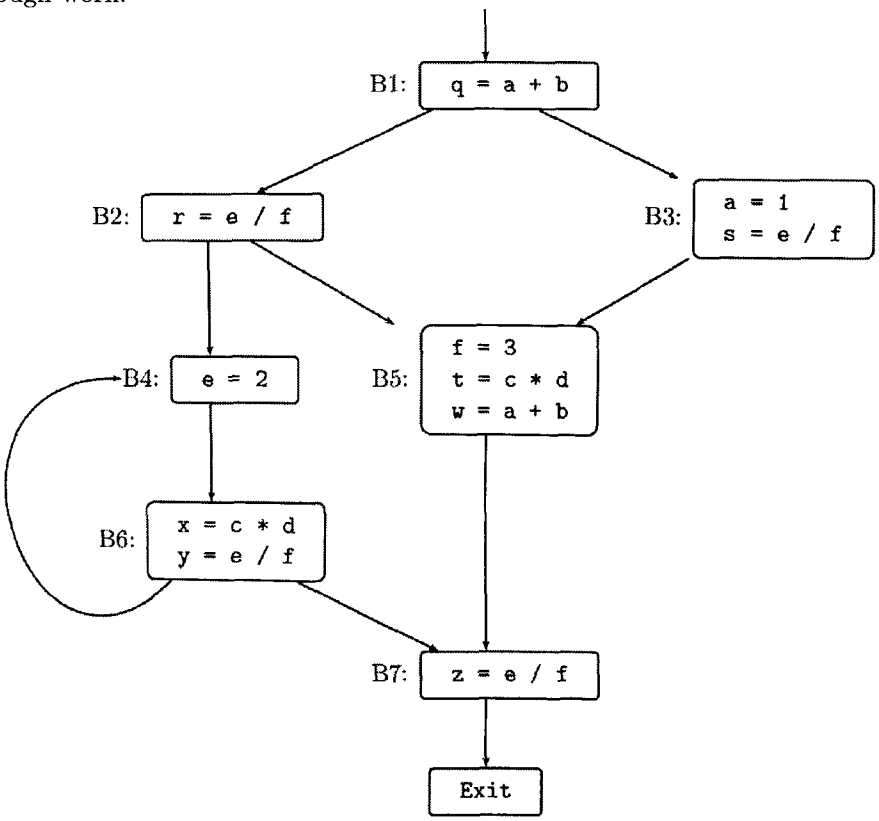
(d) Identify the leaders (and thus the basic blocks) in the 3-address code in (a).

(e) Identify the largest basic block in your optimized code (part (a)). Using a function like *getreg*, generate machine code for this for a generic microprocessor with 4 general-purpose registers. Assume that initially all variables are on the stack, and all registers are empty.

(f) Optimize your intermediate code in (a) by using whichever of the following techniques are applicable: constant folding, global common sub-expression elimination, copy propagation, dead code elimination, code motion, induction variable elimination.

[14+8+12+10+12+12=74]

4. For the following flow graph, compute the available expressions at the entry point for each basic block. Show your rough work. [12]



INDIAN STATISTICAL INSTITUTE

Semestral Examination

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

Multidimensional Searching and Computational Geometry

Date : 29.11.2013 Maximum Marks : 100 Duration : 3.5 Hours

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

1.(a) Define the following terms in the context of *Fortune's Line Sweep Algorithm* for constructing the Voronoi diagram of a set of points:
(i) beach line, (ii) site event and (iii) circle event

(b) Show that (i) the only way in which a new arc can appear on the beach line is through a site event, and (ii) the only way in which an existing arc can disappear from the beach line is through a circle event.

(c) Design an $O(n\log n)$ time algorithm which takes as input a set of points S and reports for each point in S its nearest neighbor in S.

$[(3 \times 2) + 5 + 5 = 16]$

2.(a) Assume that the convex hull of a given point set in $\mathbb{R}^2$ is already computed.

Design a randomized incremental algorithm for computing the furthest point Voronoi diagram of the point set P so that the expected time complexity is $O(n \log n)$. Analyze the worst case time complexity of your algorithm.

(b) Suppose you are given a set P of n points in $\mathbb{R}^2$. In the *minimum width annulus* problem, the objective is to compute two concentric circles of different radii (say r_1 and r_2 respectively with $r_1 > r_2$) such that (i) all the points lie in the annulus formed by the two circles, and (ii) the width of the annulus $(r_1 - r_2)$ is minimum. Suppose you have already computed the nearest and furthest point Voronoi diagram. Use these two data structures to design an $O(n^2)$ time algorithm for the minimum width annulus problem. Justify the time complexity result of your proposed algorithm.

[(15+5)+(11+4)=35]

3.(a) Given a set P of n points and a vertical line ℓ, design a linear time algorithm for computing the intersection of the line ℓ with the median level in the arrangement of dual lines of the points in P.

(b) Now, consider two sets of points, namely "red" and "black", that are linearly separable. The objective is to find a line that partitions both the sets in two equal halves.

Use your proposed method for the problem in question 3(a) to describe a prune and search strategy for solving the problem that runs in $O(n)$ time. Justify the time complexity of your proposed algorithm.

[8+(9+5)=22]

4.(a) The GEOMBASE problem is defined as follows: *given a set of points with integer coordinates on three horizontal lines $y = 0$, $y = 1$ and $y = 2$; determine whether there exists a strictly non-horizontal line that contains three points.*

Design an efficient algorithm for this problem. State and justify the time complexity of your proposed algorithm.

(b) The separator problem is defined as follows: *Given a set L of n horizontal line segments, does there exist a non-horizontal line that splits the set L into two non-empty subsets.*

Show that if you have an algorithm for the separator problem that runs in $O(f(n))$ time, then GEOMBASE problem can be solved in $\min(f(n), n \log n)$ time. [10+10=20]

5.(a) Let S be a set of n points in $\mathbb{R}^d$. Let $G = (S, E)$ be an weighted undirected complete graph where the weight of an edge (p, q) is the Euclidean distance between the points $p, q \in S$. $MST(S)$ is the minimum spanning tree of the graph G, and $SMT(S)$ is a steiner tree that connects all the points in S. Note that, there may exist several steiner trees for connecting the points in S. Show that $W(MST(S)) \leq 2W(SMT(S))$, for any steiner tree, where $W(T)$ is the sum of edge weights of the tree T.

(b) Consider the arrangement $\mathcal{A}(L)$ of a set L of n lines in $\mathbb{R}^2$. Let ℓ ($\ell \notin L$) be an arbitrary line. Show that the number of cells of $\mathcal{A}(L)$ intersected by the line ℓ is at most $O(n)$.

[10+10=20]

6.(a) Let $\theta = \frac{2\pi}{k}$, and $k \geq 8$. Consider an angular sector C^p of angle θ with apex at a point p. Let ℓ^p be a line through p that lies inside C^p. Consider two other points $q, r \in C^p$. The projection of r on ℓ^p is closer to p than the projection of q on ℓ^p. Prove that,

(i) $|pr| \leq \frac{|pq|}{cos\theta}$, and
(ii) $|rq| \leq |pq| - (cos\theta - sin\theta)|pr|$.

(b) Define Θ-graph for a set S of n points in $\mathbb{R}^d$ with $\theta = \frac{2\pi}{k}$, and $k \geq 8$. You must define all the notations, you use, very clearly.

(c) Show that, if a Θ-graph is constructed for the point set P with $k \geq 9$, then the stretch factor of a path between two nodes in Θ-graph is $\frac{1}{cos\theta - sin\theta}$, where the stretch factor of a path from p to q in the graph is the ratio of length of the shortest path from p to q in the graph and the Euclidean distance of p and q in the plane.

[(3+4)+5+6=18]

Indian Statistical Institute

Semester Examination : 2013 – 2014

Master of Technology in Computer Science, Semester III

Functional Brain Signal Processing: EEG & fMRI

Date: 30 November 2013 Maximum Marks: 100 Duration: 3 hours

Attempt all the questions. Credit will be given for precise and brief answers.

1. Describe T_1, T_2 and T_2^* weighted magnetic resonance imaging. Underlying MR physics will have to be described (illustration with diagrams might be helpful). Reason in favor of the relation $T_2^* \leq T_2 << T_1$. 7 + 7 + 3 + 3 = 20

2. Describe precise spatial localization in MR imaging in terms of *slice localization*, *frequency encoding* and *phase encoding* (diagrams might be helpful). 8 + 6 + 6 = 20

3. Describe any four fMRI artifacts. Give a brief outline about how to remove each of them. Feel free to propose your own ideas. Mathematical equations, formulations are not essential. 4 x 5 = 20

4. Describe general linear model (GLM) for processing of fMRI signals for a single subject (a two-regressor model will be good enough). Mathematical equations and their solutions in general form are required. Feel free to put forward geometric justifications wherever appropriate, perhaps with illustrative diagrams. 20

5. (a) Write a short (but content rich) note on multi-voxel pattern analysis (MVPA). 10

   (b) Mention two most predominating artifacts on EEG signals typical in an environment of simultaneous EEG-fMRI recording (none appears during EEG acquisition far outside of an fMRI scanner), with a brief explanation for each of them. Propose one scheme for removing each of the artifacts. 2 + 2 + 3 + 3 = 10

Indian Statistical Institute
Semestral Examination: 2013
Course Name: M. Tech. in Computer Science
Subject Name: Mobile Computing

Date: 30-11-2013 Maximum Marks: 100 Duration: 3 hours

Instructions: You **may** attempt **all** questions which carry a total of **110** marks. However, the maximum marks you can score is only **100**.

1. (a) What is a wireless sensor network? List some basic functionalities and characteristics of proactive and reactive sensor networks. [3+4=7]

   (b) Compare and contrast the main characteristics of hierarchical and flat topologies of wireless sensor networks. [6]

   (c) List some reasons why the traditional routing protocol defined for wireless ad hoc networks are not well suited for wireless sensor networks. Under what condition does the direct-communication routing require less energy than the minimum-transmission-energy routing in wireless sensor networks? Describe the key features of low-energy adaptive clustering hierarchy (LEACH) routing protocol in wireless sensor networks. [3+5+8=16]

   (d) What are the faults that may occur at different layers of the wireless sensor networks? Briefly discuss the self-diagnosis and cooperative diagnosis fault detection techniques in wireless sensor networks. Discuss how two-tiered network architecture helps improving reliability and prolonging lifetime of wireless sensor networks. [2+4+5=11]

2. (a) What is a cognitive radio network? [3]

   (b) Compare the fixed spectrum access and dynamic spectrum access policies in cognitive radio network. [5]

   (c) What are the cognitive capabilities that the secondary users must have to support the dynamic spectrum access? [6]

   (d) Briefly describe the opportunistic spectrum access and concurrent spectrum access models in cognitive radio network. [4+4=8]

   (e) What is a spectrum hole? How direct spectrum sensing is used to identify the spectrum hole? [3+5=8]

3. (a) What is an ad hoc network? List the main features of an ad hoc network. [3+4=7]

   (b) What is multicasting in an ad hoc network? With an example, explain the on-demand multicast routing protocol in an ad hoc network. [3+8=11]

   (c) What is flooding? Explain the concept of expected zone and request zone as defined in location-aided routing protocol in an ad hoc network. Explain how expected zone and request zone concepts help reducing route request flood in location-aided routing protocol in an ad hoc network. [2+5+5=12]

4. (a) What are hidden terminal and exposed terminal problems? [2+2=4]

   (b) When using RTS/CTS, how does an exposed terminal decide it is safe to send to another destination? [3]

   (c) When using RTS/CTS, what prevents a hidden terminal from clobbering the packets that another node is sending? [3]

INDIAN STATISTICAL INSTITUTE
M. Tech.(Computer Science) II Year, 2013-14
Semestral Examination
Pattern Recognition and Image Processing

Date: 2-12-13 Maximum Marks: 100 Duration: 195 minutes

Note: This paper carries 108 marks. Answer as much as you can.

1. Draw the solution tree for branch and bound feature selection algorithm if 3 features are to be selected from 7 features. [6]

2. Let $\underline{X}' = (X_1, X_2, X_3, X_4)$ be a random vector with dispersion matrix Σ, where $\Sigma = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 4 & -1 \\ 0 & 0 & -1 & 4 \end{pmatrix}$. Find the first two principal components of $\underline{X}$. [8]

3. Describe any two feature selection algorithms. [8]

4. Write short notes on Probabilistic separability criteria for feature selection. [5]

5. Describe the single linkage clustering algorithm. [5]

6. Describe Canny's edge detection technique for gray level images. [15]

7. Describe a method of finding line segments in a binary image using Hough transform technique. [12]

8. (a) Describe a method of thinning and show its effect in each step on the following binary image where 'x' denotes white pixel and '0' denotes a black pixel. [10+8]

0	0	x	x	x	x
0	x	x	x	x	x
0	x	x	x	x	0
0	0	x	x	x	x
0	0	x	x	x	x
0	x	x	x	x	0
0	x	x	x	0	0

(b) Describe the boundary of the above object by using 4-directional numbers. [4]

(P.T.O)

9. Describe a region based segmentation method for a gray level image using quad-tree. [12]

10. Write short notes on the following.
    (a) Salt and pepper noise.
    (b) Median filtering.
    © Skeleton of an object in an image. [3 x 5 = 15]

INDIAN STATISTICAL INSTITUTE
End-Semester Examination: 2013
Course Name: M.Tech. In Computer Science
Subject Name: Software Design and Validation

Date: Dec 3, 2013 *Maximum Marks: 100* *Duration: 3 hours*

Answer any 4 questions

---

1. [Timing Analysis] [(6 + 8 + 2) + 9 = 25 marks]

(a) Consider the following program fragment where **x** and **y** serve as inputs and **z** serves as the output. Both the inputs are positive integers, given as unsigned 8-bit numbers (when represented in binary). You may assume each assignment / return / condition-evaluation takes 1 time unit.

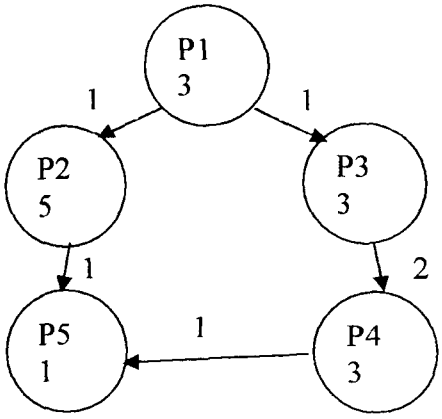
```
z = 0;
while ( x != 0) {
      if (x % 2 != 0) { z = z + y;}
      y = 2 * y;
      x = x / 2;
}
return z;
```

i) Using the timing schema WCET analysis method, derive the maximum execution time of the program fragment.

ii) Formulate the maximum execution time estimation of the program fragment using Integer Linear Programming (ILP). Clearly show the objective function and all constraints. Your ILP should perform only program path analysis and not micro-architectural modeling. The estimate produced by your ILP should be as tight as possible.

iii) Comment on how the estimate from your ILP problem will compare with the estimate you produced using timing schema.

(b) Suppose we have two processors A and B connected via a bus and five programs P1, P2, P3, P4 and P5 as shown in Figure 1. The data dependencies between the programs appear as edges in Figure 1. The computation and bus communication times are shown along the nodes and edges in Figure 1. Communication within a processor takes zero time.



Suggest a partitioning of programs P1...P5 to **A** and **B**, so that the overall execution time is minimized.

2. [Temporal Logic, Automata-based Model Checking] [(4 + 9) + 8 + 4 = 25 marks]

(a) Let AP = { a } and $\varphi = (a \wedge \mathbf{X}\, a)\ \mathbf{U} \neg a$ an LTL formula over AP

i) Is φ satisfiable? If yes, give an example run. If not, justify your answer.

ii) Construct the Büchi automaton $\mathcal{G}_\varphi$ such that $\mathcal{L}_\omega(\mathcal{G}_\varphi)$ = Words(φ). Explain your satisfiability result above on this automaton.

b) Draw a Buchi automaton that accepts the following ω-regular language:

$L = \{\sigma \in \{A, B\}^\omega \mid \sigma$ contains ABA infinitely often, but AA only finitely often$\}$

c) The language ACTL restricts CTL to using only the universal ($\forall$) quantifier. Give examples to illustrate the difference between ACTL and LTL.

3. [Communication Analysis, Scheduling] [10 + 15 = 25 marks]

(a) Deadline monotonic scheduling (DMS) is a fixed-priority preemptive scheduling algorithm that is similar to rate monotonic scheduling (RMS). In this case, priorities assigned to tasks are inversely proportional to the length of the deadline. Thus the task with the shortest deadline is assigned the highest priority, and the longest deadline task is assigned the lowest priority.

For a set of tasks, is it possible that RMS does not meet all the deadlines, but DMS can meet all the deadlines? If your answer is no, present a formal proof of the claim. If your answer is yes, present an example where this happens.

(b) Explain the communication incompatibility problem which arises from maintaining address-data correspondence in a burst while communicating over a bus. Elaborate on the converter design and comment on the storage requirements of the converter and the sender interface. Can we have a solution that avoids the central converter and manages the address-data correspondence only via sender/receiver interfaces?

4. [Functionality Validation] [5 + 6 + 14 = 25 marks]

(a) Consider the following program fragment:

```
if (x > 2) { y = x + z;} else { y = x - z;}
if (y > 0) { return 0;} else { return 1;}
```

Give one example test input (y, z) such that the code above returns 0. Characterize the set of all test inputs that cause the foregoing code to return 0. Justify.

(b) Consider the following program with two threads, which are composed asynchronously. Assume that initially A = 0, and each assignment is executed atomically. What are the possible contents of the array X when the program terminates? Explain your answer without explicitly constructing the composition.

```
Thread 1: (A := 1; A := 2; A := 3; A := 4)
Thread 2: (X[1] := A; X[2] := A; X[3] := A; X[4] := A;)
```

c) Consider the following program fragment:

```
x = 0; while (x < 100) { x = x + 1; }
```

Suppose we want to prove that (x == 100) at the end of the program. What is the initial abstract transition system we start with if we follow the abstraction-modelcheck-refinement methodology? What are the abstractions of the memory store (predicate abstractions) that we will encounter if we prove the property by abstraction refinement?

5. [Functionality Validation, Scala] [10 + 7 + 8 = 25 marks]

(a) One method of software testing for inputs with large domains is called *equivalence partitioning*. In this method, the domain of an input variable is *partitioned* into equivalence classes, so that from each equivalence class, only one test input will be tried out. We call two test inputs to be equivalent when they produce the same path in the program.

Give an example where such an equivalence partitioning will lead to efficient testing, that is, only a few test cases to try. Give an example where such an equivalence partitioning will lead to inefficient testing, that is, too many cases to try.

(b) Suppose you want to use a model checker (such as SPIN) to generate test cases of a terminating sequential program written in a C style imperative programming language. What are the temporal properties you can feed in to meet the statement coverage for test generation? Explain on an example.

(c) Prove or disprove the following statements in the Scala paradigm:

i) If call-by-value (CBV) evaluation of an expression **e** terminates, then call-by-name (CBN) evaluation of e terminates, too.

ii) If call-by-value (CBN) evaluation of an expression **e** terminates, then call-by-name (CBV) evaluation of e terminates, too.

INDIAN STATISTICAL INSTITUTE
M. Tech (CS) II year : 2013–2014
Cryptology
Semestral Examination

Date: 2. 12. 2013 Maximum Marks: 100 Time: 3 Hours

Answer any five questions.
Please answer all parts of a question at the same place.

1. Consider the Geffe generator where the outputs of three LFSRs having length 29, 37 and 47 are combined by a multiplexer to generate the key stream and then message bits are XORed with the key stream bits to generate the cipher bits. Describe a ciphertext only attack on this scheme which requires lesser time complexity than the exhaustive search. (20)

2. (a) Describe the Key Scheduling Algorithm (KSA) of RC4.
   (b) Prove that the initial bytes of the permutation S after the KSA are biased towards the secret keys.

   (5+15 = 20)

3. (a) Describe the RSA and CRT-RSA public key cryptosystems (setup, encryption and decryption) with examples taking 2 digit primes.
   (b) Briefly outline the Pollard Rho algorithm for factoring large integers.

   (15+5 = 20)

4. (a) What is the Discrete Log problem?
   (b) Explain the ElGamal Public-Key cryptosystem (setup, encryption and decryption) with examples taking a 2 digit prime.

   (5+15 = 20)

5. (a) Explain different modes of operation of a Block cipher with proper figures.
   (b) Describe how a block cipher can be used as a self synchronizing stream cipher.

   (15+5 = 20)

P.T.O.

6. (a) What are the security issues that you need to address while designing a cryptographic hash function.

   (b) Consider a hash function for which the message digest is 80 bits long. Will you consider this hash function secure? Justify with proper mathematical reasoning.

   (c) Explain the Merkle-Damgard construction for iterated hash functions.

   $(5 + 5 + 10 = 20)$

7. (a) List the points on the elliptic curve $y^2 = x^3 + x + 6$ over Z_{11}.

   (b) Describe (with examples using the elliptic curve in the previous question) how an elliptic curve modulo a prime can be used to design a public key cryptosystem. Clearly explain the encryption and decryption algorithms.

   $(10+10 = 20)$

8. Write short notes on (any two):

   (a) Advanced Encryption Standard (AES),

   (b) Message Authentication Codes (MAC),

   (c) Shank's algorithm for solving the Discrete Log Problem.

   $(10+10 = 20)$

INDIAN STATISTICAL INSTITUTE

First-Semester Examination: 2013-2014

M. Tech. (CS) 2nd Year

Artificial Intelligence

Date: 04.12.2013 Maximum Marks: 100 Duration: 3 hours

Answer all questions in brief.

1. The game of NIM is played as follows: Two players alternate in removing one, two or three coins from a stack initially containing five coins. The player who picks up the last coin loses.
   a) Draw the full game tree.
   b) Show that the player who has the second move can always win the game.
   c) Execute α-β pruning procedure on the game tree. How many terminal nodes are examined? For each cutoff, specify whether it is α-cutoff or β-cutoff. [3 + 2 + (2 + 1 + 2) = 10]

2. Answer the following:
   a) Prove that if there is a tableau proof of α from a set of premises $\sum$, then α is a logical consequence of $\sum$.
   b) Show that $(\exists x)\,(P(x) \wedge Q(x)) \rightarrow (\exists x)\,P(x) \wedge (\exists x)\,Q(x)$ is valid whereas the converse $(\exists x)\,P(x) \wedge (\exists x)\,Q(x) \rightarrow (\exists x)\,(P(x) \wedge Q(x))$ is not. [4 + 6 = 10]

3. Answer the following:
   a) Write a program in Prolog for in-order traversal of a binary tree. The traversal method stores the elements of the tree in a list.
   b) Explain with example the difference between red cut and green cut in Prolog.
   c) Describe the difference between the following two codes (i) and (ii) written in Prolog when the goal query is "grandfather(james, X)":

(i) grandfather(X,Y):- father(Z, Y), father(X, Z). father(james, robert). father(mike, william). father(william, james). father(robert, hency).	(ii) grandfather(X,Y):-father(X, Z), father(Z, Y). father(james, robert). father(mike, william). father(william, james). father(robert, hency).

   d) Write a program in Prolog for merging two ordered lists. [6 + 4 + 5 + 5 = 20]

4. Consider the following set of sentences. "Mary will get her degree only if she registers as a student and passes her examination. She has registered herself as a student. She has passed her examination." Prove that "she will get a degree" using both
   a) semantic tableaux approach; and
   b) resolution refutation method. [5 + 5 = 10]

5. Answer the following:
   a) Prove that a clause *C* is a logical consequence of a set of clauses *S* if and only if the set *S'* = *S U {~C}* is unsatisfiable.
   b) Prove using semantic tableaux approach that the following sentences are mutually consistent. "All Indian citizens who are adult have right to vote in election. Mary is an Indian citizen and has voting right. Mary is an adult."

c) What are the main features of an expert system? Explain with an example the forward chaining approach in a rule based expert system. [6 + 6 + (4 + 4) = 20]

6. Answer *any three* from the following: [3 X 10 = 30]

a) Discuss the Dempster-Shafer theory of evidence. Suppose an initial observation S_1 confirms some hypothesis h with the belief MB = 0.35. The second observation S_2 performs the same hypothesis h with the belief MB = 0.45. Find the certainty factor of the hypothesis h using two observations S_1 and S_2. [6 + 4 = 10]

b) Describe the following with suitable examples:
(i) Bayes' theorem for the probabilistic reasoning;
(ii) Simulated annealing approach. [5 + 5 = 10]

c) In *farmer-fox-goose-grain* puzzle, a farmer wishes to cross a river taking his fox, goose, and grain with him. He can use a boat which will accommodate only the farmer and one possession. If the fox is left alone with the goose, the goose will be eaten. If the goose is left alone with the grain it will be eaten. Draw a state space search tree for this puzzle. Denote left and right river banks as left-bank and right-bank, respectively. [10]

d) Let $I = <U, A>$ be a decision table, where $U = \{x1,\ldots\ldots, x6\}$ is a nonempty set of finite objects, the universe, and $A = C \cup D$ is a nonempty finite set of attributes. Here, C = {Headache, Muscle pain, Temperature} and $D = \{Flu\}$ are the sets of condition and decision attributes, respectively.

U	*Headache*	*Muscle pain*	*Temperature*	*Flu*
x1	Yes	Yes	Normal	No
x2	Yes	Yes	High	Yes
x3	Yes	Yes	Very high	Yes
x4	No	Yes	Normal	No
x5	No	No	High	No
x6	No	Yes	Very high	Yes

In the context of rough set theory, explain the following with the above example data:
i) lower and upper approximations of decision attribute,
ii) degree of dependency of a condition attribute, and
iii) reduct and core of the decision table. [(2 + 2) + 2 + (2 + 2) = 10]

Indian Statistical Institute

Semester Examination 2013-14
M. TECH. (CS) II Year
Subject: Document Processing and Retrieval
Full Marks: 100 Duration: 3 hrs.

04·12·13

(Answer all questions)

1. Generate a feature based tree classifier to recognize the following alphanumeric characters.

   E, F, 3, 5, R, P, t, f, O, Q, S [11]
2. What is Word-spotting? Discuss about a scale invariant word-spotting method for Devnagari printed words. [2+7]
3. Discuss two rotation invariant features that can be used for multi-oriented character recognition. [5+5]
4. Discuss a digital straight line based approach for multi-skew detection of different text lines from a Devnagari printed document image? [10]

5. What is halftoning? Describe the basic principles of generating halftone image by digital means? Starting from D_2 and the relations given below, recursively produce a 8 X 8 Dither halftoning matrix D_8.

$$D_2 = \begin{bmatrix} 0 & 2 \\ 3 & 1 \end{bmatrix}$$

$$U_n = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 1 & 1 & \cdots & 1 \\ . & . & . & . & . \\ . & . & . & . & . \\ 1 & 1 & 1 & \cdots & 1 \end{bmatrix}$$

$$D_n = \begin{bmatrix} 4D_{n/2} + D_2(00)U_{n/2} & 4D_{n/2} + D_2(01)U_{n/2} \\ 4D_{n/2} + D_2(10)U_{n/2} & 4D_{n/2} + D_2(11)U_{n/2} \end{bmatrix}$$

   Describe briefly the principles of error diffusion halftoning approach. [2+3+10+5]

6. a) Can all possible files of size n-bits (n divisible by 4) be compressed into files of size n/2 bits? Justify your answer. How many n-bit files can at all be compressed into files of exactly n/4- bits? [3]

   b) Define an m^{th} order Golumb code. Find the codewords for 2^{nd} and 4^{th} order Golumb codes for ten symbols expressed as natural numbers 0, 1, 2 ... 9. How to choose m for efficient coding of pixel run lengths? [4 + 10 + 3]

7. a) Describe how classifier combination is made to improve recognition accuracy. Describe sum-rule, product-rule, max-rule and Borda count approach of classifier result combination. [5 + 8]

   b) Show that in the approach of normalizing the scores of individual classifier using *Informational confidence*, the performance function $p(K_i)$ for i^{th} classifier confidence K_i is given by,

$$p(K_i) = 1 - \exp(-K_i/E)$$

   where E is a multiplying factor influencing the scale. [7]

INDIAN STATISTICAL INSTITUTE
M.Tech. (Computer Science)
Second year, First Semester
Computer Architecture
Semestral Examination, 2013-14

05·12·13

Full Marks : 100 Time : 3 Hours

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

1. a) Consider an instruction pipeline with five phases as *instruction fetch, instruction decode/register fetch, execute/address calculation, memory access* and *write back*, respectively. Referring to this pipeline structure , explain the terms 'Structural hazards' and 'Control hazards', if any, in executing the instructions on this architecture.

   b) Consider the execution of the following machine language instructions (meaning of each instruction being as specified in the code segment) on the above pipelined architecture :

LW	R1, 0(R2);	R1 ← Mem{0+R2)
SUB	R4, R1, R5 ;	R4 ← R1 – R5
AND	R6, R7, R1 ;	R6 ← R7 * R1
DIV	R10, R11, R1;	R10 ← R11 / R1

   Indicate the possible data hazards that may be encountered in executing the above code section. Can these data hazards, if any, be avoided by any possible means? If so, how? If not, then why not? **(2+(2+4)= 8)**

2. a) Mention the distinguishing features between a superscalar processor architecture and a VLIW processor architecture.
   b) Consider the following assembly language program segment for a loop executed on an instruction-pipelined superscalar architecture having pipeline stages as described in Q. 1 above, where LD stands for a 'load double word' (8 bytes) instruction, ADDD stands for addition of two double word floating point numbers, SD stands for a double word store instruction, SUBI stands for subtracting an immediate operand from a register and BNEZ stands for 'branch on not equal to zero'.

LOOP :	LD	F0,0(R1)
	ADDD	F4,F0,F2
	SD	0(R1),F4
	SUBI	R1,R1,#8
	BNEZ	LOOP

   Assume that each data element is stored in the memory as a double word and the **superscalar architecture** provides one integer functional unit and another floating point functional unit. Also assume that the LD instruction needs 2 cycles and the ADDD instruction needs 3 cycles to complete. Also, a stall of one cycle is needed after the SUBI instruction as the BNEZ instruction computes the branch address and the condition for branching in the *instruction decode/register fetch* phase of the pipeline. Now answer the following :

How many cycles per data element are needed for the above program segment and how can it be reduced by loop unrolling and appropriate rescheduling of the instructions? Give one specific example for such a rescheduled code segment with five-fold loop unrolling and the corresponding number of cycles needed per data element. **(4+(2+6)=12)**

3. a) Indicate the possible control hazards that may be encountered while dealing with loops in a program. Explain clearly, with the help of appropriate state diagram, how a two-bit prediction system can reduce the number of mispredictions in controlling loops in a program.

b) Consider a loop as given below :

```
for (i = 0; i < 100; i = i+1) {
        A[i] = A[i] + B[i];       /* S1 */
        B[i+1] = C[i] + D[i] ;  /* S2 */
}
```

Does there exist any loop-carried dependence in the above loop segment? If yes, then mention that one. What are the dependences between S1 and S2? Can the loop carried dependence, if any, be eliminated to execute the different iterations of the loop in parallel? If so, then how?

(10+10=20)

4. Explain the disadvantages of static rescheduling of instructions by compilers in order to avoid pipeline stall cycles. Can these disadvantages be removed by dynamic scheduling? Illustrate your answer by a suitable example.

Describe the basic features of Scoreboarding technique for dynamic scheduling with specific discussions on the following items : i) dealing with structural hazards, ii) dealing with RAW hazards, WAR hazards and WAW hazards, iii) pipeline stages and their broad functions, iv) different information to be kept for controlling the total operation and v) the required control logic.

(4+16=20)

5. a) Explain how 'blocking' the data segments can help in reducing the number of cache misses. Consider the following program segment meant for multiplying two N x N matrices Y and Z to produce the product matrix X by dividing the matrix in terms of blocks of size B x B.:

```
for (jj = 0; jj < N; jj = jj + B)
for (kk = 0; kk < N; kk = kk + B)
for (i = 0; i < N; i = i + 1)
        for (j = jj; j < min(jj+B-1, N); j = j+1)
{r = 0;
 for (k = kk; k < min(kk+B-1, N); k = k+1){
        r = r + Y[i][k] * Z[k][j];};
X[i][j] = X[i][j] + r;
};
```

Calculate the total number of cache misses encountered in executing the above program segment and compare it with those which may be experienced without using the concept of 'blocking', under different values of available cache size.

b) Explain the problem of cache coherence in a multiprocessor environment. Show, with the help of appropriate state diagrams, how this cache coherence problem can be solved by using a snooping coherence protocol. **(10+10=20)**

6. a) Explain the terms 'convoy' and 'chime' in connection with a vector processor architecture.

b) Show how the following code sequence lays out in convoys, assuming a single copy of each functional unit, with V1, V2, V3 and V4 being the vector registers, F0 storing a scalar, Rx and Ry storing the starting addresses of the vectors X and Y respectively :

```
LV      V1, Rx ;      Load vector X
MULVS.D V2,V1,F0;     vector-scalar multiply
LV      V3,Ry;        load vector Y
ADDVV.D V4,V2,V3;     add two vectors V2 and V3
SV      V4,Ry;        store the sum
```

c) Considering a vector processor, show how the following computation with vectors of any arbitrary length ('n' in this example) can be handled very efficiently :

```
for (i = 0; i < n; i = i+1)
        Y[i] = a * X[i] + Y[i];
```

(4+6+10=20)

Indian Statistical Institute
M.Tech.(CS), Second Year
Information and Coding Theory
End-Term Examination

Date: December 6, 2013
Time 3 hours

The question paper contains 8 questions. Total marks is 85. Maximum you can score is 75.

1. Show that Hamming distance follows the triangle inequality: for any vectors $\mathbf{x} = x_1x_2\cdots x_n$, $\mathbf{y} = y_1y_2\cdots y_n$, $\mathbf{z} = z_1z_2\cdots z_n$,

$$\text{dist}(\mathbf{x},\mathbf{y}) + \text{dist}(\mathbf{y},\mathbf{z}) \geq \text{dist}(\mathbf{z},\mathbf{x}). \qquad (5)$$

2. Show that, if a code has minimum distance d and the codeword $\mathbf{x}$ is transmitted and not more than $\frac{1}{2}(d-1)$ errors occur, and $\mathbf{y}$ is received, then $\text{dist}(\mathbf{x},\mathbf{y}) < \text{dist}(\mathbf{y},\mathbf{z})$, for all codewords $\mathbf{z} \neq \mathbf{x}$. (5)

3. What is a perfect code? Show that the binary repetition code is not a perfect code. (1 + 4 = 5)

4. Prove the Gilbert-Varshamov bound. Suppose q is a prime power and r, n, d integers satisfying

$$\binom{n-1}{0} + \binom{n-1}{1} + \cdots + \binom{n-1}{d-2} < q^r.$$

   Then, show that an $[n, n-r, d]$-code over $\mathbb{F}_q$ exists. (5)

5. (a) Construct a Hadamard matrix H_8 of order 8, starting from a Hadamard matrix of order 2.
   (b) Construct two Hadamard codes $\mathcal{C}_1$ and $\mathcal{C}_2$ from this matrix. What are the parameters of these codes?
   (c) Construct a BIBD using H_8. What are the parameters of the design?
   (d) Can you give a generalized construction of (1) BIBD and (2) Orthogonal Array from a Hadamard matrix of order $4m$? (8 + 8 + 6 + 8 = 30)

6. What are MDS codes? Prove that Reed-Solomon codes are MDS linear codes. Construct a quaternary Reed-Solomon code of length 3 and dimension 2. (2+10+8 = 20)

7. Let $\mathcal{C}$ be a cyclic code of length n, which is an ideal in $R_n = F[x]/(x^n - 1)$. Show that there is a unique monic polynomial $g(x)$ of minimal degree in $\mathcal{C}$ and that this polynomial is the generator polynomial of $\mathcal{C}$. What is the generator matrix of $\mathcal{C}$? (3+3+4 = 10)

8. Prove the BCH-bound. Let $\mathcal{C}$ be a cyclic code with generator polynomial $g(x)$, such that for some integers $b \geq 0, \delta \geq 1$,

$$g(\alpha^b) = g(\alpha^{b+1}) = g(\alpha^{b+2}) = \cdots = g(\alpha^{b+\delta-2}) = 0.$$

   Show that the minimum distance of the code is at least δ. (5)

INDIAN STATISTICAL INSTITUTE

Semestral Examination

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

Advanced Algorithms for Graph and Combinatorial Optimization Problems

Date : 6.12.2013 Maximum Marks : 100 Duration : 3.5 Hours

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

1.(a) Define the concept of a *perfect graph.*

(b) Prove that if the two graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ are perfect, then the graphs $G_1 \bigcup G_2$ and $G_1 + G_2$ are also perfect, where
$G_1 \bigcup G_2 = (V_1 \bigcup V_2, \quad E_1 \bigcup E_2)$, and
$G_1 + G_2 = (V_1 \bigcup V_2, \quad E_1 \bigcup E_2 \bigcup \{(u,v) | u \in V_1, v \in V_2\})$.

(c) Let $G = (V, E)$ be a graph with the set of vertices $V = (v_1, v_2, \ldots, v_n)$, and $h = (h_1, h_2, \ldots, h_n)$ be a vector of non-negative integers. Let us construct $H = G \bullet h$ by substituting for each vertex v_i, an independent set of h_i vertices $v_i^1, v_i^2, \ldots, v_i^{h_i}$, and joining v_i^s with v_j^t if and only if v_i and v_j are adjacent in G.

Prove that if $w(G) = \chi(G)$, then $w(H) = \chi(H)$, where $w(G)$ and $\chi(G)$ are respectively the number of vertices in the maximum clique, and the chromatic number of the graph G. [2+(5+5)+10=22]

2.(a) Formulate the set cover problem as an integer programming problem.

(b) Show that the dual fitting based analysis for the greedy set cover actually establishes an approximation gurantee of H_k for the greedy set cover algorithm, where $H_k = \sum_{i=1}^{k} \frac{1}{i}$, and k is the size of the largest among the given subsets for covering the elements of the universal set U. [5+10=15]

3.(a) Consider a set of line segments $S = \{s_1, s_2, \ldots, s_n\}$. The coordinates of the end-points of each segment are input to the program. Show that $O(n \log n)$ time is sufficient to compute the minimum number of vertical lines that stab all the members in S.

(b) If the objective is to stab the line segments in S by minimum number of horizontal and vertical lines, then formulate the problem as a integer linear programming (ILP) problem.

(c) Design a constant factor approximation algorithm using the rounding mechanism of the linear programming solution of the aforesaid ILP problem. [8+6+12=26]

4. The metric k-center problem is a variation of facility location problem, and is defined as follows:

   Given a set of n points in $\mathbb{R}^2$ with a matrix that contains distances for every pair of points. Select k points to place the facilities such that the maximum distance of a point to its nearest facility is minimized.

   In order to design a 2-approximation algorithm for this problem, the following results are used.

   (i) Let $G = (V, E)$ be a graph. The graph $G' = (V, E')$ is defined by squaring the graph G. It has V as the set of vertices; between a pair of vertices $u, v \in V$, there is an edge $(u, v) \in E'$ if there is path of length at most 2 between u and v in G. If I is the independent set of G' and D is the dominating set of G, then $|I| \leq |D|$.

   (ii) A maximal (not maximum) independent set of G' can be obtained in time polynomial in n.

   (a) Justify these results.

   (b) Design a 2-approximation algorithm for the metric k-center problem. You may or may not use the aforesaid results.

   [6+5+12=23]

5. Max-SAT problem is defined as follows:

   Given a boolean formula in CNF over n variables $x_1, x_2, \ldots, x_n$, and a weight w_i for each clause C_i, $1 \leq i \leq m$, find the truth assignment to the variables such that it maximizes the sum of weights of the satisfied clauses.

   A simple randomized algorithm for the Max-SAT problem is designed as follows: *set the variable x_i to 1 and 0, both with probability 0.5.*

   (a) Show that if W is the sum of weights of the satisfied clauses, returned by the above algorithm, then $E(W) = \frac{1}{2}OPT$, where OPT is the optimum solution.

   (b) Describe a method of derandomization to achieve a deterministic $\frac{1}{2}$-factor approximation algorithm for the Max-SAT problem. [7+8=15]

6. Consider the bin-packing problem, where a set of n objects of different sizes are given as input. The objective is to pack those objects in minimum number of bins of fixed size.

   (a) Show that, when all items are of size less than ϵ, then the first fit algorithm produces a solution of size $(1 + \epsilon)OPT + 1$, where OPT denotes the minimum number of bins required to pack the given objects.

   (b) If all the objects are of size larger than ϵ, then if there are g (a constant integer number) distinct sizes, then the problem can be solved in time polynomial in n.

   [8+8=16]

INDIAN STATISTICAL INSTITUTE

Semestral Examination: (2013 - 2014)

Course Name: M.Tech (CS) **Year: 2nd year**

Subject Name: Neural Networks & Applications

Date: December 06, 2013 **Maximum Marks: 100** **Duration: 3 hrs**

Answer all the questions.

1. Show how the Oja's model of Principal Component Analysis Network extracts the last two principal components from a set of n-dimensional samples. State clearly all the assumptions made in this regard. [30]

2. State and prove the percentron convergence theorem. [20]

3. Show how momentum factor can be incorprated in the leaning rule of a multilayer perceptron to speed up the learning process. [10]

4. a) Describe the architecture and learning rule of Kohonen's Self Oraganizing Feature Map.
b) Show how the model can be used for clustering a set of high-dimensional points.
[15+5 = 20]

5. Write short notes on any TWO of the following. [2×10 = 20]
   a) Spiking neurons
   b) Spiking neural networks
   c) Hebb rule of learning
   d) Reinforcement learning
   e) Use of artificial neural networks for function approximation
   f) Use of spiking neural networks for time series prediction.

INDIAN STATISTICAL INSTITUTE

Final Examination:(2013-2014)

MTech C.S. 2nd Year

Digital Signal Processing

Date: 09.12.2013 Maximum Marks: 100 Duration: 3 hours

Note: The marks add up to 117. The maximum you can score is 100. The exam is open-book, open-notes. Use of calculators is permitted.

1. An LTI discrete-time system has the difference equation

$$y[n] = b_1 x[n+k] + b_2 x[n+k-1] + b_2 x[n+k-3] + b_1 x[n+k-4]$$

where b_1 and b_2 are real.

(a) Determine the frequency response $H(e^{j\omega})$.

(b) Give a value of k such that $H(e^{j\omega})$ is a real function of ω.

[5+5]

2. Two causal first order LTI discrete-time systems are cascaded together, with system functions

$$H_1(z) = \frac{2 - 0.4z^{-1}}{1 + 0.7z^{-1}}, \quad H_2(z) = \frac{0.5 + z^{-1}}{1 + 0.5z^{-1}}$$

(a) Find the system function of the overall system and the impulse response.

(b) Give Direct Form II and a Parallel Form implementation.

[(3+5)+(4+4)]

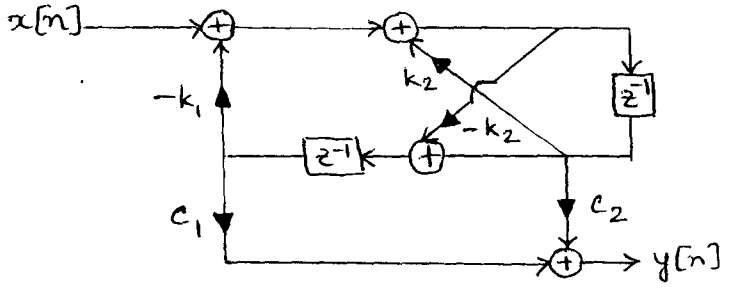
3. Let $x[n]$ be a sequence of length 9.

(a) Draw a flow graph for computing $X[1]$, $X[4]$ and $X[7]$ using one 3-point Discrete Fourier Transform (DFT).

(b) For $x[n] = \{1\ 2\ 0\ 3\ 4\ 5\ 0\ 7\ 6\}, n = 0, \cdots, 8$, calculate $X[1]$, $X[4]$ and $X[7]$ using your flow graph.

[5+5]

4. Calculate the system function of the system with block diagram given below: [10]



5. Consider a length 6 sequence $x[n]$. Let $y[n]$ be a sequence such that $y[n] = e^{j\frac{2\pi}{6}n}x[n]$, having Discrete Time Fourier Transform (DTFT) $Y(e^{j\omega})$. Let

$$Y_4[k] = Y(e^{j\omega})|_{\omega=\frac{2\pi k}{4}}, \quad k=0,\cdots,3$$

Determine the length 4 sequence $y_4[n]$ obtained as the Inverse DFT of $Y_4[k]$. [10]

6. An analog system has the transfer function

$$H(s) = \frac{4}{(s+2)(s-3)}$$

A discrete-time system is to be obtained from it using Impulse Invariance. Determine the impulse response. Also state the location of the poles and whether the system will be stable. [5+3+2]

7. Consider a sequence $x[n] = -(0.3)^n u[-n-1]$ where $u[n]$ is the unit-step sequence.

   (a) Without computing $X(z)$, determine $X(z^5)$.

   (b) Determine $(1+z^{-2})X(z^5)$, without computing $X(z)$.

   [5+5]

8. A sequence $x[n]$ of length 15 is to be filtered by an LTI system with impulse response of length 5. Determine the number of multiplications required for

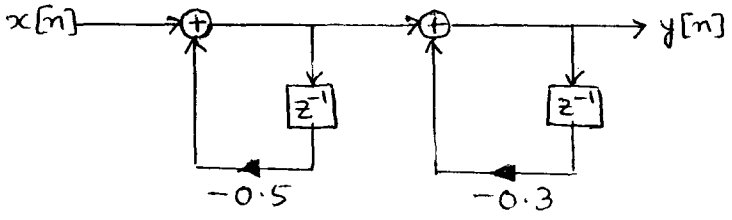
(a) A direct computation of the convolution.

(b) Using radix-2 FFT and linear convolution.

(c) Using radix-2 FFT and circular convolution.

[4+6+6]

9. For the LTI system with structure given below, compute the output noise power due to rounding off, using 4-bit, signed two's-complement fixed-point arithmetic. [10]



10. An LTI system has the following system function

$$H(z) = (1-0.8e^{j0.4\pi}z^{-1})(1-0.8e^{-j0.4\pi}z^{-1})(1-1.2e^{j0.6\pi}z^{-1})(1-1.2e^{-j0.6\pi}z^{-1})$$

Is the system minimum phase? If not, determine a minimum-phase system with the same magnitude response. [10]

11. A causal LTI FIR discrete-time system has an impulse response

$$h[n] = b_0\delta[n] + b_1\delta[n-1] + b_2\delta[n-2] + b_3\delta[n-3] + b_4\delta[n-4]$$

What values should the coefficients $b_0, \cdots, b_4$ have so that the frequency response has a constant group delay? [5]

INDIAN STATISTICAL INSTITUTE

Mid-Semestral Examination : 2013 – 14

MTech CS (2nd Year)

Computational Finance

Date: 25 February 2014 Maximum Marks: 20 Duration: 1½Hours

1. Critically explain the concepts: [2 X 2 = 4]
   a) Mutual Fund Principle
   b) Law of One Price

2. Let

$$A_{(K+1)X(K+2N)} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 1 & 1 & \cdots & 1 \\ \Delta S_1^*(\omega_1) & -\Delta S_1^*(\omega_1) & \Delta S_2^*(\omega_1) & \cdots & -\Delta S_N^*(\omega_1) & -1 & 0 & \cdots & 0 \\ \Delta S_1^*(\omega_2) & -\Delta S_1^*(\omega_2) & \Delta S_2^*(\omega_2) & \cdots & -\Delta S_N^*(\omega_2) & 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \Delta S_1^*(\omega_K) & -\Delta S_1^*(\omega_K) & \Delta S_2^*(\omega_K) & \cdots & -\Delta S_N^*(\omega_K) & 0 & 0 & \ldots & -1 \end{bmatrix}$$

and $b_{(K+1)} = (1, 0, \ldots, 0)'$. Show that

$$Ax = b, \; x \geq 0, x \in \mathbb{R}^{K+2N}$$

has a solution if and only if there exists an arbitrage opportunity in the securities market with N securities and K states of nature. [5]

3. In the two period model, explicitly solve the Consumption Investment problem for the utility function $u(w) = \ln w$. Compute the relevant expressions and solve for the optimal trading strategy when N = 1, K = 2, r = 1/9, $S_0 = 5$, $S_1(\omega_1) = 20/3$, $S_1(\omega_2) = 40/9$ and $P(\omega_1) = 3/5$. [7]

4. Prove the Put – Call parity for a European option for the multi-period market. [4]

INDIAN STATISTICAL INSTITUTE

Mid Semestral Examination

M. Tech (CS) - II Year, 2013-2014 (Semester - IV)

Topics in Algorithms and Complexity

Date : 24.02.2014 Maximum Marks : 60 Duration : 2.5 Hours

Note: Answer as much as you can, but the maximum you can score is 60.

(Q1) Consider the following algorithm for sorting. Given an input array $\mathcal{A}$ of n distinct numbers, we keep shuffling $\mathcal{A}$ randomly until it is sorted. The shuffling should generate any permutation of the input independently and equally likely.

(a) Design such a shuffling algorithm and argue why your algorithm returns any permutation independently and equally likely.

(b) Analyze the expected time complexity of the said sorting algorithm.

$[5 + 5 = 10]$

(Q2) Let A be an unsorted array on n integers. We need to search for a value a in A. Consider the following randomized strategy: choose a random index $i \in [1, n]$ and check if $A[i] = a$. If $A[i] = a$, then we terminate; otherwise, we continue the search by choosing a new random index into A. We continue choosing random indices into A until we find an index j such that $A[j] = a$, or until we have checked every element of A. Notice that, we choose from the whole set of indices every time, so that we may examine a given element more than once.

(a) Suppose, there is only one index i such that $A[i] = a$. What is the expected number of indices that must be tried before a is found and the algorithm terminates?

(b) Suppose, there are $k \geq 1$ indices i such that $A[i] = a$. What is the expected number of indices into A that must be chosen before the algorithm terminates? Your answer should be a function of n and k.

(c) Suppose, there is no index i for which $A[i] = a$. What is the expected number of indices into A that must be chosen before all elements of A have been checked and the algorithm terminates?

$[3 + 3 + 4 = 10]$

(Q3) A hypergraph $\mathcal{H} = (V, E)$ is a generalization of a graph. V is a set of n vertices. $E = \cup_{i=1}^{m} S_i$, where $S_i \subseteq V$. Each S_i is termed as a hyperedge. Notice that if all S_is are 2-element subsets of V, then a hypergraph becomes a graph.

The problem is to find *balanced coloring* of a hypergraph. Each vertex in V is to be assigned the color RED or BLUE. The *discrepancy* $\mathcal{D}(i)$ of a hyperedge S_i is defined as $|R_{S_i} - B_{S_i}|$, where R_{S_i} and B_{S_i} denote the number of red and blue vertices in the hyperedge S_i. The

discrepancy $\mathcal{D}$ of $\mathcal{H}$ is defined to be $\max_i \mathcal{D}(i)$. The goal is to assign colors to the vertices so that $\mathcal{D}$ is minimized.

If each vertex is colored RED or BLUE with probability $1/2$ independently, then obtain high probability tail bounds on $\mathcal{D}$. You can use Chernoff bounds and the union bound. [15]

(Q4) Suppose that we have an algorithm that takes as input a string of n bits. We know that the expected running time is $O(n^c)$ (where c is any constant) if the input bits are chosen independently and uniformly at random. What can Markov's inequality tell us about the worst-case running time of this algorithm on inputs of size n? [10]

(Q5) You are given a list of n candidates whom you interview one per day. The recruitment is done according to the following strategy.

Randomly permute the list of candidates and interview the candidates according to this order. Let a be the best candidate among the first $i-1$ candidates. Interview a candidate i. If i is better than a, then fire a and hire i; else retain a. The cost of hiring a candidate is C_h and the cost of interviewing a candidate is C_i; assume C_i is much less than C_h.

Analyze the above randomized algorithm. [5]

(Q6) We define an r-way cut-set in an undirected graph as a set of edges whose removal breaks the graph into r or more connected components. Explain how the randomized min-cut algorithm can be used to find minimum r-way cut-sets, and bound the probability that it succeeds in one iteration. [12]

(Q7) Consider an instance of SAT with m clauses, where every clause has exactly k literals. Design a Las Vegas randomized algorithm that finds an assignment satisfying at least $m(1-2^{-k})$ clauses. Analyze the expected running time of the algorithm. [5 + 10 = 15]

INDIAN STATISTICAL INSTITUTE

Mid-Semestral Examination: 2013-14

Subject Name : **Advanced Cryptography** Maximum Score: 40

Course Name : M.Tech. (CS) II yr. Duration: 3 Hours Date- 24.02.14

Note: Attempt all questions. Marks are given in brackets. Total marks is 48 but you can score maximum 40. Use separate page for each question.

Problem 1 (8). Assume $\mathcal{NP} = \mathcal{BPP}$. Then show that one-way function does not exist.

Problem 2 (6). Prove that if there exists a Las-Vegas algorithm for a language L then the language L must be in $\mathcal{ZPP}$.

Problem 3 (5+7=12). Prove that if $f, g : \{0,1\}^* \to \{0,1\}^*$ are efficiently computable length-preserving functions. If f is bijection and g is weak one-way then show that $f \circ g$ is weak one-way. Is the above true if we replace bijectivity of f by any weak one-way function?

Problem 4 (3+3 = 6). Let $f : \{0,1\}^* \to \mathbb{Z}$ be an efficiently computable function. Let $F(n) = 2^{-n} \sum_{x \in \{0,1\}^n} f(x)$. For any k, construct a PPT A such that

$$Pr[|A(1^n) - F(n)| \leq n^{-k}] \geq 1 - 2^{-n}.$$

Can you construct a sub-exponential algorithm A such that for any negligible function $\epsilon(n)$,

$$Pr[|A(1^n) - F(n)| \leq \epsilon(n)] \geq 1 - 2^{-n}?$$

Problem 5 (4+4 = 8). For any noticeable function η, $\exists$ polynomial p (depending on η) such that $(1-\eta)^p$ is negligible. Conversely, if $\exists$ polynomial p (depending on η) such that $(1-\eta)^p$ is negligible then η is noticeable.

Problem 6 (8). Let f be a one-way function then for any PPT A

$$Pr[A(f(U_n), I) = u_I] \leq 1 - \frac{1}{2n}$$

where $U_n = (u_1, \ldots, u_n)$ and I are indepedently and uniformly distributed over $\{0,1\}^n$ and $\{1, 2, \ldots, n\}$ respectively.

Indian Statistical Institute

Advanced Image Processing

M.Tech.(CS)-II Year: 2013-14

Full marks: 60 Time: 2 Hours

Date: 24.02.2014

Answer **any six** questions. All questions carry equal marks.

1. a) Assuming pin-hole camera model for perspective projection from 3-D to 2-D, prove that a set of parallel straight lines not perpendicular to z-axis is mapped to a set of concurrent straight lines.

   b) What is this common point called? [9+1=10]

2. a) State three basic principles of photometric model of image formation.

   b) Derive the following image formation equation:

$$g(x,y) = \int_{\infty}^{\infty} \int_{\infty}^{\infty} h(x-\alpha, y-\beta) f(\alpha,\beta) d\alpha d\beta + \eta(x,y)$$

   All terms have their usual meaning. [7+3=10]

3. Derive the expression of Wiener filter for image restoration using minimum mean-square estimation approach. [10]

4. a) State and prove the correlation theorem.

   b) Derive the Fourier transform of the Laplacian of a two-variable function $f(x,y)$. Assume that x and y are continuous variables. [7+3=10]

5. a) Describe the Hotelling transform and show that it is optimal in the least-square-error sense.

   b) Calculate the sequency of each column of Hadamard matrix of order 8.

   [(5+2)+3=10]

6. Describe the thresholding method proposed by N. Otsu. How do you extend this method to obtain multiple thresholds? [7+3=10]

7. Describe the Fast Fourier Transform algorithm and discuss its computational complexity. [8+2=10]

8. a) How do you rotate an image $f(x, y)$ using Hotelling transform?

   b) Prove that the Fourier transform of an image $f(x, y)$ is rotated by an angle θ if $f(x, y)$ is rotated by the same angle. [5+5=10]

INDIAN STATISTICAL INSTITUTE

Mid-Semester Examination: 2014 (Second Semester)
Course Name: M. Tech. (CS) 2nd Year
Subject Name: Natural Language Processing
Date: 24.02.2014 Maximum Marks: 50 Duration: 2 hours

Note: Open Book/Class-note Exam.

Q1. [15 marks]

Transliteration is the conversion of a text from one script to another. For instance, Tajmahal in Roman (English) is written as ताजमहल in Devanagari (Hindi). In case of machine translation or cross lingual information retrieval, etc., transliteration plays important role for handling proper names. Assume you have long list of proper names written in Devanagari (Hindi) and a list of many words transliterated in Roman (English) – Devanagari (Hindi) pairs as follows:

john => जॉन ; obama => ओबामा ; kamal => कमल ; tajmahal => ताजमहल ;

mall => मॉल ; mata => माता ; etc.

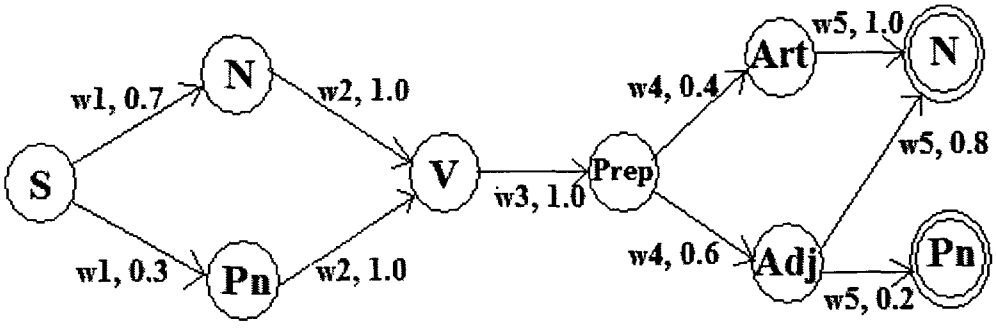
Given these resources, design an approach for Roman (English) - Devanagari (Hindi) transliteration. Write down major steps of your system and assumption on the lexicon length, memory space, etc. A word in Roman (English) will be input to your system and the corresponding transliterated word in Devanagari (Hindi) will be the output.

Q2. [10 marks]

Assume you are dealing with a language where all sentences are of five words/tokens and only six possible POS tags are there <N, Pn, V, Prep, Art, Adj>. Two groups design two different POS tagging systems with the following assumptions as follows:

Group 1, from the grammatical knowledge of the language, assumes that (i) V cannot occur in the beginning or at the end, (ii) N and Pn can occur only in the beginning or at the end, (iii) after V at least one Prep/Art/Adj should appear.

Group 2, from many tagged sentences, assumes the following Markov Chain [S: start state; double circled states are the final states] for doing the said POS tag.



You are asked to compare the two methods without testing them on a common test dataset. Do a quantitative analysis to do this comparison.

Q3. [15 marks]

In your smart phone, you have experienced the following [while typing a message you need not type the following underlined letters; they are typed automatically]:

I have ...
You did ...
You should not ...

The essence of this application is: given the previous word/words a list of current words is *predicted* by the system. Therefore, when you type a few initial characters of the current word, the intended word (from the list of predicted words) is immediately *selected.*

While designing this system, (i) which kind of data structures will you use? (ii) Explain the prediction and selection method using the data structures. (iii) Estimate the space requirement (in terms of MB) of your method. [Assume that you are dealing with a lexicon of 100,000 distinct tokens]

Q4. [15 marks]

As an employee of a security services company, you have discovered that a particular group speaks in a very peculiar language among themselves. Their language (L_1) uses 20 Greek letters as alphabet. You are amazed to see that basically they are speaking in English only. Whenever a sender sends any word in L_1, the receiver easily decodes its corresponding English word. You also invent that the group, for their communication, makes use of only 15 English words: {do, bomb, fire, carry, car, suicide, ...}. However there is hardly any one-to-one correspondence from a word in L_1 to the corresponding word in English. You gather sufficient instances of the words in L_1 and the corresponding words in English like the following:

do = {αβγδ; αγδ; αβπγ; πβγδ; απβγδ....}

carry = {θρκλ; θρτκλ; ρτκλ; κθρτ; ....}

bomb = {βθγπ; βθδγπ; βδγπ; θδγπ; ....}

From these instances, you plan to design an HMM based decoding system that takes a word in L_1 and generates the corresponding word in English. Answer the following in the context of your HMM design: (i) how many HMMs you design to achieve your goal? (ii) How do you define states of your HMM? (iii) Assuming that your HMM parameters are correctly estimated, how will you decide which English word corresponds to an input word, say "αγδπ" [just discuss about your decision taking module].

INDIAN STATISTICAL INSTITUTE

Mid Semester Examination

M. Tech (CS) - II Year, 2013-2014 (Semester - IV)

Computational Complexity

Date : 28.02.2014 Maximum Marks : 60 Duration : 2 hours

Note: Answer as much as you can, but the maximum you can score is 60.
Write your answers in the same order as that of the questions given below and all parts of a question contiguously.

(Q1) Consider the following algorithm for SAT:
"On input ϕ, try all possible assignments to the variables. Accept if any satisfy ϕ."
This algorithm clearly needs exponential time; thus SAT has exponential time complexity. Therefore, SAT is not in P. Since SAT is in NP, we can conclude that P is not equal to NP.
Argue the validity of this proof for $\mathsf{P} \neq \mathsf{NP}$. [5]

(Q2) Show that the set $\{< M > \mid M$ does not halt on any input$\}$ is not recursively enumerable. [5]

(Q3) Show that for non-negative integers m, the predecessor function $V(m) = m - 1$ if $m > 0$, and 0 for $m = 0$, is primitive recursive. [5]

(Q4) Prove that if $\mathsf{NP} \neq \mathsf{coNP}$, then $\mathsf{P} \neq \mathsf{NP}$. [5]

(Q5) Let A be the language of properly nested parentheses. For example, (()) and (()(()))() are in A but)(is not. Show that
(i) $A \in \mathsf{TIME(n)}$;
(ii)A can be accepted by a single tape Turing machine in $O(n \log n)$;
(iii) $A \in \mathsf{L}$. $[(5 + 3 + 2) = 10]$

(Q6) Show that $\mathsf{TISP}(n, \log n) \neq \mathsf{TIME}(n) \cap \mathsf{SPACE}(\log n)$. [15]

(Q7) Prove that $\mathsf{P}^{TQBF} = \mathsf{NP}^{TQBF}$. [15]

(Q8) Prove that if $3SAT$ is polynomial-time reducible to $\overline{3SAT}$, then $\mathsf{PH} = \mathsf{NP}$. [7]

(Q9) Define $\mathsf{P}_{/\mathrm{poly}}$. Comment on the relationship between P and $\mathsf{P}_{/\mathrm{poly}}$. $[(3 + 5) = 8]$

INDIAN STATISTICAL INSTITUTE
M. Tech (CS) II year : 2013–2014
Quantum Information Processing & Quantum Computation
Periodical Examination

Date: 25. 02. 2014 Maximum Marks: 50 Time: 2 Hours

Answer any part of any question. Maximum marks you can obtain is 50. The paper is of 60 marks.

Please answer all parts of a question at the same place.

1. Explain, with examples, how
   (a) an n-qubit pure state and
   (b) an n-qubit mixed state

   can be represented in matrix form. 5+5 = 10

2. (a) Describe the the Deutsch-Jozsa algorithm.
   (b) How it is related to Walsh spectrum of a Boolean function?
   (c) Consider that a Boolean function $f(x_1, x_2, x_3) = x_1 \oplus x_2 x_3$ is used in the Deutsch-Jozsa algorithm as unitary transform U_f. What do you expect in the output?
   (d) What is the importance of this algorithm in contrast to classical paradigm?

   5+3+5+2 = 15

3. (a) Describe the Grover's algorithm.
   (b) Why is this algorithm important in terms of complexity issues.

   7+8 = 15

4. Consider that there are two identical boxes, each containing three identical qubits. The qubits in one box are $|0\rangle$ and in the other box are $\cos\frac{\pi}{9}|0\rangle + \sin\frac{\pi}{9}|1\rangle$. You are allowed to measure in $\{|0\rangle, |1\rangle\}$ basis (computational basis) only. With what probability can you find out which box contains the $|0\rangle$ qubits? 10

5. Write short notes on
   (a) super dense coding,
   (b) teleportation and
   (c) remote state preparation.

   3+4+3 = 10

INDIAN STATISTICAL INSTITUTE

M. Tech. (CS) II Year (2013-14), II semester

Periodical Examination

ADVANCED PATTERN RECOGNITION

Date: 26.02.14 Duration: 150 minutes Marks: 70

Note: Answer all the questions.

1. Describe a density based clustering algorithm. [5]

2. Describe a data condensation procedure for data mining. [5]

3. (a) Describe Parzen's density estimation procedure.
   (b) Describe its generalization to the multivariate case. [5+7=12]

4. Let $x_1 = (0,0), x_2 = (0,1), x_3 = (1,0)$ and $x_4 = (1,1)$. Let $\theta_1 = \theta_3 = 1$ and $\theta_2 = \theta_4 = 2$. Let θ_i denote the class of x_i for each i. Give the result (maximum number of iterations is 16) for finding the straight line that separates the two classes by applying the Perceptron learning algorithm with learning rate as 0.5, and the initial separating straight line as $x - y = 0$. Write the result of every iteration clearly.
   [15]

5. (a) Describe the k-nearest neighbor density estimation procedure.
   (b) Derive the k-nearest neighbor decision rule using the density estimation procedure.
   (c) Describe an algorithm for reducing the size of the training sample set for k-nn decision rule. [5+6+8=19]

6. Let $f(x,y) = 9x^2 - 12xy + 5y^2 - 10$, where x and y are real numbers. Using gradient descent technique, find a local minima (x_0, y_0) of f. Is (x_0, y_0) global minima of f? Justify your answer. [14]

INDIAN STATISTICAL INSTITUTE

Mid-Semestral Examination of Second Semester (2013-2014)

M.TECH.(CS) II YEAR

Topics in Algebraic Computation

Date: 26.02.2014 Maximum marks: 70 Duration: 2 hours 30 minutes

Note: Each question carries 10 marks. Answer as much as you can. The maximum you can score is 70.

1. Exhibit two matrices which have the same characteristic polynomials but, are not similar.
2. Let A be an $n \times n$ matrix. Describe how the sweep-out method can be used to find a generalised inverse of A and estimate the time complexity of your algorithm.
3. Describe an algorithm to compute the inverse of a permutation matrix and find the corresponding time complexity.
4. Provide an overview of the importance of the (L, U, P) decomposition in relating the computational complexities of different matrix algebra computations.
5. Let

$$A = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 2 & 1 & 0 & 1 \\ 2 & 1 & 2 & -1 \\ 0 & -2 & 1 & -1 \end{pmatrix}.$$

   Find a similarity transform to convert A to a matrix in the upper Hessenberg form and hence compute the characteristic polynomial of A.
6. How is the negative wrapped convolution used in the Schonhage-Strassen integer multiplication algorithm?
7. Suppose we wish to multiply two integer polynomials of degree n, where n is a power of two. Show how to choose m and a suitable ω such that the FFT algorithm can be used for this task in $\mathbb{Z}_m$.
8. Explain what is meant by computing the reciprocal of an n-bit integer. Show that the problem of squaring an integer reduces to that of computing reciprocals.
9. Describe the input and the output of the HGCD algorithm for polynomials. Show that the asymptotic complexity of computing the GCD of two polynomials is same as that of computing the HGCD of these two polynomials.
10. Let $\phi(x)$ and $g(x)$ be two polynomials over some field. It is required to generate all possible a, b and c which defines bi-variate polynomials $T(x, y) = xy + ax + by + c$ such that $\phi(x)$ divides $T(x, g(x))$. Describe a method for achieving this.
    (Hint: Consider $x^i(g(x))^j \bmod \phi(x)$ for $0 \leq i, j \leq 1$ and reformulate the problem in terms of matrices.)
11. Factor $u(x) = x^6 + 3x^5 - 4x^4 + x^2 - 7$ modulo 2 and lift the factorisation to modulo 2^2.

INDIAN STATISTICAL INSTITUTE

Mid-Semester Examination: 2013-2014 M. Tech. (CS) Second Year

VLSI Testing and Fault-tolerance

Date: 28.2.2014 Maximum marks = 30 Credit: 30%

Time: 3 hours

Name:_____ Roll No.:_____

Instructions (Read carefully)

A. This is an **OPEN BOOK/OPEN NOTES** exam. Answer all questions; partial credit may be given for incomplete/incorrect answers.

B. Total points = 40; maximum score = 30.

C. You may write your answer on the test booklet.

---

1. (10 points) Consider the following circuit-under-test (CUT) in Figure 1 with 5 Boolean inputs x_1, x_2, x_3, x_4, x_5, which consists of two 2-input NAND gates, one XOR gate, one 2-input OR gate, producing a Boolean function F. The lines in the circuit are labeled as l_1, l_2, l_3, ..., l_9.

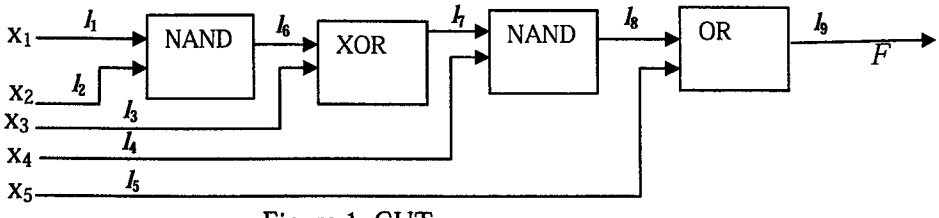


Figure 1. CUT

(a) Write down the output Boolean function F in terms of the primary input variables.

(b) Determine a test vector that detects the fault l_6 stuck-at 0.

(c) Ignore the faults in the interior of XOR blocks and consider single stuck-at 0 and 1 faults only on lines l_1, l_2, l_3, ..., l_9. Write a complete test set T that gives 100% fault coverage. Justify your argument.

(d) Show that under any single stuck-at fault f considered above, the output function F_f in the presence of fault f will contain an even number of true minterms out of 32 input combinations.

(1+1+5+3)

2. (10 points) Referring to the CUT of Figure 1, please answer the following questions.

(a) Compute the Boolean difference (dF/dl_6) of the function F with respect to line l_6, and determine all the test vectors that are capable of detecting l_6 stuck-at 0.

(b) Consider an AND-bridging fault f_b between l_3 and l_6. Derive a test vector for f_b. Justify your argument. (5+5)

3. (12 points) Consider the CUT as shown in Figure 2, with three inputs *A, B, C* and two outputs *S* and *Y*.

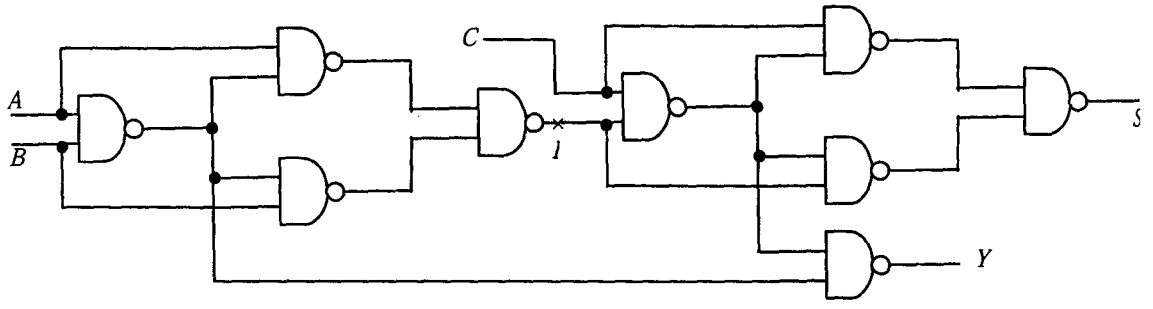


Figure 2. A circuit-under-test (CUT)

(a) List the checkpoints of the circuit with respect to each output.

(b) Consider the fault *l* stuck-at-1. Determine a test vector which detects it at the output *Y*.

(c) Consider an input vector {A B C} = 1 1 0; determine by critical path tracing, the set of all faults detected by this vector 1 1 0.

(d) Modify the circuit (by adding extra logic or extra input/outputs) such that (i) it retains the original behavior of the circuit under functional mode and (ii) for every line *L* in the modified circuit, all paths from *L* have equal inversion parity to at least one of the outputs. Your modification should aim at minimum additional cost/delay.

(e) Show how this CUT can be decomposed into minimum number of sub-circuits such that in each sub-circuit, its inputs are logically independent. How such decomposition help in test generation? Can you outline an algorithm how to achieve such decomposition? (2+2+4+2+2)

4. (8 points) Let *C* be an *n*-input, *m*-output arbitrary combinational circuit consisting of a number of AND, OR, NAND, NOR, and NOT gates, such that the sub-circuit traced back from each primary output to the primary inputs, is fan-out free. Show that no stuck-at fault in *C* is a redundant fault. In other words, for every stuck-at fault, single or multiple, at least one test vector exists. (8)

INDIAN STATISTICAL INSTITUTE

Mid-Semester Examination: (2013-2014)
M.Tech C.S., 2nd Year

Advanced Digital Signal Processing

Date: 28.2.2014 Maximum Marks: 60 Duration: 2 hours

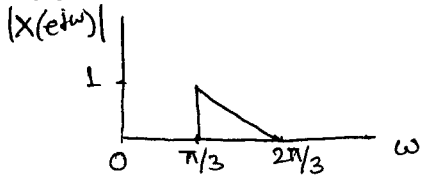
Note: The marks add up to 70. The maximum you can score is 60. The exam is open-book, open-notes. You are permitted to use calculators.

Number of question papers required: 2

Questions:

1. Design a computationally efficient realization of a decimator, using polyphase decomposition, which will decrease the sampling rate by a factor of 2.5. [10]

2. A signal $x[n]$ has the Fourier Transform shown below.



It is downsampled by a factor of 3 and then upsampled by a factor of 3. It finally passes through a filter with frequency response

$$H_1(e^{j\omega}) = \begin{cases} 1 & |\omega| < \pi/3 \\ 0 & \text{otherwise} \end{cases}$$

Sketch the output. Also sketch the output if $x[n]$, passes through a filter with frequency response

$$H_2(e^{j\omega}) = \begin{cases} 1 & |\omega| > 2\pi/3 \\ 0 & \text{otherwise} \end{cases}$$

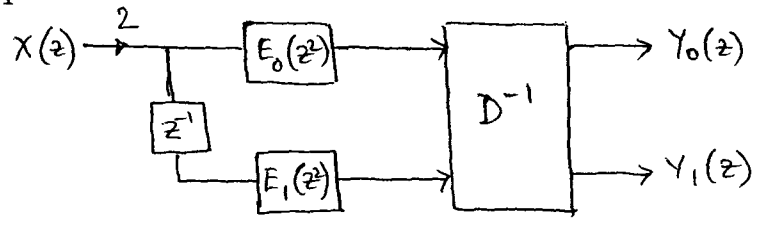
instead of the previous filter. [7.5+7.5]

3. An FIR filter has the transfer function

$$H(z) = -1 + 9z^{-2} + 16z^{-3} + 9z^{-4} - z^{-6}$$

Determine and justify if it is (a)linear phase, (b)half-band. [5+5]

4. A two-channel analysis filter bank has the structure shown with $E_0(z) = E_1(z) = 1$



Let $H_i(z) = Y_i(z)/X(z)$. Determine, $i = 0, 1$

(a) The transfer functions $H_0(z)$ and $H_1(z)$ and sketch their magnitude responses. [5+5]

(b) The transfer functions of the filters of the corresponding synthesis filter bank so that the cascade of the analysis and synthesis filter banks forms a perfect reconstruction system. [5+5]

5. A continuous-time signal consists of two sinusoids, one of 2kHz and the other of 3 kHz. It is sampled at 32 kHz. Using the DFT,

(a) Is it possible to resolve the two sinusoids using 20 samples?

(b) Is it possible to resolve the two sinusoids using a data length of 20 padded with 10 zeros?

(c) Is it possible to resolve the two sinusoids using 40 samples? In each case, justify your answer.

[5+5+5]

Indian Statistical Institute
Mid-Semester Examination (2014)
M.Tech.(CS) II Year
Computer Vision

Date: 28.02.2014 **Full Marks: 60** **Duration – 2 hours**

Answer as many questions as you like, but you can at most score 60.

1. Explain with examples the basic components in any vision system and their functions. Compare between animal and machine vision systems in general with respect to such function modules. 4+6

2. Explain the basic goal of computer vision. How is it different from computer graphics? Explain briefly how are the geometric transformations like Translation, Scaling and Rotation relevant to the computer vision problem? Hence derive the Translation, Scaling and Rotation transform matrices and their inverses respectively.

   2+2+4+12

3. Explain in your own words what do you mean by Perspective transform in the light of the pinhole model. Why do you think parallel tracks appear to converge at a distance? Derive the Perspective matrix and its inverse. Hence explain weak perspective projection and the relevance of orthographic projection.

   4+2+5+4

4. What do you mean by a shift invariant linear system? How can you explain linear image transforms in this light? Explain the advantage/s of Gaussian filtering of an image. What is the significance of combining a Laplacian operator to such Gaussian filtering to extract image properties? Does this have any significance from the angle of biological vision as well – if so, how? 3+3+3+3+3

5. Explain clearly each of the following terms and their interrelation, if any, as well as their significance in the light of the central problem of computer vision: a) illumination, b) reflectance, c) luminance, d) brightness/lightness. What do you mean by the phrase, 'recovering shapes from images'? Which are the cues that are normally helpful in such recovery? 10+2+3

INDIAN STATISTICAL INSTITUTE

Second Semestral Examination (2014)

M.Tech. (Computer Science) Second Year

Natural Language Processing

Date: 21.04.2014 **Time: 2 h 30 min.** **Maximum Marks: 50**

Notes. Answer all questions.
[Open book/note examination]

Question 1.

Suppose the character level bigram statistics for a language L_1 are available on the Web. You are also given a list of twenty thousand (20K) words which are used as named entities (e.g., person names, locations, organizations, etc.) in the language L_1.

With the help of these two language resources, design a bigram statistics based method for detection of named entities (NE) in a given text in L_1. Your method should mark a word (w_i) in the given text as w_i/NE if the word is a named entity. [10]

Question 2.

(a) Consider a simple Probabilistic Context Free Grammar (PCFG) as given below. The non-terminals are S, NP, VP, V, PP, N and P; S being the start symbol. The terminals are the words in italics.

S	→	NP VP	1.0
VP	→	V NP	0.6
VP	→	V PP	0.4
NP	→	N NP	0.5
NP	→	N	0.5
PP	→	P NP	1.0
N	→	*fire*	0.5
N	→	*flies*	0.5
V	→	*flies*	0.5
V	→	*like*	0.5
P	→	*like*	1.0

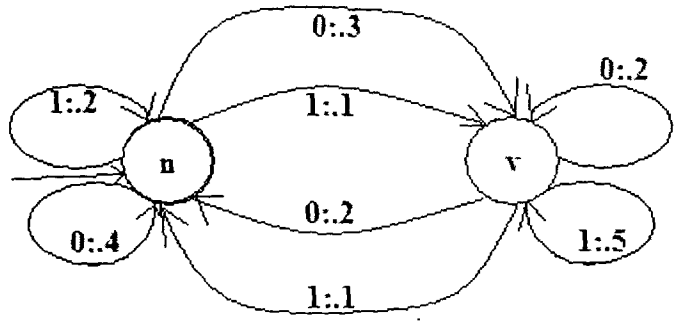
Using the above grammar, show two different parses of the sentence *fire flies like fire*. Compute the probabilities of these two parses. Compute the sentence probability.

[(2 + 2) + (2+2) + 2 = 10]

(b) Suppose for learning the probabilities of the rules of the above grammar, you plan to use a bracketed corpus instead of a corpus of normal sentences, i.e., ((*fire flies*) (*like*) (*fire*)) instead of *fire flies like fire*. Write down the modified probabilities of the rules of the PCFG so that it will give only one parse for the given sentence (i.e., the parse for the given bracketed form of the sentence) not the two parses as you got in part (a). Explain your method. [10]

Question 3.

Consider the following two-state Hidden Markov Model.



Compute the best state sequence when the input is "1110".

[10]

Question 4.

Consider a maximum entropy based text classification problem where articles are to be classified as "financial" or not. Only one feature is used: presence or absence of the word *dividend* in the article. Feature f_1 is 1 if and only if the article is in "financial" and *dividend* occurs. f_2 is the filler feature (refer Text Categorization chapter in Manning and Schutze's book, if needed).

(a) Compute the maximum entropy distribution [by computing p(0, 0), p(0,1), p(1, 0) and p(1, 1)] in form the following equation. Assume that log $\alpha_1 = 2.0$ and log $\alpha_2 = 1.0$.

$$p(\vec{x}, c) = \frac{1}{Z} \prod_{i=1}^{K} \alpha_i^{f_i(\vec{x},c)}$$

(b) Give an example empirical distribution whose maximum entropy distribution corresponds to the one in (a).

[10 + 5 = 15]

8. Consider the following block of gray levels:

6	6	2	3
9	8	4	1
8	2	3	7
4	2	7	8

Calculate the compressed and reconstructed representation of the block using Block Truncation Coding. Calculate PSNR and bpp. [8+2=10]

9. Consider the following digital signal: [1, 2, 4, 1, -1, -2, -1, 1]. Construct the tree wavelet expansion of this signal using the following wavelet filter: [0.1294, 0.2241, -0.8365, 0.4830]. [10]

10. Consider the following block of gray levels:

0	0	1	1
0	0	1	1
0	2	2	2
2	2	3	3

Construct the gray level co-occurrence matrices for angle $\theta = 0°, 45°, 90°, 135°$, considering unit pixel distance, and compute the angular second moment for each case. [(4x2)+2=10]

11. Compute the time dispersion and spectral bandwidth of the following Gaussian signal: $f(t) = e^{-\frac{t^2}{2\sigma^2}}$. Prove that the signal $f(t)$ achieves the minimum of the uncertainty inequality. You may use the following two results:

$$(i) \int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}; \text{ and } (ii) \int_{-\infty}^{\infty} x^2 e^{-x^2} dx = \frac{\sqrt{\pi}}{2}.$$

[(4+4)+2=10]

12. (a) Describe the HSI model for color image processing.

(b) Write down the expressions for converting colors from RGB to HSI models and HSI to RGB models.

(c) What is pseudocolor image processing? [2+(3+3)+2=10]

Indian Statistical Institute

Advanced Image Processing

M.Tech.(CS)-II Year, 2013-2014

Full marks: 100 Time: 3.5 Hours

Date: 21.04.2014

Answer any **ten** questions. All questions carry equal marks.

1. (a) Define hit-and-miss operator. State the condition to be satisfied by the structuring elements.

   (b) Prove that dilation and erosion are dual operators. [2+8=10]

2. (a) Define multi-scale structuring element in (i) analog domain and (ii) discrete domain.

   (b) Prove that open operator is idempotent. [2+8=10]

3. (a) State and prove the convolution theorem.

   (b) Prove that the origin of the Fourier transform of an image $f(x, y)$ can be moved to the center of its corresponding $N \times N$ frequency square by multiplying $f(x, y)$ by $(-1)^{x+y}$. [(2+5)+3=10]

4. (a) Define (i) principal axis of an image and (ii) bi-linear interpolation.

   (b) If $\overline{m}_{ij}$ denotes the (i, j)th central moment of an image f, and θ represents the slope of the principal axis, then prove that $\tan 2\theta = \frac{2\overline{m}_{11}}{\overline{m}_{20} - \overline{m}_{02}}$.

   [(2+2)+6=10]

5. (a) What are active contours? Why are they used?

   (b) Describe internal and external energy functionals in image forces. [4+6=10]

6. Derive the expression of parametric Wiener filter for image restoration using constrained least square estimation approach. [10]

7. What are Von Neumann Neighborhood and Moore Neighborhood with respect to cellular automata? Write the Grow-Cut segmentation algorithm using cellular automata. [2+8=10]

INDIAN STATISTICAL INSTITUTE

Semestral Examination: 2013 – 14

MTech CS (2nd Year)

Computational Finance

Answer ALL questions

Date: 21 April 2014 Maximum Marks: 50 Duration: 3 Hours

1. Critically explain any TWO of the following concepts: [4 X 2 = 8]
   a) Martingale
   b) State price density
   c) Filtration
   d) Stopping time.

2. a) Define the following option contracts:
   (i) Asian
   (ii) Lookback
   (iii) Barrier
   (iv) Chooser.

   For each of them, state the payoff function carefully, explaining all notation.

   [4 X 3 = 12]

3. Suppose $S_0 = 5$, $T = 3$, $r = 0$, $u = 1.5 = 1/d$ are the parameters for a Binomial model. Compute the prices of the following options:
   i) Asian Put option with exercise price = 5
   ii) Up-an-Out Barrier Call with Barrier = 8 and exercise price = 4.5
   iii) American Put option with exercise price = 6. [3 X 6 = 18]

4. Describe a common model for algorithmic trading which uses Bayesian learning. Explain all the relevant terminology. Also mention the order of computation in each time step.

   [5 + 3 + 4 = 12]

INDIAN STATISTICAL INSTITUTE

End Semestral Examination

M. Tech (CS) - II Year, 2013-2014 (Semester - IV)

Topics in Algorithms and Complexity

Date : 21.04.2014 Maximum Marks : 100 Duration : 3.5 Hours

Note: Answer as much as you can, but the maximum you can score is 100.

(Q1) Let $G = (V, \mathcal{E})$ be a graph where we assign any of the three colors {R, B, G} to vertices in G. We say an edge (u, v) is *conflict free* if the vertices u and v are assigned different colors. Consider a coloring that maximizes the number of *conflict free* edges and let this number be c^*. Notice that this is an optimization version of the 3-COLORING problem's decision version which is NP-hard.

Consider the following randomized polynomial time algorithm to solve the above problem approximately. We pick any one of the three colors uniformly at random and color a vertex. We do this for all vertices in G.

Using the above algorithm, what will be the expected number of edges that would be *conflict free*? Explain your result. [10]

(Q2) Consider the randomized algorithm for the two dimensional linear programming studied in class.

(a) Extend the algorithm so that it works for d dimensions, where $d \geq 2$.

(b) Deduce the recurrence for the time complexity as a function of n and d, where n is the number of linear constraints.

[6 + 9 = 15]

(Q3) Consider a weighted version of the MAX SAT (maximum satisfiability) problem with n Boolean variables $(x_1, \ldots, x_n)$, m clauses $(C_1, \ldots, C_m)$ and a positive weight w_i for each clause C_i, $1 \leq i \leq m$. Each clause has at least one Boolean variable. MAX SAT is a NP-Hard problem.

We design a simple randomized approximation algorithm for MAX SAT in the following way. Set each x_i to be TRUE independently with probability $1/2$.

(a) Deduce the approximation ratio of the above randomized algorithm.

(b) Derandomize the above randomized algorithm to obtain a deterministic approximation algorithm with the same approximation ration. Prove your results.

[8 + 12 = 20]

(Q4) **(a)** Show that the expected space requirement of a random skip list for a set S of size n is $O(n)$.

(b) Show that the number of levels r in a random leveling of a set S of size n has expected value $E[r] = O(\log n)$. A *leveling* with r levels of an ordered set S is a sequence of nested subsets (called levels) $L_r \subseteq L_{r-1} \subseteq \cdots \subseteq L_2 \subseteq L_1$ such that $L_r = \emptyset$ and $L_1 = S$.

(c) Show that $r = O(\log n)$ with high probability.

[4 + 4 + 7 = 15]

(Q5) A *dominating set* of an undirected graph $G = (V, E)$ is a set $U \subseteq V$ such that every vertex $v \in V - U$ has at least one neighbor in U. For a graph G with $|V| = n$, and minimum degree $\delta > 1$, show that G has a dominating set of at most $n\,[1 + \ln(\delta + 1)]/(\delta + 1)$ vertices. [15]

[Hints: You can try using techniques of probabilistic methods. Pick each vertex with certain probability p to form a set X. X may not be the dominating set. What are the other vertices you need to add to X to form a dominating set? Let that set be Y. Show that $X \cup Y$ forms a dominating set. $|X \cup Y|$ will be the size of the dominating set and would be a function of n and p. We would obviously want a dominating set of minimum size. From this idea, you can possibly find p.]

(Q6) Let $\mathcal{E}_1, \ldots, \mathcal{E}_n$ be a set of events in an arbitrary probability space, and let $G = (V, E)$ be the dependency graph for these events. Assume there exist $x_1, \ldots, x_n \in [0, 1]$ such that, for all $1 \le i \le n$,

$$\Pr(\mathcal{E}_i) \le x_i \prod_{(i,j) \in E} (1 - x_j).$$

Then, show that

$$\Pr\left(\bigcap_{i=1}^{n} \overline{\mathcal{E}_i}\right) \le \prod_{i=1}^{n} (1 - x_i).$$

[20]

(Q7) Use the Lovasz local lemma to show that, if $4\binom{k}{2}\binom{n}{k-2}2^{1-\binom{k}{2}} \le 1$, then it is possible to color the edges of K_n with two colors so that it has no monochromatic K_k subgraph. [10]

(Q8) Consider a range space $\mathcal{S} = (X, \mathcal{R})$ formed from an undirected graph $G = (V, E)$ as follows. $X = E$ and $\mathcal{R}$ is the set of all *stars* in G; a *star* is the set of edges incident to a vertex. What is the VC-dimension of $\mathcal{S}$? Prove your result. [2 + 8 = 10]

(Q9) **(a)** Consider a range space $\mathcal{S} = (X, \mathcal{R})$ with $VC(\mathcal{S}) = d \ge 2$, where $VC(\mathcal{S})$ denotes the VC-dimension of $\mathcal{S}$. Let $\mathcal{S}_h = (X, \mathcal{R}_h)$ be the range space on X in which $\mathcal{R}_h = \{(r_1 \cap \ldots \cap r_h) \mid r_1, \ldots, r_h \in \mathcal{R}\}$. Show that $VC(\mathcal{S}_h) \le 2dh \log(dh)$.

(b) Consider the range space $S = (\mathbb{R}^d, C_h)$, where C_h is the set of all convex d-polytopes with h facets. Show that $VC(S) \le 2(d + 1)h \log((d + 1)h)$. You can assume results discussed in class, but state them explicitly.

[10+5=15]

INDIAN STATISTICAL INSTITUTE

Semestral Examination of Second Semester (2013-2014)

M.TECH.(CS) II YEAR

Topics in Algebraic Computation

Date: 22.04.2014 Maximum marks: 100 Duration: 3 hours

Note : The paper contains 125 marks. Answer as much as you can. The maximum you can score is 100.

1. For the questions below, assume that the underlying field is the field of reals $\mathbb{R}$.
   (a) Show that if V and W are affine varieties, then so are $V \cup W$ and $V \cap W$.
   (b) Is $X = \{(x, x) : x \neq 1\}$ an affine variety?
   (c) Provide bi-variate polynomials f_1 and f_2 such that $\langle f_1, f_2 \rangle \neq I(V(f_1, f_2))$.
   (d) Is $\langle x + xy, y + xy, x^2, y^2 \rangle$ equal to $\langle x, y \rangle$?
   (e) Using graded LEX ordering of monomials, divide $x^3 - x^2y - x^2z + x$ by $f_1 = x^2y - z$ and $f_2 = xy - 1$.
   (f) If $I \neq \{0\}$ is an ideal and $\langle \mathrm{LT}(I) \rangle = \langle \mathrm{LT}(g_1), \ldots, \mathrm{LT}(g_t) \rangle$, then show that $I = \langle g_1, \ldots, g_t \rangle$.

   $(6 \times 5 = 30)$

2. Assume $\mathbb{F}$ to be the underlying field for the questions below.
   (a) Describe Buchberger's algorithm to compute a Groebner basis for an ideal and show that the algorithm terminates.
   (b) Show that a reduced Groebner basis of an ideal is unique.
   (c) Let I be an ideal of $\mathbb{F}[x_1, \ldots, x_n]$ and let G be a Groebner basis of I with respect to the LEX order where $x_1 > x_2 > \cdots > x_n$. Then show that for every $0 \leq \ell \leq n$, the set

   $$G_\ell = G \cap \mathbb{F}[x_{\ell+1}, \ldots, x_n]$$

   is a Groebner basis of the ℓ-th elimination ideal of I.

   $(3 \times 10 = 30)$

3. (a) Give an example of a lattice $\Lambda \subset \mathbb{Z}^2$ which is of rank 2 and $\Lambda \neq \mathbb{Z}^2$.
   (b) Let Λ be a lattice and $b_1, \ldots, b_n \in \Lambda$. Show that $\{b_1, \ldots, b_n\}$ is a basis for Λ if and only if $\mathcal{P}(b_1, \ldots, b_n) \cap \Lambda = \{0\}$.
   (c) Define the successive minima $\lambda_i(\Lambda)$, $i = 1, \ldots, n$ for a lattice Λ of rank n.
   (d) For any full-rank lattice $\Lambda \subset \mathbb{R}^n$ and (measurable) set $S \subset \mathbb{R}^n$ with $\mathrm{vol}(S) > \det \Lambda$, then exists two distinct points $z_1, z_2 \in S$ such that $z_1 - z_2 \in \Lambda$.
   (e) Define the dual of a lattice (not necessarily of full rank) and provide a geometric interpretation of the dual.

   $(5 \times 5 = 25)$

4. (a) Define an LLL-reduced basis and the LLL-algorithm to compute it.
   (b) Provide a geometric interpretation of Babai's nearest plane algorithm.

(c) Provide the basic idea for using the LLL-algorithm to compute small roots of a low degree polynomial.

(d) Show that the search version of the CVP problem can be solved in polynomial time if and only if the decisional version of the CVP problem can be solved in polynomial time.

$(4 \times 10 = 40)$

INDIAN STATISTICAL INSTITUTE
M. Tech (CS) II year : 2013–2014
Quantum Information Processing & Quantum Computation
Semestral Examination

Date: 25. 04. 2014 Maximum Marks: 100 Time: 3 Hours
Answer any 5 questions.

1. Explain pure and mixed quantum states with the help of density matrices. Provide necessary examples with one and two qubit states. [20]

2. (a) Briefly describe Pauli-X, Pauli-Y, Pauli-Z, Phase, $\frac{\pi}{8}$ and Hadamard gates.
   (b) Explain how the state $\frac{|01\rangle+|10\rangle}{\sqrt{2}}$ can be prepared with one or more of the above gates, considering $|0\rangle$ as the input.

   [12+8 = 20]

3. (a) Briefly outline the BB84 Quantum Key Distribution Protocol.
   (b) Consider that the communicated qubits are intercepted by an eavesdropper and measured in some orthogonal basis. What kind of error will be reflected to the receiver? Will the eavesdropper be able to extract any information?

   [10+(5+5) = 20]

4. (a) Define the Quantum Fourier Transform (QFT).
   (b) Describe the quantum algorithm for obtaining the QFT.

   [5+15 = 20]

5. (a) Explain how order finding is related to factorization.
   (b) Briefly explain the quantum algoritm for order finding.

   [10+10 = 20]

6. (a) Describe the quantum circuit for implementing the three-qubit bit-flip code (error correcting code).
   (b) Can it correct more than one error? Justify.

   [15+5 = 20]

7. (a) Explain how the Deutsch-Jozsa algorithm can be related to the Walsh transform of a Boolean function.
   (b) Describe how the Grover's algorithm can be exploited for amplitude amplification in Quantum domain.

   [10+10 = 20]

INDIAN STATISTICAL INSTITUTE

End Semester Examination

M. Tech (CS) - II Year, 2013-2014 (Semester - IV)

Computational Complexity

Date : 25.04.2014 Maximum Marks : 100 Duration : 3 hours

Answer as much as you can, but the maximum you can score is 100.
Write your answers to all parts of a question contiguously.

(Q1) (a) Is it decidable whether a Turing machine takes more than 2014 steps on some input? Justify your answer.

(b) A k-place boolean predicate P of k variables is said to be primitive recursive if and only if its characteristic function f is given by:

$$f(x_1, x_2, \ldots, x_k) = 1 \text{ if } P(x_1, x_2, \ldots, x_k) = \text{true}, \quad (1)$$

$$= 0 \text{ otherwise.} \quad (2)$$

Show that if Q and R are a k-place and an l-place primitive recursive predicate respectively, then so is their disjunction.

[5 + 5 = 10]

(Q2) (a) A DNF formula in (boolean) variables $x_1, x_2 \ldots, x_n$ is of the form

$$\phi = D_1 \vee D_2 \cdots \vee D_m$$

where for $1 \leq i \leq m$, $D_i = y_{i1} \wedge y_{i2} \cdots \wedge y_{ik}$, and each literal y_{ij} is one of the n variables or its negation. To which complexity class does deciding whether a DNF formula is satisfiable belong?

(b) Illustrate how to arithmetize a boolean formula.

[5 + 5 = 10]

(Q3) (a) Define space-constructibility. Give an example.

(b) Let $BIPARTITE = \{< G > \mid G \text{ is a bipartite graph}\}$. Prove that $BIPARTITE \in$ NL.

[(3 + 2) + 10 = 15]

(Q4) (a) Argue that $EXACT\text{-}INDSET \in \Sigma_2^p$.

(b) Prove or disprove that $EXACT\text{-}INDSET \in \Pi_2^p$.

(c) Sketch the proof that APSPACE = EXP.

[4 + 5 + 6 = 15]

(Q5) (a) Define (i) the class $\mathsf{P}_{/\text{poly}}$, and (ii) a P-uniform circuit family. How are these two sets related?

(b) The n-input function $majority_n : \{0,1\}^n \rightarrow \{0,1\}$ outputs 1 iff at least half of the input variables are 1's. Show that $majority_n \in \mathsf{P}_{/\text{poly}}$. What are the size and depth complexity of $majority_n$?

(c) Prove that for every $n > 1$, there exists a function $f : \{0,1\}^n \rightarrow \{0,1\}$ that cannot be computed by a circuit C of size $2^n/(dn)$ where $d > 1$ is an appropriate constant.

$[(3 * 2 + 3) + 6 + 5 = 20]$

(Q6) (a) Show that $\mathsf{ZPP} = \mathsf{RP} \cap \mathsf{coRP}$.

(b) Show that $\mathsf{BPP} \subseteq \mathsf{PSPACE}$.

$[8 + 6 = 14]$

(Q7) (a) Show that the Fredkin gate is its own inverse.

(b) Given a n-vertex graph G, the *triangle problem* is to decide whether G contains a triangle. The graph is specified by a black box that, for any pair of vertices of G, returns a bit indicating whether those vertices are adjacent in G.

(i) What is the query complexity (worst case number of queries made) in classical computation for this problem?

(ii) An edge is said to be a triangle edge if it is part of a triangle in G. What is the quantum query complexity of deciding whether a particular edge of G is a triangle edge?

(iii) Using Grover's search algorithm, sketch how to solve the triangle problem.

(iv) What is the time complexity of your algorithm?

(c) Present Shor's order-finding algorithm for a given integer N.

(d) How is the class BQP related to the classes NP and BPP?

$[3 + (2 + 3 + 4 + 2) + 5 + 4 = 23]$

(Q8) (a) Show that $\mathsf{dIP} = \mathsf{NP}$.

(b) Define a PCP-verifier. Which PCP-verifier is equal to the class EXP?

(c) Establish that the PCP theorem implies that there exists an $\epsilon > 0$ such that there is no $(1 - \epsilon)$-factor approximation algorithm for Max3SAT, unless NP = P.

$[5 + (3 + 2) + 8 = 18]$

INDIAN STATISTICAL INSTITUTE

End Semester Examination

M. Tech (CS) - II Year, 2013-2014 (Semester IV)

Computational Complexity

Appendix: Relevant Class Definitions

Date : 25.04.2014 Maximum Marks : 100 Duration : 3 Hours

Definition 1 *(The Class* DTIME*) Let* $T : \mathbb{N} \to \mathbb{N}$ *be some function. We let* $\mathsf{DTIME}(T(n))$ *be the set of all boolean functions that are computable in* $d \cdot T(n)$*-time for some constant* $d > 0$.

Definition 2 *(The Class* P*)* $\mathsf{P} = \bigcup_{c \geq 1} \mathsf{DTIME}(n^c)$.

Definition 3 *(The Class* NP*) A language* $L \subseteq \{0,1\}^*$ *is in* NP *if there exists a polynomial* $p : \mathbb{N} \to \mathbb{N}$ *and a polynomial-time TM* M *such that for every* $x \in \{0,1\}^*$,

$$x \in L \iff \exists u \in \{0,1\}^{p(|x|)} \text{ such that } M(x,u) = 1$$

If $x \in L$ *and* $u \in \{0,1\}^{p(|x|)}$ *satisfy* $M(x,u) = 1$*, then we call* u *a certificate for* x *(w.r.t. language* L *and machine* M*).*

Definition 4 *(The Class* NTIME*) For every function* $T : \mathbb{N} \to \mathbb{N}$ *and* $L \subseteq \{0,1\}^*$*, we say that* $L \in \mathsf{NTIME}(T(n))$ *if there is a constant* $c > 0$ *and a* $cT(n)$*-time NDTM* N *such that for every* $x \in \{0,1\}^*$, $x \in L \iff M(x) = 1$.

Definition 5 *(The Class* NP*)* $\mathsf{NP} = \bigcup_{c \in \mathbb{N}} \mathsf{NTIME}(n^c)$.

Definition 6 *(The Class* coNP*)* $\mathsf{coNP} = \{L \mid \overline{L} \in \mathsf{NP}\}$

Definition 7 *(Another Definition of* coNP*) For every* $L \subseteq \{0,1\}^*$*, we say that* $L \in \mathsf{coNP}$ *if there exists a polynomial* $p : \mathbb{N} \to \mathbb{N}$ *and a polynomial-time TM* M *such that for every* $x \in \{0,1\}^*$,

$$x \in L \iff \forall u \in \{0,1\}^{p(|x|)} \text{ such that } M(x,u) = 0$$

Definition 8 *(The Class* EXP*)* $\mathsf{EXP} = \bigcup_{c \geq 0} \mathsf{DTIME}(2^{n^c})$

Definition 9 *(The Class* NEXP*)* $\mathsf{NEXP} = \bigcup_{c \geq 0} \mathsf{NTIME}(2^{n^c})$

Definition 10 *(The Class* SPACE*) Let* $S : \mathbb{N} \to \mathbb{N}$ *and* $L \subseteq \{0,1\}^*$*. We define* $L \in \mathsf{SPACE}(S(n))$ *if there is a constant* c *and a TM* M *deciding* L *such that on every input* $x \in \{0,1\}^*$*, the total number of locations on the read/write tape that are at some point non-blank during* M*'s execution on* x *is at most* $c \cdot S(|x|)$.

Definition 11 *(The Class* NSPACE*) In the above definition, replace* SPACE *with* NSPACE *and the TM with NDTM.*

Definition 12 *(The Class* PSPACE*)* $\mathsf{PSPACE} = \bigcup_{c>0} \mathsf{SPACE}(n^c)$. *The class* PSPACE *is an analog of the class* P.

Definition 13 *(The Class* NPSPACE*)* $\mathsf{NPSPACE} = \bigcup_{c>0} \mathsf{NSPACE}(n^c)$. *The class* NPSPACE *is an analog of the class* NP.

Definition 14 *(The Class* L*)* $\mathsf{L} = \mathsf{SPACE}(\log n)$.

Definition 15 *(The Class* NL*)* $\mathsf{NL} = \mathsf{NSPACE}(\log n)$.

Definition 16 *(An Alternate Definition of the Class* NL*) A language* $L \subseteq \{0,1\}^*$ *is in* NL *if there exists a polynomial* $p : \mathbb{N} \to \mathbb{N}$ *and*

- *a deterministic TM* M *using at most* $O(\log |x|)$ *space on its read/write tape for every input* x *and*
- M *has a certificate tape,*

such that for every $x \in \{0,1\}^*$,

$$x \in L \iff \exists u \in \{0,1\}^{p(|x|)} \text{ such that } M(x,u) = 1$$

$M(x,u)$ *denotes the output of* M *where* x *is placed on its input tape and* u *is placed on its certificate tape.*

Definition 17 *(Polynomial Hierarchy) For* $i \geq 1$, *a language* L *is in* Σ_i^p *if* $\exists$ *a poly-time TM* M *and a polynomial* q *such that*

$$\begin{aligned} x \in L \iff & \exists u_1 \in \{0,1\}^{q(|x|)} \forall u_2 \in \{0,1\}^{q(|x|)} \ldots \\ & Q_i u_i \in \{0,1\}^{q(|x|)} M(x, u_1, u_2, \ldots, u_i) = 1 \end{aligned}$$

where Q_i *denotes* $\forall$ *or* $\exists$ *depending on whether* i *is even or odd, respectively. The polynomial hierarchy is the set* $\mathsf{PH} = \bigcup_i \Sigma_i^p$.

Definition 18 *(The Class* AP*)* $\mathsf{AP} = \bigcup_c \mathsf{ATIME}(n^c)$.

Definition 19 *(The Class* BPP*, Bounded Error Probabilistic Polynomial Time) For* $T : \mathbb{N} \to \mathbb{N}$ *and* $L \subseteq \{0,1\}^*$ *we say that a PTM* M *decides* L *in time* $T(n)$ *if for every* $x \in \{0,1\}^*$, M *halts in* $T(|x|)$ *steps irrespective of its random choices, and* $Pr[M(x) = L(x)] \geq \frac{2}{3}$, *i.e.*

$$\forall x \in L, Pr[M \text{ accepts } x] \geq 2/3 \text{ and}$$

$$\forall x \notin L, Pr[M \text{ rejects } x] \geq 2/3.$$

We let $\mathsf{BPTIME}(T(n))$ *be the class of languages decided by PTMs in* $O(T(n))$ *time and define* $\mathsf{BPP} = \bigcup_c \mathsf{BPTIME}(n^c)$.

Definition 20 *(An Alternate Definition of the Class* BPP*) A language* $L \in \mathsf{BPP}$ *if there exists a poly-time TM* M *and a polynomial* $p : \mathbb{N} \to \mathbb{N}$ *such that for every* $x \in \{0,1\}^*$, $Pr_{r \in_R \{0,1\}^{p(|x|)}}[M(x,r) = L(x)] \geq \frac{2}{3}$ *where* $r \in_R X$ *denotes that* r *was chosen from the sample space* X.

Definition 21 *(The Class* RP*) A language* $L \subseteq \{0,1\}^*$ *is said to be in* $\mathsf{RTIME}(T(n))$ *if there exists a PTM running in time* $T(n)$ *s.t.*

$$\forall x \in L,\ Pr[M(x) = 1] \geq \frac{2}{3}$$

$$\forall x \notin L,\ Pr[M(x) = 0] = 1$$

The class $\mathsf{RP} = \bigcup_{c>0} \mathsf{RTIME}(T(n))$.

Definition 22 *(The Class* coRP*) A language* $L \subseteq \{0,1\}^*$ *is said to be in* coRP *if there exists a PTM running in polynomial time s.t.*

$$\forall x \in L,\ Pr[M(x) = 1] = 1$$

$$\forall x \notin L,\ Pr[M(x) = 0] \geq \frac{2}{3}$$

Definition 23 *(Deterministic Proof Systems) We say that a language* L *has a* k*-round deterministic interactive proof system if there is a deterministic TM* V *that on input* $x, a_1, \ldots, a_i$ *runs in time polynomial in* $|x|$*, and can have a* k*-round interaction with any function* P *such that*

$$\textit{(Completeness)} \quad x \in L \Rightarrow \exists P : \{0,1\}^* \rightarrow \{0,1\}^* \mathcal{O}_V\langle V, P\rangle(x) = 1$$

$$\textit{(Soundness)} \quad x \notin L \Rightarrow \forall P : \{0,1\}^* \rightarrow \{0,1\}^* \mathcal{O}_V\langle V, P\rangle(x) = 0$$

The class dIP *contains all languages with a* $k(n)$*-round deterministic interactive proof system where* $k(n) = poly(n)$.

INDIAN STATISTICAL INSTITUTE

Final Examination: (2013-2014)
M.Tech C.S., 2nd Year

Advanced Digital Signal Processing

Date: 25.4.2014 Maximum Marks: 100 Duration: 3 hours

Note: The marks add up to 117. The maximum you can score is 100. The exam is open-book, open-notes. You are permitted to use calculators.

Number of question papers required: 2

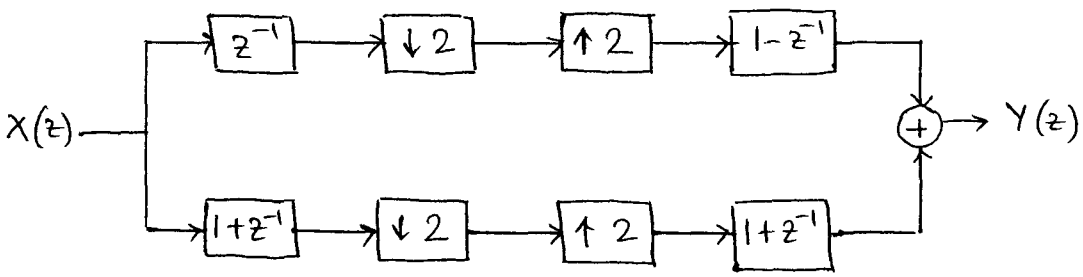
Questions:

1. The sample autocorrelation matrix of a sequence consisting of a large number of samples of a real sinusoid in noise is given by

$$R_{xx} = \begin{pmatrix} 4 & 0 & -1 \\ 0 & 4 & 0 \\ -1 & 0 & 4 \end{pmatrix}.$$

Determine the frequency and amplitude of the sinusoid. [10]

2. For the structure shown below,



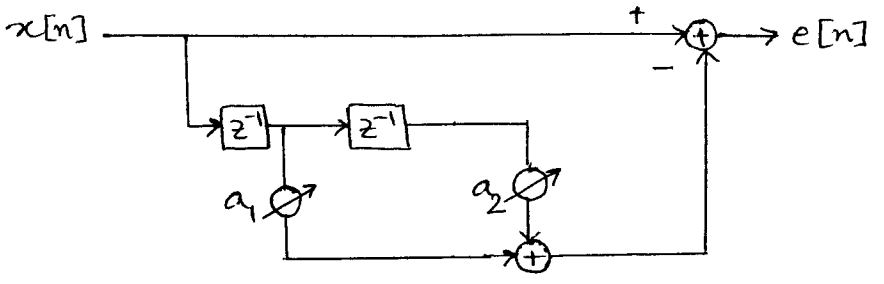
determine the overall transfer function. Is it alias-free? Justify your answer. [7+3]

3. An MA(1) process has the autocorrelation sequence

$$\gamma_{xx}(m) = \begin{cases} 6 & m=0 \\ -2 & |m|=1 \\ 0 & \text{otherwise} \end{cases}$$

Determine the coefficients of a minimum phase system that produces this process from zero mean white noise with variance 1. [7]

4. The adaptive predictor shown below



has the input

$$x[n] = \sin(n\pi/4) + w[n],$$

where $w[n]$ is a white noise sequence of variance 0.1. Obtain the expression for the performance surface, the optimum coefficients and the expression for the algorithm with μ at one-fourth of its maximum value. [5+5+5]

5. Obtain the lattice-ladder filter structure for the system with system function

$$H(z) = \frac{-4 - 1.2z^{-1} + 0.75z^{-2}}{1 - 0.3z^{-1} + 0.5z^{-2}}.$$

[10]

6. An ARMA process is generated by the difference equation

$$x[n] = 0.8x[n-1] + w[n] + 0.6w[n-1],$$

where $w[n]$ is white noise with mean zero and variance σ_w^2. Determine

(a) the system function of the whitening filter, and

(b) the power density spectrum of $x[n]$.

[5+5]

7. Using polyphase decomposition and the 'noble identities', give a computationally efficient realization of a factor-of-5 interpolator using a length 15 linear phase FIR filter. Compare the number of multiplications required by your design to that of a direct implementation.

[10+5]

8. A signal $x[n] = s[n] + w[n]$ forms the input to a Wiener filter which must estimate $s[n]$. $w[n]$ is a white noise sequence with variance 1 and uncorrelated with $s[n]$. The sequence $s[n]$ is generated as

$$s[n] = 0.6s[n-1] + v[n],$$

where $v[n]$ is a white noise sequence with variance 0.16 and uncorrelated with $w[n]$.

   (a) Set up the Wiener-Hopf equations for a filter of length 2.

   (b) Determine the optimum weights for the above filter and also the minimum mean squared error.

[10+(5+5)]

9. A band-limited continuous-time signal with a highest frequency component of 4 kHz is sampled at a rate F_s. The sampled signal is then windowed by a length-N rectangular window. For a DFT-based spectral analysis, determine the minimum and maximum values possible for F_s and the DFT length N which must be a power of 2 , if a resolution of at least 10 Hz is required. [5+5]

10. The analysis filters of a three-channel QMF bank are given by

$$[H_0(z)\ H_1(z)\ H_2(z)] = [z^{-2}\ z^{-1}\ 1] \begin{pmatrix} 2 & 4 & 1 \\ -1 & 4 & -2 \\ 2 & -1 & 2 \end{pmatrix}.$$

Determine the synthesis filters for implementing a perfect reconstruction filter bank. [10]

INDIAN STATISTICAL INSTITUTE

End-Semester Examination: 2013-2014 M. Tech. (CS) Second Year

VLSI Testing and Fault-tolerance

Date: 25.4.2014 Maximum marks = 100 Credit: 50%

Time: 3 hours

Name:_____ Roll No.:_____

Instructions (Read carefully)

A. This is an **OPEN BOOK/OPEN NOTES** exam. Answer all questions; partial credit may be given for incomplete/incorrect answers.

B. Total points = 110; maximum score = 100.

---

1. [6 + 4 + (2+3+5) = 20]

(a) Compute the expression coverage for the following expression in a run that has encountered these values for a, b, and c: (1,1,0), (0,0,1), (1,0,1), and (0,1,0):

((a ? b : c) (bc + b′c′))

(b) Draw a ***single*** Kripke structure that satisfies the CTL property AG AF (p) and the LTL property F¬p.

(c) A Gray counter has the property that successive values of the counter differ in only one bit. The exact counting sequence may vary from implementation to implementation. For example, one valid counting sequence for a 3-bit Gray counter is:

000 → 001 → 011 → 010 → 110 → 111 → 101 → 100 → 000 …

Another valid counting sequence for a 3-bit Gray counter is:

000 → 010 → 110 → 100 → 101 → 111 → 011 → 001 → 000 …

Let s [0:2] denote the 3-bit state vector of the counter.

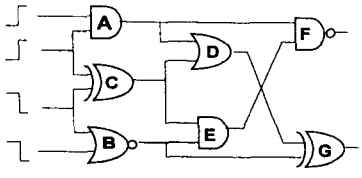
(i) Write the following properties in Linear Temporal Logic (LTL):

- *Successive states of the counter differ in exactly one bit.*
- *A visited state is re-visited every 8 clock cycles.*

(ii) Do we need any more properties for specifying a 3-bit Gray counter? If so, add more properties. If not, justify that any implementation, which satisfies these two properties is a valid 3-bit Gray counter.

2. [8 + 8 + 4 = 20]

Consider the following circuit and the input events shown in the figure below.



(a) Simulate the above circuit using the event-driven scheme. Construct timing diagrams, a timing wheel, and show how events get queued and dequeued. Assume that each gate has a unit delay.

(b) Show how the same circuit is simulated by a cycle simulator for steady-state evaluation.

(c) Justify the following statement with an example:

An event added to the event queue may or may not happen in future.

3. [8 + 4 + 8 = 20]

(a) Present an algorithm for solving the register correspondence problem and illustrate its working with an example.

(b) If no register correspondence exists, can we conclude that the sequential circuits are not equivalent? If not, give a counter-example.

(c) Explain the working principle of the SAT-based combinational equivalence checking method with an example.

4. [8+ 8 + 9= 25]

Consider the CUT as shown in the figure below, with three inputs *A, B, C* and two outputs *S* and *Y*.

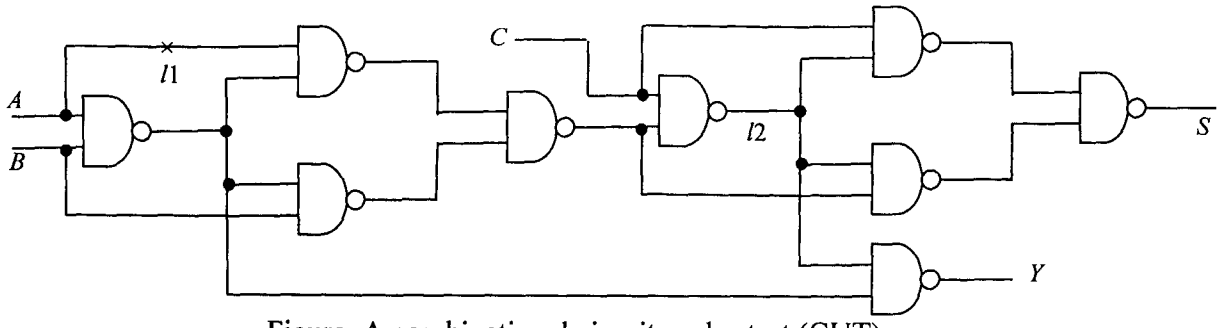


Figure. A combinational circuit-under-test (CUT)

(a) Find a test for the transition fault 0 → 1 at line *l*1.

(b) Find a test for the slow-to-fall path-delay fault along the path A → *l*1 → *l*2 → Y. Is your test robust?

(c) Append a checker circuit with the CUT such that the fault *l*1 s-a-0 and *l*2 s-a-1 become online testable.

5. [7 + 9 + 9 = 25]

Consider the sequential circuit as shown below, in which two clocked D flip-flops are on feedback paths.

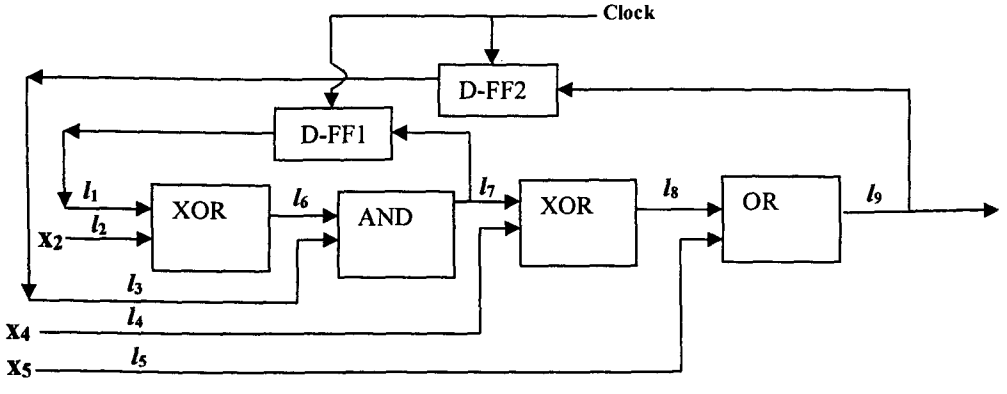


Figure. A sequential CUT

(a) Insert a scan chain and show the modified circuit with scan-in, scan-out and functional/test mode selector line.

(b) Determine a scan test vector for detecting l_6 stuck-at 1; show how the error can be observed by scan-in/scan-out mechanism.

(c) Assume both the FFs are initialized to 0. Show how a slow-rising (0 →1) transition fault at line l_6 can be tested under launch-on-capture (LoC) scheme. Write the test and explain the procedure of observing the error.

Indian Statistical Institute
Mid-Semester Examination (2014)
M.Tech.(CS) II Year
Computer Vision

Date: 25.04.14 **Full Marks: 100** **Duration – 3 hours 15 mins**

Answer as many questions as you like, but you may at most score 100.

1. a) Explain the process of straight line detection in images using the concept of Hough transform.
   b) Discuss its usefulness and limitations with respect to standard edge detection techniques like Canny and Marr-Hildreth operators.
   c) What is generalized Hough transform? **7+8+5=20**

2. a) What do you mean by modeling a camera?
   b) Briefly describe the homogeneous coordinate system and explain its necessity in modeling the camera.
   c) Write short notes on each of the following: i) projective camera, ii) affine camera, iii) weak perspective camera, and iv) orthographic camera. **3+5+(3x4)=20**

3. a) Explain, what do you mean by the terms 'motion field' and 'optical flow' and their importance in computer vision?
   b) Derive the optical flow constraint equation and demonstrate how the flow may be measured.
   c) Describe the aperture problem in the above measurement and hence show how this may be solved by using any one of the standard techniques.
   d) Justify, if the optical flow measurement problem is equally relevant for the human visual system.
   e) Discuss the role of the relevant areas in the brain which are responsible for performing the above mentioned activity. **6+8+8+4+4=30**

4. a) What do you mean by a linear shift invariant system?
   b) How can you explain linear image transforms in the light of the above?
   c) Explain the advantages of Gaussian filtering of an image, and how is it connected to the regularization problem in vision?
   d) What is the significance of combining a Laplacian operator to such Gaussian filtering in order to extract image properties?

... page 2

e) Explain how such a Laplacian of Gaussian operator is significant in biological vision. **5+5+8+6+6=30**

5. a) What do you mean by rotational invariance in digital image processing?
   b) How is the Scale Invariant Feature Transform (SIFT) algorithm related to the above concept?
   c) Hence discuss the importance of SIFT in computer vision, along with its biological motivation. **4+10+6=20**

INDIAN STATISTICAL INSTITUTE

M. Tech. (CS) II Year (2013-14), II semester

Semestral Examination

ADVANCED PATTERN RECOGNITION

Date: 29-4-14 Duration: 195 minutes Maximum marks: 100

Note: This paper carries 106 marks. Answer as much as you can.

1. Define Fisher's Discrimination Ratio (FDR). How do you find the Principal Direction using FDR for a 2- class classification problem when a training sample set is given? [3+6=9]

2. Suppose you have two m-dimensional normal populations $N(\mu_1, \Sigma)$ and $N(\mu_2, \Sigma)$ where $\Sigma = \sigma^2 I$ with prior probabilities P_1 and P_2 respectively. Given an unlabelled dataset from this setup, how do you estimate $\mu_1, \mu_2, \sigma, P_1$ using EM algorithm? [10]

3. Define a dissimilarity measure between two features. Describe a feature selection method based on the suggested dissimilarity. [4+4=8]

4. Describe a data condensation procedure. [5]

5. Suppose you have two m-dimensional normal populations $N(\mu_1, \Sigma)$ and $N(\mu_2, \Sigma)$. Let the prior probability of the first population be P, where $0<P<1$. Find the Bayes decision rule for separating the two populations and also find its probability of misclassification. [3+7=10]

6. (a) Describe a procedure for clustering using genetic algorithms.
   (b) Suppose you have two strings $a = 001101$ and $b = 001011$. Let the crossover probability be p. Find the probability of obtaining the strings a and b after crossover.
   © What is the probability of obtaining the string *001101* from *011001* by mutation if the mutation probability is q?
   (d) Describe the algorithm for roulette wheel selection scheme.
   [6+4+2+7=19]

7. (a) Define fuzzy c-partition of a dataset.
   (b) Write down the objective function for fuzzy c means (FCM) and the necessary conditions used. Describe the fuzzy c means algorithm.. [5+7=12]

(P.T.O)

8. Suppose there is one hidden layer with J number of nodes in an MLP, and sigmoid function is used as transfer function. Suppose you are using online learning algorithm. Let a training dataset be given to you and let the number of classes be 3. Then
   (a) Write down the expression for the error for MLP.
   (b) Write down the expression for the change in the connection weight joining the i-th node in the hidden layer to the second node in the output layer.

   [3+7=10]

9. Write short notes on the following.
   (a) k-fold cross validation
   (b) VC dimension
   (c) Support vectors [5+5+5=15]

10. Describe k-nearest neighbor probability density estimation procedure. [8]

INDIAN STATISTICAL INSTITUTE

Semestral Examination: 2013-14

Subject Name : **Advanced Cryptography** Maximum Score: 50

Course Name : M.Tech. (CS) II yr. Date: 30th April Duration: 3 Hours

Note: Attempt all questions. Marks are given in brackets. Total marks is 56 but you can score maximum 50. Use separate page for each question.

Problem 1 (4+4 = 8). Justify that one-way function exists if and only if pseudorandom function exists. You may use known facts.

Problem 2 (4+8=12).

1. Define a unique-commitment function C on a single bit σ.
2. Let f be a one-way permutation and b be its hard-core predicate. Define $C(r, \sigma) = (f(r), b(r) \oplus \sigma)$. Prove that $C(r, \cdot)$ is a unique-commitment.

Problem 3 (8). Show that if f_K is a pseudorandom function from D to itself then for independently chosen keys K_1, K_2, the keyed function $f_{K_1} \circ f_{K_2}$ is also a pseudorandom function.

Problem 4 (8). Let $f_K : \{0,1\}^n \to \{0,1\}^n$ be a uniform random function. Distinguish the following keyed function

$$F_{K_1,K_2,K_3}(x, y) = f_{K_3}(f_{K_1}(x) \oplus f_{K_2}(y)), \quad x, y \in \{0,1\}^n$$

from a uniform random function from $2n$ bits to n bits, where K_1, K_2, K_3 are independently chosen.

Problem 5 (5+5=10). Let $\mathcal{L} = \{(G_1, G_2) : G_1$ and G_2 are non-isomorphic groups$\}$ be the language of pair of non-isomorphic groups. Show that $\mathcal{L}$ has an interactive proof system.

Problem 6 (10). Let $G = \langle g \rangle$ be a cyclic group with generator g and $m_0 \neq m_1 \in G$. Assuming that the DDH problem is hard on G, prove that b is computationally independent with $(h, g^r, h^r \cdot m_b)$ where $(r, h, b) \in_R \{1, \cdots, |G|\} \times G \times \{0,1\}$.