

Indian Statistical Institute
Mid-Semestral Examination: 2016
Course Name: M. Tech. in Computer Science
Subject Name: Mobile Computing

Date: 17-09-2016

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. What is the difference between horizontal and vertical handover? Briefly explain the following received signal strength (RSS) based horizontal handover decision strategies: i) relative RSS with threshold, and ii) relative RSS with hysteresis and threshold. Mention three most commonly used performance evaluation metrics for vertical handover decision (VHD) algorithms. Describe a signal to interference plus noise ratio (SINR) based VHD algorithm. [2+(2+2)+3+6=15]
2. State the channel assignment problem (CAP) in the form of a generalized graph coloring problem. Explain with an example the concept of coalesced CAP. Consider the case where solution of the coalesced CAP produces zero call blocking. For this case, how the solution of the original CAP is derived from the solution of the coalesced CAP? What is perturbation-minimizing frequency assignment problem? [3+8+3+4=18]
3. Describe a 1-local distributed channel assignment algorithm for a constrained weighted hexagon graph (G, w, c_0, \dots, c_k) where 1) w is a positive integral weight vector indexed by the nodes of the hexagon graph G , and 2) c_i represents the separation constraint between pairs of nodes at graph distance i of each other, where $0 \leq i \leq k$. Assume $c_0 = a$, $c_1 = a$, $c_2 = b$, $c_i = 0 \forall i \geq 3$ and $a \geq 2b$. What is the performance ratio of your stated algorithm? [10+2=12]
4. State the difference between non-overlapping channels (NOCs) and partially overlapping channels (POCs) in IEEE 802.11 wireless local area network (WLAN)? Explain with an example the concept of *weighted conflict graph* used to model the interference in the presence of NOCs and POCs in WLAN. Describe a greedy algorithm for channel assignment using the concept of weighted conflict graph. [3+4+8=15]
5. Explain the random polling and proportional fair access methods in WLAN. [5]

INDIAN STATISTICAL INSTITUTE

Mid Semestral Examination

M. Tech (CS) - II Year, 2016-2017 (Semester - III)

Optimization Techniques

Date : 19.09.2016

Maximum Marks : 60

Duration : 3.0 Hours

Note: The question paper is of 75 marks. Answer as much as you can, but the maximum you can score is 60.

Vectors would be written in small letters with boldface, e.g. \mathbf{b} ; matrices would be written in capital letters, e.g., A . Transpose of A would be denoted by A^T and transpose of \mathbf{b} would be denoted by \mathbf{b}^T . Whenever we say that, \mathcal{P} is a linear program, we mean \mathcal{P} is of the form

$$\begin{aligned} &\text{Maximize} && \mathbf{c}^T \mathbf{x} \\ &\text{subject to} && A\mathbf{x} \leq \mathbf{b} \\ &&& \mathbf{x} \geq \mathbf{0} \end{aligned}$$

- (Q1) An investment bank is considering investments into 5 projects – A, B, C, D and E. Each project has an initial investment, an expected profit rate and an associated risk of failure expressed as a percentage of initial cost as follows:

	A	B	C	D	E
Investment (in crores)	2	1.2	0.9	2.7	1.8
Profit rate	10%	20%	20%	10%	10%
Risk rate	6%	4%	6%	5%	5%

- (i) Provide a formulation to maximize the total profits such that the bank does not invest more than 8 crores of rupees and the average failure risk is not over 5%.
- (ii) Modify your formulation if the following constraint is there on the projects – if projects A and B are chosen, then D must be chosen.

[7+8=15]

- (Q2) Convert the following program to a linear program in the standard form.

$$\begin{aligned} &\text{Maximize} && |x| + |y| + |z| \\ &\text{subject to} && x + y \leq 1 \\ &&& 2x + z = 3 \end{aligned}$$

[10]

(Q3) Let $f(x) = \max(\mathbf{c}_1^T \mathbf{x} + d_1, \mathbf{c}_2^T \mathbf{x} + d_2, \dots, \mathbf{c}_p^T \mathbf{x} + d_p)$. For such a function f , consider the mathematical program

$$\begin{aligned} &\text{Minimize} && f(x) \\ &\text{subject to} && A\mathbf{x} \leq \mathbf{b} \\ &&& \mathbf{x} \geq \mathbf{0} \end{aligned}$$

Can you convert this mathematical program to a linear program? Explain with proper arguments. [10]

- (Q4) (i) Let \mathcal{D} be the dual of \mathcal{P} . Show that the dual of \mathcal{D} is \mathcal{P} .
(ii) State and prove the weak duality theorem.
(iii) An LP can be any of the following three — feasible and bounded, feasible and unbounded, or infeasible. If the possibilities of the primal \mathcal{P} and the dual \mathcal{D} are paired together, then there can be nine possibilities. Argue with proper reasons which of the nine possibilities can occur.

[8+7+10=25]

(Q5) Use SIMPLEX method to solve the following LP:

$$\begin{aligned} &\text{Maximize} && x_1 + 2x_2 - x_3 \\ &\text{subject to} && 2x_1 + x_2 + x_3 \leq 14 \\ &&& 4x_1 + 2x_2 + 3x_3 \leq 28 \\ &&& 2x_1 + 5x_2 + 5x_3 \leq 30 \\ &&& x_1, x_2, x_3 \geq 0 \end{aligned}$$

[15]

INDIAN STATISTICAL INSTITUTE

Mid-Semester Examination: 2016-2017

M. Tech. (CS) II year

Data Mining and Knowledge Discovery

Date: 20.09.2016

Maximum Marks: 50

Duration: 2 hours

[Answer as much as you can]

1. (i) What do you understand by the "scalability" of an algorithm?
(ii) What is overfitting? How is it different from underfitting?
(iii) Outline some stopping criteria for a classifier.
(iv) Elaborate on some performance measures for evaluating classifiers.

[3+5+4+6=18]

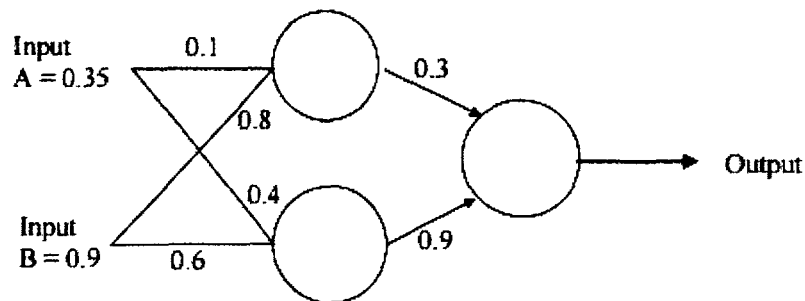
2. (i) How is decision tree different, as a classifier, with respect to neural network?
(ii) What is ROC?
(iii) Explain algorithm SPRINT, pointing out its merits and demerits in the context of data mining.

[4+4+10=18]

3. (i) What is Rainforest? How does the Rainforest framework help in classification?
(ii) What is cross validation?
(iii) What is MDL? Why is it used?

[8+3+5=16]

4. (i) Consider the simple network below:



Assume that the neurons have a sigmoid activation function and

- (a) Perform a forward pass on the network.
- (b) Perform a reverse pass (training) once (target = 0.5).
- (c) Perform a further forward pass and comment on the result.

(ii) The Back-Propagation (BP) algorithm is often used for training feed-forward neural networks. Why do we need to calculate the gradient in the BP algorithm?

P.T.O

(iii) When a BP algorithm is used, the error function must be differentiable. Why?

(iv) Explain why we prefer to use the logistic (sigmoid) function as the activation function in back-propagation networks. Is there any drawback of logistic activation function?

[(2+4+2)+2+2+2=14]

INDIAN STATISTICAL INSTITUTE

Mid-Semester Examination: 2016-2017

M. Tech. (CS) 2nd Year

Artificial Intelligence

Date: 20.09.2016

Maximum Marks: 60

Duration: 2.5 hours

Answer all questions in brief.

1. Prove that *depth-first iterative deepening* algorithm is asymptotically optimal among brute-force tree searches in terms of time, space, and length of solution. Describe the crossover and mutation operators of genetic algorithm.

[3 + 2 + 1) + (2 + 2) = 10]

2. Define monotone property of a heuristic. Prove that any monotonic heuristic is admissible. Prove that the set of states expanded by algorithm A* is a subset of those examined by breadth first search.

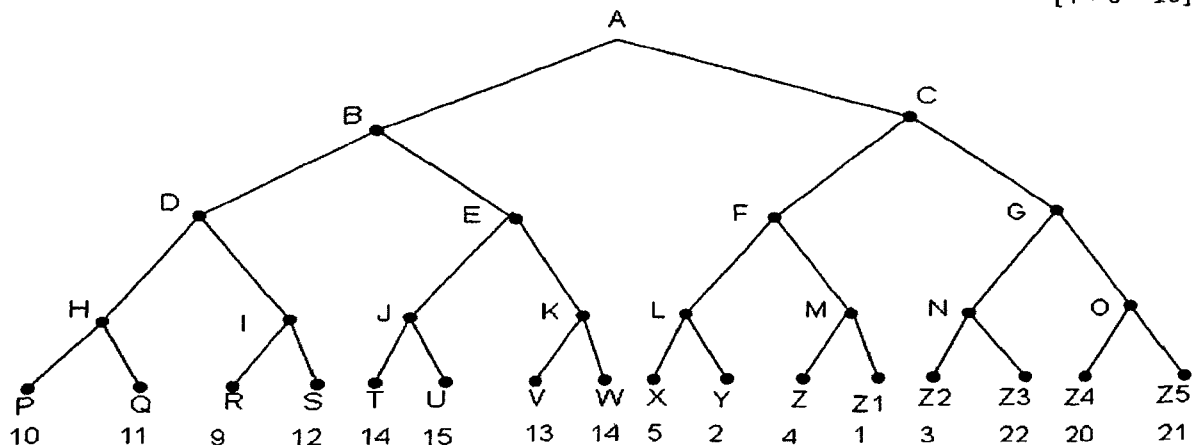
[3 + 4 + 3 = 10]

3. 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 using left-bank and right-bank to denote left and right river banks, respectively.

[10]

4. Perform the *minimax* search procedure on the game tree shown below in which the static scores are all from the first player's point of view and MAX is allowed to move first. Perform the left-to-right α - β pruning procedure on this tree and show how many nodes can be pruned away.

[4 + 6 = 10]



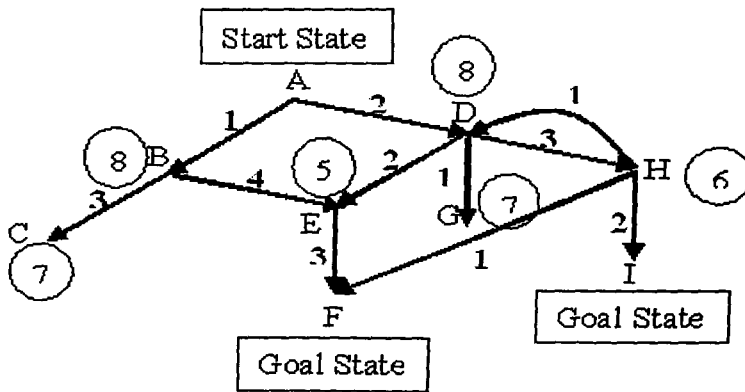
5. Solve the following cryptarithmic problem:

$$\begin{array}{r}
 \text{B A S E} \\
 + \text{B A L L} \\
 \hline
 \text{G A M E S} \\
 \hline
 \end{array}$$

[10]

6. 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]



INDIAN STATISTICAL INSTITUTE
Mid-Semester Examination: 2016
Course Name: M.Tech. In Computer Science
Subject Name: Computer Architecture

Date: 21.09.2016

Maximum Marks: 60

Duration: 3 hours

Answer Question 1 and any 3 from the rest

1. Potpourri (12 Points)**Pipelining (1 Point)**

Circle one of A, B, C, D. As pipeline depth increases, the latency to process a single instruction:

- A. decreases
- B. increases
- C. stays the same
- D. could increase, decrease, or stay the same

Program Counter (2 Points)

In the MIPS ISA, which instruction(s) do not change the program counter?

Endianness (1 Point)

Say that we were to take a processor architecture designed as big-endian and change it to be little-endian. Assuming a memory with a word-level interface, can you mention 2 example MIPS / x86 ISA instructions whose operations will be affected by this endianness change?

ISA vs Microarchitecture (2 Points)

Indicate whether each of the following design choices in a processor is a feature of the ISA or of the microarchitecture:

- Two-level global branch prediction
- Predicated instruction execution
- A floating-point unit that uses wide floating-point values for additional accuracy
- A 32-bit wide data bus to memory
- An additional set of user-visible registers

Interlocking (1 Points)

Remember that MIPS was initially designed such that hardware did not need to provide interlocking, and it was the compiler's job to reorder instructions accordingly to ensure the pipeline operates correctly. Now, consider a pipelined MIPS implementation that implements hardware based interlocking. Could compile-time instruction reordering provide any benefit in this implementation? Why? Why not?

Branch Prediction (2 Points)

A snapshot of the taken/not-taken behavior of a branch is:

T T T T T T T N N T T N N T N N T

If the branch predictor used is a 2-bit saturating counter, how many of the last ten branches are predicted correctly?

Register Dependencies (3 Points)

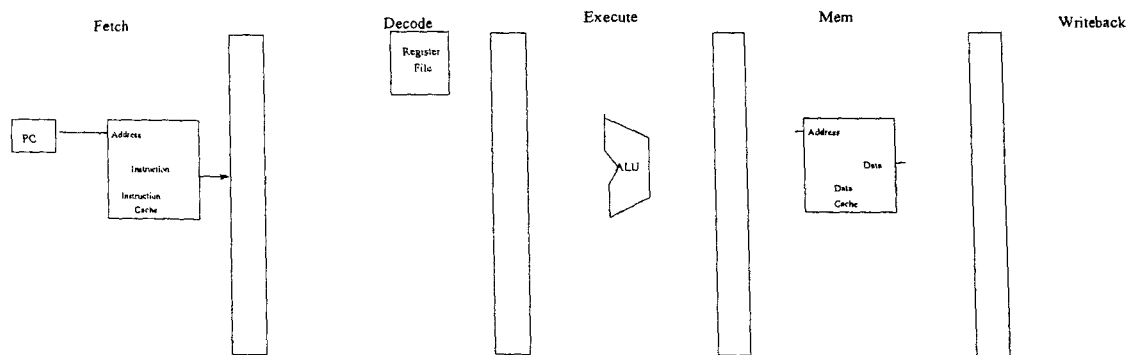
- (a) What is the fundamental cause of false register dependencies (output and anti, or write-after-read, write-after-write dependencies)?
- (b) What can be changed in the ISA, compiler, and microarchitecture to eliminate false dependencies, if at all possible, in each of the three levels above? Describe one disadvantage of each approach.

2. Fine-Grained Multithreading (8 X 2 = 16 Points)

Consider a design “Machine I” with five pipeline stages: fetch, decode, execute, memory and write-back. Each stage takes 1 cycle. The instruction and data caches have 100% hit rates (i.e., there is never a stall for a cache miss). Branch directions and targets are resolved in the execute stage. The pipeline stalls when a branch is fetched, until the branch is resolved. Dependency check logic is implemented in the decode stage to detect flow dependences. The pipeline does not have any forwarding paths, so it must stall on detection of a flow dependence.

In order to avoid these stalls, we will consider modifying Machine I to use fine-grained multithreading.

- (a) In the five stage pipeline of Machine I shown below, clearly show what blocks you would need to add in each stage of the pipeline, to implement fine-grained multithreading. You can replicate any of the blocks and add MUXes. You don't need to implement the mux control logic (although provide an intuitive name for the mux control signal, when applicable).



- (b) The machine's designer first focuses on the branch stalls, and decides to use fine-grained multithreading to keep the pipeline busy no matter how many branch stalls occur. What is the minimum number of threads required to achieve this? Why?
- (c) The machine's designer now decides to eliminate dependency-check logic and remove the need for flow-dependence stalls (while still avoiding branch stalls). How many threads are needed to ensure that no flow dependence ever occurs in the pipeline? Why?

A rival designer is impressed by the throughput improvements and the reduction in complexity that FGMT brought to Machine I. This designer decides to implement FGMT on another machine, Machine II. Machine II is a pipelined machine with the following stages.

Fetch	1 stage
Decode	1 stage
Execute	8 stages (branch direction/target are resolved in the first execute stage)
Memory	2 stages
Writeback	1 stage

Assume everything else in Machine II is the same as in Machine I.

- (d) Is the number of threads required to eliminate branch-related stalls in Machine II the same as in Machine I? If yes, why? If no, how many threads are required?
- (e) What is the minimum CPI (i.e., maximum performance) of each thread in Machine II when this minimum number of threads is used?
- (f) Now consider flow-dependence stalls. Does Machine II require the same minimum number of threads as Machine I to avoid the need for flow-dependence stalls? If no, how many threads are required?
- (g) What is the minimum CPI of each thread when this number of threads (to cover flow-dependence stalls) is used?
- (h) After implementing fine grained multithreading, the designer of Machine II optimizes the design and compares the pipeline throughput of the original Machine II (without FGMT) and the modified Machine II (with FGMT) both machines operating at their maximum possible frequency, for several code sequences. On a particular sequence that has no flow dependences, the designer is surprised to see that the new Machine II (with FGMT) has lower overall throughput (number of instructions retired by the pipeline per second) than the old Machine II (with no FGMT). Why could this be? Explain concretely.

3. Branch Prediction (4 x 4 = 16 Points)

Consider the following high level language code segment:

```
int array[1000] = { /* random values */ };
int sum1 = 0, sum2 = 0, sum3 = 0, sum4 = 0;

for (I = 0; I < 1000; I++) // LOOP BRANCH
{
    if (I % 4 == 0) // IF CONDITION 1
        sum1 += array[I]; // TAKEN PATH
    else
        sum2 += array[I]; // NOT-TAKEN PATH

    if (I % 2 == 0) // IF CONDITION 2
        sum3 += array[I]; // TAKEN PATH
    else
        sum4 += array[I]; // NOT-TAKEN PATH
}
```

Your task is to find the prediction accuracy for the LOOP BRANCH (which is taken whenever the loop repeats, and not taken when the loop exits) and both of the IF CONDITION branches inside the loop (which are taken when the if-condition is *true*, and not taken when the if-condition is *false*), for different kinds of branch predictors.

- (a) What is the prediction accuracy for each individual branch using a per-branch last-time predictor (assume that every per-branch counter starts at “not-taken”) for the following branches?
- (b) What is the prediction accuracy when a per-branch 2-bit saturating counter-based predictor is used (assume that every per-branch counter starts at “strongly not-taken”), for the following branches?
- (c) What is the prediction accuracy of IF CONDITION 1 and IF CONDITION 2, when the counter starts at
 - (i) “weakly not-taken”?
 - (ii) “weakly taken”?
- (d) What is the prediction accuracy when a two-level global branch predictor with a two-bit global history register and a separate pattern history table per branch, consisting of 2-bit saturating counters for every entry, is used? Assume that both bits of the global history register are initialized to “not-taken” and the 2-bit saturating counters in the pattern history tables are initialized to “strongly not-taken”. When calculating prediction accuracy, ignore the first 500 loop iterations.

4. Design for Performance (8 x 2 = 16 Points)

You are a programmer at a large corporation, and you have been asked to parallelize an old program so that it runs faster on modern multicore processors.

- (a) You parallelize the program and discover that its speedup over the single-threaded version of the same program is significantly less than the number of processors. You find that many cache invalidations are occurring in each core’s data cache. What program behavior could be causing these invalidations?
- (b) You modify the program to fix this first performance issue. However, now you find that the program is slowed down by a global state update that must happen in only a single thread after every parallel computation. In particular, your program performs 90% of its work (measured as processor-seconds) in the parallel portion and 10% of its work in this serial portion. The parallel portion is perfectly parallelizable. What is the maximum speedup of the program if the multicore processor had an infinite number of cores?
- (c) How many processors would be required to attain a speedup of 4?
- (d) In order to execute your program with parallel and serial portions more efficiently, your corporation decides to design a custom heterogeneous processor.
 - This processor will have one large core (which executes code more quickly but also takes greater die area on-chip) and multiple small cores (which execute code more slowly but also consume less area), all sharing one processor die.
 - When your program is in its parallel portion, all of its threads execute **only** on small cores.
 - When your program is in its serial portion, the one active thread executes on the large core.
 - Performance (execution speed) of a core is proportional to the square root of its area.
 - Assume that there are 16 units of die area available. A small core must take 1 unit of die area. The large core may take any number of units of die area n^2 , where n is a positive integer.
 - Assume that any area not used by the large core will be filled with small cores.

-
- (e) How large would you make the large core for the fastest possible execution of your program?
- (f) What would the same program's speedup be if all 16 units of die area were used to build a homogeneous system with 16 small cores, the serial portion ran on one of the small cores, and the parallel portion ran on all 16 small cores?
- (g) Does it make sense to use a heterogeneous system for this program which has 10% of its work in serial sections?
- (h) Now you optimize the serial portion of your program and it becomes only 4% of total work (the parallel portion is the remaining 96%). What is the best choice for the size of the large core in this case?

5. Value Prediction [6 + 2 + 2 + 3 + 3 = 16 points]

We discussed the idea of value prediction as a method to handle data dependences. One method of value prediction for an instruction is "last-time prediction." The idea is to predict the value to be produced by the instruction as the value produced by the same instruction the last time the instruction was executed. If the instruction was never executed before, the predictor predicts the value to be 1. Value prediction accuracy of an instruction refers to the fraction of times the value of an instruction is correctly predicted out of all times the instruction is executed.

Assume the following piece of code, which has four load instructions in each loop iteration, loads to arrays x, y, z, t:

```
// initialize integer variables c, d, e, f to zeros
// initialize integer arrays x, y, z, t

for (i=0; i<1000; i++) {
    c += x[i];
    d += y[i];
    e += z[i];
    f += t[i];
}
```

Assume the following state of arrays before the loop starts executing:

- x consists of all 0's
- y consists of alternating 3's and 6's in consecutive elements
- z consists of random values between 0 and $2^{32} - 1$
- t consists of 0, 1, 2, 3, 4, ..., 999

- (a) What is the value prediction accuracy of the aforementioned predictor for the four load instructions in the program?
- load of x[i]:
 - load of y[i]:
 - load of z[i]:
 - load of t[i]:
- (b) Can you design a predictor that can achieve higher prediction accuracy for x[i]?
- (c) Can you design a predictor that can achieve higher prediction accuracy for y[i]?
- (d) Can you design a predictor that can achieve higher prediction accuracy for z[i]?
- (e) Can you design a predictor that can achieve higher prediction accuracy for t[i]?

Indian Statistical Institute

M.Tech (CS) II

Information and Coding Theory

Mid Semester Examination

Maximum Marks: 70

Date: September 21, 2016.

Time 2.5 hours

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

1. Let X, Y and Z be joint random variables. Prove the following inequalities and find conditions for equality.

- (a) $H(X, Y|Z) \geq H(X|Z)$.
- (b) $I(X, Y; Z) \geq I(X; Z)$.
- (c) $H(X, Y, Z) - H(X, Y) \leq H(X, Z) - H(X)$.
- (d) $I(X; Z|Y) \geq I(Z; Y|X) - I(Z; Y) + I(X; Z)$.

(3+3+3+3 = 12)

2. Let X and Y be two independent integer-valued random variables. Let X be uniformly distributed over $\{1, 2, \dots, 8\}$, and let $Pr\{Y = k\} = 2^{-k}$, $k = 1, 2, 3, \dots$

- (a) Find $H(X)$.
- (b) Find $H(Y)$.
- (c) Find $H(X + Y, X - Y)$.

(1+2+2= 5)

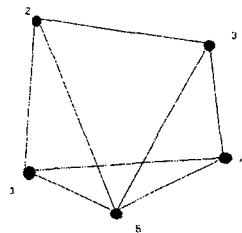
3. Let X_1, X_2, \dots be independent, identically distributed random variables drawn according to the probability mass function $p(x)$, $x \in \{1, 2, \dots, m\}$. Thus, $p(x_1, x_2, \dots, x_n) = \prod_{i=1}^n p(x_i)$. We know that $-\frac{1}{n} \log p(X_1, X_2, \dots, X_n) \rightarrow H(X)$ in probability.

Let $q(x_1, x_2, \dots, x_n) = \prod_{i=1}^n q(x_i)$, where q is another probability mass function on $\{1, 2, \dots, m\}$.

- (a) Evaluate $\lim -\frac{1}{n} \log q(X_1, X_2, \dots, X_n)$, where X_1, X_2, \dots are i.i.d. $\sim p(x)$.
- (b) Evaluate the limit of the log likelihood ratio $\frac{1}{n} \log \frac{q(X_1, \dots, X_n)}{p(X_1, \dots, X_n)}$ when X_1, X_2, \dots are i.i.d. $\sim p(x)$.

(6 + 6 = 12)

4. Consider a random walk on the following graph.



- (a) Calculate the stationary distribution.
- (b) Calculate the entropy rate.

(c) Find the mutual information $I(X_{n+1}; X_n)$ assuming that the process is stationary.

(4 + 4 + 4 = 12)

5. Let the probability of occurrence of m symbols be $\{p_1, p_2, \dots, p_m\}$. For any binary Huffman code, prove the following results:

(a) If the most probable symbol has a probability $p_1 > 2/5$, then that symbol must be assigned a codeword of length 1.

(b) If the most probable symbol has a probability $p_1 < 1/3$, then that symbol must be assigned a codeword of length ≥ 2 . (6 + 6 = 12)

6. Consider the discrete memoryless channel $Y = X + Z(\text{mod } 11)$, where

$$Z = \begin{pmatrix} 1, & 2, & 3 \\ 1/3, & 1/3, & 1/3 \end{pmatrix},$$

and $X \in \{0, 1, \dots, 10\}$. Assume that Z is independent of X .

(a) Calculate the channel capacity.

(b) For what distribution of X is this capacity achieved? (4 + 1 = 5)

7. Prove that all rates below the channel capacity C are achievable. Specifically, for every rate $R < C$, prove that there exists a sequence of $(2^{nR}, n)$ codes with maximum probability of error $\lambda^{(n)} \rightarrow 0$. (12)

INDIAN STATISTICAL INSTITUTE

Periodical Examination

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

Advanced Algorithms for Graph and Combinatorial Optimization Problems

Date: 21.9.2016

Maximum Marks: 60

Duration: 3 Hours

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

1. For any simple planar graph $G = (V, E)$ with $|V| = n$ and $|E| = m$, prove the followings:
 - (i) There is a vertex of G that has degree at most 5.
 - (ii) If $n \geq 3$ and G is bipartite then $m \leq 2n - 4$.
- (b) Consider the following algorithm for computing the *maximum independent set* (MIS) of a planar graph $G = (V, E)$.
 - Choose an appropriate vertex $v \in V$,
 - include v in the independent set S ,
 - delete v and all its neighbors from V , and
 - repeat the process on the graph G' obtained in the earlier step unless $V = \emptyset$.

For what choice of v in each step, we can get a constant factor approximation algorithm for the MIS problem of the planar graph? State the time complexity of this algorithm and justify. [(5+5)+(4+6)=20]

2. Let $G = (V, E)$ be a connected unweighted undirected planar graph, $n = |V|$. Suppose the vertices of G are partitioned according to their distances from a fixed vertex v . Let $L(i) = \text{No. of vertices which are at distance } i \text{ from } v \text{ in } G$.
 $L(0) = 1$, and let r be the maximum distance of a vertex from v . Let ℓ_1 and ℓ_2 be two levels ($\ell_1 \leq \ell_2$) such that the total number of vertices from level 0 to $L(\ell_1 - 1)$ is not exceeding $\frac{2}{3}n$, and total number of vertices from level $L(\ell_2 + 1)$ to r is not exceeding $\frac{2}{3}n$. Then show that it is possible to partition the number of vertices in G into three parts, namely A , B and C , such that there is no edge joining a vertex of A and a vertex of B , $\max(|A|, |B|) \leq \frac{2}{3}n$, and C contains no more than $L(\ell_1) + L(\ell_2) + 2(\ell_2 - \ell_1 - 1)$. [12]
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 $\ell(u) = 1 + \max_{w \in \text{adj}(u)} \ell(w)$, where $\text{adj}(u)$ 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. [4+8=12]
5. Let $G = (V, E)$ be a connected undirected graph. Let S be a vertex separator of G . Let $G_i = (V_i, E_i), i = 1, 2, \dots, t$ be the connected components of $G' = (V', E')$, where $V' = V \setminus S$. If every pair of vertices in S are connected in G , then show that (i) $\chi(G) = \max_{i=1}^t \chi(G_i)$ and (ii) $\omega(G) = \max_{i=1}^t \omega(G_i)$, where $\chi(G)$ and $\omega(G)$ are the coloring number and clique number of the graph G . [6+6=12]

Indian Statistical Institute
Mid-Semester Examination : 2016 – 2017
Master of Technology in Computer Science, Semester III
Functional Brain Signal Processing: EEG & fMRI

Date: ²⁴September 2016

Maximum Marks: 50

Duration: 2 hours

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

1. In case of the following EEG artifacts, which filters are most appropriate? Justify your answer. Multiple filters may be almost equally appropriate and in that case mention all of them with justification. 4 x 3 = 12
 - a) Eye blink.
 - b) Electrical line noise.
 - c) ECG or heart bit.
2. With the help of a diagram describe international 10-20 EEG electrode placement system. Maximum how many electrodes can be placed within the scope of this system? Justify your count. 8 + 2 + 2 = 12
3. Write short notes on the following:
 - a) Event related potential.
 - b) Mu-band EEG.
 - c) EEG forward problem. 4 x 3 = 12
4. a) Geometrically describe (you may use a suitable diagram to illustrate your point) Fisher's discriminant algorithm (FDA) for binary classification. From this, explain why the data sets should be normally distributed for FDA to work effectively. 5
b) Describe a linear support vector machine (SVM) and explain the principle of structural risk minimization. 5
c) If you are asked to classify the EEG on a visual discrimination task between a stimulus and baseline, which channels would you like to take your data from? After selecting the data, out of FDA, SVM and logistic regression, which one would you prefer to classify the EEG data and why? 1 + 3 = 4

INDIAN STATISTICAL INSTITUTE

Periodical Examination: (2016 – 2017)

M.Tech. (CS) II Year

Parallel Processing: Architectures and Algorithms

Date: 22 /09/2016

Total Marks: 72

Duration: 2 hrs

NOTE: You may answer all questions but maximum score you may attain is 60.

1. a) Show the schematic diagrams of a Parallel Random Access Machine (PRAM), and a typical Graphic Processor Unit respectively. Mention three important differences between the two.

(b) Write a parallel algorithm to multiply two $n \times n$ matrices on PRAM with CREW memory, where n is a power of 2. Use as many processors as you need to achieve a speed-up of $O(n^3/\log n)$ and to maximize utilization. Assume that both addition and multiplication take 1 time unit, and $\log n \gg 1$.

[(2+2+3)+(5+3) = 15]

2. a) For a given a program, 75% of the code is executable simultaneously by 25 processors. The rest of the code is to be executed sequentially. Find the maximum speed-up achievable when the load is:

- i) fixed, and
- ii) scalable.

b) A pipelined vector processor uses five 32-bit pipelines, each with four stages for instruction fetch, instruction decode, operand fetch and execution respectively. Instructions may require two operands at most. The clock rate is 40 ns and the memory access time is 1 μ s. Find the minimum number of memory modules required to satisfy the data demand.

[(4+4)+4=12]

3. Consider the following program segment with six instructions:

$I_1 : B = B + C$
 $I_2 : D = B * A$
 $I_3 : E = B * C$
 $I_4 : F = D + B$
 $I_5 : A = E + D$
 $I_6 : C = E * A$
 $I_7 : C = C + F$

Draw the directed acyclic program flow graph considering each statement as a process.

Show a possible scheduling of the processes on two processors following 'ETF (earliest start time first)' algorithm. Assume that the operations '+', '*', and the inter-process communication take 20, 40 and 50 time units respectively. Calculate the speed-up achieved by your scheduling.

[6+6+3 = 15]

P.T.O

4. (a) In an $N \times N$ mesh network, how many distinct shortest paths exist between a pair of vertices (x_1, y_1) and (x_2, y_2) , $0 \leq x_1, y_1, x_2, y_2 < N$. Two paths are distinct if they differ at least in one edge. How many paths of these are node disjoint (no common node, except the source and destination)?

(b) Find an expression for the *Moore bound* on the number of nodes N in a regular graph with degree d and diameter k . Explain why in reality a regular graph with degree d and diameter k can attain the bound on N for some small values of d and k only.

(c) In an N -node hypercube H_n , ($N=2^n$), each node has a single packet to route to a unique destination. Develop a routing algorithm that each node should follow to complete the whole procedure. How many communication rounds will be required to deliver all the packets? Assume that the links are bidirectional.

[5 × 3 = 15]

5. a) State if the following statements are true (T) or false (F):
- A thread block is an array of threads that can cooperate, synchronize and share data in shared memory. (T/F)
 - Warp is a group of threads that always execute same instructions simultaneously. (T/F)
 - Warp follows Multiple Instruction Multiple Thread (MIMT) programming model. (T/F)
 - `__syncthreads()` synchronizes all threads in a block. (T/F)
 - The `__device__` qualifier declares a variable that resides on the device. (T/F)

b) Write a CUDA program to compute $C = A+B$, where A , B , and C are $N \times N$ matrices. Explain how the Grids, Threads, Blocks, and Memory (Global/ Shared) can be used efficiently to get better speed-up, when N is too large.

[5+5+5=15]

INDIAN STATISTICAL INSTITUTE

Mid-Semestral Examination: (2016 - 2017)

Course Name: M. Tech. (CS)

Year: 2nd year

Subject Name: Neural Networks & Applications

Date: September 22, 2016

Maximum Marks: 50 Duration: 2 hrs

Answer all the questions.

1. Describe the steps for training a single node perceptron model. Justify the step of modifying the weight values for linearly separable two-class pattern classification problems. Show how single equation can be used for modifying the weight values for all the patterns in both the classes. [10]
2. Write a short note on self organization. [15]
3. Consider a two-input AND gate. Design a single node perceptron model with possible values of its weights so that the model can classify the input patterns of the AND gate based on its output. That is, the input patterns for which the output is zero form class C_1 , and the other input patterns are in class C_2 . Hence design a similar model for classifying the input patterns of a two-input OR gate. [15]
4. Describe different functions with appropriate explanation, which are usually used as activation functions of artificial neural networks. [10]

INDIAN STATISTICAL INSTITUTE

Periodical Examination

M. Tech (CS) - II and JRF

Multi-dimensional search and computational geometry

Date : 22.09.2016

Maximum Marks : 40

Duration : 3 Hours

Question 1: State and prove the lower-bound on the running time of Intersection of n line segments. 10

Question 2: Prove that, the vertices of a triangulated simple polygon are three colorable. 10

Question 3: (a) Prove that, in every triangulation of n points in R^2 there exists an independent set of at least $\lfloor \frac{n}{18} \rfloor$ among the vertices of maximum degree 8. Describe an algorithm that finds such a set of vertices in $O(n)$ time. 5

(b) How is the above bound useful for building planar point location data structure? 5

Question 4: (a) If the center of minimum enclosing disk is constrained to lie on the x -axis then which are the points, in Figure 1, will get pruned after first iteration for the given pairing of the points. The bisectors of the pairs are shown in the figure. 5

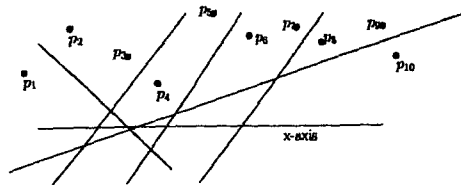


Figure 1:

(b) List the ordered set of deleted points by Graham Scan algorithm for computing the upper envelope of the convex hull of the point set shown in Figure 2. 5

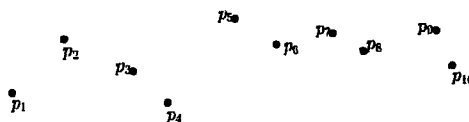


Figure 2:

Indian Statistical Institute

Mid-Semester Examination 2016-17

Course: Mtech CS 2 Subject: Cognitive Science

Date: 22/09/2016 Maximum Marks: 30 Duration : 90 minutes

Answer the following questions. Marks for each question are mentioned in bracket after each question.

1. Explain the difference between introspective and behaviorist approaches? (3)
2. What is the relation between Cognitive Science and Computer Science, if any. (3)
3. What are absolute threshold and difference threshold? (3)
4. At night, we can see big things better than small things. Why? (3)
5. Explain Weber's Law. What is Weber fraction? (3)
6. Explain with a graph/plot/figure the concept of Point of Subjective Equivalence (PSE).(3)
7. Describe the method of limits to measure difference threshold. (3)
8. What is the law of specific nerve energies? Explain its significance. (3)
9. What are 3 ways among many others by which depth can be estimated monocularly (using one eye). (3)
10. What is the relationship of Fourier's theorem to analysis of a visual scene? (3)

Indian Statistical Institute
Semester-I 2016-2017
M.Tech.(CS) - Second Year
Mid-term Examination (September 23, 2016)
Subject: Pattern Recognition and Image Processing

Total marks: 80

Maximum marks: 60

Duration: 2 Hours

Answer maximum of 60 marks (any part of any question). Answers should be precise.

1. (a) Suppose the class conditional probability density functions of class 1 (w_1) and class 2 (w_2) for any pattern x are defined as follows:

$$p(x|w_1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad \forall x;$$

$$p(x|w_2) = \frac{1}{4}, \quad -2 < x < 2.$$

Find the Bayes' decision boundary for this two-class problem by assuming the prior probabilities $P(w_1) = P(w_2) = 0.5$. (10 Marks)

- (b) Derive the conditions under which the Bayes' classifier acts like a minimum distance classifier. (5 Marks)
- (c) Describe the K-NN classifier. Write down the advantages and disadvantages of K-NN classifier. (5 Marks)
2. (a) Explain the differences between Euclidean and Mahalonobis distances. (5 Marks)
- (b) Explain the advantages and limitations of (i) k -means (ii) k -medoids and (iii) DBSCAN clustering algorithms. (6 Marks)
- (c) Consider the following data in 2D. Find the clusters (using Euclidean distance) and draw the dendrogram using

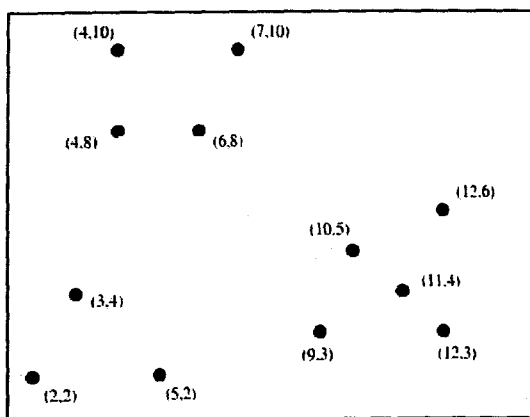


Figure : Twelve points to be clustered hierarchically

Figure 1: 2D Data.

- (i) Single linkage (3 Marks)
 - (ii) Complete linkage (3 Marks)
 - (iii) Average linkage (Group average) (3 Marks)
3. (a) How does the variance-covariance matrix capture the relationships between features? (6 Marks)
- (b) Consider the following data of two classes. Compute (i) Scatter matrices (with-in and

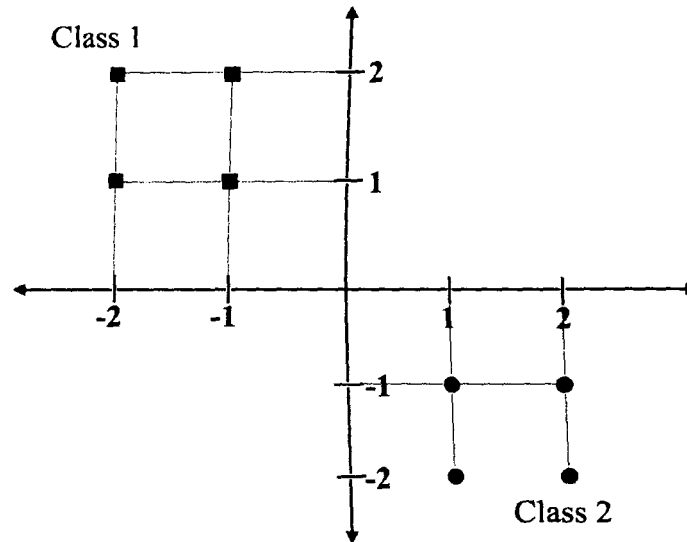


Figure 2: Data of two classes.

- between) (ii) Mixture scatter matrix (iii) Any one objective measure based on these scatter matrices for finding out the goodness of feature subsets. (9 Marks)
- (c) Describe the principal component analysis with it's pros and cons. (5 Marks)
4. (a) Discuss (i) Sequential forward selection, (ii) Sequential backward selection, and (iii) "Plus-L, minus-R" selection, strategies for feature subset selection. (6 Marks)
- (b) Derive the branch and bound tree for $D = 5$, $d = 2$ and $J(X) = \sum_{\xi_i \in X} (\frac{1}{2})^2$, where X and ξ_i are the set of features and the i^{th} feature, respectively. (10 Marks)
- (c) Explain the differences between feature selection and feature extraction. (4 Marks)

INDIAN STATISTICAL INSTITUTE

MIDTERM EXAMINATION M.TECH(CS) II YEAR

CRYPTOLOGY

Date: 23.09.2016 Maximum marks: 70 Duration: 2 hours and 30 mins.

The paper contains 80 marks. Answer as much as you can, the maximum you can score is 70.

1. (a) Define perfect secrecy.

[3]

(b) Consider a symmetric cryptosystem with key space \mathcal{K} , message space \mathcal{M} and cipher space \mathcal{C} , where $\mathcal{K} = \{K_1, K_2, K_3\}$, $\mathcal{M} = \{a, b, c\}$, and $\mathcal{C} = \{1, 2, 3, 4\}$. Suppose the encryption matrix is as follows:

	a	b	c
K_1	1	2	3
K_2	2	3	4
K_3	3	4	1

This means a when encrypted with key K_1 yields 1, etc. Given that the keys are chosen with equal probability and the plaintext distribution is

$$\Pr[a] = \frac{1}{2}, \Pr[b] = \frac{1}{3}, \Pr[c] = \frac{1}{6}.$$

Find the probability distribution on \mathcal{C} . Does this crypto-system provide perfect secrecy?

[7]

2. (a) Define a pseudorandom generator.

[2]

(b) Let R and S be nonempty finite sets and G be a pseudorandom generator (PRG) with S as the seed space and $R \times S$ as the output space. Describe the Blum-Micali construction to construct a PRG G' with seed space S and output space $R^n \times S$ for $n \geq 1$.

[3]

3. (a) State and prove the difference lemma.

[5]

(b) Consider the following two games named **G0** and **G1** and an adversary \mathcal{A} who queries q elements from $\{0, 1\}^n$.

Game G0	Game G1
<p>Initialization $bad \leftarrow false;$ $Dom \leftarrow \emptyset;$ $Ran = \emptyset;$</p> <p>Query Phase For a query $x^{(i)}$ of \mathcal{A} do the following if $x^{(i)} \notin Dom$ then $y^{(i)} \xleftarrow{\\$} \{0, 1\}^n;$ if $y^{(i)} \in Ran,$ $bad \leftarrow true;$ $y^{(i)} \xleftarrow{\\$} \{0, 1\}^n \setminus Ran;$ end if $T[x^{(i)}] \leftarrow y^{(i)};$ $Dom \leftarrow Dom \cup \{x^{(i)}\};$ $Ran \leftarrow Ran \cup \{y^{(i)}\}$ end if; return $T[x^{(i)}];$</p>	<p>Initialization $bad \leftarrow false;$ $Dom \leftarrow \emptyset;$ $Ran \leftarrow \emptyset;$</p> <p>Query Phase For a query $x^{(i)}$ of \mathcal{A} do the following if $x^{(i)} \notin Dom$ then $y^{(i)} \xleftarrow{\\$} \{0, 1\}^n;$ if $y^{(i)} \in Ran,$ $bad \leftarrow true;$ end if $T[x^{(i)}] \leftarrow y^{(i)};$ $Dom \leftarrow Dom \cup \{x^{(i)}\};$ $Ran \leftarrow Ran \cup \{y^{(i)}\}$ end if; return $T[x^{(i)}];$</p>

Find an upper bound for

$$|\Pr[\mathcal{A}^{G0} \Rightarrow 1] - \Pr[\mathcal{A}^{G1} \Rightarrow 1]|. \quad [10]$$

- (c) Let $F : \mathcal{K} \times \{0, 1\}^n \rightarrow \{0, 1\}^n$ be a permutation family and let \mathcal{A} be an adversary who asks at most q queries. Using the result above, state and justify an upper bound for

$$|\text{Adv}_F^{\text{prp}}(\mathcal{A}) - \text{Adv}_F^{\text{prf}}(\mathcal{A})|. \quad [5]$$

4. (a) With the help of a diagram show the CBC mode of operation for encrypting $4n$ bits of message with a block cipher of block length n . [2]

- (b) Assume that a sender encrypts a message of $2m$ blocks using the CBC mode. During transmission, block number m gets corrupted. After the receiver decrypts the cipher, how many blocks of the decrypted message would be corrupted? [3]

- (c) The CBC-Chain mode of operation is a CBC variant in which the IV that is used for the very first message to be encrypted is randomly selected, where as the IV used for each subsequent encrypted message is the last block of ciphertext that was generated. Show that CBC-Chain is insecure by constructing an efficient IND-CPA adversary. [10]

5. (a) Define a strong pseudorandom permutation.

[3]

- (b) Let \mathcal{K} be a non-empty finite set and let $f : \mathcal{K} \times \{0, 1\}^n \rightarrow \{0, 1\}^n$ be a pseudorandom function family. Given f , define functions $\text{Fiestel} : \mathcal{K} \times \{0, 1\}^{2n} \rightarrow \{0, 1\}^{2n}$ and $G : \mathcal{K} \times \{0, 1\}^{2n} \rightarrow \{0, 1\}^{2n}$ as follows:

```

Fiestelk(x)
  x0 || x1 ← x;
  return x1 || x0 ⊕ fk(x1);

```

```

Gk(x)
  y ← Fiestel(x);
  z ← Fiestel(y);
  return z

```

Show that G is not a pseudorandom family.

[7]

6. (a) Let $\text{Perm}(n)$ denote the set of all permutations on $\{0, 1\}^n$ and let x, y be fixed n -bit strings. Find $\Pr[\pi \xrightarrow{\$} \text{Perm}(n) : \pi(x) = y]$.

[5]

- (b) Let $P, C \in \{0, 1\}^{64}$ are given. We apply the following algorithm for exhaustive key search for DES, which on input (P, C) outputs the set of all possible keys which maps P to C on encryption.

```

Algorithm Exhaustive(P, C)
  Keys ← ∅
  for each key K ∈ {0, 1}^{56}
    if DESK(P) = C,
      Keys ← Keys ∪ {K}
    end if
  end for
  return Keys

```

If we assume that DES behaves like a random permutation, i.e, for any fixed $P, C \in \{0, 1\}^{64}$

$$\Pr[k \xrightarrow{\$} \{0, 1\}^{56} : \text{DES}_k(P) = C] = \Pr[\pi \xrightarrow{\$} \text{Perm}(64) : \pi(P) = C].$$

Then, what is the size of the set Keys returned by the algorithm Exhaustive in average?

[5]

- (c) Describe the double DES scheme. Describe a meet in the middle attack on double DES.

[10]

INDIAN STATISTICAL INSTITUTE

Mid-Semester Examination: 2016 (First Semester)

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

Subject Name: Natural Language Processing

Date: 24.09.2016 Maximum Marks: 50 Duration: 2 hours

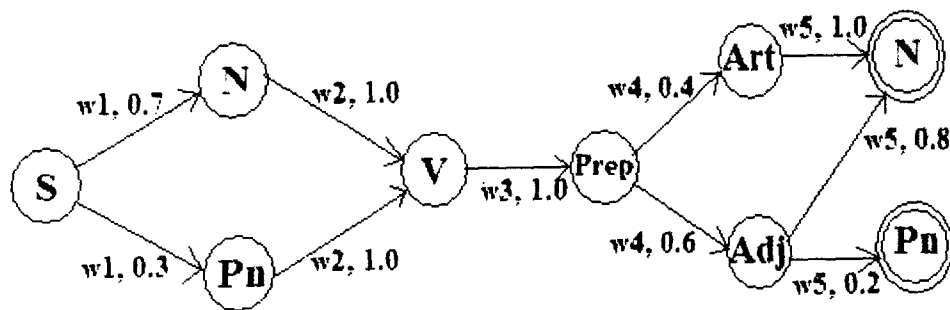
Note: **Open Book/Class-note Exam.**
 Answer all questions.

Q1. [15 marks]

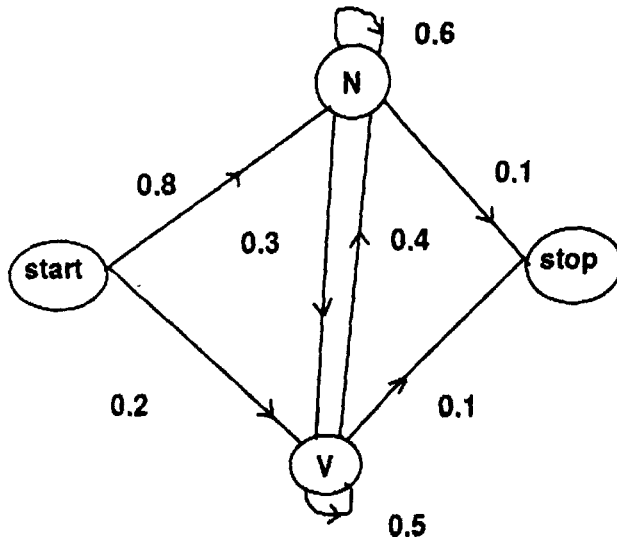
Assume you are dealing with a language where all sentences are of five words/tokens and only six possible POS tags are there $\langle N, Pn, V, Prep, Art, Adj \rangle$. Two groups design two different POS tagging systems with the following assumptions as follows:

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 are given, 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.



2. Consider the bi-gram HMM shown above. There are two tags N and V which can emit words. Two special tags start and stop cannot emit any word. Let there be only three words “I”, “You” and “He” in the vocabulary set. The word emission probabilities are given below.

$$p(\text{“I”}|N) = 0.2, p(\text{“You”}|N) = 0.4 \text{ and } p(\text{“He”}|N) = 0.4$$

$$p(\text{“I”}|V) = 0.5, p(\text{“You”}|V) = 0.4 \text{ and } p(\text{“He”}|V) = 0.1$$

The observation sequence is “He I He”.

- a. Find the probability of occurring the observation sequence using only forward probabilities (using alpha). 8
- b. Find the same using only backward probabilities (using beta). 8
- c. Assuming the values of all the parameters of the above HMM as initial values and the given observation, estimate $p(\text{“1”}|H)$ and $p(\text{“1”}|C)$ for the next iteration. 4+4

3. We have a tri-gram HMM tagger on a training set with the following POS tagged sentences
 The dog saw the cat, D N V D N
 The dog saw the saw, D N V D N

We estimate the parameters of the HMM using maximum likelihood estimation.

- a. Draw the automaton of the given HMM. 5

Consider the following sentence
 The cat saw the saw

- b. Find the most probable tag sequence for the above sentence and find the joint probability of that tag sequence with the sentence. 5+5

INDIAN STATISTICAL INSTITUTE

M.TECH. (CS) - YEAR II

2016 - 2017

LOGIC FOR COMPUTER SCIENCE

MID-SEMESTER EXAMINATION

Date: 24.09.2016
Time: 3 Hours

Marks: 40

Answer Question 1 and any 4 from the rest.

1. Let \mathcal{L} be a language whose formulas φ are given as follows:

$$\varphi := p \mid \neg\varphi \mid \varphi \times \varphi$$

where $p \in \mathcal{P}$, a countable set of propositions. Let $\mathcal{V} : \mathcal{P} \rightarrow \{0, 1\}$ be a valuation function. The truth definition of the formulas is given as follows:

$$\begin{aligned}\mathcal{V} \models p &\text{ iff } \mathcal{V}(p) = 1; \\ \mathcal{V} \models \neg\varphi &\text{ iff } \mathcal{V} \not\models \varphi; \\ \mathcal{V} \models \varphi \times \psi &\text{ iff } \mathcal{V} \models \varphi \text{ and } \mathcal{V} \models \psi.\end{aligned}$$

Consider the following axiom scheme:

$$\begin{aligned}\text{Axiom 1: } &\varphi \times (\psi \times \varphi) \\ \text{Axiom 2: } &(\varphi \times (\psi \times \chi)) \times ((\varphi) \times (\varphi \times \chi)) \\ \text{Axiom 3: } &((\neg\varphi \times \neg\psi) \times (\neg\varphi \times \psi)) \times \varphi\end{aligned}$$

and the rule:

$$\frac{\varphi \quad \varphi \times \psi}{\psi}$$

Let Γ be a set of formulas in the given language, and φ is a formula in the language. Define $\Gamma \vdash \varphi$ if there is a sequence of formulas $\varphi_1, \dots, \varphi_n$ such that φ_n is φ and each φ_i is either an instance of an axiom (mentioned above), or a member of Γ or obtained by the rule mentioned above. Define $\Gamma \models \varphi$ if for every valuation \mathcal{V} , $\mathcal{V} \models \varphi$, whenever $\mathcal{V} \models \gamma$ for all $\gamma \in \Gamma$.

Show that $\Gamma \vdash \varphi$ iff $\Gamma \models \varphi$. [20]

2. Let \mathcal{A} be a non-empty set. A binary relation $<$ on \mathcal{A} is said to be linearly ordered if $<$ is irreflexive, transitive and satisfies the law of trichotomy. The relation $<$ is dense if for any $x, y \in \mathcal{A}$, with $x < y$, there is some $z \in \mathcal{A}$ such that $x < z < y$. Find a suitable first order language and give axioms in it for a dense linear order without bounds. In addition, give axioms for an equivalence relation in the same language. [5]
3. Assume that a first order language has equality and a two place predicate symbol P . For each of the following conditions, find a sentence φ such that a structure \mathcal{A} is a model of φ iff the condition is met. [5]
- (a) The domain of the structure \mathcal{A} has exactly two members.
(b) The interpretation of P in the structure \mathcal{A} , $P^{\mathcal{A}}$ is a function on the domain of \mathcal{A} .
4. Check whether the following formulas are valid: [5]
- (a) $(\exists x\varphi \vee \psi) \leftrightarrow \exists x(\varphi \vee \psi)$, x is not free in ψ .
(b) $\forall x(\varphi \rightarrow \psi) \leftrightarrow (\varphi \rightarrow \forall x\psi)$, x is not free in φ .
5. Give examples to show that the following formulas are not valid. [5]
- (a) $\forall x(\varphi \vee \psi) \rightarrow (\forall x\varphi \vee \forall x\psi)$.
(b) $(\exists x\varphi \wedge \exists x\psi) \rightarrow \exists x(\varphi \wedge \psi)$.
6. Show that an infinite map can be colored with four colors iff every finite submap of it can be. [5]
7. There exists a set of sentences in first order logic whose models are precisely the finite sets. Prove or disprove. [5]

INDIAN STATISTICAL INSTITUTE

Mid-Semestral Examination:(2015-2016)

M.TECH (CS) II YEAR

Subject Name: Quantum Information Processing and Quantum Computation

Maximum Marks: 30

Duration: 2 hours

Date: 24/09/16

Answer any three of the following four questions

1. a) Consider the following linear operator A acting on a two dimensional Hilbert space.;

$$A = \frac{1}{2}(I + n \cdot \sigma)$$

where $n \cdot \sigma = \sum n_i \sigma_i$ ($i = x, y, z$), n is a vector in R^3 , σ 's are Pauli matrices and I is identity operator .

i) Under what condition A is a positive operator .

ii) Under what condition A is a projection operator.

b) Show that for every the density operator D with $D^2 = D$; there exists a unit vector $|\psi\rangle$, such that

$$D = |\psi\rangle\langle\psi|.$$

c) Let the initial density matrix of a qubit is $\frac{1}{2}(I+n \cdot \sigma)$. If spin measurement is performed along some vector m , what is the probability for the spin up result?

(2+2)+3+3

2. a) Consider following two states $|0\rangle$ and $|\psi\rangle = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$, where $|0\rangle$ and $|1\rangle$ are two orthogonal vectors. Show that these two states can not be reliably cloned.

b) Consider a Swap operator U_s which acts in the following way;

$$U_s|\psi\rangle \otimes |\phi\rangle = |\phi\rangle \otimes |\psi\rangle$$

for all possible states $|\psi\rangle, |\phi\rangle$. Then show that U_s can not be written as

$$U_s = U_1 \otimes U_2$$

where U_1 and U_2 are acting on particle 1 and particle 2 respectively.

c) Show how the swap operation (gate) can be realized by using three C-Not gates.

4 + 3 + 3

3. a) Let Alice and Bob share the following state;

$$|\psi\rangle_{AB} = \frac{1}{\sqrt{3}}(|0\rangle_A \otimes |0\rangle_B + \sqrt{\frac{2}{3}}|1\rangle_A \otimes |1\rangle_B)$$

where $|0\rangle$ and $|1\rangle$ are eigen states of σ_z .

(i) Show that the state can not be written as $|\phi\rangle \otimes |\chi\rangle$.

(ii) Find the probability of transforming the state to a maximally entangled state by local operations and classical communications.

b) How much communication can be made by sending just a qubit? Discuss how 2 bits of communication can be made by sending a qubit if it is in a maximally entangled state with another qubit held by the receiver.

(1 + 4) + 5

4. a) Consider a normalized qubit state ;

$$|\psi\rangle = a|0\rangle + b|1\rangle$$

where a and b are complex numbers and $|a|^2 + |b|^2 = 1$

Then show that

$$\frac{1}{2}I = \frac{1}{4} \sum_{i=0,x,y,z} \sigma_i |\psi\rangle \langle \psi| \sigma_i$$

where σ_0 is the identity operator I .

b) Describe how an unknown quantum state of a qubit can be teleported to distant laboratory by using maximally entangled state as resource.

c) Argue how no-cloning principle is not violated in quantum teleportation.

3 + 6 + 1

INDIAN STATISTICAL INSTITUTE

Mid Semestral Examination:(2016-2017)

MTech C.S. 2nd Year

Digital Signal Processing

Date: 24.9.2016

Maximum Marks: 60

Duration: 2 hours

Note: The marks add up to 73. The maximum you can score is 60. Use of calculators is permitted.

1. Let $x[n] = 0.7(1 + [-1]^n)u[n - 3]$ where $u[n]$ is the unit step sequence. Obtain the z-transform and sketch its poles and zeros. Determine the Discrete-time Fourier Transform if it exists. [8+4+3]

2. Consider a system with system function

$$H(z) = \frac{z^{-2}}{(1 - 1.2z^{-1})(1 - 0.3z^{-1})}$$

Determine

- (a) Its impulse response if the system is known to be causal.
(b) Its impulse response if the system is known to be stable.

[5+5]

3. Let

$$x[n] = \begin{cases} n & 0 \leq n \leq 4 \\ 0 & \text{otherwise} \end{cases}$$

Denoting the DTFT of $x[n]$ by $X(e^{j\omega})$, compute (or sketch), without actually evaluating $X(e^{j\omega})$, the inverse DTFT of the following:

- (a) $e^{j3\omega} X(e^{j\omega})$
(b) $X(e^{j(\omega+\pi)})$
(c) $\frac{dX(e^{j\omega})}{d\omega}$

[6+6+6]

4. An LTI system with impulse response

$$h[n] = (0.7)^n u[n]$$

is cascaded with a discrete-time ideal low pass filter with cut-off at $\pi/4$.

- (a) Is the overall system LTI, causal and stable? Justify your answer in each case.
- (b) Determine the frequency response of the overall system.

[(4+4+4)+5]

5. Consider an analog signal $x(t) = \cos(2\pi 50t)$, $0 < t < \infty$. The signal is sampled at a sampling rate of 200 Hz to produce $x[n]$. Determine the discrete-time signal $x[n]$ and sketch $|X(e^{j\omega})|$. Now, let $x[n]$ be downsampled by a factor of 2 to yield $y[n]$. Sketch $|Y(e^{j\omega})|$. [(3+5)+5]

INDIAN STATISTICAL INSTITUTE

First-Semester Examination: 2016-2017

M. Tech. (CS) 2nd Year

Artificial Intelligence

Date: 02.12.2016

Maximum Marks: 100

Duration: 3 hours

Answer any ten questions. All questions carry equal marks.

1. (i) Prove that the *depth-first iterative deepening* algorithm is asymptotically optimal among brute-force tree searches in terms of time, space, and length of solution.
(ii) Prove that any monotonic heuristic is *admissible*.

[(3 + 2 + 1) + 4 = 10]

2. Solve the following cryptarithmic problem:

$$\begin{array}{r} \text{B A S E} \\ + \text{B A L L} \\ \hline \text{G A M E S} \\ \hline \end{array}$$

[10]

3. 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 and show that the player who has the second move can always win the game.
b) 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.

[5 + 5 = 10]

4. 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:

- a) A tile may move into an adjacent empty cell with unit cost.
b) 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 irrespective of 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

P.T.O

the cost of the solution?

[3 + 7 = 10]

5. Explain the concepts of belief and plausibility with an example in the context of Dempster-Shafer theory of evidence. Suppose an initial observation S_1 confirms some hypothesis h with the belief $MB = 0.3$. The second observation S_2 confirms the same hypothesis h with the belief $MB = 0.5$. Find the certainty factor regarding the hypothesis h by two observations S_1 and S_2 .
[6 + 4 = 10]
6. Define information gain. Consider the following decision table and calculate the information gain for four condition attributes, namely, *Credit history*, *Debt*, *Collateral*, and *Income*, considering *Risk* as the decision attribute. Explain which condition attribute will be selected first, based on information gain.

<i>Credit history</i>	<i>Debt</i>	<i>Collateral</i>	<i>Income</i>	<i>Risk</i>
bad	high	none	\$0 to \$15K	high
unknown	high	none	\$15K to \$35K	high
unknown	low	none	\$15K to \$35K	moderate
unknown	low	none	\$0 to \$15K	high
unknown	low	none	over \$35K	low
unknown	low	adequate	over \$35K	low
bad	low	none	\$0 to \$15K	high
bad	low	adequate	over \$35K	moderate
good	low	none	over \$35K	low
good	high	adequate	over \$35K	low
good	high	none	\$0 to \$15K	high
good	high	none	\$15K to \$35K	moderate
good	high	none	over \$35K	low
bad	high	none	\$15K to \$35K	high

[2 + 6 + 2 = 10]

7. In the context of rough set theory, define and explain the following with the decision table in Question 6:
- lower and upper approximations of the decision attribute, and
 - degree of dependency and significance of a condition attribute.
- [6 + 4 = 10]
8. (i) 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.

P.T.O

(ii) 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."

[6 + 4 = 10]

9. (i) Prove that if α is a logical consequence of a set of premises Σ , then there is a tableau proof of α from Σ .

(ii) Prove that a clause C is a logical consequence of a set of clauses S if and only if the set $S' = S \cup \{\sim C\}$ is unsatisfiable.

[5 + 5 = 10]

10. 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]

11. (i) Prove that if a well formed formula α is a tableau provable, then α is valid.

(ii) Write a program in Prolog for merging two ordered lists.

[5 + 5 = 10]

12. (i) Describe the Bayes' theorem for probabilistic reasoning.

(ii) Explain with examples the differences between red cut and green cut in Prolog.

[6 + 4 = 10]

INDIAN STATISTICAL INSTITUTE

Semestral Examination

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

Advanced Algorithms for Graph and Combinatorial Optimization Problems

Date: 2.12.2016

Maximum Marks: 100

Duration: 3.5 Hours

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

1. Show that the minimum vertex cover of a bipartite graph G can be obtained by computing the maximum matching in G . [10]
2. Consider the problem of computing the *maximum independent set* of an undirected and unweighted tree T , denoted by $MIS(T)$. We fix a vertex r as the root. Now, define T_v as the subtree rooted at v (surely $T_r = T$). Define $I(v) = |MIS(T_v)|$. Justify the following recurrence:

$$I(v) = \max \left(\sum_{u \in C(v)} I(u), 1 + \sum_{w \in GC(v)} I(w) \right),$$

where $C(v)$ is the set of children of v in T and $GC(v)$ is the set of grand-children of v .

- (a) Using this recurrence or otherwise, design an algorithm for computing $MIS(T_r)$ in $O(n)$ time specifically writing the steps of the algorithm, where n is the number of vertices in T .
 - (b) Justify the time complexity of your algorithm. [10+5=15]
3. Consider a reduction of the vertex cover problem for a graph $G = (V, E)$ to a max-cut problem of another graph $G' = (V', E')$, where the vertices and edges of G' are $V' = V \cup \{w\}$ and $E' = E \cup E''$. Here w is a newly added vertex, and E'' consists of $degree(v) - 1$ parallel edges from each vertex $v \in V$ to the newly added vertex w . Now, prove the following claim:

The cut $(U, (V \setminus U) \cup \{w\})$ will be a max-cut in G' if and only if U is the maximum sized vertex cover in G . [12]

4. Let G be an n -vertex planar graph. Given an ϵ satisfying $0 < \epsilon \leq 1$, one can design an algorithm \mathcal{A} taking $O(n \log n)$ time for finding some set C of $O(\sqrt{\frac{n}{\epsilon}})$ vertices whose removal leaves G with no connected component with more than ϵn vertices. It can be proved that such a set C always exists.
 - (a) Using \mathcal{A} or otherwise, design a polynomial time algorithm to compute an independent set I of the graph G such that the relative error, $\frac{|I^*| - |I|}{|I^*|}$, is $O\left(\frac{1}{\sqrt{\log \log n}}\right)$, where I^* is a maximum independent set of G .
 - (b) Justify the approximation factor and time complexity of your algorithm. [8+(4+3)=15]
5. (a) If I and D denote respectively any *maximal* independent set and a *minimum* dominating set of a graph G , then show that $|I| \leq |D|$.
 - (b) Use (a) to propose a two factor approximation algorithm for the k -center problem. Justify the time complexity and approximation factor of your algorithm. [5+10=15]

6. Consider a set of polygonal objects P_1, P_2, \dots , with a total of n vertices inside an axis-parallel rectangular area in \mathbb{R}^2 . The objective is to stab these objects by minimum number of horizontal and vertical lines.

Using LP-rounding technique, design a 2-factor approximation algorithm for this problem.

Justify the approximation factor and the time complexity of your algorithm. $[10+(5+5)=20]$

7. (a) Define k -tree and partial k -tree.
- (b) Show that a subgraph G' of a chordal graph G with the size of the largest clique $w(G) = k + 1$ is a partial k -tree.
- (c) Also show that if G' is a partial k -tree, then it has a tree decomposition of width at most k (clearly explain all the steps in your proof). $[(2+2)+8+12=24]$

Indian Statistical Institute

Semestral Examination: 2016

M. Tech. in Computer Science

Mobile Computing

Date: 2-12-2016

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) State the main difference between cooperative and local spectrum sensing techniques in cognitive radio networks. [6]
(b) What are the different types of cognitive capabilities with which a cognitive radio user should be equipped to support dynamic spectrum access? [6]
(c) Briefly describe a cooperative spectrum sensing technique in cognitive radio networks. [8]
2. (a) What is flooding? Explain how *expected zone* and *request zone* concepts help to reduce route request flooding in location-aided routing (LAR) protocol in an ad hoc network. [3+9=12]
(b) What is multicasting in an ad hoc network? Explain how multicast routes are established and updated by the source in on-demand multicast routing protocol in an ad hoc network. [3+9=12]
(c) What are the main differences between proactive and reactive routing protocols in an ad hoc network. [6]
3. (a) Describe how cluster-heads are determined in low-energy adaptive clustering hierarchy (LEACH) protocol in wireless sensor networks. [10]
(b) Describe the directed diffusion (DC) routing in wireless sensor networks (WSNs). [10]
(c) Write an approximation algorithm for the following minimum relay node placement problem in wireless sensor networks:
Given a set of sensor nodes S deployed within a region and a uniform communication radius d , the problem is to place a set of relay nodes R such that the whole network G is connected. The objective of the problem is to minimize $|R|$, where $|R|$ denotes the number of relay nodes in R . [10]
4. (a) Briefly describe the four main types of *device tier* communications in 5G cellular networks. [12]
(b) Formulate the overall throughput optimization problem of device-to-device (D2D) communications as a nonlinear integer programming problem. [12]
(c) Briefly describe the issue of *mode selection* in cellular network assisted D2D communications. [6]

INDIAN STATISTICAL INSTITUTE

Semestral Examination

M. Tech (CS) - II Year, 2016-2017 (Semester - IV)

Optimization Techniques

Date : 05.12.2016

Maximum Marks : 100

Duration : 3.5 Hours

Note: The question paper is of 120 marks. Answer as much as you can, but the maximum you can score is 100.

Vectors are written in lower case with boldface, e.g. \mathbf{b} ; matrices are written in upper case, e.g., A . Transpose of A is denoted by A^T , and transpose of \mathbf{b} is denoted by \mathbf{b}^T . Whenever we say that, \mathcal{P} is a linear program, we mean \mathcal{P} is of the form

$$\begin{aligned} \text{Maximize} \quad & \mathbf{c}^T \mathbf{x} \\ \text{subject to} \quad & A\mathbf{x} \leq \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0} \end{aligned}$$

(Q1) Give the mathematical program formulations for the following:

- (i) Suppose that in an application, the feasible solutions must obey the disjunction of the two constraints: $x \geq a$ or $y \geq b$. Show how you can encode this as a set of constraints.
- (ii) Suppose that we have a constraint of the form $x \in \{a_1, a_2, \dots, a_n\}$. Show how you can encode this as a set of constraints. [5+5=10]

(Q2) Let $\mathcal{J} = \{J_1, J_2, \dots, J_n\}$ be a set of n jobs to be executed on m processors. Let $a_{ij} \in \{0, 1\}$, $i = 1, \dots, n$, $j = 1, \dots, m$, indicate whether job i can be processed on the j -th processor; $a_{ij} = 1$ indicates job i can be processed on the j -th processor, 0, otherwise. Any job J_i has a deadline $d_i > 0$; job J_i must complete execution within d_i time units after the beginning of its execution. For each task $J_i \in \mathcal{J}$, we have a processing time τ_i , a positive integer, and an amount R_{ji} of the j -th resource it needs, $j = 1, \dots, r$. There are B_j units of the j -th resource available at all times, $j = 1, \dots, r$. We also have a precedence relation $(\mathcal{J}, \mathcal{A})$, given in terms of a directed acyclic graph, such that $(J_i, J_j) \in \mathcal{A}$ means that J_i must complete its execution before J_j starts its own. Our goal is to minimize the finishing time of the schedule.

Express this problem as a mathematical program, preferably an Integer Linear Program. [15]

- (Q3) Let $P, Q \subseteq \mathbb{R}^n$ be convex sets and let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a strictly convex function. Suppose that x^* is an optimum solution to $\min\{f(x) \mid x \in P \cap Q\}$ and x^* lies in the interior of Q . Show that x^* is also an optimum solution to $\min\{f(x) \mid x \in P\}$. [15]
- (Q4) (i) Define a *totally unimodular* matrix (TUM).
(ii) Let A be a TUM. Is $[A|A]$ a TUM? If yes, prove it. Else, give a counterexample.
(iii) Show that the incidence matrix of a bipartite graph is a TUM. [2+5+8=15]
- (Q5) (i) Write mathematical programs for *maximum flow* and *minimum cut* problems. Are both of these linear programs?
(ii) Using duality, prove the max-flow min-cut theorem.
(iii) Let (X, \bar{X}) be the minimum cut. Using the complementary slackness conditions, show that arcs going from X to (\bar{X}) are saturated with flow and the reverse arcs carry no flow. [(2+2+2)+10+9=25]
- (Q6) Consider the following rock-paper-scissors game between Player A and Player B. Player A has three options to play: rock, paper and scissors and Player B has two options: rock and paper. The payoff matrix is as follows:

	rock	paper
rock	0	-2
paper	2	0
scissors	-2	2

Formulate the linear programs to find out the optimal mixed strategies for Player A and Player B. There is no need to solve the linear programs. [10]

- (Q7) Describe the steps of ellipsoid method of solving linear programs by bringing out the following characteristics of the algorithm:
- (i) how can an optimization problem be converted to a feasibility problem?
(ii) how is the starting ellipsoid chosen?
(iii) The ellipsoid volumes keep shrinking in the algorithm by a ratio that is bounded above by $e^{-1/(2n+2)}$, where n denotes the dimension of the polyhedron corresponding to the feasibility region. Analyze the time complexity of the ellipsoid method showing how this shrinking ratio plays a role. [5+(3+2+5)=15]
- (Q8) (i) Show how the solution to a system of linear equations $Ax = b$ is related to the minimization of the function given in the quadratic form $f(x) = \frac{1}{2}x^T Ax - b^T x + c$ by bringing out the nature of the matrix A .
(ii) Using the nature of the matrix A deduced above, describe the method of *conjugate direction* for solving the unconstrained minimization problem given by $f(x)$. [5+10=15]

INDIAN STATISTICAL INSTITUTE

Final Examination:(2016-2017)

MTech C.S. 2nd Year

Digital Signal Processing

Date: 5.12.2016

Maximum Marks: 100

Duration: 3 hours

Note: The marks add up to 114. The maximum you can score is 100. Use of calculators is permitted.

1. $x[n]$ and $y[n]$ are real-valued sequences of length N each. Consider $v[n] = x[n] + jy[n]$. Let $V[k]$, $X[k]$ and $Y[k]$ denote the N -point DFTs of $v[n]$, $x[n]$ and $y[n]$ respectively.

- (a) Explain how you would obtain $X[k]$ and $Y[k]$ from $V[k]$.
- (b) Now consider a sequence $a[n]$ of length $2N$ such that $a[2n] = x[n]$ and $a[2n + 1] = y[n]$. Express $A[k]$ the $2N$ -point DFT of $a[n]$ in terms of $X[k]$ and $Y[k]$.

[5+5]

2. Draw a flow-graph for computing the DFT of a sequence of length 15 using three 5-point DFTs and some multipliers. [10]
3. A filter has a system function

$$H(z) = \frac{2 - 0.4z^{-1} - 0.6z^{-2}}{1 + 0.2z^{-1} - 0.15z^{-2}}$$

Give realizations of this function in Direct Form II, Cascade and Parallel forms (one each). [5+5+5]

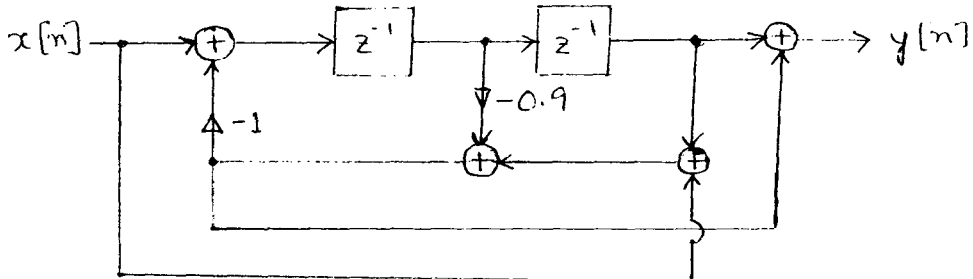
4. (a) The following causal IIR system function is designed using the impulse invariance method with $T = 0.2\text{sec}$. Determine the corresponding causal analog impulse response.

$$H(z) = \frac{3z}{z - e^{-1.5}} + \frac{4z}{z - e^{-1.8}}$$

- (b) Determine the weighting function $W(\omega)$ to be used to design a Type I linear-phase FIR highpass filter using the Parks-McClellan method to meet the following specifications: $\omega_p = 0.7\pi$, $\omega_s = 0.55\pi$, $\delta_p = 0.03808$ and $\delta_s = 0.0112$.

[10+5]

5. For the structure shown below



- Determine the system function $H(z)$. Can this be used as an all-pass filter? Justify your answer.
- Determine the output noise variance due to product round-off before addition. Assume all coefficients to be represented as fractions in a signed two's-complement fixed-point format with a wordlength of $B + 1$ bits. Assume that the multiplier of -1 does not generate any noise, and B is a positive integer.
- The filter $H(z)$ is now cascaded with another filter with system function

$$G(z) = \frac{-0.5 + 0.2z^{-1} + z^{-2}}{1 + 0.2z^{-1} - 0.5z^{-2}}$$

which is realized using the Direct Form II canonical structure. Calculate the output noise variance of the combined structure due to product round-off as above.

[(5+3)+10+8]

6. Consider the sequence

$$x[n] = \begin{cases} 1 & 0 \leq n \leq 8 \\ 0 & \text{otherwise} \end{cases}$$

Its z-transform $X(z)$ is sampled at seven points $\omega_k = 2\pi k/7$, $0 \leq k \leq 6$ on the unit circle, yielding the frequency samples

$$\tilde{X}[k] = X(z)|_{z=e^{j2\pi k/7}}, \quad 0 \leq k \leq 6$$

Determine, without computing $\tilde{X}[k]$, the periodic sequence $\tilde{x}[n]$ and its period, whose Discrete Fourier Series coefficients are given by $\tilde{X}[k]$.
[10]

7. (a) Determine the output of the filter with impulse response $h[n] = \alpha^n u[n]$ for an input $x[n] = \beta^n u[n]$, where $u[n]$ is the unit-step sequence.

(b) Evaluate

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

without actually computing $X(e^{j\omega})$ where $X(e^{j\omega})$ is the DTFT of $x[n] = n, 0 \leq n \leq 4$.

[5+8]

8. A causal LTI FIR discrete-time system has impulse response $h[n] = a_1\delta[n-2] + a_2\delta[n-1] + a_3\delta[n] + a_4\delta[n+1]$. Can it have zero phase? Justify your answer. [5]
9. The linear convolution of two sequences of lengths 110 and 1300 is to be computed using 128 point DFTs and IDFTs. Determine the smallest number of DFTs and IDFTs required for the task, using the overlap-save approach.

[10]

INDIAN STATISTICAL INSTITUTE

Semestral Examination: (2015 – 2016)

M.Tech. (CS) II Year

Parallel Processing: Architectures and Algorithms

Date: 6/12/2016

Total Marks: 116

Duration: 3 hrs

NOTE: You may answer all questions but the maximum marks you may score is 100.

1. Show the schematic diagram of a Parallel Random Access Machine (PRAM), and mention the underlying assumptions of the model. Compare the SIMD model with the PRAM model.
Show that the CRCW model of PRAM with N processors can always be simulated on the EREW model having N processors with an $O(\log N)$ -fold increase in the processing time. What is the associated space complexity?
[3+3+4+2=12]

2. a) Prove that a binary tree with $(N-1)$ nodes is not contained in an N -node hypercube H_n , ($N=2^n$), for $n > 2$.
b) Draw the block diagram of a 3-stage non-blocking $N \times N$ Clos network using $(n \times m)$ switches at the input stage, $N = n.r$, where r is an integer. Prove that $(2n-1)$ middle-stage switches are sufficient for non-blocking operation of the network.
[6 +4+6 = 16]

3. Describe Batcher's odd-even merge procedure for merging two sorted sequences of length n and prove its correctness.
Draw a network for sorting 8 elements using Batcher's odd-even merge technique.
Derive an expression for the time required for sorting n elements by Batcher's odd-even merge technique.
[6+4+5+5=20]

4. Given two polygons Q and R with m and n edges respectively, $m \leq n$, describe a parallel algorithm to find if Q and R intersects in $O(\log n)$ time. Assume that the polygons are represented by the edges. Find the AT-cost of the algorithm.
[12 + 4=16]

5. Consider an $n \times n$ mesh of processors where the boundary processors in row 1 and column 1, are only capable of handling input operations. Design an $O(n)$ algorithm for multiplying two $n \times n$ matrices A and B on this architecture. Justify that this is the fastest possible algorithm for the given architecture.
[12 + 4 = 16]

6. Explain how the Newton's method for solving non-linear equations can be implemented efficiently on a CRCW SM MIMD computer to solve the equation $f(x) = 0$. Assume that the equation $f(x) = 0$ has one and only one root in the given interval (a, b) . Write down the procedure, and mention why concurrent-read and concurrent-write options are needed for the memory.
[12+4 =16]

P.T.O

7. Given the following CUDA Kernel,

```
1  __global__ void function(const int *array,int *partial,const size_t N)
2  {
3  extern __shared__ int shared_data[1024];
4  unsigned int thread_id=threadIdx.x;
5  unsigned int index = blockIdx.x * blockDim.x + threadIdx.y;
6  if(index<N)
7      shared_data[thread_id]=array[index];
8  __syncthreads();
9  for(int offset=blockDim.x/2;offset>0;offset >>=1)
10 {
11     if(thread_id<offset)
12     {
13         if(shared_data[thread_id+offset]>shared_data[thread_id])
14             shared_data[thread_id]=shared_data[thread_id+offset];
15     }
16 }
17 if(thread_id==0)
18     partial[blockIdx.x]=shared_data[0];
19
20     __syncthreads();
21 }
```

- Find out the syntax errors in the code.
- Make necessary corrections. What will be the output of the corrected version?
- Describe in brief the operations performed by each of the steps.
- Is this an optimized code? Justify your answer.

[6+6+3+5=20]

INDIAN STATISTICAL INSTITUTE

Semestral Examination:(2016-2017)

M.TECH (CS) II YEAR

Subject Name: Quantum Information Processing and Quantum Computation

Maximum Marks: 60

Duration: 3 hours

Date: 07.12/2016

Answer any five of the following six questions

1. Let each of Alice, Bob and Charlie, staying in separate places, be supplied three bits $\{x_1, x_2, x_3\}$, $\{y_1, y_2, y_3\}$, $\{z_1, z_2, z_3\}$ respectively where $x_i, y_i, z_i \in \{0, 1\}$. The bits satisfy the following condition;

$$x_i \oplus y_i \oplus z_i = 1, \quad i = \{1, 2, 3\}.$$

They have to compute the following function;

$$f(x, y, z) = x_1.y_1.z_1 \oplus x_2.y_2.z_2 \oplus x_3.y_3.z_3$$

(i) Show that there is a classical protocol by which they can compute the function by using 3 bits of communication.

(ii) Explain how a quantum protocol can be successful using just two bits of communication.

[4 + 8]

2. a) Consider a function $f : \{0, 1\}^n \rightarrow \{0, 1\}$

Consider a unitary gate U_f which acts in the following way:-

$$U_f|x\rangle|y\rangle = |x\rangle|y \oplus f(x)\rangle$$

where $x \in \{0, 1\}^n$ and $y \in \{0, 1\}$

Show that

$$U_f|x\rangle\frac{1}{\sqrt{2}}[|0\rangle - |1\rangle] = (-1)^{f(x)}|x\rangle\frac{1}{\sqrt{2}}[|0\rangle - |1\rangle].$$

b) Consider a function $f : \{0, 1\}^n \rightarrow \{0, 1\}$, where the function f is either constant or balanced ($f(x) = 0$ for half of the possible input values). Describe a quantum algorithm by which the function can be shown to be either constant or balanced by just one implementation of a quantum circuit.

[4 + 8]

3. The function $f : \{0, 1\}^n \rightarrow \{0, 1\}$ is such that;
 $f(x) = 1$, for $x = \omega$ and $f(x) = 0$ for $x \neq \omega$

- a) In the classical world, how many queries are required to find ω ?
- b) Show how a quantum algorithm can provide a quadratic speed up for this search problem.

[2+10]

4. Consider a 2 to 1 function $f : \{0, 1\}^n \rightarrow \{0, 1\}^n$. The function has a period given by an n -bit string a : that is

$$f(x) = f(y) \text{ iff } y = x \oplus a$$

- a) Discuss how hard it is to find the period a in the classical world.
- b) Show that there is a quantum algorithm by which the period can be found in polynomial time.

[3 + 9]

5.a) Consider a qubit unitary operator U where $U = U_1 \sigma_x U_2 \sigma_x U_3$ where U_1 , U_2 and U_3 are also unitary operators acting on a qubit. Show how (controlled) $C - U$ can be realized by two qubits CNot gates and single qubit gates.

b) Consider the following three qubit quantum state

$$|\psi\rangle = \frac{1}{\sqrt{2^3}} \sum_{y=0}^{2^3-1} e^{2\pi i \omega y} |y\rangle$$

where $\omega = .x_1 x_2 x_3$ in the binary system i.e. $x_i \in \{0, 1\}$ for $i = 1, 2, 3$. Show how ω can be determined by using a quantum circuit (use diagram).

[5 + 7]

6. a) Show how phase error in computational basis can be converted to bit error.
b) Discuss why there can not be a quantum error correcting code with less than 5 qubits.
c) Show how using the following set of stabilizer operators

$$\{Z_1Z_2, Z_2Z_3, Z_4Z_5, Z_5Z_6, Z_7Z_8, Z_8Z_9, X_1X_2X_3X_4X_5X_6, X_4X_5X_6X_7X_8X_9\},$$

one can find a nine qubits quantum code that can correct one qubit error.

[2+3+7]

Course Name : M.TECH. (CS) YEAR II
 Subject name : LOGIC FOR COMPUTER SCIENCE
 Date : 07.12.2016 Maximum Marks : 50 Duration : 3 hours
 Note, if any : Open class reference book, class notes examination

Answer any 10 questions. Notations are used as in the class.

1. Let $X = \{p_n : n \in \mathbb{N}\} \cup \{\wedge\}$. Show that the set of all finite strings over X is countable. [5]
2. Show that the following formulas are theorems of classical propositional logic:
 - (a) $(\alpha \vee \beta) \leftrightarrow \neg(\neg\alpha \wedge \neg\beta)$
 - (b) $((\neg\alpha \rightarrow \neg\beta) \wedge (\neg\alpha \rightarrow \beta)) \rightarrow \alpha$ [2.5 + 2.5 = 5]
3. Give conjunctive and disjunctive normal forms of the following formulas:
 - (a) $((\alpha \rightarrow \beta) \rightarrow \beta) \rightarrow \beta$
 - (b) $(\alpha \rightarrow (\alpha \wedge \neg\beta)) \wedge (\beta \rightarrow (\beta \wedge \neg\alpha))$ [2.5 + 2.5 = 5]
4. (a) Give a criterion for a formula in conjunctive normal form to be a tautology. Justify your answer.
 (b) Let Γ be a maximal consistent set of formulas in classical propositional logic. Show that $\Gamma \vdash \alpha$ iff $\alpha \in \Gamma$. [2.5 + 2.5 = 5]
5. Are the following sets of formulas independent? Justify your answer.
 - (a) $\{\alpha, \beta, \alpha \rightarrow \gamma, \gamma \rightarrow \beta\}$
 - (b) $\{\alpha \rightarrow \beta, \beta \rightarrow \gamma, \gamma \rightarrow \alpha\}$ [2.5 + 2.5 = 5]
6. In each of the following cases, find a suitable first order language and give axioms in it for the given collections of structures:
 - (a) Sets of size 3
 - (b) Equivalence relation
 - (c) Bipartite graphs [1.5 + 1.5 + 2 = 5]
7. Check whether the following formulas are valid:
 - (a) $\forall x \exists y (R(x, y) \rightarrow R(y, y)) \rightarrow \exists y (R(y, y) \rightarrow R(y, y))$.
 - (b) $(\exists x. P(x) \rightarrow \exists y \forall z. R(z, f(y))) \rightarrow ((\exists x. P(x) \rightarrow \forall y \neg \forall z. R(z, f(y))) \rightarrow \forall x \neg P(x))$. [2.5 + 2.5 = 5]
8. Consider a first order language \mathcal{L} with equality whose vocabulary consists of only a two-place predicate symbol P . Let \mathfrak{A} be a finite structure corresponding to the language \mathcal{L} . Let \mathfrak{B} be another structure corresponding to the language \mathcal{L} such that \mathfrak{A} and \mathfrak{B} are elementarily equivalent. Are these two structures isomorphic? Justify your answer. [5]
9. (a) Let Σ be a set of first order sentences, and κ be a class of first order structures. Show that (i) $\Sigma \subseteq Th(Mod(\Sigma))$, and (ii) $\kappa \subseteq Mod(Th(\kappa))$
 (b) Is the theory of infinite sets complete? Justify your answer. [3 + 2 = 5]
10. Check by resolution method whether the following holds:

$$\{\exists x(\sigma(x) \wedge \forall y(\varphi(y) \rightarrow \alpha(x, y))), \forall x(\sigma(x) \rightarrow \forall y((\varphi(y) \wedge \eta(y)) \rightarrow \neg\alpha(x, y))), \forall x((\varphi(x) \wedge \psi(x, a)) \rightarrow \forall z(\sigma(z) \rightarrow \alpha(z, x)))\} \models \forall x((\varphi(x) \wedge \psi(x, a)) \rightarrow \neg\eta(x))$$
 [5]
11. Show that in the structure $(\mathbb{N}, +)$ with equality, the following relations are definable:
 - (a) $<$: (b) Zero: (c) Successor. [2 + 1 + 2 = 5]
12. What are the definable subsets of \mathbb{N} in $\mathfrak{N}_S = (\mathbb{N}, 0, S)$? Give justifications. [5]

INDIAN STATISTICAL INSTITUTE

Semestral Examination (2016)

M. Tech. (Computer Science) Second Year

Natural Language Processing

Date: 08.12.2016

Time: 2 h 30 min

Total Marks: 57

Maximum Score: 50

NOTE: Answer all questions.

1. (a) Assume the monogram and bigram statistics for a language with the alphabet set: $\{\alpha, \beta, \gamma, \epsilon, \phi, \gamma\}$ are as follows:

Monogram statistics: each symbol is equally likely, i.e. each occurs with $1/6$ probability.

Bigram relative frequencies in the language (frequencies of the other bigrams are wither NOT known or Zero):

$$\alpha\beta = 1.5, \beta\gamma = 0.75, \beta\epsilon = 0.25, \epsilon\phi = 0.125, \alpha\gamma = 0.0625.$$

Using the above statistics **as much as possible**, calculate the probability of the following

sentence: $\gamma \alpha \beta \alpha \gamma \beta$

[7]

- (b) Direction: Label the following statement as True or False. Briefly justify your answer [no marks will be given if the justification is not proper]

Consider the English word, **bank** and its four possible translations in Hindi: कूल, बैंक, किनारा, and ढाल.

Statement: If in half of the cases, the translation of **bank** is बैंक, the following is the set of final probabilities given the maximum entropy model:

$$p(\text{कूल} \mid \text{bank}) = 1/4; p(\text{बैंक} \mid \text{bank}) = 1/4; p(\text{किनारा} \mid \text{bank}) = 1/4; p(\text{ढाल} \mid \text{bank}) = 1/4;$$

[3]

2. (a) Consider the following sentence-aligned (German – English) corpus:

das haus – the house

das buch – the book

ein buch – a book

Calculate the lexical probabilities (i.e. word alignments) after the first iteration of the EM algorithm.

[4]

(b) Consider the following English sentence:

BLEU is designed to approximate human judgement at a corpus level, and performs badly if used to evaluate the quality of individual sentences.

The above English sentence is sent to two different English-Hindi machine translation systems and their outputs are given below:

System 1: ब्लेउ एक कोष के पर मानव मूल्यांकन अनुमानित है । यह अलग-अलग वाक्यों की मान का मूल्यांकन करने के लिए इस्तेमाल अगर बुरी तरह से करता है ।

System 2: ब्लेउ एक कोष के स्तर पर मानव मूल्यांकन करने के लिए लगभग तैयार है और व्यक्तिगत वाक्यों की गुणवत्ता का मूल्यांकन करने के लिए लागू है जब खराब प्रदर्शन कर रहा है ।

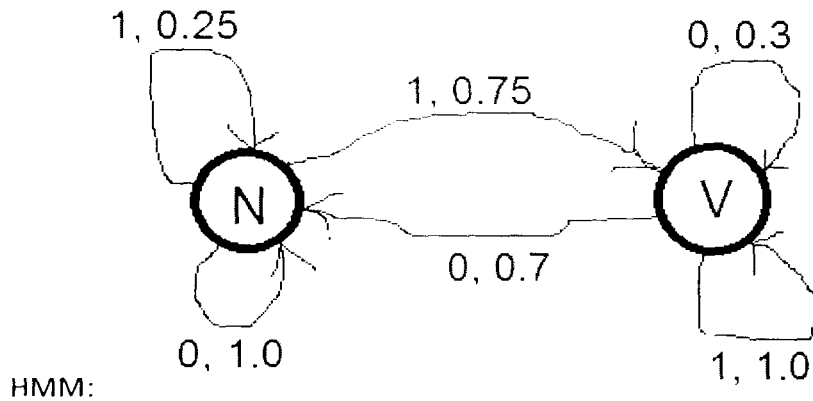
Evaluate the BLEU score for both the systems using the following Gold Reference.

Gold: ब्लेउ एक कार्पस के पर मानव निर्णय करने के लिए तैयार है . और अलग-अलग वाक्यों की गुणवत्ता का मूल्यांकन करने के लिए अच्छा परिणाम देता नहीं है ।

[Note: Consider up to 3-grams only in computing the required BLEU score]

[6]

3. (a) We can reuse HMM training algorithm for training a probabilistic regular grammar as we can have a regular grammar for an HMM (and vice versa) as follows:



Regular Grammar for the given HMM is:

$N \rightarrow 0N \quad 1.0$
 $N \rightarrow 1N \quad 0.25$
 $N \rightarrow 1V \quad 0.75$
 $V \rightarrow 0V \quad 0.3$
 $V \rightarrow 0N \quad 0.7$
 $V \rightarrow 1V \quad 1.0$

Construct the corresponding HMM for the following Regular Grammar:

$N \rightarrow 0V \quad 0.4$
 $N \rightarrow 1V \quad 0.1$
 $N \rightarrow 0N \quad 0.3$
 $N \rightarrow 0 \quad 0.2$
 $V \rightarrow 1N \quad 0.2$
 $V \rightarrow 0 \quad 0.8$

[9]

(b) Following the above regular grammar (for which you have been asked to construct the corresponding HMM), show two parse trees for the following string: **010**. Which parse is more likely in the language?

$$[(2 + 2) + 2 = 6]$$

4. (a) Why does Recurrent Neural Network (RNN) suffer from exploding/vanishing gradient?

(b) Explain a simple technique to deal with exploding gradient problem.

(c) Why do we need to use forget gate in the LSTM architecture?

$$[4 + 2 + 4 = 10]$$

5. (a) Why is it impractical to use softmax layer in Continuous Bag of Words (CBOW) and skipgram models?

(b) How does the (i) hierarchical softmax and (ii) negative sampling mitigate the above problem?

(c) Let there be a total of V words in a corpus given by w_1, w_2, \dots, w_V . Let p_i denote the probability of word w_i calculated using hierarchical softmax. Show that $\sum_{i=1}^V p_i = 1$.

$$[4 + (3 + 3) + 2 = 12]$$

INDIAN STATISTICAL INSTITUTE

Third Semester Examination

Course Name: MTech Computer Science

Subject: Cognitive Science

Date: 09/12/2016 Maximum Marks: 50 Duration : 90 minutes

Answer the following questions.

1. In information theoretical terms as well as cognitive science, what is chunking? Explain with example.(5)
2. What are primacy and recent effects in short term memory? (5)
3. What is binocular rivalry? Describe any two different types of binocular rivalry based on stimulus property type. (5)
4. Describe serial and parallel search with graphs in the light of Treisman and Gelade's Visual Search paradigm. (5)
5. What is REM sleep? What happens to duration of REM sleep as night goes on? (5)
6. What are hypersomnia and sleep apnoea? (5)
7. Why does the sensory homunculus look like the way it does – what explains the proportional shape of its parts? (5)
8. What is correspondence bias or fundamental attribution error? Explain with example. (5)
9. What is self-serving bias? Explain with example. (5)
10. Explain consciousness from the stand-point of complexity theory. (5)

Indian Statistical Institute
Semester-I 2016-2017
M.Tech.(CS) - Second Year
December 13, 2016
Subject: Patter Recognition and Image Processing

Total marks: 120

Maximum marks: 100

Duration: 3 Hours

Answer a maximum of 100 marks (any part of any question). Answers should be precise.

1. (a) Discuss the basic differences between the working principles of single layer perceptron and minimum distance classifier. **(5 Marks)**
(b) Explain: (i) Confusion matrix, (ii) Precision, (iii) Recall, (iv) F-measure, and (v) Accuracy. **(10 Marks)**
(c) What is cross validation? Discuss different cross validation techniques. **(5 Marks)**
2. (a) The following table contains five sample data items with the distance between the elements indicated in the table entries. Draw the dendrograms and make the clusters using

	A	B	C	D	E
A	0	1	2	2	3
B	1	0	2	4	3
C	2	2	0	1	5
D	2	3	1	0	3
E	3	3	5	3	0

- (i) Single linkage
(ii) Complete linkage, and
(iii) Average linkage clustering algorithms. **(12 Marks)**
- (b) Cluster the data set {2, 4, 10, 12, 3, 20, 30, 11, 25} using k-means clustering algorithm. Assume $k = 2$ and
(i) initial means as 3 and 18; and
(ii) initial means as 10 and 20,
for clusters 1 and 2, respectively. **(8 Marks)**
3. (a) Consider a dataset in which every pattern is represented by a set of 10 features. The goal is to identify a subset of 5 or less number of features which gives the best classification accuracy on this dataset. Find out how many feature subsets would be considered for (i.e., the number of times the criterion function will be invoked) by each of the following feature selection algorithms before finding out a solution.
(i) Exhaustive search. **(4 Marks)**
(ii) Sequential Backward Selection (SBS) search. **(4 Marks)**
- (b) Discuss on filter and wrapper approaches for feature selection with their advantages and disadvantages. **(6 Marks)**

- (c) What is (i) Sequential forward selection, (ii) Plus-L, minus-R selection: and (iii) Branch and bound algorithm with respect to feature selection? **(6 Marks)**
4. (a) Construct a Huffman code for the five letters "E A M N T", which are listed in decreasing order of frequency of use. Use this Huffman code to decode '0111000101010000110'. **(10 Marks)**
- (b) Discuss on (i) histogram equalization, and (ii) histogram specification. **(10 Marks)**
5. (a) Consider the image segment shown below. Let V be the set of gray-level values used to define adjacency.

		3	1	2	1	— q
		2	2	0	2	
		1	2	1	1	
p —	—	1	0	1	2	

- (i) Let $V = \{0, 1\}$ and compute D_4 (city-block), D_8 (chess-board) and D_m (city-block, if possible, otherwise chess-board)-distances between p and q .
- (ii) Repeat for $V = \{1, 2\}$.
- (10 Marks)**
- (b) Discuss on (i) Mean, (ii) Median, (iii) Laplacian, (iv) Sobel, and (v) Robert filters. **(10 Marks)**
6. (a) Explain: (i) Global, (ii) Otsu's and (iii) Adaptive histogram thresholding methods for image segmentation. **(6 Marks)**
- (b) Define (i) Fourier transform, and (ii) inverse Fourier transform. **(6 Marks)**
- (c) Discuss on (i) Hough transform, (ii) Region growing, (iii) Region splitting and merging, and (iv) Run length coding. **(8 Marks)**

Indian Statistical Institute
Semester Examination : 2016 – 2017
Master of Technology in Computer Science, Semester III
Functional Brain Signal Processing: EEG & fMRI

Date: 14 December 2016

Maximum Marks: 100

Duration: 3 hours

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

1. What is the net magnetic field in a voxel of an MR image? How many different types of structural MR imaging are possible and what they are? Describe each of the imaging modes. Describe how grey scale image of a voxel is created. 3 + 2 + 10 + 5 = 20
2. What is functional MR imaging? Which among the structural and the functional MR imaging has poorer signal to noise ratio and why? Describe how fMRI depends on four major parameters or aspects. 5 + 3 + 3 x 4 = 20
3. Describe in detail instantaneous phase of a signal. What is phase synchronization? How statistical significance of phase synchronization between two signals is determined? Try to combine phase synchronization with statistical significance in a way that the value of phase synchronization remains between 0 and 1 (both inclusive). 6 + 2 + 12 = 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, possibly with illustrative diagrams. 20
5. (a) We know in most real life signals Fourier transformation is not 'perfect', in the sense that resultant Fourier series will not converge to the signal at every time point – please explain this point. 10
(b) However, Fourier transformation is everywhere in MR imaging and it is the most instrumental theoretical tool behind precise localization of the image. Why then the Fourier transformation works so perfectly in case of MR imaging? A detailed explanation is required incorporating Larmour equation, slice selection (tell only the role of FT), frequency encoding (only the role of FT) and phase encoding (only the role of FT). 10

Indian Statistical Institute

M.Tech (CS) II

Information and Coding Theory

Semester Examination

Maximum Marks: 80

Date: December 11, 2016

Time 3 hours

The question paper contains 7 questions. Total marks is 90. Maximum you can score is 80. Unless otherwise mentioned, all notations are the same as presented in class

1. Show that Hamming distance follows the triangle inequality: for any vectors $\mathbf{x} = (x_1, x_2, \dots, x_n)$, $\mathbf{y} = (y_1, y_2, \dots, y_n)$, $\mathbf{z} = (z_1, z_2, \dots, z_n)$,

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

2. Show that Hamming codes are perfect single error-correcting codes. (5)

3. (a) Construct a BIBD from a Hadamard matrix of order $4m$. What are the parameters of the design?

(b) Construct a Hadamard matrix of order 12.

(c) Construct a $(7, 16, 8)$ Hadamard code. (5+10+10=25)

4. What are MDS codes? Prove that Reed-Solomon codes are MDS linear codes. Write a decoding algorithm for Reed-Solomon codes and prove the correctness of the algorithm.

(2 + 5 + 5 + 8 = 20)

5. Consider a double-error correcting binary BCH code of length 15. What are the parameters of the code? Write the (a) generator polynomial (b) parity check matrix of the code.

(1 + 5 + 4 = 10)

6. Show that a (n, m) -Cauchy-Reed-Solomon code can tolerate at most m -erasures. Construct a $(5, 2)$ -Cauchy-Reed-Solomon code over $GF(2^3)$. (5 + 10 = 15)

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)

INDIAN STATISTICAL INSTITUTE

First Semester Examination: 2016-2017

M. Tech. (CS) II year

Data Mining and Knowledge Discovery

Date: 14.12.2016

Maximum Marks: 100

Duration: 3 hours

[Answer as much as you can]

1. (i) What is Minkowski distance?
(ii) What is Jaccard coefficient?
(iii) What are some of the major clustering approaches?
(iv) What are some of the limitations of k-means clustering?
(v) Describe algorithm PAM. [3+3+4+4+6=20]

2. (i) What is the advantage of density-based clustering?
(ii) Describe algorithm DBSCAN.
(iii) What is hierarchical clustering?
(iv) Enumerate the different ways of computing inter-cluster similarity. [4+8+4+6=22]

3. (i) What are support and confidence in the context of association rules?
(ii) How does FP-tree help in rule mining? Explain tree construction with an example.
(iii) Define some cluster validity indices. [5+9+6=20]

4. (i) Compare and contrast (providing specific differences, advantages and disadvantages) between the following optimization algorithms: Adagrad, RMSprop, Adadelta.
(ii) What is the purpose of momentum? What are the differences between standard Momentum and Nesterov Momentum?
(iii) Explain why early-stopping prevents overfitting? Would you expect it to work in the context of linear regression?
(iv) When you are training a neural network with some form of stochastic gradient descent, the learning rate needs to be chosen. (a) What problem will result from using a learning rate that is too large, and how can one detect that? (b) What is the problem if the learning rate is too small, and how can that be detected?
(v) In mini-batch gradient descent which is more important: "the number of mini-batch updates" or "visiting all the training data", and why?
(vi) Why is over-fitting of a network with a small training set considered to be a good sanity check? [6+5+4+3+3+2=23]

P. P. C.

5.

- (i) Show that the softmax regression model is over-parameterized, and in the special case where $k = 2$ (number of classes) it reduces to logistic regression.
- (ii) Discuss, in brief, different feature rescaling techniques.
- (iii) What is the *vanishing gradient problem* in neural networks? Why do we need to initialize weights carefully?
- (iv) Consider the following CNN example:
Input volume size: $32 \times 32 \times 3$
Convolution layer:
 - # convolution filters: 10
 - Filter size: 5×5
 - Stride: 1
 - Pad: 2
 - (a) What will be the output volume size?
 - (b) What will be the number of parameters in this layer?
- (v) Consider that you have a set of digits 2's and 5's from the MNIST dataset. How would you discriminate 2's and 5's using only one autoencoder without any labels?
- (vi) How do we choose the size of filter for Convolutional layers? What are the types of activation functions for the filters? [$5+3+3+(2+2)+5+3=23$]

6. Write short notes on any three of the following:

[$7+7+7=21$]

- (i) Big data mining
- (ii) Hadoop
- (iii) Recommendation Systems
- (iv) PageRank
- (v) Distributed data mining

INDIAN STATISTICAL INSTITUTE
End-Semester Examination: 2016-17

Course Name: M.Tech. in Computer Science, Second Year
Subject Name: Computer Architecture

Date: 15. 12. 2016

Maximum Marks: 100

Duration: 3 hours

Answer **any four** questions:

1. **Vector Processors:**

[5 x 5 = 25]

Consider the execution of a program on a vector computer with the following latencies for various instructions:

- VLD (Vector load) and VST (Vector store): 50 cycles for each vector element; fully interleaved and pipelined.
- VADD (Vector add): 4 cycles for each vector element (fully pipelined).
- VMUL (Vector multiply): 16 cycles for each vector element (fully pipelined).
- VDIV (Vector divide): 32 cycles for each vector element (fully pipelined).
- VRSHF (Vector right shift): 1 cycle for each vector element (fully pipelined).

Assume the following:

- The machine has an in-order pipeline.
 - The machine supports chaining between vector functional units.
 - In order to support 1-cycle memory access after the first element in a vector, the machine interleaves vector elements across memory banks. All vectors are stored in memory with the first element mapped to bank 0, the second element mapped to bank 1, etc.
 - Each memory bank has an 8KB row buffer.
 - Vector elements are 64 bits in size.
 - Each memory bank has two ports (so that two load / store instructions can be active simultaneously), and there are two load/store functional units available.
- a) What is the minimum power-of-two number of banks required in order for memory accesses to never stall? Assume a vector stride of 1.
- b) The machine (with as many banks as you calculated above) executes the following program (assume that the vector stride is set to 1 as earlier).

```
VLD  V1 ← A
VLD  V2 ← B
VADD V3 ← V1, V2
VMUL V4 ← V3, V1
VRSHF V5 ← V4, 2
```

It takes 111 cycles to execute this program. What is the vector length?

- c) If the machine did not support chaining (but could still pipeline independent operations), how many cycles would be required to execute the same program? Explain your work.
- d) Consider the following modified design of the vector processor. The number of banks is reduced by a factor of 2 from the number of banks you found in part (a) above. Since loads and stores might stall due to bank contention, an arbiter is added to each bank so that pending loads from the oldest instruction are serviced first. How many cycles does the program take to execute the same program on this modified machine with chaining?

- e) Consider now the design of a second generation vector processor, with a multicore machine in which 4 vector processors share the same memory system. The number of banks is scaled up by 4 (compared to the answer you found in (a) above) to match the memory system bandwidth to the new demand. However, when the new machine is simulated with a separate vector program running on every core, the average execution time is longer than if each individual program ran on the original single-core vector machine with $\frac{1}{4}$ the number of banks. Explain why this could be happening. Suggest a modification in the shared memory hierarchy that can you make to this second generation system to alleviate this problem.

$$[(3 + 4 + 3 + 10) + 5 = 25]$$

2. **DRAM:**

- a) A memory system has 4 channels, each channel has 2 ranks of DRAM chips. Each memory channel is controlled by a separate memory controller. Each rank of DRAM contains 8 banks. A bank contains 32K rows. Each row in one bank is 8KB. The minimum retention time among all DRAM rows in the system is 64ms. In order to ensure no data is lost, every DRAM row is refreshed once per 64ms. Every DRAM row refresh is initiated by a command from the memory controller which occupies the command bus on the associated memory channel for 5 ns and the associated bank for 40ns.

Consider a span of 1.024 seconds. We define *utilization* (of a resource such as a bus or a memory bank) as the fraction of total time for which a resource is occupied by a refresh command. Answer the following questions. For each question, you can leave your answer in simplified form in terms of powers of 2 or powers of 10.

- I) How many refreshes are performed by the memory controllers during the 1024 second period in total across all four memory channels?
- II) Compute the *data bus* and *command bus* utilization, across all memory channels that is directly caused by DRAM refreshes.
- III) Compute the *bank* utilization (averaged across all banks) that is directly caused by DRAM refreshes.
- IV) The system designer wishes to reduce the overhead of DRAM refreshes in order to improve system performance and reduce the energy spent in DRAM. A key observation is that not all rows in the DRAM chips need to be refreshed every 64ms. The row refresh needs are summarized in the table below:

Required Refresh Rate	Number of Rows (overall)
64ms	2^5
128ms	2^9
256ms	All other rows

Given this distribution, if all rows are refreshed only as frequently as required to maintain their data, how many refreshes are performed by the memory controller during the 1.024 second period in total across all memory channels? What is the average bank utilization (averaged across all banks) in this case? Comment on the extra overhead needed to implement this improvement.

- b) Comment on the benefits of a bank and rank organization over a simple monolithic structure for a DRAM chip. Be precise and to the point in your answer.

3. **Caches:**

$$[(3 + 2) + (3 + 2) + (6 + 3 \times 3) = 25]$$

a) Consider the following design choices (assume same associativity) for the L1 cache:

Choice 1: A sectored 64KB cache with 64-byte blocks and 8-byte sub-blocks

Choice 2: A non-sectored 64KB cache with 8-byte blocks

- State one definite advantage of Choice 1 over Choice 2, and one definite advantage of Choice 2 over Choice 1.
- Can you comment which of the choices has faster cache hit latency? Justify.

b) Consider another alternative design choice for the cache: instead of directly using a number of bits to index into the cache, take those bits and form the index using a hash function implemented in hardware that randomizes the index.

- What type of cache misses can this idea potentially reduce? Explain briefly why.
- What is a disadvantage of this idea other than the additional hardware cost of the hash function?

c) A processor has a 4-way set associate L1 cache that can hold 4 blocks in total. The access latency to this cache is 1 cycle. The replacement policy is LRU. The processor is known *not to employ any prefetching* mechanism. The processor also has a 16-way set associate L2 cache that can hold 128 blocks in total. The access latency to this cache is 20 cycles.

A programmer writes a test program that in a loop repeatedly accesses only the following data cache blocks (assume billions of iterations are run):

A, B, C, D, E, F

where A, B, ..., F are different cache block addresses.

In the steady state (i.e. after the loop has executed for a few iterations), the programmer finds out that the **average memory access time** is **1** cycle.

Now, the programmer writes another program that in a loop repeatedly accesses only the following data cache blocks:

A, B, C, D, E, F, G, H

In the steady state (i.e. after the loop has executed for a few iterations), the programmer finds out that the **average memory access time** is **20** cycles.

I) Comment on the size of the victim cache that is present in the processor. Comment on the access latency of the victim cache and how it is accessed.

II) Based on the above information, comment (with proper reasoning) on the average memory access time in each of the following cases:

➤ A program that in a loop repeatedly accesses *only* the following data cache blocks:
A, B, C, D, E

➤ A program that in a loop repeatedly accesses *only* the following data cache blocks:
A, B, C, D, E, F, G

➤ A program that in a loop repeatedly accesses *only* the following data cache blocks:
A, B, C, D, E, F, G, H, I

4. Prefetching and Memory Scheduling:

[5 + 5 + 5 + (5 + 5) = 25]

- a) Consider the following design of a prefetcher of a machine with a single core, L1 and L2 caches and a DRAM memory system. You are required to examine different prefetcher designs and analyze the trade-offs involved.

You run an application that has the following access pattern to memory (the ones below are cache block addresses):

A A+1 A+2 A+7 A+8 A+9 A+14 A+15 A+16 A+21 A+22 A+23 A+28 A+29 A+30...

Assume this pattern continues for a long time. For all parts of this question, you are required to compute the prefetch accuracy, coverage and bandwidth overhead after the prefetcher is trained and is in steady state. Therefore, *exclude the first six requests from all computations.*

Also, if there is a request already outstanding to a cache block, a new request for the same cache block will not be generated. The new request will be merged with the already outstanding request.

- I) Consider a stride pre-fetcher that observes the last three cache block requests. If there is a constant stride between the last three requests, it prefetches the next cache block using that stride. Compute the coverage and accuracy of your stride prefetcher for this application.
- II) Consider an alternative design of the stride pre-fetcher that, on a cache block access, prefetches the next N cache blocks.

The coverage and accuracy of this prefetcher are 66.67% and 50% respectively for the above application. What is the value of N?

- III) What is the bandwidth overhead of this next-N-block prefetcher for the above application? Recall that we define the bandwidth overhead of a prefetcher as:

$$\frac{\text{Total number of cache block requests with the prefetcher}}{\text{Total number of cache block requests without the prefetcher}}$$

- b) Describe briefly the basic philosophy of the FR-FCFS policy that is implemented in modern DRAM controllers. Identify one weakness of this policy and propose an improvement to address the weakness.

5. Out-of-order, Superscalar processors:

[12 + 7 + 6 = 25]

- a) Consider the state of the Register Alias Table (RAT) and Reservation Stations (RS) for an out-of-order execution engine that employs Tomasulo's algorithm. The out-of-order machine in this problem behaves as follows:

- The frontend of the machine has a 1-cycle fetch stage and a 1-cycle decode stage. The machine can fetch 1 instruction per cycle, and can decode 1 instruction per cycle.
- The machine dispatches 1 instruction per cycle into the reservation stations, in program order. Dispatch occurs during the decode stage.
- An instruction always allocates the first reservation station that is available (in top-

to-bottom order) at the required functional unit.

- When a value is captured (at a reservation station) or written back (to a register) in this machine, the old tag that was previously at that location is *not cleared*; only the valid bit is set.
- When an instruction in a reservation station finishes execution, the reservation station is cleared.
- Both the adder and multiplier are fully pipelined. An add instruction takes 2 cycles. A multiply instruction takes 4 cycles.
- When an instruction completes execution, it broadcasts its result. A dependent instruction can begin execution in the next cycle if it has all of its operands available.
- When multiple instructions are ready to execute at a functional unit, the oldest ready instruction is chosen.

Initially, the machine is idle and the reservation stations are empty. 5 instructions are then fetched, decoded, and dispatched into the reservation stations. When the final instruction has been fetched and decoded, one instruction has already been written back. Given below is the state of the machine at this point, after the fifth instruction has been fetched and decoded.

Register Alias Table (RAT)

Reg	V	Tag	Value
R0	1		13
R1	0	A	8
R2	1		3
R3	1		5
R4	0	X	255
R5	0	Y	12
R6	0	Z	74
R7	1		7

ADD Reservation Station

	Src1			Src2		
	Tag	V	Value	Tag	V	Value
A	-	1	5	Z	0	-
B						
C						

MUL Reservation Station

	Src1			Src2		
	Tag	V	Value	Tag	V	Value
X	A	1	8	-	1	7
Y	X	0	-	-	1	13
Z	-	1	3	-	1	8

Based on the information above, determine the 5 instructions that have been dispatched into the machine in program order. The source registers for the instructions can be specified in any order. Present the instructions in the following format:

Opcode destination ← source1, source2

- b) What is the fundamental difference between dependencies through registers and dependencies through memory? Does this pose any additional problem for Out-of-order execution? How does an Out-of-order execution engine handle this problem?
- c) Provide two reasons why a superscalar microarchitecture could provide higher performance than a *same-width* VLIW microarchitecture. Provide two reasons why a VLIW microarchitecture is simpler than a *same-width* superscalar microarchitecture.

6. **Cache Coherence and Interconnects:** [3 + 2 + (3 + 3 + 2 + 2) + 5 + 5 = 25]

- a) Consider a multiprocessor system with private L1 caches and a shared L2 cache. If shared data is read by all cores every cycle, but written once per 1000 cycles by a single core, what kind of coherence protocol (**update or invalidate**) would you use? Explain your answer.
- b) Consider using the directory based coherence mechanism for implementing cache coherence. For a cache block A, the bit vector stored in the directory is all zeros. What does this tell you about the cache block?
- c) Consider a prototype processor design, which has two cores and uses the MESI cache coherence protocol for each core's private L1 caches. There are some bugs in the design of the processor's coherence modules. Specifically, the *BusRead* and *BusWrite* signals on the module occasionally *do not get* asserted when they should have, but data still gets transferred correctly to the cache. Fill in the entries in the table below, with a \checkmark , if, for each MESI state, the missing signal has no effect on correctness. If correctness may be affected, fill in a **X**. Justify briefly your answers.

State	BusRead	BusWrite
M		
E		
S		
I		

- d) Consider a multiprocessor system with 512 processors. Each processor has a 1 Megabyte private writeback cache with 64-byte cache blocks. The main memory size is 1 Gigabyte. Compute the number of bits of state you need in the entire system for coherence protocol implementation for each of the following protocols:

- Snoopy bus based MESI cache coherence protocol
- Directory based cache coherence protocol

Where do these bits reside for each protocol implementation above? Justify which of the above protocols you would choose for this system.

- e) Consider a scenario where you need to connect 625 processors, and you are considering three different topologies: bus, point-to-point network, mesh. Mention one disadvantage of each and justify which of the above you will choose for this system.

7. Write short notes on the following: [10 + 7 + 8 = 25]

- a) Runahead execution

- b) Stall Time Fairness Memory Scheduling
- c) Tightly coupled and Loosely coupled Multiprocessors

INDIAN STATISTICAL INSTITUTE

Semestral Examination : (2016 - 2017)

Course Name : M. Tech. (CS)

Year : 2nd year

Subject Name : Neural Networks & Applications

Date : December 16, 2016 Maximum Marks : 100 Duration : 3 hrs 30 mins

Answer all the questions.

1. State and prove perceptron convergence theorem. [5 + 15 = 20]

2. Consider two sets of labeled sample points in a three dimensional feature space, where the two sets correspond to two classes - class A and class B. Class A points are distributed within a sphere of centre (0, 0, 0) and radius 5 cm, while the samples in class B are in the region between two spheres, both being centred at (0, 0, 0), having the radii 5 cm and 8 cm respectively.
 - a) Write down, from your intuition, the equation of a classifying boundary separating the samples in classes A and B.
 - b) Write down a suitable transformation function so that the classes A and B, containing the corresponding sample points, become linearly separable in the transformed feature space.
 - c) Describe how the problem in (b) above can be mapped into the framework of a Radial Basis Function Neural Network (RBFNN).
 - d) Derive the learning rule for training this RBFNN. [2 + 4 + 10 + 14 = 30]

3. Consider the example of a network involving a single weight w , for which the cost function is
$$E(w) = k_1(w - w_0)^2 + k_2$$
where w_0 , k_1 and k_2 are constants. A backpropagation algorithm with momentum term is used to minimize $E(w)$. Explain the way in which the inclusion of the momentum constant α influences the learning process. [15]

4. Consider a problem of classifying a set of 224×224 colored (Red-Green-Blue) images into TEN different classes, using a Convolutional Neural Network (CNN) model with the following architecture.

- (a) The model consists of FIVE pooling layers, *viz.*, POOL_1, POOL_2, POOL_3, POOL_4 and POOL_5, each of which follows TWO convolution layers. Consider RELU being incorporated in convolution layers.
- (b) There are THREE fully connected layers, the last one being treated as the output layer.
- (c) Each convolution layer performs 3×3 convolutions with stride 1 and pad 1.
- (d) Each pooling layer performs 2×2 max pooling with stride 2 and no padding.
- (e) The number of filters in each of the convolution layers is as follows:
 - i) below POOL_1 is 64;
 - ii) below POOL_2 is 128;
 - iii) below POOL_3 is 256;
 - iv) below POOL_4 is 512; and
 - v) below POOL_5 is 512.
- (f) There are 4096 nodes in each of the first and second fully connected layers.

Now calculate the following:

- A. The number of nodes in each layer, starting from the input layer to the output layer.
- B. The number of weights between each pair of successive layers. [13 + 22 = 35]

INDIAN STATISTICAL INSTITUTE

Semestral Examination : (2016 - 2017)

Course Name : M. Tech. (CS)

Year : 2nd year

Subject Name : Neural Networks & Applications

Date : December 16, 2016 Maximum Marks : 100 Duration : 3 hrs 30 mins

Answer all the questions.

1. State and prove perceptron convergence theorem. [5 + 15 = 20]

2. Consider two sets of labeled sample points in a three dimensional feature space, where the two sets correspond to two classes - class A and class B. Class A points are distributed within a sphere of centre (0, 0, 0) and radius 5 cm, while the samples in class B are in the region between two spheres, both being centred at (0, 0, 0), having the radii 5 cm and 8 cm respectively.
 - a) Write down, from your intuition, the equation of a classifying boundary separating the samples in classes A and B.
 - b) Write down a suitable transformation function so that the classes A and B, containing the corresponding sample points, become linearly separable in the transformed feature space.
 - c) Describe how the problem in (b) above can be mapped into the framework of a Radial Basis Function Neural Network (RBFNN).
 - d) Derive the learning rule for training this RBFNN. [2 + 4 + 10 + 14 = 30]

3. Consider the example of a network involving a single weight w , for which the cost function is
$$E(w) = k_1(w - w_0)^2 + k_2$$
where w_0 , k_1 and k_2 are constants. A backpropagation algorithm with momentum term is used to minimize $E(w)$. Explain the way in which the inclusion of the momentum constant α influences the learning process. [15]

4. Consider a problem of classifying a set of 224×224 colored (Red-Green-Blue) images into TEN different classes, using a Convolutional Neural Network (CNN) model with the following architecture.
- (a) The model consists of FIVE pooling layers, viz., POOL_1, POOL_2, POOL_3, POOL_4 and POOL_5, each of which follows TWO convolution layers. Consider RELU being incorporated in convolution layers.
 - (b) There are THREE fully connected layers, the last one being treated as the output layer.
 - (c) Each convolution layer performs 3×3 convolutions with stride 1 and pad 1.
 - (d) Each pooling layer performs 2×2 max pooling with stride 2 and no padding.
 - (e) The number of filters in each of the convolution layers is as follows:
 - i) below POOL_1 is 64;
 - ii) below POOL_2 is 128;
 - iii) below POOL_3 is 256;
 - iv) below POOL_4 is 512; and
 - v) below POOL_5 is 512.
 - (f) There are 4096 nodes in each of the first and second fully connected layers.

Now calculate the following:

- A. The number of nodes in each layer, starting from the input layer to the output layer.
- B. The number of weights between each pair of successive layers. [13 + 22 = 35]

INDIAN STATISTICAL INSTITUTE

ENDTERM EXAMINATION M.TECH(CS) II YEAR

CRYPTOLOGY

Date: December 16, 2016 Maximum marks: 100 Duration: 3 hours.

The paper contains 125 marks. Answer as much as you can, the maximum you can score is 100.

1. (a) Show that the function family $f_k(x) = k \oplus x$ is not a PRF family.
 - (b) Describe by a diagram how to encrypt $4n$ bits of message with a block cipher of n bit block length using the CBC mode of operation.
 - (c) Describe briefly some advantages of the counter (CTR) mode of operation over the CBC mode.
 - (d) Suppose $f : \{0, 1\}^m \rightarrow \{0, 1\}^m$ is a preimage resistant bijection. We define a compression function $h : \{0, 1\}^{2m} \rightarrow \{0, 1\}^m$ as follows. Given, $x \in \{0, 1\}^{2m}$, write $x = x' || x''$ where $x', x'' \in \{0, 1\}^m$. Then define $h(x) = f(x' \oplus x'')$. Prove that h is not second pre-image resistant.
 - (e) Describe briefly the three different ways by which an authenticated encryption scheme can be obtained by generically composing a CPA secure encryption scheme and a secure message authentication code.
 - (f) Explain why 2 is not a valid value for the encryption exponent of RSA.
 - (g) Write a recursive version of the Euclids algorithm to find $\gcd(a,b)$, where $a \geq b \geq 0$.
- [3 × 7 = 21]

2. (a) Consider the following variant of DES called DESW

$$\text{DESW}_{K,K_1} = \text{DES}_K(m \oplus K_1),$$

where K and K_1 are 56 bits and 64 bits long respectively. Thus DESW has a total key length of 120 bits and has a block length of 64 bits. Show that the effective key length of DESW is much less than 120 bits, in particular you have to show that exhaustive key search on DESW can be done with just a little more effort than in DES.

[7]

- (b) Consider a block cipher BLOCK, whose encryption algorithm is described below

```

Algorithm BLOCK( $K, M$ ) /*  $|M|=64, |K| = 160$  bits */
 $K \leftarrow K_0 || K_1 || K_2 || K_3 || K_4$ ; /*  $|K_i| = 32$  bits */
 $L_0 || R_0 \leftarrow M$ ;
for  $i = 0$  to 4,
     $R_{i+1} \leftarrow L_i \oplus f(K_i)$ ;
     $L_{i+1} \leftarrow R_i$ ;
end for
return  $L_5 || R_5$ ;

```

Where $f : \{0, 1\}^{32} \rightarrow \{0, 1\}^{32}$ is a bijection.

- i. Write the decryption algorithm for BLOCK.
- ii. Show that if for any m randomly chosen from $\{0, 1\}^{64}$ if we know that $c = \text{BLOCK}(K, m)$ then we can decrypt any cipher created by the key K .

[3+7=10]

3. (a) Given a message authentication code Ψ , define the forgery advantage of an adversary attacking Ψ .

[5]

- (b) Let $F : \mathcal{K} \times \mathcal{M} \rightarrow \{0, 1\}^n$ be a pseudorandom function. Suppose F is being used as a message authentication code. Let \mathcal{A} be a PRF adversary for F and \mathcal{B} be an adversary who tries to forge F . Derive a relationship between the advantages of \mathcal{A} and \mathcal{B} .

[7]

- (c) Let F be a pseudo-random function. Show that the following constructions are insecure as message authentication codes (in each case $K \in \{0, 1\}^n$ is the private key) by constructing an efficient adversary with a high forgery advantage:

- i. To authenticate a message $m = m_1 || m_2 || \dots || m_\ell$ where $m_i \in \{0, 1\}^n$, compute $t = F_K(m_1) \oplus F_K(m_2) \oplus \dots \oplus F_K(m_\ell)$ as the tag.
- ii. To authenticate a message $m = m_1 || m_2 || \dots || m_\ell$ where $m_i \in \{0, 1\}^n$, do the following:

$$\begin{aligned}
 r &\stackrel{\$}{\leftarrow} \{0, 1\}^n \\
 t &\leftarrow F_k(r) \oplus F_k(m_1) \oplus F_k(m_2) \oplus \dots \oplus F_k(m_\ell) \\
 &\text{send } (r, t)
 \end{aligned}$$

[6]

4. (a) Describe the square and multiply algorithm to compute modular exponentiations.

[5]

- (b) Describe briefly with justifications how to compute the following using the above algorithm:

- i. The multiplicative inverse of an element in \mathbb{Z}_p^* .
- ii. The square root of a quadratic residue in \mathbb{Z}_p^* , where $p \equiv 3 \pmod{4}$.

[8]

5. (a) Let \mathcal{K}, \mathcal{D} be finite non-empty sets and $\{f_k\}_{k \in \mathcal{K}}$ be a family of functions, where for each $k \in \mathcal{K}$, $f_k : \mathcal{D} \rightarrow \{0, 1\}^n$. State when the family f is said to be an ϵ -almost universal and an ϵ -almost XOR universal family.

[5]

- (b) Let $\{\text{Poly}_h\}_{h \in \mathbb{F}_{2^n}}$ be a family of functions, where for each $h \in \mathbb{F}_{2^n}$, $\text{Poly}_h : \mathbb{F}_{2^n}^m \rightarrow \{0, 1\}^n$ defined by the rule

$$\text{Poly}_h(x_1, x_2, \dots, x_m) = x_1 h^{m-1} \oplus x_2 h^{m-2} \oplus \dots \oplus x_{m-1} h \oplus x_m,$$

where the multiplication is in the finite field \mathbb{F}_{2^n}

- i. Given $x_1, x_2, \dots, x_m, h \in \mathbb{F}_{2^n}$, describe an algorithm to compute $\text{Poly}_h(x_1, x_2, \dots, x_m)$. Your algorithm should not require more than m finite field multiplications.
- ii. Show that $\{\text{Poly}_h\}_{h \in \mathbb{F}_{2^n}}$ is an $\frac{m-1}{2^n}$ -almost universal hash family.
- iii. Show that $\{\text{Poly}_h\}_{h \in \mathbb{F}_{2^n}}$ is not an almost XOR universal (AXU) family.
- iv. Describe how the family $\{\text{Poly}_h\}_{h \in \mathbb{F}_{2^n}}$ can be modified to make it a AXU family. Justify your answer.

[5+5+3+2=15]

6. (a) Let $N = pq$ be a product of two distinct primes. Show that if $\phi(N)$ and N are known, then it is possible to compute p and q in time polynomial in bit length of N .

[5]

- (b) Suppose Bob has an RSA cryptosystem with modulus N and encryption exponent e_b and Charlie has a RSA cryptosystem with the same modulus N but an different encryption exponent e_c . Suppose also $\gcd(e_b, e_c) = 1$. Now, Alice encrypts the same plaintext x to send to Bob and Charlie, i.e., she computes $y_b = x^{e_b} \pmod{N}$ and $y_c = x^{e_c} \pmod{N}$, and then sends y_b to Bob and y_c to Charlie. Show that an adversary who intercepts both y_b and y_c can recover the plaintext x .

[7]

7. (a) Describe the Decision Diffie Hellman problem.

[3]

- (b) Show that the Decision Diffie Hellman problem is not hard in the multiplicative group \mathbb{Z}_p , for any odd prime p .

[7]

8. A set of users A_1, A_2, \dots, A_n and B wish to generate a secret *conference key*, i.e. all valid users should know the key, but an eavesdropper should not be able to obtain any information regarding the key. They decide to use the following protocol: Let p be a public prime and $g \in \mathbb{Z}_p^*$ be of order q , where q is a large prime such that $q|(p-1)$. The element g is also public. Now, B selects $b \xleftarrow{\$} \{1, 2, \dots, q-1\}$ and computes $y = g^b \in \mathbb{Z}_p^*$. Each user A_i picks a secret $a_i \xleftarrow{\$} \{1, 2, \dots, q-1\}$ and computes $x_i = g^{a_i} \in \mathbb{Z}_p^*$. User A_i sends x_i to B . User B responds to user A_i by sending $z_i = x_i^b \in \mathbb{Z}_p^*$.

- (a) Show that A_i given z_i (and a_i) can determine y .
- (b) Explain why y can be securely used as a conference key. You need to explain why at the end of the protocol all parties A_1, A_2, \dots, A_n and B know y , and also explain informally why an adversary cannot determine y .
- (c) Prove part (b). You need to show the following: If there exists an efficient algorithm \mathcal{A} that given the public values in the above protocol, outputs y , then there also exists an efficient algorithm \mathcal{B} that breaks the computational Diffie-Hellman assumption in the subgroup of \mathbb{Z}_p^* generated by g . Use algorithm \mathcal{A} as a subroutine for your algorithm \mathcal{B} .

[2+4+8=14]

INDIAN STATISTICAL INSTITUTE

Semestral Examination

M. Tech (CS) - II and JRF

Multi-dimensional search and Computational Geometry

Date : 16.12.2016

Maximum Marks : 50

Duration : 3 Hours

Question 1: Build the 2D-tree T for the points shown in Figure 1 with respect to the given partition and show the nodes in T that will be visited while searching for the Query Range Q . [5+5]

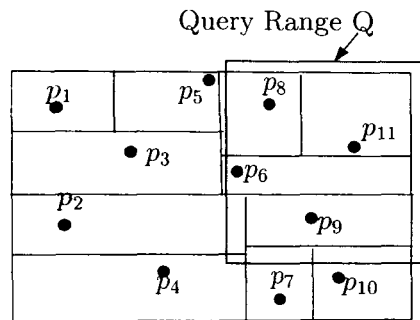


Figure 1: 2D Tree

Question 2: Build a priority search tree T for the points shown in Figure 2 and show the nodes in T that will be visited while searching for the Query Range Q . [5+5]

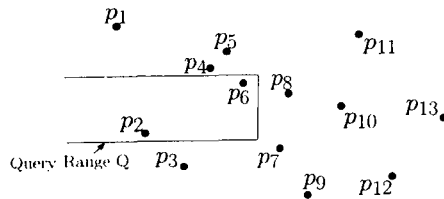


Figure 2: Priority Search Tree

Question 3: Prove that the minimum spanning tree of a set of points P is a subgraph of the Delaunay triangulation. 10

P.T.O

Question 4: Given an arrangement A of a set L of n lines, and given a line ℓ that is not in L , the zone of ℓ in A , denoted $Z_A(\ell)$, is the set of faces whose closure intersects ℓ . Prove that the total number of edges in all the cells of the zone $Z_A(\ell)$ is at most $O(n)$. 10

Question 5: We are given a set P of n points in R^d . We wish to solve the Euclidean Traveling Salesman Problem (TSP) on P . Give a $(2 + \epsilon)$ -approximation algorithm of the TSP that runs in $O(n \log n + (\frac{12}{\epsilon})^d n \log n)$ time and uses $O((\frac{12}{\epsilon})^d n)$ space, when d is a constant. 10

INDIAN STATISTICAL INSTITUTE
First Semester, 2016-2017
M.Tech (Computer Science)
Pattern Recognition and Image Processing (Back Paper)

Date: 13.01.2017

Maximum Marks: 100

Time: 3 hours

Note: Answer a maximum of six questions.

(1)

I. Define

- a) 4-adjacency,
- b) 8- adjacency and
- c) m- adjacency between pixels of an image.

II. Explain the following

- a) Histogram equalization
- b) Log transformation
- c) Power-law transformation
- d) Contrast stretching
- e) Bit-plane slicing

(5+10=15 Marks)

(2)

- I. Show that the application of a 3×3 -sized local averaging mask can be replaced by a 1×3 and a 3×1 masks applied sequentially. Compare the amount of additions that are needed in both the cases.
- II. Depict the 3×3 Sobel gradient masks in X and Y-directions. Show, for one of the Sobel masks, that it can be separated as above into two one-dimensional masks.

(8+7=15 Marks)

(3)

- I. Find the Fourier transform of the sequence $f(0)=0, f(1)=1, f(2)=0, f(3)=1$. Then calculate the inverse Fourier transform and compare the result with the original sequence. Draw the Fourier spectrum.
- II. Explain: (a) Run length coding, (b) Block truncation coding, and (c) Huffman coding with examples. Discuss their advantages and disadvantages.

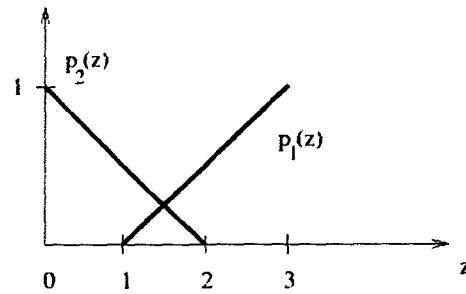
(9+6=15 Marks)

(4)

- I. Explain: (a) Global thresholding, (b) Region splitting and merging, (c) Multiple histogram thresholding methods for image segmentation.
- II. Suppose that an image has the intensity distribution as shown in the figure, where $p_1(z)$ corresponds to the intensity distribution of the objects and $p_2(z)$ corresponds to the intensity distribution of the background. Assuming that $P_1 = P_2$ (here P is a prior probability), find the optimal threshold between the object

and the background pixels.

(9+6=15 Marks)



(5)

- I. Given a one dimensional data set $\{1, 5, 8, 10, 2\}$, use the agglomerative clustering algorithm with complete linkage (Euclidean distance) to establish a hierarchical grouping relationship. What are the clusters at each level?
- II. Describe the (a) DBSCAN, (b) Divisive, (c) K-Means, (d) K-Medoids and (e) Agglomerative clustering methods.

(10+10=20 marks)

(6)

- I. Describe the algorithm of Principal Component Analysis.
- II. Explain: (a) Branch and Bound, (b) Sequential Forward, (c) Sequential Backward and (d) "Plus-L, minus-R" Selection (LRS) strategies.

(4+16=20 marks)

(7)

- I. Describe the following classifiers
 - a. Bayes'
 - b. K-NN
 - c. Minimum distance
- II. Differentiate between supervised and unsupervised learning.

(15+5=20 marks)

INDIAN STATISTICAL INSTITUTE

First-Semester Examination: 2016-2017
(Back Paper)

M. Tech. (CS) 2nd Year

Artificial Intelligence

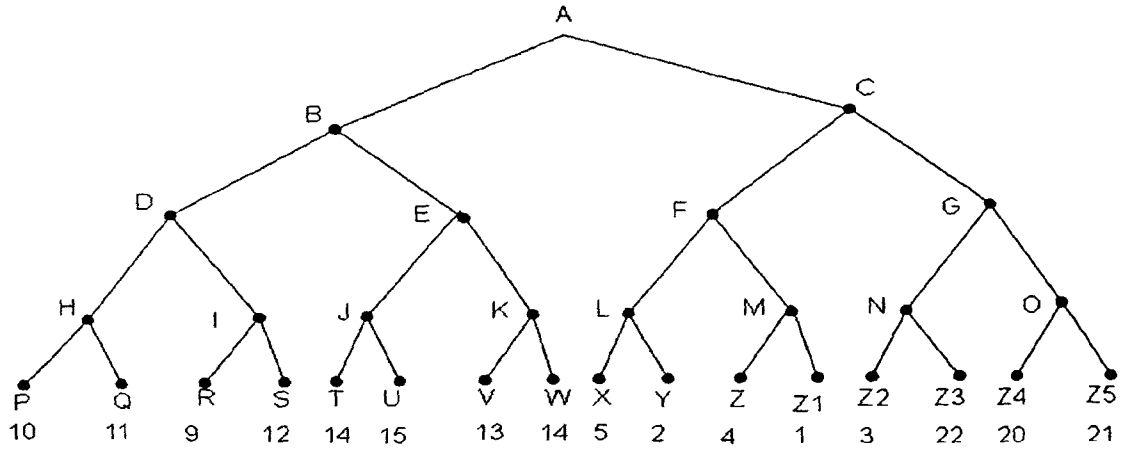
Date: 13/01/2017

Maximum Marks: 100

Duration: 3 hours

Answer all questions in brief.

1. a) Prove that if a well formed formula α is a tableau provable, then α is valid.
b) Show that "Property dealers are not happy" is tableau provable from the following set of statements: "Sales of property goes down if share market is rising. Property dealers are not happy if share market is rising. Share market is rising."
c) What are the different steps to convert a predicate expression into a clausal form?
d) Prove using semantic tableaux that the following sentences are mutually inconsistent: "All musicians are singers. A teacher is not a singer. Mary is a teacher. Mary is a musician".
[6 + 5 + 4 + 5 = 20]
2. Discuss the Stanford certainty theory with an example. What is the main difference between Dempster-Shafer theory of evidence and Bayesian reasoning?
[7 + 3 = 10]
3. a) Explain with an example the importance of cut in Prolog.
b) Write a program in Prolog for post-order traversal of a binary tree. The traversal program stores the elements of the tree in a list.
c) Write a program in Prolog to check whether a given element belongs to a binary tree.
[4 + 6 + 5 = 15]
4. Explain genetic algorithm. What are the differences between genetic algorithm and simulated annealing approach?
[5 + 5 = 10]
5. a) Explain with suitable examples the differences between (*any one of the following*)
(i) steepest ascent hill climbing approach and local beam search
(ii) red cut and green cut in Prolog
b) Describe the following with suitable examples (*any one of the following*)
(i) decision tree
(ii) fuzzy reasoning
c) What do you mean by validity and inconsistency in a well formed formula expressed in propositional logic?
[3 X 5 = 15]
6. Perform the *minimax* search procedure on the game tree shown below in which the 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 away.
[3 + 6 + 6 = 15]



7. Let $I = \langle U, A \rangle$ be a decision table, where $U = \{x_1, \dots, x_6\}$ is a nonempty set of finite objects, and $A = C \cup D$ is a nonempty finite set of attributes. Here, $C = \{\text{Headache, Muscle pain, Temperature}\}$ is the set of condition attributes and $D = \{\text{Flu}\}$ is the set of decision attributes.

U	Headache	Muscle pain	Temperature	Flu
x_1	Yes	Yes	Normal	No
x_2	Yes	Yes	High	Yes
x_3	Yes	Yes	Very high	Yes
x_4	No	Yes	Normal	No
x_5	No	No	High	No
x_6	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 the decision attribute,
- (ii) degree of dependency of a condition attribute, and
- (iii) reduct and core of the decision table.

$$[(3 + 3) + 3 + (3 + 3) = 15]$$

INDIAN STATISTICAL INSTITUTE

Second Semestral Examination (Supplementary): (2016 - 2017)

Course Name: M. Tech. (CS)

Year: 2nd year

Subject Name: Neural Networks & Applications

Date: 13 / 01 / 17 Maximum Marks: 50 Duration: 2 hrs

Answer all the questions.

1. Describe the operation of a Radial Basis Function neural network. [20]
2. Describe the operation of a principal component analysis network that incorporates Hebbian learning rule only. [10]
3. Derive expression(s) for the updated weights, in terms of the current ones, under backpropagation learning. [20]

Indian Statistical Institute

M.Tech (CS) II

Information Security and Assurance

Mid Semester Examination

Total Marks:60

Date: February 20, 2017

Time 2.5 hours

The question paper contains 6 questions. Total marks is 60. Maximum you can score is 50.

1. Prove that if the DDH problem is hard relative to group G , then the El Gamal encryption algorithm is CPA-Secure. (10)
2. Construct a one-time signature scheme using hash functions. Define and prove the security of the algorithm. (5 + 5 = 10)
3. Consider Shamir's secret sharing scheme. How can you protect against a user who shares a fake share? Write an algorithm to protect against a cheating dealer. A cheating dealer gives inconsistent shares to the users. (2 + 8 = 10)
4. Let $S = \{4, 8, 1, 34, 45, 10, 2, 16, 32, 5, 11, 6, 13, 9, 15, 19\}$. The set S is stored in an untrusted server. Describe a data structure to store S , such that integrity is preserved. How will a user who does not possess S check if the element $2 \in S$? (5 + 5 = 10)
5. How can you distribute pairwise keys between n users using polynomials? What is the security of this scheme? (7 + 3 = 10)
6. How can you construct Schnorr Signature scheme using Schnorr identification scheme? Discuss the security of Schnorr Signature scheme. (7+3=10)

INDIAN STATISTICAL INSTITUTE
M. Tech. (CS) II Year (2016-17): II Semester
Periodical Examination
ADVANCED PATTERN RECOGNITION

Date: 20-02-2017

Duration: 150 minutes

Marks: 60

Note: Answer all the questions

1. Describe the Bayes decision rule for 3-class classification problem. Show that it minimizes the probability of misclassification. [2+10=12]

2. Let

$$\begin{aligned} p_1(x) &= x, \quad x \in [0,1] \\ &= (2 - x), \quad x \in (1,2] \quad \text{and} \\ &= 0 \quad \text{otherwise} \end{aligned}$$

$$\begin{aligned} p_2(x) &= 1 + \frac{4}{5}\left(x - \frac{9}{4}\right); \quad x \in [1, \frac{9}{4}], \\ &= \frac{4(3-x)}{3}; \quad x \in (\frac{9}{4}, 3], \\ &= 0; \quad \text{otherwise} . \end{aligned}$$

be probability density functions for classes 1 and 2 respectively. Let the prior probability of class 1 be P. (a) Find the Bayes decision rule for the classification problem and find its probability of misclassification. (b) Assume that the following rule is used for classification.

x is put in class1 if $x < 1.5$. Otherwise, it is put in class 2.

Find the probability of misclassification for this rule, and show that it is more than the probability of misclassification of Bayes Decision rule. [(5+5)+(5+5)=20]

3. (a) Describe the crossover and mutation operations for genetic algorithm.
(b) Describe the algorithm for roulette wheel selection scheme for genetic algorithms.
(c) Hence, or otherwise, Describe the basic steps in the elitist model of Genetic algorithm, and show its convergence to the optimal solution as the number of iterations goes to infinity. [(6+4) + 4 +(4+10)= 28]

Information Retrieval

Mid Sem Examination
M.Tech CS 2nd Year, 2nd Semester
Full marks: 80
Time: 150 minutes

21 February 2017

Instruction: Each question carries 10 marks. For every answer, you must explain the solution.

1. Suppose we know that there are exactly 20 relevant documents in the corpus for a particular query. The precision achieved by a retrieval system for that particular query at recall points 0.1, 0.2, 0.3 and 0.4 are 1.0, 0.75, 0.5 and 0.5 respectively. Then, calculate the precision (p@5) achieved by the system after the top 5 ranked documents, or show that it cannot be inferred from the given information.
2. Given a query, a particular search system ranks the documents by a combination of their static authority score (such as PageRank) and the tf.idf style scores for the query terms. However, the ordering of the documents in the posting lists must be fixed at indexing time. Which (if any) among the below orderings of the documents in the posting lists would allow the system to perform a linear merge of the posting lists at query time? Note that there may be none, or more than one right answers.
 - (a) Ordered by descending authority scores
 - (b) Ordered by ascending document id
 - (c) Ordered by descending tf.idf scores
 - (d) Ordered by descending document id
3. Recall that a permuterm index stores all rotations of each word after adding a special ending character \$ at the end of the word. Write down all rotations generated this way from words **examination** and **explanation**. Would any of these words be found as a match for the wildcard query **ex*tion** when the permuterm index is used? Explain your answer.
4. For each of the following scenarios, explain whether skip lists would be useful in reducing the response time of a search engine.
 - (a) General Boolean retrieval where the queries are Boolean expressions using AND, OR and NOT operators involving the terms in the corpus.
 - (b) The web search query “**to be or not to be**” seeking documents containing exactly the phrase as in the query. Assume the index contains the position information of the term corresponding to each document present in the posting list of the term.

5. In Rocchio's algorithm, the modified query q_m is obtained from the original query q_0 by:

$$q_m = \alpha q_0 + \beta \frac{1}{|D_r|} \sum_{d_j \in D_r} d_j - \gamma \frac{1}{|D_{nr}|} \sum_{d_j \in D_{nr}} d_j$$

where D_r is the set of known relevant documents and D_{nr} is the set of known non-relevant documents.

Each of the scenarios below is a form of relevance feedback. For each of them, what can you say about the sets D_r , D_{nr} , and a possible set of values for the parameters α , β and γ ? Explain your answers.

- (a) User submits query to the search engine. Engine returns 10 results. Beside each result, there is a link see similar pages.
 - (b) User submits query to the search engine. Engine returns 10 results. Beside each result, there is a like button. As each like button is clicked, the results are automatically refreshed by the engine.
6. Suppose L is a list of n (document id, score) pairs sorted by document ids. Assume $k \ll n$. What is the complexity of the most efficient algorithm to determine the top k documents, ranked by score, from L ? Describe the algorithm and analyze the complexity briefly.
7. Given a very large set of integers distributed in a Hadoop cluster, how would you find the *median* of the integers? Do not assume that the total number of integers are known. Write the *map* and *reduce* algorithms with explanation of why your algorithm would work.
8. In the PageRank computation framework, a random surfer is assumed to have started at any page *with equal probability*. At every iteration she either follows an outlink with probability β , or teleports (disappears and starts again at any node) with probability $1 - \beta$. If the surfer follows an outlink, she follows any of the available outlinks with *with equal probability*. If she teleports, she starts at any node again with *equal probability*. The PageRank models a static quality or authority score, independent of the query, for every document. The query specific score (for example, based on tf.idf scores of the query terms) are then combined with the PageRank score to compute a ranking of the documents.

Suppose, in addition to the quality and term weighting scores, the engine needs to take the timestamp of a document into account; that is, a more recent document should be ranked higher than an older one, if the PageRank and term weighting score combinations for both documents are similar. Can you modify the PageRank computation approach so that this parameter is also modeled in the quality score? If yes, describe how and why your approach would work. If no, explain why it cannot be done within the PageRank framework.

INDIAN STATISTICAL INSTITUTE

Mid-Semester Examination: (2016-2017)

M.Tech C.S., 2nd Year

Advanced Digital Signal Processing

Date: 22.2.2017

Maximum Marks: 60

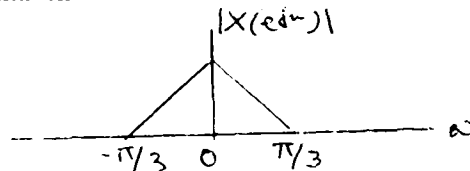
Duration: 2 hours

Note: The marks add up to 67. The maximum you can score is 60. You are permitted to use calculators.

Number of question papers required: 2

Questions:

1. 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

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

Sketch the output obtained.

[10]

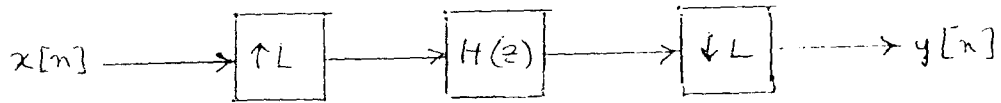
2. An FIR filter has the transfer function

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

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

3. Consider a cascade multirate system shown below with $H(z)$ having a polyphase decomposition

$$H(z) = \sum_{k=0}^{L-1} z^{-k} E_k(z^L)$$



Show that the overall system has a system function equal to $E_0(z)$. Is it time invariant? Justify your answer. [10+5]

4. A two-channel perfect reconstruction (PR) QMF filter bank has analysis filters $H_0(z)$ and $H_1(z)$ and synthesis filters $G_0(z)$ and $G_1(z)$. For a fixed $H_0(z)$, suggest possible choices for the other filters in terms of $H_0(z)$, justifying how the conditions for perfect reconstruction and no-aliasing are satisfied. Determine if the filterbank arising as a result of switching the sets of analysis and synthesis filters will also form a PR system. [10+5]
5. Design

- (a) a one-stage, and
 (b) a two-stage

decimator to decrease the sampling rate from 40 kHz to 2 kHz. The decimator is to be designed as an equiripple FIR filter with a passband edge at 800 Hz., and passband ripple = stopband ripple = 0.002 . Use Kaiser's formula to estimate the order of the FIR filters used:

$$N \cong \frac{-20 \log_{10}(\sqrt{\delta_p \delta_s}) - 13}{14.6(\omega_s - \omega_p)/2\pi}$$

[5+12]

INDIAN STATISTICAL INSTITUTE
 Mid-Semester Examination
 M. Tech. (CS) II year (2nd Sem): 2016–2017
 Advanced Cryptology

Date: 22. 02. 2017

Total Marks : 120

Time : 3 Hours

Answer as much as you can. Maximum you can score is 100.

For notational simplicity, we write \mathbb{F} to denote the finite field $\text{GF}(2^n)$ and \oplus to denote the usual field addition. We write $X \stackrel{\$}{\leftarrow} \mathcal{S}$ to mean that X is uniformly distributed over \mathcal{S} and $X \leftarrow^* \mathcal{S}$ to denote that X follows arbitrary distribution. We use $\langle x \cdot r \rangle$ to denote the inner product of two vectors x and r .

1. (a) Let us consider the following set

$$\text{Func} := \{f : \mathbb{F} \rightarrow \mathbb{F}\}.$$

Let us consider that f and g are two random functions sampled uniformly and independently from Func. Now we define the following function:

$$h : \mathbb{F} \times \mathbb{F} \rightarrow \mathbb{F},$$

defined as $h(x, y) = f(x) \oplus g(y)$, where $x, y \in \mathbb{F}$. Is the function h a secure PRF? If yes, analyse and show the security bound of the PRF. If not, demonstrate the attack.

- (b) Define one way function. Prove that if one way function exists then $P \neq NP$.
 (c) State Hoeffding inequality. Let $f : \{0, 1\}^* \rightarrow \{0, 1\}$ be a function. Let

$$F(n) := \frac{\sum_{x \in \mathbb{F}} f(x)}{2^n}$$

denote the average value of f over \mathbb{F} . Let $p(\cdot)$ be a polynomial. Present a probabilistic polynomial time (PPT) algorithm, that on input 1^n , will output an estimate $A(n)$ of $F(n)$, such that

$$\Pr \left[|F(n) - A(n)| > \frac{1}{p(n)} \right] < \frac{1}{2^n}.$$

(Hint: You may consider that the algorithm samples sufficiently many random points $v \in \mathbb{F}$ and then takes the average of f over the sampled v values. Then apply Hoeffding inequality.)

[5 + (1 + 4) + (2 + 8)]

2. (a) Explain with proper justification whether the following functions are negligible or noticeable.

(a) $n^{-\log n}$.

(b) n^{-5} .

(c) $\mu(n) = \begin{cases} 2^{-n}, & \text{if } n \text{ is even;} \\ n^{-3}, & \text{if } n \text{ is odd.} \end{cases}$

- (b) If $f(n)$ is a non-negligible function and $g(n)$ is a negligible function, then prove that $h(n) := f(n) - g(n)$ is a non-negligible function.
 (c) Define strong one way function. Define weak one way function. Justify the necessity of including the length of the output in the definition.

[(3 + 2 + 3) + 5 + (2 + 2 + 3)]

3. (a) Prove that if f is a length preserving one way function, then $g : \mathbb{F} \rightarrow \mathbb{F}$ defined as $g(x) := f(x) \oplus x$ is not a one way function.

P.T.O

(b) Let f be a one way function. We define another function $g : \mathbb{F} \rightarrow \mathbb{F}$, defined as

$$g(x) = \begin{cases} f(x), & \text{if } x \neq 0^n; \\ 0^n, & \text{if } x = 0^n. \end{cases}$$

Prove that g is a one way function.

(c) If $f : \mathbb{F} \rightarrow \mathbb{F}$ is a one way function, then prove that $g : \mathbb{F} \times \mathbb{F} \rightarrow \mathbb{F} \times \mathbb{F}$ defined as $g(x_1, x_2) := 0^n \| f(x_1)$ is a one way function.

[8 + 7 + 5]

4. (a) Prove that if there exists a weak one way function, then there exists a strong one way function.

(b) Define hard core predicate of a function. Justify why there cannot exist a single hard core predicate which is hard core for all one way functions.

[15 + (2 + 3)]

5. (a) Prove that if $R \stackrel{\$}{\leftarrow} \mathbb{F}$, and $X \stackrel{*}{\leftarrow} \mathbb{F}$, where R and X are statistically independent, then $R \oplus X \stackrel{\$}{\leftarrow} \mathbb{F}$.

(b) Let $f : \mathbb{F} \rightarrow \mathbb{F}$ be a one way function. Let us consider $g : \mathbb{F} \times \mathbb{F} \rightarrow \mathbb{F} \times \mathbb{F}$ defined as $g(X, R) = (f(X), R)$, where $R \stackrel{\$}{\leftarrow} \mathbb{F}$. Let $hc : \mathbb{F} \times \mathbb{F} \rightarrow \{0, 1\}$ be a function defined as $hc(X, R) = \langle X \cdot R \rangle$. Now prove the following.

If there exists a PPT algorithm \mathcal{A} and a polynomial $p(n)$ such that

$$\Pr[\mathcal{A}(g(X, R)) = hc(X, R)] \geq \frac{3}{4} + \frac{1}{p(n)},$$

then there exists another PPT algorithm \mathcal{A}' and a polynomial $p'(n)$ such that

$$\Pr[\mathcal{A}'(f(X)) \in f^{-1}(f(X))] \geq \frac{1}{p'(n)}.$$

Write down $p'(\cdot)$ in terms of $p(\cdot)$.

[5 + 15]

6. (a) If $X_1, \dots, X_{t(n)} \sim \text{Ber}\left(\frac{1}{2} + \epsilon(n)\right)$ are i.i.d (independent and identically distributed) random variables, then prove that

$$\Pr\left[\sum_i X_i \geq \frac{t(n)}{2}\right] \geq \left(1 - \frac{1}{2n}\right)$$

using Chernoff Bound. Write down $t(n)$ in terms of n and ϵ .

(b) How can you construct pairwise independent random variables from n many independent random variables? Justify why your constructed variables are pairwise independent?

(c) If $X_1, \dots, X_m \sim \text{Ber}\left(\frac{1}{2} + \frac{\epsilon(n)}{2}\right)$ are identically distributed pairwise independent random variables, then prove that

$$\Pr\left[\sum_i X_i \geq \frac{m}{2}\right] \geq \left(1 - \frac{1}{2n}\right)$$

using Chebyshev's inequality. Write down the expression for m in terms of n and $\epsilon(n)$.

[7 + (3 + 3) + 7]

Mid-Semester Examination (2017)

MTech (CS) – II

Subject: Computer Vision

Date: 23.02.2017 Time: 2 hours 30 mins (2.30-5.00 pm)

(Answer all the questions)

Full Marks: 60

1. Explain briefly how the geometric transformations like Translation, Scaling and Rotation are relevant to central problem of computer vision? Hence derive the transformation matrices for Translation, Scaling and Rotation and also their respective inverses. 4+12
2. Explain the advantage/s of Gaussian filtering an image. What is the significance of combining a Derivative operator to such Gaussian filtering in extracting image properties? Explain also the advantage of using Laplacian operator as the Derivative Operator & derive a digital mask for the Laplacian. 3+3+4+4
3. Using the pinhole camera model, explain what you mean by perspective projection. What is thin lens approximation? Derive the corresponding thin lens formula. 6+2+4
4. Explain the concept of receptive field. Hence explain the basic model for a spatial filter representing the visual receptive field. 5+5
5. A camera captures a rectangular object at a distance of 20 m. The dimension of the object is 50 cm x 25 cm. The object is captured in the camera in a size 100 pixels x 50 pixels. Find the focal lengths f_x and f_y of the camera. What is the field of view had the image size been 200 pixels x 100 pixels? 5+3

Indian Statistical Institute

Advanced Image Processing

M.Tech.(CS)-II Year: 2016-17

Full marks: 60

Time: 2 Hours

Date: 23.02.2017

Answer any six questions. All questions carry equal marks.

1. (a) Given a pin-hole camera with focal length λ , state the perspective projection as linear transformation to map 3-D world point to 2-D image point.

(b) Prove that the distant objects appear smaller due to perspective projection.

[5+5=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. (a) Prove that

$$(A \oplus B)^c = A^c \ominus \tilde{B}$$

where \tilde{B} is the geometric reflection of B .

(b) Prove that

$$(A \bullet B) \bullet B = A \bullet B.$$

[5+5=10]

4. (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]

5. Describe the fast Fourier transform algorithm and discuss its computational complexity.

[8+2=10]

6. (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.

$$[(5+2)+3=10]$$

7. Describe the thresholding method proposed by N. Otsu. How do you extend this method to obtain multiple thresholds?

$$[7+3=10]$$

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-Semestral Examination : 2016 – 17
 MTech CS (2nd Year)
 Computational Finance

Date: 24 February 2017

Maximum Marks: 30

Duration: 2 Hours

1. Define the following: [3 X 3 = 9]
- a) Dominant strategy
 - b) Second order continuous parameter stochastic process
 - c) Martingale

2. Assuming $V_0 > 0$, the discounted return is

$$R_n^* = [S_n^*(1) - S_n^*(0)]/S_n^*(0) \text{ for } n = 1, \dots, N$$

Show that

(a) $G^* = \sum_{n=1}^N H_n S_n^*(0) R_n^*$

(b) $R_n^* = \frac{R_n - R_0}{1 + R_0}$ for $n = 1, \dots, N$.

- (c) Q is a risk neutral probability if and only if $E_Q[R_n^*] = 0$ for $n = 1, \dots, N$.

[3+3 + 4 = 10]

3. In the two period model, explicitly solve the Consumption Investment problem for the utility function $u(w) = \frac{1}{\gamma} w^\gamma$ where $\gamma < 1$. Show that the

Lagrange Multiplier

$$\lambda = v^{-(1-\gamma)} \{E[(L/B_1)^{-\frac{\gamma}{1-\gamma}}]\}^{(1-\gamma)}$$

the optimal attainable wealth

$$W = \frac{v(L/B_1)^{-1/1-\gamma}}{E[(L/B_1)^{-\gamma/1-\gamma}]}$$

and the optimal objective value is $E[u(W)] = \lambda v/\gamma$.

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 $P(\omega_1) = 3/5$. [4 + 3 + 2 + 4 = 13]

INDIAN STATISTICAL INSTITUTE
Mid-Semestral Examination: 2016 – 17

Course Name: M. TECH CS - II

Subject: Computer Graphics

Date: 24.02.2017

Maximum Marks: 60

Duration: 2 hrs.

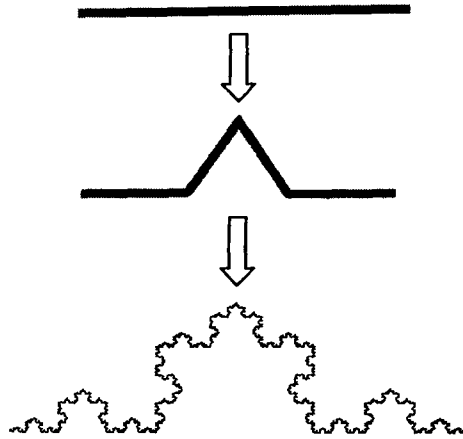
Answer as much as you can.

1. a) Calculate a 4×4 matrix to 3D rotate points about the vector $A = [1 \ -2 \ 1]^T$ by 60° . (5)
 - b) Find the transformation matrix that transforms the square ABCD whose center is at (2, 2) to half of its size, with center still remaining at (2,2). Coordinates of the vertices of the square ABCD are A(0,0), B(0,4), C(4,4) and D(4,0). Find the coordinate of the vertices of the new square. (6)
 - c) "Shearing along both X and Y directions at the same time (simultaneously) is not equivalent to the composition of pure shear along X-axis followed by a pure shear along Y-axis" - prove or disprove the statement. (4)
2. a) A window is given with the lower left corner at (1,2) and the upper right corner at (16,12). Coordinates of the vertices of a triangle in this window are given by (3,2), (10,7.5), and (5,5). Find the screen coordinates of the vertices of the same triangle when mapped into a viewport with corners (100,100) and (400,200). (5)
 - b) Illustrate the Cohen Sutherland clipping algorithm by clipping a line between points $P_1 (70, 20)$ and $P_2 (100, 10)$ against a square window with lower left hand corner (50,10) and upper right hand corner (80, 40). (10)
3. a) Derive the Bresenham's algorithm for drawing a line with only integer operations for the case $-1 \leq m \leq 0$, m denoting slope of the line. (10)
 - b) A unit square is transformed by 2×2 transformation matrix. The resulting position vectors of the vertices are the column vectors of the following matrix:
$$\begin{pmatrix} 0 & 2 & 8 & 9 \\ 0 & 3 & 4 & 1 \end{pmatrix}$$

What is the transformation matrix? (5)
4. a) Consider a raster system with the resolution of 1024×768 pixels and the color palette calling for 65,536 colors. What is the minimum amount of video RAM that the computer must have to support the above-mentioned resolution and number of colors? (2)
 - b) You want to join two Bézier curves. The first curve must start at (0,0,0) with a tangent vector of (1,1,1). The second curve must end at (5,0,-5) with a tangent vector of (-1,-1,-1). Give the geometry matrices for each curve such that the two curves join at (2,5,0) and have C^1 continuity. (8)

P.T.O

- c) What do you mean by fractal dimension? Calculate the dimension of the Koch Curve, formed as shown below and explain your steps: (2+3 = 5)



5. Develop a scan conversion algorithm to draw an ellipse whose axes are parallel to the X and Y axes of the reference coordinate system and the axes parameters are $r_x = 6$, $r_y = 8$ (r_x denoting length of the semi-minor axis parallel to X-axis and r_y denoting the length of the semi-major axis along Y axis). Explain your steps. (15)

6. Write short notes on any two of the following:

$$7\frac{1}{2} \times 2 = 15$$

- Sutherland-Hodgman Polygon Clipping Algorithm.
- Liang-Barsky Line Clipping Algorithm.
- Perspective Projections.

INDIAN STATISTICAL INSTITUTE
M. Tech. (CS) II Year (2016-17), II semester
Semestral Examination
ADVANCED PATTERN RECOGNITION

Note: This paper carries 107 marks. Answer as much as you can.

Date: 18.04.17

Duration: 210 minutes

Maximum Marks: 100

1. (a) Describe k-nearest neighbor probability density estimation procedure.
(b) Derive k-nearest neighbor decision rule using the density estimation procedure.
(c) Describe an algorithm for reducing the size of the training set for k-nearest neighbor rule. [4+6+6 = 16]
2. (a) Describe k-nearest neighbor based data condensation algorithm.
(b) Describe a dissimilarity measure between two features and state a feature selection algorithm using it. [4+7=11]
3. Suppose you have two 2-dimensional normal populations $N(\mu_1, \Sigma)$ and $N(\mu_2, \Sigma)$ where $\mu_1 = (0.0, 0.0)$, $\mu_2 = (1.0, 2.0)$, $\Sigma = \begin{pmatrix} 1 & -0.5 \\ -0.5 & 2 \end{pmatrix}$. Let the prior probabilities of the populations be 0.5 and 0.5.
 - (a) Find the Mahalanobis distance between the two populations.
 - (b) Find the Bayes decision rule for separating the two populations. Also find its probability of misclassification in terms of standard normal probabilities. [5+(5+7)=17]
4. Let $x_1 = (0.0, 0.0)$, $x_2 = (1.0, 0.0)$, $x_3 = (0.0, 1.0)$ and $x_4 = (2.0, 2.0)$. Let there be two classes and let the first two points belong to class 1, and the rest of them belong to class 2. Let $x + y = 0.5$ be the starting straight line for Perceptron algorithm. Let $\lambda = 0.05$ be the given learning rate. Apply Perceptron algorithm and with maximum number of iterations as 12 to find the discriminating straight line between the classes. Show the result after 12 iterations. [15]
5. Describe Fuzzy C means algorithm. [15]
6. (a) Describe the basic steps of gradient descent optimization algorithm.
(b) Let $f(x, y) = 4x^2 - 12xy + 9y^2 - 10$ where x and y are real numbers. Minimize f by using gradient descent technique. [3+10=13]
7. Suppose there are c classes with prior probabilities as P_1, P_2, \dots, P_c . Show that the misclassification probability for Bayes decision rule is less than or equal to $1 - \max\{P_1, P_2, \dots, P_c\}$. [5]

(P.T.O)

8. Describe DBSCAN algorithm. [5]

9. Write short notes on the following.

(a) VC dimension

(b) Support vector machine.

[5+5=10]

Indian Statistical Institute

M.Tech (CS) II

Information Security and Assurance

Semester Examination

Maximum Marks: 100

Date: April 18, 2017

Time 3.5 hours

The question paper contains 7 questions. Total marks is 105. Maximum you can score is 100.

1. Consider the elliptic curve $y^2 = x^3 + 2x + 2$ over \mathbb{Z}_{17} and a point $P = (5, 1)$ on the curve. Compute $2P$. What is the order of the elliptic curve? (5 + 10 = 15)
2. You want to share a video with all students who attended the course on “Information Security and Assurance”. Write an efficient algorithm to encrypt the video, such that only the above students can decrypt and view the content, and no one else can decrypt the video. What is the communication and computation complexity of your algorithm? Describe the security model and prove the security with respect to the model. (6 + 4 + 3 + 7 = 20)
3. Alice has input bits x_1, x_2 and Bob has input bits y_1, y_2 . Both Alice and Bob want to compute $z = (x_1 \wedge y_1) \vee (x_2 \wedge y_2)$, securely. Describe a step-by-step procedure for computing z , such that Alice does not know y_1, y_2 , Bob does not know x_1, x_2 and none other than Alice and Bob knows z . (15)
4. N communicating nodes $\mathcal{N} = \{n_1, n_2, \dots, n_N\}$ collect data and send to an aggregator node n_a . The aggregator node has to verify the integrity of data collected from \mathcal{N} . How can you do so using BLS signatures? What is the complexity of the verification algorithm? Can you design an efficient algorithm to reduce the verification time? What is the new time complexity of the new verification algorithm? (3 + 4 + 4 + 4 = 15)
5. (a) What are the challenges in storing data in untrusted servers, for example clouds?
(b) I have stored a large file F in an untrusted server. I do not have a local copy on my disk. How can I verify the integrity of F without downloading the whole file?
(c) What is the communication and computation complexity of the procedure. (5 + 10 + 5 = 20)
6. What is Pedersen commitment scheme? How is it different from El Gamal Encryption algorithm? (5 + 5 = 10)
7. Write (with proper justification) an interactive proof for graph isomorphism problem. (5 + 5 = 10)

**Indian Statistical Institute
Semester Examination (2017)
M.Tech (CS) II Year
Computer Vision**

Date: 20.04.17

Full Marks: 100

Duration – 3 hours

Attempt all the questions

1. (a) Compare the human eye and the CCD camera.
(b) Explain the difference between the CCD plane and the image buffer plane in the camera.
(c) Derive a relation between the CCD plane coordinates and the buffer coordinates.
(d) A very small celestial object is being captured through a telescope fitted with a CCD camera that forms a digital image of size 1000×1000 pixel². The object is localized at the pixel $(r, c) = (250, 650)$ where r and c represent the r^{th} row and c^{th} column respectively, the origin $(1, 1)$ being at the top left corner. The CCD plane is 100×100 mm² and the focal length is 1 m. Find the three-dimensional unit vector pointing at the celestial object (the optical center is assumed to be in the center of the pixel image).

4+4+6+6=20

2. (a) What do you mean by scene (object) radiance and image irradiance?
(b) Prove that image irradiance varies directly as the fourth power of the angle subtended by the object patch at the center of the lens with respect to the optical axis.

6+14=20

3. (a) Derive the Perspective transformation equations of the camera co-ordinates (x, y) in terms of the world co-ordinates (X, Y, Z) using the method of homogenization.
(b) Show that the world point cannot be uniquely determined from the camera co-ordinates by applying inverse transformation.
(c) Explain the statement in (b) above in the light of perspective projection.
(d) How can you solve the problem in (b) using two cameras with a lateral shift between them.

6+5+4+5=20

4. (a) Explain which of the following constraints viz. brightness constancy, reflectance constancy and smoothness play a role in estimating: (i) optical flow and (ii) lightness value.

.....contd. to page 2

(b) What is the importance of corner points as compared to edge points in an image?

(c) Derive the matrix equation of the Harris corner detector. $(5+5)+4+6=20$

5. Write short notes on any four:

(a) Hough Transform

(b) Helmholtz reciprocity condition

(c) Camera Calibration

(d) Scale space

(e) Raw Primal Sketch

$4 \times 5 = 20$

Information Retrieval

End Sem Examination
M.Tech CS 2nd Year, 2nd Semester
Full marks: 100
Time: 180 minutes

21 April 2017

1. **Term weighting:** Consider the following formulae for tf.idf for a term t in a document d , where $freq(t, d)$ is the number of times t occurs in d , $df(t)$ is the number of documents in which t occurs, N is the total number of documents in the corpus, and $\max_x(freq(x, d))$ denotes the maximum number of times any term x occurs in d . Let $sign(y)$ be the sign of a real number y (the sign is 1 for positive numbers, -1 for negative numbers and 0 if $y = 0$).

$$tf.idf_1(t, d) = freq(t, d) \times \frac{1}{df(t)}$$

$$tf.idf_2(t, d) = \left(0.5 \times sign(freq(t, d)) + \frac{0.5 \times freq(t, d)}{\max_x(freq(x, d))} \right) \times \log \frac{N}{df(t)}$$

- (a) Explain intuitively, but in detail with valid reasons, why the second formula is better than the first one.
- (b) What can be the maximum value of $tf.idf_2(t, d)$? Explain your answer.

[6 + 4 = 10 marks]

2. **Question Answering:** Recall that the AskMSR system worked by rewriting questions to form several queries so that one of the formed queries would potentially find an exact match in the corpus. Given the question: *Where was Sachin Tendulkar born?*

- (a) Describe how the AskMSR system would work for the above question. In particular, describe the procedure of rewriting the question to form queries and write down the queries generated from the question accordingly.
- (b) In the corpus, what would be a likely matching text which would help the system to answer the question. For your information, Sachin Tendulkar was born in Mumbai.

[7 + 3 = 10 marks]

3. **Index Compression:** Recall that the variable byte encoding scheme encodes integers using 1 byte units, with the first bit of every byte being used as an indicator and rest of the 7 bits as payload.

- (a) Give an example of an integer which would take more than 8 bytes to be encoded by the variable byte encoding.
- (b) How are posting lists encoded using variable byte encoding? Consider the posting list: 920, 1022, 1030. What would be the encoding of this posting list? Provide necessary explanation for arriving at your answer.

[5 + (3 + 5) = 13 marks]

4. **Social Network Graph Mining:** After the betweenness of every edge in a social network graph is computed, describe two approaches for clustering the network into communities. Also explain if any one of the approaches is likely to be more efficient than the other in practice, for large social networks.

[4 + 4 + 4 = 12 marks]

Indian Statistical Institute

Advanced Image Processing

M.Tech.(CS)-II Year, 2016-2017

Full marks: 100

Time: 3 Hours

Date: 21.04.2017

Answer any ten questions. All questions carry equal marks.

1. Suggest a morphological algorithm for computing distance transform of a binary image. [10]
2. Define conditional dilation, and hence, obtain geodesic distance between two points. [10]
3. State with the help of examples alone the three criteria of multi-scale processing which are satisfied by opening by reconstruction, but not by conventional opening. [10]
4. Derive the expression of parametric Wiener filter for image restoration using constrained least square estimation approach. [10]
5. Consider the following digital signal: [3, 5, 2, 4, 6, 4, 2, 4]. Construct the tree wavelet expansion of this signal using Haar wavelet transform. [10]
6. Consider the following block of gray levels:

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

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

7. (a) Derive the expression of *principal axis* of an object in an image.
(b) Define mutual information. How do you compute mutual information between two images, which are co-registered. [5+(2+3)=10]

8. Consider the following block of gray levels:

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

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

9. Consider the block of gray levels in Question 8. Encode the above gray levels with strings of 0's and 1's based on Huffman coding. Calculate the average code-word length. [8+2=10]

10. 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:

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

[(4+4)+2=10]

11. (a) Describe the RGB and HSI models for color image.

(b) What is pseudocolor image?

[(2+6)+2=10]

12. Write a short note on skeleton by influence zone.

[10]

13. Write short notes on (i) image registration; and (ii) fast inverse wavelet transform.

[5+5=10]

INDIAN STATISTICAL INSTITUTE
End-Semestral Examination: 2016 – 17

Course Name: M. TECH CS - II

Subject: Computer Graphics

Date: 26.04.2017

Maximum Marks: 100

Duration: 3 hrs.

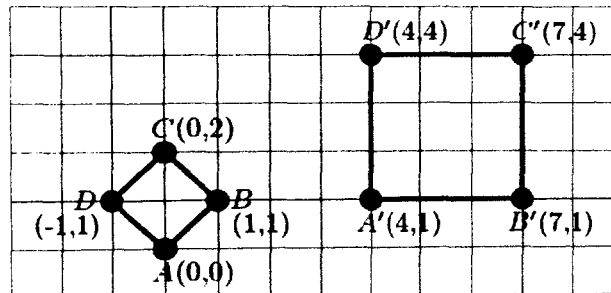
Answer as much as you can.

1. a) A cubic Bézier curve has the following mathematical form: $C(t) = \sum_{i=0}^3 P_i b_i(t)$

i) Show that $\sum_{i=0}^3 b_i(t) = 1$ for all t .

ii) Suppose \mathbf{R} is a rotation matrix. Show that $\mathbf{R}C(t)$ is given by a Bézier curve that has control points $\mathbf{R}P_i(t)$.

b) Consider the following figure:



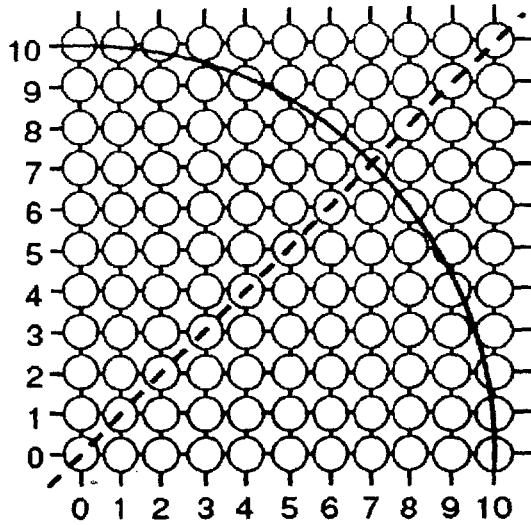
i) Give a matrix or product of matrices that will transform the figure ABCD into the figure A'B'C'D'.

ii) Show what happens if the same transformation is applied to figure A'B'C'D'.

(5+5)+(5+5) = 20

2. a) Give the parametric equation for a point P along a ray starting at S with direction d . Let t_a and t_b be two solutions to this equation corresponding to the intersections of the ray with a sphere. Which is closer to S ? Let t_i correspond to the intersection closer to S . Give expressions for (i) the radius of the sphere and (ii) the surface normal of the sphere at that intersection point in terms of S , d , t , and the center C of the sphere.

b) Demonstrate the mid-point circle algorithm by using the pixel grid shown below and filling out the table below. Assume that the symbols used in the table retain their usual meanings.



Pixel Grid

k	p_k	(x_{k+1}, y_{k+1})	$2x_{k+1}$	$2y_{k+1}$
0				
1				
2				
3				
4				
5				
6				

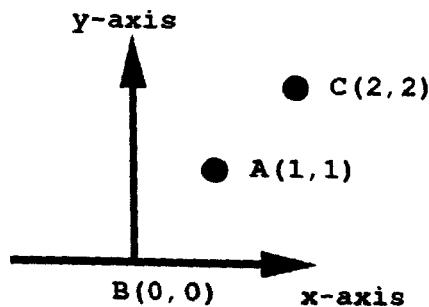
Table to fill in

$(3+4+5)+8= 20$

3. a) Can a 2D rotation (R) and a scaling (S) transformation be commutative i.e. $R.S = S.R$? If yes, under what conditions?
- b) The intensity (I) at a point due to diffuse (Lambertian) reflection and ambient light can be estimated using the following equation:

$$I = I_a k_a + f_A I_p k_d (\vec{N} \cdot \vec{L}).$$






Here the symbols bear their usual meanings. Explain the terms in the above equation. Given the constants $I_a = 2, k_a = 0.2, I_p = 10, k_d = 0.4$, and one light source at point A (1, 1), and f_A is the inverse of the distance of point B from the location of light source (See the figure below). Calculate the intensity at point B of the figure using the above information. What will be the intensity at point B using the above equation when the light source is moved from point A to C as shown in the Figure below.



- c) Suppose, at a particular point, the light source vector is given (in 3D Cartesian coordinates) by $L = (1, 0, 0)$ and the normal vector is given by $N = [0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}]$. Compute the reflection vector for a perfectly reflective surface.

$5+(3+5)+7= 20$

4. a) Consider the following 2×2 blocks corresponding to 5 intensity levels for halftoning.

- Level 0: Intensities 0.0 – 0.2 
- Level 1: Intensities 0.2 – 0.4 
- Level 2: Intensities 0.4 – 0.6 
- Level 3: Intensities 0.6 – 0.8 
- Level 4: Intensities 0.8 – 1.0 

Derive and sketch the resulting halftone image on the following example image (with normalized intensity values shown in each cell). Show your steps.

0.80	0.28	1.00	0.02	0.50	0.20
0.25	0.05	0.01	0.05	0.08	0.90
0.20	0.25	0.25	0.45	0.70	0.60
0.22	0.25	0.60	0.90	0.90	0.80
0.30	0.45	0.80	0.90	1.00	0.90
0.40	0.20	0.61	0.61	0.75	0.95

b) What is the fundamental difference between halftoning and dithering? Consider the following dither pattern and the same 5 intensity levels as was provided for part (a). Now derive and sketch the dithered image for the image matrix shown below. Explain your steps.

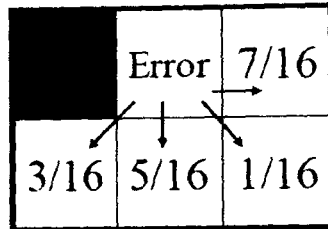
0.2	0.6
0.8	0.4

Dither Pattern

0.34	0.45	0.56	0.67
0.42	0.65	0.78	0.88
0.21	0.33	0.57	0.77
0.01	0.22	0.44	0.55

Image matrix with normalized intensity values.

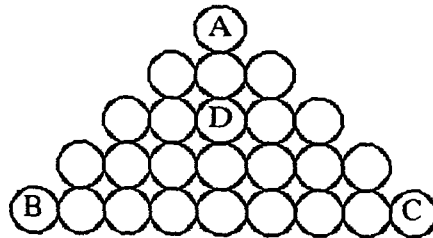
5. a) Assume we can display levels in increments of 50. Consider the following pattern for error diffusion:



Now illustrate how the second, third and fourth pixels in first row of the following image submatrix change due to error diffusion following the above pattern:

	68	76	109
18	34	56	78

- b) Assume a Gouraud shading model for the pixel-triangle below. Pixel A has been assigned the color (90,80,100), pixel B is color (70, 60, 60) and pixel C is (100, 80, 80). What color should be assigned to pixel D? Explain.



- c) Consider the colours X , Y , and Z of a hypothetical colour model are defined in terms of the R , G , and B primaries as:

$$X = 0.3R + 0.6G + 0.1B$$

$$Y = 0.6R - 0.3G - 0.3B$$

$$Z = 0.2R - 0.5G + 0.3B$$

- (ii) Write the equations that map from three values in XYZ colour space into RGB colour space.
 (iii) Can the colours X , Y , Z be considered as primaries? Justify your answer.
- d) Recall that an example of a one point perspective transformation (in which one principle axis (z) pierces the projection plane) is:

$$\mathbf{P}^{1pt} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{1}{d} & 0 \end{bmatrix}$$

Now briefly explain, with a neat figure, the functionality of the following projection matrix which uses a parameter q :

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \frac{\sin(q)}{d} & 0 & \frac{\cos(q)}{d} & 0 \end{bmatrix}$$

$$5+5+(3+3)+4= 20$$

6. Write short notes on any two of the following:

$$10 \times 2 = 20$$

- a) BSP Trees.
- b) Image Morphing.
- c) Z-buffer Algorithm.
- d) CMYK Color Model.

INDIAN STATISTICAL INSTITUTE
Semester Examination
M. Tech. (CS) II year (2nd Sem): 2016-2017
Advanced Cryptology

Date: 26. 04. 2017

Total Marks: 120

Time: 3.5 Hours

Answer as much as you can. Maximum you can score is 100.

1. (a) Consider the following two keyed functions where N is called **nonce**:

Function $\text{CTR1}_K(N, M)$	Function $\text{CTR2}_K(N, M)$
(1) $M \leftarrow M_1 \dots M_l$	(1) $M \leftarrow M_1 \dots M_l$
(2) for $j = 1$ to l	(2) for $j = 1$ to l
(3) $X_j \leftarrow E_K(N \oplus \langle j \rangle_{128})$	(3) $X_j \leftarrow E_K(N \langle i \rangle_{32})$
(4) $C_j \leftarrow X_j \oplus M_j$	(4) $C_j \leftarrow X_j \oplus M_j$
(5) return $C_1 \dots C_l$	(5) return $C_1 \dots C_l$

Consider block size of the block cipher E_K to be 128 bits and nonce size for CTR1_K is 128 bits and for CTR2_K is 96 bits. $\langle x \rangle_b$ denotes the b -bit binary representation of integer x .

- i. Find a PRF distinguishing attack for CTR1_K if \mathcal{A} is not allowed to query with the same nonce, i.e., \mathcal{A} cannot make two queries (N, M) and (N', M') such that $N = N'$.
 - ii. Find a PRF distinguishing attack for CTR2_K if \mathcal{A} is allowed to query with the same nonce.
- (b) Let Π be a uniform random permutation over $\{0, 1\}^n$. Consider the following keyed function from $\{0, 1\}^n \rightarrow \{0, 1\}^{2n} : f_\Pi(x) := (\Pi(x) \oplus \Pi(x \oplus 1), \Pi(x) \oplus \Pi(x \oplus 2))$. Find a PRF distinguisher for the above construction.
- (c) Consider the following encryption scheme: $\mathcal{E}_K(M_1, M_2) := E_K(M_1) || E_K(aM_1 + bM_2)$, where a and b are two non-zero field elements. Find a PRF distinguisher that is allowed to make *only a single query* for the above encryption scheme.

[(5 + 5) + 6 + 5=21]

2. (a) What is (uniform) random function and (uniform) random permutation? If Γ is a uniform random function from \mathcal{D} to \mathcal{R} , then show that $\forall x \notin \{x_1, \dots, x_q\} \subseteq \mathcal{D}$ and $\forall y_1, \dots, y_q \in \mathcal{R}$,

$$\Pr[\Gamma(x) = y \mid \Gamma(x_1) = y_1, \dots, \Gamma(x_q) = y_q] = |\mathcal{R}|^{-1}.$$

If Π is a uniform random permutation over \mathcal{D} then show that $\forall x \notin \{x_1, \dots, x_q\} \subseteq \mathcal{D}$ and $\forall y \notin \{y_1, \dots, y_q\} \subseteq \mathcal{D}$,

$$\Pr[\Pi(x) = y \mid \Pi(x_1) = y_1, \dots, \Pi(x_q) = y_q] = (|\mathcal{D}| - q)^{-1}.$$

- (b) Define the following:
- (i) (q, ϵ) -PRF (ii) (q, ϵ) -PRP (iii) (q, ϵ) -SPRP.
- (c) "If a keyed permutation F_K is SPRP secure then F_K is PRP secure"-Justify. Is the reverse direction of the above statement true? - Justify.

[4 + 4 + 3 + (3+6)=20]

3. (a) Prove that 3-round LR function is a permutation. Prove that 2-round LR function is not a secure PRP.

- (b) Consider the following keyed function:

$$F_f(L, M, R) := (f(L) \oplus M, R, L)$$

where L, M, R are all n bit strings, and $f : \{0, 1\}^n \rightarrow \{0, 1\}^n$.

F.T.O

- (i) Prove that F_f is a permutation from $3n$ bits to $3n$ bits.
- (ii) Show that $F^3 := F_{f_3} \circ F_{f_2} \circ F_{f_1}$ is not a secure PRP, where f_1, f_2 and f_3 are independent uniform random functions.
- (iii) Show that $F^4 := F_{f_4} \circ F_{f_3} \circ F_{f_2} \circ F_{f_1}$ is not a secure PRP, where f_1, f_2, f_3 and f_4 are independent uniform random functions.

[(3+2) + (2+3+5)=15]

4. (a) Prove that the number of queries q needed to distinguish the uniform random function from a uniform random permutation is $\Omega(2^{n/2})$.
- (b) Consider a uniform random function Γ from $\{0, 1\}^{128}$ to $\{0, 1\}^{128}$ and a uniform random permutation Π over $\{0, 1\}^{128}$. What is the maximum distinguishing advantage of these two if maximum 2^{42} queries are allowed?

[8+4=12]

5. (a) Consider the following keyed function

$$F_\pi(x) = \pi(x) \oplus x$$

where π is a permutation over $\{0, 1\}^n$ and x is a n -bit string.

- (i) Give a lower bound on the number of queries you need to distinguish this function from a uniform random function.
 - (ii) Give an upper bound of the PRF advantage for F_π .
- (b) Show that PRF advantage of the sum function based on two independent random functions i.e.

$$F_{f_1, f_2}(x) := f_1(x) \oplus f_2(x) \oplus c$$

is 0 where f_1 and f_2 are two independent uniform random functions, where c is a fixed non-zero constant.

- (c) Show that PRF advantage of the sum function based on two independent random permutations i.e.

$$F_{\pi_1, \pi_2}(x) := \pi_1(x) \oplus \pi_2(x)$$

is at most $q^2/2^{n+1}$ where π_1 and π_2 are two independent uniform random permutations.

[(8+8)+5+5=26]

6. (a) Show that there exist 2^n distributions X_1, \dots, X_{2^n} over $\{0, 1\}^n$ such that for every i , X_i and X_{i+1} are computationally indistinguishable, but X_1 and X_{2^n} are easy to distinguish.
- (b) If $f: \mathcal{X} \rightarrow \mathcal{Y}$ be a function and X_0, X_1 be two random variables over \mathcal{X} , then show that

$$\Delta(f(X_0), f(X_1)) \leq \Delta(X_0, X_1).$$

- (c) Let X and Y be two random variables defined over a set \mathcal{S} such that $\forall x \in \mathcal{S}_0 \subseteq \mathcal{S}, \Pr[X = x] \geq (1 - \epsilon_1) \Pr[Y = x]$ and $\Pr[Y \notin \mathcal{S}_0] \leq \epsilon_2$. Prove that

$$\Delta(X, Y) \leq \epsilon_1 + \epsilon_2.$$

- (d) What are the requirements for a ZK proof? Which requirement is satisfied with the concept of a *simulator* and how?

[6 + 6 + 6 + (3 + 5)=26]

INDIAN STATISTICAL INSTITUTE

End Semester Examination: (2016-2017)

M.Tech C.S., 2nd Year

Advanced Digital Signal Processing

Date: 28.4.2017

Maximum Marks: 100

Duration: 3 hours

Note: The marks add up to 114. The maximum you can score is 100. You are permitted to use calculators.

Questions:

1. The spectrum of a discrete time signal $x[n]$ must be estimated using the Discrete Fourier Transform (DFT) and a rectangular window. Let

$$x[n] = \cos(\pi n/16) + \cos(3\pi n/64) + \cos(2\pi n/8)$$

- (a) Determine the number of peaks that can be resolved using a window length of 64 and DFT of the same length.
- (b) Determine the minimum length that the window must have in order to resolve all 3 peaks. Assume that the DFT has the same length.

[8+5]

2. Let

$$x[n] = \begin{cases} \cos(2\pi(1/16)n) & \text{if } 0 \leq n \leq 15 \\ \cos(2\pi(5/16)n) & \text{if } 16 \leq n \leq 31 \\ 0 & \text{otherwise} \end{cases}$$

Sketch the sampled Short-time Fourier Transform $X[rR, k]$ obtained using a rectangular window of size 16, a DFT length of 16, $R = 16$, for $r = 0, 1$ (~~2 separate sketches~~). [10]

3. A bandlimited continuous-time signal with $|\Omega| \leq 2\pi(10^3)$ rad/s, is sampled at the Nyquist rate for 20 s. For the sequence obtained,
- (a) determine the minimum DFT length (implemented using a radix-2 FFT algorithm) for estimating the spectrum using the periodogram approach such that the uniformly spaced frequency locations are no more than 2 Hz apart.

- (b) determine the maximum possible reduction in variance that can be obtained by the method of averaging periodograms (no overlap), using the DFT length determined in (a) above.

[8+5]

4. An ARMA process has a power spectral density given by:

$$(4/9) \frac{5 - 2z - 2z^{-1}}{10 - 3z^{-1} - 3z}$$

- (a) Determine a stable and causal filter for generating this process from a white noise process.
- (b) Give the difference equation satisfied by the corresponding whitening filter.

[8+4]

5. An AR(1) process $s[n]$ has an autocorrelation sequence $\gamma_{ss}(m) = 0.6^{|m|}$ is generated by passing white noise $v[n]$ through an LTI system. Using the Yule-Walker equations, determine the filter coefficient and the variance of $v[n]$. Now, consider a signal

$$x(n) = s(n) + w(n)$$

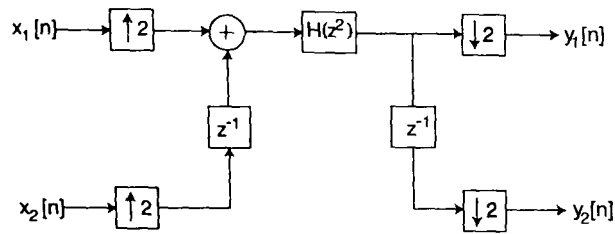
where $w(n)$ is a white noise process with variance 1. Assume that $\{v[n]\}$ and $\{w[n]\}$ are uncorrelated. Design a length-2 Wiener filter to estimate $s(n)$ from $x(n)$.

[8+10]

6. Consider the transfer functions $H_0(z) = 1 + az^{-1} + z^{-2}$ and $H_1(z) = 1 + az^{-1} + bz^{-2} + az^{-3} + z^{-4}$

- (a) Determine if these are linear phase and half-band.
- (b) Design an alias-free, perfect reconstruction filter bank using these filters as the analysis filters and show that system performs as desired. State explicitly the conditions that must be satisfied by a and b .

[(5+5)+(10+4)]



7. For the structure shown below,

- (a) Determine $Y_1(z)/X_1(z)$ and $Y_2(z)/X_2(z)$
- (b) Determine $y_1[n]$ and $y_2[n]$ if $H(z)$ is an all-pass filter and

$$x_1[n] = x_2[n] = \begin{cases} n & \text{if } 0 \leq n \leq 4 \\ 0 & \text{otherwise} \end{cases}$$

[8+8]

8. Let a scale space be denoted by V_j and the wavelet space by W_j and assume that they satisfy the requirements for multiresolution (a higher value of j indicates higher resolution). Consider two functions $f(t) \in V_0$ and $g(t) \in V_1$. Determine if $g(t) - f(t) \in W_1$ and justify your answer. [8]

INDIAN STATISTICAL INSTITUTE

Semestral Examination: 2016 – 17

MTech CS (2nd Year)

Computational Finance

Date: 28 April 2017

Maximum Marks: 100

Duration: 3 Hours

1. Let

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

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

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

has a solution if and only if there exists an arbitrage opportunity in the securities market with N securities S_i ($i = 1, \dots, N$) and K states of nature ω_j ($j = 1, \dots, K$). S_i^* 's are discounted (by the bank process) values. [20]

2. Prove the Put – Call parity of European option for the multi-period market. Is the same relation true for American options? – Prove or refute logically. [6 + 6 = 12]

3. Define the following option contracts:

- (i) Lookback
- (ii) Barrier
- (iii) Chooser

For each of them, state the payoff function carefully, explaining all notation. [3 X 6 = 18]

4. Prove directly from the definition of Ito integrals that

$$\int_0^t s dW_s = tW_t - \int_0^t W_s ds$$

[15]

5. (a) Use Ito's formula to write the following stochastic processes X_t in the standard form

$$dX_t = u(t, \omega)dt + v(t, \omega)dW_t$$

for suitable choices of $u \in R, v \in R$

a) $X_t = W_t^2$

b) $X_t = 2 + t + e^{W_t}$

(b) Check whether $X_t = t^2 W_t - 2 \int_0^t s W_s ds$ is a Martingale.

[(3 + 4) + 8 = 15]

6. Clearly explain the *Greedy* algorithm for multi-venue order allocation in the context of dark pool trading. Demonstrate its complexity and worst case scenario computation time?

[10 + 15 = 20]

INDIAN STATISTICAL INSTITUTE

Supplementary Examination: (2016-2017)
M.Tech C.S., 2nd Year

Advanced Digital Signal Processing

Date: 3.5.2017

Maximum Marks: 60

Duration: 2 hours

Note: The marks add up to 70. The maximum you can score is 60.

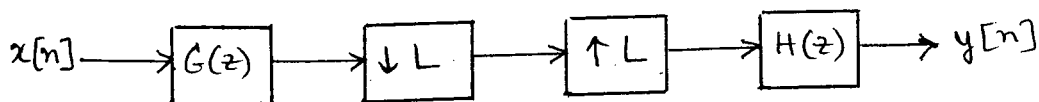
Questions:

1. Determine (with justification) if the down-sampler is linear and time-invariant. [5+5]
2. Design a two-channel alias-free perfect reconstruction quadrature-mirror filter bank with $H_0(z) = z^{-1} + 0.5z^{-2}$. Verify your design. [15]
3. Show that

$$\sum_{k=0}^{M-1} H(zW^k) = ME_o(z^M) \quad \text{where} \quad H(z) = \sum_{k=0}^{M-1} z^{-k} E_k(z^M)$$

[10]

4. 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. [15+5]
5. Show that the structure given below is an LTI system. Also obtain the condition which the product $H(z)G(z)$ should satisfy if the system has to be an identity system. [8+7]



INDIAN STATISTICAL INSTITUTE
Mid-Semestra. Examination: 2018-17

Course Name: M. TECH CS - II

Subject: Computer Graphics

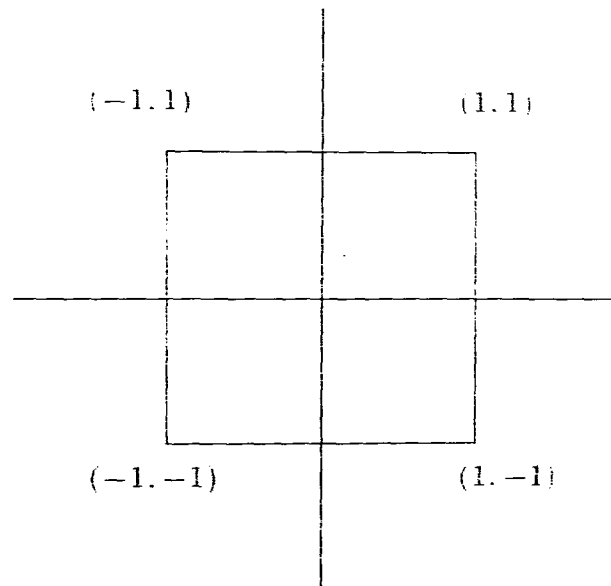
Date: 05/04/2017

Maximum Marks: 60

Duration: 2 hrs.

Answer as much as you can.

1. a) Here is a test pattern in the plane:



Draw its image after being transformed by each of the following (individually, not cumulatively) and give brief explanations for each transform.

(i)
$$\begin{bmatrix} \cos 60^\circ & \sin 60^\circ & 0 \\ -\sin 60^\circ & \cos 60^\circ & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

(ii)
$$\begin{bmatrix} 1 & 0 & 1.5 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

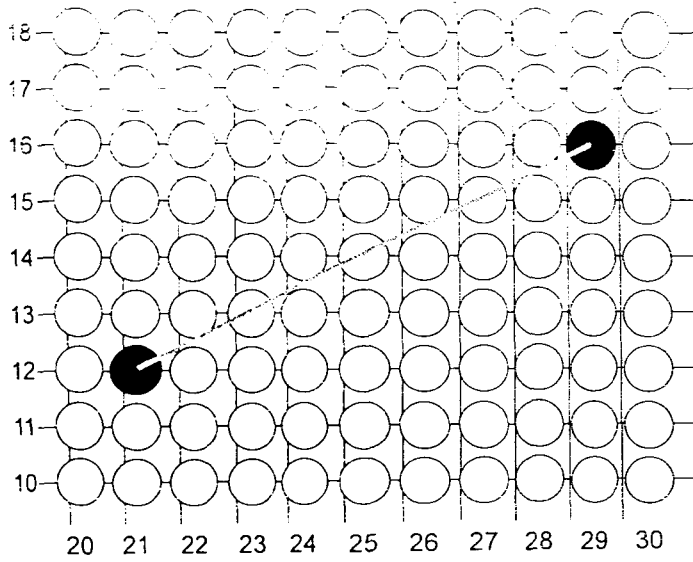
(5+5 = 10)

(b) Determine the form of a transformation matrix for a reflection about an arbitrary line with equation $y = mx + b$. (5)

2. a) Consider the line which starts at (4, 13) and ends at (22, 3).

i) How many pixels will there be in this line?

- n) With the DDA (Digital Differential Analyzer) algorithm, what will be the amount added to the secondary component each time through the loop (incrementing value)? Explain. (3+4 = 7)
- n) Derive the window to viewport mappings for $W = (0, 100, 50, 200)$ and $V = (20, 70, 10, 80)$. (5)
- c) Suppose an RGB raster system is to be designed using a 10 inch by 12 inch screen with a resolution of 150 pixels per inch in each direction. If we want to store 8 bits per pixel in the frame buffer, how much storage (in bytes) do we need for system the frame buffer? (3)
3. a) Consider a raster system with the resolution of 1024×768 pixels and the color palette calls for 65,536 colors. What is the minimum amount of video RAM that the computer must have to support the above-mentioned resolution and number of colors? (5)
- b) You want to join two Bézier curves. The first curve must start at $(0,0,0)$ with a tangent vector of $(1,1,1)$. The second curve must end at $(5,0,-5)$ with a tangent vector of $(-1,-1,-1)$. The two curves should be joined at $(2,5,0)$ and have C^1 continuity. Give the geometry matrices for each curve. (10)
4. a) Given an eyepoint at $(0,0,0)$, a pixel at $(-3, 2, -4)$, and a sphere of radius 5 with its center at $(-5, 10, -10)$, does the ray from the eyepoint through the pixel intersect the sphere? If so, what is the intersection point? (6)
- b) Define the frustum of vision to be centered on the z-axis with the eyepoint at $(0, 0, -d)$ and with the window on the x: y-plane having its center at $(w_{cx}, w_{cy}, 0)$, width $2 \cdot w_{sx}$, and height $2 \cdot w_{sy}$. The far clipping plane is to be parallel to the window and intersect the z-axis at $(0, 0, f)$.
- (i) Give the equations (in the window coordinate system) of the six clipping planes. Adjust the signs in your equations (if necessary) so that a positive value denotes "visible" and a negative value denotes "invisible".
- (ii) What is the perspective matrix appropriate for this formulation of the problem. (6+3 = 9)
5. Illustrate the steps of the Bresenham line drawing algorithm for a line going from $(21,12)$ to $(29,16)$. Refer to the following pixel grid (see the following figure) and fill out the table shown below. Show all your calculations. Here p_k refers to the decision parameter at the k -th step. (15)



k	p_k	(x_k, y_k)
0		
1		
2		
3		
4		
5		
6		
7		
8		

6. Write short notes on any two:

$$7\frac{1}{2} \times 2 = 15$$

- a) A-buffer Algorithm.
- b) Backface Culling
- c) Sutherland-Hodgman Clipping
- d) Constructive Solid Geometry Methods