INDIAN STATISTICAL INSTITUTE
Semester Examination
M. Tech. (CS) II year (2nd Sem): 2015–2016
Advanced Cryptology

Date: 21. 04. 2016 Maximum Marks : 80 Time : 3 Hours

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

1. (a) Define a hard-core predicate for a function.
   (b) Can a function have a hard-core predicate, if the function is not one-way? Justify.
   (c) Does there exist a predicate that is hard-core for all one-way functions? Justify.

   $[4 + (2 + 4) + (2 + 4) = 16]$

2. (a) Show that the existence of one of OWF or PRG implies the existence of the other.
   (b) State Goldreich-Levin theorem. Prove the theorem for the special case when the h oracle for a fixed $x \in \{0,1\}^n$, on input $r \in \{0,1\}^n$, gives an output that matches with $\langle x, r \rangle$ with probability $1 - negl(n)$.

   $[(4 + 4) + (2 + 6) = 16]$

3. Show that the existence of PRG implies the existence of secure bit commitment scheme (describe the construction and provide a brief proof). $[8 + 8 = 16]$

4. (a) How can you use 1/2-OT to construct a $\binom{2}{1}$-OT?
   (b) What is the IND-ID-CPA game? Is the basic IBE scheme discussed in class both IND-ID-CPA and CCA secure?

   $[8 + (4 + 4) = 16]$

5. (a) How is verifiable secret sharing (VSS) different from ordinary secret sharing?
   (b) Describe Feldman's VSS scheme.
   (c) What are its drawbacks?

   $[4 + 8 + 4 = 16]$

6. (a) How to handle format preserving encryption (FPE) when the encrypted data is outside the domain?
   (b) In cycle-walking, derive the expected length of a cycle.
   (c) Why are variants of Feistel commonly exploited for FPE?

   $[4 + 4 + 2 = 10]$

INDIAN STATISTICAL INSTITUTE

M. Tech. (CS) II Year (2015-16), II semester

Semestral Examination

ADVANCED PATTERN RECOGNITION

Date: 22.04.2016 Duration: 210 minutes Maximum marks: 100

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

1. Suppose you are given a training set for a two class classification problem. It is also known that the classes are linearly separable. Describe perceptron algorithm for finding a separating hyperplane between the two classes. [12]

2. Define a dissimilarity measure between two features. Describe a feature selection method based on the suggested dissimilarity. [4+4=8]

3. Suppose you have two m-dimensional normal populations $N(\mu_1, \Sigma)$ and $N(\mu_2, \Sigma)$. Let the prior probability of the first population be P, where $0<P<1$. Find the Bayes decision rule for separating the two populations and also find its probability of misclassification. [3+7=10]

4. Describe the basic steps of the elitist model of genetic algorithm, and show that it provides an optimal solution as the number of iterations goes to infinity. [15+10=25]

5. (a) Define fuzzy c-partition of a dataset.
   (b) Write down the objective function for fuzzy c means (FCM) and the necessary conditions used. Describe the fuzzy c means algorithm. [5+9=14]

6. Suppose there is one hidden layer with J number of nodes in an MLP, and sigmoid function is used as transfer function. Suppose you are using online learning algorithm. Let a training dataset be given to you and let the number of classes be 3. Then
   (a) Write down the expression for the error for the MLP.
   (b) Write down the expression for the change in the connection weight joining the i-th node in the hidden layer to the second node in the output layer. [3+7=10]

7. Write short notes on the following.
   (a) Kernel functions
   (b) VC dimension
   (c) Support vectors [5+5+5=15]

(P.T.O)

8. (a) Describe k-nearest neighbor probability density estimation procedure.
(b) Derive the k-nearest neighbor decision rule from the density estimation procedure. [6+7=13]

INDIAN STATISTICAL INSTITUTE

End Semester Examination: (2015-2016)
M.Tech C.S., 2nd Year

Advanced Digital Signal Processing

Date: 22.4.2016 Maximum Marks: 100 Duration: 3 hours

Note: The marks add up to 114. The maximum you can score is 100. The exam is open-book, open-notes. You are permitted to use calculators.

Questions:

1. Let

$$x[n] = \begin{cases} \cos(2\pi 0.1n) & \text{if } 0 \leq n < 30 \\ \cos(2\pi 0.4n) & \text{if } 30 \leq n < 60 \end{cases}$$

Sketch the Short-time Fourier Transform obtained using

(a) A rectangular window of size 10, a 10 pt DFT and no overlap.

(b) A rectangular window of size 10, a 10 pt DFT and 50% overlap.

You may give a series of 2-D sketches – 6 for the first case and 11 for the second. [8+8]

2. An AR(1) process $s[n]$ satisfies the difference equation

$$s[n] = 0.8s[n-1] + v[n]$$

where $\{v[n]\}$ is a white noise sequence with variance $\sigma_v^2 = 0.36$. Let

$$x[n] = s[n] + w[n]$$

where $\{w[n]\}$ is a white noise sequence with variance $\sigma_w^2 = 0.81$. Assume that $\{v[n]\}$ and $\{w[n]\}$ are uncorrelated.

(a) Design a Wiener filter of length 2 to estimate $\{s[n]\}$.

(b) Determine the minimum MSE achieved.

(c) Indicate, with justification how the filter coefficients would change if $\sigma_w^2 \approx 0$.

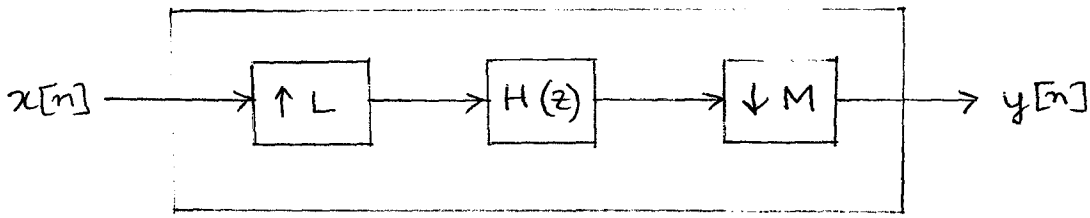
[8+5+3]

3. A zero-mean, discrete-time signal $x[n]$ of length N has the autocorrelation sequence $\mathcal{E}\{x[n]x[m]\} = \sigma_x^2\delta[n-m]$ where $\delta[n]$ is the unit impulse.

   Let the N-point DFT of $x[n]$ be $X[k]$, $k = 0, \cdots, N-1$. Determine $\mathcal{E}\{X[p]X^*[q]\}$ in terms of p and q.

   [8]

4. In the structure given below, $H(z)$ represents an LTI system. Under what condition is it possible to replace the system enclosed by the rectangle by an equivalent LTI system? Justify your answer. Determine the transfer function of this equivalent LTI system. [5+5]



5. An AR(1) process is defined by the difference equation

$$x(n) = 0.6x(n-1) + w(n)$$

   where $w(n)$ is a white noise process with variance 0.64. Determine

   (a) The one-step predictor coefficient.

   (b) The power spectral density of $x[n]$.

   (c) The autocorrelation sequence values for lags 0 and 1, using the Yule Walker equations.

   [2+5+8]

6. Design (giving justification) a two-channel alias-free perfect reconstruction quadrature-mirror filter bank with $H_0(z) = z^{-1} + 0.5z^{-2}$. [10]

7. An ARMA process has power spectral density

$$\Gamma_{xx}(z) = 4\frac{(1-2z^{-1})(1-1/2z^{-1})}{(1-3z^{-1})(1-1/3z^{-1})} \quad 1/3 < |z| < 3$$

(a) Determine the system function of a stable filter for generating $x[n]$ from white noise. Is it causal as well?

(b) Give the lattice ladder representation of the above filter.

[(5+3)+8]

8. (a) Show that

$$\sum_{m=-N}^{N} r_{xx}(m)e^{-j2\pi fm} = \frac{1}{N}|\sum_{n=0}^{N-1} x[n]e^{-j2\pi fn}|^2$$

where $r_{xx}(m) = \frac{1}{N}\sum_{n=0}^{N-1} x(n+m)x^*(n)$

(b) Determine the periodogram $I[k]$ for a sequence $x[n] = (0.7)^n u[n]$ using a rectangular window of size 10.

[10+5]

9. Consider the FIR transfer function

$$H(z) = -3 + 19z^{-2} + 32z^{-3} + 19z^{-4} - 3z^{-6}$$

Determine (with justification) if it is

(a) A linear phase filter,

(b) A half-band filter.

[3+5]

Indian Statistical Institute
Semester Examination (2016)
M.Tech (CS) II Year
Computer Vision

Date: 26.04.16 **Full Marks: 100** **Duration – 3 hours 15 mins**

Answer as many questions as you like, but you may at most score 100.

1. Consider a pinhole camera. Its focal length is 20 mm. A scene point is located at the world co-ordinates (X, Y, Z) = (40, 30, 20), where the unit is in meter. Now compute the following:
   (a) The two-dimensional image co-ordinates
   (b) The Field of View (FoV) if the image plane is 40 mm x 40 mm.
   (c) The distance where the camera needs to be placed from a building 200 m high so that the FoV is completely utilized. 6+4+5=15

2. Derive the following, each with respect to transformation from an initial point having Cartesian world co-ordinates (X, Y, Z):
   (a) The generalized translation matrix and its inverse
   (b) The generalized scaling matrix and its inverse
   (c) The rotation matrix and its inverse, around any one of the X-, Y- or Z-axes.

   Hence answer the following questions:

   (i) Explain the necessity for introduction of a Homogenous Co-ordinate System both for world and camera co-ordinates, in order to derive the perspective transformation matrix or its inverse (for simplicity, assume an ideal situation where the world and camera co-ordinates coincide).

   (ii) Next, consider a more realistic situation, where instead of the idealized one considered for (i), so that there may be two translation vectors of the camera, one for the camera gimbal from the world co-ordinate origin, and one for imaging plane with respect to the gimbal center (the camera may undergo either tilt or pan movement):
      I. What is the camera model for such a realistic situation?
      II. How is this realistic camera model different from the pinhole camera model?
      III. How do you calibrate this camera?

   (5+6+7)+6+(6+3+7)=40

.....page 2

3. Explain how depth is perceived through stereo imaging with the help of the lateral stereo model. Hence explain the concept of disparity. 12+3=15

4. (a) What are the basic assumptions of the retinex theory?
   (b) Explain each of the concepts of irradiance, radiance and reflectance in the context of the retinex theory.
   (c) Why and how was the Laplacian operator utilized in the retinex theory?
   (c) Derive the relation between scene radiance and image irradiance.
   (d) Explain the concept of Bi-directional Reflectance Distribution Function.
   4+(3x3)+6+10+6=35

5. (a) Explain the concept of scale-space in computer vision.
   (b) Why did the SIFT algorithm use the Difference of Gaussian function?
   (c) Explain the concept of raw primal sketch in the context of visual receptive field. 5+5+5=15

Indian Statistical Institute

M.Tech (CS) II

Information Security and Assurance

Semester Examination

Maximum Marks: 80

Date: April 26, 2016
Time 3 hours

The question paper contains 6 questions. Total marks is 90. Maximum you can score is 80.

1. (a) What is Proof-of-work?
   (b) Why is it implemented in Bitcoin protocol?
   (c) How is it implemented in Bitcoin protocol?

   $(2 + 4 + 6 = 12)$

2. How are transaction graphs represented in Bitcoin? (10)

3. What anonymization techniques are used in Bitcoin protocol? (8)

4. Consider a sensor network with many sensors which sense data. Further, consider a few cluster heads to aggregate data. The aggregated data is then send to a base station.
   (a) Design an efficient signature scheme to authenticate the data sensed by multiple sensors. Calculate the communication and computation overheads. One should attempt to minimize communication and computation costs.
   (b) What encryption technique will the sensors use, such that the cluster head can still aggregate the encrypted data and then send to the base station for further application. Describe the encryption scheme.

   $((7+3) + 10 = 20)$

5. Under what assumption is the Boneh-Gentry-Waters broadcast encryption scheme secure? Prove it by defining an appropriate security model. (15)

6. (a) What are the challenges in storing data in untrusted servers, for example clouds?
   (b) Describe a general construction for data auditing in an outsourced server. Can you use Merkle trees? If yes, how? Can you suggest a more efficient alternative?

   $(5 + 10 + 10 = 25)$

INDIAN STATISTICAL INSTITUTE

Semestral Examination of Second Semester (2015-2016)

M.TECH.(CS) II YEAR

Topics in Algebraic Computation

Date: 27.04.2014 Maximum marks: 100 Duration: 3 hours

Note: The paper contains 120 marks. Answer as much as you can. The maximum you can score is 100.

1. In the following $\mathbb{F}$ is an algebraically closed field.
   (a) Is $X = \{(x, x) : x \neq 1\}$ an affine variety?
   (b) Using graded LEX ordering of monomials, divide $x^3 - x^2y - x^2z + x$ by $f_1 = x^2y - z$ and $f_2 = xy - 1$.
   (c) Show that for any variety V, $I(V)$ is a radical ideal.
   (d) If I is an ideal of $\mathbb{F}[x_1, \ldots, x_n]$, show that $V(I)$ is an affine variety.
   (e) If $I \neq \{0\}$ is an ideal and $\langle \mathrm{LT}(I) \rangle = \langle \mathrm{LT}(g_1), \ldots, \mathrm{LT}(g_t) \rangle$, then show that $I = \langle g_1, \ldots, g_t \rangle$.
   (f) Show that I has a reduced Groebner basis.

   $(6 \times 5 = 30)$

2. (a) Describe Buchberger's algorithm to compute a Groebner basis for an ideal and show that the algorithm terminates.
   (b) Briefly describe the main ideas behind solving a system of multivariate polynomial equations.

   $(10 + 10 = 20)$

3. (a) Let $\phi(x)$ and $g(x)$ be two polynomials over the finite field $\mathbb{F}_q$, where q is a prime power. It is required to generate all possible a, b and c which defines bi-variate polynomials $T(x, y) = xy + ax + by + c$ such that $\phi(x)$ divides $T(x, g(x))$. Describe a method for achieving this. (Hint: Consider $x^i(g(x))^j \bmod \phi(x)$ for $0 \leq i, j \leq 1$ and reformulate the problem in terms of matrices.)
   (b) Factor $u(x) = x^6 + 3x^5 - 4x^4 + x^2 - 7$ modulo 2 and lift the factorisation to modulo 2^2.
   (c) Describe the basic idea of Hensel lifting in the context of factorisation of polynomials over the integers.

   $(10 + 10 + 10 = 30)$

4. (a) For a sequence $s_0, s_1, \ldots$ generated by an LFSR with connection polynomial $C(D)$ of degree L, define $S(D) = s_0 + s_1D + s_2D^2 + \cdots$. Show that $C(D)S(D)$ is a polynomial of degree less than L.
   (b) If some LFSR of length L generates the sequence $s_0, s_1, \ldots, s_{N-1}$ but not the sequence $s_0, s_1, \ldots, s_{N-1}, s_N$, then show that any LFSR that generates the latter sequence has length L' satisfying $L' \geq N + 1 - L$.

(c) Describe a simple list decoding algorithm for Reed-Solomon code.

(d) Derive the key equation for decoding of Goppa codes and show how solving it will lead to the actual decoding.

(e) Define Hamming code and describe the decoding algorithm for this code.

$(8 \times 5 = 40)$

INDIAN STATISTICAL INSTITUTE

Semestral Examination: 2015 – 16

M.Tech. CS (2nd Year)

Algorithmic Game theory

Date: 29 April 2016 Maximum Marks: 100 Duration: 3 Hours

This paper carries 120 marks. You can attempt any part of any question.

The maximum you can score is 100.

Q1:

a) Define the notion of PLS-completeness.

b) Prove that the problem of computing a Pure Nash Equilibrium of a symmetric congestion game is PLS-complete. [4 + 16 = 20]

Q2:

a) Consider an atomic selfish routing game in which all players have the same source vertex and sink vertex (and each controls one unit of flow). Assume that edge cost functions are non-decreasing, but *not affine*. Prove that a (pure-strategy) Nash equilibrium (i.e., an equilibrium flow) can be computed in polynomial time.

b) Prove that in an atomic selfish routing network of parallel links, every equilibrium flow minimizes the potential function.

c) Show by giving example that b) does not hold in general networks, even when all players have a common source and sink vertex. [7 + 7 + 6 = 20]

Q3: Consider a multi-commodity network $G = (V, E)$, where for $i = 1, 2,..,k$, $r_i > 0$ units of traffic travel from an origin $s_i \in V$ to a destination $t_i \in V$.

a) Define the notion of a flow and an equilibrium flow for multi-commodity networks.

b) Prove that every multi-commodity network with any arbitrary cost function in a set C has Price of Anarchy at most the Pigou bound.

c) Prove that if C is the set of cost functions of the form $c(x) = ax + b$ with $a, b > 0$, then the Pigou bound is 4/3. [6 + 7 + 7 = 20]

Q4:

a) It is known that finding a Nash equilibrium in a two-person zero-sum game is significantly easier than that in general two-person games. Now consider a three-person zero-sum game, that is, a game in which the rewards of the three players always sums to zero. Show that finding a Nash equilibrium in such games is at least as hard as that in general two-person games.

b) Show that in an equilibrium of a non-degenerate game, all pure best responses are played with positive probability. [13 + 7 = 20]

Q5: Consider a setting with n agents and m goods where each agent is endowed with a bundle of goods and a linear utility function that specifies the utility that this agent derives from consuming a bundle. The value of a non-empty coalition S of agents can be defined as the maximum total utility that agents in S can derive by optimally redistributing their endowments. Model this setting as a TU game. Does this game always have a nonempty core? [14 + 6 = 20]

Q6:

a) Consider a single-item auction with at least three bidders. Prove that awarding the item to the highest bidder, at a price equal to the third-highest bid, yields an auction that is not dominant-strategy incentive compatible (DSIC).

b) Consider an auction with k identical goods, with at most one given to each bidder. There are n bidders whose valuations are i.i.d. drawn from a regular distribution F. Describe the optimal auction in this case. Which of the following does the reserve price depend on: k, n, and/or F?

[10 + 10 = 20]

Indian Statistical Institute

Advanced Image Processing

M.Tech.(CS)-II Year, 2015-2016

Full marks: 100 Time: 3 Hours

Date: 29.04.2016

Answer Question 1 and any **six** from rest of the questions.

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

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

2. (a) Define morphological dilation and erosion operations.

   (b) Prove that dilation and erosion are dual operations.

   (c) State and prove the idempotent property of morphological open operation.

   [2+6+7=15]

3. (a) What is the difference between tree wavelet expansion and packet wavelet analysis of an image?

   (b) Consider the following digital signal: [1, 2, 4, 1, -1, -2, -1, 1]. Construct the tree wavelet expansion of this signal using the following wavelet filter: [0.1294, 0.2241, -0.8365, 0.4830]. [3+12=15]

4. (a) How is intensity of an image block represented using Block Truncation Coding (BTC)?

   (b) Consider the following block of gray levels:

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

   Calculate the compressed and reconstructed representation of the block using BTC. Calculate PSNR and bpp. [5+(8+2)=15]

5. Derive the expression of parametric Wiener filter for image restoration using constrained least square estimation approach. [15]

6. Consider the following block of gray levels:

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

Construct the gray level co-occurrence matrices for angle $\theta = 0°, 45°, 90°, 135°$, considering unit pixel distance, and compute the angular second moment for each case. [10+5=15]

7. (a) Consider the block of gray levels in Question 6. Encode the above gray levels with strings of 0's and 1's based on Huffman coding. Calculate the average code-word length.

(b) Define mutual information. How do you compute mutual information between two images, which are co-registered. [(8+2)+(2+3)=15]

8. 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:

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

[(6+6)+3=15]

9. Write short notes on (i) Digitization; (ii) image registration; and (iii) Thinning. [5+5+5=15]

M. TECH. (CS) - II YEAR

DATA MINING
END SEM EXAMINATION

Date : 08/12/2015

Time: 3 Hours

Full marks: 50

1. We know that if $\phi : \mathbb{R}^n \to \mathbb{R}^m$ maps n dimensional data points to m dimensionial data points, and $K : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is a function such that $K(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i) \bullet \phi(\mathbf{x}_j) \; \forall \mathbf{x}_i, \mathbf{x}_j \in \mathbb{R}^n$, then K works as a valid SVM kernel. By definition, we say that K is a *valid kernel* if for any arbitrary finite subset $\{\mathbf{x}_1, \cdots, \mathbf{x}_k\} \subset \mathbb{R}^n$, the $k \times k$ matrix $K(i,j) = K(\mathbf{x}_i, \mathbf{x}_j)$ is positive semidefinite (Recall: A symmetric $k \times k$ real valued matrix A is called positive semidefinite if for all $\mathbf{v} \in \mathbb{R}^k, \mathbf{v}^T A \mathbf{v} \geq 0$).

   (a) Suppose K_1 and K_2 are two *valid* SVM kernels. Show that $K_1 K_2$, defined as $(\mathbf{x}, \mathbf{y}) \mapsto K_1(\mathbf{x}, \mathbf{y}) K_2(\mathbf{x}, \mathbf{y})$, is also a *valid* SVM kernel.

   (b) Is $K(\mathbf{x}, \mathbf{y}) = 1 - ||\mathbf{x}|| \cdot ||\mathbf{y}||$ a valid SVM kernel? Prove or disprove.

(5+ 5=10 Marks)

2. From an undirected graph $G(V, E)$, a graph G' can be obtained in the following way.

   - If XY is an edge of G, then XY is considered to be a node of G'. Note that XY and YX represent the same node of G'.
   - If XY and YZ are edges in G, then there is an edge between the nodes XY and YZ in G'. In other words, *nodes* in G' have an edge between them if the corresponding *edges* in G have a *node* of G common in them.

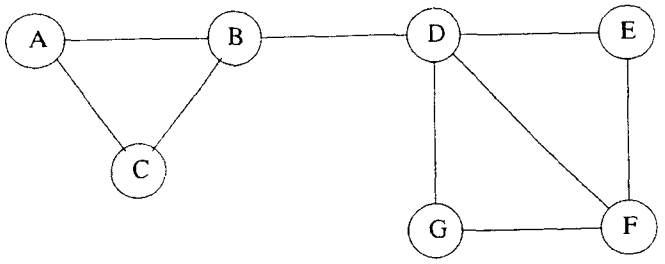


Figure 1: A tiny social network graph

   (a) Apply the above construction to the graph G shown in Figure 1 and draw the resulting graph. If we apply the above construction to a network graph G of friends (say Facebook), what is the interpretation of the edges of the resulting graph G'?

   (b) How is the degree of a node XY in G' related to the degrees of the nodes X and Y in G?

(5 + 5 = 10 Marks)

3. Suppose the Web consists of a clique (set of nodes with all nodes have links to all other nodes) of n nodes and a single additional node which every node of the clique links to, as illustrated by an example in Figure 2 with $n = 4$.

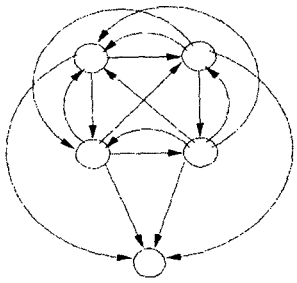


Figure 2: Example of the Web with a clique with 4 nodes and a single additional node

Determine the PageRank (computed with teleporting, with teleporting probability β) of each page, as a function of n and β.

(5 Marks)

4. Suppose A is a 5×6 term - document matrix, representing a dataset with 5 terms and 6 documents. Assume that the term weights of A are simply the term frequencies. That is, if $A(i, j) = 4$ then the i-th term occurs 4 times in the j-th document. The singular value decomposition $A = U\Sigma V^T$ of A, is given as below:

$$\begin{pmatrix} -0.2 & -0.5 & 0.2 & 0 & -0.8 \\ -0.2 & -0.7 & 0.4 & 0 & 0.6 \\ -0.6 & 0.1 & -0.2 & -0.7 & 0 \\ -0.7 & 0.2 & -0.1 & 0.7 & 0 \\ -0.1 & 0.5 & 0.9 & -0.1 & 0 \end{pmatrix} \begin{pmatrix} 5.2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3.7 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3.6 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1.1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.2 & 0 \end{pmatrix} \begin{pmatrix} -0.4 & -0.1 & -0.7 & -0.6 & -0.2 & -0.2 \\ 0.1 & 0.4 & 0.2 & -0.1 & 0.3 & -0.8 \\ -0.2 & 0.7 & -0.2 & 0 & 0.5 & 0.4 \\ -0.7 & -0.3 & 0.4 & -0.1 & 0.4 & 0.1 \\ 0.4 & 0 & 0.4 & -0.8 & 0 & 0.3 \\ 0.4 & -0.3 & 0.4 & 0.3 & -0.6 & -0.6 \end{pmatrix}$$

Suppose the 5 terms are $t_1, \cdots, t_5$, with $t_1 = obama, t_3 = modi$, and $t_5 = kolkata$. It is known that one among t_2 and t_4 is $india$, the other is usa.

(a) If the documents are numbered $d_1, \cdots, d_6$ in the order represented by the matrices U, Σ and V^T, which (one, more than one, or none) documents are likely to be about Obama's visit to India while Modi is the prime minister?

(b) Estimate the frequency of the term $kolkata$ in d_2, or explain why it cannot be done.

(5 + 5 = 10 Marks)

5. A simple estimate of the communication cost of a mapreduce algorithm is the the summation of the size of the input to all the tasks. Using this principle, compute and compare the communication costs of the matrix - matrix multiplication methods, namely (as done in class)

(a) Method 1: a natural join followed by a grouping with aggregation

(b) Method 2: a single mapreduce process

(5 + 5 = 10 Marks)

6. Recall the Flajolet - Martin algorithm for counting distinct elements in a stream. For a hash function h and any element a in the stream, the tail length of h and a is defined as the number of zeros at the end of $h(a)$. If R is defined to be the maximum tail length seen in a stream at any point of time, then 2^R is used as an estimate for the number of distinct elements in the stream.

   Since the estimate using one hash function would not be very accurate, several hash functions $h_1, \cdots, h_k$ are used and the estimates 2^{R_i} corresponding to each h_i, for $i = 1, \cdots, k$ are combined using mean or median or both to get a more accurate estimate.

   Analyze and explain what would be the shortcomings if only mean or only median of 2^{R_i}s is used as an estimate.

   (5 Marks)

7. Perform a hierarchical clustering of the one-dimensional set of points 1, 4, 9, 16, 25, 36, 49, 64, 81, assuming clusters are represented by their centroid (average), and at each step the clusters with the closest centroids are merged.

   (5 Marks)

8. Give an example of a dataset (with a distance measure) and a selection of k initial centroids such that in the process of the K-means algorithm when the points are reassigned to their nearest centroid at the end, at least one of the initial k centroids is reassigned to a different cluster. You must demonstrate the K-means algorithm on your example data.

   (5 Marks)

INDIAN STATISTICAL INSTITUTE

Semestral Examination

M. Tech (CS) - II Year, 2015-2016 (Semester - I)

Optimization Techniques

Date : 10.12.2015 Maximum Marks : 100 Duration : 3.5 Hours

Note: The question paper is of 125 marks. Answer as much as you can, but the maximum you can score is 100.

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 $\mathbf{b}^T$. The i-th element of a vector $\mathbf{b}$ will be written as b_i, and the (i,j)-th element of a mtrix A will be written as A_{ij}. 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) Let $S = \{s_1, s_2, \ldots, s_n\}$ be a set of line segments drawn on a two dimensional grid. For each segment $s_i \in S$, its both the end-points are grid points, and these are given as input. The problem is to compute the minimum number of axis-parallel (horizontal and vertical) lines that can hit all the members in S. Formulate this problem as an integer programming problem.

[10]

(Q2) Let $P, Q \subset \mathbb{R}^n$ be convex sets and let $f : \mathbb{R}^n \to \mathbb{R}$ 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]

(Q3) Deduce the dual of the following LP, where A is an $m \times n$ matrix, $\mathbf{x}$ is an m-dimensional vector, and $\mathbf{y}$ is an n-dimensional vector:

$$\begin{aligned} \text{Minimize} \quad & \mathbf{x}^T A\mathbf{y} \\ \text{subject to} \quad & \textstyle\sum_{j=1}^{n} y_j = 1 \\ & \mathbf{y} \geq \mathbf{0} \end{aligned}$$

[7]

(Q4) Player I has a red card of value 8 and a blue card of value 1. Player II has a red card of value 2 and a blue card of value 7. The players simultaneously choose a card to play. If the chosen cards are of the same color, Player I wins. Player II wins if the cards are of different colors. The amount won, in rupees, is equal to the number on the winners card. Compute the payoff matrix, the value of the game and the optimal mixed strategies of the players.

[5+3+10=18]

(Q5) State and prove the *minimax theorem* for the two player zero-sum game using duality.

[5+10=15]

(Q6) (a) Let us define the two types of problems as follows:

LP Problem: Given an $n \times n$ integer matrix A, an $m \times 1$ integer vector $\mathbf{b}$ and an $n \times 1$ integer vector $\mathbf{c}$, find an $n \times 1$ vector $\mathbf{x} \geq 0$ that minimizes $\mathbf{c}^T\mathbf{x}$ subject to the condition that $A\mathbf{x} = \mathbf{b}$ is satisfied, or report that no such vector $\mathbf{x}$ exists that satisfies $\mathbf{x} \geq 0$ and $A\mathbf{x} = \mathbf{b}$, or report that the set $\{\mathbf{c}^T\mathbf{x} | A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq 0\}$ has no lower bound.

LI Problem: Given an $n \times n$ integer matrix A, an $m \times 1$ integer vector $\mathbf{b}$, does there exist an $n \times 1$ vector $\mathbf{x}$ that satisfies $A\mathbf{x} \leq \mathbf{b}$.

Show that there is a polynomial time algorithm for a LP problem if and only if there is a polynomial time algorithm for the LI problem.
Note: By an integer matrix (resp. vector) we mean a matrix (resp. vector) whose all elements are integers.

(b) Let $n \geq 2$, $S(0,1) = \{\mathbf{x} \in I\!R^n | \mathbf{x}^T\mathbf{x} \leq 1\}$, and
$E = \{\mathbf{x} \in I\!R^n | (\mathbf{x} - \mathbf{t})^T B^{-1} (\mathbf{x} - \mathbf{t}) \leq 1\}$, where

$$\mathbf{t} = \begin{bmatrix} -\frac{1}{n+1} \\ 0 \\ 0 \\ \cdots \\ 0 \end{bmatrix}_{n \times 1} \quad \text{and} \quad B = \begin{bmatrix} \frac{n^2}{(n+1)^2} & 0 & 0 & \cdots & 0 \\ 0 & \frac{n^2}{n^2-1} & 0 & \cdots & 0 \\ 0 & 0 & \frac{n^2}{n^2-1} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \frac{n^2}{n^2-1} \end{bmatrix}_{n \times n}$$

Show that

(i) B is positive definite and so E is an ellipsoid,

(ii) The hemisphere $HS = \{\mathbf{x} | \mathbf{x}^T\mathbf{x} \leq 1 \text{ and } x_1 < 0\} \subseteq E$, and

(iii) $\frac{\text{volume}(E)}{\text{volume}(S(0,1))} < 2^{-\frac{1}{2(n+1)}}$.

(c) State the general idea of the Ellipsoid algorithm for linear programming.

[10+(3+5+5)+5=28]

(Q7) Consider the following quadratic programming problem:

Minimize $\quad \mathbf{c}^T\mathbf{x} + \frac{1}{2}\mathbf{x}^T H\mathbf{x}$,
subject to $\quad A\mathbf{x} \leq \mathbf{b}, \qquad \mathbf{x} \geq 0$,

where $\mathbf{c}$ is an $n \times 1$ vector, $\mathbf{b}$ is an $m \times 1$ vector, H is an $n \times n$ matrix and A is an $m \times n$ matrix with real values, which are input to the quadratic programming problem. Describe Wolfe's method to solve this problem.

[12]

(Q8) (a) Suppose $f(x)$ is a multivariate function which is twice differentiable at $\overline{x}$. If $\bigtriangledown f(\overline{x}) = 0$ and the Hessian matrix $H(\overline{x})$ is positive definite, then show that $\overline{x}$ is a strict local minima.

(b) Minimize $f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$ using Newton's method. You may take the starting point as $x_1 = (0,0)^T$.

[10+10=20]

Indian Statistical Institute

Semester Examination : 2015 – 2016

Master of Technology in Computer Science, Semester III

Functional Brain Signal Processing: EEG & fMRI

Date: 11/12/2015 Maximum Marks: 100 Duration: 3 hours

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

1. Describe T_1, T_2 and T_2^* relaxation time. 4 + 2 + 4 = 10

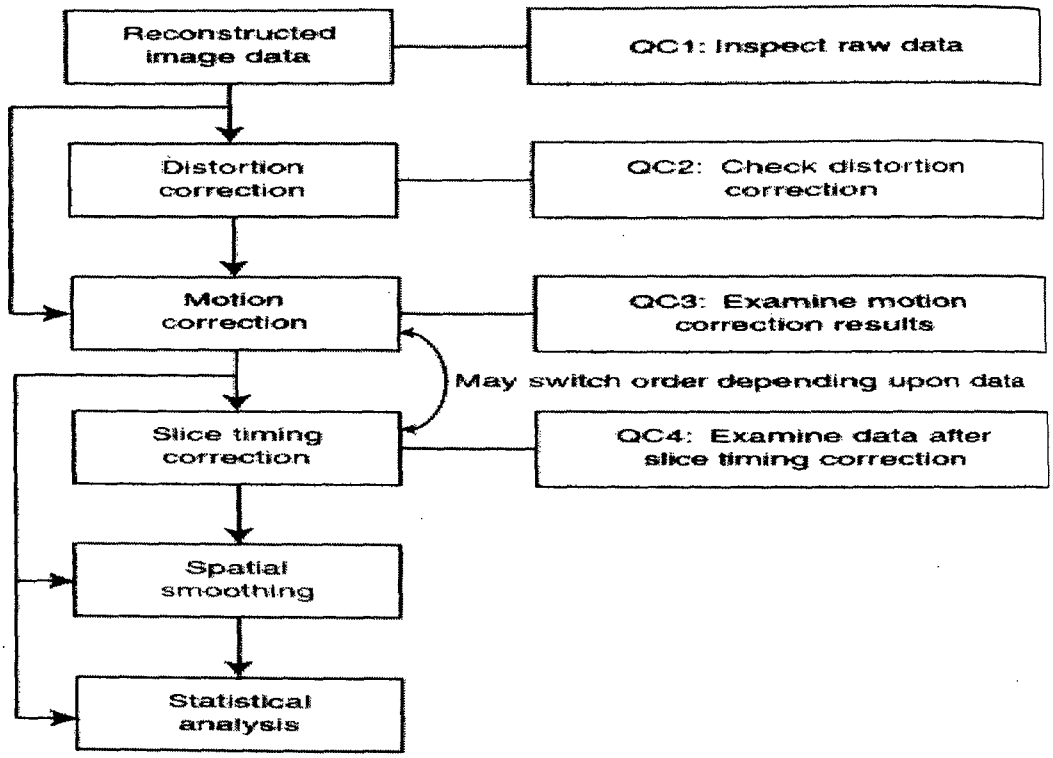
2. Consider the following table:

Tissue	T_1	T_2
Fat	241	85
Brain, white matter	683	90
Brain, gray matter	813	100
CSF	2500	1400

If we want to take an MR image with 128 axial slices, what should be the minimum time required for the best quality of the image and why? 10

3. Describe the general principle of spatial localization in an MR imaging by applying gradient magnetic field along any one of the three spatial axes. Describe it in only one dimension (Notions of slice location, phase encoding and frequency encoding are not required). Mention what is net magnetic field and how that is sensed by a solenoid and then applying appropriate transformation the spatial localization is determined. 10

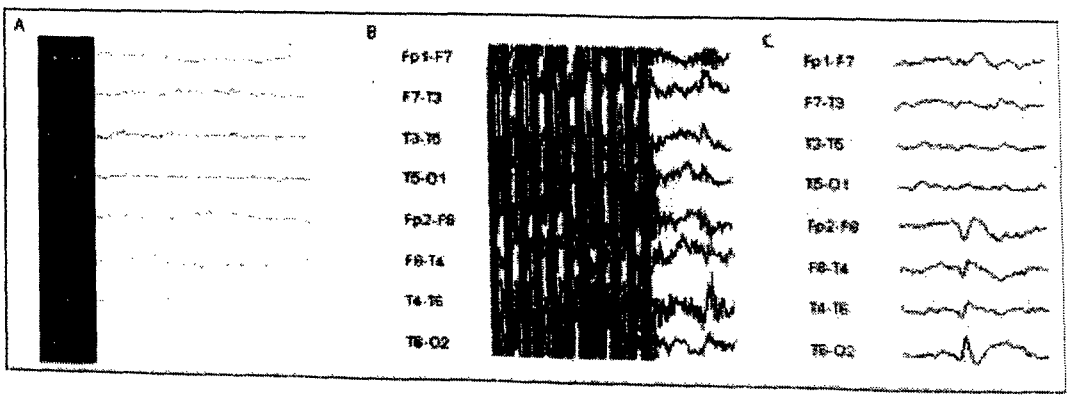
4. The following flow chart gives the usual <u>preprocessing</u> steps involved in processing the fMRI signals. Not all the steps are followed all the time. Identify the step that is always essential and explain why. Describe a possible algorithm for this step (feel free to invent your own, that need not be from the text books, but you will have to give a short logical description). 4 + 6 = 10



5. Any fMRI signal comes with significant ambiguity, because most part of the brain is active for most of the time. How then the specific activation regions are to be identified? Discuss this with an example. 10

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

7. Identify artifacts in the following EEG signals acquisitioned in an fMRI environment. You can mark and name them in the question paper itself and staple that with the answer sheet. Explain the reasons behind occurring these artifacts in the EEG signals recorded in an fMRI environment. 10



8. BOLD activation depends on three prominent attributes in the brain. Name them with a brief explanation how they affect the BOLD signal. $1 + 3 \times 3 = 10$

9. Briefly describe the mathematical formulation of general linear model (GLM) for processing of fMRI. Give only the very basic mathematical equations (put them together in a matrix form). Solutions are not required. 10

10. Briefly describe any four practices followed during simultaneous EEG-fMRI acquisition so as to minimize artifacts in one or other of the signals. $2.5 \times 4 = 10$

INDIAN STATISTICAL INSTITUTE

First Semestral Examination (2015)

M. Tech. (Computer Science) Second Year

Natural Language Processing

Date: 11.12.2015 **Time: 2 hours** **Marks: 50**

NOTE:

- Answer all questions.
- Questions are grouped into two parts, PART-I (Marks: 15) and PART-II (Marks: 35).
- PART-I has to be answered on the question paper itself and must be submitted **within** 20 **minutes** after the examination begins.
- Use the given answer sheet for answering PART-II questions only.

PART-II

Q2. Direction: Label the following statements as True or False. Briefly justify your answer **[no marks will be given if the justification is not proper]** **[Marks: 5 + 5 + 5 = 15]**

(i)

S	→	NP VP				
VP	→	V	|	V NP		
NP	→	Det N	|	N		
V	→	*sells*				
Det	→	*the*				
N	→	*sport*	|	*Dasgupta*	|	*items*

Statement: The above Grammar (G) is **unambiguous** because there is only one rule for the Start Symbol, S; hence, for any sentence (accepted by G) only one parse tree is generated.

(ii) Consider the English word, **bank** and its one of the 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(कूल | bank) = 1/4; p(बैंक | bank) = 1/4; p(किनारा | bank) = 1/4; p(ढाल | bank) = 1/4;

(iii) The definition of **suffix**: a morpheme added at the end of a word to form a derivative (e.g. *-ation*, *-fy*, *-ing*, *-itis*). Consider the following algorithm:

- Arrange all the English Dictionary Head words into a Trie structure, T
- C is an English corpus
- Suffix list L = {}
- For each surface word, w ∈ C
    1. Trace w into T: trace is successful if a leaf node is reached and unsuccessful if no leaf node is reached.
    2. If the trace is successful:
        - 2.1 Break w as x + y where the substring, x is traced to reach the leaf node and y denotes the remaining substring (note that y can be null).
        - 2.2 $L = L \cup \{y\}$
- Output L

Statement: The above algorithm produces all the valid suffixes in English.

Q3. Assume that a sufficiently large language corpus is divided into six parts (r_1, r_2, ..., r_6). You observe the numbers of pronouns in three particular parts (say, in r_i, r_j and r_k) are 200, 100 and 300. So the observed mean($\bar{n}$) is 200. For distribution of pronouns, let p_i denote the probability of observing pronouns in r_i and n_i denote the number of pronouns observed in r_i. Find one possible distribution (write down the values of p_i's and n_i's) of the pronouns in the six parts which is consistent with the following two constraints and the distribution should have **maximum entropy** over the other possible distributions:

$$\text{Constraint 1: } \bar{n} = \sum_{i=1}^{6} n_i p_i \text{ ; Constraint 2: } \sum_{i=1}^{6} p_i = 1$$

Justify your answer. **[Marks: 10]**

Q4. Consider the Probabilistic Context Free Grammar (PCFG) as given below. The non-terminals are S, NP, VP, V, PP, N and P; S being the start symbol. The terminals are the words in italics.

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

Using the above grammar, there are two possible parses of the sentence *fire flies like fire*. (i) Show these parses, (ii) Compute the probabilities of these two parses, (iii) Compute the sentence probability. **[Marks: (2 + 2) + (2+2) + 2 = 10]**

INDIAN STATISTICAL INSTITUTE

Semestral Examination : 2015 – 16

MTech CS (2nd Year)

Computational Finance

Date: 4 December 2015 Maximum Marks: 100 Duration: 3 Hours

1. Let

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

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

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

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

2. In the two period model, explicitly solve the Consumption Investment problem for the utility function $u(w) = \frac{1}{\gamma} w^{\gamma}$ where $0 < \gamma < 1$. Show that the Lagrange Multiplier is

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

the optimal attainable wealth is

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

and the optimal objective value is E[u(W)] = λv/γ. Here v is the initial wealth, L is the state price density and B_1 is the bank process.

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 $P(\omega_1) = 3/5$.

[6 + 4 + 4 + 6 = 20]

3. In the context of high frequency trading distinguish between *feature selection* and *feature engineering* with an example. Explain the use of such techniques in Algorithmic trading. [10 + 10 = 20]

4. (a) What is a *Kaplan-Meier* estimator? Illustrate its use in algorithmic trading.

(b) What is the main difference between conventional asset markets and *dark pools*? [(5 + 7) + 8 = 20]

5. What is a *Levy* process? When do you need a Levy process to model financial asset prices? Explain with examples of real life situations.

Distinguish between such models and the conventional *Black-Scholes* model with an example of a financial asset price model.

[5 + 5 + 10 = 20]

INDIAN STATISTICAL INSTITUTE
Mid-Semester Examination
M. Tech. (CS) II year (2nd Sem): 2015–2016
Advanced Cryptology

Date: 22. 02. 2016 Maximum Marks : 50 Time : 2.5 Hours

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

1. (a) What is the common philosophy behind the bounds like Markov's inequality, Chebyshev's inequality and Chernoff Bound?
   (b) Define pairwise independence and total independence. Does one imply the other?
   (c) State Chernoff-Hoeffding bound.

   [2 + (4 + 2) + 2]

2. (a) Can a probabilistic Turing machine be viewed as a deterministic Turing machine? If yes, how? If no, why?
   (b) What is the containment relationship amongst P, PP, BPP and P/Poly?

   [(2 + 4) + 4]

3. (a) Show that if a positive function $\epsilon(n) \in O(1/p(n))$ $\forall$ positive polynomials $p(n)$, then $\epsilon(n)$ is negligible.
   (b) Prove that a positive function $\epsilon(n)$ is negligible if and only if $\forall$ positive polynomials $p(n)$ the product $p(n)\epsilon(n)$ converges to zero.
   (c) Define a function that is neither negligible, nor noticeable, nor overwhelming. Justify your claim.

   [2 + 4 + 4]

4. (a) Define a weak one-way function, explicitly using all the required quantifiers.
   (b) Logically negate your above statement to derive a definition of a *not weak one-way* function.
   (c) Define a trapdoor one way permutation. Given an example.

   [3 + 3 + (2 + 2)]

5. (a) Show that two ensembles of probability distributions may be computationally indistinguishable, but statistically distinguishable.
   (b) What is a hybrid random variable? What is its use?

   [6 + (2 + 2)]

6. (a) Prove that a sequence is next-bit unpredictable if and only if it is pseudo-random.
   (b) If we have an n-bit to $(n+1)$-bit PRG, can we use it to make an n-bit to 2^n-bit PRG?

   [8 + 2]

Roll No: _____

Information Retrieval Mid Sem Examination

M.Tech CS 2nd Year, 2nd Semester
Full marks: 60
Time: 90 minutes

22 February 2016

Instruction: Write your roll number at the top. Use the institute provided answer sheets for rough work, if any. Staple the question paper and the rough paper and submit both. For each question, tick the answer of your choice. For each question, 3 marks would be awarded for the correct answer, zero for not attempting and -1 for a wrong answer.

1. To compress index and vocabulary we need to use
   (a) Lossless compression
   (b) Lossless compression for vocabulary and lossy compression for posting lists
   (c) Lossy compression
   (d) Lossy or lossless depending on storage limitations
2. To be able to answer phrase queries, an inverted index needs to store
   (a) In every posting list, the tf.idf scores of the term in each document
   (b) In each posting list, the positions of the term for every document in which the term occurs
   (c) Skip pointers in every posting list
   (d) For every term, the list of other terms that appear after the term across all documents
3. Which of the following is not an important reason (in other words, the least important among these) for relevance feedback being very little used in web search?
   (a) Relevance feedback is hard to explain to the common user, thus makes the user interface complex
   (b) Most web search users want to finish the search in a single or minimal interaction
   (c) Relevance feedback would expand the query and hence make the search engine's response time several times slower
   (d) Most web search users are not interested in enhancing recall and relevance feedback is essentially a recall enhancing strategy
4. Suppose L is a list of n (document id, score) pairs sorted by document ids. What is the complexity of the most efficient algorithm to determine the top k documents ranked by score from L? Assume $k << n$.
   (a) $O(n \log n)$
   (b) $O(n + k \log n)$
   (c) $O(n + n \log k)$

(d) $O(n^2)$

5. In practice, the inverted index for a text document collection can be compressed with a reasonable compression factor because

   (a) Text data can be usually compressed well.

   (b) The posting lists may have very large document ids, but the gaps between consecutive document ids are often small. Hence the numbers in the lists can be encoded with less number of bits.

   (c) The document vectors can be encoded as a vector of numbers, hence they would take up less space.

   (d) The text documents contain characters, each of which takes up 1 byte. In the index, each word is represented by a term-id which is an integer.

6. Suppose, in a text document collection, the most frequent term occurs 1,000,000 times in total. What would be a reasonable estimate for the total number of occurrences of the 100th most frequent term in the same collection?

   (a) About 100

   (b) Cannot be estimated at all

   (c) About 10,000

   (d) About 1,000

7. A document is said to be relevant to a query if

   (a) According to the perception of user or a domain expert, it satisfies the information need expressed by the query

   (b) The terms are present in consecutive positions in the document

   (c) The terms in the query are present in the document

   (d) The terms in the query are present in the doucment with high frequency

8. Suppose there are exactly 10 relevant documents for a 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, 1.0, 0.75 and 0.5 respectively. Then, the precision (p@5) achieved by the system after the top 5 ranked documents is

   (a) 0.69

   (b) 0.625

   (c) 0.6

   (d) Cannot be determined from the given information

9. If the length of each posting list is assumed to be n, then in the worst case, the NRA algorithm may have to scan up to

   (a) Up to length $O(n \log n)$

   (b) Up to length $O(n)$

   (c) Up to length $O(\sqrt{n})$

   (d) Up to length $n/2$

10. Suppose A is a term-document matrix, each row corresponding to a term, each column corresponding to a document. The entries are the tf.idf scores. The matrix $C = (c_{ij})$ is defined as $C = AA^T$. Which of the following statements is not true?

(a) If the (i, j)-th entry of C is relatively higher than most other entries in the same row, then it indicates that the i-th term and the j-th term occur together in many documents.

(b) C is symmetric matrix, that is, $c_{ij} = c_{ji}$ for all i and j.

(c) The diagonal entries of C are greater than or equal to all other entries in the corresponding row. In other words, $c_{ii} \geq c_{ij}$ for all $j \neq i$.

(d) If σ is a singular value of A then σ^2 is an eigenvalue of C.

11. 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. Which orderings of the documents in the posting lists would allow the system to perform a linear merge?

(a) Ordered by document id or by descending authority scores

(b) Ordered by document id or by descending tf.idf scores

(c) Ordered by descending static authority scores or by descending authority scores

(d) Ordered by descending tf.idf scores or by ascending tf.idf scores

12. Suppose A is a term-document matrix obtained from a text document collection. The matrix $C = AA^T$ is the matrix representing co-weights between terms and $C = U\Sigma^2U^T$ is the SVD of C. In practice, the least sparse (most dense) matrix among the following would be (assume the fraction of zero entries in a matrix is the measure of its sparseness)

(a) C

(b) U

(c) Σ

(d) A

13. Stemming should definitely not be used for

(a) A field which stores the abstract of a paper in a paper dataset

(b) A field which stores a long question in a question-answer dataset

(c) A field which stores the names of persons in a resume dataset

(d) A field which stores the answer to a question in a question-answer dataset

14. Using a variable byte encoding with a 1 byte unit, first bit being the indicator (1 indicating this is the last byte) and the next 7 bits being the payload, the bit represenation of 128 would be

(a) 11000000

(b) 11111111

(c) 00000001 10000000

(d) 00000000 00100000

15. In a vector space model, if (the inverse of) Euclidean distance is used to measure the similarity of document vectors, then

(a) The similarity of every document with a very long document will be high, because of the length of the vector corresponding to the long document

(b) One long document and one short document on the same topic would be far apart although they should be considered similar

(c) It would be ideal because vectors close to each other would be considered as similar

(d) Documents of the same length will be similar to each other though that is not desired

16. Suppose $d(s_i, s_j)$ denote the edit distance between two strings s_i and s_j and $|s|$ denote the length of the string s. Which of the following statements is not correct?

    (a) $d(s_i, s_j) \leq |s_i| + |s_j|$

    (b) $d(s_i, s_j) \geq abs(|s_i| - |s_j|)$

    (c) $d(s_i, s_j) \leq \max\{|s_i|, |s_j|\}$

    (d) $d(s_i, s_j) \geq \min\{|s_i|, |s_j|\}$

17. 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. Which of the scenarios below most accurately correspond to the *find other similar documents like this one* feature in a search engine?

    (a) $\beta = \gamma = 1, |D_r| = 0, |D_{nr}| = 1$

    (b) $\alpha = 1, \beta = \gamma = 0$

    (c) $\gamma = 0, |D_r| = 1, |D_{nr}| = 0$

    (d) $\alpha = \beta = \gamma = 1/3, |D_r| = |D_{nr}| = 1$

18. Which of the following statements is not correct?

    (a) Stemming may lower the precision in some cases

    (b) Skip lists are not useful for queries of the form `X OR Y`

    (c) Stemming may increase the size of the vocabulary

    (d) For phrase-queries, bi-word index may produce some false-positives

19. Which of the following is not an entry generated from the term `midsem` in the permuterm index?

    (a) m$midse

    (b) idsem$m

    (c) mi$dsem

    (d) sem$mid

20. In which framework, documents d_is are ranked by the expression $P(Relevant = 1|d, q)$, for a given query q?

    (a) Retrieval using standard relevance feedback.

    (b) Retrieval using the vector space model

    (c) Language model based retrieval

    (d) Probabilistic information retrieval

INDIAN STATISTICAL INSTITUTE

M. Tech. (CS) II Year (2015-16), II semester

Periodical Examination

ADVANCED PATTERN RECOGNITION

Date: 23-2-2016 Duration: 150 minutes Marks: 70

Note: Answer all the questions.

1. Let $(x_i, y_i); i = 1,2,..., n$ be the given data points, where $x_i, y_i \in \Re \forall i$. Derive the expressions for regression line of y on x, and the error value of the model. [10]

2. Describe the crossover and mutation operations in Genetic algorithms. [8+5=13]

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

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

5. (a) Describe the k-nearest neighbor density estimation procedure.
   (b) Derive the k-nearest neighbor decision rule from the above procedure.
   (c) Describe an algorithm for reducing the size of the training sample set for k-nn decision rule. [5+7+8=20]

Mid-Semester Examination (2016)

MTech (CS)

Subject: Computer Vision

Date: 24.02.2016 Time: 2 hours 15 mins (2.30-4.45 pm)

(Answer as many questions as you can, but the **maximum that you can score is 60**)

1. Compare in terms of similarities and dissimilarities between a camera based machine vision system and an eye based biological vision system. Using the pinhole camera model, explain what you mean by perspective projection. What is thin lens approximation? Derive the corresponding thin lens formula.

   5+5+3+7=20

2. Explain the concept of receptive field. Hence explain the basic model for a spatial filter representing the receptive field. How is such a spatial filter implemented in computer vision? Explain how the Marr-Hildreth operator is related to the above-mentioned receptive field model? Derive the mathematical expression corresponding to this operator and show that it is rotationally symmetric. How is "raw primal sketch" implemented through this operator?

   4+5+4+4+5+3=25

3. Draw a neat diagram to geometrically derive the camera co-ordinates for perspective transform corresponding to any world co-ordinate (X, Y, Z). Next, derive the perspective transformation matrix. Using the same, derive the camera co-ordinates and verify the results from geometric derivation. Then compute the inverse perspective transformation matrix, and hence show that depth information is imperative to calculate the other way round.

   5+4+5+4+3=21

Indian Statistical Institute

M.Tech (CS) II

Information Security and Assurance

Mid Semester Examination

Maximum Marks: 70

Date: February 24, 2016
Time 2.5 hours

The question paper contains 7 questions. Total marks is 70. Maximum you can score is 60.

1. Alice and Bob want to establish a common key between them.
   (a) How can they do so using Diffie-Hellman key exchange protocol?
   (b) How can they do so, using a polynomial based technique?
   (c) Compare the (i) security (ii) computation cost, (iii) communication cost of (a) and (b) above. (4 + 4 + 6 = 14)

2. A group key enables a group of n users to communicate securely with one another in the group using a common key. Can a polynomial based scheme be used to construct a group key? If so, how? What is the security of such a scheme? (5 + 3 = 8)

3. 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)

4. Show that the plain RSA signature algorithm is insecure by demonstrating two attacks on it. How will you modify this algorithm to counter these attacks? State any assumptions that you make. (6 + 4 = 10)

5. Prove that the El Gamal encryption scheme is not CCA-secure. (8)

6. How can Bloom filters be used to construct password checker? Can you suggest an alternate technique. (6+2 = 8)

7. How can you construct Schnorr Signature scheme using Schnorr identification scheme? Discuss the security of Schnorr Signature scheme. (8+4=12)

INDIAN STATISTICAL INSTITUTE

Mid-Semestral Examination : 2015 – 16

MTech CS (2^{nd} Year)

Algorithmic Game theory

Date: 25 February 2016 Maximum Marks: 30 Duration: 2.30 Hours

This paper carries 32 marks. Attempt all questions. The maximum you can score is 30.

1. Prove that if $(x1;y1)$ and $(x2;y2)$ are mixed Nash equilibria of a two-player zero-sum game, then so are $(x1;y2)$ and $(x2;y1)$. [6]

2. Present linear programming formulations for the following problems:
   a. Given the payoff matrix of a two-player zero-sum game, compute a mixed-strategy Nash equilibrium.
   b. Given a k-player game in normal form, compute a correlated equilibrium.

   [4 + 4 = 8]

3. Show that the Randomised Weighted Majority (RWM) algorithm with update rule $w_i^t = w_i^{t-1}(1-\eta)^{l_i^{t-1}}$ achieves the same external regret bound as for the Polynomial Weighted (PW) algorithm, for losses in [0, 1]. [8]

4. Consider an n player game in which each player has 2 strategies. For this problem, think of the strategies as "on" and "off." For example, the strategy can be either to participate or not to participate in some event. Furthermore, assume that the game is symmetric, in that all players have the same payoff functions, and that the payoff for a player depends only on the strategy of the player and the number of people playing strategy "on." So the game is defined by $2n$ values: $u_{on}(k)$ and $u_{off}(k)$, which denote the payoff for playing the "on" and "off" strategies, assuming that k of the other players chose to play "on" for $k=0,\ldots,n-1$.

   Give a polynomial time algorithm to find a correlated equilibrium for such a game. Note that the input to this problem consists of the $2n$ numbers above. As usual, polynomial means polynomial in this input length. You may use the fact that linear programming is solvable in polynomial time. [10]

Indian Statistical Institute

Advanced Image Processing

M.Tech.(CS)-II Year: 2015-16

Full marks: 60 Time: 2 Hours

Date: 26.02.2016

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

1. Assuming pin-hole camera model, prove that perspective projection maps a straight line in 3D to a straight line in 2D. [10]

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

   b) Derive the equation:

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

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

3. Derive the expression of non-parametric Wiener filter in spatial domain using minimum mean-square approach. [10]

4. a) State and prove the correlation theorem.

   b) Show that the Fourier transform of the autocorrelation function of $f(x)$ is its power spectrum. [7+3=10]

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

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

   [(5+2)+3=10]

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

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

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

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

[(2+2)+6=10]

9. Define (i) path, (ii) connected component, (iii) vanishing point, (iv) medial axis, and (v) restoration. [5X2=10]

INDIAN STATISTICAL INSTITUTE

Mid-Semester Examination: (2015-2016)
M.Tech C.S., 2nd Year

Advanced Digital Signal Processing

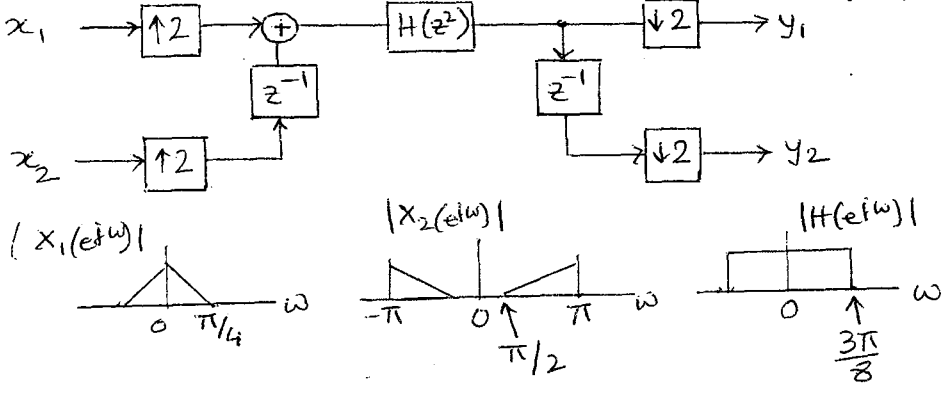
Date: 26.2.2015 Maximum Marks: 60 Duration: 2 hours

Note: The marks add up to 69. The maximum you can score is 60. The exam is open-book, open-notes. You are permitted to use calculators.

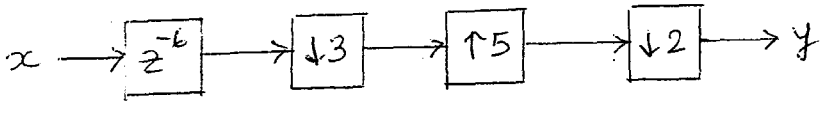
Number of question papers required: 3

Questions:

1. For the multirate system shown below and $X_1(e^{j\omega})$, $X_2(e^{j\omega})$ and $H(e^{j\omega})$ as indicated, sketch the outputs $Y_1(e^{j\omega})$ and $Y_2(e^{j\omega})$. [8+8]



2. Obtain an expression for the output in terms of the input, for the structure given below:



[10]

3. Design a two-stage interpolator (interpolation factor of each stage to be decided by you) which will increase the sampling rate from 3 kHz to 60 kHz. The decimation filter is to be designed as an equiripple filter with a passband edge at 1.2 kHz and passband and stopband ripples of 0.02 and 0.01 respectively. Compare the computational complexity with that of a single stage design. [10+5]

4. Show that

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

where

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

[10]

5. A continuous-time signal $x_c(t)$ has a Fourier Transform such that $X_c(j\Omega) = 0$ for $|\Omega| \geq 2\pi(8000) rad/s$. Assume that it is sampled at the Nyquist rate.

   (a) For a STFT computation, what should be the window length, L, so that it corresponds to a $15ms$ section of the signal?

   (b) If the DFT length N is set equal to L, what is the spacing in Hz between the DFT samples $X[k]$?

   (c) If the window offset is chosen to be $L/4$, how many windows would be obtained?

   (d) If the DFT length N is set equal to $2L$, and the corresponding zero-padded DFT is denoted as $X_p[k]$, what is the relationship between $X[k]$, $k = 0, \cdots, L-1$ and $X_p[k]$, $k = 0, \cdots, 2L-1$

[4+5+4+5]

INDIAN STATISTICAL INSTITUTE

Mid-Semestral Examination of Second Semester (2015-2016)

M.TECH.(CS) II YEAR

Topics in Algebraic Computation

Date: 26.02.2016 Maximum marks: 60 Duration: two and half hours

Note : The paper contains 80 marks. Answer as much as you can. The maximum you can score is 60.

1. (a) Let A be an $n \times n$ tridiagonal matrix. Describe an efficient algorithm to find the null space of A.

   (b) Let A be an $n \times n$ matrix which has at most 10 non-zero entries in each row. Describe an efficient way to represent A. Let P be an $n \times n$ permutation matrix. Describe an $O(n)$ time algorithm to compute PA.

   (c) Let A be an $n \times n$ matrix with entries from $\mathbb{F}[x]$, where $\mathbb{F}$ is a field. Define the k-th determinantal divisor $d_k(A)$ and the determinantal rank $\rho(A)$ of A. Show

   i. $d_k(A)$ divides $d_{k+1}(A)$.

   ii. If A and B are equivalent, then $\rho(A) = \rho(B)$.

   $(10 + 10 + 10 = 30)$

2. (a) Describe the main algebraic ideas due to which the DFT of an n-tuple can be computed in $O(n \log n)$ ring operations.

   (b) Provide a pseudo-code for an in-place computation using the FFT algorithm.

   (c) Let $a(x)$ and $b(x)$ be two polynomials of degrees 10 each having integer coefficients which are in the range $[-15, 15]$. Determine an appropriate setting to apply the FFT to compute $a(x)b(x)$. Explain all your steps.

   $(13 + 10 + 7 = 30)$

3. (a) Let $a(x)$ and $b(x)$ be polynomials of degrees $3m$ each. Show that $a(x)b(x)$ can be computed using 5 multiplications of polynomials of degree m and an additional $O(m)$ steps.

   (b) Consider the following recurrence:

   $$\begin{aligned} T(n) &= 1 && \text{if } n = 1; \\ T(n) &= aT(n/c) && \text{if } n > 1. \end{aligned}$$

   Here a and c are positive real numbers with $a > c$. Show that if n is a power of c, then $T(n) = O(n^{\log_c a})$.

   $(10 + 10 = 20)$

Information Retrieval
M.Tech CS II 2015-16 End Sem

Time: 3 Hours | Full marks: 50

Date: 19.04.2016

1. Recall the construction of modified query q_m from an initial 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, D_{nr} is the set of known non-relevant documents, and α, β, γ are non negative parameters.

   (a) Under what conditions (in terms of α, β, γ, D_r and D_{nr}) would the modified query q_m be the same as the original query q_0? Explain with justification.

   (b) In all other cases, is q_m closer to the centroid of the relevant documents than q_0 is? Prove your answer.

(4 + 4 = 8 Marks)

2. (a) Suppose A and B are two question answering (QA) systems. For a particular question q, the two systems return ranked lists of 5 answers each, out of which some are right (marked R) and some are wrong (marked W), as follows:

   - System A: R W W W R
   - System B: W R R W W

   Which system would be preferred as a QA system and why? Explain with an appropriate measure of effectiveness for QA systems.

(3 Marks)

   (b) The AskMSR question answering (QA) system rewrites factoid natural language queries into several forms and performs a *phrase search* hoping to get a sentence which would contain the answer. Suppose a given question is:

Who is the founder of the Indian Statistical Institute?

   Then,

   i. What would be the phrase queries from the above question that the system would form by rewriting?

   ii. Which (if any, or none) of the phrase queries are expected to match a sentence in some document containing the answer? Give an example of such a sentence, if any.

(2 + (2 + 1) = 5 Marks)

3. Compare the effectiveness of the variable byte encoding and the Gamma encoding for (a) an integer less than 20 and (b) an integer greater than 1 billion, of your choice. Can your findings be generalized for the cases of small and large integers? Justify your answer.

((3 + 3) + 3 = 9 Marks)

4. Suppose a retrieval system has an inverted index I_{old}. When any new document or update arrives, it updates them into a new index I_{new}. Periodically, the two indexes are merged to construct a single index I which subsequently would become I_{old}. When a query arrives, the system can perform search over the two indexes separately, get ranked results and then merge the ranked lists to produce a single final ranked list of documents. Assume that the ranking from any index is performed using the score of a document d for a query q, which is defined as the sum of the tf.idf scores of d in the posting lists of the terms in q.

$$\text{score}(q, d) = \sum_{w \in q} \text{tf.idf}(w, d)$$

Similarly, the score of a document d in the final ranked list would be the sum of its scores in the two ranked lists obtained from the two indexes.

Consider the two scenarios: (1) the retrieval is performed from the two indexes and then ranked lists are merged, and (2) the two indexes are merged to construct a single index and retrieval is performed from the merged index. Analyze if the ranked lists would necessarily be the same in the two scenarios for any given query. Justify your answer.

(6 Marks)

5. Suppose a query suggestion system uses query logs and relies on within session query reformulation information to suggest queries. Each record in the query log contains the *IP address* of the user, the *timestamp* of the query, the *query text*, the *browser information* and the *operating system*. However, the system does not have information about the user ids (such as Google id when the users are logged in). Come up with a method with which the system would be able to identify each query session (as accurately as possible) and describe your method justifying each major step.

(6 Marks)

6. Suppose you want to suggest context sensitive spelling corrections (where the misspelled word may actually be a valid dictionary word but wrong in the context where it is used) in queries or in documents. For example, the sentence *"I want to go their"* should be written as *"I want to go there"*, or *"India will loose the game"* should actually be *"India will lose the game"*. Design an algorithm to suggest such corrections and describe the major steps of your algorithm.

(6 Marks)

7. The Heap's law estimates the vocabulary size M (the number of distinct terms) of a document collection as a function of the size of the collection as:

$$M = kT^b$$

where typically $30 \leq k \leq 100$ and b is approximately 0.5. Also, the Zipf's law estimates the collection frequency f_i of the i-th most frequent term as

$$f_i = c/i$$

for some collection specific constant c.

According to Heap's law, does the vocabulary size converge to some large number or go to infinity? Analyze the same using Zipf's law.

(4 + 4 = 6 Marks)

8. Security is an important consideration for retrieval systems in corporations. A low-level employee should not be able to find the salary roster of the corporation, but authorized managers need to be able to search for it. The search results shown to the user must not contain documents they are barred from opening; the very existence of a document can be sensitive information. One option to take care of this problem is to check for every document if the user is allowed to view it before finally displaying it as a search result. However, this slows down response time.

Suppose in a company there are users with three kinds of access privileges. Suggest an approach (of maintaining the user - document access control information, and accordingly a method to retrieve documents when a user submits a search query) other than the option described above to implement access control in search for the documents of the company. Would your approach has some serious challenges in some scenarios? Explain.

(4 + 2 = 6 Marks)

Indian Statistical Institute

Mid-Semester Examination 2015-2016
M. TECH.(CS) II Year
Subject: Document Processing and Retrieval
Full Marks: 50 Duration: 2 hrs.

Date: September 21 2015

(Answer all questions)

1. (a) Define compression ratio, compression gain, lossless and lossy coding, Codec. 1X5=5

(b) How many different files of size N bits can be generated? Out of those how many files can at most be compressed into files of N/4 bits so that the original N-bits file could be uniquely retrieved? 2+5=7

2. (a) What is Run-Length Encoding (RLE)? For a string of N characters undergoing RLE, give an expression of compression factor that may be achieved in general. Give some examples with reasons where RLE is expected to succeed. What are the scanning modes for generating run lengths in digital image? 2+5+2+1=10

3. (a) Define information and entropy of symbols. How entropy is related to data compression? Define redundancy in data in terms of entropy. 2+2+2=6

(b) Describe how m^{th} order Golumb code can be generated and its decoding approach. How the value of m can be decided for a set of n data? What kind of data the Golumb code suitable for? 4+2+2=8

(c) Generate 6^{th} order (m = 6) Golumb code for first 12 run-Length symbols (i.e. n = 0, 1, 2, 3, ., ., ., 11). 4

4. Write down the block diagram of a document analysis system. What is a skew of a document image? Describe an approach for skew detection of a document image. 3+2+5=10

INDIAN STATISTICAL INSTITUTE

Mid Semestral Examination

M. Tech (CS) - II Year, 2015-2016 (Semester - III)

Optimization Techniques

Date : 21.09.2015 Maximum Marks : 60 Duration : 3.0 Hours

Note: The question paper is of 80 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

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

(Q1) Let $\mathcal{D}$ be the dual of $\mathcal{P}$. Show that the dual of $\mathcal{D}$ is $\mathcal{P}$. [5]

(Q2) Show that $\mathcal{P}$ has an optimal solution if and only if the following set of constraints has a feasible solution.

$$\begin{aligned} A\mathbf{x} &\leq \mathbf{b} \\ A^T\mathbf{y} &\geq \mathbf{c} \\ \mathbf{c}^T\mathbf{x} &\geq \mathbf{b}^T\mathbf{y} \\ \mathbf{x}, \mathbf{y} &\geq \mathbf{0} \end{aligned}$$

[10]

(Q3) (i) State and prove the weak duality theorem.

(ii) Argue with proper reasons that for a primal linear program $\mathcal{P}$ and its dual $\mathcal{D}$, exactly one of the following possibilities can occur:

(a) Neither $\mathcal{P}$ nor $\mathcal{D}$ has a feasible solution.

(b) $\mathcal{P}$ is unbounded and $\mathcal{D}$ has no feasible solution.

(c) $\mathcal{P}$ has no feasible solution and $\mathcal{D}$ is unbounded.

(d) Both $\mathcal{P}$ and $\mathcal{D}$ have a feasible solution, and their optimal solutions are same.

State results that you might need to argue for the above answer.

[5+10=15]

(Q4) Solve the following linear program using the simplex method.

$$\begin{array}{lll} \text{Minimize} & 12x_1 + 3x_2 + 4x_3 & \\ \text{subject to} & 4x_1 + 2x_2 + 3x_3 & \geq 2 \\ & 8x_1 + x_2 + 2x_3 & \geq 3 \\ & x_1, x_2, x_3 & \geq 0 \end{array}$$

[20]

(Q5) Deduce the conditions on A, $\mathbf{b}$ and $\mathbf{c}$ so that the primal linear program $\mathcal{P}$ and its dual $\mathcal{D}$ are the same linear program. [5]

(Q6) Check whether the vector $[3 \quad -1 \quad 0 \quad 2]$ is an optimal solution to the linear program without using simplex to solve the linear program.

$$\begin{array}{lll} \text{Maximize} & 6x_1 + x_2 - x_3 - x_4 & \\ \text{subject to} & x_1 + 2x_2 + x_3 + x_4 & \leq 7 \\ & 3x_1 + x_2 - x_3 & \leq 11 \\ & x_2 + x_3 + x_4 & = 3 \\ & x_1, x_2, x_3, x_4 & \geq 0 \end{array}$$

[10]

(Q7) In the *set cover* problem, we have an universe $\mathcal{U} = \{u_1, \ldots, u_n\}$ of n elements. Let $\mathcal{S} = \{S_1, \ldots, S_m\}$ be a set of m sets, where each set $S_i \subseteq \mathcal{U}$. Each set S_i has a weight $w_i \geq 0$. The problem in *set cover* is to find a minimum weight collection of subsets of $\mathcal{S}$ that covers all elements of $\mathcal{U}$.

(a) Write an integer linear program (ILP) for the *set cover problem* using decision variables x_i to indicate whether the set S_i is included in the solution or not.

(b) Relax the above ILP and round the optimal solution of the linear program as follows: given the optimal solution $\mathbf{x}^*$ of the linear program, we include the subset S_i in our solution if and only if $x_i^* > \frac{1}{f}$, where f is the maximum number of sets in which any element appears and x_i^* is the i-th component of $\mathbf{x}$.

For this rounding scheme, show that the set generated is a set cover and is an f-factor approximation algorithm.

[5+10=15]

INDIAN STATISTICAL INSTITUTE

Mid-Semestral Examination : (2015 - 2016)

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

Subject Name : Neural Networks & Applications

Date : September 22, 2015 **Maximum Marks : 50** **Duration : 2 hrs**

Answer all the questions.

1. Explain the main characteristics of Artificial Neural Networks. [5]

2. Explain the main computational differences between the von Neumann computer and Artificial Neural Networks. [5]

3. Describe briefly, along with appropriate real life examples, the basic types of artificial neural network learning algorithms. [10]

4. Consider two sets of labeled sample points in a three dimensional feature space. 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 in the transformed feature space, become linearly separable.
   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 the above RBFNN. For deriving this learning rule, consider only the sets of above labeled samples, but not their distribution.

   [2 + 4 + 10 + 14 = 30]

INDIAN STATISTICAL INSTITUTE

Periodical Examination: (2015 – 2016)

M.Tech. (CS) II Year

Parallel Processing: Architectures and Algorithms

Date: 22/09/2015 Total Marks: 70 Duration: 2 hrs

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

1. a) Show the schematic diagram of a *shared-memory* SIMD computer mentioning its features. Define the different theoretical models of *shared-memory* assumed in PRAM model.

   b) Given a very large unsorted file consisting of n distinct entries, it is required to search if a given input x is present in the file. Write the steps to solve it on an EREW SM PRAM with N processors, $N < n$. Find the worst-case time complexity. Can you improve the average-case complexity using some flag for early-termination with EREW model? If not, can you choose any other model of SM for it? What will be the problem, in case the items of the file are not distinct? How can you resolve it?

   [(2+2)+8 = 12]

2. (a) State Bernstein's conditions for parallelism.

   (b) Consider the following program segment with six instructions:

   P_1 : A = B * C
   P_2 : D = B + A
   P_3 : E = A + C
   P_4 : F = D * E
   P_5 : B = E + A
   P_6 : C = D * A

   Draw the data dependence graph considering each statement as a process. Ignore the resource dependencies. Show a possible scheduling of the processes on two processors following 'ETF (earliest start time first)' algorithm. Assume that the operations '+', '*', and inter-process communication take 10, 40 and 30 time units respectively. Calculate the speed-up achieved by your scheduling.

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

3. Answer in brief:

   a) Prove that H_n, a hypercube of order n, is Hamiltonian, for $n \geq 2$.

   b) Find an expression for the *Moore bound* on the number of nodes in a regular graph with degree d and diameter k.

   c) Prove that if each node of an $N \times N$ mesh contains one packet to be routed to a unique destination, following farthest-first strategy, routing can be completed in $(2N-2)$ steps. Assume that links are bidirectional.

   d) Prove that $(2m-1)$ middle-stage switches are sufficient for non-blocking operation of a 3-stage $N \times N$ Clos' network using $(n \times m)$ switches at the input stage, where N is an integral multiple of n.

   [6 × 4 = 24]

P.T.O

4. a) Identify the function the given CUDA kernel code computes explaining the steps.

b) Is the given code correct? If not, find the errors and correct it.

c) With the knowledge of the underlying GPU architecture of the given machine, how can you optimize the performance of this CUDA kernel in terms of speed-up? Justify your answer in brief.

```
// Setup the execution configuration

#define  dim 2
dim3 dimBlock(dim, dim);
dim3 dimGrid(N / dim,  N / dim);

// Kernel code
__global__ void kernel(float* A, float* B, float* C, int N)
{
 __shared__  float A_s[dim][dim];
 __shared__ float B_s[dim][dim];

int tx = threadIdx.x;
int ty = threadIdx.y.

int Row = blockIdx.y * dim + threadIdx.x;
int Col = blockIdx.x * dim + threadIdx.y;
float result = 0;

 for (int m = 0; m < N/dim; ++m)
 {

 A_s[ty][tx] = A[Row*N + (m*dim + tx)];
 B_s[ty][tx] = B[Col + (m*dim + ty)*N];

 for (int k = 0; k < dim; ++k)
{
 result += A_s[ty][k] * B_s[k][tx];
__ Synchthreads();
}
 C[Row*N+Col] = result;
__ Synchthreads();

}
```

[4+6+6=16]

Indian Statistical Institute

M.Tech (CS) II

Information and Coding Theory

Mid Semester Examination

Maximum Marks: 70

Date: September 23, 2015.
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_1 and X_2 be identically distributed but not necessarily independent. Let

$$\rho = 1 - \frac{H(X_2|X_1)}{H(X_1)}.$$

   (a) Show that $\rho = \frac{I(X_1;X_2)}{H(X_1)}$.
   (b) Show that $0 \leq \rho < 1$.
   (c) When is $\rho = 0$?
   (d) When is $\rho = 1$?

(4+2+3+3 = 12)

2. Under what condition does $H(X|g(Y)) = H(X|Y)$ hold? (6)

3. Consider a sequence of n binary random variables $X_1, X_2, \ldots, X_n$. Each n-sequence with an even number of 1's has probability $2^{-(n-1)}$ and each n-sequence with an odd number of 1's has probability 0. Find the mutual informations

$$I(X_1; X_2), I(X_2; X_3|X_1), \ldots, I(X_{n-1}; X_n|X_1, \ldots, X_{n-2}).$$

(6)

4. Let $X_1, X_2, \ldots$ be independent, identically distributed random variables drawn according to the probability mass function $p(x), x \in \{1, 2, \ldots, m\}$. Thus, $p(x_1, x_2, \ldots, x_n) = \Pi_{i=1}^{n} p(x_i)$. We know that $\frac{1}{n} \log p(X_1, X_2, \ldots, X_n) \to H(X)$ in probability.
   Let $q(x_1, x_2, \ldots, x_n) = \Pi_{i=1}^{n} q(x_i)$, where q is another probability mass function on $\{1, 2, \ldots, m\}$.
   (a) Evaluate $lim \quad \frac{1}{n} \log q(X_1, X_2, \ldots, X_n)$, where $X_1, X_2, \ldots$ are i.i.d. $\sim p(x)$.
   (b) Evaluate the limit of the log likelihood ratio $\frac{1}{n} \log \frac{q(X_1,\ldots,X_n)}{p(X_1,\ldots,X_n)}$ when $X_1, X_2, \ldots$ are i.i.d. $\sim p(x)$.

(6 + 6 = 12)

5. Find the entropy rate of the Markov chain associated with a random walk of a (i) Bishop (ii) Rook on a 8×8 chessboard. (6+6=12)

6. Give an interpretation of $D(p||q)$. Prove it. (3+7=10)

7. Find the (a) binary (b) ternary Huffman codes for the random variable X with probabilities

$$p = \left(\frac{1}{21}, \frac{2}{21}, \frac{3}{21}, \frac{4}{21}, \frac{5}{21}, \frac{6}{21}\right).$$

(6 + 6 = 12)

INDIAN STATISTICAL INSTITUTE

Mid-Semester Examination: 2015-2016

M. Tech. (CS) 2nd Year

Artificial Intelligence

Date: 23.09.2015 Maximum Marks: 40 Duration: 2 hours

Answer all questions in brief.

1. 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 without regard for the position of the empty cell. Define the problem as a state space graph problem and find a sequence of moves that will transform the initial configuration to a goal configuration. What is the cost of the solution? [2 + 6 = 8]

2. Solve the following cryptarithmetic problem:

```
  B A S E
+ B A L L
----------
G A M E S
----------
```

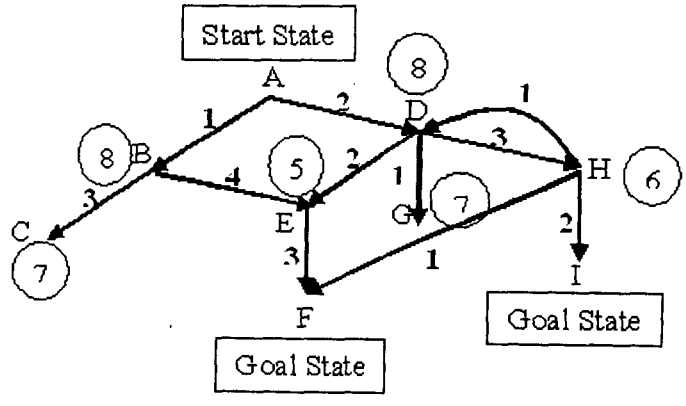
[8]

3. 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. Prove that any monotonic heuristic is admissible. [(3 + 2 + 1) + 2 = 8]

4. 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. [4 + 4 = 8]

P.T.O

5. 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). [4 + 4 = 8]



INDIAN STATISTICAL INSTITUTE
Mid-Semester Examination
M. Tech. (CS) II year (1st Sem): 2015–2016
Quantum Information Processing and Quantum Computation

Date: 24. 09. 2015 Maximum Marks : 50 Time : 2.5 Hours

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

1. (a) If an atom with a magnetic moment $\vec{m}$ enters a magnetic field $\vec{B}$, it experiences a force

$$\vec{F} = \vec{\nabla}(\vec{m} \cdot \vec{B}),$$

where

$$\vec{\nabla} = \frac{\partial}{\partial x}\hat{i} + \frac{\partial}{\partial y}\hat{j} + \frac{\partial}{\partial z}\hat{k}.$$

From the above formula, explain the results of the Stern-Gerlach experiment.

(b) What happens when we place a photon counter behind one of the two slits in Young's double-slit experiment?

[6 + 4]

2. (a) Prove that the eigenvalues of a Hermitian operator are all real.

(b) If A is Hermitian, show that e^{iA} is unitary.

[5 + 5]

3. (a) Starting from the Schrödinger equation

$$i\hbar\frac{d}{dt}|\psi(t)\rangle = H|\psi(t)\rangle,$$

show that the time-evolution operator is unitary, under the assumption that the Hamiltonian is time-independent.

(b) How is the above mechanism used in designing quantum gates?

[7 + 3]

4. (a) What is quantum entanglement?

(b) Prove formally that $\frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)$ cannot be written as a tensor product of two separable states.

(c) If A does local measurement on an entangled state shared with B, then the post-measurement state of B is immediately determined (assuming the same basis of projective measurements), even without the actual measurement. Does this violate the special theory of relativity, by allowing faster-than-light travel?

[2 + 4 + 4]

5. (a) Briefly describe the mathematical formulation of quantum teleportation.

(b) Does quantum teleportation violate the no-cloning theorem?

[6 + 4]

6. (a) Why is super-dense called the inverse of quantum teleportation?

(b) Name a universal quantum gate and derive the corresponding unitary matrix from the functional definition of the gate.

[4 + 6]

Indian Statistical Institute
Semister-I 2015-2016
M.Tech.(CS) - Second Year
Mid-term Examination (24 September, 2015)
Subject: Patter Recognition and Image Processing

Total marks: 80 **Maximum marks:** 60 **Duration:** 2 Hrs

Answer maximum of 60 marks (any part of any question). Simple calculators can be used. Answers should be precise.

1. (a) Describe the following classifiers with their advantages and disadvantages.
   (i) Bayes' classifier. (**5 Marks**)
   (iii) K-Nearest Neighbour (KNN) classifier. (**5 Marks**)
   (b) Derive the equations of decision boundaries for the minimum distance classifier with the following dataset. (**10 Marks**)

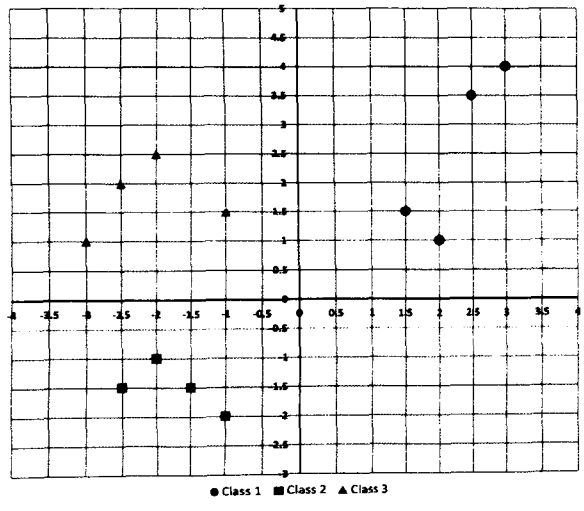


Figure 1: Scatter plot of training patterns for three classes.

2. (a) Define the Sum-of-Squared-Error clustering criterion to cluster n samples into k clusters; and interpret the criterion. (**5 Marks**)
   (b) Explain the limitations of $k-$means and $k-$medoids clustering algorithms. (**6 Marks**)
   (c) Consider the following proximity matrix $P = [p_{ij}]$ with Euclidean distance.

$$P = \begin{bmatrix} 0 & 2 & 3 & 7 & 8 \\ 2 & 0 & 1 & 4 & 5 \\ 3 & 1 & 0 & 4 & 5 \\ 7 & 4 & 4 & 0 & 1 \\ 8 & 5 & 5 & 1 & 0 \end{bmatrix};$$

where p_{ij} is the proximity or distance between i^{th} and j^{th} patterns. Find the clusters using

P·T·O

(i) Single Linkage (**3 Marks**)

(ii) Complete Linkage (**3 Marks**)

(iii) Average Linkage (Group Average) (**3 Marks**)

3. (a) Define variance-covariance matrix and confusion matrix (multi-class). (**6 Marks**)

   (b) Discuss on a class separability measure. (**6 Marks**)

   (c) Find the first and second principal component vectors for the data $\{(1,0),(0,1),(-1,0),(0,-1),(2,0),(0,2),(-2,0),(0,-2)\}$. (**8 Marks**)

4. (a) Discuss filter and wrapper approaches for feature selection with their advantages and disadvantages. (**6 Marks**)

   (b) Consider a dataset in which every pattern is represented by a set of 12 features. The goal is to identify a subset of 4 features or less which gives the best performance on this dataset. Find out how many feature subsets would be considered (i.e., the number of times the criterion function will be invoked) by each of the following feature selection algorithms before identifying a solution.

   (i) Exhaustive search. (**3 Marks**)

   (ii) Sequential Backward Selection (SBS) search. (**3 Marks**)

   (c) Discuss branch and bound algorithm for feature selection with advantages and disadvantages. (**8 Marks**)

Indian Statistical Institute
Mid-Semestral Examination: 2015
Course Name: M. Tech. in Computer Science
Subject Name: Mobile Computing

Date: 26-09-2015 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. Classify the vertical handover decision (VHD) algorithms based on the main handover decision criteria used. Define handover failure probability. Describe a signal to interference plus noise ratio (SINR) based VHD algorithm. [3+2+5=10]

2. Discuss the advantages and disadvantages of RSS based handover decision strategies. Consider a situation where an MS is moving from BS_1 to BS_2. The mean signal strength received at the MS from BS_1 and BS_2 is shown in Figure 1.

   (a) Suppose the MS uses **relative signal strength with threshold** handoff strategy. At what point handoff will occur if the threshold used by the MS is i) T_1, ii) T_2, and iii) T_3? Justify your answer.

   (b) Suppose the MS uses **relative signal strength with hysteresis and threshold** handoff strategy. At what point handoff will occur if the threshold and hysteresis margin used by the MS are T_3 and h respectively? Justify your answer. [2+(2+2+2)+2=10]

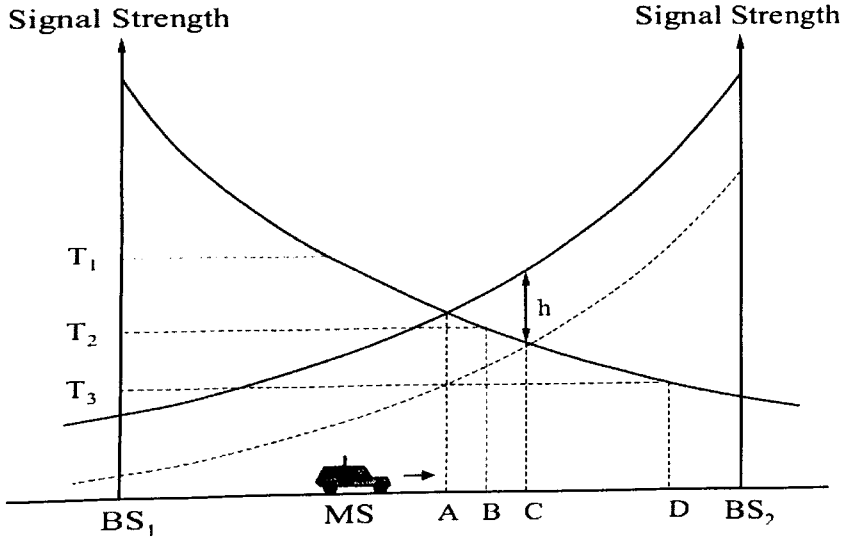


Figure 1: Signal strength and hysteresis between two adjacent BSs for potential handoff.

3. Formulate the perturbation-minimizing frequency assignment problem (PMFAP) as a mathematical optimization problem. Explain with an example the forced assignment with rearrangement (FAR) operation used in PMFAP. [5+5=10]

4. Briefly explain the channel assignment problem (CAP) in cellular network. Consider the 21-node cellular graph as shown in Figure 2. The demand vector $W = (w_i)$ is shown in Figure 2 where the label $[x]$ associated with a node represents the demand of that node. The frequency separation matrix $C = (c_{ij})$ is given as: $c_{ij} = 2$ if nodes i and j are adjacent, $c_{ij} = 1$ if nodes

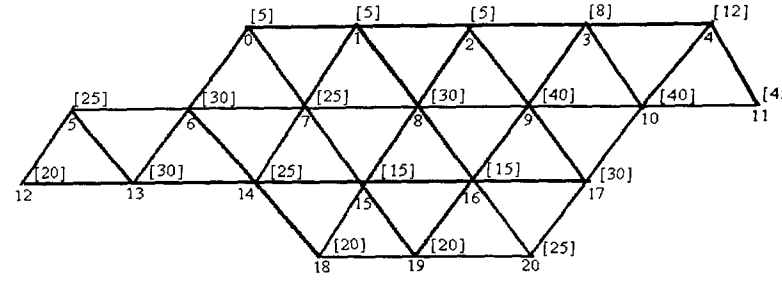


Figure 2: The demand vector.

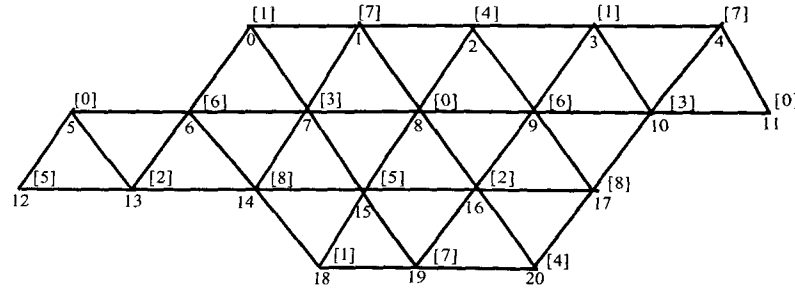


Figure 3: The homogeneous single-channel assignment.

assignment of the cellular graph as shown in Figure 3 where the label $[y]$ associated wi a node indicates the frequency assigned to that node. Construct the `coalesced CAP` usi the given single-channel assignment, demand vector and the frequency separation matr Suppose, given the bandwidth, you solved the coalesced CAP by applying some algorith Suppose your solution produces zero call blocking. How can you derive the solution of t original CAP from the solution of the coalesced CAP for this case of zero call blocking?

[4+8+8=

5. State the `STA` to `AP` association problem in `WLAN`. State the limitations of random polli access method. How proportional fair access method overcomes these limitations? When association is said to be max-min fair? Discuss about the possible throughput improveme that may be achieved by considering the frequency assignment and association jointly o considering them in succession.

[3+2+2+3+5=

INDIAN STATISTICAL INSTITUTE

Mid-SemestralExamination : 2015 – 16

MTech CS (2nd Year)

Computational Finance

Date: 28 September 2015 Maximum Marks: 30 Duration: 2Hours

This paper carries 34 marks. Attempt all questions. The maximum you can score is 30.

1. Critically explain the concepts: [3 X 2 = 6]
   a) Mutual Fund Principle
   b) Law of One Price

2. Assuming $V_0 > 0$, the discounted return is
   $$R_n^* = [S_n^*(1) - S_n^*(0)]/S_n^*(0) \text{ for } n = 1,\ldots,N$$
   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,…,N.
   (c) Q is a risk neutral probability if and only if $E_Q[R_n^*] = 0$ for n = 1,…,N.
   [4 + 3 + 5 = 12]

3. In the two period model, explicitly solve the Consumption Investment problem for the utility function $u(w) = \ln w$. Compute the relevant expressions and solve for the optimal trading strategy when N = 1, K = 2, r = 1/9, $S_0 = 5$, $S_1(\omega_1) = 20/3$, $S_1(\omega_2) = 40/9$ and $P(\omega_1) = 3/5$. [10]

4. Prove the Put – Call parity for a European option for the multi-period market. [6]

INDIAN STATISTICAL INSTITUTE
Mid Semestral Examination: 2015-16
M. Tech. (CS) II Year
Cognitive Science

Date: 28.09.2015 **Maximum Marks:** **Duration: 1 $\frac{1}{2}$ Hours**

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

1 Explain the concept of absolute threshold and describe the various methods of measuring it by examples. **(6)**

2 Write briefly about the 3 levels of David Marr's information processing task with respect to vision **(3)**

3 What are bottom-up processing and top-down processing? **(2)**

4 At night, we can see big things better than small things. Why? **(3)**

5 Explain the concept of pitch with respect to sound **[2]**

6 Does prolonged exposure to a sub-way (metro/underground) t5rain at 20 feet distance cause hearig loss? **[1]**

7 What is the difference between conduction he4aring loss and nerve hearing loss? **[2]**

8 Describe4 an experiment to investigate whether heat and cold are separate sensations, medicated through separate channels. **[4]**

9 Describe two ways by which depth can be estimated monocularly (using one eye). **[2]**

10 How are 'greebles' like faces?

Pick the correct anser. There is only ONE correct answer per question. They are for one mark each.

1 Which of the following methods is not used in measuring difference threshold?

a. Method constant stimuli
b. Method of limits
c. Method of adjustment

2. Which of the following brain regions are associated more with science perception than object perception?

a. FFA
b. PPA
c. LOC

****

Indian Statistical Institute

Mid-Semester Examination (2015-2016)
M.Tech. (CS) II
Advanced Algorithms for Graph and Combinatorial Optimization Problems

Date: September 28, 2015 Maximum Marks: 60 Time: 2.5 hours

There are seven questions. Answer as many as you can.
Marks allotted to each question are indicated within parentheses near the right margin.

1. You are asked to design a survey for a set of n customers of a company, which sells k types of products. The survey must satisfy the following:
   - a customer i can be asked questions only about the product types he has bought, and the number of questions must be in the range c_i to c_i';
   - for each product type j, there must be between p_j to p_j' distinct customers who are asked about it.

   (i) Design an algorithm to determine whether a feasible survey can be designed. (9)
   (ii) Prove its correctness. (4)
   (iii) Analyze the worst case time complexity of your algorithm. (4)

2. In Orlin's algorithm for finding the maximum flow in a given flow network,
   (i) define a critical arc; (4)
   (ii) derive the number of critical nodes over all capacity scaling phases. (4)

3. For a non-bipartite graph G, let b be a blossom in an augmenting path form its node u with respect to a matching M. Prove that if there is an augmenting path from u in G/b (the graph obtained from G by contracting its blossom b), then there is an augmenting in G from u with respect to M. (5)

4. Determine a smallest (with respect to number of vertices) imperfect graph G such that $\chi(G) = \omega(G)$. (10)

5. A caterpillar tree C is a tree in which all the vertices are adjacent to a fixed path P of C.
   (*i*) Prove that a tree is an interval graph if and only if it is a caterpillar tree. (5)
   (*ii*) Formulate an efficient algorithm to recognize whether a given tree is an interval graph. (5)

6. Let $G = (V, E)$ be a graph. The total graph $T(G) = (V^t, E^t)$ of G is the graph with the set of vertices $V^t = V \cup E$ and the set of edges $E^t = \{uv | u \text{ and } v \text{ are incident in } G\}$.

   (*i*) Draw the total graph of K_4. (3)

   (*i*) Prove that G is chordal only if $T(G)$ is chordal. (4)

   (*ii*) Is the converse true? Justify your answer. (3)

7. Let $G = (V, E)$ be a split graph.

   (*i*) Give necessary and sufficient conditions for which G is a tree. (2)

   (*ii*) Let V be the disjoint union of a clique A and an independent set B. Prove that if G is isomorphic to its complement $\bar{G}$, then $|A| = |B|$. (5)

   (*iii*) Let $\Sigma = [a_1, a_2, ..., a_n]$ be a degree sequence. Define $\bar{a_i} = n - 1 - a_{n-i+1}$. Show that Σ is graphic if and only if $\bar{\Sigma} = [\bar{a_1}, \bar{a_2}, ..., \bar{a_n}]$ is graphic. (3)

INDIAN STATISTICAL INSTITUTE
Mid-Semester Examination
M. Tech. (CS) II year (1st Sem): 2015–2016
Cryptology

Date: 29. 09. 2015 Maximum Marks : 50 Time : 2.5 Hours

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

1. (a) Derive the size of the keyspace of Hill cipher with block-length n bits, when the modulus is a prime p.
   (b) What is the rationale behind Kerckhoff's principle?

   [6 + 4]

2. (a) What is the adversarial indistinguishability game?
   (b) Prove that an encryption scheme is perfectly secure, if and only if for every adversary A, the probability of A's winning the above game is 1/2.

   [2 + 8]

3. (a) When is an encryption scheme called computationally secure?
   (b) Give formal definitions of PRP, PRF and PRG.

   [4 + 6]

4. (a) Assume that a block cipher $E_k()$ with key length n bits has no better attack than exhaustive key search with complexity 2^n. Does the cascaded construction $E_{k_2}(E_{k_1}())$, where k_1, k_2 are two independent n-bit keys, provide a security with key search complexity 2^{2n}?
   (b) What is the assumption between the relationship of the random variables in the piling-up lemma?
   (c) Why in TDES the second operation is DES^{-1} instead of DES?

   [4 + 2 + 2]

5. (a) In AES mix-column step, two polynomials are multiplied modulo a polynomial $g(x)$ of degree 4. Do we require $g(x)$ to be irreducible? If so, why?
   (b) How would you test if a polynomial of degree n is primitive or not? What is the time complexity of your algorithm?

   [4 + 6]

6. (a) Can a PRF be used as a MAC? Can a MAC be used as a PRF? Justify (no formal proof required).
   (b) Prove that if a hash function is collision resistant, then it is also second pre-image resistant and pre-image resistant.

   [4 + 6]

Data Mining
Mid Sem Examination

Time: 2 hours and 30 minutes

30 September 2015

For questions 1 and 2, a correct answer would get 3 marks, a wrong answer would get -1 and no answer would get zero. In addition to choosing the correct alternative, you must provide an explanation. A correct choice without any explanation would be considered as no answer. A correct choice with an entirely wrong explanation would be considered as a wrong answer. All other questions carry 5 marks each. The maximum you can score is 30.

1. Suppose for every element in a data stream, you have to perform a lookup to decide if the element is present in a pre-defined set S of elements. Then, size of the bit array you would use is

   (A) Exactly same as $|S|$

   (B) Approximately same as $|S|$

   (C) Significantly more than $|S|$

   (D) Significantly less than $|S|$

2. To classify news articles and tweets into a set of pre-defined categories, which models of Naive Bayes classification would be more appropriate?

   (A) Multinomial model for both

   (B) Bernoulli's model for both

   (C) Bernoulli's model for news articles and multinomial model for tweets

   (D) Multinomial model for news articles and Bernoulli's model for tweets

3. Suppose the elements of an $m \times n$ shingles - document matrix are residing in a distributed file system. A single processor can access the data element (cell) wise, in the form of triplets (i, j, s_{ij}), where i is the shingle id, j is the document id and $s_{ij} \in \{0, 1\}$ indicates whether or not the i-th shingle is present in the j-th document. However, for the same document or same shingle, all entries may not be available together in any single processor. Suppose a hash function $h : \mathbb{N} \to \mathbb{N}$, which is to be used (instead of the more expensive random permutation) to create the min-hash signatures of the n documents, is given. Using MapReduce, how would you create the min-hash values for the documents? Write the *map* and the *reduce* functions and briefly explain why your algorithm would work.

4. What is the expected number n_t of triangles in a random undirected graph with n nodes and m edges, for sufficiently large n and m? In a social network graph of the same size, if it is known that the number of triangles is three times that of the expected average n_t, what can you say about the average number of mutual friends between any two friends?

5. Suppose a training set has n (where $n > 100$) data points and the accuracy of a classifier is measured by k-fold cross validation for some $k > 5$. For a fixed dataset and fixed classifier, will this experiment produce the same accuracy if it is performed several times? If your answer is yes, explain why. If your answer is no, then describe one sufficient condition (with respect to the nature of the data, k and n) under which the experiment will produce the same accuracy no matter how many times it is performed. If no such condition is possible, justify why.

   Recall k-fold cross validation: First the dataset is divided into k random equal partitions. Then each k partition is used as a test set while the classifier is trained using the remaining $k-1$ partitions. The average accuracy over these k iterations is considered as the overall accuracy.

6. Suppose an $m \times n$ term-document matrix $A = (a_{ij})$ is given where the m rows correspond to m terms $(t_1, \ldots, t_m)$, the n columns correspond to n documents $(d_1, \ldots, d_n)$ and a_{ij} is the TF.iDF score of the term t_i in document d_j. The documents are column vectors in the m-dimensional term space. Given any query, it is also considered to be a vector in the same m-dimensional space.

   A search system needs to rank the documents by cosine similarity of the query and the document vectors. However, given any query, it would be very inefficient to compute its similarity with all the n documents, so it would be better to use an inverted index. Describe what would you store for each term in the inverted index. Also describe how would you process a query to compute the similarities and rank the documents.

   Hint: Note that the cosine similarity between a document d and a three term query (t_1, t_2, t_3) would not be the sum of the scores of d in the rows of t_1, t_2 and t_3, so simply storing the rows of the matrix A as lists in the inverted index will not suffice.

7. Consider the dataset illustrated as a scatterplot in Figure 1, where there are two attributes (X and Y) for each data point and there are two classes, namely A and B. Although the number of points drawn in the figure is below 100, assume that there are at least 1000 data points and the figure is just an illustration of how the points are distributed. Given such a training data, what would be a preferred classifier between decision tree and linear discriminant analysis (LDA)? Explain clearly, if necessary using the figure and assuming any reasonable scale for the X and Y axes.

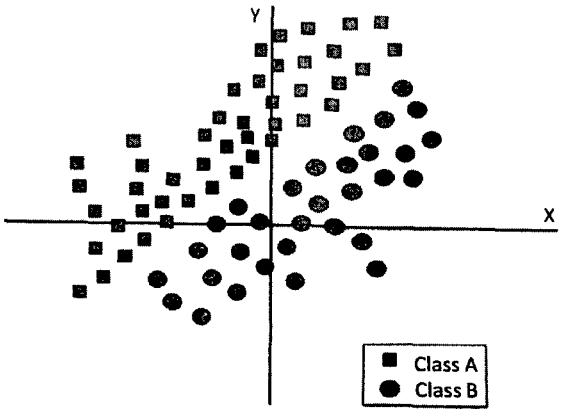


Figure 1: Dataset shown as scatterplot

Indian Statistical Institute

Mid-Semester Examination : 2015 – 2016

Master of Technology in Computer Science, Semester III

Functional Brain Signal Processing: EEG & fMRI

Date: 01/09/15 Maximum Marks: 100 Duration: 3 hours

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

1. Each cortical source of the scalp EEG is contributing to all the channels and each channel is contributed by all the sources – explain. What is dipole source model? How many parameters do you need to describe a dipole and why? 3 + 4 + 1 + 2 = 10

2. What is impulse response system? What is system function? An impulse response filter is completely described by its system function – explain. 2 + 2 + 6 = 10

3. What is low-pass filter and what is high-pass filter? How would you combine these two filters in order to eliminate 50 Hz line noise in a recorded EEG signal? How would you make a band-pass filter with the help of a low-pass and a high-pass filter? 2 + 2 + 4 + 2 = 10

4. Indentify two possible reasons for each of the following observations:
   (a) On a data set PCA works nicely, but ICA does not.
   (b) On another data set ICA works nicely but PCA does not. 2x(2.5) + 2x(2.5) = 10

5. Let $x(n) = \sum_{n=-\infty}^{\infty} a_n \cos(2\pi nt) + b_n \sin(2\pi nt), \quad -\infty < t < \infty$ be a signal. Determine a_n and b_n for any n (mathematical derivation required). 5 + 5 = 10

6. Write short notes on the following:
   (i) 10-20 EEG electrode placement system
   (ii) P300
   (iii) k-means clustering
   (iv) Power spectrum
   4 x 5 = 20

7. Carefully observe Fig 1 below. It is a real human EEG signal. The background and foreground signals have been shown in the figure. Design an algorithm by which the background signal can be automatically separated out from the foreground signal by an electronic computer. Present an analysis of the algorithm for its (a) time complexity and (b) space complexity (that is, the maximum RAM requirement). Marks will depend on the correctness and time and memory efficiency of the algorithm. You can describe the

algorithm in plain English in a step by step manner with justification for each step wherever necessary. 20 + 5 + 5 = 30

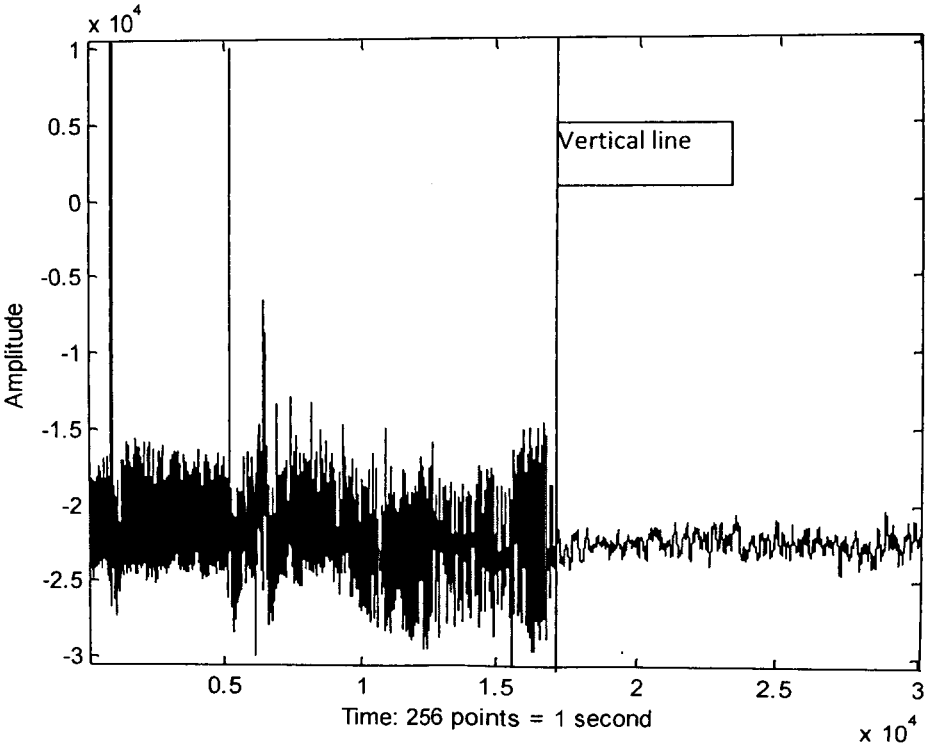


Fig 1. A human EEG signal. The signal on the left side of the vertical line is the foreground and that on the right is the background.

INDIAN STATISTICAL INSTITUTE

Academic Session: 2015-16

MID-TERM EXAMINATION (Date of Exam: 01.10.2015)

Course name: M. Tech. (CS), Second Year

Subject: Natural Language Processing

Time: 2½ hours **Full Marks: 58** **Maximum Scoring Marks: 50**

Note: Attempt all questions

---

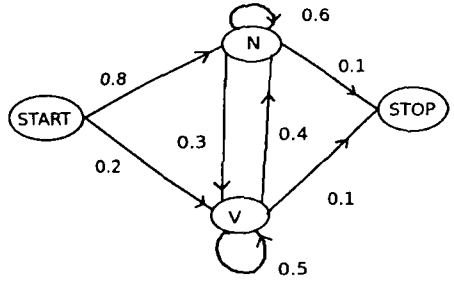
1. Consider the bi-gram HMM shown below. There are two tags N and V which can emit words. Two special tags START and STOP are also there which are incapable of emitting words. Let there are only three symbols `I', `YOU' and `HE'. The symbol emission probabilities are given below.

p(`I'|N) = 0.2, p(`You'|N) = 0.4 and p(`He'|N) = 0.4

p(`I'|V) = 0.5, p(`You'|V) = 0.4 and p(`He'|V) = 0.1

Given these initial parameters and an observation "HE I HE", estimate all symbol emission parameters after the first iteration of Baum-Welch algorithm.

[Marks: 15]



2. Prove that under uniform distribution over all words in the vocabulary, the perplexity of a language model is equal to the vocabulary size.

[Marks: 5]

3. There is a real-time Android keyboard application that can predict the next word (three most likely candidates) from a given input text sequence. Instead of taking a word from the list three candidates, if a user types an incomplete word, the application tries to predict a word close to the user input (i.e. the incomplete word). The application also supports real-time incremental learning i.e. the system learns from the users' input.

Assume that you are asked to develop this application. What would be your algorithms (in pseudo code) if the incremental learning mechanism is (a) **not** considered and (b) **considered** as a feature of the application

In each case, your resource requirements should be stated clearly and you should also mention how the resource requirements can be easily handled in practice. Steps of your algorithm should be properly explained in the form of pseudo code.

[Marks: 10 + 10 = 20]

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

das haus – the house

das buch – the book

ein buch – a book

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

[Marks: 10]

(b) Now, assume that there is a prior knowledge that the German word "ein" only corresponds to English word "a". What would be your method to incorporate this knowledge into IBM model 1?

[Marks: 8]

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

30/11/2015

(Answer all questions)

1. Generate a feature based tree classifier for the recognition of the following printed alph characters (R, B, I, J, 2, 7, e,c, 6, 9, E, F).

2. What is a multi-script document? Discuss about the possible OCR technologies document. Write a rotation and scale invariant feature extraction method for mult character recognition.

3. What are the differences between holistic recognition and segmentation-based recognit do you mean by real-word errors and non-word errors? How 'split-character errors' di 'run-on-errors'?

4. What is Word-spotting? Describe a suitable method for line segmentation from a h document image.

5. What is the tilt of a word image? How tilt differs from skew? Describe a Hough trans approach for skew detection of a Roman document.

6. Write a method to separate graphics and text portions from a document image.

7. Given a set of symbols and their probabilities of occurrence, describe a method to g Shannon-Fano codes? (Explain with an example). Under what condition on symbol p does this code produce highest compression? How efficient is this code compared coding?

8. (a) Using an example describe the approach of generating Huffman codes and er positive properties. When is a small variance Huffman-code preferable and why?
(b) Given eight symbols A, B, C, D, E, F, G and H with probabilities 1/30, 1/30, 1/30 5/30, 5/30, and 12/30, respectively. Draw any three different Huffman trees with h height 6 each (total 6 trees) and calculate the average code size for [6+6]

9. Describe the basics of One-dimensional Facsimile coding scheme including docum resolution, run-length coding of a line with termination and make-up codes, EOI beginning and end codes, fill bits etc. (Code table is not needed).

Indian Statistical Institute

End-Semester Examination (2015-2016)

M.Tech. (CS) II

Advanced Algorithms for Graph and Combinatorial Optimization Problems

Date: 30.11.2015 Maximum Marks: 100 Time: 3 hours

Answer as much as you can. The maximum you can score is 100 marks. Marks allotted to each question are indicated within square brackets near the right margin. This paper has two pages.

1. Let $G = (V_G, E_G)$ and $H = (V_H, E_H)$ be two graphs. The strong product $G \boxtimes H$ is the graph with vertex set $V_{G \boxtimes H} = \{(u, v) | u \in V_G \text{ and } v \in V_H\}$. Two vertices (u, v) and (u', v') of $G \boxtimes H$ are adjacent if either of the following holds:

   (i) $u = u'$ and $vv' \in E_H$,

   (ii) $uu' \in E_G$ and $v = v'$,

   (iii) $uu' \in E_G$ and $vv' \in E_H$.

   Prove the following: (χ, ω, α, κ being the chromatic number, the clique number, the independence number and the clique cover number respectively)

   (a) $\chi(G \boxtimes H) \geq max\{\chi(G), \chi(H)\}$,

   (b) $\omega(G \boxtimes H) = \omega(G)\omega(H)$,

   (c) $\alpha(G \boxtimes H) \geq \alpha(G)\alpha(H)$,

   (d) $\kappa(G \boxtimes H) \leq \kappa(G)\kappa(H)$. [4 + 5 + 2 + 4 = 15]

2. A (0, 1)-valued matrix M is said to have the circular 1's property for columns if its rows can be permuted in such a way that the 1's in each column occur in a circular consecutive order (i.e., similar to being wrapped around a cylinder).

   Design an algorithm to test whether a given $m \times n$ (0,1)-valued matrix M with f non-zero entries has the circular 1's property. [9+6=15]

   Analyze the worst case time complexity of your algorithm.

3. (a) Prove that the complement of the 3-dimensional cube Q_3 is nonplanar.

   (b) Trace the Hopcroft and Tarjan's path addition algorithm for planarity testing on the Petersen graph. Show the DFS numbering of the vertices and their values for *lowpoint*1 and *lowpoint*2. [5+10=15]

4. (a) Present a randomized algorithm to obtain a global min-cut for a given undirected graph.

   (b) Analyse the probability of success and hence the worst case time complexity.

[3+5=8]

5. (a) Give an approximation algorithm for the set cover problem, Derive the performance guarantee of your algorithm. What is its worst case time complexity?

   (b) Formulate the vertex cover problem in terms of the set cover problem.

[(5+3+3)+4=15]

6. Let $G = (V, E)$ be a complete undirected graph with edge costs satisfying the triangle inequality, and let k be a positive integer. The problem is to partition V into sets V_1, V_2, ..., V_k so that the costliest edge between two vertices in the same set is minimized, i.e., minimize $\max_{1 \leq i \leq k, u,v \in V_i} cost(u, v)$. Design a 2-factor approximation algorithm for this problem, together with a tight example. [6+(4+2)=12]

7. Consider the following algorithms for the knapsack problem where the size of each object is at most B, the bound on the size of the knapsack. Sort the objects by decreasing ratio of value to size. Let this order be $a_1, a_2, \ldots, a_n$.

   (a) Show that by greedily picking the objects in that order, the performance of the algorithm can be arbitrarily bad.

   (b) Instead, from the sorted list, first find the lowest k such that the size of the first k objects exceeds B. Now, among the two sets $\{a_1, a_2, \ldots, a_{k-1}\}$ and $\{a_k\}$, pick the one which has more value. Show that this algorithm achieves an approximation factor of 2.

[4+6=10]

8. For the bin packing problem,

   (a) illustrate an instance for which the First-Fit decreasing algorithm gives $\frac{11}{9} \cdot OPT$;

   (b) prove that for any $\epsilon > 0$, there is no approximation algorithm having a guaranatee of $\frac{3}{2} - \epsilon$ for this problem, assuming $\mathbb{P} \neq \mathbb{NP}$;

   (c) give its ILP formulation and LP relaxation.

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

9. (a) Present an exact LP-relaxation for the maximum matching problem for a given bipartite graph.

   (b) Give an example of a chordal graph which is not an inerval graph.

[5+4=9]

INDIAN STATISTICAL INSTITUTE

First-Semester Examination: 2015-2016

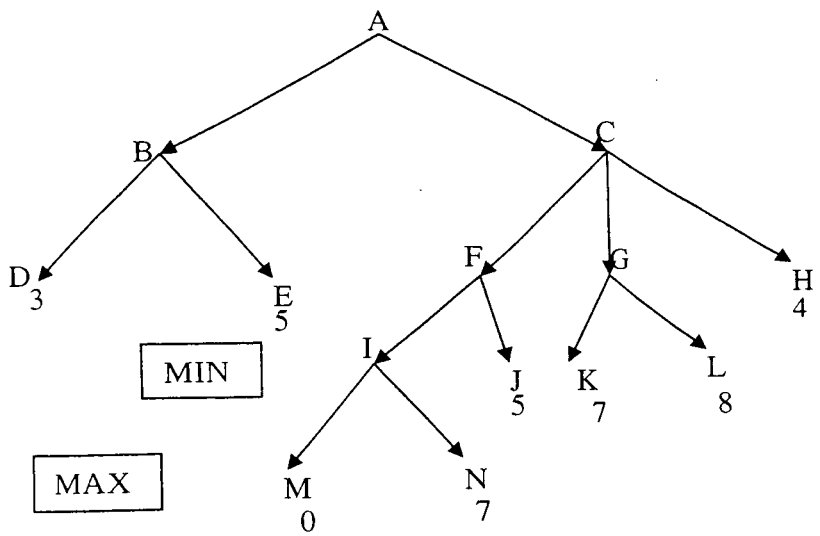
M. Tech. (CS) 2nd Year

Artificial Intelligence

Date: 02·12·2015 Maximum Marks: 100 Duration: 3 hours

Answer all questions.

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



(ii) Define the constraint satisfaction problem and solve the following cryptarithmetic problem:

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

[(3 + (3 + 3) + 1) + (2 + 8) = 20]

2. (i) Consider the following planning problem. The monkey-and-bananas problem is faced by a monkey in a laboratory with bananas hanging out of reach from the ceiling. A box is available that will enable the monkey to reach the bananas if he climbs on it. Initially the monkey is at A, the bananas at B, and the box at C. The monkey and box have height Low, but if the monkey climbs onto the box, he will have height High, the same as the bananas. The actions available to the monkey include Go from one place to another, Push an object from one place to another, ClimbUp onto or ClimbDown from an object, and Grasp or Ungrasp an object. Grasping results in holding the object if the monkey and object are in the same place at the same height.

- Express the initial state description.
- Express the six actions in STRIPS-style definitions.
- Suppose the monkey wants to fool the scientists, who are off to tea, by grabbing the bananas, but leaving the box in its original place. Write this as a general goal (i.e., not

P·T·O

assuming that the box is necessarily at C) in the language of situation calculus.

(ii) In *farmer-fox-goose-grain* puzzle, a farmer wishes to cross a river taking his fox, goose, and grain with him. He can use a boat which will accommodate only the farmer and one possession. If the fox is left alone with the goose, the goose will be eaten. If the goose is left alone with the grain it will be eaten. Draw a state space search tree for this puzzle. Denote left and right river banks as left-bank and right-bank, respectively. [(3 + 6 + 3) + 8 = 20]

3. Answer the following:
   a) 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.
   b) Prove using semantic tableaux approach that the following sentences are mutually consistent. "All Indian citizens who are adult have right to vote in election. Mary is an Indian citizen and has voting right. Mary is an adult."
   c) Prove that if α is a logical consequence of a set of premises $\sum$, then there is a tableau proof of α from $\sum$.
   d) 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. [6 + 4 + 5 + 5 = 20]

4. (i) 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.

   (ii) Show a formulation in propositional logic of the following problem: given a directed graph, does it contain a Hamiltonian cycle? [Note: A Hamiltonian cycle in a graph is a closed path that visits each node, other than the first, exactly once]. [5 + 5 + 10 = 20]

5. Answer the following:
   a) Write a program in Prolog for post-order traversal of a binary tree. The traversal method stores the elements of the tree in a list.
   b) Describe the Bayes' theorem for probabilistic reasoning.
   c) Let $I = \langle U, A \rangle$ be a decision table, where $U = \{x1, \ldots\ldots, x7\}$ is a nonempty set of finite objects, the universe, and $A = C \cup D$ is a nonempty finite set of attributes. Here, $C = \{A1, A2\}$ and D are the set of condition and decision attributes, respectively.

U	*A1*	*A2*	*D*
x1	16-30	50	yes
x2	16-30	0	no
x3	31-45	1-25	no
x4	31-45	1-25	yes
x5	46-60	26-49	no
x6	16-30	26-49	yes
x7	46-60	26-49	no

In the context of rough set theory, explain the following with the above example data:
i) lower and upper approximations of decision attribute,
ii) boundary region of decision attribute, and
iii) degree of dependency and significance of a condition attribute.

[5 + 5 + (2 + 2) + 2 + (2 + 2) = 20]

INDIAN STATISTICAL INSTITUTE

Semestral Examination : (2015 - 2016)

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

Subject Name : Neural Networks & Applications

Date : November 30, 2015 **Maximum Marks : 100** **Duration : 3 hrs**

Answer all the questions.

1. Show geometrically how a single node perceptron model of artificial neural networks classifies a set of two dimensional patterns distributed in linearly separable two classes. [25]

2. Consider a Multilayer Perceptron (MLP) with two hidden layers for classifying n-dimensional P patterns distributed in k classes. Assume that the numbers of nodes in 1st and 2nd hidden layers are l and m respectively. Derive expressions for amount of updation for the weights associated with the links in the MLP, considering batch mode learning. [25]

3. Consider Oja's model of Principal Component Analysis Network (PCAN). Derive how this model can be used to extract n principal components from a set of n-dimensional patterns. Establish the relations, if any, among the parameters associated with these principal components. Assume that the number of patterns is greater than n. [10+20 = 30]

4. Explain Hopfield model of artificial neural networks. Derive an expression for the estimate for capacity of storing patterns in the model. [20]

INDIAN STATISTICAL INSTITUTE
Semester Examination
M. Tech. (CS) II year (1st Sem): 2015–2016
Cryptology

Date: 01. 12. 2015 Maximum Marks : 40 Time : 2.5 Hours

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

1. (a) Explain how the period of an LFSR sequence is related to the nature of the connection polynomial.
   (b) What will be the problem if the state transition matrix of an LFSR is singular?

   [7 + 3]

2. (a) Define non-linearity of a Boolean function.
   (b) If I keep all n-variable Boolean functions inside a bag and randomly pick up one Boolean function from this bag, what is the probability that the function is non-linear? Justify.
   (c) Construct a 3-variable Boolean function that is not correlation-immune.

   [2 + 3 + 5]

3. (a) State and Prove Chinese Remainder Theorem.
   (b) Show that the subset-sum problem defined over a super-increasing knapsack is not NP-complete.

   [(2 + 4) + 4]

4. (a) Explain the common-exponent and the common-modulus attacks on the basic RSA scheme.
   (b) How can these attacks be avoided?

   [(4 + 4) + 2]

5. (a) Show a forgery attack on the basic RSA-based digital signature scheme.
   (b) How can this attack be prevented?
   (c) What do you mean by the security of a key exchange protocol?

   [4 + 2 + 4]

Indian Statistical Institute

M.Tech (CS) II

Information and Coding Theory

Semester Examination

Maximum Marks: 80

Date: December 3, 2015
Time 3 hours

The question paper contains 8 questions. Total marks is 90. Maximum you can score is 80. Unless otherwise mentioned, all notations are the same as presented in class

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

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

and $X \in \{0, 1, \ldots, 10\}$. 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)

2. 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 $\lambda^{(n)} \to 0$. (10)

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

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

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

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

5. (a) Construct a Hadamard matrix H_8 of order 8, starting from a Hadamard matrix of order 2.

   (b) Construct two Hadamard codes $\mathcal{C}_1$ and $\mathcal{C}_2$ from this matrix. What are the parameters of these codes?

   (c) Construct a BIBD using H_8. What are the parameters of the design?

   (d) Construct a Hadamard matrix of order 12. (8 + 8 + 6 + 8 = 30)

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

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

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

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

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

INDIAN STATISTICAL INSTITUTE

First Semester, 2015-2016

M.Tech (Computer Science)II-Year

Pattern Recognition and Image Processing

Date: 03.12.2015 Maximum Marks: 100 Time: 3 hours

Note: Answer a maximum of six questions.

1)

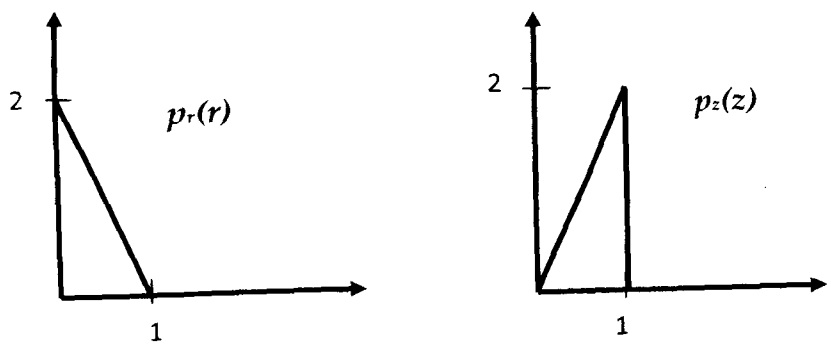
I. Define

a) 4-adjacency,

b) 8- adjacency and

c) m- adjacency between pixels of an image.

II. An image has the gray level PDF $p_r(r)$ shown in the following diagram. It is desired to transform the gray levels of this image to the specified $p_z(z)$ form shown. Assuming continuous variables, find the transformation (in terms of r and z) to accomplish this. (2x3+14=20 Marks)



I. Show that the application of a 3×3-sized local averaging mask can be replaced by an 1×3 and a 3×1 masks applied sequentially. Compare the amount of additions that are needed in both 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+(2+5)=15 Marks)

2)

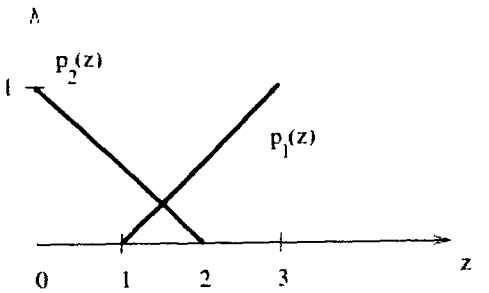
I. Find the Fourier transform of the sequence f(0)=0, f(1)=1, f(2)=0, f(3)=1. Then calculate the inverse Fourier transform and compare the result with the original sequence. Draw the Fourier spectrum.

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)

(3)

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)



(4)

I. Given a one-dimensional data set {-6, -1, 7, 12, 16, 22}, use the agglomerative clustering algorithm with average linkage (Euclidean distance) to establish a hierarchical grouping relationship. What are the clusters at each level?

II. Describe one similarity and one dissimilarity measures with respect to clustering.

III. Explain the importance of initialization of centres in K-means clustering using examples. (10+6+4=20 Marks)

(5)

I. Describe the method of Principal Component Analysis for dimensionality reduction with advantages and disadvantages.

II. Derive the Branch and Bound tree for D=6; d=3 for feature selection with the criterion function

$$J(\bar{x}) = \left(\sum_{\zeta_i \epsilon \bar{x}} i \right)^2 ;$$

where $\zeta_i \epsilon \bar{x}$ is the i^{th} feature in the feature set $\bar{x}$. (5+10=15 Marks)

(6)

I. Given a two-class discrimination problem in which the features follow a two-variate normal distribution with the following parameters. Find the Bayes' decision boundary and explain its significance. Assume equal prior probabilities for the classes.

$$\mu_1 = [1, -1]^T, \mu_2 = [3, 2]^T \text{ and } \Sigma_1 = I;\ \Sigma_2 = 2I.$$

μ_i and Σ_i denote the mean vector and covariance matrix of the i^{th} class, i=1, 2.

II. Discuss over-fitting vs under-fitting with respect to classification. Explain different cross validation techniques for assessing classification performance. (10+5+5=20 Marks)

INDIAN STATISTICAL INSTITUTE

Semestral Examination: (2015 – 2016)

M.Tech. (CS) II Year

Parallel Processing: Architectures and Algorithms

Date: 04/12/2015 Maximum Marks: 100 Duration: 3 hours

1. a) Given the choices i) 2D torus, ii) binary hypercube and iii) cube-connected cycle, for interconnecting 64-nodes of a multiprocessor system, select the best one based on the parameter W /(d.D.E), where d is the degree of any node, D is the diameter, W is the bisection width and E is the total number of edges of the network.

   b) Show an embedding of a 5×3 mesh on a 16-node hypercube with dilation 2.

   [12+6=18]

2. Describe an algorithm for finding minimum spanning tree of a weighted graph which will run on an EREW SM SIMD computer. Comment on the optimality of the above algorithm.

   [10+5=15]

3. 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 an interval (a, b). [15]

4. (a) Draw the block diagram of an 8×8 omega network using 2×2 switches.
   (b) Given a BPC permutation P: $x_2\, x_1\, x_0 \rightarrow x_1\, x_0\, x_2$, show the path matrix of P for 8×8 omega network. Hence draw the conflict graph and find out the minimum number of passes required to route P in omega network and the corresponding partitions of the paths.
   Is P admissible in the omega network with one additional stage? If yes, then show the corresponding path matrix.

   [4+(4+4+2+4)=18]

5. a) Describe Batcher's odd-even merge algorithm for merging two sorted sequences and prove its correctness.
   b) Draw a network for sorting 8 elements by Batcher's odd-even merge technique.
   c) Derive an expression for the time required for sorting n elements by Batcher's odd-even merge technique.

   [(4+3)+3+6=16]

6. a) Write a CUDA program to find transpose of a 1024×1024 matrix using shared memory.
   b) Mention briefly the role of tile size in determining the speed up in the above program.

   [14+4=18]

(The questions are in Page 2 of this document)

INDIAN STATISTICAL INSTITUTE

Final Examination : 2015-2016

Course Name: Cognitive Science

Subject Name : Computer Science, MTech CS II

Date : 5/12/15 Maximum Marks : 40 Duration: 100 minutes

Note, if any : Answer the following questions. Marks for each question are mentioned in bracket after each question.

Instructor: Garga Chatterjee, CVPR Unit

1. What are cognitive heuristics? (1)

2. What is correspondence bias? What is another name for it? (2+1)

3. Describe an experiment that demonstrates self-serving bias. (4)

4. What is correspondent inference theory with reference to attributional logic (mention all 4 conditions) ? (5)

5. In attributional logic, what is discounting principle and augmenting principle? Explain with examples. (4)

6. Describe the famous experiment (including result) on visual attention using Posner's flanker task (1984). (3)

7. Describe the Stroop color naming task and explain its result. (4)

8. In the multiple salience maps that have been proposed in brain-based attentional models, which parts of the cortex have a. Visual salience map b. Map for top-down task relevant information. (1+1)

9. What is binocular rivalry? Name atleast aspects of 3 aspects of visual input/image that based on which binocular rivalry can occru? (4)

10. Briefly describe the 5 stages of a typical sleep cycle. (5)

11. Should anyone be forced against their will to a 'lie-detection' polygraph test? From your ethical stand-point, argue your case. Write down your views. (2)

12. Why is it incorrect to infer that a subject is telling a lie simply from a polygraph test? (3)

INDIAN STATISTICAL INSTITUTE
Semester Examination
M. Tech. (CS) II year (1st Sem): 2015–2016
Quantum Information Processing and Quantum Computation

ate: 07. 12. 2015 Maximum Marks : 40 Time : 2.5 Hours

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

1. (a) What is measure-and-resend attack on BB84 protocol?
   (b) Derive an expression of success probability of determining the correct key by a measure-and-resend attacker.

   [4 + 6]

2. (a) What is the non-identity square-root of a one-qubit identity gate?
   (b) What is the difference between CNOT and CCNOT gates?
   (c) Design a swap gate using only CNOT gates.

   [3 + 4 + 3]

3. (a) Show that $\forall \mathbf{x} \in \{0,1\}^n$,

$$H^{\otimes n} |\mathbf{x}\rangle = \frac{1}{2^n} \sum_{\mathbf{y} \in \{0,1\}^n} (-1)^{\mathbf{x}\cdot\mathbf{y}} |\mathbf{y}\rangle .$$

   (b) What are the interpretations of different measurement outputs in Deutsch-Jozsa algorithm?

   [6 + 4]

4. (a) How is Grover's search problem different from the satisfiability problem?
   (b) Can the solution of one of the above two problems be used to solve the other?
   (c) What is the geometric interpretation of Grover's algorithm?

   [2 + 3 + 5]

5. (a) Show that factoring can be reduced to order-finding.
   (b) What are the implications of Shor's algorithm in the domain of cryptography and security against a quantum adversary?

   [6 + 4]

Indian Statistical Institute
Semestral Examination: 2015
M. Tech. in Computer Science
Second year
Mobile Computing

Date: 07/12/15 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) Describe how randomized rotation of cluster-heads is used to evenly distribute the energy load among the sensors in Low-Energy Adaptive Clustering Hierarchy (LEACH) protocol in wireless sensor networks. [10]

   (b) i. Describe the direct communication and Minimum Transmission Energy (MTE) routing protocols in wireless sensor networks. [10]

   ii. Consider a linear network consisting of n sensor nodes and the Base Station (BS) as shown in Figure 1 where the distance between two consecutive nodes is r. The distance between node n and the BS is also r. Circles denote the sensor nodes and the square denotes the BS.

   A. Derive expressions for total energy expended in the system for transmitting a k-bit message from node 1 to the BS using direct communication and MTE protocols. [10]

   B. Under what condition does the direct communication routing require less energy than the MTE routing in wireless sensor networks? [5]

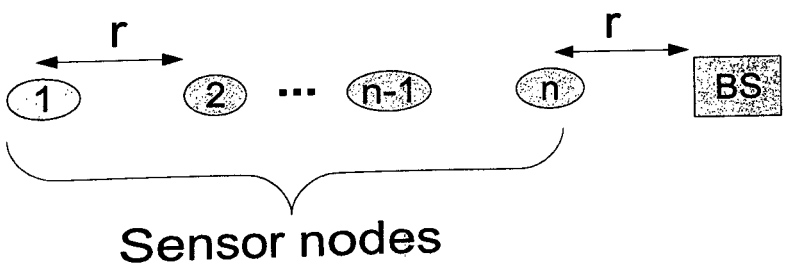


Figure 1: Linear network with n nodes and the base station.

2. (a) Describe a fault detection technique and a fault recovery technique for wireless sensor networks. [5+5=10]

   (b) What is attribute-based addressing in wireless sensor networks? [5]

   (c) How soft and hard thresholds are used in Threshold sensitive Energy Efficient Network protocol (TEEN) in wireless sensor networks? [10]

   (d) i. Write an approximation algorithm for the following minimum relay node placement problem in wireless sensor networks:
   Given a set of sensor nodes S in 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 2-connected. The objective of the problem is to minimize $|R|$, where $|R|$ denotes the number of relay nodes in R. [15]

   ii. Comment whether your approximation algorithm works in the scenario where communication radii of sensor nodes and relay nodes are different. [5]

P.T.O

3. (a) Briefly describe the Fixed Spectrum Access (FSA) and Dynamic Spectrum Access (DSA) policies in cognitive radio networks. [5+5=10]

(b) What are the different types of cognitive capabilities with which a cognitive radio user may be equipped to support DSA. [5]

(c) Describe a direct spectrum sensing technique and an indirect spectrum sensing technique in cognitive radio networks. [5+5=10]

(d) What are the important features that any routing protocol for cognitive radio networks should have? [5]