資格考 演算法

1. (20%) Solving the recurrence $T(n) = 2T(\lfloor\sqrt{n}\rfloor) + \log_2 n$ using big-O notation as tight as possible.
2. (20%) Consider the following two problems in which we are given a directed graph $G=(V, E)$ and vertices $u, v \in V$.

   Unweighted shortest path problem: Find a path from u to v consisting of the fewest edges.

   Unweighted longest simple path problem: Find a path from u to v consisting of the most edges.

   (a) (10%) Determine which problem can be solved using the dynamic-programming in polynomial time.

   (b) (10%) Determine which problem cannot be solved using the dynamic-programming in polynomial time.
3. (30%) Consider the problem of finding the 5-vector $x=(x_i)$ that satisfies

$$\begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & -1 \\ -1 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} \leq \begin{pmatrix} 0 \\ -1 \\ 1 \\ 5 \\ 4 \\ -1 \\ -3 \\ -3 \end{pmatrix}$$

   . Determine how many solutions to this problem.
4. (30%) Show the lower bound for any comparison sort algorithm.

Graph Theory 資格考

1. (20%) For a set $S \subseteq N$ of size n, determine the number of spanning trees with vertex set S.

2. (20%) Show that an edge is a cut-edge if and only if it belongs to no cycle.

3. (20%) Prove or disprove: If T is a minimum-weight spanning tree of a weighted graph G, then u,v-path in T is a minimum-weight u,v-path in G.

4. (20%) Show that the minimum number of edges in a connected graph with n vertices is n-1.

5. (20%) Prove or disprove: Every tree has at most one perfect matching.

5/2008 博士班資格考:　機率與統計 Show All Details.

1. If the RVs X and Y are joint normal with zero mean, find the estimate of parameter y given x.　(20%)

2. (30%)
   Consider the R.V. with the density function as

$$f(x)=\begin{cases}\frac{1}{\beta}e^{-x/\beta} & x\geq 0\\ 0 & otherwise\end{cases}$$

   Suppose that independent observations $x_1, x_2, x_3, \ldots, x_n$ are taken from this distribution. Then
   a. What is the likelihood function
   b. Find the MLE for the parameter, β

3. (20%)
   Let X and Y be two RV's with finite variances. Show that $E^2\{XY\}\leq E^2\{X\}E^2\{Y\}$.

4. (30%)
   a. Write out the *p.d.f.* of a Poisson process with parameter a.
   b. Find the Characteristic function of a Poisson process.
   c. Use your result in b. to find the mean and the variance of a Poisson process.

Cryptography

Ph.D Qualify Examination　　　　Close Book　　　　2008. May

1. Define the following terms: (20%)

(1) Security Attack.　　(2) Security Mechanism.

(3) Security Service.　　(4) Steganography.

(5) [illegible]　　(6) [illegible]

(7) Data Confidentiality.　　(8) Data Integrity.

(9) Non-repudiation.　　(10) Euler Totient Function.

2. Discuss the pros and cons of private-key crypto system and public-key crypto[illegible]. What are the usefulness of them in practice? (20%)

3. List and show how the trapdoor one-way functions are used to construct modern cryptography? (20%)

4. Is cryptography the only way to achieve network security? What is the limitation of modern cryptography? What else is/are also important? (20%)

5. Explain all what you know about security. (20%)

DBMS Qualify Exam

1. (30%) Consider the following schema for a suppliers-and-parts database:

SUPPLIER(<u>SupNo</u>, SName, Status, City)

PART(<u>PartNo</u>, Color, Weight, City)

PROJECT(<u>ProjNo</u>, PName, City)

SHIPMENT(<u>SupNo, PartNo, ProjNo</u>, Qty)

Answer the following queries in SQL.

(a) Get the total weight for the parts that have a "red" color.

(b) Get the total quantity for the red parts that are supplied by any supplier and used in any project.

(c) Get the suppliers that supply a project with the parts being all red parts.

2.(30%) Answer the above queries in relational algebra.

3. (15%) A relation, R(A, B, C, D, E, F, G), whose attributes satisfy the functional dependencies: $(BC \rightarrow A, D, E, F, G)$, $(C \rightarrow E)$, $(D \rightarrow F, G)$, $(D \rightarrow B)$

Normalize the above relation to make it satisfy

(a) 2NF

(b) 3NF

(c) BCNF

Note: Don't make unnecessary normalization unless it is required.

4.(15%)

(a) Explain what serializability is.

(b) What is two-phase locking protocol?

(c) Is two-phase locking a necessary or sufficient condition for serializability? Explain why.

5. (10%) In performing an update operation, the record to be updated must be locked (as indicated in concurrency control protocols). Then, can a query access these locked records during query processing? If no, how can a query result be accurate with these records being locked and cannot be accessed? If yes, then will the access to these locked items violate the concurrency control protocol (because the lock on the records is broken)? How does a DBMS manage this problem?

Digital Image Processing

PhD Qualification Examination

Department of CSIE, NCKU

May. 2008

1. Color image processing is an important technique supporting many applications in medical image analysis. Please give two examples in medical image processing by using color technique. Please explain as detailed as possible. (20%)

2. Please define and describe morphological dilation, erosion, open, close, and hit or miss transformation operations for binary images. (12%) Please use the above–mentioned operations to extract edges (object boundaries) from binary image. (8%)

3. If an image undergoes uniform linear motion and is blurred with linear motion, you are asked to restore the image from motion blur. Based on the concept of linear shift-invariant system, please explain how you can restore the image with inverse filtering or Wiener filtering. (10 %) If the image moves in the x-direction only, at a rate given by $x_0 = at/T$. When t=T, the image has been displaced by a total distance a. With $y_0(t) = 0$, please obtain the degradation function of the blur. (10 %)

4. What is the deformable model (or snake) algorithm? (8%) What is the watershed segmentation algorithm? Please describe them as thorough as possible. (7%) Please also describe their major similarities and differences? (5%)

5. Please describe the convolution theorem in discrete Fourier transform. (6%) Please explain how to perform image filtering on the spatial and frequency domain. (6%) What are the differences between the two filtering operations? Are they equivalent? (8%)

Qualification Exam (May 2008)

Pattern Recognition

1. (15%) Explain the following terms
   (a) Bayesian Information Criterion
   (b) Fisher's Linear Discriminant
   (c) The Bias-Variance Decomposition

2. (25%) The broad class of distributions (also called the *exponential family*) of a random vector $\mathbf{x}$ given parameters $\boldsymbol{\eta}$ is written by

$$p(\mathbf{x} \mid \boldsymbol{\eta}) = h(\mathbf{x}) g(\boldsymbol{\eta}) \exp\{\boldsymbol{\eta}^T \mathbf{u}(\mathbf{x})\}.$$

   Show that the Gaussian distribution belong to the exponential family and find the corresponding $\boldsymbol{\eta}$, $h(\mathbf{x})$, $\mathbf{u}(\mathbf{x})$ and $g(\boldsymbol{\eta})$. How can you apply the maximum likelihood principle to find the parameters $\boldsymbol{\eta}$ from training data $X = \{\mathbf{x}_1, \cdots, \mathbf{x}_N\}$ and obtain the *sufficient statistics* $\sum_n \mathbf{u}(\mathbf{x}_n)$?

3. (15%) What is the objective function of perceptron algorithm? Explain the perceptron convergence theorem.

4. (25%) A general linear regression model for expressing multidimensional data $\mathbf{x} = [x_1, \cdots, x_D]^T$ is written by

$$y(\mathbf{x}, \mathbf{w}) = \omega_0 + \sum_{j=1}^{M-1} \omega_j \varphi_j(\mathbf{x}) = \mathbf{w}^T \varphi(\mathbf{x}),$$

   where $\mathbf{w} = [\omega_0, \cdots, \omega_{M-1}]^T$ are regression parameters and $\varphi = [\varphi_0, \cdots, \varphi_{M-1}]^T$ are basis functions. Given a set of input variables and target variables $\{\mathbf{x}_n, t_n\}_{n=1}^N$, show your procedure of finding the least squares solution and illustrate the relationship between least squares and maximum likelihood. In addition, illustrate the relationship between the regularized least squares and the Bayesian linear regression.

5. (20%) Laplace approximation is used to find a Gaussian approximation to a probabilistic distribution function $p(\mathbf{z})$ over a set of M continuous random variables $\mathbf{z}$. Given a distribution denoted by $p(\mathbf{z}) = \frac{f(\mathbf{z})}{\int f(\mathbf{z}) d\mathbf{z}}$, show your procedure of finding the approximated distribution $q(\mathbf{z})$ expressed by

$$q(\mathbf{z}) = \frac{|A|^{1/2}}{(2\pi)^{M/2}} \exp\left\{-\frac{1}{2}(\mathbf{z}-\mathbf{z}_0)^T A(\mathbf{z}-\mathbf{z}_0)\right\},$$

where $A = -\nabla\nabla \ln f(\mathbf{z})\big|_{\mathbf{z}=\mathbf{z}_0}$ and $\mathbf{z}_0$ is the mode of $p(\mathbf{z})$.