

## 作業系統 資格考題 (102 年度第 1 學期)

1. (20%) Briefly describe the microkernel and its advantages.
2. (30%) Please describe multi-level queue scheduling and multi-level feedback queue scheduling. Please also describe the difference between them.
3. (30%) Describe the actions of handling a page fault and the actions of handling a TLB miss. What is the difference between these two types of actions?
4. (10%) Briefly describe the *processor affinity* in SMP systems.
5. (10%) Implement *lock acquisition/release* by using the *swap* instruction. The *swap* instruction exchanges the content of two memory addresses atomically. Please write the pseudo code describing jobs of lock acquisition before entering the critical section and jobs of lock release after leaving the critical section.

**2013 Nov. NCKU CSIE PH.D. Qualification Examination**  
**Computer Architecture**

1. With dynamic hardware prediction for reducing branch costs, what is the disadvantage of a simple 1-bit branch-prediction buffer for a branch that is almost always taken. Explain why the 2-bit prediction scheme can remedy this disadvantage. Also, explain what is correlated predictors by illustrating an example. (20 points)
2. Explain the following synchronization primitives: atomic exchange, test-and-set, and fetch-and-increment. Also, explain what is the pair of instructions, load linked (LL) and store conditional (SC) and how this pair of instructions can be used to implement atomic exchange and fetch-and-increment. (15 points)
3. The classical approach to improving cache behavior is to reduce miss rates. Please summarize the techniques that can reduce miss rates. (20 points)
4. Describe two major instruction set characteristics that can further divide general purpose register (GPR) instruction set architecture into three classes, based on whether the instruction operands are used explicitly or implicitly. And show the advantages and disadvantages of these three further divided classes. (15 points)
5. Describe what are the RAW, WAW, and WAR hazards. (15 points)
6. Explain why the cache miss penalties increase as the processor becomes much more faster than DRAMs? Briefly describe five optimizations to reduce cache miss penalty. (20 points)

9/2013 博士班資格考： 機率與統計 Show All Details.

1. (20%) A binomial parameter  $p$  describes the probability of success. The corresponding probability of failure is therefore  $q=1-p$ . In a random sample of 4000 families which subscribing newspaper in Taipei, it was found that 1400 families subscribing to Daily News. Please find the 95% confidence interval for the actual proportion of families subscribing to Daily News.

2. (20%) Consider the R.V. with the density function as

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

Suppose that independent observations  $x_1, x_2, x_3, \dots, x_n$  are taken from this distribution. Then,

- What is the likelihood function
- Find the MLE for the parameter,  $\beta$

3. (20%) A RV  $X$  has the distribution as

$$P(X = k) = pq^k, \quad k = 0, 1, \dots$$

Find  $\Gamma(z)$  and use this to find the mean and variance given  $p=0.3$  and  $q=0.7$ .

4. (20%) Let  $X$  be a random variable and  $Y=g(X)$ . Prove that

(a)  $E[Y] = E[g(x)]$ .

(b)  $f_Y(y) = \frac{f_X(x_1)}{|g'(x_1)|} + \frac{f_X(x_2)}{|g'(x_2)|} + \dots + \frac{f_X(x_n)}{|g'(x_n)|} + \dots$ , where

$$y = g(x_1) = g(x_2) = \dots$$

5. (20%) Given a sequence of RVs  $x_1, x_2, \dots, x_n$ ,
- Give the definition of the correlation of two of the RVs.
  - Given the definition of the correlation matrix of these RVs.
  - Prove that the correlation matrix is non-negative definite.
  - In what condition that the correlation matrix is symmetry

密碼學博士資格考

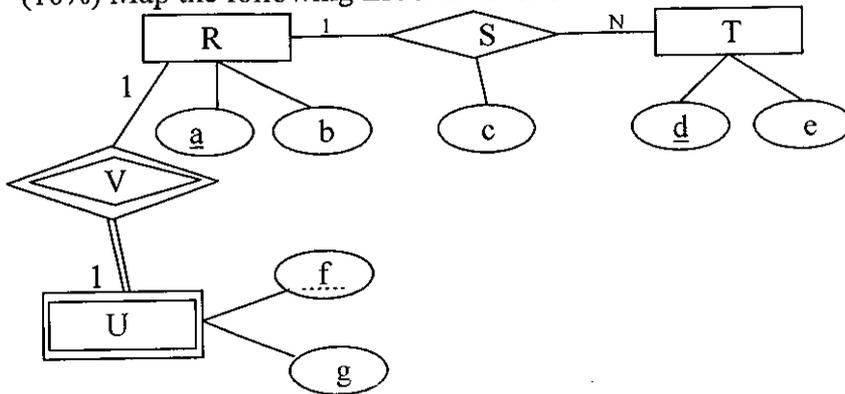
Cryptography PhD Qualified Exam.

(Close book, 2013/10/18)

- (1). (20%) What is Steganography? Describe a way to Provide steganography.
  
- (2). (40%) What are the differences between modern cryptography and quantum cryptography.
  
- (3). (40%) How to provide computer network security. Please write your answer in a systematic way.

# 2013 DBMS Qualify Exam

1. (10%) Map the following ER schema to relations.



2. (40%) Given a database schema as follows.

S(S#, Sname, Status, City) /\* This is a relation for Supplier \*/

P(P#, Pname, Color, Weight, City) /\* This is a Part relation \*/

J(J#, Jname, City) /\* This is a Project relation \*/

SPJ(S#, P#, J#, Quantity)

Answer the following queries in SQL.

- (a) Get the names of the suppliers that do not supply any part to any project.
- (b) Get the names of the projects that are supplied by more than 2 suppliers or all the parts used in the project are white parts.

Answer the following queries in relational algebra.

- (c) Get the names of the suppliers supplying parts that contains all the parts project P3 uses.
- (d) For each project, list the number of suppliers that supply this project, and the average weight of the parts used in this project.

背面有題

3. (10%) A relation,  $R(A, B, C, D, E, F)$ , whose attributes satisfy the functional dependencies:

$(A, B \rightarrow C, D, E, F), (A \rightarrow C), (B \rightarrow D), (E \rightarrow E)$

Normalize the above relation to make it satisfy

(a) 2NF

(b) 3NF

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

4. (10%) Give a formal definition for the left-outer join operation.

5. (30%) Explain the following terms

(a) Referential Integrity Constraint.

(b) Impedance mismatch.

(c) Materialization of relational view.

(d) The wait-die protocol (for breaking a deadlock)

(e) Cascading rollback.

(f) Recovery protocol based on immediate update.

## Qualify Exam - Data Mining

Notice: 1) Close book. 總分: 100 分; 2) 可以中文或英文作答。

(共 1 頁)

1. (20%) Answer the following questions:
  - A. Explain what are *closed itemsets* and *maximal itemsets*, respectively.
  - B. Give an efficient way to do incremental mining of sequential patterns.
  - C. Explain how "*K-fold cross validation*" works for evaluating a classification model.
  - D. Explain what is "overfitting" problem in classification modeling and how to avoid it in using decision tree for classification.
  
2. (25%) Answer the following questions on frequent itemsets mining:
  - A. Give the algorithms of *Apriori* and *FP-Growth*, respectively.
  - B. Compare both methods in terms of execution efficiency and memory space requirement, respectively.
  
3. (25%) Answer the following questions on clustering methods:
  - A. Explain in details how *k-means* and *DBSCAN* work respectively.
  - B. Give comparisons on advantages and drawbacks between the above two clustering methods.
  - C. Given a dataset  $D$ , suppose two clustering results  $R1$  and  $R2$  are obtained using *k-means* and *DBSCAN*, respectively. Explain how to compare the quality of the clustering results  $R1$  and  $R2$  through cluster validation methods.
  
4. (20%) Given a dataset  $D$  of  $m$  records, with attributes  $\{A_1, A_2, \dots, A_n\}$  and a class  $C$  ( $C$  is in categorical type) in each record, where the value of any  $A_i$  ( $1 \leq i \leq n$ ) is in categorical (nominal) type. Answer the following questions about classification modeling:
  - A. If the distribution of class  $C$  in  $D$  is uniform, in general cases, how would you rank the order of *Decision Tree*, *SVM (Support Vector Machine)* and *CBA (Classification Based on Association)* in terms of the accuracy and execution efficiency in building a classification model on  $D$ ? Explain why.
  - B. Suppose the distribution of class  $C$  in  $D$  is highly imbalanced such that the *Recall* of the classification result is low. Without changing the classification methods, what is the most effective approach to deal with dataset  $D$  for improving the *Recall*?
  
5. (10%) Answer the following questions:
  - A. Explain in details how "*Ensemble*" and "*Boosting*" methods work for building a classifier, respectively.

102. 11.1 資格考 VLSI

1. Explain the following terms in detail: (80%)

- (a) critical path
- (b) setup time
- (c) clock skew problem
- (d) RTL code
- (e) hold time
- (f) soft IP
- (g) functional simulation
- (h) power dissipation (power consumption)

2. Describe the difference between full custom and Cell-based design flow. (20%)

演算法資格考 Nov. 2013

1. (20%) Solve  $T(n) = T(\frac{n}{3}) + T(\frac{2n}{3}) + \Theta(n)$  using  $\Theta$ .
2. (20%) Determine the height of  $n$ -node heap, and provide a proof for your answer.
3. (20%) Present the well-known counting sort algorithm and analyze the complexity.
4. (20%) Describe a  $\Theta(n \lg n)$ -time algorithm that, given a set  $S$  of  $n$  integers and another integer  $x$ , determine whether or not there exist two elements in  $S$  whose sum is exactly  $x$ .
5. (20%) Explain why the statement, "The running time of algorithm A is at least  $O(n^2)$ ," is meaningless.

Ph.D. Qualify Examination 2013  
Theory of Computation

- This examination is closed books.
- Please turn off your cell phones.
- Remember that there are 2 pages of the qualify examination.
- Answer all questions as possible. You may have a partial score if you answer in the correct direction.

1. Deterministic Finite Acceptors (DFAs) (20 pts. 10 pts each)
  - (a) Construct a Deterministic Finite Automata to accept the set of all strings over  $\{0, 1\}$  that do not contain the substring "101".
  - (b) Construct a DFA with  $k$  states over  $\{0, 1\}$  which accepts the language  $L_k = \{0, 1\}^* 0^{k-1} 0^*$ . (10 pts)
2. Nondeterministic Finite Acceptors (NFAs) (20pts. 10 pts each)

Draw Non-deterministic Finite Automata to accept the following sets of strings over  $\{a, b\}$ :

  - (a) Construct an NFA to accept that "All" strings end in "aaa".
  - (b) Construct an NFA with "three states" that accepts the language  $\{ab, abc\}^*$ .
3. Prove that the language is nonregular: (15 pts)  
 $L = \{a^n b a^{n+1} : n \geq 0\}$ .
4. Suppose  $r_1$  and  $r_2$  are regular expressions, and  $L(r_1)$  and  $L(r_2)$  are regular languages, correspondingly. Prove that  $L(r_1 + r_2)$  is also a regular language. (15 pts)  
hint: Note that  $L(r_1 + r_2) = L(r_1) \cup L(r_2)$ . You can consider the solution by using NFA.
5. Construct an NPDA that accepts the following language on  $\{a,b,c\}$  (use an NPDA with 2 states): (10 pts)  
 $L = \{w : n_a(w) + n_b(w) = n_c(w)\}$ .

6. Fill the following languages into the language hierarchy (If  $L_i$  is a regular language and also a context-free language, please fill  $L_i$  in the set of regular languages): (20 pts)

$$L_1 = \{a^n b^n : n \geq 0\},$$

$$L_2 = L(a^* b^*),$$

$$L_3 = \{a^n b^n c^n : n \geq 0\},$$

$$L_4 = \{a^n w w^R a^n : n \geq 0, w \in \{a, b\}^*\},$$

$$L_5 = \{ab, ad, a\},$$

$$L_6 = \{ww : w \in \{a, b\}^*\},$$

$$L_7 = \{a^{n!} : n \geq 0\},$$

$$L_8 = \{a^n b^j a^j b^n : n \geq 0, j \geq 0\},$$

$$L_9 = \{a^n b^m c^{n+m} : n \geq 0, m \geq 0\},$$

$$L_{10} = \{a^3 b^n c^n : n \geq 0\}.$$

