國 立 清 華 大 學 命 題 紙

99 學年度 服務科學研究所乙組(服務系統組) 碩士班入學考試

科目 計算機概論 科目代碼 4801 共 3 頁，第 1 頁 *請在【答案卷卡】作答

Part I: Short Essay Questions (簡答題) 5 題，共 36 分。

1. Describe the advantages and disadvantages of direct memory access (DMA). Consider a system that needs to service both time-sharing and real-time tasks. **(6 points)**

2. A database is defined as follows: "A database, representing some aspects of the real world, consists of a self-describing collection of integrated data that are shared by various application programs of specific purposes." Please explain what does it mean by "self-describing collection of data." Why is it important? **(5 points)**

3. The execution of an instruction on a computer can be divided into the instruction cycle and the execution cycle. What are the two main steps performed in the instruction cycle? **(5 points)**

4. There is a performance problem associated with the display of bit-map terminals. When we scroll the screen, it will trigger a request for redrawing the screen. One way is copying all the bits on the screen to store in memory, but doing this will put a huge load on the CPU. Some video cards have some special feature in handling this scrolling screen situation without copying the content of the screen. How can this be done? Please describe. **(10 points)**

5. Race Condition may happen when multiple processes try to access a common resource, thus execution results depending on some shared state. Please *use one example to illustrate the Race Condition* for multiple processes execution when they both trying to increment the value of a global variable.
   What can be done to avoid the undesirable results of Race condition in this example. **(10 points)**

Part II: Problem Solving Questions 4大題，共64分。

6. Assume that we are dealing with the problem of finding the intersection of two **sorted arrays** A and B (i.e., finding elements that exist in both arrays). We further assume that the elements in A are distinct and the elements in B are distinct. In your answers to the following questions, you can use any standard algorithm, provided you state exactly what the algorithm does.

   a) In this question, we assume that the two arrays are relatively short (e.g., $|A| = |B| = 100$). Please develop an efficient algorithm (in pseudo-codes) that finds and prints the intersection of two sorted arrays. What is the time complexity of your algorithm as a function of $|A|$, when $|A|$ and $|B|$ are approximately equal? **(10 points)**

   b) Now assume that the size of B is much larger than that of A (e.g., $|A| = 100$ and $|B| = 1000000$). Please develop a better-suited algorithm (in pseudo-codes) that finds and prints the intersection of these two sorted arrays. What is the time complexity of your algorithm as a function of $|A|$ and $|B|$ in the case $|A| << |B|$? **(10 points)**

7. Apple's iPod claims to have the capability of preventing from skipping effect. That is, when you shake the iPod device, there will be no effect on the music it plays. iPod provides skip protection by playing music from its solid-state cache. When all the music stored in the cache has played, iPod's hard drive starts copying more music to the cache. Shaking iPod while the hard drive is copying music can cause songs to skip. Assuming a disk is with two cylinders, each with 6 tracks and 64 sectors/track, each sector containing 512 data bytes, and a rotation rate of 7200 RPM. A seek takes 20 msec on average. The size of solid-state cache is 4 MB. In the condition of continuous playing, ***what is the time it takes to fill up the solid-state cache***. In the newer generation of the iPod, what would you suggest to make the skip protection works better. Please ***give two recommendations***. Note: You may explicitly make assumption of any additional variables you think are necessary in doing the calculation in answering the question. **(10 points)**

8. Google uses a particular method in ranking the search results, which is called PageRank. Google describes PageRank as follows: "PageRank relies on the uniquely democratic nature of the web by using its vast link structure as an indicator of an individual page's value. In essence, Google interprets a link from page A to page B as a vote, by page A, for page B. But, Google looks at more than the sheer volume of votes, or links a page receives; it also analyzes the page that casts the vote. Votes cast by pages that are themselves "important" weigh more heavily and help to make other pages important."

a) Using a very simple calculation, suppose there are four pages A, B, C, and D. In its simplest form, assuming the total PageRank is 1. Thus by evenly distributing among the four pages, each page has initially a PageRank value of 0.25. When Pages A, B, C all point to Page D, then Page D will have a PageRank value of 0.75. Now consider a group of ten pages. Each page can have only one link pointing toward other page (or outbound links). And each page can receive unlimited number of links pointing toward itself (or inbound links). Please ***draw the links among this group of ten pages (your answer will contain the structure of the links)***, such that this resultant link structure will contain a page that has the highest possible PageRank value among all other types of link structure. Also ***calculate the highest PageRank value for that page***. **(10 points)**

b) Based on our experiences, one can agree that the earlier (or higher) a site appears in the search results list, the more visitors it will receive from the search engine. Thus there is an interest for web sites to move their sites to appear in the higher position. In this question, consider one search engine employing the PageRank as the underlying ranking mechanism. What are the methods you will use to move one web site higher in the search result list. Please name ***three different methods***. For each method, you need to write down ***What to do, and Why?*** Note that your answer should be specifically responding to the underlying ranking mechanism we describe here, not any general situation. **(10 points)**

9. The following table shows a relation called GRADE-REPORT for a university. Assume that the primary key of this relation is {Student-ID, Course-ID}. Consider the following functional dependencies existed in this relation:
   - Student-ID → Student-Name, Major (assume that a student can have only one major)
   - Course-ID → Course-Title, Instructor-Name, Instructor-Office (assume a course is taught by only one instructor)
   - Student-ID, Course-ID → Grade
   - Instructor-Name → Instructor-Office (assume the names of instructors are unique).

GRADE-REPORT Relation

Student-ID	Student-Name	Major	Course-ID	Course-Title	Instructor-Name	Instructor-Office	Grade
12345	Lee	MIS	MIS 331	Database	Codd	B104	A
12345	Lee	MIS	MIS 341	System Analysis	Parsons	B317	B
22334	Baker	Acct	MIS 480	Expert Systems	Codd	B104	C
22334	Baker	Acct	Acct 201	Accounting	Miller	H310	B
22334	Baker	Acct	Mkt 300	Marketing	Bennett	B212	A

a) Please apply the 2nd normal form (2NF) normalization process to decompose the GRADE-REPORT relation into a set of 2NF relations. **(6 points)**

b) Based on the answer to a), please apply the 3rd normal form (3NF) normalization process to decompose each of the 2NF relations that you produce in a) into a set of 3NF relations. **(8 points)**