國立成功大學 108 學年度碩士班招生考試試題

系　　所：數據科學研究所

考試科目：計算機概論　　考試日期：0223，節次：2

※ 考生請注意：本試題不可使用計算機。　請於答案卷(卡)作答，於本試題紙上作答者，不予計分。

1. 單選題

(1) (4%) Which of the following does not store data permanently?

(A) ROM

(B) RAM

(C) Hard Disk

(D) USB

(2) (4%) Which of the following statements are correct?

S1. Registers in CPU is used to store intermediate data and instructions.

S2. Program counter keeps track of the memory address of the instruction that is to be executed next.

S3. Speed of CPU is also known as clock speed, which is the number of instructions executed by CPU in one second.

S4. Memory Address Register (MAR) acts as an interface between CPU and memory. When CPU issues a Read Memory command, instruction is fetched and placed in MAR.

(A) S1 and S4

(B) S2 and S4

(C) S3 and S4

(D) S2 and S3

(3) (4%) Which of the following statements are **not** correct?

(A) A computer stores all data in binary.

(B) The statement age = age + 1 increases the value that is in the age variable by 1.

(C) A byte can hold any number between 0 and 255.

(D) A sequence of 8 bits, like 00111110, could be interpreted as a number, but it cannot be interpreted as a letter.

(4) (4%) Regarding the features of recursive functions, which is **not** correct?

(A) One or multiple base cases and one or multiple recursive cases.

(B) The function calling itself at some point.

(C) Testing for a base case before calling a recursive case.

(D) Recursion is memory-intensive since it tends to declare many local variables.

(5) (4%) Regarding public key encryption, which of the following statements are **not** collect?

(A) A message encrypted by the public key can only be decrypted by the secret key.

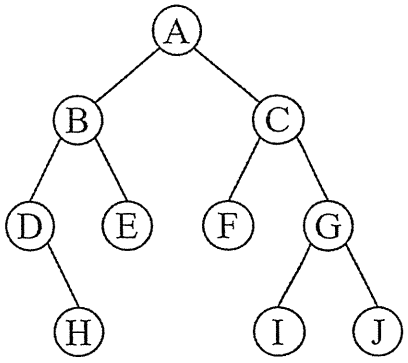
(B) A message encrypted by the secret key can only be decrypted by the public key.

(C) A message encrypted by the public key can also be decrypted by the public key.

(6) (4%) Which of the following network services in the Internet use Transmission Control Protocol (TCP) as the transport layer protocol?
(A) World Wide Web (WWW)
(B) FTP software
(C) Telnet
(D) LINE text messages
(E) All of above.

(7) (4%) Linear search is very inefficient compared to binary search when facing which of the following data?
(A) Small and sorted arrays.
(B) Small and unsorted arrays.
(C) Large and sorted arrays.
(D) Large and unsorted arrays.

(8) (4%) Traverse the given tree below using Inorder, Preorder, and Postorder traversals, which one is **not** correct?



(A) Preorder: A B D H E C F G I J
(B) Postorder: H D E B A F I J G C
(C) Inorder: D H B E A F C I G J

(9) (4%) Which of the following statement about data structures is correct?
(A) Every binary search tree with n nodes has height $O(\log n)$
(B) Let T be a minimum spanning tree of graph G. Then, for any pair of nodes s and t, the shortest path from s to t in G is the path from s to t in T.
(C) Let T be a complete binary tree with n nodes. Finding a path from the root of T to a given node $v \in T$ using Breadth-First Search takes $O(\log n)$ time.
(D) All of above are incorrect.

(10) (4%) Which of the following statements about database systems is correct?

S1. Database management system (DBMS) is a software used for management, maintenance and retrieval of data stored in a database.

S2. Candidate key is a set of one or more attributes that uniquely identifies tuple within a relation.

S3. Foreign key is a non-key attribute, whose values are derived from the primary key of some other table is known as foreign key.

S4. All attribute combinations inside a relation that can serve as primary key are called primary key.

(A) S1 and S2

(B) S2 and S3

(C) S1 and S3

(D) S2 and S4

2. (9%) Explain the algorithm of **Quick Sort**, where you can create an example to describe how the algorithm works. Then please show its **worst**-case and **average**-case time complexity.

3. Answer the following questions about Operating Systems.

(1) (3%) What are the differences between starvation and deadlock?

(2) (3%) What are the differences between process and thread?

(3) (3%) CPU scheduling is a process which allows one process to use the CPU while the execution of another process is on hold (in waiting state) due to unavailability of any resource like I/O etc, thereby making full use of CPU. The aim of CPU scheduling is to make the system efficient, fast and fair. You are asked to provide **two CPU scheduling** algorithms, and briefly describe each of them.

4. Answer the questions on artificial intelligence and data science.

(1) (5%) What is Turing test?

(2) (5%) Briefly describe logistic regression.

(3) (5%) What is deep learning? What does "deep" mean? **Why** "deep"?

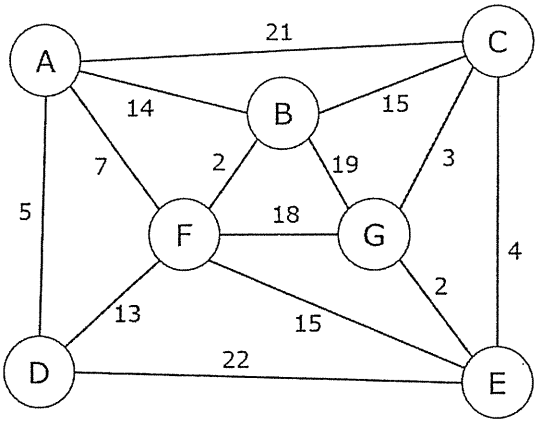
5. (4%) Consider a database with the table created by the following SQL statement

```
CREATE TABLE G (                -- G is short for Grades
    sid INT,                    -- sid is the ID of student
    class CHAR(20),             -- class name
    grade INT,                  -- grade score between 0 and 100
    PRIMARY KEY (sid, dept)
);
```

Answer the question by providing valid SQL query: Find the student IDs (sid) of all students who have received grade scores of 85 or above in at least 3 classes.

6. Expression representation and translation.
   (1) (3%) Draw a Binary Tree for the expression: A * B – (C + D) * (P / Q)
   (2) (3%) Translate infix expression, A * (B + D) / E – F * (G + H / K), into its equivalent postfix expression.

7. Given a weighted graph below, compute and draw **minimum spanning trees** by the following two different algorithms.



(1) (4%) **Prim's** algorithm, starting from node A. During your algorithm, if two unvisited nodes have the same distance, use alphabetical order to determine which one is selected first.
(2) (4%) **Kruskal's** algorithm. When you are selecting two edges with the same weight, select the edge that comes alphabetically **last** (e.g., select E-F before B-C. Also, select A-F before A-B).

8. Suppose you create a hash table of size 7 with hash function $h(k) = k \mod 7$. Draw the resulting table after inserting the following numbers in the given order: 19, 26, 13, 48, 17, for each of the three settings listed below.
(1) (3%) Handle collisions using **separate chaining**.
(2) (3%) Handle collisions using **linear probing**.
(3) (3%) Handle collisions using **double hashing** with a second hash function: $h'(k) = 5 - (k \mod 5)$.