國立成功大學
110學年度碩士班招生考試試題

編　號：268

系　所：數據科學研究所

科　目：計算機概論

日　期：0202

節　次：第 2 節

備　註：不可使用計算機

編號： 268 國立成功大學 110 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：0202，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. (40%) In the following statements, please specify if the statement is **True** or **False**. If the statement is **True**, explain why it is True. If it is **False**, give the correct answer or explain why.

(1) ■ Assumed a programming language is column-major based. The run-time of an algorithm adopted column-major scheme is the same as that adopted row-major.

(2) ■ RSA encryption is asymmetric cryptography, and it includes three keys: 1)public key, 2) private key, and 3) paired key.

(3) ■ Only Floyd-Warshall and Bellman-Ford can be used on a negative-weighted graph in shortest path algorithms, whereas the Dijkstra algorithm cannot.

(4) ■ Consider a Python code below, and the output will be "[[0], [1], [2], [3], [4]]"

(a) ◆ print([[x] for x in range(5)])

(5) ■ Shell sort (diminishing increment sort) has the following properties:

(a) ◆ Worst Case: $O(n^2)$, Average: $O(n^2)$, and is a Stable Sorting.

(6) ■ Huffman Tree is the most common tree structure for fast search purposes.

(7) ■ The following two combinations can determine the unique Binary Tree Structure:

(a) ◆ Give Preorder and Postorder sequences

(b) ◆ Give Postorder and Inorder sequences

(8) ■ *k*-nearest neighbor (KNN) search is the simplest classifier determined by finding the minimum Elcudiean distance *D(a, b)* between the test sample between each training sample. Given the two datasets below, KNN can conduct precisely the same results without any modification.

$$D(a,b) \updownarrow \sqrt{(a_1 \angle b_1)^2 \vee (a_2 \alpha b_2)^2 (\ldots ((a_n . \ b_n)^2}$$

Dataset 1

a_1	a_2	a_3	Label
4	4	0	Apple
5	5	1	Apple
7	6	1	Banana
4	3	0	Apple
6	7	1	Banana
5	8	1	Banana
5	6	1	Apple

Dataset 2

b_1	b_2	b_3	Label
4	40	0	Apple
5	50	1	Apple
7	60	1	Banana
4	30	0	Apple
6	70	1	Banana
5	80	1	Banana
5	60	1	Apple

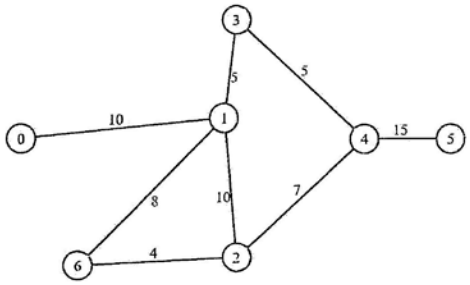
國立成功大學 110 學年度碩士班招生考試試題

系　　所：數據科學研究所

考試科目：計算機概論　　考試日期：0202，節次：2

2. (10%) The sentence "THIS IS A BOOK" can be compressed by Huffman Coding. Please answer the following questions in detail:
   (1)■ (3%) Draw the Huffman Tree
   (2)■ (4%) Find the bit-code representation of Huffman Coding for the sentence.
   (3)■ (3%) Find the compression ratio (Assumed each character is stored using eight bits of information).

3. (12%) Spanning tree.
   (1)■ (4%) Please describe how Kruskal's algorithm works.
   (2)■ (8%) Give a graph below. Please find the minimum cost spanning tree by Kruskal's algorithm.



4. (8%) Consider a pseudo code for preorder below. Please **FIND** the **pseudo-codes** for postorder and inorder tree traversal.

```
void preorder( BTREE_NODE_p_t node )
    if ( node != NULL )
        visit( node )
        preorder( node->left )
        preorder( node->right )
```

5. (5%) Divide and Conquer is a recursive algorithm applied to many problems. If the problem is too big, then recursively divide the problem into smaller subproblems until the problem size is small enough to solve. Please identify which following sorting algorithm(s) is/are divide and conquer algorithms based on the above description.
   (1)■ Merge sort
   (2)■ Insertion sort
   (3)■ Selection sort
   (4)■ Quick sort

國立成功大學 110 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：0202，節次：2

6. (16%) Answer the following questions on data science.
   (1)■ (3%) Please give one real-world application for logistic regression.
   (2)■ (3%) What is overfitting? Please give an example to explain why and how.
   (3)■ (4%) Consider two well-trained machine learning models. Please give a strategy to determine which model is better.
   (4)■ (6%) Consider noisy training data consisting of a lot of missing values. Please give the solution to the following two conditions of missing values.
      (a)◆ Most values of one of the features are missing (either nan or empty)
      (b)◆ Less than 5% of the features are missing for some samples (say, <10%).

7. (4%) Explain what weak AI and strong AI are.

8. (5%) Consider a recursive tree insertion algorithm below. The problem below aims to complete the following two tasks (tips: node behaviors):
   (1)■ Blank (B1) in the program below should be_____.
   (2)■ Blank (B2) in the program below should be_____.

```
Algorithm: recursiveInsert (node, key, value)
Input: a tree node (root of subtree), key-value pair

  if key < node.key

    if _____(B1)
      recursiveInsert (node.left, key, value)
      //In this case, find the corresponding key recursively.
      return
    endif
    Create new node and append to node.left; // Otherwise, append key here.

  else // Otherwise, go right.

    if _____(B2)
      recursiveInsert (node.right, key, value)
      return
    endif
    Create new node and append to node.right;

  endif
```