國立成功大學

115學年度碩士班招生考試試題

編　號：130

系　所：電機工程學系

科　目：資料結構

日　期：0203

節　次：第 2 節

注　意：1.不可使用計算機
2.請於答案卷(卡)作答，於試題上作答，不予計分。

Problem 1: Priority Queue (20 pts)

A hospital emergency department treats patients based on urgency (severity score).

(a) [4 pts] In terms of Time Complexity, state the primary advantage of using a Priority Queue over a standard FIFO Queue for finding the next patient to treat.

(b) [4 pts] Assume a **Max-Heap**. Insert the following patients in this order (Severity Scores):

$$P1(3), \quad P2(5), \quad P3(2), \quad P4(4)$$

Draw the final heap structure (tree view) below.

(c) [2 pts] Starting from the result in (b), show the **array representation** of the heap after **two** `extract-max` operations have been performed.

(d) [6 pts] The time complexity of Heap operations is determined by the tree structure. Answer the following:

1. What is the **height** of a binary heap with n nodes?
2. In a worst-case `Insert`, how many swaps occur relative to the height?
3. If we used a **Sorted Array** instead, what would be the complexity of `Insert`?

(e) [4 pts] To prevent "starvation" of low-severity patients, the system increases a patient's priority based on wait time. Name or briefly describe **one** mechanism to implement this efficiently.

Problem 2: Lists (Undo/Redo) (20 pts)

A text editor supports Undo/Redo operations.

(a) [6 pts] **Complexity Analysis:**
Suppose the text cursor is positioned in the **middle** of a large document (n characters). The user types a character and then immediately presses `Undo`.

1. State the Time Complexity of this `Undo` for an **Array List**.
2. State the Time Complexity of this `Undo` for a **Linked List**.
3. Briefly explain **why** the Array List has this complexity:

(b) [4 pts] Draw the logical structure of the list after the following sequence. **Clearly indicate where the "Current" pointer is located.**
Sequence: `Type A` → `Type B` → `Type C` → `Undo` → `Undo` → `Type D`

(c) [4 pts] A user performs an `Undo`, and then immediately types a new character. What happens to the "Redo" history? (Answer in 1 sentence).

(d) [6 pts] Compare the Time Complexity in the table below:

Operation	Linked List	Array List
Insert at Current Pointer		
Access i-th element		

Problem 3: Decision Trees (20 pts)

A Decision Tree is a flowchart used to make decisions. It has three main parts:
Decision Nodes: *Boxes that ask a question about the data (e.g., "Is Income High?").*
Branches: *Arrows that connect nodes, representing the answer (e.g., "Yes" or "No").*
Leaf Nodes: *The final endpoints that give the prediction (e.g., "Approve" or "Deny").*
You are using a Decision Tree to automate loan approvals. Start at the top (Root). Answer the question, follow the matching branch, and repeat until you reach a Leaf.

(a) [4 pts] Which specific part of the tree contains the **final answer/prediction**?

(b) [5 pts] Consider the following simple logic tree:

- **Root Question:** "Is Income High?"
  - If **No**: Go to Result → [**Deny Loan**]
  - If **Yes**: Go to Next Question → "Has Criminal Record?"
    - If **Yes**: Go to Result → [**Deny Loan**]
    - If **No**: Go to Result → [**Approve Loan**]

Task: A new customer arrives with **High Income** but has a **Criminal Record**. Trace their path. What is the final decision?

(c) [5 pts] The algorithm must choose the best question to ask. Compare these two questions for a group of 4 people (2 Good Credit, 2 Bad Credit):

- **Question A:** Splits group into {2 Good Credit} and {2 Bad Credit}.
- **Question B:** Splits group into {1 Good, 1 Bad} and {1 Good, 1 Bad}.

Which question is better? Briefly explain why using the word "separate" or "mix".

(d) [4 pts] Why is a "short" tree generally preferred over a very "deep" tree for a simple problem?

(e) [2 pts] If a tree is too complex and memorizes the training data perfectly but fails on new customers, what is this error called? .

Problem 4: Sorting & Routing (20 pts)

(a) [4 pts] To minimize maximum lateness in Package delivery logistics, packages are sorted by deadline. What is the name of this scheduling strategy?

(b) [4 pts] You are running **Insertion Sort** on the array: `[45, 10, 30, 25, 60]`. Show the state of the array after the element 10 has been processed and inserted into the sorted portion.

(c) [5 pts] Select the most appropriate sorting algorithm for the following scenarios:

1. The data is already **nearly sorted**.
2. The dataset is huge (cannot fit in RAM).

(d) [4 pts] The road network has **non-negative** travel times. Name the standard shortest-path algorithm used here.

(e) [3 pts] If the road network has **negative** edge weights (e.g., time gains/discounts), which algorithm must be used instead?

Problem 5: Algorithm Analysis (20 pts)

You are analyzing an algorithm called "Iterative Centroid Clustering" (commonly known as K-Means). The algorithm takes a set of data points and k starting points called "Centroids." It groups the data by repeating these two steps:

1. **Assignment Step:** For every data point, calculate its distance to each Centroid. Assign the point to the logical group of the **closest** Centroid.
2. **Update Step:** For each logical group, calculate the **average** (mean) position of all points assigned to that group. Move the Centroid to this new average position.

Termination Rule: The algorithm stops repeating these steps when the Centroid positions no longer change (i.e., the new position is identical to the old position).

Data Provided for Analysis:

- **Data Points:** $A(1,1)$, $B(2,1)$, $C(10,10)$, $D(11,11)$.
- **Initial Centroids:** Centroid 1 is at $C_1(1,0)$. Centroid 2 is at $C_2(10,11)$.

Questions:

(a) [4 pts] Based on the algorithm description above, is this a *supervised* learning method (where the user tells the algorithm the correct answer beforehand) or an *unsupervised* method (where the algorithm discovers the structure on its own)? Explain your answer in one sentence.

(b) [4 pts] **Execution (Assignment Step):** Calculate the squared Euclidean distance from Point $A(1,1)$ to C_1 and from Point $A(1,1)$ to C_2. Based on these distances, which Centroid is Point A assigned to?
Hint: Squared Distance $= (x_1 - x_2)^2 + (y_1 - y_2)^2$

(c) [5 pts] **Execution (Update Step):** Assume the Assignment Step resulted in the following groups:

- Group 1 contains points $\{A(1,1), B(2,1)\}$.
- Group 2 contains points $\{C(10,10), D(11,11)\}$.

Calculate the (x, y) coordinates of the **New Centroids** using the "Update Step" rule.

(d) [4 pts] According to the "Termination Rule" in the text, how will the computer know when to stop running the loop?

(e) [3 pts] Look at the mathematical definition of the **Update Step** (calculating an average). What mathematical error would occur if a Centroid ends up with **zero** points assigned to it?