

Algorithm Middle Exam.

107.11.16

/* 答案紙 請寫上：學號 姓名 ；試卷共 9 頁 總分 100

1. (6%) What is the definition of Algorithm? (三個條件)

有限步驟, 有效率計算, 有終止

2. (4%) Design a Turing to verify "Can 30 be divided by 8?" (執行過程要寫出來 !
(2%))

3. (6%) The following statements are the merge-sort,

Merge-sort (A, p, r)

1. $q \leftarrow \lfloor (p+r)/2 \rfloor$ $O(n)$

2. Merge-sort (A, p, q)

 $O(n)$

3. Merge-sort (A, q+1, r)

 $T(\frac{n}{2})$

4. Merge (A, p, q, r) /* 將兩組排列好的序列合併成一排列好的序列

(1) 改正程式錯誤

 $O(n)$

(2) Use the following sequence of 8 data: 3, 1, 7, 5, 6, 2, 4, 8 to illustrate each step of merge-sort. (正確的演算法)

(3) And analyze the time complex of the algorithm.

✓ 4. (6%) 解 $T(n) = 2T(n/2) + n \lg n$ $n=2^k, k=\lg n$ 5. (8%) 我們可以用 Divide-and-Conquer 方法, 將兩個 $n \times n$ 矩陣相乘, 分成 4 個 $n/2 \times n/2$ 小矩陣相乘後合併, 如下:

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \quad B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}, \quad C = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix},$$

so that we rewrite the equation $C = A \cdot B$ as

$$\begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \cdot \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}.$$

Equation (4.10) corresponds to the four equations

$$C_{11} = A_{11} \cdot B_{11} + A_{12} \cdot B_{21},$$

$$C_{12} = A_{11} \cdot B_{12} + A_{12} \cdot B_{22},$$

$$C_{21} = A_{21} \cdot B_{11} + A_{22} \cdot B_{21},$$

$$C_{22} = A_{21} \cdot B_{12} + A_{22} \cdot B_{22}.$$

將會有 8 個小矩陣相乘與 4 個 $n^2/4$ 加法。若採 Strassen's method 則拆成 7 個小矩陣相乘與 18 個 $n^2/4$ 加法, 請問, 優勢在哪? (列出分析過程才有分數)

6. (12%) (1) (4%) 何謂 max-heap? 此種資料結構有何優勢? 可應用在哪裡?
- (2) (4%) 在 BUILD-MAX-HEAP(A) 副程式中, 為何第 2 行程是要為 $\text{for } i \leftarrow \lfloor \text{length}[A]/2 \rfloor \text{ downto } 1$? 改為 $\text{for } i \leftarrow 1 \text{ to } \lfloor \text{length}[A]/2 \rfloor$ 可以嗎? 此副程式時間複雜度為多少? 推導過程
- (3) (4%) Illustrate each step of the heap sort by using the following example $A[6] = \{3, 2, 9, 6, 4, 8\}$. 說明: 1. 初始樹; 2. 建好 heap tree; 3. Heap sort 前兩筆資料排序步驟

HEAPSORT(A, n)

BUILD-MAX-HEAP(A, n)

for $i \leftarrow n$ downto 2

do exchange $A[1] \leftrightarrow A[i]$

MAX-HEAPIFY(A, 1, $i - 1$)

BUILD-MAX-HEAP(A)

1 heap-size[A] \leftarrow length[A]

2 for $i \leftarrow \lfloor \text{length}[A]/2 \rfloor$ downto 1

3 do MAX-HEAPIFY(A, i)

MAX-HEAPIFY(A, i)

1 $l \leftarrow \text{LEFT}(i)$

2 $r \leftarrow \text{RIGHT}(i)$

3 if $l \leq \text{heap-size}[A]$ and $A[l] > A[i]$

4 then largest $\leftarrow l$

5 else largest $\leftarrow i$

6 if $r \leq \text{heap-size}[A]$ and $A[r] > A[\text{largest}]$

7 then largest $\leftarrow r$

8 if largest $\neq i$

9 then exchange $A[i] \leftrightarrow A[\text{largest}]$

10 MAX-HEAPIFY(A, largest)

$$T(n) = T\left(\frac{n}{2}\right) + cn$$

$$T\left(\frac{n}{2}\right) = \left(T\left(\frac{n}{2^2}\right) + \frac{cn}{2}\right) + cn$$

$$= T\left(\frac{n}{2^2}\right) + \frac{1}{2}cn + cn$$

$$= \left(T\left(\frac{n}{2^3}\right) + \frac{cn}{2^2}\right) + \frac{1}{2}cn + cn$$

$$= T\left(\frac{n}{2^3}\right) + \frac{1}{2^2}cn + \frac{1}{2}cn + \frac{1}{2}cn$$

$$= T\left(\frac{n}{2^k}\right) + \left[\sum_{i=0}^{k-1} \left(\frac{1}{2^i}\right)\right]cn$$

$$= T\left(\frac{n}{2^k}\right) + \left[\frac{1}{\frac{1}{2}}\right]cn$$

$$= T(1) + cn = O(n)$$

$$k=1 \quad k=2 \quad k=3$$

$$\frac{1}{2^0} \quad \frac{1}{2^0} + \frac{1}{2^1} \quad \frac{1}{2^0} + \frac{1}{2^1} + \frac{1}{2^2}$$

7. (4%) 試寫一個產生亂數(隨機)序列演算法之虛擬碼?

RANDOMIZE-IN-PLACE(A)

1 $n \leftarrow \text{length}[A]$

2 $i \leftarrow 1$ to n

3 do swap $A[i] \leftrightarrow A[\text{RANDOM}(i, n)]$

8. (8%) The following algorithm is counting sort. If array A is following:

For $i \leftarrow 0$ to k

Do $C[i] = 0$

For $j \leftarrow 1$ to n

Do $C[A[j]] \leftarrow C[A[j]] + 1$

For $i \leftarrow 1$ to k

Do $C[i] = C[i] + C[i - 1]$

For $j \leftarrow n$ downto 1

Do $B[C[A[j]]] \leftarrow A[j]$

$C[A[j]] \leftarrow C[A[j]] - 1$

若 $A[] = \{3, 2, 1, 5, 3, 2, 2, 1\}$

(1) (4%) Please write the initial state and each state of array A & B & C in the step of algorithm.

(2) (4%) 為何此種方法又稱為 stable sorting? 是因為哪一行程式造成此現象?

9. (6%) 下列為一個 $O(n)$ 的 selection algorithm

SELECT algorithm:

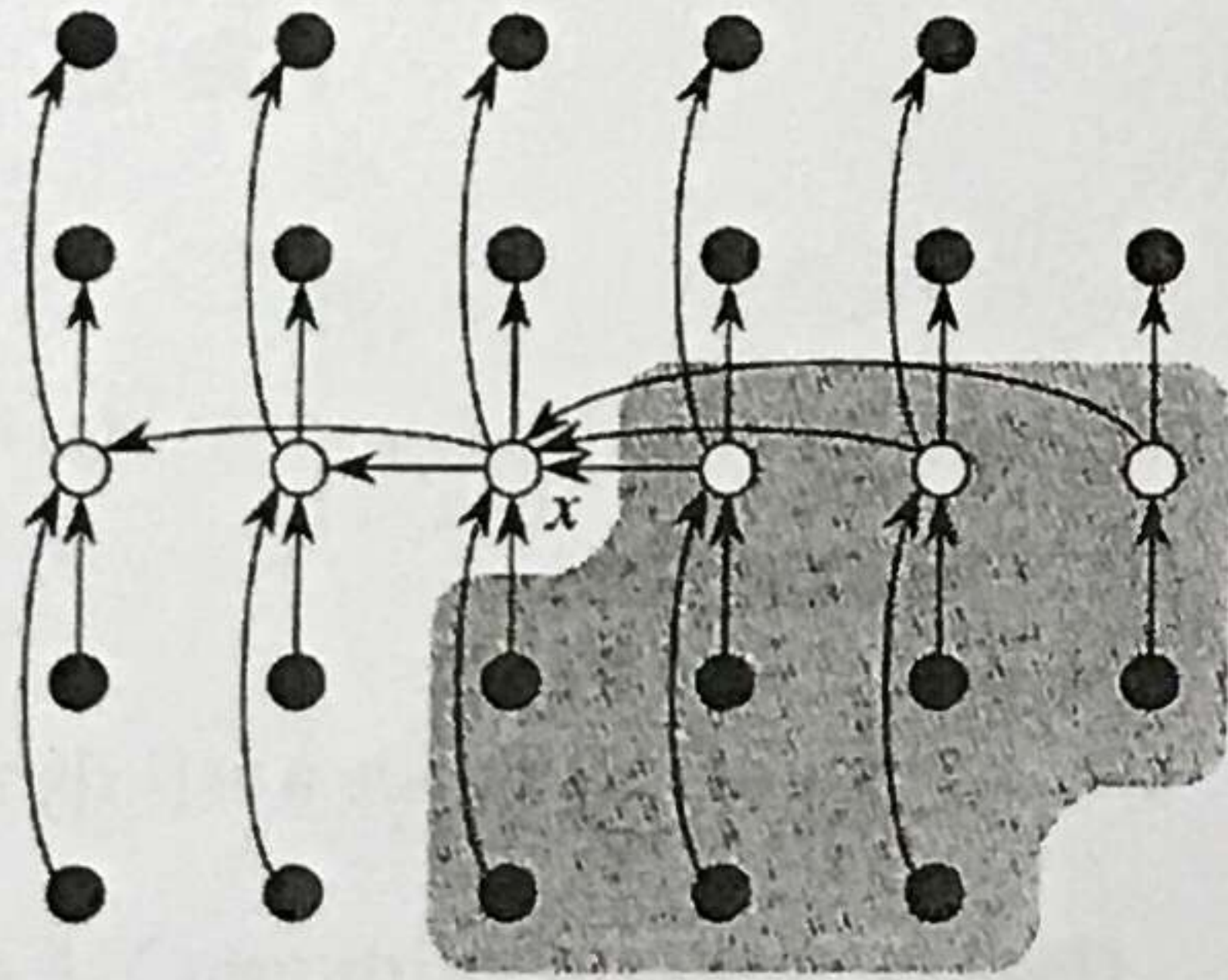
1. Divide the n elements of the input array into $\lfloor n/5 \rfloor$ groups of 5 elements each and at most one group made up of the remaining $n \bmod 5$ elements. $O(n)$

2. Find the **median** of each of the $\lfloor n/5 \rfloor$ groups by the first insertion sorting the elements of each group and then picking the median from the sorted list of group elements. $O(n)$

3. Use the **SELECT** recursively to find the median x of the $\lfloor n/5 \rfloor$ medians found in step 2. $T(\lfloor \frac{n}{5} \rfloor)$

4. Partition the input array around the median-of-medians x using the modified version of **PARTITION**. (以 x 來分(x 為第 k 小), 前段 $k-1$ 個, 後半段 $n-k$ 個) $O(n)$

5. If $i = k$, then return x . Otherwise, use **SELECT** recursively to find the i th smallest on the low side if $i < k$, or $(i - k)$ th smallest elements on the high side if $i > k$. $T(\frac{7}{10}n + b)$



(1) (2%) 請說明步驟 4 所選出 x , 其大小排序位置範圍為? (以 n 筆資料中, x 介在: $R1 \leq x \leq R2$, $R1, R2 = ?$)

(2) (4%) 設此方法時間為 $T(n)$, 分析此方法, 單獨每一步驟之時間複雜度? 與總時間複雜度?

10. (20%) The matrix-chain multiplication problem is: given a chain $\langle A_1, A_2, \dots, A_n \rangle$ of n matrices, where for $i = 1, 2, \dots, n$, fully parenthesize the product

A_1, A_2, \dots, A_n in a way that minimizes the number of scalar multiplications.

Use DP to solve this problem is as follows.

Step 1: The structure of an optimal parenthesization(括號)

Suppose that an optimal parenthesization of $A_i A_{i+1} \dots A_j$ splits the product between A_k and A_{k+1} . Thus, we can build an optimal solution to an instance of the matrix-chain multiplication problem by splitting the problem into two subproblems, finding optimal solutions to subproblem instances, and then combining these optimal subproblem solutions.

Step 2: A recursive solution

We define the cost of an optimal solution recursively in terms of the optimal solutions to subproblems. Let $m[i, j]$ be the number of scalar multiplications needed to compute the matrix $A_{i..j}$; for the full problem, the cost of a cheapest way $A_{1..n}$ is $m[1, n]$.

A dimension of matrix A_i is $P_{i-1} \times P_i$; the matrix product $A_{i..k} A_{k+1..j}$ is $P_{i-1} \times P_k \times P_j$. Then,

$$m[i, j] = m[i, k] + m[k + 1, j] + p_{i-1} p_k p_j .$$

一般式：

$$m[i, j] = \begin{cases} 0 & \text{if } i = j , \\ \min_{i \leq k < j} \{m[i, k] + m[k + 1, j] + p_{i-1} p_k p_j\} & \text{if } i < j . \end{cases}$$

Let $s[i, j] = k$ such that $m[i, j] = m[i, k] + m[k + 1, j] + p_{i-1} p_k p_j$.

Step 3. Computing the optimal costs

We can use a tabular, bottom up approach, compute the optimal cost. For a matrix A_i has **dimension** $P_{i-1} \times P_i$. The input sequence $\langle P_0, P_1, \dots, P_n \rangle$, where **length** $[P] = n+1$.

We use a **table** $m[1..n, 1..n]$ for storing $m[i, j]$ and a table $s[1..n, 1..n]$ for recording the index k achieve the optimal cost in computing $m[i, j]$.

The procedure could be:

$$m[2, 5] = \min \begin{cases} m[2, 2] + m[3, 5] + p_1 p_2 p_5 = 0 + 2500 + 35 \cdot 15 \cdot 20 = 13000, \\ m[2, 3] + m[4, 5] + p_1 p_3 p_5 = 2625 + 1000 + 35 \cdot 5 \cdot 20 = 7125, \\ m[2, 4] + m[5, 5] + p_1 p_4 p_5 = 4375 + 0 + 35 \cdot 10 \cdot 20 = 11375 \end{cases} = 7125.$$

PRINT-OPTIMAL-PARENS(s, i, j)

```

1  if  $i = j$ 
2      then print "A";
3  else print "("
4      PRINT-OPTIMAL-PARENS( $s, i, s[i, j]$ )
5      PRINT-OPTIMAL-PARENS( $s, s[i, j] + 1, j$ )
6      print ")"

```

(1) (8 %) 此問題所有括號位置有幾種可能？(用 order 表示)。請說明 Step 3 演算法中，line 4, 5, 6, 8 之時間複雜度？

(2) (12%) 若 $A_1(10, 5), A_2(5, 10), A_3(10, 10), A_4(10, 5), A_5(5, 10)$

求 Matrix $m(i, j)$ & $s(i, j)$ & 最佳解 (如何括號乘？最少乘法次數？)為何？

11. (20%) We want to construct a binary search tree whose expected search cost is smallest. We call such a tree an *optimal binary search tree*.

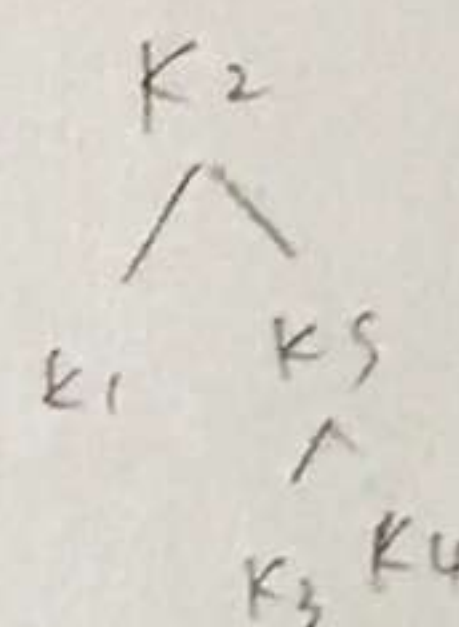
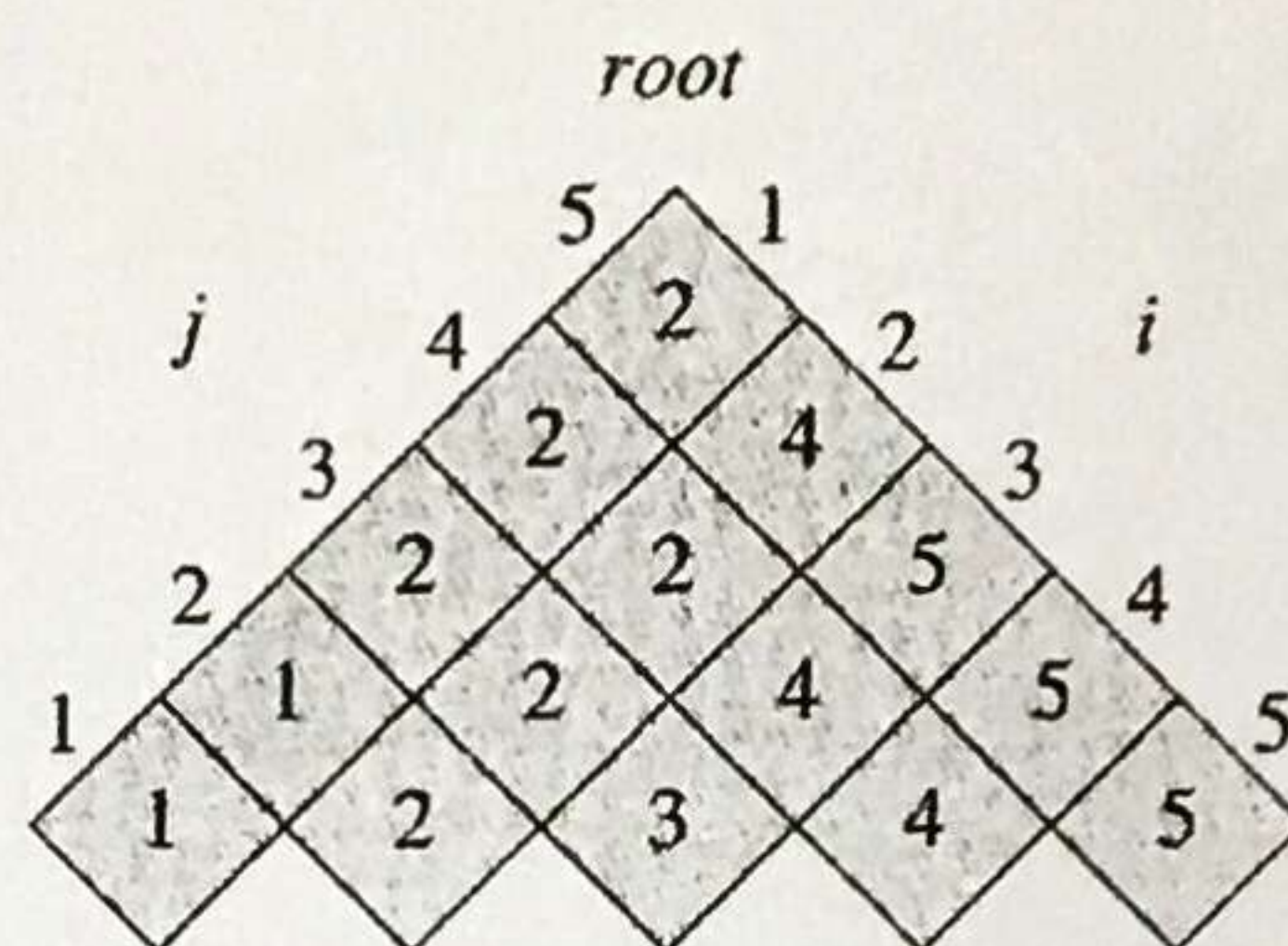
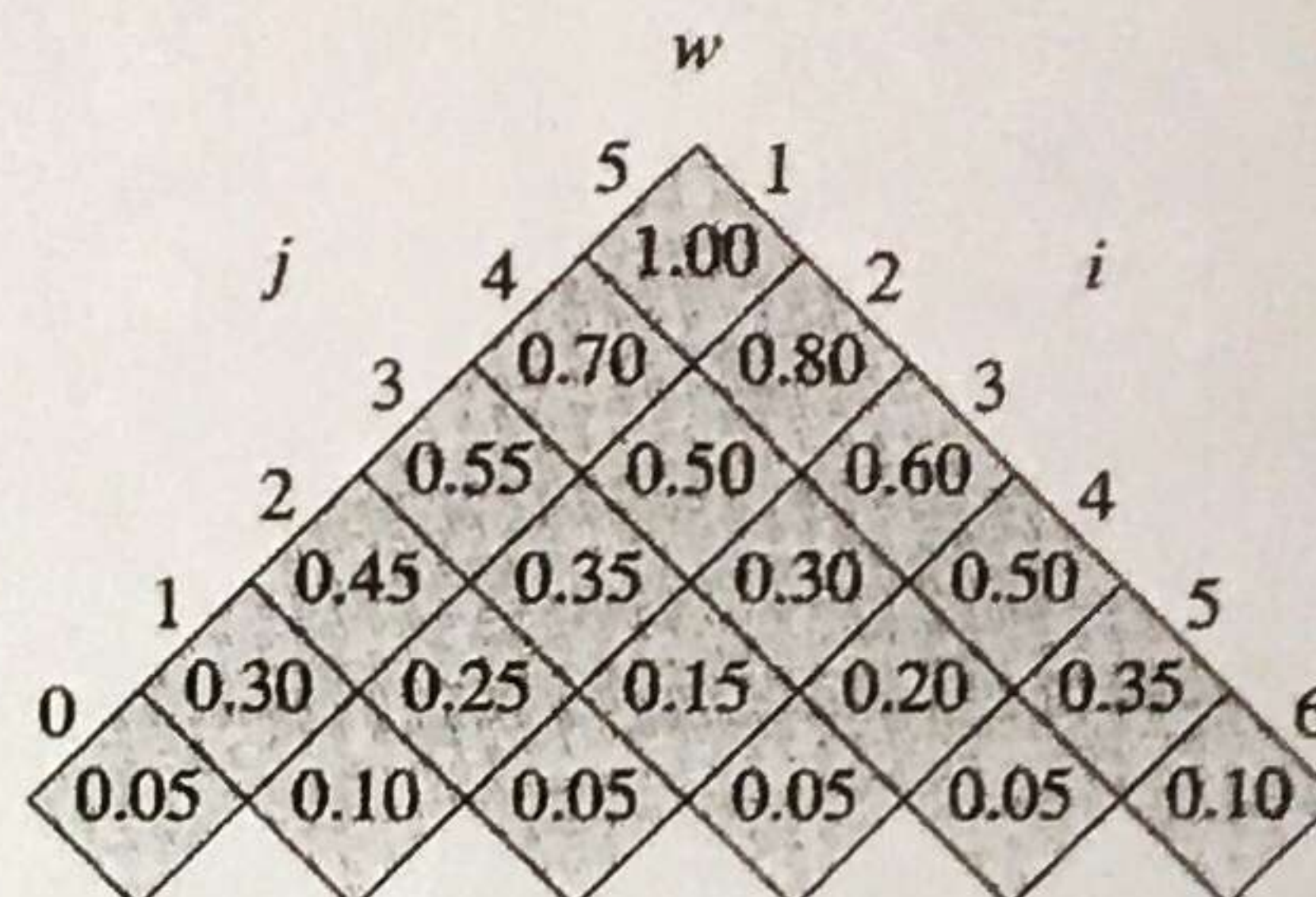
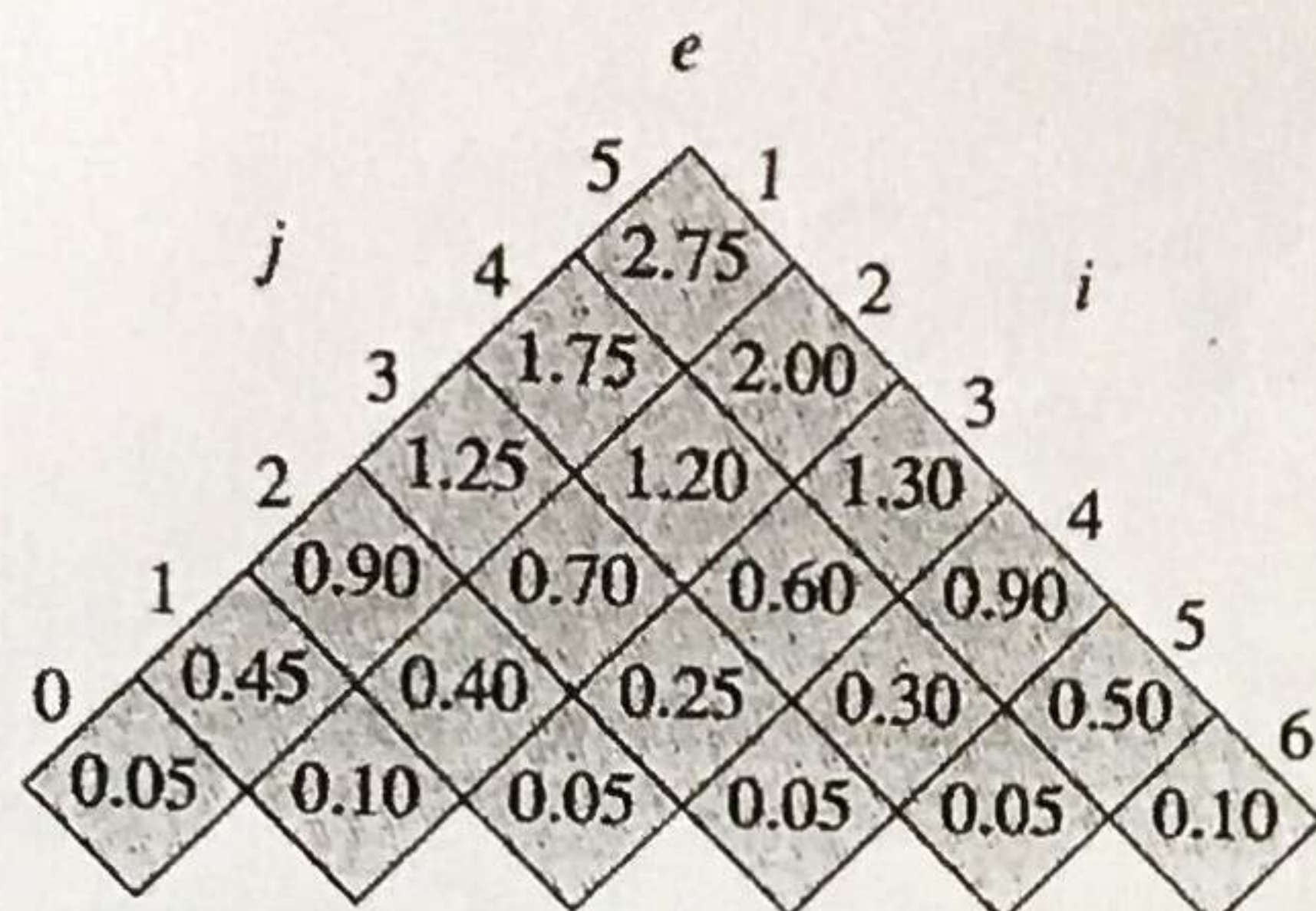
OPTIMAL-BST(p, q, n)

```

1  for  $i \leftarrow 1$  to  $n + 1$ 
2      do  $e[i, i - 1] \leftarrow q_{i-1}$ 
3      do  $w[i, i - 1] \leftarrow q_{i-1}$ 
4  for  $l \leftarrow 1$  to  $n$ 
5      do for  $i \leftarrow 1$  to  $n - l + 1$ 
6          do  $j \leftarrow i + l - 1$ 
7               $e[i, j] \leftarrow \infty$ 
8               $w[i, j] \leftarrow w[i, j - 1] + p_j + q_j$ 
9              for  $r \leftarrow i$  to  $j$ 
10                 do  $t \leftarrow e[i, r - 1] + e[r + 1, j] + w[i, j]$ 
11                     if  $t < e[i, j]$ 
12                         then  $e[i, j] \leftarrow t$ 
13                              $root[i, j] \leftarrow r$ 
14  return  $e$  and  $root$ 

```

注意虛線框，老師有可能改掉。
以考試的虛線框為全盤參考



Print-optimal-BST($root, i, j$)

1. if $i \geq j$ then print "root", $root(i, j)$
2. print "left root", Print-optimal-BST ($i, root(i, j)$)
3. print "right root", Print-optimal-BST ($root(i, j)+1, j$)
4. Else return.

現有 5 筆資料，每筆資料被讀取機率為 p_i ，落在資料間隔機率為 q_i ，如下表：

i	0	1	2	3	4	5
p_i		0.1	0.15	0.15	0.1	0.05
q_i	0.1	0.05	0.05	0.1	0.1	0.05

求 $e(i, j)$, $w(i, j)$, $root(i, j)$ 與 optimal binary search tree 為何？



考試科目：

演算法

班

別：

數位三

學號：

510333020

姓

名：

鄭家欣

1. 有限的步驟, 有效率的計算, 有終止性

2. $\begin{array}{r} 2 \overline{) 30} \ 0 \\ 2 \overline{) 15} \ 1 \\ 2 \overline{) 7} \ 1 \\ 2 \overline{) 3} \ 1 \\ 1 \end{array}$

 $Q = \{q_0, q_1, q_2, q_3, q_r, q_n\}$ q_0 : initial state $\Gamma = \{0, 1, b\}$ $I \subseteq \Gamma$ b : blank, in $\Gamma - I$ q_r : Yes state q_n : No state

	0	1	b
q_0	$(q_0, 0, R)$	$(q_0, 1, R)$	(q_0, b, L)
q_1	$(q_2, 0, R)$	$(q_n, 1, S)$	
q_2	$(q_3, 0, R)$	$(q_n, 1, S)$	(q_n, b, S)
q_3	$(q_r, 0, S)$	$(q_n, 1, S)$	(q_n, b, S)

Ans: No, $q_0 \rightarrow q_1 \rightarrow q_2 \rightarrow q_n$
(執行這過程) 30 can't be divided by 8.

3. (1) Merge-sort(A, P, r)
if $p < r$
2 then $q \leftarrow \lfloor (p+r)/2 \rfloor$ $O(1)$
3 Merge-sort(A, P, q) $O(1)$
4 Merge-sort(A, q+1, r) $T(\frac{n}{2})$
5 Merge(A, P, q, r) $O(n)$

(3) $T(n) = T(\lceil \frac{n}{2} \rceil) + T(\lfloor \frac{n}{2} \rfloor) + cn, n \geq 2$

$$= 2T(\frac{n}{2}) + cn$$

$$= 2[2T(\frac{n}{2^2}) + \frac{cn}{2}] + cn$$

$$= 2^2 T(\frac{n}{2^2}) + 2cn$$

$$= 2^2 [2T(\frac{n}{2^3}) + \frac{cn}{2^2}] + 2cn$$

Let $n=2^k$ $= 2^3 T(\frac{n}{2^3}) + 3cn$

$k = \lg n$

$= 2^k T(\frac{n}{2^k}) + kcn = nT(1) + \lg n \cdot cn, \therefore T(n) = O(n \lg n)$

3. (2) $[3, 1, 7, 5, 6, 2, 4, 8]$ 4. $T(n) = 2T(\frac{n}{2}) + n \lg n$

$$= 2[2T(\frac{n}{2^2}) + \frac{n}{2} \lg \frac{n}{2}] + n \lg n$$

$$= 2^2 T(\frac{n}{2^2}) + n \lg n - n \lg 2 + n \lg n = 2^2 T(\frac{n}{2^2}) + 2n \lg n - n$$

$$= 2^2 [2T(\frac{n}{2^3}) + \frac{n}{2^2} \lg \frac{n}{2^2}] + 2n \lg n - n$$

$$= 2^3 T(\frac{n}{2^3}) + n \lg n - n \lg 2^2 + 2n \lg n - n$$

$$= 2^3 T(\frac{n}{2^3}) + 3n \lg n - 3n$$

Let $n=2^k$
 $k = \lg n$

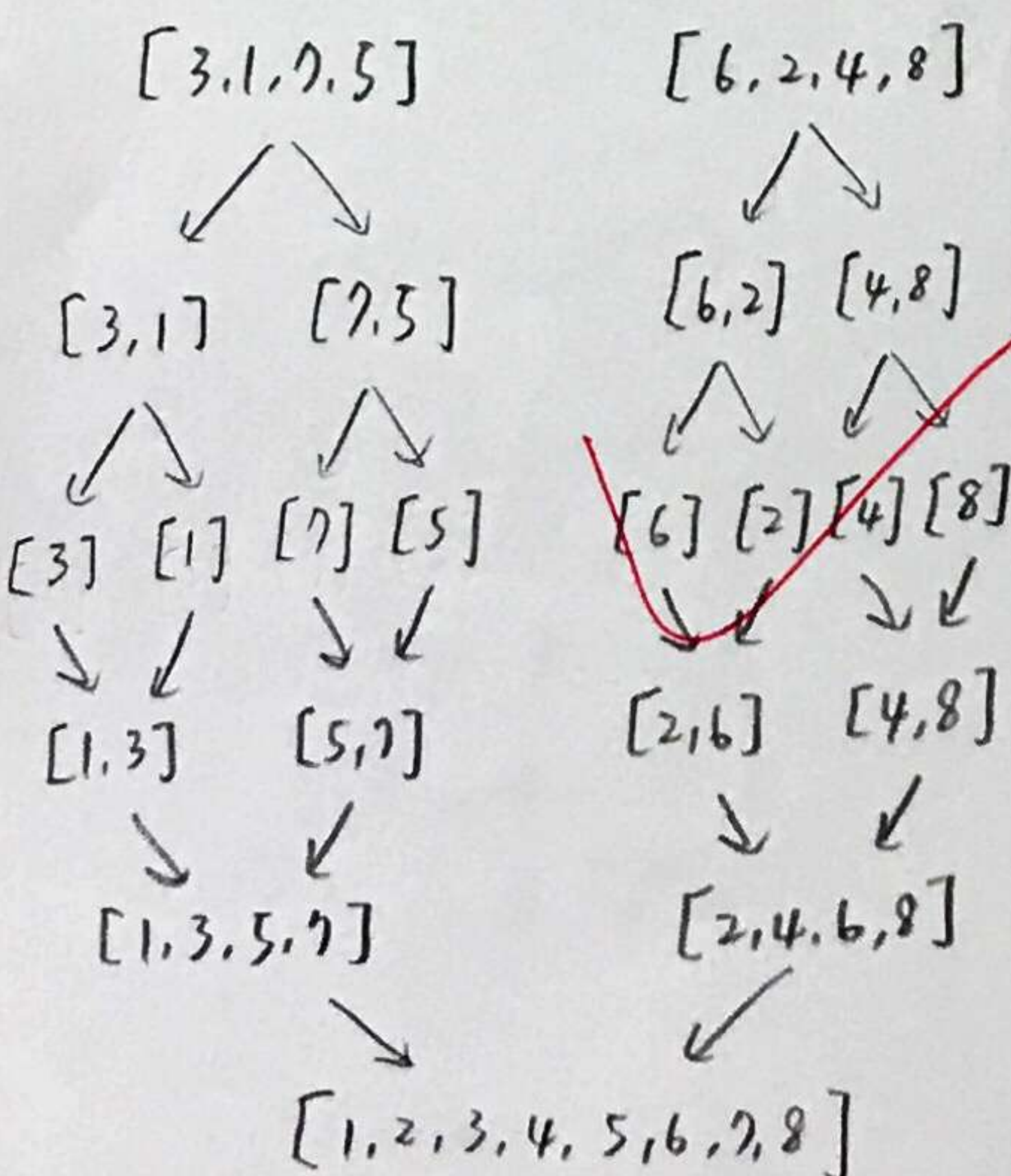
$$= 2^k T(\frac{n}{2^k}) + k n \lg n - \left[\frac{k(k+1)}{2} \right] n$$

$$= n T(1) + \lg n \cdot n \lg n - \left[\frac{\lg^2 n + \lg n}{2} \right] n$$

$$= cn + \frac{1}{2} n \lg^2 n - \frac{1}{2} n \lg n$$

$$= cn + \frac{1}{2} n \lg^2 n + \frac{1}{2} n \lg n$$

$$\therefore T(n) = O(n \lg^2 n)$$



5.
(8%) $T(n) = O(1) + 8T(\frac{n}{2}) + O(n^2)$

$= 8T(\frac{n}{2}) + O(n^2)$

(8个乘法, 4个 $\frac{n^2}{4}$ 加法 = $O(n^2)$)

$T(n) = O(n^3)$ (此時, divide-and-conquer 並未節省時間)

* 改進:

Strassen's method

$T(n) = \begin{cases} O(1) & \text{if } n=1 \\ 7T(\frac{n}{2}) + O(n^2) & \text{if } n>1 \end{cases}$

(18个加法)

$T(n) = \dots O(n^{\log_2 7}) = O(n^{2.81})$

6. (1) heap tree: 幾乎是完整二元樹 (只有最後一層沒滿), 資料插入、取出最大值等, 只需要 $O(\log n)$, 可用於多作業系統中。
max-heap 是最大堆積, 能將最大值堆積到 root, 也就是 index 1 的位置。工作排程管理。

(2) ① 要從子樹開始檢查回 root 才能把最大值換到根, \therefore 不行改為 for $i \leftarrow 1$ to $\lfloor \text{length}(A)/2 \rfloor$
② 若從 index 1 開始往下檢查, 有可能最大值就不在 root, 只能保證父節點都比子節點大

(2) ③ 高度 h , y 取為 $\frac{n}{2^{h+1}}$, \therefore 總花費時間為:

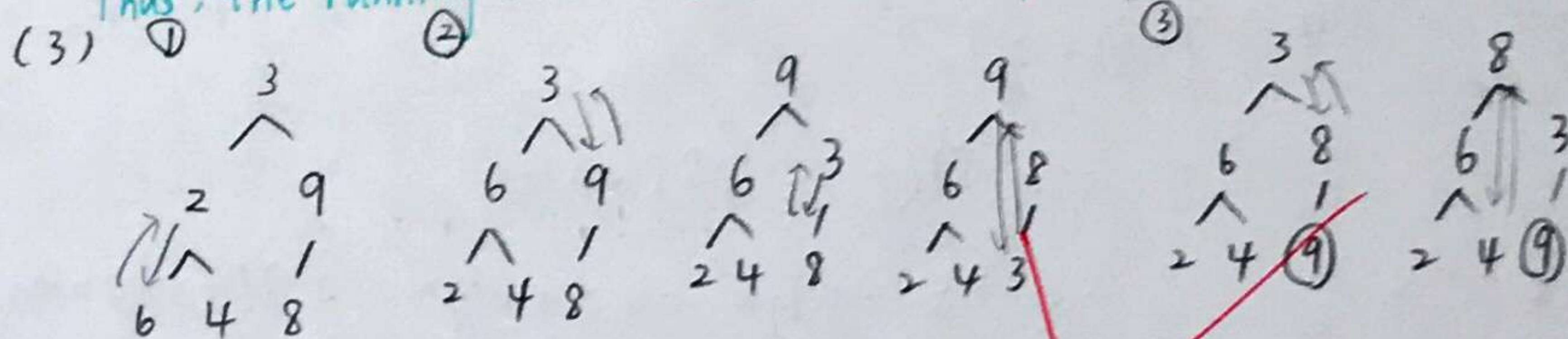
$\sum_{h=0}^{\lfloor \log n \rfloor} \lceil \frac{n}{2^{h+1}} \rceil O(h) = O(n \sum_{h=0}^{\lfloor \log n \rfloor} \frac{h}{2^h})$

$O(n \sum_{h=0}^{\lfloor \log n \rfloor} \frac{h}{2^h}) = O(n \sum_{h=0}^{\infty} \frac{h}{2^h})$
 $= O(n)$

The last summation can be evaluated by substituting $x = \frac{1}{2}$ in the formula (A.8), which yields.

$\sum_{h=0}^{\infty} \frac{h}{2^h} = \frac{\frac{1}{2}}{(1-\frac{1}{2})^2} = 2$

Thus, the running time of Build-Max-Heap can be bounded as



7. RANDOMIZE-IN-PLACE (A)

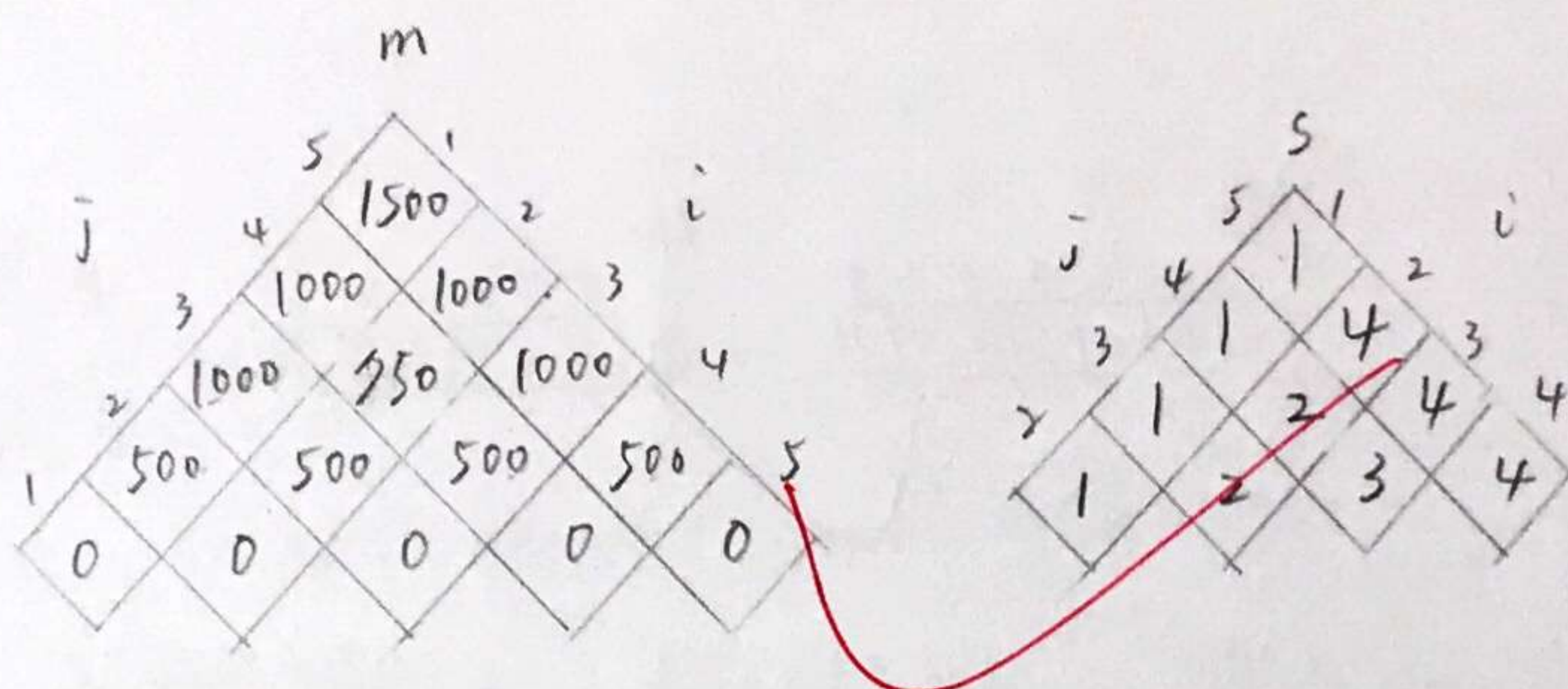
1 $n \leftarrow \text{length}[A]$

2 $i \leftarrow 1$ to n

3 do swap $A[i] \leftrightarrow A[\text{RANDOM}(i, n)]$

10, (2)

- $A_1(10, 5)$
- $A_2(5, 10)$
- $A_3(10, 10)$
- $A_4(10, 5)$
- $A_5(5, 10)$



$$m[1,2] = m[1,1] + m[2,2] + 10 \times 5 \times 10 = 500$$

$$m[2,3] = m[2,2] + m[3,3] + 5 \times 10 \times 10 = 500$$

$$m[3,4] = m[3,3] + m[4,4] + 10 \times 10 \times 5 = 500$$

$$m[4,5] = m[4,4] + m[5,5] + 10 \times 5 \times 10 = 500$$

$$m[1,3] = m[1,1] + m[2,3] + 10 \times 5 \times 10 = 0 + 500 + 500 = 1000$$

$$= m[1,2] + m[3,3] + 10 \times 10 \times 10 = 500 + 0 + 1000 = 1500$$

$$m[2,4] = m[2,2] + m[3,4] + 5 \times 10 \times 5 = 0 + 500 + 250 = 750$$

$$= m[2,3] + m[4,4] + 5 \times 10 \times 5 = 500 + 0 + 250 = 750$$

$$m[3,5] = m[3,3] + m[4,5] + 10 \times 10 \times 10 = 0 + 500 + 1000 = 1500$$

$$= m[3,4] + m[5,5] + 10 \times 5 \times 10 = 500 + 0 + 500 = 1000$$

$$m[1,4] = m[1,1] + m[2,4] + 10 \times 5 \times 5 = 0 + 750 + 250 = 1000$$

$$= m[1,2] + m[3,4] + 10 \times 10 \times 5 = 500 + 500 + 500 = 1500$$

$$= m[1,3] + m[4,4] + 10 \times 10 \times 5 = 1000 + 0 + 500 = 1500$$

$$m[2,5] = m[2,2] + m[3,5] + 5 \times 10 \times 10 = 0 + 1000 + 500 = 1500$$

$$= m[2,3] + m[4,5] + 5 \times 10 \times 10 = 500 + 500 + 500 = 1500$$

$$= m[2,4] + m[5,5] + 5 \times 5 \times 10 = 750 + 0 + 250 = 1000$$

$$m[1,5] = m[1,1] + m[2,5] + 10 \times 5 \times 10 = 0 + 1000 + 500 = 1500$$

$$= m[1,2] + m[3,5] + 10 \times 10 \times 10 = 500 + 1000 + 1000 = 2500$$

$$= m[1,3] + m[4,5] + 10 \times 10 \times 10 = 1000 + 500 + 1000 = 2500$$

$$= m[1,4] + m[5,5] + 10 \times 5 \times 10 = 1000 + 0 + 500 = 1500$$

$$(A_1((A_2(A_3 A_4)) A_5)), 1500 = \text{?}$$

