國立陽明交通大學 112 學年度碩士班考試入學招生試題

科目：資料結構與演算法(1101)　　　　考試日期：112 年 2 月 6 日 第 1 節

系所班別：資訊聯招　　　　

【不可使用計算機】＊作答前請先核對試題、答案卷(試卷)與准考證之所組別與考科是否相符！！

Part I - Multiple choice questions (選擇題 (單選/多選)，此部份使用答案卡作答)

For questions 1 to 16, if a question has multiple answers, then a correct response needs to contain all of them.

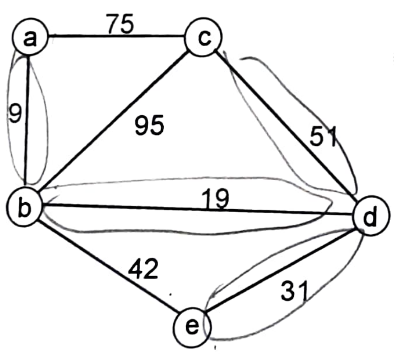
1. (3%) Insert the following nodes 4, 5, 6, 1, 2, 3, 7 in sequence into an AVL tree. Which of the following statement(s) is (are) true?

(A) Node of 4 is the root node.
(B) Node of 5 and Node 7 are at the same level.
(C) Node of 6 is the ancestor of Node 7.
(D) Node of 7 has no sibling.

2. (3%) Given the resultant time complexity of T(n) in big-O notation, where in all cases, T(1) = 1, which of the following statement(s) is (are) true?

(A) Let $T(n) = 2T(n-1) + O(1)$. Then $T(n) = O(2^n)$
(B) Let $T(n) = T(n/2) + O(1)$. Then $T(n) = O(\log(n))$
(C) Let $T(n) = T(n-1) + O(n)$. Then $T(n) = O(n^2)$
(D) Let $T(n) = 2T(n/2) + O(n)$. Then $T(n) = O(n \log(n))$

3. (3%) Please refer to the following graph and solve the minimum spanning tree (MST) using **Kruskal's algorithm**. Which of the following statement(s) is(are) true?



(A) The sum of the weight of all edges in the MST is 110.
(B) Node 'a' connects to node e through node 'b' and 'd' in the MST.
(C) The sum of the weight of all edges connecting node 'd' in MST is 50.
(D) There are 4 edges in total in the MST.

4. (3%) Which of the following statement(s) about Kruskal algorithm and Prim's algorithm is(are) correct?

(A) Kruskal algorithm's concentration is on vertices.
(B) Prim's algorithm is better when there are many more edges than vertices.
(C) Kruskal algorithm can't use on negative-weighted-undirected graph.
(D) For Prim's algorithm, choose different vertex as start vertex may get different total weight of the minimum spanning tree.

5. (3%) Insert the following nodes {6,15,3,9,4,10,8} in order into a binary search tree. Which of the following is(are) correct?

(A) Pre-order: 4,3,8,10,9,15,6
(B) In-order: 3,4,6,8,9,10,15
(C) Finding the node of 8 will go through three edges
(D) The nodes of 4 and 9 are at the same level

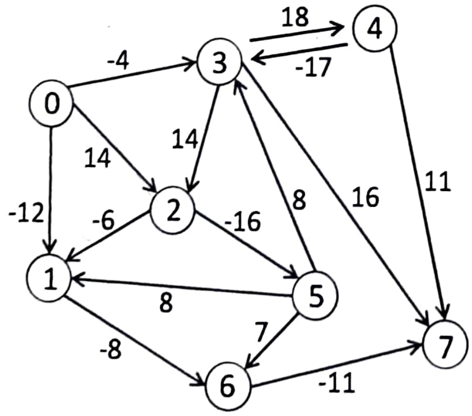
6. (3%) The hash function [h(k) = k mod m] is used to map an integer key to one of the slots, indexed 0, ..., m-1, of a hash table of m slots. Let m = 7. The table is initially empty. After inserting the keys 8, 12, 3, 11, 1, 2 into the table in the given order, which of the following is(are) correct?

(A) Consider using linear probing to handle collisions, then key 2 will be in the slot indexed 6.
(B) Consider using quadratic probing to handle collisions, i.e., let the i-th probe position for a value k be given by the function $h(k, i) = h(k) + i^2$, then key 2 will be in the slot indexed 2. 6
(C) Consider using the function $h(k, i) = h(k) + i + i^2$ to handle collision, then key 2 will be in the slot indexed 2.
(D) Consider using chaining to handle collision, then key 2 will be in the slot indexed 2.

7. (3%) Which of the following statement(s) is(are) true?

(A) A directed graph can be represented as either using an adjacency matrix or an adjacency list.
(B) Breadth-first search (BFS) and Depth-first search (DFS) are algorithms for traversing a graph or tree. BFS typically uses a queue to store the nodes that need to be visited, while DFS typically uses a stack.
(C) BFS can be used to find the shortest path between two nodes in a graph.
(D) DFS can be used to determine whether a graph is strongly connected and to check whether a graph contains a cycle.

8. (4%) Refer to the directed weighted graph below, and use the Bellman Ford's algorithm to find the shortest paths from node 0 to any other nodes in the graph. Which of the following statement(s) is(are) true?



(A) The total weight of the minimal shortest path is -31.
(B) The total weight of the maximal shortest path is 10.
(C) The total weight of all shortest paths from node 0 to the other nodes is -49.
(D) The total weight of the last second minimal shortest path is -20.

9. (3%) Given the following pseudo code, it sorts the elements of array Y in ascending order and stores the sorted elements to array Z. Y has 2n numbers, n > 0. Let Y[ ] = {1, 2, 3, …, n, 1, 2, 3, …, n}. The array index is zero based. The pseudo code is as follows.

```
k = 0, i = 0, j = n
while i < n and < 2n
   if ( ___L1___ ) {
     ___L2___
     i = i + 1
   } else {
     Z[k] = Y[j]
     j = j + 1
   }
   k= k + 1
end while
while j < 2n
   Z[k] = Y[j]
   k= k + 1, j = j + 1
end while
```

Which of the following instructions should be placed in (a) and (b), respectively?

(A) `L1: Y[i] <  Y[j] ;  L2: Z[ k ] = Y[i+1]`
(B) `L1: Y[i] <= Y[j] ;  L2: Z[ k ] = Y[i]`
(C) `L1: Y[i] <  Y[j] ;  L2: Z[ k ] = Y[i]`
(D) `L1: Y[i] <= Y[j] ;  L2: Z[ k+1 ] = Y[i]`

10. (3%) The following question is about a red-black tree with M nodes, where M >= 3. Node definition is used. For the M nodes, the external and internal nodes are included. The function ceil(x) returns the smallest possible integer value which is greater than or equal to the given argument x. Which of the following statement(s) is(are) true?

(A) There are at least ceil((M−1)/3) red nodes.
(B) There are at least ceil(M / 2) black nodes.
(C) The height difference is at most one between the left and right subtrees of a given internal node.
(D) It is impossible to delete all the black nodes until the root is deleted.

11. (3%) There is a max binary heap with height h, where h >=2. The height of the root is 0. The pop operation is used to remove an element. After 15 elements are stored to the heap, the new height of the heap is H. The function ceil(x) returns the smallest possible integer value which is greater than or equal to the given argument x. Which of the following statement(s) is(are) correct?

(A) It takes at most $(2^{H+1} - 1)$ pop operations to remove all the nodes.
(B) Let G = ceil(H/2). It takes at most 2^G pop operations to remove three levels of the heap.
(C) To remove the second smallest element, it takes N pop operations, where N is inside $[2^{H-1}, 2^H]$.
(D) |H-h| is smaller than or equal to 2.

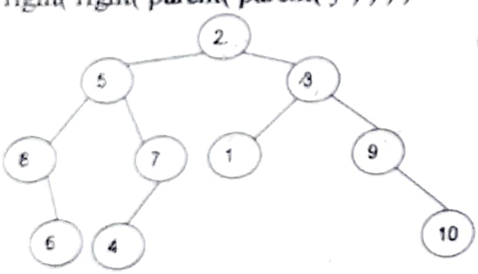
12. (3%) A min binary heap has one element which is 10. Now the following elements are stored to the heap in order: 5, 4, 3, 2, 9, 7, 8, 6. After that, node 5 is deleted. Which of the following statement(s) is(are) correct?

(A) The parent of node 9 is 7 and 8 is the child of 6
(B) Node 4 is not the child of node 5 and 7 is not the parent of node 8
(C) The children of node 3 are nodes 4 and 6
(D) Node 2 does not have a parent and node 8 does not have any children

13. (3%) S is an ordered sequence of numbers. S is formed by two ordered sequences {2n-1, 2n-3, ..., 1} and {2, 4, 6, ..., 2n}. Let S = {(2n-1), 2, (2n-3), 4, (2n-5), 6, ..., 3, (2n-2), 1, 2n}, where n >= 10. The elements of S are stored into an empty stack one by one in the given order. After that the elements are popped from the stack one by one. At the X-th step, (n-1) is popped from the stack while (n+1) is popped from the stack at the Y-th step. Assume that s and t are two different positive elements of S, and s + t = 2n. Which of the following statement(s) is(are) correct?

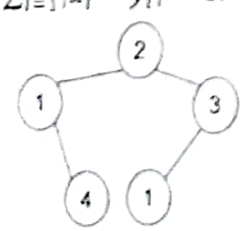
(A) X = n or Y = n.
(B) |Y-X| = 2.
(C) X must be larger than Y for any positive integer n.
(D) It takes at most n pop operations to pop one of the two elements, s and t.

14. (3%) The following question is a about a binary tree. Denote x as a node. The function parent(x) returns the parent of x. Furthermore, left(x) and right(x) return the left child and right child of node x, respectively. w and y are nodes of a binary tree which is shown below. left(parent(parent(right(w)))) is 5. right(right(parent(parent(y)))) is 9. What are w and y?

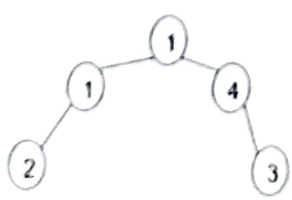


(A) w is 3 and y is 8.
(B) w is 1 and y is 5.
(C) w is 1 and y is 8.
(D) w is 3 and y is 7.

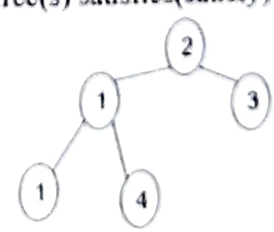
15. (3%) This question considers binary trees with 5 nodes. The preorder and post order traversal results are stored in the ordered lists $\{x_1, x_2, x_3, x_4, x_5\}$ and $\{y_1, y_2, y_3, y_4, y_5\}$, respectively. Furthermore, $\sum_{i=1}^{5}|x_i - y_i| = 6$. Which of the following binary tree(s) satisfies(satisfy) the condition?



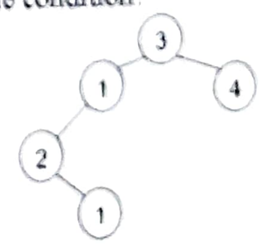
(A)



(B)



(C)



(D)

16. (4%) Initially, array X[ ] = {n, 2, 3, 4, 5, ..., n-1, n}, n >= 4. The array index is zero-based. Insertion sort is used to sort the elements of X in descending order. Given the following pseudo code, how many times is Line L performed?

```
j = 1
while j < n
    b = X[j]
    k = j - 1
    while k >= 0 and X[k] < b
        X[k+1] = X[k]          // Line L
        k = k - 1
    end while
    X[k+1] = b
    j = j + 1
end while
```

(A) n+2
(B) n(n-1)/2
(C) (n-1)(n-2)/2
(D) n^2

Part II – Non-choice questions (非選擇題)

For questions 17 to 26, make sure your answer is written in a neat and tidy way. Otherwise it will not be graded.

17. Consider processing a set of n items. The algorithm consists of two parts, first part takes $O(n \log n)$ time, second part takes $\Theta(n \log n)$ time. What is the total time for processing the set of n items? Give me your reasons? Use the most precise asymptotic notation to present the time required. (2%).

18. **Linear time algorithm** for the Selection problem, (the computational problem to find the kth largest number among n numbers):

(a) Typical implementation is to divide the n numbers into $\lceil \frac{n}{5} \rceil$ groups. We then find the medians of the $\lceil \frac{n}{5} \rceil$ groups, and then find the median of those medians. After $(\Theta(n) + T(\lceil n/5 \rceil))$ time, we can drop a set S of numbers because S does not contain the answer. Show that $|S| \geq n/4$ where $|S|$ is the size of S. (4%).

(b) Can we divide the n numbers into $\lceil \frac{n}{7} \rceil$ groups instead of $\lceil \frac{n}{5} \rceil$ groups? Why or why not? (5%).

19. Given n points on a line, the n points are the n users who need wifi signal. Suppose that we have only one kind of wifi router and we don't care about the number of wifi routers needed, but we have to minimize the power consumption of the routers. Power consumption is presented as the distance that signal of a router can reach.

(a) **Design a Divide and Conquer** algorithm to determine the minimum power consumption. (4%).

(b) What is the time required for your Divide and Conquer algorithm? (2%).

20. Stack operations are

- `Push(item)`: `S.push(item)` pushes `item` into the stack `S`.
- `Pop()`: `S.Pop()` removes top of the stack `S`.
- `isEmpty()`: `S.isEmpty()` answers whether the stack `S` is empty.

The questions are about implementing a FIFO queue `Q` by using two stacks `S1` and `S2`.

(a) (2%). Describe the algorithm `enQueue(item)` (insert `item` into `Q`) using pseudo code.

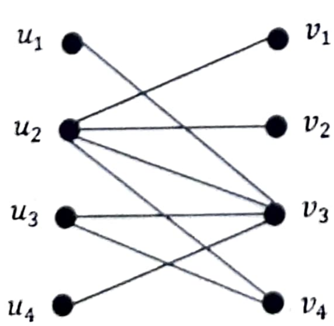
(b) (2%). Describe the algorithm `deQueue()` (delete an item from the `Q`) using pseudo code.

(c) (4%). Show that `deQueue` takes $O(1)$ amortized cost.

21. Let $G = (V, E)$ be an undirected graph. A vertex cover U is a vertex subset such that, any edge in E has at least one endpoint vertex in U.

(a) (2%). Let U be a vertex cover and M be a matching for G. Prove or disprove that, $|M| \leq |U|$.

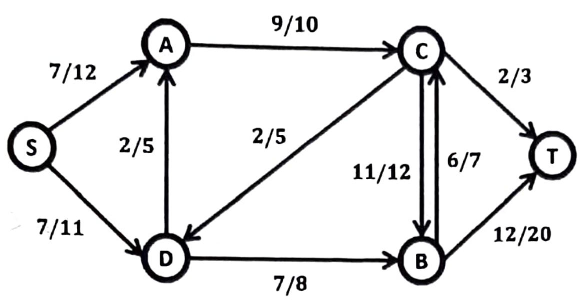
(b) (2%). Consider the following graph. Identify a maximum-size matching and a minimum-size vertex cover for it.



22. Consider the residual flow network with a source vertex S and a sink vertex T given in the following figure. Apply the Ford-Fulkerson algorithm to it.

(a) (2%). Write down the augmenting paths used by the Ford-Fulkerson algorithm.

(b) (2%). Identify a minimum S-T cut for the network.



23. Let $a_1, a_2, \ldots, a_n$ be a sequence of integers and assume that $a_0 = a_{n+1} = -\infty$. For any i with $1 \le i \le n$, we define a_i to be a peak, if $a_i \ge \max(a_{i-1}, a_{i+1})$.

(a) (2%). Identify all peaks in the given sequence $6, 7, 4, 3, 2, 1, 4, 5$.

(b) (2%). Write down a procedure that finds a peak for $a_1, \ldots, a_n$ in $O(\log n)$ time.

(c) (2%). Briefly justify your answer in (b).

24. (2%). Let $G = (V, E)$ be a complete graph with edge length function w. Prove or disprove the statement: The cost of any minimum spanning tree (MST) for (G, w) is always no more than the cost of any TSP tour for (G, w).

25. Consider the Knapsack problem with n items and a knapsack size W.
Let $(a_1, b_1), (a_2, b_2), \ldots, (a_n, b_n)$ be the size and profit of the items in order, i.e., a_i is the size of the i^{th} item and b_i is its profit.

(a) (2%). Consider the following simple greedy algorithm.

- $\mathcal{C} \longleftarrow \emptyset$, $\mathcal{S} \longleftarrow \{1, 2, \ldots, n\}$.
- Repeat until $\mathcal{S} = \emptyset$, do
  - Find $i \in \mathcal{S}$ with the maximum profit-cost ratio. That is,

$$\frac{b_i}{a_i} = \max_{j \in \mathcal{S}} \frac{b_j}{a_j}.$$

  - Add i to $\mathcal{C}$ if $a_i + \text{size}(\mathcal{C}) \le W$.
  - Remove i from $\mathcal{S}$.

【不可使用計算機】＊作答前請先核對試題、答案卷(試卷)與准考證之所組別與考科是否相符！！

Does this algorithm correctly compute an optimal solution for the Knapsack problem? If so, briefly justify its correctness. If not, provide a counter-example with the smallest number of items.

(b) (2%). For any i and w with $0 \le i \le n$ and $w \ge 0$, define $A(i, w)$ to be the maximum profit we can obtain, if we are using only the first i items and the knapsack size is w. Consider the following recurrence formula for $A(i, w)$.

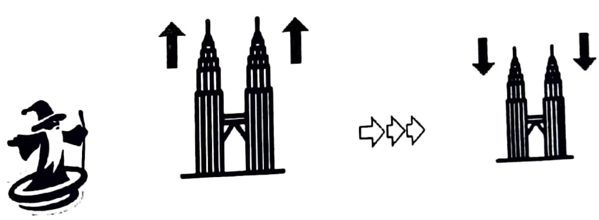
$$A(i,w) = \begin{cases} -\infty, & \text{if } w < 0, \\ 0, & \text{if } w \ge 0,\ i = 0, \\ \max(\ A(i-1, w),\ A(i-1, w-a_i) + b_i\ ), & \text{if } w \ge 0,\ i > 0. \end{cases}$$

Suppose that $W = 5$ and we have 3 items with size and profit $(4, 20)$, $(2, 8)$, $(3, 14)$. Use the recurrence to calculate the value $A(i, w)$ for all $0 \le i \le 3$ and $0 \le w \le 5$ in a table, as shown below. What is the optimal profit?

	0	1	2	3	4	5
$i = 0$						
$i = 1$						
$i = 2$						
$i = 3$						

(column label: w)

26. Consider the following optimization problem. Alex is a wizard and he wants to cast spells to grow the height of a tower. Alex has n types of spells, denoted by (a_i, b_i, k_i) for $1 \le i \le n$, where $a_i, b_i \ge 1$.



If Alex casts the i^{th}-type of spell in the beginning of a round, the height of the tower grows by a_i immediately. Then, when it comes to the end of each round for the next k_i rounds, including the round Alex casts the spell, the height of the tower shrinks by b_i.

The effect of the spells can stack. So, at the end of each round, the height the tower shrinks is equal to the sum of b_i for all spell i that is still in effect.

Alex can cast at most one spell in each round, and each spell can be cast at most once. Alex can choose any subset of spells he wants and cast them in any order he likes. The tower starts at height zero. The problem is to compute the *maximum record height* of the tower that Alex can make.

(a) (2%) Suppose that $n = 4$ and the spells are $(5, 3, 2)$, $(10, 9, 2)$, $(20, 33, 1)$, and $(30, 115, 1)$. What is the maximum record height of the tower Alex can make?

(b) (3%) Define the following notions. For any $1 \le i, j \le n$,

- Let $A(i, j)$ denote the maximum record height of the tower, if (i) only the first i spells are considered and (ii) j spells are in effect when the value $A(i, j)$ is attained.
- Let $B(i, j)$ denote the maximum record height of the tower, if (i) only the first i spells are considered, (ii) j spells, including the i^{th} spell, are in effect when the value $B(i, j)$ is attained, and (iii) the i^{th} spell is cast at the first round.

$A(i, j)$ and $B(i, j)$ are defined to be $-\infty$ if it is not possible. Based on the optimal substructure, write down the recurrence formula for $A(i, j)$ and $B(i, j)$, as is done in Question 25-(b). If needed, you may state explicitly and assume that (a_i, b_i, k_i) are sorted in a particular order.