系所組別： 電機工程學系丁組、電腦與通信工程研究所甲組

考試科目： 資料結構

考試日期：0220，節次：2

※ 考生請注意：本試題 □可 ☑不可 使用計算機　　請勿在本試題紙上作答，否則不予計分

※ 本試題共計 4 頁。作答時可不必抄題，但請必須將各題之題號標註清楚。

※ Note that throughout the paper, the height of a binary tree with only one node is defined as 1 while an empty binary tree has height of zero.

1. 是非題 (16 分；每題作答正確則得 2 分，作答錯誤則得 0 分並倒扣 1 分，未作答得 0 分)
For each question from (1-1) to (1-8), mark T (or ○) if the statement given is true, mark F (or ×) otherwise. Note that you will earn 2 points for each correct answer and lose 1 point for each incorrect answer. You earn no point if you do not give any answer.

(1-1) If a problem can be solved by a recursive algorithm, then the problem can be solved by an iterative algorithm with the same time complexity.

(1-2) For sequential search, searching a data item in a data list sorted by the key values shows better worst case performance than searching in an unsorted data list.

(1-3) The ADT (Abstract Data Type) of sparse matrix can be used for the ADT of general matrix.

(1-4) Hashing can be applied to developing a sorting algorithm of time complexity no worse than $O(n\log_2 n)$ where n is the number of data items.

(1-5) The relation among all the members of a family can be represented by a directed graph.

(1-6) Heap sort algorithm belongs to the category of selection sorting.

(1-7) Given a graph G , if there is a vertex v of G such that DFS(v) spanning tree is also the BFS(v) spanning tree of G, then G is also a tree structure.

(1-8) When creating a B-tree from a list of data items, the data items must be already sorted by the key values.

2. 選擇題 (16 分，每小題 2 分)
For each question from (2-1) to (2-8), please give all appropriate items as the answer (2 points each). Note that there might be two or more appropriate items and you must give all such items to get the points. In case you believe that no appropriate item is listed, then give F as your answer.

(2-1) Listed below are statements about *array* in computer programming and data structures:
① *Array* can be used to implement the data type of character strings of variable length.
② *Array* can be used to implement the data type of complex numbers in mathematics.
③ The components of an *array* do not need to be of the same data type and the same size.
④ *Array* can be used to implement linked list and *array* can be implemented by linked list.
⑤ The number of dimensions of an *array* is only limited by available memory storage.
Based on the above, which is/are true?
(A) ①④ are correct.　(B) ②④ are correct and ③ are incorrect.
(C) ②⑤ are incorrect.　(D) ①⑤ are correct and ④ are incorrect.
(E) ① is correct and ③ is incorrect.

(2-2) Which of the following statements is/are true?
(A) Priority queue is a data structure with FIFO property.
(B) Given a binary tree T of n_0 leaves, the threaded binary tree of T has $(2 \cdot n_0)$ threads.
(C) Given an AVL tree T of n nodes, the time complexity of doing LL and RR rotations on T are the same as that of doing LR and RL rotations on T.
(D) $= x + y\ z$ is the prefix representation of the infix expression $x = y + z$.
(E) None of the above is true.

系所組別： 電機工程學系丁組、電腦與通信工程研究所甲組

考試科目： 資料結構　　考試日期：0220，節次：2

※ 考生請注意：本試題 □可 ☑不可 使用計算機

2. 選擇題 (16 分，每小題 2 分) [續]

(2-3) In sorting n ($n > 2$) data items using a comparison-based sorting algorithm, if k denotes the number of key comparisons and m denotes the number of data movements incurred from maintaining the order of the data items, then which of the following is/are correct?

(A) $(n \cdot \log_2 n) \le m \le n(n-1)/2$　(B) $(n \cdot \log_2 n) \le k \le n(n-1)/2$　(C) $0 \le m \le k \le n^2$

(D) If $m = 0$, then $k = (n-1)$　(E) If $k = n(n-1)/2$, then $m = (k-1)$

(2-4) Given a triangular matrix T of dimension $n \times n$ and an array U of m components. U is used to store the non-zero components of T. If $T[0][1]$ is stored into $U[0]$ and $T[1][4]$ is stored into $U[n+1]$, then which of the following is/are true?

(A) T is a lower triangular matrix and $T[i][j] \neq 0$, $0 \le i < j < n$.

(B) $T[k][k] = 0$, $0 \le k < n$.

(C) $T[4][9]$ is stored into U $[4n-5]$.

(D) The value stored into U $[3n+7]$ is the value of $T[3][6]$.

(E) The non-zero components of T are stored into array U by the column-major order.

(2-5) Which of the following information can determine the structure of a binary tree T ?

(A) T is a complete binary tree and the output of doing in-order traversal on T.

(B) Consider T as a graph and its adjacency matrix.

(C) The output of doing level-order traversal and post-order traversal on T.

(D) The output of doing pre-order traversal on T and the number of degree-1 nodes of T.

(E) T has no degree-1 node and the output of doing pre-order traversal on T.

(2-6) Regarding to Dijkstra's algorithm in solving the shortest path problem, which of following statements are true?

(A) The Dijkstra's algorithm solves the shortest path problem on edge-weighted multi-graph in which all the weights are non-negative.

(B) The Dijkstra's algorithm belongs to the category of greedy algorithms.

(C) When implementing the Dijkstra's algorithm, the data structure of priority queue is used.

(D) For the Dijkstra's algorithm, using adjacency list and adjacency matrix show the same worst-case time complexity.

(E) The Dijkstra's algorithm can be used to find the minimum cost spanning tree.

(2-7) Which of the following statements on sorting is/are not true?

(A) The data items to be sorted must be of the same length.

(B) Comparison-based sorting algorithms always show time complexity better than those of non-comparison-based sorting algorithms.

(C) Whether the sorting algorithm is stable or not may depend on the implementation.

(D) If the data items to be sorted are already sorted at the beginning of sorting process, then this sorting case will produce the best case performance of the sorting algorithm.

(E) In binary tree sort, the output is produced from the pre-order traversal of a binary tree with the height of $O(n)$.

(2-8) In searching a data item with key value k in a set of n data items, which is/are true?

(A) Hashing is most efficient approach if the hash table has n buckets.

(B) Binary search tree is better than binary search if the data items are not sorted.

(C) Binary search is better than sequential search if the data items are not sorted.

(D) The time complexity of a successful search is the same as that of a failed search.

(E) The actual searching time depends only on k and n.

※ 考生請注意：本試題 □可 ☑不可 使用計算機

3. On merge sort, please do the following:
   (3-1) Derive its time complexity using recurrence relation. (6 points)
   (3-2) Is merge sort a stable sorting algorithm? Prove your answer. (6 points)

4. Given the complexity of F_1 is $O(n^2)$ and that of F_2 and F_3 are respectively $O(n \cdot \log_2 n)$ and $O(\log_2 n)$. If G is a compound function of F_1, F_2 and F_3 and $G = F_1 + F_2 \cdot F_3$, then what is the complexity of G ? (6 points)

5. The C function popMinHeap() given below is supposed to delete the smallest integer from a min heap which is stored in an integer array heap[]. After executing popMinHeap(), the heap remains as a min heap structure. Variable n indicates the number of keys contained in array heap[]. Unfortunately, popMinHeap() has some bugs. Please correct all the bugs in the program. DO NOT rewrite the entire program, just show your modification with the line number. (10 points)

```
1   int popMinHeap(int *n, int* heap)
2   {   int parent, child, min, tmp;
3       min = heap[1];      tmp = heap[n--];
4       parent = 1;         child = 2;
5       while (child < *n) {
6           if ((child <= *n) && (heap[child] > heap[child + 1]))
7               child++;
8           if (tmp > heap[child]) break;
9           heap[parent] = heap [child];    parent = child;
10          child = parent*2; }
11      heap[parent] = tmp;
12      return min;
13  }
```

6. For each of the terms below, please give its definition or explanation. (9 points, 3 points each)
   (1) Eulerian walk (2) Strongly connected graph (3) Balance factor of AVL tree

7. Suppose you are given a set of alphabets and their frequencies below. (5 points)

Alphabet	a	e	h	l	p
Frequency	15K	10K	2K	4K	7K

   (1) Construct the Huffman encoding tree step-by-step based on the table. (3 points)
   (2) Encode the message "alpha" using the Huffman encoding tree. (2 points)

(背面仍有題目,請繼續作答)

※ 考生請注意：本試題 □可 ☑不可 使用計算機

8. 複選題：(15 points)

(8-1) Given a preorder traversal sequence of a binary search tree T as

41, 20, 8, 32, 25, 63, 55, 49, 44, 68, 65.

Which of the following statement(s) is (are) true? (5 points)

(A) The number of leaf nodes in the tree T is 4.
(B) If the tree T is traversed by postorder traversal, the first node visited is 8.
(C) The height of T is 4.
(D) The tree T is not an AVL tree.
(E) The tree T cannot be uniquely determined.

(8-2) Given an undirected graph $G(V, E)$ where the number of vertexes is n and the number of edges is m. Which of the following statement(s) is (are) true about the graph G? (5 points)

(A) If G is represented by adjacency matrix, the space complexity is $O(n^2)$.
(B) If G is represented by adjacency list, the space complexity is $O(m)$.
(C) If G is represented by adjacency list, the time complexity to determine whether G is connected is $O(m)$.
(D) If G is connected, then $n \geq m + 1$.
(E) If G is represented by adjacency list, the time complexity for determining whether every vertex in G has at least two paths to any other vertex in G is $O(n^3)$

(8-3) Arrange the following functions by growth rate in increasing order.

n, $\log(n!)$, n^2, $n \log(\log n)$, 2^{1000}, 2^n, $n^2 \log n$, n^3, $n!$, $(\log n)^{\log n}$

Based on the increasing order arrangement, which of the following statement(s) is (are) true? (5 points)

(A) 2^n is located at the last position.
(B) n^2 log n is located at the fifth position.
(C) $(\log n)$log n is located before n^3.
(D) $\log(n!)$ is located before n^2.
(E) n^3 is located at the seventh position.

9. Given a hash table of 11 buckets, labeled from 0 to 10. Each bucket can hold only one key. The hash function h(key) = key % 11. Chaining is used in the hash scheme. Now, suppose the hash table is initially empty and then the numbers "3, 41, 15, 36, 74, 58, 91, 45, 48, 64" are hashed sequentially into the hash table. (11 points)

(1) Show the hash table after inserting the sequence of numbers. (3 points)

(2) What are the average comparisons if each number of the sequence is looked up once? (3 points)

(3) In general, what is the worse-case time complexity for hashing in such a hash scheme? (2 points)

(4) If the worse-case time complexity of hashing need to be improved to $O(\log n)$, how ? (3 points)