國立中央大學98學年度碩士班考試入學試題卷

所別：資訊工程學系碩士班　科目：資料結構與演算法　共 2 頁　第 1 頁
*請在試卷答案卷（卡）內作答
所別：軟體工程研究所碩士班

1. [Hashing]
Assume that the data and key are both integers. The hash table size (HSize) is 10. The index runs from 0 to 9. Please fill in the content of the hash table if numbers 6, 3, 35, 15, 27, 95, and 23 are sequentially inserted into the hash table using the following scheme. If a number can not be inserted to the hash table, please indicate it.

(a) HF: key mod HSize, open addressing with linear probing. (4%)

0	1	2	3	4	5	6	7	8	9

(b) HF: key mod HSize, open addressing with quadratic probing. (4%)

0	1	2	3	4	5	6	7	8	9

(c) HF: key mod HSize, open addressing with double hashing, where HF2 = 7- (key mod 7). (4%)

0	1	2	3	4	5	6	7	8	9

(d) Some numbers may not be able to insert to the table even though there are still spaces in the hash table for quadratic probing or double hashing. What can we do to reduce the chance of this mishap to happen when we design the hash table? (1%)

2. [Sorting]
(a) Given the sequence of numbers: 8, 1, 7, 2, 6, 3, 5, 4 we intend to sort them into ascending order. How many swaps does it take to sort this sequence via Insertion sort? (3%)
(b) Given the sequence of numbers: 8, 1, 7, 2, 6, 3, 5, 4 we intend to sort them into ascending order. How many recursive calls does it take by the Merge sort to get the job done? (3%)
(c) Given the sequence of numbers: 15, 14, 13, 12, 11, 9, 3, 1, we intend to sort them into ascending order. What is the required storage space for the counting sort? Assume the range of input to be 1…15. Count each cell of each array counting sort uses, including the input array. (3%)
(d) Given the sequence of numbers: 15, 14, 13, 12, 11, 9, 3, 1, we intend to sort them into ascending order. What is the result of the first pass if the radix sort is used on this input? (3%)

3. [Huffman code]
Given the frequency of characters as follows, use Huffman coding to solve the following problem:

Character	A	B	C	D	E
Frequency	22	3	10	5	60

(a) Suppose D is known to be encoded as "0001" Show how A, B, C, D, and E are encoded. (5%)

(b) Give two 6-bit sequences of 0's and 1's such that none of them is a valid encoded message. (5%)

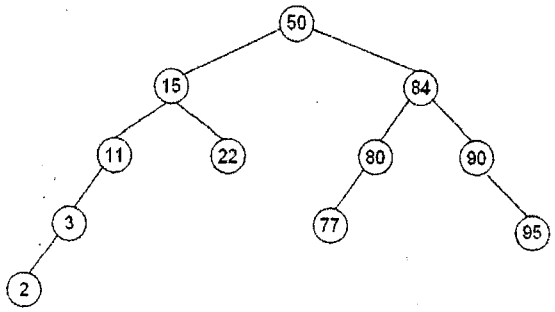
參考用

注意：背面有試題

國立中央大學98學年度碩士班考試入學試題卷

所別：資訊工程學系碩士班　科目：資料結構與演算法　共 2 頁　第 2 頁
*請在試卷答案卷（卡）內作答
所別：軟體工程研究所碩士班

4. [AVL Tree]
Given the following figure in which the key value of the root is 50, answer the following questions:



(a) Use the post-order tree walk algorithm to print the key of each node. (3%)
(b) Is the tree a binary search tree? Why? (3%)
(c) We can remove a node to make the tree in the AVL form. What node is it? And why is it? (3%)
(d) Suppose we don't want to delete any node. Please show how you can use tree rotation operations to balance the tree. (3%)
(e) Following the above question, the tree should be balanced now. Assume a new node 13 is added to the tree right after the rotation takes place. Show how you can keep the tree in the AVL form. (3%)

5. Given n numbers $x_1, x_2, \ldots, x_n$, consider the problem of computing $d[i, j] = x_i + x_{i+1} + \ldots + x_j$, for all $i \leq j$. A naïve algorithm by computing each $d[i, j]$ independently will take $\Theta(n^3)$ time. Derive an efficient way to solve this problem in $O(n^2)$ time. (10%)

6. The input is a sequence of n integers with many duplications, such that the number of distinct integers in the sequence is $O(\log n)$.
(a) Design a sorting algorithm to sort such sequences using at most $O(n \log\log n)$ comparisons in the worst case. (10%)
(b) Why is the lower bound of sorting $\Omega(n \log n)$ not satisfied in this case? (5%)

7. Let $G=(V, E)$ be a weighted connected undirected graph where V represents the set of vertices and E represents the set of edges. A spanning tree of G is an undirected tree $S = (V, T)$ where T is a subset of E. The total weight of a spanning tree is the sum of all weights of T. A minimum spanning tree of G is a spanning tree of G with the smallest total weight. Write a greedy algorithm to find the minimum spanning tree of a given weighted connected undirected graph G. (10%)

8. In a 2D plane, we say that a point (x_1, y_1) dominates (x_2, y_2) if $x_1 > x_2$ and $y_1 > y_2$. A point is called a maximal point if no other point dominates it. Given a set of n points, the maxima finding problem is to find all of the maximal points.
(a) Write a divide and conquer algorithm to solve the maxima finding problem with the time complexity of $O(n \log n)$. (10%)
(b) Show that your algorithm is indeed of the time complexity $O(n \log n)$. (5%)

注意：背面有試題