

诚信应考,考试作弊将带来严重后果!

# 华南理工大学期末考试

## 《 Data Structure 》A 试卷

- 注意事项: 1. 考前请将密封线内填写清楚;  
2. 所有答案请直接答在试卷上;  
3. 考试形式: 闭卷;  
4. 本试卷共十大题, 满分 100 分, 考试时间 120 分钟。

题号	一	二	三	四	五	六	七	八	九	十	十一	总分
得分												
评卷人												

1. Select the correct choice. (20 scores, each 2 scores)
- (1) An algorithm must be or do all of the following EXCEPT: ( B )  
(A) Partially correct (B) Ambiguous (C) terminate (D) Concrete steps
- (2) Pick the growth rate that corresponds to the most efficient growing algorithm as n gets larger: ( B )  
(A)  $4n^3$  (B)  $5n^2 \log n$  (C)  $3n!$  (D)  $2^n$
- (3) If a data element requires 12 bytes and a pointer requires 4 bytes, then a linked list representation will be more space efficient than a standard array representation when the fraction of non-null elements is less than about: ( B )  
(A) 1/3 (B) 3/4 (C) 3/5 (D) 2/3
- (4) Which is the realization of a data type as a software component: ( A )  
(A) An abstract data type (B) A real data type  
(C) A type (D) A data structure
- (5) The most effective way to reduce the time required by a disk-based program is to: ( B )  
(A) Improve the basic I/O operations. (B) Minimize the number of disk accesses.  
(C) Eliminate the recursive calls. (D) Reduce main memory use.
- (6) In the hash function, collision refers to ( B ).  
(A) Two elements have the same key value.  
(B) Different keys are mapped to the same position of hash table.  
(C) Two records have the same requiring number.  
(D) Data elements are too much.
- (7) Given an array as  $A[m][n]$ . Supposed that  $A[0][0]$  is located at  $644_{(10)}$  and  $A[4][4]$  is stored at  $676_{(10)}$ . “ $_{(10)}$ ” means that the number is presented in decimals. Then the element  $A[2][2]_{(10)}$  is at position:  
( B )  
(A) 692 (B) 660 (C) 650 (D) 708

- (8) Which statement is not correct among the following four: ( **A** )
- (A) The number of empty sub-trees in a non-empty binary tree is one less than the number of nodes in the tree.
  - (B) The Mergesort is a stable sorting algorithm.
  - (C) The root of a binary tree transferred from a general tree has only left child.
  - (D) A sector is the smallest unit of allocation for a record, so all records occupy a multiple of the sector size.
- (9) Tree indexing methods are meant to overcome what deficiency in hashing?  
( **D** )
- (A) Inability to handle range queries.
  - (B) Inability to maximum queries
  - (C) Inability to handle queries in key order
  - (D) All of above.
- (10) Assume that we have eight records, with key values A to H, and that they are initially placed in alphabetical order. Now, consider the result of applying the following access pattern: F D F G E G F A D F G E if the list is organized by the transpose heuristic, then the final list will be ( **B** ).
- (A) A F C D H G E B
  - (B) A B F D G E C H
  - (C) A B F G D C E H
  - (D) A H F D G E C B

2. Fill the blank with correct C++ codes: (16 scores)

- (1) Given an array storing integers ordered by value, modify the binary search routines to return the position of the first integer with the least value greater than K when K itself does not appear in the array. Return ERROR if the greatest value in the array is less than K: (10 scores)

```
// Return position of lest element >= K
int newbinary(int array[], int n, int K) {
    int l = -1;
    int r = n;           // l and r beyond array bounds
    while (l+1 != r) {    // Stop when l and r meet
        int i = (l+r) / 2;    // Check middle of remaining subarray
        if (K < array[i]) r=i;    // In left half
        if (K == array[i]) return i;    // Found it
        if (K > array[i]) l=i;    // In right half
    }
    // K is not in array or the greatest value is less than K
    if K < array[n-1] or r != n
        then return r;    // the first integer with the least value greater than K
                           // when K itself does not appear in the array
    else return ERROR;    // the greatest value in the array is less than K
}
```

(2)

The height of the shortest tree and the tallest tree with both  $n$  nodes is respectively 2 or  $n(n < 2)$  and  $n$ , suppose that the height of the one-node tree is 1. A 3-ary full tree with  $n$  internal nodes has  $3n+1$  nodes. (6 scores)

3. Please calculate the number of binary trees in different shape with 6 nodes in total, and 6 nodes? (4 scores)

2 nodes: 2 shapes

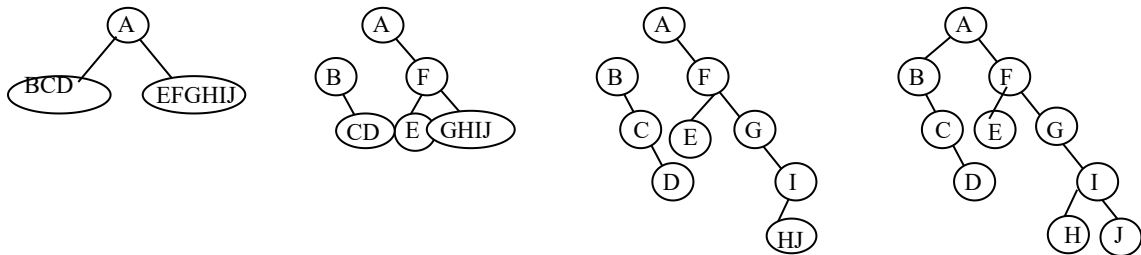
3 nodes: root with 1 and 2 can be allocated to left and right so  $1+2+2=5$

4 nodes: 3 can be allocated as 0,3;1,2;2 so  $5+5+2+2=14$

5 nodes:  $14+14+5+5+4=42$

6 nodes:  $32+32+14+14+5*2*2=132$ ,  $C_{2n}^n/n+1$

4. A certain binary tree has the postorder enumeration as DCBEHJIGFA and the inorder enumeration as BCDAEFGHIJ. Try to draw the binary tree and give the postorder enumeration. (The process of your solution is required!!!) (6 scores)



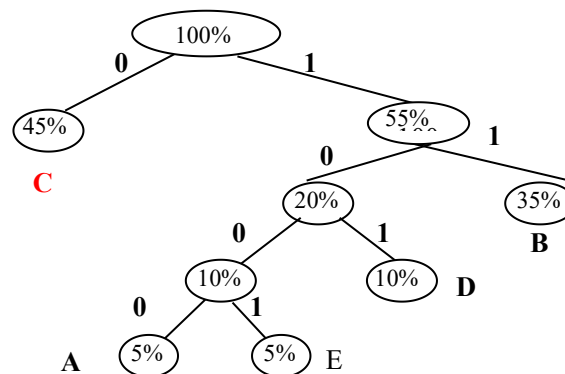
preorder enumeration: ABCDFEGIHJ

5. Design an algorithm to transfer the score report from 100-point to 5-point, the level E corresponding score  $<60$ ,  $60 \sim 69$  being D,  $70 \sim 79$  being C,  $80 \sim 89$  as B, score  $\geq 90$  as A. The distribution table is as following. Please describe your algorithm using a decision tree and give the total path length. (6 scores)

Score in 100-point	0-59	60-69	70-79	80-89	90-100
Distribution rate	5%	10%	45%	35%	5%

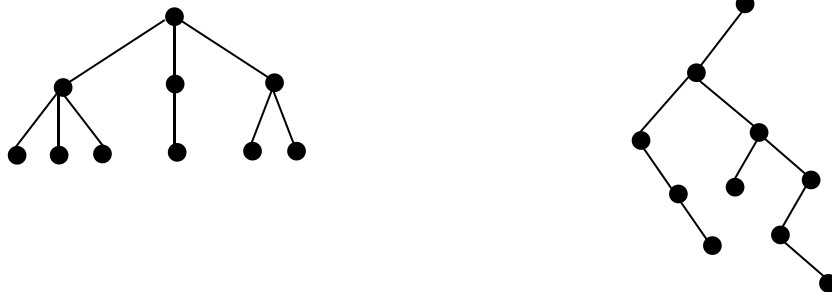
solution:

the design logic is to build a Huffman tree



Total length:  $4 * 10\% + 10\% * 3 + 35\% * 2 + 45\% = 1.85$ , the 0-false,1-true as the logic branches.

6. Recovery a general tree as following transferred from a binary tree. (4 scores)



7. Trace by hand the execution of Quicksort algorithm on the array:  $\text{int } a[] = \{44, 77, 55, 99, 66, 33, 22, 88, 79\}$ . The pivot is 66 in the first pass, the second is 55 and 88, the third is 33 and 77, the four is 22, and so on till the algorithm is finished. (6 scores)

initial: 44, 77, 55, 99, 66, 33, 22, 88, 79

pass 1: 44 22 55 33 66 99 77 88 79

pass 2: 44 22 33 55 66 79 77 88 99

pass 3: 33 22 44 55 66 77 79 88 99

pass 4: 22 33 44 55 66 77 79 88 99

final sorted array:

22 33 44 55 66 77 79 88 99

8. (a) Describe simply the main tasks of the two phases of external sorting. (4 scores) The task of first phase is to break the files into large initial runs by replacement selection; the second phase is to merge the runs together to form a single sorted run file. (b) Assume that working memory is 512KB broken into blocks of 4096 bytes (there is also additional space available for I/O buffers, program variables, etc.) What is the expected size for the largest file that can be merged using replacement selection followed by two passes of multi-way merge? Explain how you got your answer. (4 scores)

Since working memory is 512KB and the blocksize is 4KB, the working memory holds 128 blocks. The expected runlength is 1024KB, so a single pass of multiway merge forms runs of length  $1024\text{KB} * 128 = 128\text{MB}$ . The second pass then forms a run as large as  $128\text{MB} * 128 = 16\text{GB}$ .

9. Assume a disk drive is configured as follows. The total storage is approximately 675M divided among 15 surfaces. Each surface has 612 tracks; there are 144 sectors/track, 512 byte/sector, and 16 sectors/cluster. The disk turns at 7200rpm (8.33ms/r). The track-to-track seek time is 20 ms, and the average seek time is 80 ms. Now how long does it take to read all of the data in a 380 KB file on the disk? Assume that the file's clusters are spread randomly across the disk. A seek must be performed each time the I/O reader moves to a new track. Show your calculations. (The process of your solution is required!!!) (6 cores)

A cluster holds  $16 \times 0.5K = 8K$ . Thus, the file requires  $380/8=47.5$  clusters.

The time to read a cluster is seek time to the cluster+ latency time + (interleaf factor  $\times$  rotation time). Average seek time is defined to be 80 ms. Latency time is  $0.5 \times 8.33$ , and cluster rotation time is  $47.5 \times (16/144) \times 8.33$ . Seek time for the total file read time is

$$47 \times (80 + 0.5 \times 600/72 + (16/144) \times 600/72) + (80 + 0.5 \times 600/72 + (8/144 \times 600/72)) = 4083.98 \text{ms}$$

10. Using closed hashing, with double hashing to resolve collisions, insert the following keys into a hash table of eleven slots (the slots are numbered 0 through 10). The hash functions to be used are H1 and H2, defined below. You should show the hash table after all eight keys have been inserted. Be sure to indicate how you are using H1 and H2 to do the hashing. ( The process of your solution is required!!!)

$$H_1(k) = 3k \bmod 11$$

$$H_2(k) = 7k \bmod 10 + 1$$

Keys: 22, 31, 18, 35, 44, 13, 1, 67.

(8 scores)

Answer:

$H_1(22)=0$ ,  $H_1(31)=5$ ,  $H_1(18)=10$ ,  $H_1(35)=6$ , no conflict

When  $H_1(44)=0$ ,  $H_2(44)=9$   $(0+9 \times 1) \% 11=9$ , so 44 enters the 9<sup>th</sup> slot;

$H_1(13)=6$ ,  $H_2(13)=2$   $(6+1 \times 2) \% 11=8$ , so 13 enters the 8<sup>th</sup> slot;

$H_1(1)=3$ , so 1 enters 3 ;

$H_1(67)=3$ ,  $H_2(67)=10$   $(3+2 \times 10) \% 11=1$  so 67 enters 1 (pass by 2)

22		67	1		31	35		13	44	18
0	1	2	3	4	5	6	7	8	9	10

11.

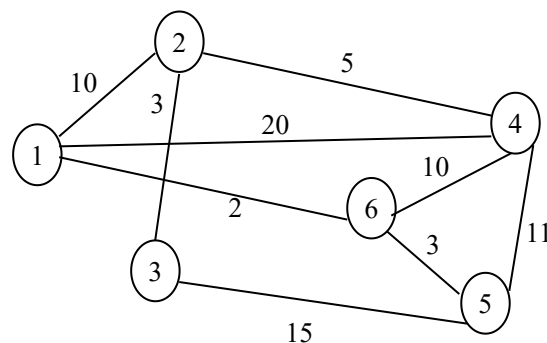


Figure 1 Example graph

(a) find the shortest paths from Vertex1 to all the other vertices. (3)

(b) Use Kruskal's algorithm to find the minimum-cost spanning tree. (3)

- (a) 1 to 2: 10 (1,2);  
 1 to 3: 13 (1,2,3);  
 1 to 4: 12 (1,6,4);  
 1 to 5: 5 (1,6,5);  
 1 to 6: 2 (1,6);

(b)

