座位号 专业 学院 学号 姓名

密 封 线 内 不 答 题

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 》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	总分
得 分											
评卷人											

1. Select the correct choice. (20 scores, each 2 scores)

(1) Pick the growth rate that corresponds to the most inefficient algorithm as n gets large: (C)
(A) $2n^3$ (B) 2^n (C) n! (D) $10n^2\log n$

(2) An algorithm must be or do all of the following EXCEPT: (B)
(A) Partially correct (B) Ambiguous (C) Terminate (D) Concrete steps

(3) If a data element requires 8 bytes and a pointer requires 2 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) 1/4 (B) 4/5 (C) 4/7 (D) 3/4

(4) Which statement is not correct among the following four: (C)
(A) The Quick-sort is an unstable sorting algorithm.
(B) The number of empty sub-trees in a non-empty full binary tree is one more than the number of nodes in the tree.
(C) The worst case for my algorithm is n becoming larger and larger because that is the slowest.
(D) A cluster is the smallest unit of allocation for a file, so all files occupy a multiple of the cluster size.

(5) Which of the following is a true statement: (C)
(A) A general tree can be transferred to a binary tree with the root having both left child and right child.
(B) In a BST, the node can be enumerated sorted by a preorder traversal to the BST.
(C) In a BST, the left child of any node is less than the right child, but in a heap, the left child of any node could be less than or greater than the right child.
(D) A heap must be full binary tree.

(6) The golden rule of a disk-based program design is to: (B)
(A) Improve the basic operations. (B) Minimize the number of disk accesses.

(C) Eliminate the recursive calls. (D) Reduce main memory use.

(7) Given an array as A[m] [n]. Supposed that *A* [0] [0] is located at $644_{(10)}$ and *A* [2] [2] is stored at $676_{(10)}$, and every element occupies one space. "$_{(10)}$" means that the number is presented in decimals. Then the element *A* [5] [5]$_{(10)}$ is at position:

(D)

(A) 692 (B) 695 (C) 650 (D) 724

(8) If there is 0.5MB working memory, 4KB blocks, and yield 128 blocks for working memory. By the multi-way merge in external sorting, the average run size and the sorted size in one pass of multi-way merge on average are separately (C)?

(A) 0.5MB, 128 MB (B) 2MB, 512MB

(C) 1MB, 128MB (D) 1MB, 256MB

(9) Which algorithm is used to generate runs in classic external sorting? (D)

(A) Quick-sort (B) Bubble sort

(C) Insertion sort (D) Replacement selection

(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 G F A D F G, if the list is organized by the Transpose heuristic, then the final list will be (D).

(A) G F D A C H B E (B) A G F D B C E H

(C) F D G A E C B H (D) A B F G D C E H

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

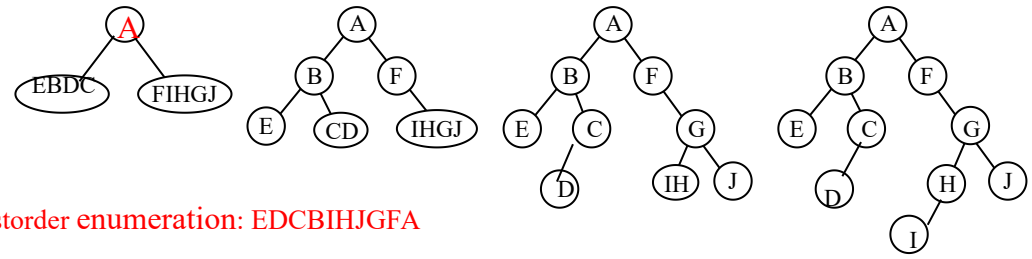
(1) Given an array storing integers ordered by distinct value without duplicate, modify the binary search routines to return the position of the integer with the smallest value greater than K when K itself does not appear in the array. Return ERROR if the greatest value in the array is less than K: (10 scores)

```
// Return position of smallest 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_____;    // Look at middle of 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）_// the greatest value in the array is not less than K with r updated
       return   r   ;   // when K itself does not appear in the array
    else     return  ERROR;   // the integer with the greatest value less than K
}
```

(2) The height of a complete binary tree with k nodes is _⌈log2(k+1)⌉(1 node tree has hight 1)_ (2 scores)

(3) The number of different shapes of binary trees with 5nodes is _42_.(2 scores)

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



Postorder enumeration: EDCBIHJGFA

4. Determine Θ for the following code fragments in the average case. Assume that all variables are of type int. (6 scores)

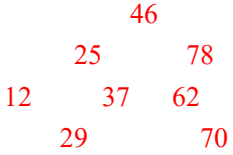
(1) sum=0;
```
for (i=0; i<5; i++)
  for (j=0; j<n; j++)
    sum++;
```
solution : Θ___(n)_____

(2) sum = 0;
```
for(i=1;i<=n;i++)
  for(j=n;j>=i;j--)
    sum++;
```
solution : Θ__(n²)_____

(3) sum=0;
```
if (EVEN(n))
 for (i=0; i<n; i++)
   sum++;
else
   sum=sum+n;
```
solution : Θ___(n)_____

5. Trace by hand the execution of creation a binary search tree with the input sequence as : {46，25，78，62，12，37，70，29} which is empty tree initially. (6 scores)

Solution: BST obtained with data inserted one by one

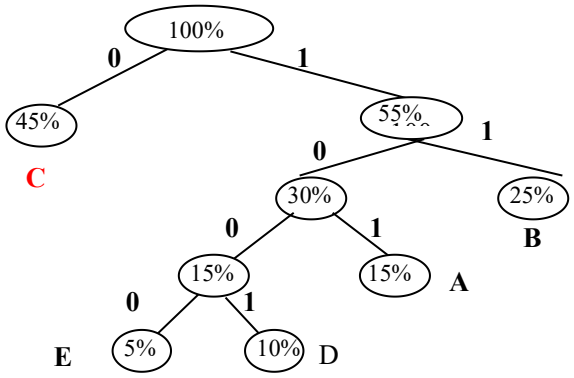


6. Design an algorithm to transfer the score report from 100-point to 5-point, the level E corresponding score<60, 60～69 being D, 70～79 being C, 80～89 as B， score>=90 as A. The distribution table is as following. Please describe your algorithm using a decision tree and give the total path length. (9 scores)

Score in 100-point	0-59	60-69	70-79	80-89	90-100
Distribution rate	5%	10%	45%	25%	15%

solution:

the design logic is to build a Huffman tree



Total length: 4 * 5% +10% * 4 + 15 %* 3 + 25% * 2 + 45%= 2.00, the 0-false,1-true as the logic branches.

7. Assume a disk drive is configured as follows. The total storage is approximately 1.35G divided among 15 surfaces. Each surface has 612 tracks; there are 144 sectors/track, 1024 byte/sector, and 16 sectors/cluster. The interleaving factor is four. The disk turns at 7200rmp (8.33 ms/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 360 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!!!) (9 scores)

Solution：

The first question is how many clusters the file requires?

A cluster holds 16*1K = 16K. Thus, the file requires 360/16=22.5 clusters=22complete cluster and 8k（8 sectors）

The time to read a cluster is seek time to the cluster+ latency time + (interleaf factor × rotation time).

Average seek time is defined to be 80 ms. Latency time is 0.5 * 8.33 ms (60/7200≈ 8.33ms), and cluster rotation time is 4* (16/144)*8.33.

Seek time for the total file read time is

22* (80 + 0.5 * 8.33+ 4 * (16/144)*8.33) +(80 + 0.5 * 8.33+ 4 * (8 /144)*8.33)≈ 2019.095ms

8. 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!!!)

H1(k) = 3k mod 11 H2(k) = 7k mod 10+1

Keys: 22, 41, 53, 46, 30, 13, 1, 67. (9 scores)

Solution：

H1(22)=0, H1(41)=2, H1(53)=5, H1(46)=6, no conflict

When H1(30)=2, H2(30)=1 （2+1*1）%11=3，so 30 enters the 3rd slot;

H1(13)=6, H2(13)=2 (6+1*2)%11=8, so 13 enters the 8th slot;

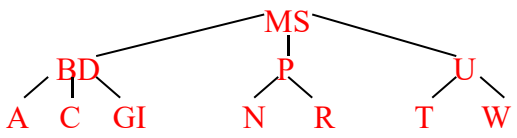
H1(1)=3, H2(1)=8 (3+5*8)%11= 10 so 1 enters 10 (pass by 0, 8, 5, 2);

H1(67)=3, H2(67)=10 (3+2*10)%11= 1 so 67 enters 1(pass by 2)

22	67	41	30		53	46		13		1
0	1	2	3	4	5	6	7	8	9	10

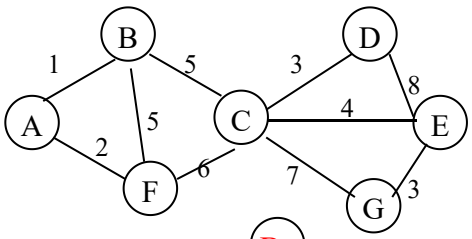
9. You are given a series of records whose keys are chars. The records arrive in the following order: C, S, D, T, A, M, P, I, B, W, N, G, U, R. Show the 2-3 tree that results from inserting these records.（the process of your solution is required!!!）(9 scores)

Solution：



10.

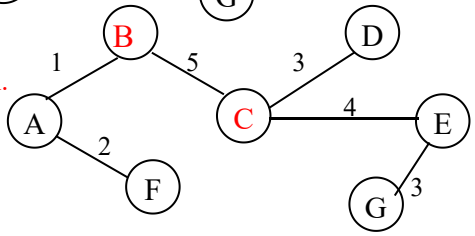
The following graph is a communication network in some area, whose edge presents the channel between two cities with the weight as the channel's cost. How to choose the cheapest path that can connect all cities? And how to get cheapest paths connecting each city-pair? You can draw all choices if there is more than one path. (10 scores)



Solution：

Draw the MST: It is not a Hamilton path.

Get the shortest paths between pairs.



	A	B	C	D	E	F	G
A		1	B: 6	BC:9	BC: 10;	2	BC:13
B	1		5	C: 8;	C:9;	A: 3;	C:12;
C	B: 6	5		3	4	6	7
D	BC:9	C: 8	3		C:7;	C:9	C:10;
E	BC: 10;	C:9;	4	C:7;		C:10	3
F	2	A: 3;	6	C:9	C:10		C:13;
G	BC:13	C:12;	7	C:10;	3	C:13;	