座位号 专业 学院 学号 姓名 密 封 线 （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B

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) An algorithm must be or do all of the following EXCEPT: (C)
(A) Correct (B) No ambiguous (C) General steps (D) terminate

(2) Pick the growth rate that corresponds to the most efficient algorithm as n gets large: (B)
(A) n^4 (B) $100n^3\log n$ (C) n! (D) 2^n

(3) If a data element requires 6 bytes and a pointer requires 3 bytes, then a standard array representation will be more space efficient than a linked list representation when the fraction of non-null elements is more than about: (D)
(A) 1/3 (B) 1/2 (C) 3/4 (D) 2/3

(4) 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) A general tree can be transferred to a binary tree with the root having 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.

(5) We use the parent pointer representation for general trees to solve (C) problem?
(A) Exact-match query (B) General tree traversal
(C) Determining if two nodes are in the same tree (D) Shortest paths

(6) The most effective way to reduce the time required by a disk-based program is to: (B)
(A) Improve the basic operations. (B) Minimize the number of disk accesses.
(C) Sorting the data of file. (D) Reduce main memory use.

(7) In the following sorting algorithms, which is the best one to find the first 10 biggest elements in the 1000 unsorted elements? (B)

(A) Insert sort. (B) Heap sort.
(C) Quicksort. (D) Shell sort.

(8) Given an array as A[m][n]. Supposed that $A[0][0]$is located at $544_{(10)}$ and $A[2][2]$is stored at $576_{(10)}$, and every element occupies one space. "$_{(10)}$" means that the number is presented in decimals. Then the element $A[3][3]_{(10)}$ is at position:

(A)

(A) 592 (B) 595 (C) 550 (D) 608

(9) Which queries supported by both of the hashing and tree indexing method? (D)

(A) Range queries. (B) Queries in key order
(C) Minimum or maximum queries (D). Exact-match queries

(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 Move-to-front heuristic, then the final list will be (B).

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

2. Fill the blank with correct C++ codes: (20 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 largest value less than K when K itself does not appear in the array. Return ERROR if the least value in the array is greater than K: (12 scores)

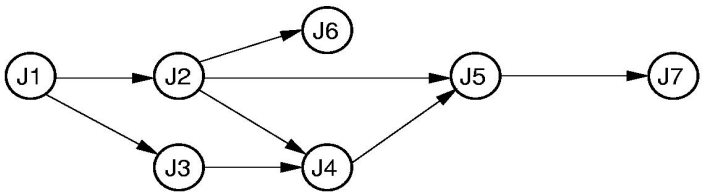
```
// Return position of lest element >= K
int newbinary(int array[], int n, int K) {
  int l = -1;
  int r = n;                    // l and r beyond array bounds
  while (l+1 != r) {            // Stop when l and r meet
    __int i=（l+r）/2_____;      // Check middle of remaining subarray
    if (K < 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 least value in the array is greater than K
  if  _K> array[0] (or l!=-1)
      then    return  _l;   //the smallest value in the array is not greater than K with l updated
      else        return  ERROR;    // the least value in the array is greater than K
                          }
```

(2) A full 8-ary tree with 100 internal nodes has _701__ leaves. (4 scores)

(3) The number of different shapes of binary trees with 6 nodes is _132_. (4 scores)

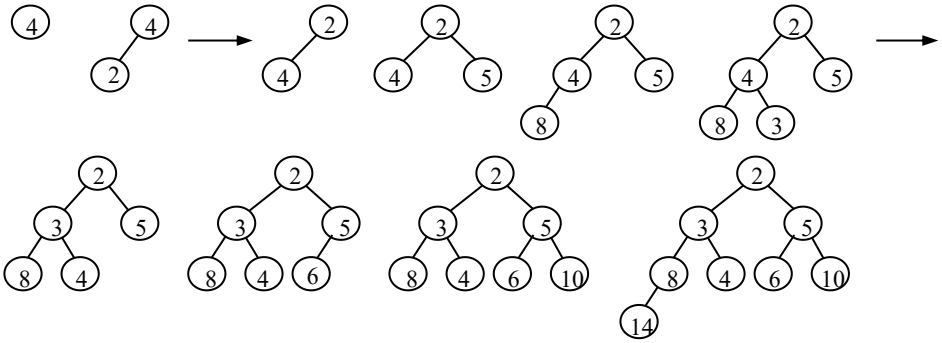
3. Show an acceptable topological sort to a series of jobs with some prerequisites

showed as following figure. (4 scores)



J1, J3, J2, J6, J4, J5, J7

4. Show the min-heap that results from running buildheap on the following values stored in an array: 4, 2, 5, 8, 3, 6, 10, 14. (6 scores)



5. Trace by hand the execution of Radix-sort algorithm on the array: int a[] = {265 301 751 129 937 863 742 694 076 438}. (6 scores)

initial: 265 301 751 129 937 863 742 694 076 438
pass 1: [] [301 751] [742] [863] [694] [265] [076] [937] [438] [129]
pass 2: [301] [] [129] [937 438] [742] [751] [863 265] [076] [] [694]
pass 3: [075] [129] [265] [301] [438] [] [694] [742 751] [863] [937]
final sorted array: 075 129 265 301 438 694 742 751 863 937

6. (10 scores)

(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 256KB broken into blocks of 8192 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. (6 scores) Since working memory is 256KB and the blocksize is 8KB, the working memory holds 32 blocks. The expected runlength is 512KB, so a single pass of multiway merge forms runs of length 512KB*32=16MB. The second pass then forms a run as large as 16MB*32=512MB.

7. 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 interleaving factor is 3. The disk turns at 7200rmp (8.3ms/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!!!) (6cores)

A cluster holds 16*0.5K = 8K. Thus, the file requires 360/8=45clusters.

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 *8.3, and cluster rotation time is 3 * (16/144)*8.3.Seek time for the total file read time is

45* (80 + 0.5 * 8.3+ 3 * (16/144)*8.3) = 3911.25

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. (10 scores)

$H_1(22)=0$, $H_1(41)=2$, $H_1(53)=5$, $H_1(46)=6$, no conflict

When $H_1(30)=2$, $H_2(30)=1$ （2+1*1）%11=3，so 30 enters the 3rd slot;

$H_1(13)=6$, $H_2(13)=2$ (6+1*2)%11=8, so 13 enters the 8th slot;

$H_1(1)=3$, $H_2(1)=8$ (3+5*8)%11= 10 so 1 enters 10 (pass by 0, 8, 5, 2);

$H_1(67)=3$, $H_2(67)=10$ (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. (11scores)

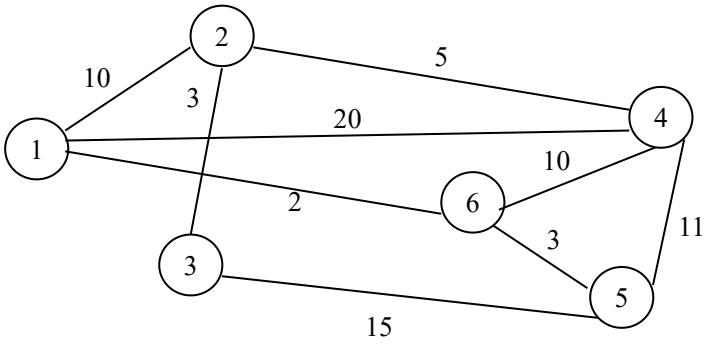


Figure 1 Example graph

(1) Draw the adjacency matrix representation and adjacency list representation for the graph of the figure-1. And if a pointer requires four bytes, a vertex label requires two bytes, and an edge weight requires two bytes, which representation requires more space for this graph? (8 scores)

(a) adjacency matrix (3 scores)

	1	2	3	4	5	6
1		10		20		2
2	10		3	5		
3		3			15	
4	20	5			11	10

```
5 |       15 11     3     |
6 | 2       10  3         |
-------------------------
```

(b) adjacency list: (3 scores)

1 -> 2(10) -> 4(20) -> 6(2) -> \

2 -> 1(10) -> 3(3) -> 4(5) -> \

3 -> 2(3) -> 5(15) -> \

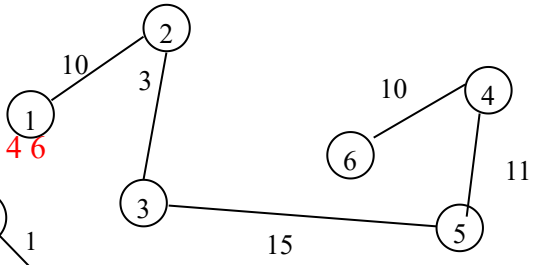
4 -> 1(20) -> 2(5) -> 5(11) -> 6(10) -> \

5 -> 3(15) -> 4(11) -> 6(3) -> \

6 -> 1(2) -> 4(10) -> 5(3) -> \

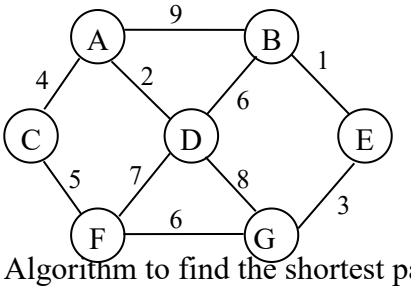
(c) (2 scores)Space of adjacency matrix: 2*36= 72 (bytes); Space of adjacency list 4*6+(2+2+4)*16= 152 (bytes); So adjacency list requires more space for this graph.

(2) Show the DFS tree for the example graph, starting at Vertex 1. (3 scores)



Order that Vertexs are processed: 1 2 3 5 4 6

10. (7 scores)

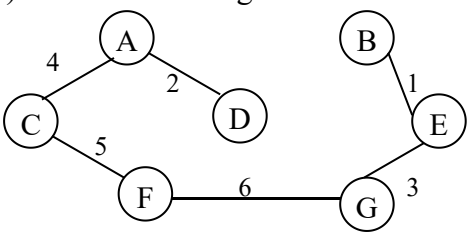


(1) Use Dijkstra's Algorithm to find the shortest paths from C to all other vertices. (4)

C to A: 4 (C,A); CF: 5(C,F); CD: 6(C,A,D); CB: 12(C,A,D,B);

CG:11 (C,F,G); CE: 13(C,A,D,B,E)

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



OR