

# 上海科技大学 2018 年攻读硕士学位研究生

## 招生考试试题

科目代码：991

科目名称：数据结构与算法

考生须知：

1. 本试卷满分为 150 分，全部考试时间总计 180 分钟。
2. 所有答案必须写在答题纸上，写在试题纸上或草稿纸上均无效。
3. 每道题的中文部分均已翻译为英文，考生可在中英文中任选一种语言作答。

### 1. True or False (5 problems, 2 points each) 判断题（5 题，每题 2 分）

Please indicate in the answer sheet whether each statement is true or false. Write down “T” for being true and “F” for being false.

请在答题纸上写明下列每个命题的真假。真则写“T”，假则写“F”。

1. Let  $f(n) = n^3 - 4n + 4$  and  $g(n) = 5n^3 - 100$ , then  $f(n) + g(n)$  is  $\Omega(n^3)$  and  $f(n) \cdot g(n)$  is  $o(n^6)$ .  
若函数  $f(n) = n^3 - 4n + 4$  以及  $g(n) = 5n^3 - 100$ , 则  $f(n) + g(n)$  是  $\Omega(n^3)$  并且  $f(n) \cdot g(n)$  是  $o(n^6)$ .
2. Using a simple uniform hashing function  $h$  to hash  $n$  distinct keys into an array of length  $m$ , the expected cardinality of  $\{k, l: k \neq l \text{ and } h(k) = h(l)\}$  is  $n/m$ .  
用简单均匀的哈希函数将  $n$  个不同的 keys 映射到一个长度为  $m$  的数组，集合  $\{k, l: k \neq l \text{ and } h(k) = h(l)\}$  的期望大小是  $n/m$ .
3. A directed acyclic graph with  $n$  nodes has at most  $n(n-1)/2$  edges.  
一个有  $n$  个节点的有向无环图最多有  $n(n-1)/2$  条边。
4. In any depth-first search of a graph  $G$ , if the finishing time of  $u$  is later than the finishing time of  $v$  for two vertices  $u$  and  $v$  in  $G$ , and  $u$  and  $v$  are in the same DFS tree, then  $u$  is an ancestor of  $v$  in the depth first tree.  
在图深度优先遍历 DFS 算法中，对于图  $G$  任意两点节点  $u$  和  $v$ ，如果  $u$  的结束时间大于  $v$  的结束时间，并且  $u$  和  $v$  在同一个 DFS 树中，那么在此 DFS 树中  $u$  是  $v$  的先驱。
5. Given a boolean formula  $F$  of length  $n$  defined over 100 variables, deciding if  $F$  is satisfiable is NP-complete, assuming  $P \neq NP$ .

给定一个长度为  $n$ 、包含 100 个变量的布尔公式  $F$ , 判断  $F$  是否可满足是 NP-complete, 假设  $P \neq NP$ .

## 2. Multiple Choices – Select One (15 problems, 2 points each) 单选题 (15 题, 每题 2 分)

Each question has only one correct choice. Please indicate the correct choice in the answer sheet.  
每题只有一个正确选项。请在答题纸上写下正确选项的序号。

1. Suppose a stack is implemented with a linear linked list that has just one pointer variable pointing to the first element of the list (the top of the stack). Which of the following correctly gives the time complexity of the (1) push, and (2) pop operations in this implementation?

假设一个栈由一个线性链表实现, 其中仅有一个指针指向链表的第一个元素 (栈顶)。

下面哪一个关于进栈和出栈操作的算法复杂度是正确的?

- A. (1)  $O(1)$  (2)  $O(1)$
- B. (1)  $O(1)$  (2)  $O(n)$
- C. (1)  $O(n)$  (2)  $O(1)$
- D. (1)  $O(n)$  (2)  $O(n)$
- E. (1)  $O(\log n)$  (2)  $O(1)$

2. Which of the following is most efficient for updating a list that contains integers and is of predefined size?

- A. A stack
- B. A linked list
- C. An array
- D. A sequential file
- E. A binary search tree

下面哪种数据结构对更新由固定长度的整数组成的列表效率最高:

- A. 栈
- B. 链表
- C. 数组
- D. 顺序文件
- E. 二叉搜索树

3. Suppose that stacks and queues are provided opaque data types, offering only operations to add elements, to remove elements, and to test for emptiness. Suppose that a programmer wants to count the number of elements in a given stack or queue  $C$ , which is currently in some state  $t$ , using only one auxiliary stack or queue  $D$ . The structures  $C$  and  $D$  can be used in any way possible based on the methods they offer, but  $C$  must be restored to its state  $t$  after counting its elements. Counting elements as described above is possible for which of the

following data types?

- I. C is a queue and D is a queue.
  - II. C is a stack and D is a stack.
  - III. C is a queue and D is a stack.
- 
- A. None
  - B. I and II only
  - C. I and III only
  - D. II and III only
  - E. I, II, and III

假设栈和队列是已提供的非透明数据类型, 仅实现了元素增加、元素删除、以及是否为空的测试。一个程序员要计算一个栈或者队列 C 的元素个数, 而 C 当前的状态是 t 并且只能使用一个辅助的栈或队列 D。C 和 D 可以被任何合理的方式使用, 但计数之后 C 必须恢复到原来的状态 t。下面哪些选项可以实现上述的计数操作?

- I. C 是队列并且 D 是队列
  - II. C 是栈并且 D 是栈
  - III. C 是队列并且 D 是栈
- 
- A. 无选项
  - B. I 和 II
  - C. I 和 III
  - D. II 和 III
  - E. I, II 和 III
- 
4. A hash function  $h$  maps 16-bit inputs to 8-bit hash values. What is the largest  $k$  such that in any set of 1,000 different inputs, there are at least  $k$  inputs that  $h$  maps to the same hash value?
- 一个哈希函数  $h$  将 16 比特的输入映射到一个 8 比特的哈希值。假设给定任意一个包含 1000 个不同输入的集合, 都至少有  $k$  个输入会被  $h$  映射到同一个哈希值。这个  $k$  的最大值是多少?
- A. 3
  - B. 4
  - C. 64
  - D. 256
- 
5. Which of the following data structure is most efficient to schedule jobs on a shared computer, which keeps track of the jobs to be performed and their relative priorities?
- A. Priority queues
  - B. Array

- C. Linked list
- D. Stack

以下哪种数据结构可以用来有效的调度一个共享计算机上的工作，其可以跟踪需要完成的工作和它们的相对优先级？

- A. 优先队列
- B. 数组
- C. 链表
- D. 堆栈

6. What is the time-complexity to insert an element into a  $n$ -element priority queue?

在一个长度为  $n$  的优先队列中插入一个新元素的复杂度是？

- A.  $O(1)$
- B.  $O(\lg n)$
- C.  $O(n)$
- D.  $O(n \lg n)$

7. A full binary tree is a rooted tree in which every internal node has exactly two children. How many internal nodes are there in a full binary tree with 50 leaves?

一棵满二叉树是一棵有根树，其所有非叶节点都有两个孩子节点。一棵有 50 个叶节点的满二叉树有几个非叶节点？

- A. 25
- B. 49
- C. 50
- D. 51
- E. 100

8. What are the indexes of the child nodes whose parent index is  $n$  in a binary tree stored in an array? Assume the index of the root node is 0.

在一棵存储于数组的二叉树中，索引号为  $n$  的节点的孩子节点的索引号是多少？假定根节点的索引号为 0。

- A.  $n \times 2, n \times 2 + 1$
- B.  $n^2 + 1, n^2 + 2$
- C.  $n \times 2 + 1, n \times 2 + 2$
- D.  $2^n, 2^n + 1$

9. Suppose we have a disjoint set (with the forest representation) containing 6 elements. Initially its parent array is [0,1,2,3,4], i.e., each element belongs to a different set. Which of the following can be the parent array after we perform a sequence of union-by-rank (or

equivalently, union-by-height) operations?

假设我们有一个包含 6 个元素、使用森林表示法的不相交集合（并查集）。其初始父亲数组为[0,1,2,3,4]，即每个元素属于一个不同的集合。当我们进行一系列按秩合并（或等价地，按高度合并）操作之后，下列哪一个可能是最终的父亲数组？

- A. [1,1,1,4,0]
- B. [2,0,2,0,2]
- C. [1,4,2,0,0]
- D. [3,2,4,4,4]

10. Kruskal's algorithm and Prim's algorithm are two classical algorithms for finding a minimum spanning tree in a graph. Which of the following must be true?

- i. Kruskal's algorithm is a greedy algorithm.
- ii. Kruskal's algorithm is a dynamic programming algorithm.
- iii. Prim's algorithm is a greedy algorithm.
- iv. Prim's algorithm is a divide and conquer algorithm.

- A. i and iii
- B. i and iv
- C. ii and iii
- D. ii and iv

Kruskal 算法和 Prim 算法是计算图中最小生成树的两个经典算法，以下哪些项是肯定正确的？

- i. Kruskal 算法是一种贪心算法。
- ii. Kruskal 算法是一种动态规划算法。
- iii. Prim 算法是一种贪心算法。
- iv. Prim 算法是一种分治算法。

- A. i 和 iii
- B. i 和 iv
- C. ii 和 iii
- D. ii 和 iv

11. Given a weighted, directed graph  $G$  represented by an adjacency matrix  $W=(w_{ij})$  where  $w_{ij}$  is the weight of the edge  $(i,j)$ . Assume that the graph contains no negative-weight cycles and has  $n>3$  vertices, if one uses the matrix multiplication based algorithm to solve the all-pairs shortest-paths problem, let  $W^i$  for  $i \geq 0$  denotes the matrix after  $i$ -times of matrix multiplication, which of the following must be the matrix containing the shortest-path weights?

- i.  $W^2$ , ii.  $W^{n-1}$ , iii.  $W^{2n-1}$ , iv.  $W^{2n}$

- A. i

- B. ii
- C. ii and iii
- D. ii, iii and iv

给定一个以邻接矩阵表示的带权有向图  $W=(w_{ij})$ , 其中  $w_{ij}$  表示  $(i,j)$  的权值。假设该图不包含非负权值环, 并且含有  $n>3$  个节点, 如果采用基于矩阵乘法的算法解决所有节点对间最短路径问题, 以  $W^i$  表示矩阵  $i$  次乘法后的结果, 以下那些矩阵肯定包含最短路径的权值?

i.  $W^2$ , ii.  $W^{n-1}$ , iii.  $W^{2n-1}$ , iv.  $W^{2n}$

- A. i
- B. ii
- C. ii 和 iii
- D. ii, iii 和 iv

12. Which of the following must be true?

- i. Given a directed graph  $G=(V,E)$ , an associated graph of  $G$  is a graph  $G'=(V',E')$ , where the vertices in  $V'$  are the strongly connected components of  $G$  and for vertices  $S,T \in V'$ ,  $(S,T) \in E'$  if and only if there exist  $u \in S$  and  $v \in T$  such that  $(u,v) \in E$ . Then,  $G'$  is a directed acyclic graph.
- ii. Consider two positively weighted graphs  $G=(V,E,W)$  and  $G'=(V,E,W')$  with the same vertices  $V$  and edges  $E$  such that, for any edge  $e \in E$ , we have  $W(e) = W'(e) * W'(e)$ . For any two vertices  $u,v \in V$ , any shortest path between  $u$  and  $v$  in  $G'$  is also a shortest path in  $G$ .
- iii. Consider two positively weighted graphs  $G=(V,E,W)$  and  $G'=(V,E,W')$  with the same vertices  $V$  and edges  $E$  such that, for any edge  $e \in E$ , we have  $W(e) = W'(e) + W'(e)$ . For any two vertices  $u,v \in V$ , any shortest path between  $u$  and  $v$  in  $G'$  is also a shortest path in  $G$ .

- A. only i
- B. i and ii
- C. i and iii
- D. ii and iii

以下哪些项是肯定正确的?

- i. 给定一个有向图  $G=(V,E)$ ,  $G$  的关联图是一个图  $G'=(V',E')$ , 其中  $V'$  中的节点是  $G$  的强连通分量, 对于任意一对节点  $S,T \in V'$ ,  $(S,T) \in E'$  当且仅当存在  $u \in S$  和  $v \in T$  满足  $(u,v) \in E$ .  $G'$  是一个有向无环图。
- ii. 考虑带权图  $G=(V,E,W)$  和  $G'=(V,E,W')$ , 所有  $W$  和  $W'$  中的权值为正数, 如果  $G$  和  $G'$  有相同的节点和边, 并且对于任意边  $e$  满足  $W(e) = W'(e) * W'(e)$ , 那么对于任意节点对  $u,v \in V$ ,  $G'$  中的  $u$  和  $v$  之间的最短路径同样是  $G$  中的  $u$  和  $v$  之间的最短路径。
- iii. 考虑带权图  $G=(V,E,W)$  和  $G'=(V,E,W')$ , 所有  $W$  和  $W'$  中的权值为正数, 如果  $G$  和  $G'$  有相同的节点和边, 并且对于任意边  $e$  满足  $W(e) = W'(e) + W'(e)$ , 那么对于

任意节点对  $u, v \in V$ ,  $G'$  中的  $u$  和  $v$  之间的最短路径同样是  $G$  中的  $u$  和  $v$  之间的最短路径。

- A. i
- B. i 和 ii
- C. i 和 iii
- D. ii 和 iii

13. What is the worst case time complexity for quicksort to sort an array of  $n$  elements?

对包括  $n$  个元素的数组采用 quicksort 算法进行排序, 其最坏时间复杂度是?

- A.  $O(n \log n)$
- B.  $O(n)$
- C.  $O(n^2)$
- D.  $O(\log n)$

14. When is insertion sort a good choice for sorting an array?

- A. When the array has many elements.
- B. When the array has only a few elements out of place.
- C. When the inputs are strings instead of numbers.
- D. Insertion sort is never a good choice for sorting an array.

对于数组排序, 以下哪种情况下应该选择插入排序?

- A. 当数组很大时。
- B. 当数组中只有少数元素不在其排序后的位置时。
- C. 当数组内容是字符串而非数值时。
- D. 插入排序永远不是一个好的选择。

15. If an algorithm works by solving non-overlapping subproblems, the algorithm is called:

- A. Recursive
- B. Dynamic programming
- C. Divide and conquer
- D. Greedy

如果一个算法通过将问题划分成若干个没有重叠的子问题进行解决, 那么该算法称为:

- A. 递归法
- B. 动态规划
- C. 分治法
- D. 贪心算法

**3. Multiple Choices – Select One or More (5 problems, 2 points each) 多选题 (5 题, 每题 2 分)**

Each question may have one or more correct choices. Please indicate all the correct choice(s) in the answer sheet.

每题有一个或多个正确选项。请在答题纸上写下所有正确选项。

1. Read the following statements, and choose the statements that is (are) correct:

阅读下面的陈述, 并选择正确的选项:

- A.  $\log_2 n = O(\log_{10} n)$
- B.  $\log_{10} n = O(\log_2 n)$
- C.  $n^2 = O(2^n)$
- D.  $2^n = O(n^2)$
- E. None of the above

2. In the character-coding problem, a file contains only characters a, b, c with frequencies 45, 13, 12 respectively, which of the following codewords is an optimal character code for the file?

在一个字符编码问题中, 如果一个文件只由 a, b, c 组成, 它们出现的频率分别是 45, 13, 12, 以下哪些编码是最优的?

- A. a:000, b:010, c:101.
- B. a:000, b:010, c:100.
- C. a:00, b:01, c:10.
- D. a:1, b:01, c:00.

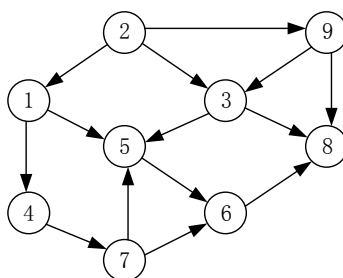
3. Given a binary tree whose pre-order traversal is ABCDE and post-order traversal is CBEDA, which of the following can be its in-order traversal?

一棵二叉树的前序遍历是 ABCDE, 后序遍历是 CBEDA, 它的中序遍历有可能是哪个?

- A. BACDE
- B. BCAED
- C. CBAED
- D. ACBED

4. Consider the following directed graph.

考虑以下有向图。





Which are topological sorts of the nodes of the above graph?

以下哪一个或哪些是上图的拓扑排序?

- A. 2,1,9,3,4,7,5,6,8
- B. 2,1,9,4,5,7,3,6,8
- C. 2,9,1,4,7,3,5,6,8
- D. 2,1,4,7,5,9,3,6,8
- E. 2,9,1,4,3,7,5,6,8

5. Which of the following is true of problems solvable by dynamic programming?

- A. They cannot be solved by a greedy algorithm.
- B. They have the optimal substructure property.
- C. They can be divided into overlapping subproblems.
- D. The time and space complexities of the dynamic programming algorithm are equal.

对于可以通过动态规划可以解决的问题, 以下哪些表述是正确的?

- A. 这些问题无法通过贪心算法解决。
- B. 这些问题具有最优子结构性质。
- C. 这些问题可以划分成若干个重叠子问题。
- D. 动态规划的时间复杂度和空间复杂度相等。

#### 4. Stack (20 points) 栈 (20 分)

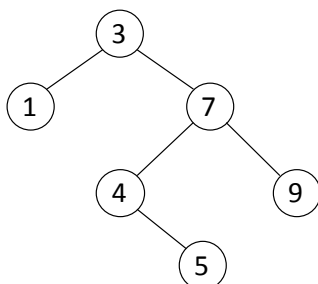
Design a stack which, in addition to *push* and *pop*, also has a function *min* which returns the minimum element. *Push*, *pop* and *min* should all operate in  $O(1)$  time. Please describe your design.

设计一个栈, 除了提供入栈和出栈操作, 还提供 *min* 函数, 返回栈里的最小元素。所有三个函数都应该具有  $O(1)$  的复杂度。请描述你的设计。

#### 5. Binary Search Tree (20 points) 二叉搜索树 (20 分)

Consider the following binary search tree:

根据以下二叉搜索树回答:



1) Starting from an empty binary search tree, the insertion of which of the following sequences of integer keys could produce the binary search tree above?

从一个空的二叉搜索树开始, 以下哪个插入序列可以生成以上二叉搜索树? (4 分)

- A. 3, 7, 1, 5, 9, 4
- B. 3, 7, 4, 1, 5, 9
- C. 3, 4, 7, 5, 9, 1
- D. 3, 1, 7, 9, 5, 4

2) What is the in-order traversal of the tree?

以上二叉树的中序遍历是? (4 分)

3) Draw the updated tree when the root of key 3 is deleted from the tree.

画出把树中根节点 3 去掉后的二叉搜索树。(6 分)

4) Draw the updated tree when the node of key 3 is re-inserted back to the tree.

画出把 3 节点重新插入到第 3) 小问所得树之后的二叉搜索树。(6 分)

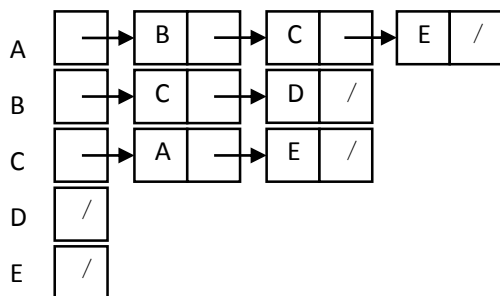
## 6. Graph Representation (20 points) 图的表示 (20 分)

Given the following adjacent-list representation of a graph,

- 1) draw the graph;
- 2) draw the adjacent-matrix representation of the graph.

给定一个图的邻接表表示,

- 1) 画出该图; (10 分)
- 2) 画出该图的邻接矩阵表示。(10 分)



## 7. Graph Traversal (20 points) 图的遍历 (20 分)

Consider the following modified traversal algorithm for graphs:

考虑以下修改后的图遍历算法:

- 1: Visited = { };
- 2: ModTraversal (v: Node, G: Graph) {
- 3:     // assume   G.V is the set of nodes in the graph G
- 4:     //           G.E is the set of edges in the graph G
- 5:     Queue active =  $\emptyset$ ;   // a queue that is first in first out
- 6:     **forall** (v,w) **in** G.E **do** { // an ascending alphabetical

```

// order of w' s
7:      if (not  $w \in \text{Visited}$ ) then {
8:          Visit ( $w$ );
9:           $\text{Visited} = \text{Visited} \cup \{w\}$ ;
10:         ENQUEUE( $\text{active}, w$ );
11:     };
12: };
13: while ( $\text{active} \neq \emptyset$ ) do {
14:      $w = \text{DEQUEUE}(\text{active})$ ;
15:     ModTraversal ( $w, G$ );
16: };
17: }
```

- 1) Consider the following graph: in what order are the nodes "visited" by the modified traversal at Line 8? The traversal shall be called by

```

Visited = Visited  $\cup$  { $a$ }; // Visited={ $a$ }
ModTraversal ( $a, G$ );
```

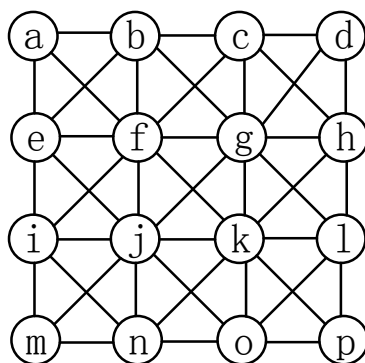
( $a$  being the start node for the traversal, and edges connecting from the same node are visited in an ascending order at Line 6 according to the target nodes of the edges. For example,  $a$  has three edges ( $a,b$ ), ( $a,e$ ), ( $a,f$ ), since  $b < e < f$ , ( $a,b$ ) should be visited before ( $a,e$ ) and ( $a,f$ ) at Line 6, while ( $a,e$ ) should be visited before ( $a,f$ ) at Line 6)

给出下图用上述遍历算法遍历时的节点访问顺序, 遍历算法采用以下方式调用:

```

Visited = Visited  $\cup$  { $a$ }; // Visited={ $a$ }
ModTraversal ( $a, G$ );
```

( $a$  是遍历的起始节点; 当存在多个同一个节点出发的边时, 上述算法行 6 以边的目标节点的升序访问。比如, 连接节点  $a$  的边有三条 ( $a,b$ ), ( $a,e$ ), ( $a,f$ ), 由于  $b < e < f$ , ( $a,b$ ) 必须在 ( $a,e$ ) 和 ( $a,f$ ) 前访问, ( $a,e$ ) 必须在 ( $a,f$ ) 前访问) (8 分)



- 2) Draw the spanning tree computed by ModTraversal of ( $a$ )  
画出上述算法在上图的遍历生成树 (6 分)
- 3) Assumes that the input graph  $G$  is represented using adjacency lists, set operations (e.g., membership checking and element adding) and ENQUEUE and DEQUEUE can be done in

constant time (i.e.,  $O(1)$ ), give the worst-time complexity (Big O notation) of ModTraversal in terms of  $|V|$  and  $|E|$ .

假设上述算法的输入图是以邻接链表表示, 集合的操作 (检查是否包含某个节点和集合中加入节点) 和入队、出队都在常数时间内完成 (即  $O(1)$ ), 使用  $|V|$  和  $|E|$  表示上述遍历算法的最坏时间复杂度 (大 O 表示法)。(6 分)

### 8. Sorting (20 points) 排序 (20 分)

Suppose we are given an array of  $n$  numbers, where each number is at most  $k$  positions away from its position if we sort the array, for a constant  $k \leq n$ . For example, given the array  $[2\ 3\ 1\ 4\ 6\ 5\ 7\ 9\ 8]$ , each value is at most 2 positions away from its sorted position. Give an algorithm to sort the array running in  $O(n \log k)$  time. Write pseudo-code for your algorithm and also explain it in words.

假设给定一个长度为  $n$  的数组, 并且数组中每个值的位置距离排序后该值的位置不超过  $k$  (小于或等于  $k$ ),  $k \leq n$ 。比如数组  $[2\ 3\ 1\ 4\ 6\ 5\ 7\ 9\ 8]$ , 每个值的位置距离其排序后的位置不超过 2。设计一个最坏时间复杂度为  $O(n \log k)$  的排序算法, 以伪代码形式给出, 并解释该程序。

计算机/软件工程专业

每个学校的

考研真题/复试资料/考研经验

考研资讯/报录比/分数线

免费分享



微信 扫一扫

关注微信公众号

计算机与软件考研