

上海科技大学 2019 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：

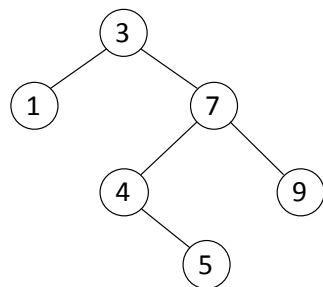
1. 本试卷满分为 150 分，全部考试时间总计 180 分钟。
 2. 所有答案必须写在答题纸上，写在试题纸上或草稿纸上均无效。
 3. 每道题的英文部分均已翻译为中文，考生可在中英文中任选一种语言作答。
-

1. True or False (10 problems, 2 points each) 判断题（10 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。真则打“√”，假则打“×”。

1. In a circular linked list, some link fields may be null. ()
在循环链表中，某些链接域可能为空。()
2. Given any functions $f(n)$ and $g(n)$, it is possible to have both $f(n) = \Omega(g(n))$ and $f(n) = o(g(n))$. ()
给定任意函数 $f(n)$ 和 $g(n)$ ， $f(n) = \Omega(g(n))$ 和 $f(n) = o(g(n))$ 可能同时成立。()
3. A good hash function of a hash table satisfies the assumption of simple uniform hashing. ()
一个好的哈希函数需满足简单均匀。()
4. The following tree is a binary search tree. ()
下列树是二叉搜索树。()



5. The number of nodes in a tree can be more than twice the number of leaf nodes. ()
一棵树的节点个数有可能大于叶节点个数的两倍。()
6. The space complexity of an adjacency-matrix representation of a graph is independent of the number of edges in the graph. ()
图的邻接矩阵表示的空间复杂度与图的边数无关。()
7. In the breadth-first search procedure of graphs, each vertex must be enqueued exactly once. ()
在图的广度遍历算法中, 每个节点恰好仅入队一次。()
8. Given a weighted, directed graph with nodes numbered $1, 2, \dots, n$, finding the length of a shortest path from vertex 1 to vertex n or determining that no such path exists can be done in polynomial time. ()
给定一个带权有向图, 图的节点为 $1, 2, \dots, n$, 要计算节点 1 和 n 之间的最短路径或判断这样的路径不存在, 这个问题可以在多项式时间内解决。()
9. Given a graph, deciding whether it has a clique of 100 nodes is NP-complete, assuming $P \neq NP$. ()
给定一个图, 确定图中是否存在一个正好包含 100 个节点的团是一个 NP-complete 的问题, 假设 $P \neq NP$ 。()
10. Given a graph, deciding whether it is possible to remove half the nodes in the graph so that the remaining graph is 3-colorable is solvable in polynomial time, assuming $P \neq NP$. ()
给定一个图, 确定是否可以去除图中一半的节点后使得该图可以三染色, 这可以在多项式时间内完成, 假设 $P \neq NP$ 。()

2. Multiple Choices – Select One (10 problems, 2 points each) 单选题 (10 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. In what kind of storage structure for strings can one easily insert, delete, concatenate, and

rearrange substrings? ()

- A. Fixed length storage structure
- B. Linked list storage
- C. Variable length storage with fixed maximum
- D. Array list storage
- E. None of the above

哪种字符串的存储结构可以方便地进行插入, 删除, 并联以及重新安排子字符串? ()

- A. 固定长度的存储结构
- B. 链表存储结构
- C. 具有固定最大长度的变长存储结构
- D. 数组存储结构
- E. 以上都不是

2. In which of the following are the functions listed in order of increasing asymptotic size? ()

下面哪个选项中的函数是按照渐进大小的增序排列? ()

- A. $2^n, n + (\log n)^2, n^3, 10n^2, n \log n$
- B. $n + (\log n)^2, n^3, 10n^2, n \log n, 2^n$
- C. $n + (\log n)^2, n \log n, 10n^2, n^3, 2^n$
- D. $n + (\log n)^2, 10n^2, n \log n, n^3, 2^n$
- E. $2^n, n^3, 10n^2, n \log n, n + (\log n)^2$

3. We store 10 elements with keys {5,28,19,15,20,12,33,17,10,18} into a hash table with collisions resolved by chaining (with linked list). The hash table has 9 slots and the hash function is defined as $h(k) = k \bmod 9$. What is the maximum length of the linked lists in the hash table? ()

存储 10 个元素到一个哈希表, 这 10 个元素的 key 是{5,28,19,15,20,12,33,17,10,18}。哈希表总共有 9 个 slots, 哈希函数是 $h(k) = k \bmod 9$, 并用链表解决冲突。哈希表中最长的链表长度是? ()

- A. 1
- B. 2
- C. 3
- D. 4

4. 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 500 leaves? ()

满二叉树的所有中间节点都有两个孩子节点。一个有 500 个叶子节点的满二叉树有多少个中间节点? ()

- A. 250
- B. 499
- C. 500
- D. 501
- E. 1,000

5. Which of the following statements about trees is false? ()

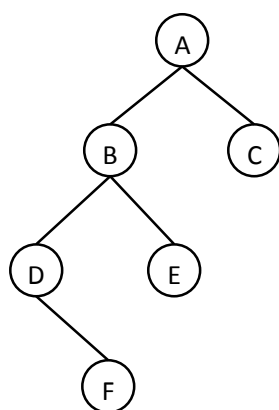
- A. An internal node may have no child
- B. Adding an edge to a tree creates a cycle
- C. Not every node has a parent node in a tree
- D. The height of a tree with one node is 0

以下哪一个关于树的描述是错误的? ()

- A. 非叶节点有可能没有孩子
- B. 在一棵树中加一条边会形成一个环
- C. 一棵树中并非每个节点都有父亲节点
- D. 一棵包含一个节点的树高度为 0

6. What is the in-order traversal of the following binary tree? ()

下面二叉树的中序遍历是什么? ()



- A. DFBEAC
- B. ABDFEC
- C. FDEBCA
- D. FDEBAC

7. Prim's algorithm is one algorithm for finding a minimum spanning tree in a graph and the Floyd-Warshall algorithm can be used to solve the all-pairs shortest-paths problem on a directed graph. Which of the following must be true? ()
- i. Prim's algorithm is a greedy algorithm.
 - ii. Prim's algorithm is a dynamic programming algorithm.
 - iii. Floyd-Warshall algorithm is a greedy algorithm.
 - iv. Floyd-Warshall algorithm is a dynamic programming algorithm.

Prim 算法是一种计算图的最小生成树算法, Floyd-Warshall 算法可用于计算有向图中所有节点对之间的最短路径。以下哪些表述肯定是正确的? ()

- i. Prim 算法是一种贪心算法。
 - ii. Prim 算法是一种动态规划算法。
 - iii. Floyd-Warshall 算法是一种贪心算法。
 - iv. Floyd-Warshall 算法是一种动态规划算法。
- A. i 和 iii
B. i 和 iv
C. ii 和 iii
D. ii 和 iv
8. Among the following problems concerning a given graph G, which is currently known to be solvable in linear time? ()
- A. Finding a longest simple cycle in an undirected graph
 - B. Computing the strongly connected components of a directed graph
 - C. Solving the single-source shortest-paths problem of a weighted, directed graph G
 - D. Finding a largest clique in an undirected graph G

以下哪一个图的问题可以在线性时间内解决? ()

- A. 找到一个无向图中的最长简单环
 - B. 计算有向图的强连通分量
 - C. 解决带权有向图的单源最短路径问题
 - D. 找到一个无向图的最大团
9. Which of the following is true of merge sort and quicksort? ()
- A. Both are divide and conquer algorithms

- B. Both take $\Theta(n)$ time in each divide step, given an input array of size n
- C. Both have the same worst-case time complexity
- D. All of the above are true

下列关于归并排序和快速排序的陈述哪一项是对的? ()

- A. 都是用分治法
- B. 给定一个长度为 n 的输入数组, 都是需要 $\Theta(n)$ 的时间进行每一次切分
- C. 都有相同的最坏时间复杂度
- D. 以上都对

10. Which of the following is the closest to the minimum number of comparisons between array elements that any comparison-based algorithm needs to perform to find both the minimum and maximum values in an input array of size n ? ()

任何一个基于比较的算法在一个长度为 n 的数组中同时找出最大和最小值, 至少需要数组元素之间的比较的次数与下列哪个最接近? ()

- A. n
- B. $1.5n$
- C. $2n$
- D. n^2

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. Consider using a linked list to represent a sparse matrix. Only non-zero elements are stored in the list. What information should be stored in the nodes of the linked list? ()

- A. Row number
- B. Column number
- C. Value of the matrix element
- D. Number of non-zero elements
- E. Number of zero elements

假设使用一个链表来表示一个稀疏矩阵, 其中只存储非零元素。在链表的节点中需要存

储什么信息? ()

- A. 行号
- B. 列号
- C. 矩阵元素的值
- D. 非零元素的个数
- E. 零元素的个数

2. In a character-coding problem, if a file contains only characters 'a', 'b', and 'c' with frequencies 45, 13, and 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

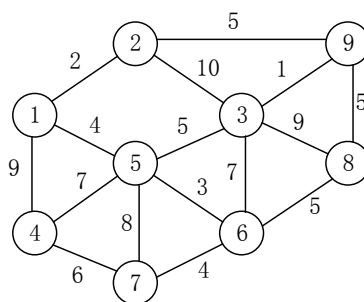
3. Given an undirected graph with n nodes and m edges, select the correct statement(s) below. ()

- A. $m = n - 1$ if the graph is a tree
- B. $n = m - 1$ if the graph is a tree
- C. $m < n$ if the graph is a forest
- D. $n < m$ if the graph is a forest
- E. None of the above

给定一个具有 n 个节点和 m 条边的无向图, 选择下列正确的称述? ()

- A. 如果该图是一棵树, 那么 $m = n - 1$
- B. 如果该图是一棵树, 那么 $n = m - 1$
- C. 如果该图是一个森林, 那么 $m < n$
- D. 如果该图是一个森林, 那么 $n < m$
- E. 以上都不正确

4. Consider the following weighed, undirected graph. Which of the following must be true? ()



- A. The above graph has four minimum spanning trees.
- B. There is at least one minimum spanning tree containing the edge (6,8)
- C. There is at least one minimum spanning tree containing the edge (5,6)
- D. All the minimum spanning trees contain the edge (2,9)
- E. All the minimum spanning trees contain the edge (4,7)

考虑以下带权无向图，以下哪一个或哪些表述是正确的？（ ）

- A. 上图有四个最小生成树
- B. 存在一个包含边 (6,8) 的最小生成树
- C. 存在一个包含边 (5,6) 的最小生成树
- D. 所有最小生成树都包含边 (2,9)
- E. 所有最小生成树都包含边 (4,7)

5. Write $X \leq_p Y$ to mean that a problem X is polynomial-time reducible to problem Y . Suppose that a problem X is in P , a problem Y is in NP , and $X \leq_p Y$. Which of the following **must** be true? ()
- A. X is in NP
- B. X is NP -complete
- C. Y is NP -complete
- D. If a problem Z is in P , then $Z \leq_p X$

用 $X \leq_p Y$ 表示问题 X 可以多项式时间规约到问题 Y 。假设问题 X 属于 P , 问题 Y 属于 NP , 同时 $X \leq_p Y$ 。下列哪些项一定为真？（ ）

- A. X 属于 NP
- B. X 是 NP -complete
- C. Y 是 NP -complete
- D. 假定 Z 属于 P , 那么 $Z \leq_p X$

4. Lists (25 points) 列表 (25 分)

Given the following array of distinct keys (the index of the array starts from 0):

给定如下不含重复关键字的数组 (数组的索引从 0 开始):

K1	K2	K3	K4	K5
----	----	----	----	----

answer the following questions.

回答以下问题:

- 1) What is the average time complexity of inserting a new key into the array? Use Θ notation to represent the complexity and assume that the number of existing keys in the array is N and the array is not full. (5 points)

在如上数组中插入一个新的数据的平均时间复杂度是多少? 用 Θ 表示这个时间复杂度, 同时假设这个数组原来存放有 N 个数据、而且不满。(5 分)

- 2) Use two arrays to represent the keys in the above array as a singly-linked list (but keep the same order as in the array), and fill in the values in the two arrays (Hint: one array stores the keys and the other stores the pointers to the next key, use "/" to represent the NULL pointer). (5 points)

用两个数组将上述给定的数组中的数据表示成一个单链表 (需保持数据在原数组中的顺序), 在两个数组中填写对应的信息 (提示: 其中一个数组存放关键字, 另一个存放对应的指向下一个关键字的指针, 用 "/" 表示 NULL 指针)。(5 分)

Keys:	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
Pointers:	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>

- 3) Remove key K3 from the above two-array representation, and draw the new status of the two arrays (leave a cell blank if it is not occupied by the list). (5 points)

在以上两个数组表示的链表中去掉关键字 K3, 画出去掉之后两个数组的状态 (不被链表占用的空格留空)。(5 分)

Keys:	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
Pointers:	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>

- 4) Following the result of (3), if you are only given the array index of K4 and asked to remove K4 in $\Theta(1)$ time complexity, draw the new status of the two arrays after the removal. (5 points)

根据 3) 的结果, 只给定了 K4 在关键字数组里的索引, 要求把 K4 用 $\Theta(1)$ 时间复杂度从链表中删除, 画出删除后的两个数组状态。(5 分)

Keys:

--	--	--	--	--

Pointers:

--	--	--	--	--

- 5) Assume that the given keys in the array satisfy a linear order $K1 > K2 > K3 > K4 > K5$. Use a complete binary tree to represent this order as a binary search tree and then represent this binary search tree as an array. Draw the final array (the length of the array is 5). Note: Here a complete binary tree is a binary tree which is completely filled on all levels except possibly the lowest, which is filled from the left up to a point. (5 points)

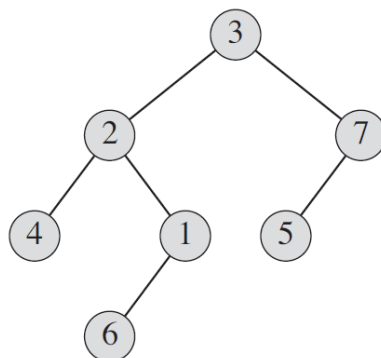
假设给定的关键字满足一个线性关系: $K1 > K2 > K3 > K4 > K5$ 。用一个完全二叉树将这个线性关系表示成一个二叉搜索树,然后将这个二叉搜索树用一个数组来表达,最后画出这个数组(数组的长度是5)。备注:在这里完全二叉树是一个二叉树,其中每一层(可能除最后一层之外)都完全填充,而最后一层从左到右填到某一个位置为止。(5分)

--	--	--	--	--

5. Binary Trees (25 points) 二叉树 (25 分)

- 1) Write down the preorder, inorder, postorder, and breadth-first traversals of the following binary tree in the form of a sequence of nodes. (9 points)

请写出下面这棵二叉树的前序、中序、后序和广度优先遍历,以节点序列的形式给出。(9分)



- 2) Draw the binary tree whose preorder is 1,3,2,7,6,5,4 and whose inorder is 1,2,3,4,5,6,7. (8 points)

一棵二叉树的前序遍历是 1,3,2,7,6,5,4, 中序遍历是 1,2,3,4,5,6,7。请画出这棵树。(8分)

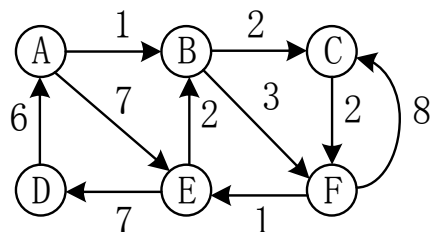
- 3) Draw the binary tree whose postorder is 1,3,2,7,6,5,4 and whose inorder is 1,2,3,4,5,6,7. (8 points)

一棵二叉树的后序遍历是 1,3,2,7,6,5,4, 中序遍历是 1,2,3,4,5,6,7。请画出这棵树。(8分)

6. Graph (25 points) 图 (25 分)

Given the following weighted directed graph,

给定如下带权有向图,



- 1) draw the adjacent-list representation of the graph; (5 points)
画出该图的邻接表表示; (5 分)
- 2) draw the adjacent-matrix representation of the graph; (5 points)
画出该图的邻接矩阵表示; (5 分)
- 3) draw one breadth-first tree of the graph starting from the source A; (5 points)
画出该图的以 A 为出发节点的一棵广度优先树; (5 分)
- 4) draw one shortest-paths tree of the graph rooted at the source A. (10 points)
画出该图的以 A 为出发节点的最短路径树。(10 分)

7. Quicksort (25 points) 快速排序 (25 分)

Suppose we apply the quicksort algorithm to an array E of n integers, to sort the n elements in ascending order. The time complexity is measured in the number of comparisons between array elements, and each comparison takes constant time.

假设我们把快速排序算法用在一个包含 n 个整数的数组 E , 将这 n 个元素排为升序。我们用元素之间的比较次数来衡量时间复杂度, 而每一次元素间的比较需要常数时间。

- 1) Suppose a version of quicksort always chooses the left-most element in an array to be the pivot. Give a worst-case input array of size n for this quicksort algorithm, represented as a permutation of all integers in set $\{1, 2, \dots, n\}$. Derive the asymptotic time complexity of the algorithm on this input in terms of Θ . (8 points)

假设一个快速排序算法的版本总是选数组里最左边的元素作为基准。对这样一个快速排序算法, 找到一个最坏情况的包含 n 个元素的输入数组, 用集合 $\{1, 2, \dots, n\}$ 中的所有整数的一个排列来表示。对这样的输入, 推导这个算法的渐近时间复杂度, 用 Θ 来表示。(8 分)

- 2) As a recursive algorithm, quicksort tends to call itself for many small subarrays, which would compromise the time efficiency. It is known that the Insertion sort algorithm is fast on small input arrays. Describe a method to modify the quicksort algorithm by using Insertion sort, such that the above issue can be addressed. Suppose each subarray on

which Insertion sort is applied has no more than k elements. Show that the modified algorithm has average-case time complexity of $O(nk + n \lg(n/k))$. (9 points)

作为一个递归算法, 快速排序算法一般要对很多小的子数组进行递归调用, 从而影响了时间效率。而插入排序算法在小数组上运行效率很高。描述一个修改快速排序算法的方法, 利用插入排序算法来解决上述的问题。假设插入排序算法处理的每一个子数组的长度不超过 k 个元素。推导证明这个修改过的快速排序算法的平均时间复杂度是 $O(nk + n \lg(n/k))$ 。(9分)

- 3) Suppose that quicksort uses the median of $E[\text{first}]$, $E[\text{middle}]$ and $E[\text{last}]$ as the pivot. Show that the number of comparisons between array elements performed by this quicksort algorithm is no more than $\frac{n^2}{4} + o(n^2)$ in the worst case. (8 points)

假设快速排序算法总是从输入数组的第一个、最后一个、和中间位置这三个元素中选择中位数作为基准。证明该快速排序算法最坏情况下完成排序需作不超过 $\frac{n^2}{4} + o(n^2)$ 次元素间的比较。(8分)