

南京邮电大学
2006 年攻读硕士学位研究生入学考试
数据结构 试题

考生注意：本试卷共 4 页。所有答题均写在答题纸上（包括单选和填空题），请务必准确标明所答题的题号和填空号。

一、单选题（每题 3 分，共 30 分）

1. 可以使用大O记号表示一个算法的时间复杂度。下列标识中不正确的是 D。
A. $n^2 + 2n = O(n^3)$ B. $n\log_2 n + 2n = O(n^2)$
C. $n^2 + n\log n = O(n^2\log_2 n)$ D. $n^2 + n\log_2 n = O(n\log_2 n)$
2. 下列术语中， B 与数据的存储结构无关。
A. 循环队列 B. 堆栈 C. 散列表 D. 单链表
3. 设线性表非空，采用下列 D 所描述的链表可以在 $O(1)$ 时间内在表尾插入一个新结点。
A. 带表头结点的单链表，一个链表指针指向表头结点
B. 带表头结点的单循环链表，一个链表指针指向表头结点
C. 不带表头结点的单链表，一个链表指针指向表的第一个结点
D. 不带表头结点的单循环链表，一个链表指针指向表的第一个节点
4. 设主串为“abceabceabceabceab”，字串为“abceabcea”。则在KMP匹配第一趟失配后下一趟匹配开始时，子串指针指示的字符是 A。
A. a B. b C. c D. e
5. 二叉树中第 5 层上的结点个数最多为 C，假定根节点层次为 1。
A. 8 B. 15 C. 16 D. 32
6. 用DFS遍历一个有向无环图，并在DFS算法退栈返回时打印当前顶点，则输出的顶点序列是 C。
A. 拓扑有序的 B. 无序的
C. 逆拓扑有序的 D. 按顶点编号次序的
7. 下面 A 算法可用于求无向图的所有连通分量
A. 广度优先遍历 B. 拓扑排序
C. 求最短路径 D. 求关键路径
8. 设有以元素 10, 9, 20, 6, 8, 23, 21, 17 为叶结点的 8 路合并胜方树。在输出一个元素后，将有一个新元素补充到相应的叶结点中。在重构的胜方树中，应有 D 个元素需要修正。
A. 1 B. 2 C. 3 D. 4
9. 在一棵二叉搜索树上搜索一个元素的平均时间复杂度为 D。

- A. $O(n^2)$ B. $O(n)$ C. $O(n\log_2 n)$ D. $O(\log_2 n)$

10. 下面关于倒排文件的说法中正确的是 B。

- A. 倒排文件是对主关键字建立索引的
B. 倒排文件是对次关键字建立索引的
C. 倒排序文件的优点是维护简单
D. 采用倒排文件是为了节省存储空间

二、填空题（每题 6 分，共 48 分）

1. 设有二维数组 $A[0..5, 0..6]$ ，从地址 $Loc(A)$ 开始顺序存放，每个元素占 2 个字节。元素 $A[4, 4]$ 在行优先方式下的存储地址为 (1)，在列优先方式下的存储地址为 (2)。

解答：

(1) $Loc(A) + 2 \times 7 \times 4 + 2 \times 4 = Loc(A) + 64$

(2) $Loc(A) + 2 \times 6 \times 4 + 2 \times 4 = Loc(A) + 56$

2. 已知一算术表达式的中缀形式为 $(A+B)*C-E/D$ ，其后缀形式为 (3)。另有一算术表达式的后缀形式为 $2\ 6\ 4\ -\ *\ 2\ /$ ，每个操作数均为 0 至 9 的一位数。此表达式的值为 (4)。

解答：

(3) $AB+C*ED/-$

(4) $2*(6-4)/2=2$

3. 使用如图 1 所示的哈夫曼树实现对字符集 $\{F, G, H, I, J\}$ 的哈夫曼编码。

已知编码后得到的码文为：011101011101，则利用此哈夫曼树进行译码得到的电文为 (5)。设该字符集中各字符的使用频率各不相同，则其中使用频率最小的两个字符应当是 (6)。

解答：

(5) G I H J G

(6) I J

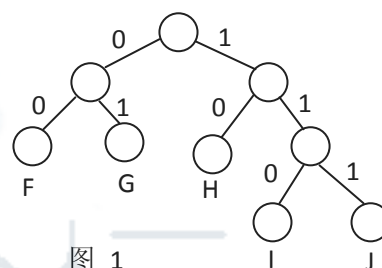


图 1

4. 设有向图如图 2 所示。则所有可能的拓扑有序序列为 (7)。

解答：

(7) ① ② ④ ⑤ ③ ⑥

① ④ ② ⑤ ③ ⑥

① ④ ⑤ ② ③ ⑥

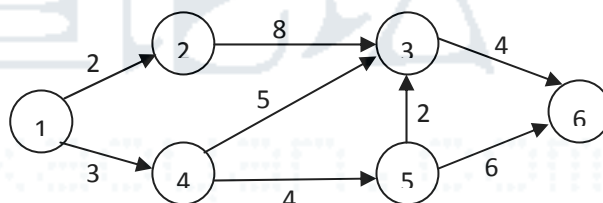


图 2

5. 已知对一棵二叉树的先序和中序遍历的结点次序可以唯一确定该二叉树。设某棵二叉树的先序和中序次序分别为 ABCDEFG 和 BDCAEGF，则其后序次序为 (8)。此外，已知后序和 (9) 序遍历次序也可以唯一确定一棵二叉树。

解答：

(8) D C B G F E A

(9) 中

6. 引入B- 树的最根本原因是____(10)____。在有n个元素的m阶 B- 树中，搜索一个元素的算法需要访问外存的次数至多为____(11)____。

解答：

(10)为了减少外搜索的磁盘访问次数，为修改过程提供简单的平衡算法

$$(11)\log_{(\frac{m}{2}\text{取上整})} \left\lceil \frac{N+1}{2} \right\rceil + 1$$

7. 分别采用直接插入排序和快速排序方法对下面所列出的四个序列进行排序（由小到大）。使得直接插入排序时间最长的序列是____(12)____，使得快速排序时间最短的序列是____(13)____

- A. 10, 20, 30, 40, 50, 60, 70
B. 70, 60, 50, 40, 30, 20, 10
C. 40, 10, 30, 20, 60, 50, 70
D. 40, 20, 10, 30, 50, 70, 60

解答：

(12)B

(13)C D

8. 数据的逻辑结构是指____(14)____，数据的存储结构是指____(15)____。

解答：

(14)对数据逻辑关系的描述

(15)数据在存储器中的存储方法

三、解答题（每题 8 分，共 40 分）

1. 当采用行三元组表存储稀疏矩阵实现矩阵快速转置算法时，需要附设两个一维数组，设为 num 和 k，其中 num[col]表示原矩阵第 col 列中非零元素个数。现有稀疏矩阵如右图所示

(1) 请给出它的转置前后的行三元组表示的示意图；

(2) 计算数组 num 和 k 的元素值。

$$\begin{pmatrix} 0 & 12 & 0 & 0 & 0 \\ -1 & 0 & 13 & 18 & 0 \\ 11 & 20 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

2. 求解下列关于散列表的问题

(1) 除法（除留余数法）散列函数是最常用的一种散列函数，请写出散列函数的一般形式。

(2) 设有散列表 Ht 如下，现采用二次探查法解决冲突。已知 $H(38) = 5$ ， $H(25) = 3$ ，请将 38 和 25 依次插入表中适当位置。请在答题纸上画出完整的插入两个新元素后的散列表。

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|----|---|---|----|----|---|----|----|----|---|---|----|
| Ht | | | 13 | 14 | | 16 | 17 | 18 | | | |

3. 已知 AOE 网如第二题第 4 小题图 2 所示。
- (1) 计算每个顶点代表的事件可能的最早发生时间。
 - (2) 列出计算顶点 4 所代表的事件允许的最迟发生时间的计算式和计算结果。
 - (3) 计算关键路径长度，并列出关键路径的顶点序列。
4. 从空二叉平衡树开始，依次插入：19，25，43，16，18，22。请画出每插入一个元素后的二叉平衡树。
5. 设有向图如图 3 所示，
- (1) 画出邻接表存储表示的示意图；
 - (2) 求它的全部强连通分量。

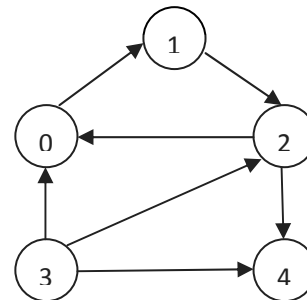


图 3

四、算法设计题（共 32 分）

解题要求：

- (1) 只允许使用 pascal，C 和 C++ 语言中的一种语言描述数据结构和算法。
- (2) 算法描述中不允许直接调用教材上以实现的过程或函数。
- (3) 要求对算法的每一条语句加明确注释。

1. （14 分）

设有 n 个元素保存在一维数组 A 中。枚举排序的基本思想是借助于一个一维数组，设为 $count$ 。 $count[k]$ 记录在初始序列 A 中，比第 k 个元素小的元素个数。因此 $count[k]$ 也表明了第 k 个元素在有序序列中的位置。请设计一个算法计算数组 $count$ 的值。分析此算法的最坏、最好和平均情况时间复杂度。

2. （18 分）

设一棵 n 个结点的完全二叉树采用顺序存储结构，保存在一维数组 A 中。试设计一个递归算法，复制该完全二叉树，得到一棵新的采用普通二叉链表存储的二叉树。二叉链表的每个结点有三个域： $lChild$ ， $rChild$ 和 $element$ 。算法返回所构造的新二叉树的根结点地址。