

2004 年上海海运学院攻读硕士学位研究生入学考试试题

(答案必须做在答题纸上, 做在试题上不给分)

考试科目: 数 据 结 构

一、判断题 (本题 20 分, 每小题 2 分)

- 1 为了很方便地插入和删除数据, 可以使用双向链表存放数据。
- 2 两个栈共享一片连续内存空间时, 为了提高内存利用率, 减少溢出机会, 应把两个栈的栈底分别设在这片内存空间的两端。
- 3 数组是同类型值的集合。
- 4 若一棵二叉树的树叶是某子树中序遍历序列中的第一个结点, 则它必是该子树后序遍历序列中的第一个结点。
- 5 采用链地址法解决冲突时, 若规定插入总是在链首, 则插入任一个元素的时间是相同的。
- 6 在查找树 (二叉排序树) 中插入一个新结点, 总是插入到叶子结点的下面。
- 7 用邻接矩阵存储一个图时, 在不考虑压缩存储的情况下, 所占用的存储空间大小与图中顶点的个数有关, 而与图的边数无关。
- 8 散列表的平均查找长度只与表的长度有关, 而与处理冲突的方法无关。
- 9 顺序查找法适用于存储结构为顺序或链表存储的线性表。
- 10 在外部排序时, 利用选择树方法在能容纳 m 个记录的内存缓冲区中产生的初始归并段的平均长度为 $2m$ 个记录。

二、填空题 (本题 25 分, 每小题 5 分)

- 1 表长为 n 的顺序存储的线性表, 当在任何位置上插入或删除一个元素的概率相同时, 插入一个元素所需移动元素的平均个数为 (1), 删除一个元素所需移动元素的平均个数为 (2)。
- 2 广义表 $A = (a, b, (c, d), (e, (f, g)))$, 取表头和表尾函数分别为 $head()$ 和 $tail()$, 则 $tail(head(tail(tail(A)))) =$ (1), 而从表中取出原子项 e 的运算为 (2)。
- 3 假定一棵树的广义表表示为 $A(B(E), C(F(H, I, J), G), D)$, 则该树的度为 (1), 深度为 (2), 终端结点 (叶子) 的个数为 (3), C 结点的双亲结点为 (4), 其子孙有 (5) 个结点。
- 4 由带权为 3, 9, 6, 2, 5 的 5 个叶子结点构成一棵哈夫曼(Huffman)树, 则其带权路径长度为 (1)。
- 5 在堆排序、快速排序和归并排序中, 若只从存储空间考虑, 则应首先选取 (1) 方法, 其次选取 (2) 方法, 最后选取 (3) 方法; 若只从排序结果的稳定性考虑, 则应选取 (4) 方法; 若只从平均情况下排序最快考虑, 则应选取 (5) 方法; 若只从最坏情况下排序最快并且要节省内存考虑, 则应选取 (6) 方法。

三、选择题 (本题 25 分, 每小题 5 分)

- 1 在一个单链表中, 已知 *q 结点是 *p 结点的前趋结点, 若在 *q 和 *p 之间插入 *s 结点, 则须执行操作 (A)_. 在双向循环链表中, 在 p 所指的结点之后插入 s 指针所指的结点, 其操作是 (B)_.

A: (1) $s \rightarrow \text{next} = p \rightarrow \text{next}; p \rightarrow \text{next} = s;$ (2) $q \rightarrow \text{next} = s; s \rightarrow \text{next} = p;$
(3) $p \rightarrow \text{next} = s \rightarrow \text{next}; s \rightarrow \text{next} = p;$ (4) $p \rightarrow \text{next} = s; s \rightarrow \text{next} = q;$

B: (1) $p \rightarrow \text{next} = s; s \rightarrow \text{prior} = p; p \rightarrow \text{next} \rightarrow \text{prior} = s; s \rightarrow \text{next} = p \rightarrow \text{next};$
(2) $s \rightarrow \text{prior} = p; s \rightarrow \text{next} = p \rightarrow \text{next}; p \rightarrow \text{next} = s; p \rightarrow \text{next} \rightarrow \text{prior} = s;$
(3) $p \rightarrow \text{next} = s; p \rightarrow \text{next} \rightarrow \text{prior} = s; s \rightarrow \text{prior} = p; s \rightarrow \text{next} = p \rightarrow \text{next};$
(4) $s \rightarrow \text{prior} = p; s \rightarrow \text{next} = p \rightarrow \text{next}; p \rightarrow \text{next} \rightarrow \text{prior} = s; p \rightarrow \text{next} = s;$

- 2 有一个二维数组 $A[1..6, 0..7]$, 每个数组元素用相邻的 6 个字节存储, 存储器按字节编址, 那末, 这个数组的体积是 (A) 个字节。假设存储数组的首地址是 0, 则若按行存储, 数组元素 $A[2, 4]$ 的第一个字节的地址是 (B); 若按列存储, 数组元素 $A[5, 7]$ 的第一个字节的地址是 (C)。

A~C: (1) 12 (2) 46 (3) 72 (4) 96 (5) 156 (6) 276 (7) 282 (8) 288

- 3 树是结点的有限集合, 它 (A) 根结点, 记为 T, 其余的结点分成 m ($m > 0$) 个 (B) 的集合 T_1, T_2, \dots, T_m , 每个集合又都是树, 此时结点 T 称为 T_i 的父结点, T_i 称为 T 的子结点 ($1 \leq i \leq m$)。一个结点的子结点数称为该结点 (C)。二叉树与树是两个不同的概念, 二叉树也是结点的有限集合, 它 (D) 根结点。可以把树的根结点的层数定义为 1, 其他结点的层数等于其父结点所在层数加上 1。令 T 是一棵二叉数, K_i 和 K_j 是 T 中子结点数小于 2 的结点中的任意两个, 它们所在的层数分别为 λK_i 和 λK_j , 当关系式 $|\lambda K_i - \lambda K_j| \leq 1$ 一定成立时, 则称 T 为一棵 (E)。

A、D: (1) 有 0 个或 1 个 (2) 有 0 个或多个
(3) 有且只有 1 个 (4) 有 1 个或 1 个以上

B: (1) 互不相交 (2) 允许相交 (3) 允许叶结点相交 (4) 允许树枝结点相交

C: (1) 权 (2) 维数 (3) 次数 (4) 序

E: (1) 丰满树 (2) 查找树 (3) 平衡树 (4) 完全树

- 4 对于一个具有 n 个顶点的图, 若采用邻接矩阵表示, 则该矩阵的大小 (A); 对于一个具有 n 个顶点和 e 条边的无向图, 若采用邻接表表示, 则表向量的大小为 (B), 所有顶点邻接表的结点总数为 (C), 采用邻接表存储的图的深度优先遍历算法和广度优先遍历算法分别类似于二叉树的 (D) 和 (E)。

A: (1) n (2) n^2 (3) $(n-1)^2$ (4) $(n+1)^2$

B、C: (1) n (2) $n+1$ (3) $n-1$ (4) $n+e$ (5) e (6) $e/2$ (7) $2e$

D、E: (1) 先序遍历 (2) 中序遍历 (3) 后序遍历 (4) 按层遍历

- 5 按所载的信息的形式, 文件可分为记录式文件和流式文件。流式文件在逻辑上是字符的集合。为了提高记录式文件的存取效率, 往往采用索引技术。索引的本质是按某种“标准”, 将记录进行分类或排序, 通常这个“标准”即是记录的 (A)。如果索引文件很大, 还可以对此索引文件再次索引, 直至建立起多级索引, 多级索引机制都是以 (B) 为基础。建立“次索引”是与索引十分类似的另一种基本检索方法, 也称为 (C), 其中的内容是属性值的全部记录的地址。

A: (1) 物理块地址 (2) 关键字值 (3) 属性值集合 (4) 指针

B: (1) 树 (2) 链表 (3) 队列 (4) 有向图

C: (1) 散列表 (2) 查找顺序表 (3) 属性地址表 (4) 倒排表

四、(15 分) 已知模式串 $S = \text{"abcaabbabacabaacbacba"}$, 模式为 $p = \text{"abcabaa"}$, 试画出按 KMP 算法 (模式匹配的一种改进算法) 进行模式匹配时每一趟的匹配过程。

五、(本题 15 分, 第 1 小题 6 分, 第 2 小题 9 分) 假设一个森林的两种遍历如下:

先序遍历 ACBGDEHFJI

后序遍历 CGBAHEJFI

1 画出该森林;

2 画出与该森林对应的二叉树以及它的中序线索树的存储方式。

六、(本题 15 分, 第 1 小题 6 分, 第 2 小题 9 分) 已知有向图是一个网络图, 它的邻接矩阵表示如下:

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|----------|----------|----------|----------|----------|----------|
| 1 | ∞ | 2 | 3 | ∞ | ∞ | ∞ |
| 2 | ∞ | ∞ | 2 | 4 | 2 | ∞ |
| 3 | ∞ | ∞ | ∞ | 1 | 3 | ∞ |
| 4 | ∞ | ∞ | ∞ | ∞ | ∞ | 4 |
| 5 | ∞ | ∞ | ∞ | 1 | ∞ | 5 |
| 6 | ∞ | ∞ | ∞ | ∞ | ∞ | ∞ |

1 画出该有向图以及它的邻接表存储方式;

2 若上述网络图是某一工程作业的 AOE 网, 试计算每个结点的最早开工时间和最迟开工时间, 并求出它的关键路径和完成该工程所需的最少时间。

七、(20 分) 在内部排序的过程中, 通常需要对待排序的关键码进行多遍扫描, 采用不同的排序方法, 会产生不同的中间结果。

设要将序列 (Q、H、C、Y、P、A、M、S、R、D、F、X) 中的关键码按字母排序的升序重新排序。请分别使用冒泡排序、初始步长为 4 的 Shell 排序、归并排序和以第一个元素为分界元素的快速排序进行第一趟扫描的过程和结果, 以及使用堆排序进行初始建堆的过程和结果。

八、(15 分) 设给定的散列表存储空间为 $H[1..m]$, 每个 $H(i)$ 单元可存放一个记录, $H[i]$ ($1 \leq i \leq m$) 的初始值为 Null, 选取的散列函数为 $H(R.key)$, 其中 $R.key$ 为 R 记录的关键字, 解决冲突方法为“线性探测法”, 试编写一个函数, 将某记录 R 填入到散列表 H 中。(为了简单, 只考虑记录仅包含一个 Key 域的情况)