



2008 年哈工大计算机科学与技术专业 854 考研真题

I. 数据结构（含高级语言）部分（75分）

一、填空（每题2分，共20分）

1. 已知一个线性表有 n 个元素，其中每个元素的数据占 8 个字节，假设一个指针的大小为 4 个字节，如果采用有 30 个元素的数组存储，那么当数组中有效元素个数满足 (1) 条件时，数组的存储效率比不带头结点的单链表更高。
2. 给定 14 个字母，假设它们的权值都相等，采用 huffman 编码，则每个字母的平均代码长度是 (2)。
3. 按 C 语言的运算符优先级，中缀表达式“ $A \& \& B \parallel (E > F)$ ”的等价后缀形式为 (3)。
4. 设按顺时针方向移动的循环队列 $Q[N]$ 的头尾指针分别为 F 、 R ，头指针 F 总是指在队列中的第一个元素的前一位置，尾指针 R 在最后一个元素的位置，则队列中的元素个数为 (4)。
5. 从空二叉树开始，严格按照 BST（二叉查找树）的插入算法，逐个插入关键字 {18, 73, 10, 5, 68, 99, 27, 41, 32, 25} 构造出一颗 BST，对该 BST 按照先根遍历得到的序列为 (5)。
6. 将两个长度为 m 的有序序列归并为一个有序序列，最少需要做 (6) 次关键字比较，最多需要做 (7) 次关键字比较。
7. 散列查找中，(8) 现象称为冲突，(9) 现象称为聚集。
8. 设可用的内存单元可处理 4 个记录，采用 4 路归并的选择树法生成由小到大的初始归并段，对有 12 个记录在案的文件，产生的第一个初的归并段长度为 (10) 个。
9. 在两种求图的最小生成树的算法中，(11) 算法适合于边稀疏的图的最小生成树。
10. 已知一个序列为 {21, 39, 35, 12, 17, 43}，则利用堆排序方法建立的初始堆为：(12)。

二、判断（每题1分，共9分）

1. 倒排文件只能按关键字的顺序存储。(①)
2. 堆的存储表示可能是链接式的，也可以是顺序的。(②)
3. 在 AOE 网中，任何一个关键活动的延迟，都会使整个工程延迟。(③)
4. 有环路的有向图不能进行拓扑排序。(④)
5. 对无向图进行一次深度优先搜索可以访问到图中的所有顶点。(⑤)
6. 大根堆的最大元素应该在堆顶，即根结点。(⑥)
7. 归并排序的平均时间复杂度为 $O(n \log n)$ ，最坏为 $O(n^2)$ 。(⑦)
8. 栈总是在栈底删除元素。(⑧)
9. 分块查找只适合静态查找，不适合动态查找。(⑨)

三、问答题（每题8分，共16分）

1. 许多文献中认为常用的排序算法是快速排序算法，而不是归并排序，你是如何理解的？



2. 在包含 n 个关键字的线性表中进行顺序查找,若查找第 i 个关键字的概率为 P_i 且分布如下:

$$P_1=1/2, P_2=1/4, \dots, P_{n-1}=1/2^{(n-1)}, P_n=1/2^n;$$

求: (1) 查找成功的平均查找长度。(2) 查找失败情况下的平均查找长度。

四、算法设计题(每题15分, 共30分)

1. 设二叉树结点表示的数据元素类型为Elementtype,二叉树用左右链表示。一棵二叉树的最大枝长和最小枝长分别如下定义:

最大枝长就是二叉树的层数; 最小枝长就是离根结点距离最近的叶结点到根路径上的

边数。

请设计一个算法, 同时求出一棵二叉树的最大和最小枝长。

2. 设计一查找无环路有向图第 i 对顶点间“最长简单路径”(所谓最长简单路径是指该简单路

径包含边最多)的算法, 即以一个无环路有向图作为输入, 对于每个顶点如果它们之间存

在简单路径, 则输出其中最长的, 否则输出为空。

备注: 09-12 年 854 数据结构真题请参考 408



2013 年哈工大计算机科学与技术专业 854 考研真题

一、单项选择题：(1-10 小题，每小题2 分，共20 分。在每小题给出的四个选项中，请选出一项最符合题目要求的。)

1. 深(高)度为6(根的层次为1)的完全二叉树至少有()结点。
A. 64 B. 32 C. 31 D. 63
2. 若具有 n 个结点， k 条边的非连通无向图是森林($n > k$)，则该森林中必有()棵树。
A. k B. n C. $n-k$ D. $n+k$
3. 若无向图 G 有 n 各顶点，其邻接矩阵为 $A[1 \cdots n, 1 \cdots n]$ ，且压缩存储在 $B[1 \cdots k]$ ，则 k 的值至少为()。
A. $n(n+1)/2$ B. $n^2/2$ C. $(n-1)(n+1)/2$ D. $n(n-1)/2$
4. 下列排序算法中，()算法可能会出现下面情况：在最后一趟(遍)开始之前，所有元素都不在其最终的位置上。
A. 堆排序 B. 冒泡排序 C. 选择排序 D. 插入排序
5. 文件有 m 各初始归并段，采用 k 路归并时，所需要的归并趟(遍)数是()。
A. $\lceil \log_2 k \rceil$ B. $\lceil \log_2 m \rceil$ C. $\lceil \log_k m \rceil$ D. $\lceil \log_m k \rceil$
6. 下述编码中那一组不是前缀码()。
A. (00, 01, 10, 11) B. (0, 1, 00, 11)
C. (0, 10, 110, 111) D. (1, 01, 000, 001)
7. 设高度为 h 的二叉树只有度为0 和度为2 的结点，则此类二叉树中所包含的结点数为()。
A. $2h$ B. $2h-1$ C. $2h+1$ D. $h+1$
8. 在某棵树中，结点 M 和 N 是结点 P 的第 i 和 $i+1$ 个孩子，则在这棵树的二叉树表示中，结点 M 与 N 的关系是()。
A. M 、 N 具有同一双亲 B. M 是 N 的左孩子
C. N 是 M 的左孩子 D. N 是 M 的右孩子
9. 具有 n 各结点的二分(折半)查找判定树，查找失败的外部结点(失败结点)共有()个。
A. $n-1$ B. n C. $n+1$ D. $\log_2 n$
10. 在含有 n 各关键字的最小堆(堆顶元素最小)中，关键字最大的记录有可能存储在()。
A. $\lfloor n/2 \rfloor$ B. $\lfloor n/2 \rfloor - 1$ C. 1 D. $\lfloor n/2 \rfloor + 2$

二、填空题：(11-15 小题，每空1 分，共10 分。)

11. 对于一个长度为 n 的顺序存储的线性表，在第一个元素前插入元素的时间复杂度为____；如果插入任何位置的几率相同，那么插入操作的平均时间复杂度为____。



12. 稀疏矩阵常用的两种存储方法是____和____。
13. 后缀算术表达式 $9\ 2\ 3\ +\ -\ 8\ 2\ /\ -$ 的值为____。中缀算术表达式 $(3+4*X)-2*Y/3$ 对应的后缀算术表达式为____。
14. 具有 $2n$ 个结点的完全二叉树，含有____个度为1 的结点，____个度为2 的结点。
15. 在一棵高度为 h 的B 树中，叶子结点处于第____层，当向该B 树中插入一个新关键字时，为查找插入位置需读取____个结点。

三、解答题：（16-17 小题，每小题10 分，共20 分。）

16. 一颗二叉树T 的前序和中序遍历序列分别为：C, A, B, D, E, F 和C, B, D, A, F, E。请回答下列问题：

- ①画出二叉树T；
- ②简要概括由任意二叉树的前序和中序遍历序列构造二叉树的方法；
- ③画出二叉树T 的后续线索二叉树。

17. 已知加权有向图的邻接矩阵如下图所示，如需在其中一个顶点建立娱乐中心，请回答下列问题：

- ①利用Floyd 算法求出每对顶点的最短距离，并用矩阵表示；
- ②求每个顶点的偏心度；
- ③确定娱乐中心应选哪个顶点。

	a	b	c	d
a	0	1	∞	∞
b	∞	0	2	∞
c	∞	∞	0	2
d	∞	4	3	0

第 17 题图

四、算法设计题(18-19 小题，共 25 分)

按以下要求设计算法：

- (1)描述算法设计的基本思想；
- (2)根据设计思想，采用 C 或 C++或 Java 语言描述算法；
- (3)分析算法时间复杂度和空间复杂度。

18. (12 分) 已知一个按升序排好的数组和一个数字，请设计一个尽可能高效的算法 Findsum, 在数据组中查找两个数，使得它们的和正好等于已知的那个数字，例如数组 1, 2, 4, 6, 7, 11 和数字 11。由于 $4+7=11$ ， 因此输出 4 和 7。如果存在多对这样的数字，输出任意一对 即可。

19. (13 分) 在森林的孩子-兄弟表示法(二叉链表表示法)存储结构中，每个结点的结构为：

firstchild	data	rightsib
------------	------	----------

的结构为：

请设计一个算法 LeavesCounter，求森林的叶子结点数。



2014 年哈工大计算机科学与技术专业 854 考研真题

I. 数据结构部分(共 75 分)

一、单项选择题: (1-10小题, 每小题2分, 共20分。在每小题给出的四个选项中, 请选出一项最符合题目要求的)

1. 有一个 100×90 整型数的稀疏矩阵, 非0元素有10个, 设每个整型数占2字节, 则用三元组表示该矩阵时, 所需的字节数为(1)。
A. 60 B. 66 C. 180 D. 33
2. 下列内部排序算法中, 其比较次数与序列初始状态无关的是(2)。
A. 快速排序 B. 直接插入排序 C. 二路归并 D. 选择排序
3. 若度数为 m 的哈夫曼树中, 其叶子结点的个数为 n , 则非叶子结点的个数为(3)。
A. $n-1$ B. $n/(m-1)$ C. $(n-1)/(m-1)$ D. $(n+1)/(m+1)-1$
4. 长度为12有序表, 按折半查找法对该表进行查找, 以等概率查找表内各元素, 则查找成功时所需要的平均比较次数为(4)。
A. $35/12$ B. $37/12$ C. $39/12$ D. $43/12$
5. 设有 K 个关键字互为同义词, 若用线性探测法把这 K 个关键字存入散列表中, 至少要进行(5)次探测。
A. $K-1$ B. K C. $K+1$ D. $K(K+1)/2$
6. 有 n 个初始归并段, 采用 K 路归并时, 所需要的归并遍数是(6)。
A. $\log_n k$ B. $\log_2 k$ C. $\log_2 n$ D. $\log_k n$
7. 有 n 个顶点, e 条边的有向图采用邻接存储, 若删除与顶点 v_i 相关的所有边, 其时间复杂度为(7)。
A. $O(n)$ B. $O(e)$ C. $O(\max(n,e))$ D. $O(n \cdot e)$
8. 在平衡二叉树中插入一个结点造成不平衡。设最低的不平衡结点为A, 并已知插入后A的左子树根的平衡度为0, 右子树根的平衡度为1, 则应作(8)型的调整达到平衡。
A. LL B. LR C. RL D. RR
9. 一棵具有 n 个非叶子结点完全二叉树的线索树, 含有多少条线索(9)。
A. $2n+1$ 或 $2n$ B. $2n+2$ 或 $2n+1$ C. $2n+1$ 或 $2n-1$ D. $2n+2$ 或 $2n-2$
10. 在某森林的二叉树表示中, 结点M和结点N是同一父节点的左儿子和右儿

子, 则在该森林中(10)。

A.M、N 具有同一双亲

B.M、N 可能没有共同祖先

C.M是N的儿子

D.M是N的左兄弟

二、填空题:(11-18 小题, 每空 1 分, 共 10 分)

11. 高度为 h 的完全二叉树至少有 11-1 个结点。

12. 对 N 个结点的 k 叉树($k \geq 2$)的 k 叉链表中有 12-1 空指针。

13. 对具有 n 个元素的顺序存储的有序表和顺序存储的无序表进行顺序查找, 在等概率的情况下, 查找不成功时的平均查找长度分别为 13-1、13-2。

14. 在 m 阶B-树中, 当有关键字插入导致相关结点分裂时, 原结点上有 14-1 个关键字。

15. 以比较为基础的排序的时间复杂度的下界是 15-1。

16. 完全二叉树的顺序存储序列为ABCDEFGH, 其后序遍历的序列为 16-1。

17. 在AOE网络中, 关键活动是指 17-1, 缩短 17-2 活动的持续时间, 可以提前完成工程。

18. 求最短路径的 Dijkstra 算法和求最小生成树的 Prim 算法之间的主要区别 18-1。

三、简答题:(19-20 小题, 共 20 分)

19. (12分) 从大规模数据(例如1亿个数)表中取前100个最大值, 给出一种高效算法并描述算法思想, 阐述选择该算法的理由。

20. (8分) 给出判断一个有向图是否存在拓扑排序的算法:给出图-1 所示有向图的拓扑序列。

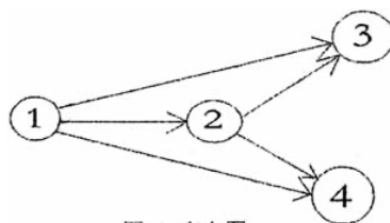


图-1 有向图

四、算法设计题:(21-22 小题, 共 25 分)

按以下要求设计算法:

(1)描述算法设计的基本思想;



(2)根据设计思想,采用 C 或 C++或 Java 语言描述算法;

(3)分析算法时间复杂度和空间复杂度

21. (13分) 二叉树采用左右链存储,完成下列算法,要求算法尽可能高效,分析算法时间和空间复杂度:

- (1) 判断二叉树是否为完全二叉树;
- (2) 输出二叉树从右向左数第K个叶结点。

22. (12分) 设计一种数据结构,满足栈的性质,实现下列 3 个操作:

- (1) Push(v): 将v加入到栈;
- (2) Pop(): 删除栈顶元素并返回此元素
- (3) Maxelement(): 返回栈中最大元素;

让它们的时间复杂度都为 $O(1)$ 。





2015 年哈工大计算机科学与技术专业 854 考研真题

一、单项选择题

1. 一颗树的结点数等于所有结点的入度数之和 ()
A. -1 B.+1 C. $\times 2$ D. $\div 2$
2. 下列冒泡排序算法中, 空间复杂度最大的是 ()
A. 冒泡排序 B.快速排序 C. 归并排序 D.堆排序
3. 设 $S[n]$ 为一数组, 现利用 $S[n]$ 设计 2 个栈 s_1 和 s_2 , 为了合理利用数组空间, 最佳方案为 ()
A. $S[0]$ 为 s_1 的栈底, $S[n]$ 为 s_2 的栈底
B. $S[0]$ 为 s_1 的栈底, $S[n-1]$ 为 s_2 的栈底
C. $S[0]$ 为 s_1 的栈顶, $S[n-1]$ 为 s_2 的栈底
D. $S[0]$ 为 s_1 的栈顶, $S[n]$ 为 s_2 的栈底
4. 用直接插入排序方法对下面四个序列进行排序 (由小到大), 元素比较次数最少的是 ()
A. 94, 32, 40, 90, 80, 46, 21, 69
B. 32, 40, 21, 46, 69, 94, 90, 80
C. 21, 32, 46, 40, 80, 69, 90, 94
D. 90, 69, 80, 46, 21, 32, 94, 40
5. 无向图的邻接表是 () 矩阵
A. 上三角 B.下三角 C.对称 D.反对称
6. 在二叉排序树的存储结构中, 关键字值最小的结点 ()
A. 左指针一定为空 B.右指针一定为空
C.左右指针均为空 D.左右指针均不为空
7. 散列表的地址区间为 0-17, 散列函数为 $H(K) = K \bmod 17$ 。采用线性探测法处理冲突, 并将关键字序列 62, 48, 72, 38, 8, 18, 59 依次存储到散列表中。元素 59 存放在散列表中的地址是 ()
A. 8 B.9 C.10 D.11
8. 无向图 $G=\{V,E\}$, 其中
 $V=\{a,b,c,d,e,f\}$, $E=\{(a,b),(a,e),(a,c),(b,e),(c,f),(f,d),(e,d)\}$, 对该图进行广度优先遍历, 得到的顶点序列正确的是 ()
A. a,b,e,c,d,f B. a,c,f,e,b,d C. a,e,b,c,f,d D. a,e,d,f,c,b
9. 设有向无环图 $G=(V,E)$, 顶点集 $V=\{V_0, V_1, V_2, V_3, V_4\}$, 边集 $E=\{<V_0, V_1>, <V_0, V_2>, <V_0, V_3>, <V_1, V_3>, <V_2, V_4>, <V_3, V_4>\}$ 。若从顶点 V_0 开始对图进行拓扑排序, 则可能得到的不同拓扑序列排序个数是 ()
A.1 B.2 C.3 D.4
10. 对于一颗满二叉树, 共有 n 个结点和 m 个叶节点, 高度为 h (设根为第 1 层), 则 ()
A. $n=h+m$ B. $h+m=2h$ C. $h=m-1$ D. $m=2^h-1$



二、填空题

11. 若循环单链表表示的队列长度为 n ，头指针指向最后一个节点，则其入队的时间复杂度为_____，出队的时间复杂度为_____。
12. Prim 算法的时间复杂度为_____，它适合求解_____图。
13. 高度为 4 的 3 阶 B-树，对于内结点，关键字最多为_____，最少为_____。
14. 有一个 169 项的表，欲采用等分区间顺序查找法，则每块的理想长度是_____，平均查找长度是_____。
15. 在堆排序过程中，调整堆的时间复杂度为_____，堆排序算法整个过程的时间复杂度为_____。

三、简答题

16. 简答 Floyd 算法的基本设计思想，试举例说明，为什么 Floyd 算法可以有负权值的边，但不能有包含负权值边组成的回路。
17. 设有一组初始记录关键字为 (45、80、48、40、22、78)，写出每一步构造平衡二叉树的步骤，构造完之后写出其先序遍历序列。

四、算法设计题

按以下要求设计算法：

- (1) 描述算法设计的基本思想；
 - (2) 根据设计思想，采用 C 或 C++或 Java 语言描述算法；
 - (3) 分析算法时间复杂度和空间复杂度。
18. 设 H1、H2 为两个链表的头指针，编写算法 Judge 判断两个单链表是否有交叉，要求效率尽量高，并分析时间空间复杂度。
 19. 二叉树以二叉链表为存储结构，分别写出在二叉树中查找值为 X 的节点算法 Search 及求 X 所在的结点在树中层数的算法 Level1。



2016 年哈工大计算机科学与技术专业 854 考研真题

I. 数据结构部分

一、单项选择题

1. 具有 2015 个结点的完全二叉树，叶子节点个数为 (1)。
A.1006 B.1007 C.1008 D.1009
2. 算术表达式 $a*(b+c)-d$ 化为后缀表达式为 (2)。
A.abcd+*- B.abc+*-d C.abc+*d- D.ab+c*d-
3. 某队列允许在某两端进入入队操作，但仅允许在一端进行出队操作，元素 a、b、c、d、e 依次入队，则不可能得到的顺序是 (3)。
A.bacde B.dbace C.dbcae D.ecbad
4. 在有 N 个叶子结点的哈夫曼树中，其结点总数为 (4)。
A.不确定 B. $2N-1$ C. $2N+1$ D. $2N$
5. 一个只有度为 0 和 2 的二叉树，有 5 个叶子结点，该二叉树最小层数是 (5)，最大层数是 (5)。
A.3, 4 B.4, 5 C.5, 6 D.6, 7
6. 若邻接表中有奇数个边表结点，则一定是 (6)。
A.图中有奇数个结点
B.图中有偶数个结点
C.图为无向图
D.图为有向图
7. 下面给出的四种排序中 (7) 排序是不稳定性排序。
A.插入 B.冒泡 C.二路归并 D.堆
8. 当初始序列已按关键字有序时，用直接插入算法进行排序，需要比较次数为 (8)。
A. $n-1$ B. $\log_2 n$ C. $2\log_2 n$ D. n^2
9. 设散列表中有 m 个存储单元，散列函数 $H(key)=key \% p$ ，则 p 最好选择 (9)。
A.小于等于 m 的最大奇数
B.小于等于 m 的最大素数
C.小于等于 m 的最大偶数
D.小于等于 m 的最大合数
10. 哈希查找中 k 个关键字具有同一哈希值，若用线性探测法将这 k 个关键字对应的记录存入哈希表中，至少要进行 (10) 次探测。
A.k B.k+1 C. $k(k+1)/2$ D. $1+k(k+1)/2$

二、填空题

11. 在单链表中设置头结点的作用是_____，对于一个具有 n 个结点的单链表，在已知的结点 *p 后插入一个新结点的时间复杂度为_____，在给定值为 x 的结点后插入一个新结点的时间复杂度为_____。
12. 完整的存储一个 $n*n$ 对称矩阵所需的存储单元数为_____，采用上三角形式的压缩方式，存储一个对称矩阵所需的存储单元数为_____。
13. 设一棵二叉树中的中序遍历序列为 DBCA，后序遍历序列为 DBAC，则这棵



二叉树的前序遍历序列为_____。

14. 设图有 37 条边，其最小的顶点数为_____，最多的顶点数为_____。

15. 解决 Hash 冲突的两种常用方法为_____、_____。

三、简答题

16. 设有一优先级队列插入时可以任意优先级的元素，删除时只能删除最大优先级元素，写出用以下存储结构实现时的时间复杂度。

	插入时间复杂度	删除时间复杂度
无序数组		
有序数组		
无序单链表		
有序双链表		
最大堆		

17. 设有一组关键字 $T=(12,2,16,30,8,28,4,10,20,6,18)$ ，设内存工作区可存放 4 条记录，用置换选择排序的方法生成初试归并段：

(1) 一共可以生成多少个初始归并段？

(2) 对其进行归并排序，写出每一步结果，并给出最终输出结果。

四、算法设计题

按下列要求设计算法：

(1) 描述算法设计的基本思想；

(2) 根据设计思想，采用 C 或 C++ 或 Java 语言描述算法；

(3) 分析算法时间复杂度和空间复杂度。

18. 有一长度为 n 的整数序列 $a[n]$ ，其元素按先严格递减后平桥递增的顺序排列，请设计一个时间上尽可能高效的算法 FindMin() 查找最小值元素，并分析时间复杂度和空间复杂度。

19. 二叉树采用二叉链表表示法，并增加一个你父结点指针 parent 域，以及 flag 域（可以取 0, 1, 2 三个值，取 0 时表示访问根结点，取 1 表示访问左孩子，取 2 时表示访问右孩子），在该存储结构上，设计一个算法实现非递归不用栈的后序遍历算法。



2017 年哈工大计算机科学与技术专业 854 考研真题

一、选择（1-10 小题）

1. 设 n 是描述问题规模的非负整数，下面程序片段的时间复杂度是（）
 $x = n * n;$
 $\text{while}(x \geq 1)$
 $\quad x = x / 2;$
A. $O(\log_2 n)$ B. $O(n)$ C. $O(n \log_2 n)$ D. $O(n^2)$
2. 需要分配一个较大的存储空间并且插入和删除操作不需要移动元素，满足以上特点的线性表存储结构是（）
A. 单向链表 B. 静态链表 C. 线性链表 D. 顺序表
3. 已知字符串 S 为“ababcabcacbab”，模式串 t 为“abcac”。若采用 KMP 算法进行模式匹配，则需要（）遍（趟）匹配，就能确定 t 是 S 的子串
A. 3 B. 4 C. 5 D. 6
4. 已知某棵二叉树的前序序列是 1, 2, 3, 4，则不可能为该二叉树的中序序列的是（）
A. 1, 2, 3, 4 B. 2, 3, 4, 1 C. 1, 4, 3, 2 D. 3, 1, 4, 2
5. 将森林 F 转换为对应的二叉树 T , F 中任何一个没有右兄弟的结点，在 T 中（）
A. 没有左子树 B. 没有右子树 C. 左右子树都没有 D. 以上都不对
6. 一个含有 n 个顶点和 e 条边的无向图，在其邻接矩阵存储结构中共有（）个零元素
A. e B. $2e$ C. $n^2 - 2e$ D. $n^2 - e$
7. 在一棵高度为 2 的 7 阶 B 树中，所含关键字的个数最少是（）
A. 5 B. 7 C. 8 D. 14
8. 设待排序的元素个数为 n ，则基于比较的排序算法在最坏情况下的时间复杂度的下界为（）
A. $\log_2 n$ B. n C. $n \log_2 n$ D. n^2
9. 下面关于 B 树和 B+树的叙述中，不正确的是（）
A. B 树和 B+树都能有效地支持随即检索 B. B 树和 B+树都能有效地支持顺序检索
C. B 树和 B+树都是平衡的多路树 D. B 树和 B+树都可以用于文件的索引结构
10. 若待排序关键字序列在排序前已按其关键字递增顺序排序，则采用（）方法比较次数最少
A. 插入排序 B. 快速排序 C. 堆排序 D. 选择排序

二、填空题（11-19 小题，每空 1 分，共 10 分）

11. 在一棵 n 个结点的二叉树中，所有结点的空子树个数为_____
12. 若二叉树的一个叶子结点是其某子树的中序遍历序列中的第一个结点，则它必是该子树的后序遍历序列中的第_____个结点。



13. 在有 n 个选手参加的单循环赛中，总共将进行____场比赛。
14. 在有 4033 个叶结点的完全二叉树中，叶子结点的个数为____个。
15. 一个有向图 G_1 的反向图是将 G_1 的所有向边的方向取反而得到的有向图 G_2 。若 G_1 和 G_2 的邻接矩阵分别为 A 、 B ，则 A 与 B 的关系为____。
16. n 个顶点 e 条边的无环路有向图，若采用邻接表作为存储结构，则拓扑排序算法的时间复杂度为____。
17. 在 10 阶 B 树根结点所包含的关键字最多为____个，最少为____个。
18. 在具有 12 个结点的二叉平衡树 (AVL 树) 中，查找 AVL 树中的一个关键字最多需要____次比较。
19. 对初态有序的表，最省时间的排序算法是____。

三、简答题 (20-21 小题，共 20 分)

20. 在 n 个数据中找出前 K 个最大元素，可以采用堆排序或者败者树来实现，分别说明上述两种实现方法的基本步骤，并分析每种方法的时间复杂度和空间占用情况。
21. 假设举办一个 1000 人参加的学术会议，作为会议报道组的负责人，你收到会务组为每名参会者开具的包含其英文名字的注册费发票，同时还会收到为每位参会者提供的印有其英文名字的参会胸牌和其他会议资料，请回答以下问题。
 - 1) 如何有效地把每个参会者的注册费发票和参会的胸牌等其他会议资料放在一起形成一份参会资料？
 - 2) 如何在会议报道日更有效第把每份参会资料发给参会者？要求：说明你所用到的主要技术和关键步骤。

四、算法设计题 (22-23 小题，共 25 分)

按以下要求设计算法：

- (1) 描述算法设计的基本思想；
 - (2) 根据设计思想，采用 C 或 C++ 或 Java 语言描述算法；
 - (3) 分析算法时间复杂度和空间复杂度。
22. 给定一个 n 个整数的无序数组 A ，设计一个时间和空间尽可能高效的算法：
`int findtheKMin(int A[], int n, int k)`, 找出其中第 k 小的整数。
 23. 给定一颗 n 个结点的二叉排序树 (即 BST)，每个结点均存放一个整数，其结点格式为 `[lchild][data][rchild]`。令 $half = (BST \text{ 中的最大值} + BST \text{ 中的最小值}) / 2$ 。设计一个算法 `int findNearMid (BinTree *root)`，完成 1) 找出 BST 中最大值和最小值以计算 $half$ 的值；2) 返回大于 $half$ 且与 $half$ 相差最小的结点值。



2018 年哈工大计算机科学与技术专业 854 考研真题

I. 数据结构部分(共 75 分)

一、单项选择题: (1-10 小题, 每小题 2 分, 共 20 分。在每小题给出的四个选项中, 请选出一项最符合题目要求的)

1. 一个算法执行时间是 $2n^3 + 3n^2 \log_2 n + 4n$, 其时间复杂度为 (1)。
A. $O(n^3)$ B. $O(\log_2 n)$ C. $O(n \log_2 n)$ D. $O(n^2)$
2. 设一组初始记录关键字序列为(315, 205, 674, 924, 627, 893), 则用基数排序需进行 (2) 趟的分配和回收才是其能变成有序数列。
A. 5 B. 4 C. 3 D. 6
3. 某二叉树的前序遍历为 3124, 中序为 1234, 则其层序序列为 (3)。
A. 2143 B. 3124 C. 3412 D. 3241
4. 设无向图 G 中有 n 个顶点 m 条边, 则其对应的邻接表中, 顶点表结点和边表结点的个数分别为 (4)。
A. m, n B. n, m C. $n, 2m$ D. $2m, n$
5. 已排好序的 18 个元素放在一维数组 $A[19]$ 中, 第一个元素放 $A[1]$ 中, 现进行二分(折半)查找, 则查找 $A[3]$ 的比较序列下标依次为 (5)。
A. 123 B. 9523 C. 953 D. 9423
6. 设有 n 个关键字具有相同的散列值, 则用线性探测法把这 n 个关键字映射到初始为空的散列表中需做 (6) 次线性探测。
A. n^2 B. $n(n+1)$ C. $n(n+1)/2$ D. $n(n-1)/2$
7. 无向图的邻接矩阵对应的二维数组是 A , 现将其上三角(即主对角线及以上)元素按行优先顺序压缩存储在一个足够大的一位数组 Sa 中, 若 $Sa[0] = A[0][0]$, $Sa[17] = A[2][2]$, 则矩阵元素 $A[5][3]$ 的值存放在一位数组 Sa 的第 (7) 个单元中。
A. 25 B. 26 C. 27 D. 33
8. 对一棵 m 阶的 B 树, 不正确的是 (8)。
A. 树的单个节点至少有 m 棵子树
B. 根节点至少有 2 棵子树
C. 除根节点和失败节点外, 所有节点至少有 $\lceil m/2 \rceil$ 棵子树
D. 所有的终端节点(失败结点)都位于同一层
9. 设一组初始记录关键字序列为(50, 40, 95, 20, 15, 70, 60, 45), 则以增量 $d=4$ 的一趟

希尔(shell)排序, 结束前4条记录关键字为(9)。

- A. 40,50,20,95 B. 15,40,60,20 C. 15,20,40,45 D. 45,40,15,20

10. 以下四组数据分别存储在一维数组中, 堆结构的选项是(10)。

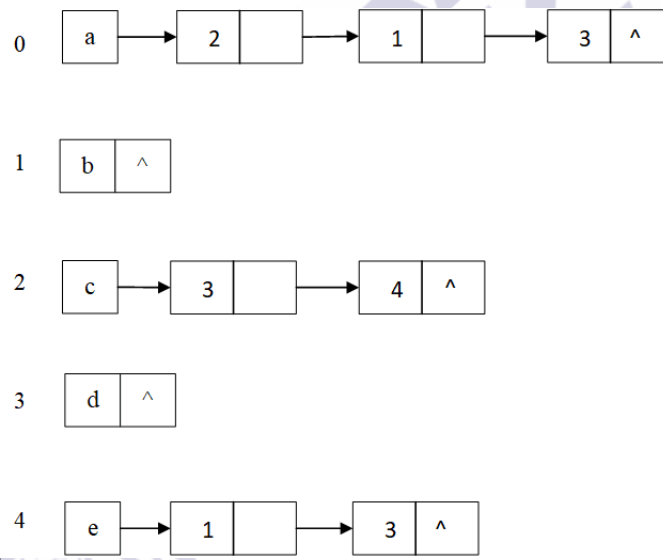
- A. 10,15,56,20,30 B. 10,20,56,15,30
C. 15,20,56,10,30 D. 56,20,10,15,30

二、填空题:(11-15 小题, 每空 1 分, 共 10 分)

11. 具有120个结点的完全二叉树, 其高度的最小值和最大值分别是____、____。

12. 弗洛伊德和迪杰斯特拉算法的适用条件分别为____、____。

13. 有向图邻接表存储结构如图所示:从顶点a的深度优先(DFS) 搜索和广度优先搜索(BFS), 输出序列分别是____和 ____。



14. 对于含有 n 个顶点 m 条边的无向带权连通图,利用克鲁斯卡尔算法求解最小生成树的时间复杂度是____,最小生成树包含____条边。

15. 对 n 个记录的文件进行快速排序, 平均情况所需要的辅助存储空间和时间复杂度分别为____和____。

三、简答题:(16-17 小题, 共 20 分)

16. (12分) 关键字的输入顺序为2,4,6,8, 10,12,14, 16, 18,分别建立初始为空的二叉搜索树(BST)和二插平衡树(AVL),要求:

(1) 画出所建立的BST,并计算其在等概率情况下搜索成功的平均查找长度ALS_b;

(2) 画出所建立的AVL,并计算在等概率的情况下搜索成功的平均查找长度ASL_a;



(3) 比较(1)和(2), 说明数据分布所建二叉树的结构形态关系, 以及搜索效率

率的影响。

17. (8分) 已知某文件预处理后, 得到5个初始归并段, 归并段中的每个数据占一个词盘读写单位, 初始归并段长度分别为(20,30, 10,5,30), 若采用二路平衡归并, 请设计一个读写磁盘次数最少的归并方案, 要求: 画出最佳归并树, 并计算磁盘的读写次数

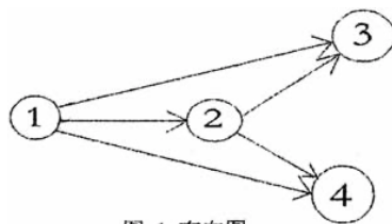


图-1 有向图

四、算法设计题:(18-19 小题, 共 25 分)

按以下要求设计算法:

- (1)描述算法设计的基本思想;
- (2)根据设计思想, 采用 C 或 C++或 Java 语言描述算法;
- (3)分析算法时间复杂度和空间复杂度

18. (13分) 在一个长度为 n 整数序列中, 奇数元素和偶数元素各占一半, 存放在数组 $A[n]$ 中。请设计一个时间和空间尽可能高效的算法NewSequence($\text{int } A[], \text{int } n$), 重新安排这些整数, 使奇数元素存放在奇数单元, 偶数元素存放在偶数单元, 并说明算法的时间和空间复杂度。

22. (12分) 已知二叉树BT采用左右键表示法(亦称二叉链表)作为其存储结构, 二叉树的结点结构如下: $[\text{lchild}][\text{data}][\text{rchild}]$ 。

请给出二叉树的类型Btree定义, 并设计一个非递归算法FirstNode (Btree BT), 直接返回BT的后序遍历的第一个被访问的结点, 说明你所设计算法的时间复杂度。

“直接”的含义是, 不能通过后序遍历得到二叉树的后序序列, 然后返回后序序列的第一个结点。



2019 年哈工大计算机科学与技术专业 854 考研真题（回忆版）

一、选择题（10 道，缺 2 道）

1. 向栈内输入 $a_1, 2, 3, 4, 5, 6$, 输出顺序应该类似于 $a_3, 4, 2, 1, 5, 6$ 这样的顺序, 问栈的容量
2. 高度为 5 的 4 阶 b 树, 所含的关键字的最小数目是多少
3. 以一个序列建立堆排序, 例如 $463954\cdots$, 问建立的堆初始排序是哪个 (这个比较狡猾没说是最大堆还是最小堆)
4. 在拓扑排序中, v 在 u 前面, 则选项中哪一个是不可能的, 有一个是有一条 u 到 v 的路径
5. prim 算法求最小生成树, $VT=\{1, 2, 3\}$, $ET=\{(1, 2) (2, 3)\}$, 接下来的最小代价边要从哪个序列里找
6. 平衡二叉树问题, A 是最低的不平衡点, A 的左子树平衡度为 1, 右子树平衡度为 0, A 需如何调整成平衡二叉树
7. n 个顶点组成无向图, 各顶点的度中最大是多少
8. 初始有序的哪种排序方法更好 (冒泡)

二、填空题（9 道题 10 个空，缺 1 道 1 空）

1. 算法的时间复杂度不仅跟问题的规模有关, 还跟实例的输入有关
2. 使用 2 路归并 n 个数, 需要 次
3. 给出后序序列, 中序序列, 求前序序列
4. kmp 算法求串 $abaabc$ 的 next 数组
5. 给了 5 个数, 求组成的哈夫曼树的 wpl
6. 哈希散列要解决 和 问题。
7. 散列表中散列函数 $hash=k \bmod 11$, 空间是 14, 给出了 4 个已知数的散列地址, 问若使用拉链法处理冲突, 则某一个数在表中的位置
8. 快速排序不适合什么情况

三、简答题（2 道）（都是图。。。）

1. 有 n 个人, m 对朋友关系, 朋友的朋友也是朋友, 这样组成 1 个朋友圈, 问怎么求有多少个朋友圈 (实质就是无向图求连通分量)
2. 一项工程, 有 11 项项目, 每个项目有完成所需时间及前置任务 (或者 10 项, 记得不是很准, 那个图类似于哈工大的教材 149 页的图, 其中试卷中 a_6 的前置任务 a_6 , 这个应该是印错了, 可能是 a_3)。求:
 - (1) 使用合适模型表示
 - (2) 最少需多久完成整个项目, 即求关键路径
 - (3) 给出如何求最少时间的方法

四、算法题（2 道）

1. 构造一种特殊的队列结构, 其中每个结点都有权值, 每一次从队列中取出的节点都是最高优先权, 设计算法输出



-
- (1) 最小权值
 - (2) 插入一个数
 - (3) 删除最小权值，要求时间复杂度在 $\log n$ 以内（大致思路模仿前几年栈限制在 $O(1)$ 思路的题，堆+栈，用空间换时间）
2. 一棵二叉树，求两个结点间的最小长度，即他们各自到最近共同祖先的路径长之和（这道题是国防科技大学软件工程真题的原题）。

