

山东工业大学

一九九六年攻读硕士学位研究生入学考试试题

报考专业: 计算机应用

考试科目: 数据结构

一、(20分) 填空 (每空0.5分)

1. (5分) 从人们把计算机作为数据处理机的观点看, 数据是对现实世界的事物采用计算机能够 输入、处理 和 输出 的形式所进行的描述。多媒体数据主要包括 文本、图形 和 声音 等基本形式, 它强调多媒体信息的 集成性。数据结构概念一般包括: 数据之间的 联系, 数据在计算机中的 组织 和数据的 操作 三个方面。

2. (5分) 对数据间关系的描述是数据的逻辑结构, 形式地可以用一个二元组 $B = (K, R)$ 来表示, 其中 K 是结点的 集合, R 是 关系 的有穷集合。当需要描述结点的组成时, 可以采用类似 PASCAL 语言中的 记录 的方法。设一片相邻的存储单元的整体记作 M 。现有逻辑结构 $B = (K, R)$, 要把 B 存储在计算机中, 首先必须建立一个从 K 的结点到 M 的单元的映象 $S: K \rightarrow M$, 即对于每一个 $k \in K$ 都有唯一的 $Z \in M$ 使得 $S(k) = Z$, Z 为 K 结点所占存储空间中的始单元。同时这个映象应具有明显地或隐含地体现 顺序 的能力。顺序的存储方法把 相邻 的结点存储在 相邻 的存储单元里。链接的存储方法是给结点附加上 指针 字段。散列的存储方法的主要思想是根据结点的 关键字 来确定它的 存储位置。

3. (5分) 数据的运算是定义在数据的 逻辑 结构上的, 但运算的具体实现要在 物理 结构上进行。一个算法是 有限 的有穷集合, 它应该具有下列特性 确定性, 可行性, 有穷性 和 有效性。数据结构的的评价基本上有两条标准: 一条标准是作为问题的参数的函数来计算 时间复杂度, 另一条标准是数据运算的 空间复杂度。

4. (5分) 从用户的观点看, 文件的逻辑结构通常可以区分为两类: 一类是如 dBASE 中数据库文件那样的文件组织结构, 称为 有序 文件; 另一种是诸如用各种文字处理软件编辑成的文本文件, 称为 无序 文件。从文件在存储器上的存放方式来看, 文件的物理结构往往可区分为三类, 即 顺序, 链接 和 索引。B+ 树适用于组织 索引 的索引结构, m 阶 B+ 树每个结点至多有 m 个儿子, 除根结点外每个结点至少有 $\lceil \frac{m}{2} \rceil$ 个儿子, 根结点至少有 2 个儿子, 有 k 个儿子的结点必有 k 个关键码。

二、(10分)

1. (6分) 设有关键码序列 10, 20, 35, 40, 44, 51, 65, 70, 85, 91, 93, 95。试按照最大关键码复写原则绘出相应的 2 阶 B+ 树。

2. (4分) 填空 (每空1分)

B- 树只适用 顺序 检索, 不适于 随机 检索, 而 B+ 树把所有的关

键码都存在_____结点上,这就为_____检索也提供了方便。

三、(20分)

1. (5分) 填空 (每空1分)

堆是一种有用的数据结构。堆排序是一种_____排序,堆实质上是一棵_____结点的层次序列。对含 n 个元素的序列进行排序时,堆排序的时间复杂度是 $O(n \log n)$,所需的附加存储结点是 $O(1)$ 。关键码序列 05, 23, 16, 68, 94, 72, 71, 73 是否满足堆的性质_____。

2. (7分) 设 n 为结点个数, datatype 为结点信息类型。为了进行堆排序,定义:

```
TYPE node = RECORD
    key: integer;
    info: datatype
END;

VAR heap: ARRAY[1..n] of node;
    l, r, i, j: 0..n;
    x: node;
```

在下面的算法描述中填入正确的内容,使其实现1964年Floyd提出的建堆筛选法,要求堆建成后便找到了最小的关键码。

筛选算法 sift(l, r, heap)

步1. [准备] $i \leftarrow l$; $j \leftarrow$ _____; $x \leftarrow \text{heap}[i]$

步2. [过筛]

循环 当 _____ 时反复执行

(1) 若 $j < r$ 且 $\text{heap}[j].\text{key} > \text{heap}[j+1].\text{key}$ 则 _____

(2) 若 _____ 则 $\text{heap}[i] \leftarrow \text{heap}[j]$; _____; _____

否则跳出循环

步3. [结束]

$\text{heap}[i] \leftarrow$ _____

3. (8分) 回答问题 (每题4分)

(1) 设待排序的结点个数是 n 。试问堆排序算法在完成一次 sift 建堆,并且取走找到的最小关键码之后,是否还需要对于 $n-1$ 个关键码从头开始建堆?为什么?

(2) 假定采用 sift 建堆算法。试问堆排序算法采用了怎样的节省存储空间措施?堆排序完成后,heap 中保存了关键码值的升序排列还是降序排列?

四、(20分)

根据定义,二叉树由结点的有限集合构成,这个有限集合或者为空集,或者由一个根结点及两棵不相交的左子树和右子树(它们也是结点的集合)组成。因此,可以通过二叉树实现集合运算。设二叉树结点的形式为

llink	info	rlink
-------	------	-------

1. (8分) 试写一基于二叉树的集合算法, 判定 $x \in A$ 的真实性。其中 A 是有限集合, A 的二叉树表示已经建立并且记为 T , x 与 $info$ 为同类型数据。

2. (8分) 试写一基于二叉树的算法, 实现集合运算 $A = A \cup B$ 。其中 A , B 为同类型数据元素的集合, 并且 A 和 B 的二叉树表示均已建立, 分别记作 TA 和 TB 。

3. (4分) 目前大多数微型计算机仍然采用16位字长, 即双字节长度。若集合 X 恰巧含有 16 个字符元素, 并且只给定一个字的存储空间, 试述一种表示 X 的方法。在你提出的方法中, 如何实现集合的交运算, 即 $X \cap Y$, 其中 X 和 Y 都是只有 16 个元素的集合, 并且各占一个字空间。

五、(10分)

背包问题可以简化为: 设某个背包可以放入的物品重量为 S 。现有 n 件物品, 重量分别为 w_1, w_2, \dots, w_n 。问能否从这 n 件物品中选择若干件放入此背包, 使得放入的重量之和正好为 S 。设 w_i 均为正整数, $i=1, 2, \dots, n$, 并且已顺序地存放在数组 W 中, 则下面是一个实现背包问题的递归算法。

布尔函数 $Knap(S, n)$

若 $S = 0$

则 $Knap \leftarrow true$

否则若 $(S < 0)$ 或 $(S > 0 \text{ 且 } n < 1)$

则 $Knap \leftarrow false$

否则若 $Knap(S - W[n], n-1) = true$

则 $print(W[n]);$

$Knap \leftarrow true$

否则 $Knap \leftarrow Knap(S, n-1)$

给定 $W = [1, 4, 4, 5, 7]$ 。试模拟求出背包问题 $Knap(10, 5)$ 为真的解, 即按照执行顺序列出 $print$ 语句的输出结果(包括数组元素的下标和数组元素的值。只列出元素值, 不列出对应下标者不给分)。

六、(10分)

1. (2分) 填空(每空0.5分)

二维图形是计算机处理的一类重要数据, 在磁盘上主要有两种存储方式, 它们是 _____ 和 _____。前者的主要优点是 _____, 后者的主要优点是 _____。

2. (3分) 城市交通图是一种二维图形, 其拓扑结构可以用网(边上带权图)表示。试述出三种城市交通图的内存存储结构。

3. (5分) 下面是一种关于有向图拓扑排序方法的叙述:

对于任意有向图, 做

步1 从图中选择一个入度为 0 的结点, 例如 V , 输出之。

步2 从图中删掉 V 及其所有的出边。

反复执行这两个步骤，直到所有的结点都输出了，也就是整个拓扑排序完成了。

试问：上述方法表述正确吗？为什么？

七、(10分)

1. (6分) 设待排序的关键码分别为 28, 13, 72, 85, 39, 41, 6, 20。按二分法插入排序算法已使前 7 个记录有序，中间结果如下：

6	13	28	39	41	72	85	20
↑			↑			↑	
l=1			m=4			r=7	

试在此基础上，沿用上述表达方式，给出继续采用二分法插入第八个记录的比较过程。

2. (4分) 回答问题(每问 2 分)

(1) 使用二分法插入排序所要进行的比较次数，是否与待排序的记录的初始状态有关？

(2) 在一些特殊情况下，二分法插入排序比直接插入排序要执行更多的比较。这句话对吗？

6 13 28 39

↑ ↑ ↑

l. m. 4

19

6 13 28 39

↑

l y m.

~~6 13 28 39~~

~~19~~
~~20~~

13 19 39 41

↑ ↑

y l.

6 13 28 39

↑ ↑

y m. l