



中国矿业大学（北京）2019 年招收攻读硕士学位研究生

入学考试模拟试题（二）

考试科目：计算机技术数据结构 考试时间：180 分钟 总分：150 分

命题时间：2018 年 11.27 日 命题人：育明教育考研考博研究室

考生注意：

1. 本试题的答案必须写在规定的答题纸上，写在试卷上一律不给分。
2. 考试结束后，将答题纸和试卷一并装入试卷袋内。

一、 问答题(75 分)

数据结构是一门研究非数值计算的程序设计问题中计算机的操作系统对象以及它们之间的关系和操作等的学科。假如目前需要抽象出一个部门员工之间的隶属关系，该部门目前有部门主任一名，组长 1-3 名，每组成员 3-5 名，主任主管各组长，各组长负责组内组员工作分配，请你就自己所学，对这个部门的隶属关系进行数据结构的抽象，并对抽象过程进行说明。（15 分）

考点：考察对于数据结构抽象的能力。

分析：本题的入手应该从数据结构的基本思想入手，可以看到隶属关系这个提示词，并且题意描述中已经明确说明了每一级有什么人，这是一种典型的树形结构，并且根据数据结构的抽象过程，形成一个二元组的描述即可得分，也就是 $\text{Department} = (P, R)$ ，一个典型的二元组结构。



答案:

参考课本 53 页, 例 1-5

答出二元组得 10 分, 答出完整抽象过程得 5 分。

1. 请分别说明栈、队列和数组它们之间的共同点和不同点。(10 分)

考点: 基本概念考察

答案:

相同点:

这三种结构都是线性的, 是源自线性结构这种数据结构, 它们都是在线性表基础上进行了限制 (2 分)

它们之中存储的数据元素必须是同一种数据类型 (2 分)

不同点:

数组一旦被定义, 它的维度和维界就不能再改变, 但是栈和队列可以根据需求再扩充 (2 分)

栈是后进先出的线性表, 删除和插入只能在表尾进行; 队是先进先出的线性表, 插入在表尾进行, 删除在表头进行; 数组是可以完成随机存取的线性表 (4 分)

在计算机中, 为了处理字符串, 人们使用了“串”来完成这个工作。在后续的发展中, 随着对字符串处理的深入研究, 不同的方法被相继提出, 如 KMP 这一被人们熟知并广泛应用的算法。(15 分)

简述串与线性表的异同。(2 分)



字符串是专门针对字符的一种线性结构，与线性表相同，都是线性结构的，只不过它的研究对象只限于字符串。

请简述 KMP 算法的原理。(8 分)

参照课本 81-82 页。答题时应包括基本思路以及相关算法流程，本题如果没有算法流程应该扣 2 分。

请试求模式串“ababaca”的 next[j]。(5 分)

[0, 0, 1, 2, 3, 0, 1]

答出过程得满分，少过程扣 2 分

简述二叉树常用的几种遍历方式，并选择其中一种排序方式进行详细描述。(10 分)

考点：考察二叉树的性质以及遍历方式，以及相关遍历方式的算法

分析：先答二叉树的性质，再从先序、中序、后序、层序这几种方式入手去答，并选择一种方式进行算法描述。

答案：二叉树一种树形结构，其特点是每个结点至多只有两棵子树（即二叉树不存在度大于 2 的结点），并且二叉树有左右子树之分，其次序不能任意颠倒。

与树相似，二叉树也以递归的形式定义，二叉树是 n ($n \geq 0$) 个结点的有限集合。二叉树有四种遍历方式，分别为先序遍历、中序遍历、后序遍历、层序遍历。(4 分)

参照王道书上二叉树那一章，选择一种遍历方式，并对算法进行描述。(6 分)

1. 动态查找表的特点是什么？试说明两种建立在动态查找表上的查



找方法（20分）

考点：动态查找表的特点以及相关动态查找表的理解

分析：首先对于动态查找表的定义进行描述，然后再举出两个例子进行描述，需要对动态查找的几个方法有一定掌握

答案：

动态查找表特点：表结构本身是在查找过程中动态生成的，即对于给定值 key，如表中存在关键字等于 key 的元素，则查找成功，返回该元素，否则插入关键字 key 的记录。（4分）

举例：

二叉排序树：在二叉树的基础上形成的一种查找树，具有二叉树的性质。特点是若左子树不空，则左子树上所有的结点 key 值小于它的根结点 key 值，若右子树不空，则右子树上所有的结点的 key 值大于根结点的 key 值，左右子树也满足二叉排序树的定义。（3分）

查找方法：当二叉排序树不空是，首先将给定值和根结点的关键字 key 值进行比较，若相等，则查找成功，返回，若不相等，则根据比较结果，分别在左子树或者右子树上继续查找，直到找到或者未找到；如果未找到，就需要进行插入操作，将未找到的值按照插入规则插入到二叉排序树中，动态的对树进行扩充。（5分）

B-树：特点是对于 m 阶的 B-树，树中的每个结点至多有 m 棵子树，若根结点不是叶子结点，至少有两颗子树，除根结点之外的所有非终端结点至少有 棵子树，所有的非终端结点包含以下信息 $(n, A_0, K_1, A_1, K_2, A_2, \dots, K_n, A_n)$ ，其中 K_i 为关键字，



$K_i < K_{i+1}$, A_i 为指向子树根结点的指针, 且 A_{i-1} 所指子树中所有结点的关键字均小于 K_i ; A_n 所指的树中所有结点关键字均大于 K_n , $n()$ 为关键字的个数。所有叶子结点都出现在同一层上, 并且不带信息。(3分)

查找方法: 首先从根开始, 根据结点指针找到根结点, 根据所查关键字与结点上的关键字比较, 找到相等值, 则查找成功, 否则, 就顺着对应的指针找到下层结点, 直到叶子结点, 若始终未找到与所查关键字相等的值, 就就要进行相应的插入操作, 动态的扩充树。(5分)

二、求解以下问题 (75分)

1. 已知一个带有表头结点的单链表, 结点结构为

Date	Link
------	------

假设该链表只给出了头指针 list。再不改变链表的前提下, 请设计一个尽可能高效的算法, 查找链表中倒数 k 个位置上的结点 (k 为正整数)。若查找成功, 算法输出该结点的 date 域的值, 并返回 1; 否则, 只返回 0。(15分)

要求:

(1) 描述算法的基本设计思想 (5分)

(2) 描述算法的详细实现步骤 (8分)

使用 c、c++、java 描述算法过程

(3) 根据设计思想和实现步骤, 计算算法的时间复杂度和空间复杂度。(2分)

参考王道 2.3.7 节 (37 页左右) 2009 年计算机联考真题, 暴力解



决得分最多 15 分

2. 求解下面有向图的有关问题（15 分）

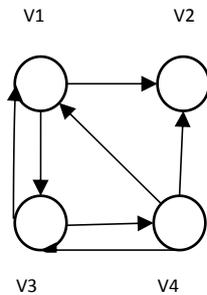
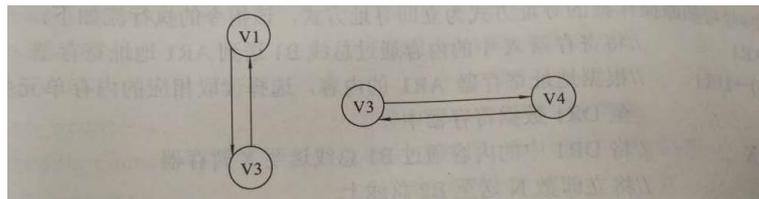
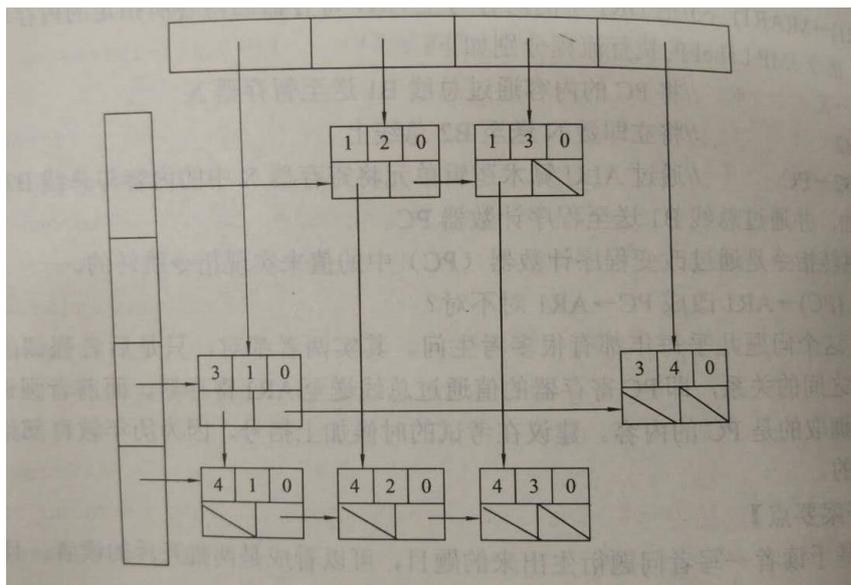


图 1. 有向图 G1

(1) 判断此有向图是否有强连通分量，如果有，请画出。（3 分）



(2) 画出此图的十字链表存储结构。（5 分）



(3) 简述基于图的深度优先搜索策略，并判断一个以邻接表存储的

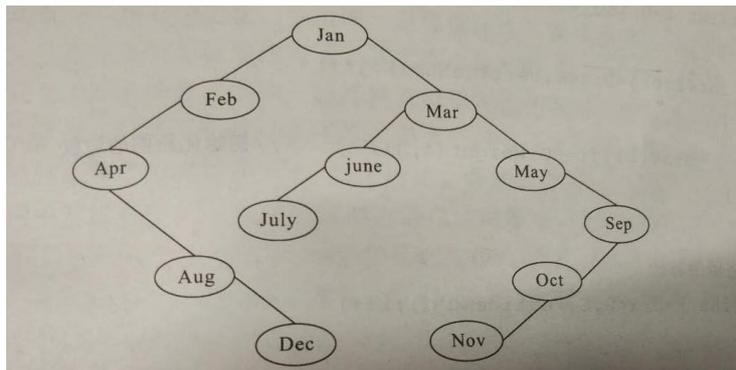


有向图是否存在定点 v_i 到 v_j 的路径的基本步骤。(7分)

(3) 将起始结点入栈并标记, 将与此结点相邻的结点依次入栈并标记, 如果相邻结点有目标结点 j 则输出成功, 否则出栈一个结点, 将与此结点相邻的结点依次入栈并标记, 直到栈空, 返回失败。

3. 已知一个长度为 12 的表 {Jan, Feb, Mar, Apr, May, June, July, Aug, Sep, Oct, Nov, Dec}; (10分)

- (1) 是按照表中元素的顺序依次插入一棵初始为空的二叉排序树 (字符之间以字典序比较大小), 请画出最终对应的二叉排序树。(5分)



- (2) 若对表中的元素先进行排序构成有序表 (字典序), 试求在等概率情况下对此有序表进行检索成功的平均检索长度。(5分)

(2) 按字典序对表中元素进行排序, 得: Apr, Aug, Dec, Feb, Jan, July, June, Mar, May, Nov, Oct, Sep, 其查找成功的平均查找长度为: $(5+7+7+4+3+10+7+4+6+8+7+6) / 12 = 74/12 = 6.2$ 。

查找成功时的平均查找长度为: $(6+5+7+6+4+9+7+3+6+6+5+4) / 12 = 68/12 = 5.7$ 。

4. 画出相应的散列表, 当发生冲突时, 采用线性探测法解决。假定每个记录的查找概率相等, 计算查找成功和查找失败时的平均查



找长度 ASL。(5 分) 画出相应的散列表，当发生冲突时，采用链式有一节点的关键字序列 $F=\{129, 72, 180, 105, 147, 96, 45, 69\}$ ，散列函数为： $H(key)=key \bmod 11$ ，其中 key 为关键字，散列地址空间为 $0 \sim 10$ 。(20 分)

要求：

(1) 画出相应的散列表，当发生冲突时，采用线性探测法解决。

假定每个记录的查找概率相等，计算查找成功和查找失败时的平均查找长度 ASL。(5 分)

(1) 采用线性探测法处理冲突建立的散列表如下：

$$H(129)=129 \bmod 11=8$$

$$H(72)=72 \bmod 11=6$$

$$H(180)=180 \bmod 11=4$$

$$H(105)=105 \bmod 11=6 \text{ 冲突 } H_1(105)=(105+1) \bmod 11=7$$

$$H(147)=147 \bmod 11=4 \text{ 冲突 } H_1(147)=(147+1) \bmod 11=5$$

$$H(96)=96 \bmod 11=8 \text{ 冲突 } H_1(96)=(96+1) \bmod 11=9$$

$$H(45)=45 \bmod 11=1$$

$$H(69)=69 \bmod 11=3$$

综上所述，散列表如表 4-7 所示。

表 4-7 线性探测法处理冲突建立的散列表

下标	0	1	2	3	4	5	6	7	8	9	10
k		45		69	180	147	72	105	129	96	
探测次数		1		1	1	2	1	2	1	2	

装填因子 $\alpha=8/11$ 。

$$ASL_{succ}=(5 \times 1+2 \times 3)/8=11/8$$

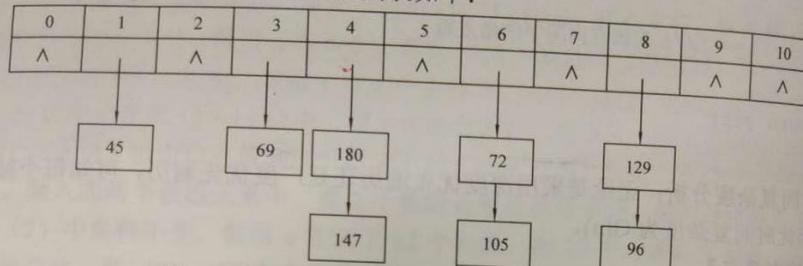
$$ASL_{unsucc}=(1+2+1+8+7+6+5+4+3+2+1)/11=40/11$$

(2) 画出相应的散列表，当发生冲突时，采用链式地址法解决。

假定每个记录的查找概率相等，计算查找成功和查找失败时的平均查找长度 ASL (只将与关键字比较次数计算在内)。(5 分)



(2) 采用链地址法处理冲突建立的散列表如下:



$$ASL_{succ} = (5 \times 1 + 3 \times 2) / 8 = 11/8$$

只找一次 不成功一定会找完链长

$$ASL_{unsucc} = (0 + 1 + 0 + 1 + 2 + 0 + 2 + 0 + 2 + 0 + 0) / 11 = 8/11$$

空就不比

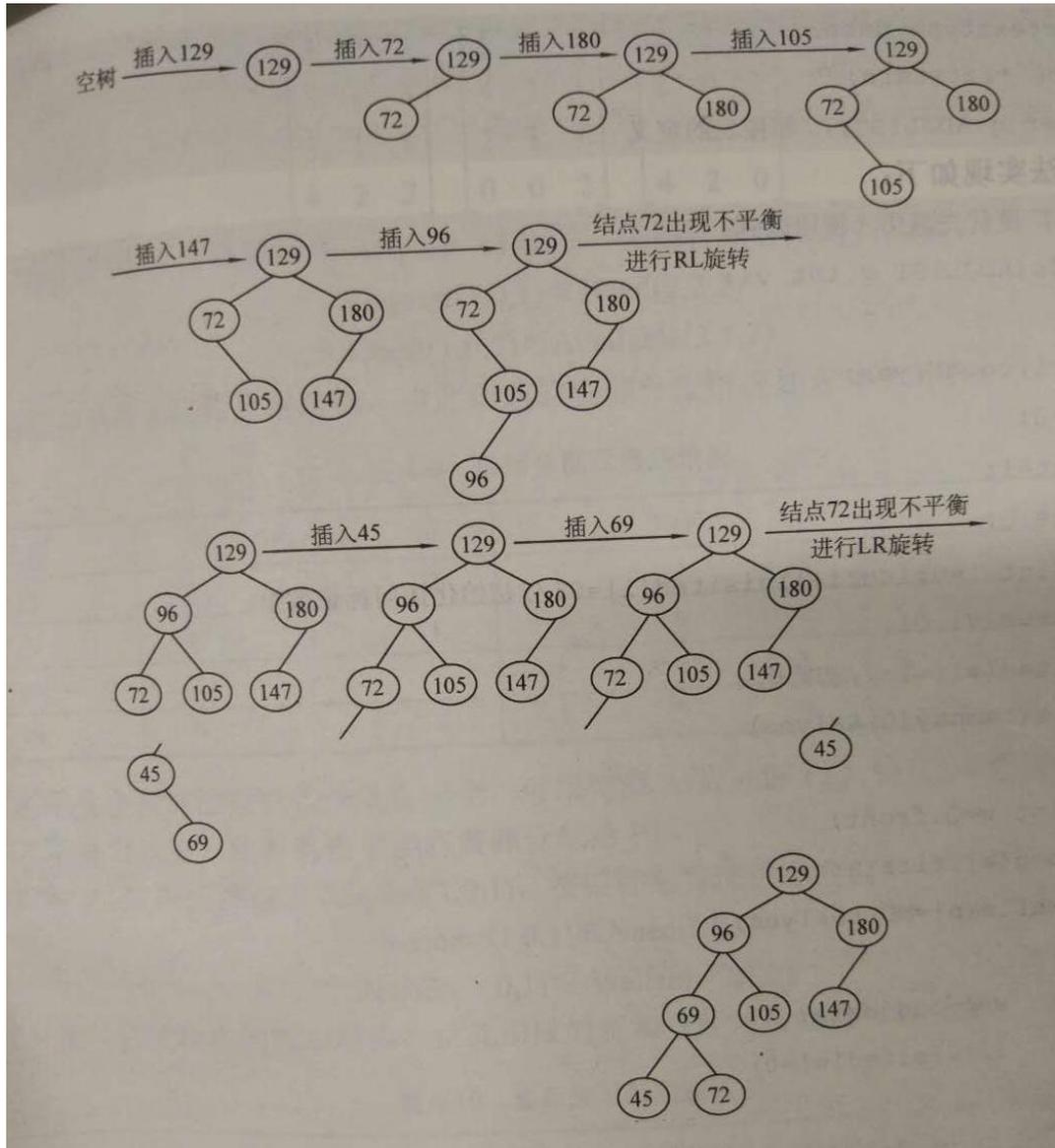
说明: 本题中已经明确指出采用链地址法, 只将与关键字的比较计算在内, 因此上述

ASL_{unsucc} 的解法是正确的。如果本题要求将空指针的比较也包括在内, 则

$$ASL_{unsucc} = (1 + 2 + 1 + 2 + 3 + 1 + 3 + 1 + 3 + 1 + 1) / 11 = 19/11$$

提醒: 散列表查找成功的平均查找长度与查找不成功的平均查找长度是不一样的计算方式, 千万要注意!

(3) 试按照各关键字在序列 F 中的次序将他们一次插入一棵初始为空的平衡二叉树中, 画出详细步骤并说明。(10 分)



5. 设记录的关键字集合 $key = \{50, 40, 66, 88, 72, 8, 20\}$ 。(15分)

(1) 写出用快速排序各趟排序之后的状态。(5分)

(2) 把关键字集合 key 调整成堆顶元素区最小值的堆(只写出简要过程)。(5分)

(3) 说明上述两种排序分别基于哪种基本思想? 稳定性如何?
(5分)

(4) 答案参考历年真题的那个本子上排序那一章 (03, 三)

+[07. 12]

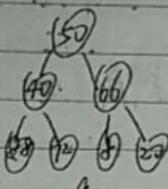
07. 三]. 设记录的关键字集合 $key = (50, 40, 66, 88, 72, 8, 20)$

- 1). 写出用快速排序法各趟排序之后的状态。
- 2). 把关键字集合 key 调整成堆顶元素, 取最小值的堆 (写出简要调整过程)。
- 3). 说明上述两种排序分别基于哪种基本思想? 其稳定性如何?

解: 第一趟: 20, 40, 8, 50, 72, 88, 66

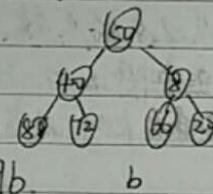
第二趟: 8, 20, 40, 50, 66, 72, 88

107. 如图 A 是初始序列



② 筛选从第一个元素开始, 由于 $66 > 8$, 则交换, 如图 b 为:

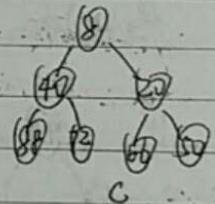
交换后的图形如图 b.



③ 由于 8 不大于其左右子树根的值, 筛选后的图不变, 即为图 b.

④ 筛选第一个元素的值, 由于 $50 > 8$, 且 $50 > 20$, 则筛选后如图 C.

图 C 即为堆顶元素为最小值的堆.



3). 快速排序: 以一个“枢轴”为中心, 将序列分成两部分, 枢轴轴的一边全是比它小 (或者小于等于) 的, 另一边比它大 (或者大于等于) 的。

堆排序: 代表堆的这棵完全二叉树的根结点的值是最大 (或最小) 的, 因此将一个无序序列调整为一个堆, 就可以找出这个序列的最大值 (或最小值), 然后新找出的这个值交换到序列的最后 (或最前), 这样有序元素增加 1, 无序元素减少 1, 对新的无序序列重复这样的操作, 实现了排序。

都不稳定。