

本本
做题本集结地

联系我: benbteam@126.com

考研408 2026
WD·数据结构
综合题 做题本

安 静 做 题 · 认 真 提 分

目 录

1.1.3	数据结构的基本概念应用题(原书 P3)	2
1.2.3	算法和算法评价应用题(原书 P6)	2
2.2.3	线性表的顺序表示应用题(原书 P19)	3
2.3.7	线性表的链式表示应用题(原书 P42)	13
3.1.4	栈应用题(原书 P68)	26
3.2.5	队列应用题(原书 P84)	28
3.3.6	栈和队列的应用应用题(原书 P96)	31
4.2.4	串的模式匹配应用题(原书 P119)	32
5.1.4	树的基本概念应用题(原书 P127)	34
5.2.3	二叉树的概念应用题(原书 P135)	35
5.3.3	二叉树的遍历和线索二叉树应用题(原书 P152)	38
5.4.4	树、森林应用题(原书 P177)	48
5.5.3	树与二叉树的应用应用题(原书 P190)	50
6.1.2	图的基本概念应用题(原书 P203)	52
6.2.6	图的存储及基本操作应用题(原书 P213)	52
6.3.4	图的遍历应用题(原书 P225)	58
6.4.6	图的应用应用题(原书 P248)	62
7.2.4	顺序查找和折半查找应用题(原书 P276)	70
7.3.4	树形查找应用题(原书 P300)	75
7.4.3	B 树和 B+树应用题(原书 P317)	80
7.5.5	散列(Hash)表应用题(原书 P329)	81
8.2.4	插入排序应用题(原书 P344)	85
8.3.3	交换排序应用题(原书 P352)	85
8.4.3	选择排序应用题(原书 P363)	89
8.5.4	归并排序、基数排序与计数排序应用题(原书 P376)	94
8.6.3	各种内部排序算法的比较及应用应用题(原书 P384)	96
8.7.6	外部排序应用题(原书 P393)	98

注：所有题目已有偿拜托专人核查并修正错误。

本文件是「做题本集结地」账号制作并免费分享。如您觉得它对您有帮助，欢迎将该文件分享给其他需要的朋友。打个小广：如果您还没有找到信任的打印店铺，可以来公众号小铺子看一眼。质量好价格便宜，并且相对于免费分享版，会有额外的氪金福利。欢迎选购。

1.1.3 数据结构的基本概念应用题(原书 P3)

(01) 对于两种不同的数据结构,逻辑结构或物理结构一定不相同吗?

(02) 试举一例, 说明对相同的逻辑结构, 同一种运算在不同的存储方式下实现时, 其运算效率不同。

1.2.3 算法和算法评价应用题(原书 P6)

(01) 分析下列各程序段,求出算法的时间复杂度。

(1)

```
i = 1, k = 0;
while( i < n-1){
    k = k + 10* i;
    i++;
}
```

(2)

```
y=0;
while((y+1)*(y+1) <= n)
    y= y + 1;
}
```

(3)

```
for(i=0, i<n, i++)
    for(j=0, j<m, j++)
        a[i][j] = 0;
```

by 本本

2.2.3 线性表的顺序表示应用题(原书 P19)

(01) 设计一个算法,从顺序表中删除具有最小值的元素(假设唯一)并由函数返回被删元素的值。空出的位置由最后一个元素填补,若顺序表为空,则显示出错信息并退出运行。

(02) 设计一个高效算法,将顺序表 L 的所有元素逆置,要求算法的空间复杂度为 $O(1)$ 。

(03) 对长度为 n 的顺序表 L , 编写一个时间复杂度为 $O(n)$ 、空间复杂度为 $O(1)$ 的算法, 该算法删除顺序表中所有值为 x 的数据元素。

(04) 从顺序表中删除其值在给定值 s 和 t 之间 (包含 s 和 t , 要求 $s < t$) 的所有元素, 若 s 或 t 不合理或顺序表为空, 则显示出错信息并退出运行。

(05) 从有序顺序表中删除所有其值重复的元素, 使表中所有元素的值均不同。

(06) 将两个有序顺序表合并为一个新的有序顺序表, 并由函数返回结果顺序表。

(07) 已知在一维数组 $A[m+n]$ 中依次存放两个线性表 $(a_1, a_2, a_3, \dots, a_m)$ 和 $(b_1, b_2, b_3, \dots, b_n)$ 。编写一个函数, 将数组中两个顺序表的位置互换, 即将 $(b_1, b_2, b_3, \dots, b_n)$ 放在 $(a_1, a_2, a_3, \dots, a_m)$ 的前面。

(08) 线性表 $(a_1, a_2, a_3, \dots, a_n)$ 中的元素递增有序且按顺序存储于计算机内。要求设计一个算法, 完成用最少时间在表中查找数值为 x 的元素, 若找到, 则将其与后继元素位置相交换, 若找不到, 则将其插入表中并使表中元素仍递增有序。

(09) 给定三个序列 A 、 B 、 C , 长度均为 n , 且均为无重复元素的递增序列, 请设计一个时间上尽可能高效的算法, 逐行输出同时存在于这三个序列中的所有元素。例如, 数组 A 为 $\{1, 2, 3\}$, 数组 B 为 $\{2, 3, 4\}$, 数组 C 为 $\{-1, 0, 2\}$, 则输出 2。要求:

- 1) 给出算法的基本设计思想。
- 2) 根据设计思想, 采用 C 或 $C++$ 语言描述算法, 关键之处给出注释。
- 3) 说明你的算法的时间复杂度和空间复杂度。

(10) 【2010 统考真题】设将 $n(n>1)$ 个整数存放于一维数组 R 中。设计一个在时间和空间两方面都尽可能高效的算法。将 R 中保存的序列循环左移 $p(0<p<n)$ 个位置, 即将 R 中的数据由 $(X_0, X_1, \dots, X_{n-1})$ 变换为 $(X_p, X_{p+1}, \dots, X_{n-1}, X_0, X_1, \dots, X_{p-1})$ 。要求:

- 1) 给出算法的基本设计思想。
- 2) 根据设计思想, 采用 C 或 $C++$ 或 $Java$ 语言描述算法, 关键之处给出注释。
- 3) 说明你所设计算法的时间复杂度和空间复杂度。

(11) 【2011 统考真题】一个长度为 $L(L \geq 1)$ 的升序序列 S , 处在第 $\lceil L/2 \rceil$ 个位置的数称为 S 的中位数。例如, 若序列 $S_1 = (11, 13, 15, 17, 19)$, 则 S_1 的中位数是 15, 两个序列的中位数是含它们所有元素的升序序列的中位数。例如, 若 $S_2 = (2, 4, 6, 8, 20)$, 则 S_1 和 S_2 的中位数是 11。现在有两个等长升序序列 A 和 B , 试设计一个在时间和空间两方面都尽可能高效的算法, 找出两个序列 A 和 B 的中位数。

要求:

- 1) 给出算法的基本设计思想。
- 2) 根据设计思想, 采用 C 或 $C++$ 或 $Java$ 语言描述算法, 关键之处给出注释。
- 3) 说明你所设计算法的时间复杂度和空间复杂度。

(12) 【2013 统考真题】已知一个整数序列 $A = (a_0, a_1, \dots, a_{n-1})$, 其中 $0 \leq a_i < n (0 \leq i < n)$ 。若存在 $a_{p_1} = a_{p_2} = \dots = a_{p_m} = x$ 且 $m > n/2 (0 \leq p_k < n, 1 \leq k \leq m)$, 则称 x 为 A 的主元素。例如 $A = (0, 5, 5, 3, 5, 7, 5, 5)$, 则 5 为主元素; 又如 $A = (0, 5, 5, 3, 5, 1, 5, 7)$, 则 A 中没有主元素。假设 A 中的 n 个元素保存在一个一维数组中, 请设计一个尽可能高效的算法, 找出 A 的主元素。若存在主元素, 则输出该元素; 否则输出 -1。要求:

- 1) 给出算法的基本设计思想。
- 2) 根据设计思想, 采用 C 或 C++ 或 Java 语言描述算法, 关键之处给出注释。
- 3) 说明你所设计算法的时间复杂度和空间复杂度。

(13) 【2018 统考真题】给定一个含 $n(n \geq 1)$ 个整数的数组, 请设计一个在时间上尽可能高效的算法, 找出数组中未出现的最小正整数。例如, 数组 $\{-5, 3, 2, 3\}$ 中未出现的最小正整数是 1; 数组 $\{1, 2, 3\}$ 中未出现的最小正整数是 4。要求:

- 1) 给出算法的基本设计思想。
- 2) 根据设计思想, 采用 C 或 C++ 语言描述算法, 关键之处给出注释。
- 3) 说明你所设计算法的时间复杂度和空间复杂度。

(14) 【2020 统考真题】定义三元组 (a, b, c) (a, b, c 均为整数) 的距离 $D = |a - b| + |b - c| + |c - a|$ 。给定 3 个非空整数集合 S_1 、 S_2 和 S_3 , 按升序分别存储在 3 个数组中。请设计一个尽可能高效的算法, 计算并输出所有可能的三元组 (a, b, c) ($a \in S_1, b \in S_2, c \in S_3$) 中的最小距离。例如 $S_1 = \{-1, 0, 9\}, S_2 = \{-25, -10, 10, 11\}, S_3 = \{2, 9, 17, 30, 41\}$, 则最小距离为 2, 相应的三元组为 $(9, 10, 9)$ 。要求:

- 1) 给出算法的基本设计思想。
- 2) 根据设计思想, 采用 C 语言或 C++ 语言描述算法, 关键之处给出注释。
- 3) 说明你所设计算法的时间复杂度和空间复杂度。

2.3.7 线性表的链式表示应用题(原书 P42)

(01) 在带头结点的单链表 L 中, 删除所有值为 x 的结点, 并释放其空间, 假设值为 x 的结点不唯一, 试编写算法以实现上述操作。

(02) 试编写在带头结点的单链表 L 中删除一个最小值结点的高效算法 (假设该结点唯一)。

(03) 试编写算法将带头结点的单链表就地逆置, 所谓“就地”是指辅助空间复杂度为 $O(1)$ 。

(04) 设在一个带头结点的单链表中, 所有结点的元素值无序, 试编写一个函数, 删除表中所有处于给定的两个值 (作为函数参数给出) 之间的元素 (若存在)。

(05) 给定两个单链表, 试分析找出两个链表的公共结点的思想 (不用写代码)。

(06) 设 $C = \{a_1, b_1, a_2, b_2, \dots, a_n, b_n\}$ 为线性表, 采用带头结点的单链表存放, 设计一个就地算法, 将其拆分为两个线性表, 使得 $A = \{a_1, a_2, \dots, a_n\}, B = \{b_n, \dots, b_2, b_1\}$ 。

(07) 在一个递增有序的单链表中,存在重复的元素。设计算法删除重复的元素,例如 (7,10,10,21,30,42,42,42,51,70) 将变为 (7,10,21,30,42,51,70)。

(08) 设 A 和 B 是两个单链表 (带头结点), 其中元素递增有序。设计一个算法从 A 和 B 中的公共元素产生单链表 C , 要求不破坏 A 、 B 的结点。

(09) 已知两个链表 A 和 B 分别表示两个集合, 其元素递增排列。编制函数, 求 A 与 B 的交集, 并存放于 A 链表中。

(10) 两个整数序列 $A = a_1, a_2, a_3, \dots, a_m$ 和 $B = b_1, b_2, b_3, \dots, b_n$ 已经存入两个单链表中, 设计一个算法, 判断序列 B 是否是序列 A 的连续子序列。

(11) 设计一个算法用于判断带头结点的循环双链表是否对称。

(12) 有两个循环单链表, 链表头指针分别为 $h1$ 和 $h2$, 编写一个函数将链表 $h2$ 链接到链表 $h1$ 之后, 要求链接后的链表仍保持循环链表形式。

(13) 设有一个带头结点的非循环双链表 L , 其每个结点中除有 pre 、 $data$ 和 $next$ 域外, 还有一个访问频度域 $freq$, 其值均初始化为零。每当在链表中进行一次 $Locate(L, x)$ 运算时, 令值为 x 的结点中 $freq$ 域的值增 1, 并使此链表中的结点保持按访问频度递减的顺序排列, 且最近访问的结点排在频度相同的结点之前, 以便使频繁访问的结点总是靠近表头。试编写符合上述要求的 $Locate(L, x)$ 函数, 返回找到结点的地址, 类型为指针型。

(14) 设将 $n(n>1)$ 个整数存放到不带头结点的单链表 L 中, 设计算法将 L 中保存的序列循环右移 $k(0<k<n)$ 个位置。例如, 若 $k=1$, 则将链表 $\{0,1,2,3\}$ 变为 $\{3,0,1,2\}$ 。要求:

- 1) 给出算法的基本设计思想。
- 2) 根据设计思想, 采用 C 或 $C++$ 语言描述算法, 关键之处给出注释。
- 3) 说明你所设计算法的时间复杂度和空间复杂度。

(15) 单链表有环,是指单链表的最后一个结点的指针指向了链表中的某个结点 (通常单链表的最后一个结点的指针域是空的)。试编写算法判断单链表是否存在环。

- 1) 给出算法的基本设计思想。
- 2) 根据设计思想, 采用 C 或 $C++$ 语言描述算法, 关键之处给出注释。
- 3) 说明你所设计算法的时间复杂度和空间复杂度。

(16) 设有一个长度 n (n 为偶数) 的不带头结点的单链表, 且结点值都大于 0, 设计算法求这个单链表的_{最大孪生和}。孪生和定义为一个结点值与其孪生结点值之和, 对于第 i 个结点 (从 0 开始), 其孪生结点为第 $n-i-1$ 个结点。要求:

- 1) 给出算法的基本设计思想。
- 2) 根据设计思想, 采用 C 或 C++ 语言描述算法, 关键之处给出注释。
- 3) 说明你的算法的时间复杂度和空间复杂度。

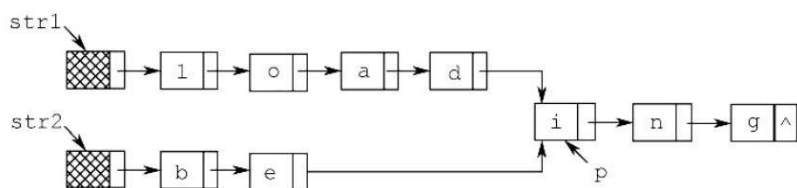
(17) 【2009 统考真题】已知一个带有表头结点的单链表, 结点结构为

data	link
------	------

假设该链表只给出了头指针 *list*。在不改变链表的前提下, 请设计一个尽可能高效的算法, 查找链表中倒数第 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 *data* 域的值, 并返回 1; 否则, 只返回 0。要求:

- 1) 描述算法的基本设计思想。
- 2) 描述算法的详细实现步骤。
- 3) 根据设计思想和实现步骤, 采用程序设计语言描述算法 (使用 *C*、*C++* 或 *Java* 语言实现), 关键之处请给出简要注释。

(18) 【2012 统考真题】假定采用带头结点的单链表保存单词,当两个单词有相同的后缀时,可共享相同的后缀存储空间,例如, *loading* 和 *being* 的存储映像如下图所示。



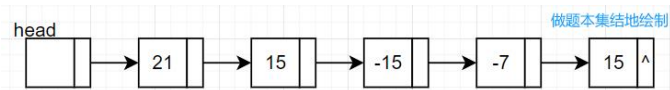
设 *str1* 和 *str2* 分别指向两个单词所在单链表的头结点,链表结点结构为

data	link
------	------

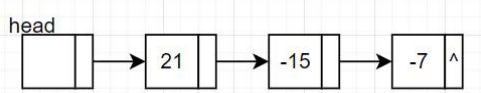
请设计一个时间上尽可能高效的算法,找出由 *str1* 和 *str2* 所指向两个链表共同后缀的起始位置 (如图中字符 *i* 所在结点的位置 *p*)。要求:

- 1) 给出算法的基本设计思想。
- 2) 根据设计思想,采用 *C* 或 *C++* 或 *Java* 语言描述算法,关键之处给出注释。
- 3) 说明你所设计算法的时间复杂度。

(19) 【2015 统考真题】用单链表保存 m 个整数, 结点的结构为 $[data][link]$, 且 $|data| \leq n$ (n 为正整数)。现要求设计一个时间复杂度尽可能高效的算法, 对于链表中 $data$ 的绝对值相等的结点, 仅保留第一次出现的结点而删除其余绝对值相等的结点。例如, 若给定的单链表 $head$ 如下:



则删除结点后的 $head$ 为



要求:

- 1) 给出算法的基本设计思想。
- 2) 使用 C 或 $C++$ 语言, 给出单链表结点的数据类型定义。
- 3) 根据设计思想, 采用 C 或 $C++$ 语言描述算法, 关键之处给出注释。
- 4) 说明你所设计算法的时间复杂度和空间复杂度。

(20) 【2019 统考真题】设线性表 $L = (a_1, a_2, a_3, \dots, a_{n-2}, a_{n-1}, a_n)$ 采用带头结点的单链表保存, 链表中的结点定义如下:

```
typedef struct node
{
    int data;
    struct node *next;
}NODE;
```

请设计一个空间复杂度为 $O(1)$ 且时间上尽可能高效的算法, 重新排列 L 中的各结点, 得到线性表 $L' = (a_1, a_n, a_2, a_{n-1}, a_3, a_{n-2}, \dots)$ 。要求:

- 1) 给出算法的基本设计思想。
- 2) 根据设计思想, 采用 C 或 $C++$ 语言描述算法, 关键之处给出注释。
- 3) 说明你所设计的算法的时间复杂度。

3.1.4 栈应用题(原书 P68)

(01) 有 5 个元素, 其入栈次序为 A, B, C, D, E , 在各种可能的出栈次序中, 第一个出栈元素为 C 且第二个出栈元素为 D 的出栈序列有哪几个?

(02) 若元素的入栈序列为 A, B, C, D, E , 运用栈操作, 能否得到出栈序列 B, C, A, E, D 和 D, B, A, C, E ? 为什么?

(03) 栈的初态和终态均为空, 以 I 和 O 分别表示入栈和出栈, 则出入栈的操作序列可表示为由 I 和 O 组成的序列, 可以操作的序列称为合法序列, 否则称为非法序列。

1) 下面所示的序列中哪些是合法的?

A. $IOHIOIOO$

B. $IOOIOHIO$

C. $IIHIOIOIO$

D. $IIIOOIOO$

2) 通过对 1) 的分析, 写出一个算法, 判定所给的操作序列是否合法。若合法, 返回 *true*, 否则返回 *false* (假定被判定的操作序列已存入一维数组中)。

(04) 设单链表的表头指针为 L , 结点结构由 $data$ 和 $next$ 两个域构成, 其中 $data$ 域为字符型。试设计算法判断该链表的全部 n 个字符是否中心对称。例如 xyx 、 $xyyx$ 都是中心对称。

(05) 设有两个栈 $S1$ 、 $S2$ 都采用顺序栈方式, 并共享一个存储区 $[0, \dots, \text{maxsize} - 1]$, 为了尽量利用空间, 减少溢出的可能, 可采用栈顶相向、迎面增长的存储方式。试设计 $S1$ 、 $S2$ 有关入栈和出栈的操作算法。

3.2.5 队列应用题(原书 P84)

(01) 若希望循环队列中的元素都能得到利用,则需设置一个标志域 tag ,并以 tag 的值为 0 或 1 来区分队首指针 $front$ 和队尾指针 $rear$ 相同时的队列状态是“空”还是“满”。试编写与此结构相应的入队和出队算法。

(02) Q 是一个队列, S 是一个空栈, 实现将队列中的元素逆置的算法。

(03) 利用两个栈 $S1$ 和 $S2$ 来模拟一个队列,已知栈的 4 个运算定义如下:

$Push(S, x);$	// 元素 x 入栈 s
$Pop(S, x);$	// S 出栈并将出栈的值赋给 x
$StackEmpty(S);$	// 判断栈是否为空
$StackOverflow(S);$	// 判断栈是否为满

如何利用栈的运算来实现该队列的 3 个运算 (形参由读者根据要求自己设计)?

$Enqueue;$	// 将元素 x 入队
$Dequeue;$	// 出队, 并将出队元素存储在 x 中
$QueueEmpty;$	// 判断队列是否为空

(04) 【2019 统考真题】请设计一个队列, 要求满足: ①初始时队列为空; ②入队时, 允许增加队列占用空间; ③出队后, 出队元素所占用的空间可重复使用, 即整个队列所占用的空间只增不减; ④入队操作和出队操作的时间复杂度始终保持为 $O(1)$ 。请回答:

- 1) 该队列是应选择链式存储结构, 还是应选择顺序存储结构?
- 2) 画出队列的初始状态, 并给出判断队空和队满的条件。
- 3) 画出第一个元素入队后的队列状态。
- 4) 给出入队操作和出队操作的基本过程。

3.3.6 栈和队列的应用应用题(原书 P96)

(01) 假设一个算术表达式中包含圆括号、方括号和花括号 3 种类型的括号,编写一个算法来判别表达式中的括号是否配对,以字符“\0”作为算术表达式的结束符。

4.2.4 串的模式匹配应用题(原书 P119)

(01) 在字符串模式匹配的 *KMP* 算法中,求模式的 *next* 数组值的定义如下:

$$next[j] = \begin{cases} 0, & j=1 \\ \max\{k \mid 1 < k < j \text{ 且 } 'p_1 \cdots p_{k-1}' = 'p_{j-k+1} \cdots p_{j-1}'\}, & \text{集合不为空} \\ 1, & \text{其他情况} \end{cases}$$

- 1) 当 $j=1$ 时, 为什么要取 $next[1]=0$?
- 2) 为什么要取 $\max\{k\}$, k 最大是多少?
- 3) 其他情况是什么情况, 为什么取 $next[j]=1$?

(02) 设有字符串 $S = 'aabaabaabaac'$, $P = 'aabaac'$ 。

- 1) 求出 P 的 *next* 数组。
- 2) 若 S 作主串, P 作模式串, 试给出 *KMP* 算法的匹配过程。

5.1.4 树的基本概念应用题(原书 P127)

(01) 含有 n 个结点的三叉树的最小高度是多少?

(02) 已知一棵度为 4 的树中, 度为 0,1,2,3 的结点数分别为 14,4,3,2, 求该树的结点总数 n 和度为 4 的结点个数,并给出推导过程。

(03) 已知一棵度为 m 的树中, 有 n_1 个度为 1 的结点, 有 n_2 个度为 2 的结点 …… 有 n_m 个度为 m 的结点, 问该树有多少个叶结点?

5.2.3 二叉树的概念应用题(原书 P135)

(01) 在一棵完全二叉树中, 含有 n_0 个叶结点, 当度为 1 的结点数为 1 时, 该树的高度是多少? 当度为 1 的结点数为 0 时, 该树的高度是多少?

(02) 一棵有 n 个结点的满二叉树有多少个分支结点和多少个叶结点? 该满二叉树的高度是多少?

(03) 已知完全二叉树的第 9 层有 240 个结点, 则整个完全二叉树有多少个结点? 有多少个叶结点?

(04) 一棵高度为 h 的满 m 叉树有如下性质: 根结点所在层次为第 1 层, 第 h 层上的结点都是叶结点, 其余各层上每个结点都有 m 棵非空子树, 若按层次自顶向下, 同一层自左向右, 顺序从 1 开始对全部结点进行编号, 试问:

- 1) 各层的结点个数是多少?
- 2) 编号为 i 的结点的双亲结点 (若存在) 的编号是多少?
- 3) 编号为 i 的结点的第 k 个孩子结点 (若存在) 的编号是多少?
- 4) 编号为 i 的结点有右兄弟的条件是什么? 其右兄弟结点的编号是多少?

(05) 已知一棵二叉树按顺序存储结构进行存储, 设计一个算法, 求编号分别为 i 和 j 的两个结点的最近的公共祖先结点的值。

(06) 【2016 统考真题】若一棵非空 $k(k \geq 2)$ 叉树 T 中的每个非叶结点都有 k 个孩子, 则称 T 为正则 k 叉树。请回答下列问题并给出推导过程。

1) 若 T 有 m 个非叶结点, 则 T 中的叶结点有多少个?

2) 若 T 的高度为 h (单结点的树 $h = 1$), 则 T 的结点数最多为多少个? 最少为多少个?

5.3.3 二叉树的遍历和线索二叉树应用题(原书 P152)

(01) 若某非空二叉树的先序序列和后序序列正好相反,则该二叉树的形态是什么?

(02) 若某非空二叉树的先序序列和后序序列正好相同,则该二叉树的形态是什么?

(03) 假设二叉树采用二叉链表存储结构, 设计一个非递归算法求二叉树的高度。

(04) 二叉树按二叉链表形式存储, 试编写一个判别给定二叉树是否是完全二叉树的算法。

(05) 假设二叉树采用二叉链表存储结构存储, 试设计一个算法, 计算一棵给定二叉树的所有双分支结点个数。

(06) 设树 B 是一棵采用链式结构存储的二叉树, 编写一个把树 B 中所有结点的左、右子树进行交换的函数。

(07) 假设二叉树采用二叉链存储结构存储, 设计一个算法, 求先序遍历序列中第 $k(1 \leq k \leq \text{二叉树中结点数})$ 个结点的值。

(08) 已知二叉树以二叉链表存储, 编写算法完成: 对于树中每个元素值为 x 的结点, 删除以它为根的子树, 并释放相应的空间。

(09) 在二叉树中查找值为 x 的结点, 试编写算法 (用 C 语言) 打印值为 x 的结点的所有祖先, 假设值为 x 的结点不多于一个。

(10) 设一棵二叉树的结点结构为 $(LLINK, INFO, RLINK)$, $ROOT$ 为指向该二叉树根结点的指针, p 和 q 分别为指向该二叉树中任意两个结点的指针, 试编写算法 $ANCESTOR(ROOT, p, q, r)$, 找到 p 和 q 的最近公共祖先结点 r 。

(11) 假设二叉树采用二叉链表存储结构, 设计一个算法, 求非空二叉树 b 的宽度 (具有结点数最多的那一层的结点个数)。

(12) 设有一棵满二叉树 (所有结点值均不同), 已知其先序序列为 pre , 设计一个算法求其后序序列 $post$ 。

(13) 设计一个算法将二叉树的叶结点按从左到右的顺序连成一个单链表, 表头指针为 $head$ 。二叉树按二叉链表方式存储, 链接时用叶结点的右指针域来存放单链表指针。

(14) 试设计判断两棵二叉树是否相似的算法。所谓二叉树 T_1 和 T_2 相似, 指的是 T_1 和 T_2 都是空的二叉树或都只有一个根结点; 或者 T_1 的左子树和 T_2 的左子树是相似的, 且 T_1 的右子树和 T_2 的右子树是相似的。

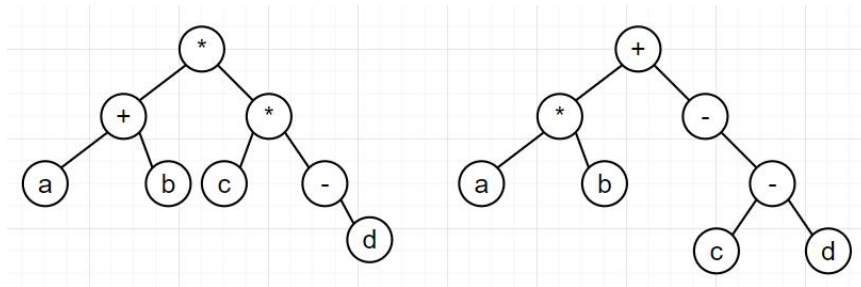
(15) 【2014 统考真题】二叉树的带权路径长度 (WPL) 是二叉树中所有叶结点的带权路径长度之和。给定一棵二叉树 T , 采用二叉链表存储, 结点结构为

left	weight	right
------	--------	-------

其中叶结点的 $weight$ 域保存该结点的非负权值。设 $root$ 为指向 T 的根结点的指针, 请设计求 T 的 WPL 的算法, 要求:

- 1) 给出算法的基本设计思想。
- 2) 使用 C 或 $C++$ 语言, 给出二叉树结点的数据类型定义。
- 3) 根据设计思想, 采用 C 或 $C++$ 语言描述算法, 关键之处给出注释。

(16) 【2017 统考真题】请设计一个算法, 将给定的表达式树 (二叉树) 转换为等价的中缀表达式 (通过括号反映操作符的计算次序) 并输出。例如, 当下列两棵表达式树作为算法的输入时:



输出的等价中缀表达式分别为 $(a+b) * (c*(-d))$ 和 $(a*b) + (-(c-d))$ 。

二叉树结点定义如下:

```
typedef struct node{
    char data[10];           //存储操作数或操作符
    struct node *left, *right;
}BTree;
```

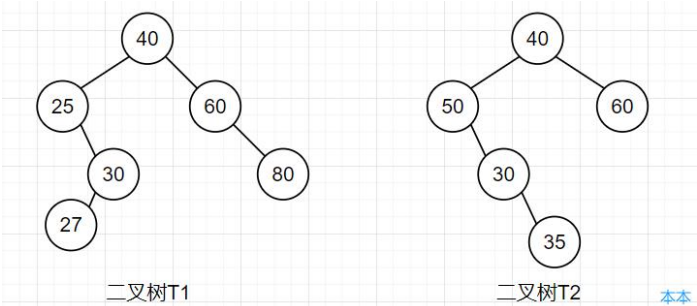
要求:

- 1) 给出算法的基本设计思想。
- 2) 根据设计思想, 采用 C 或 C++ 语言描述算法, 关键之处给出注释。

(17) 【2022 统考真题】已知非空二叉树 T 的结点值均为正整数, 采用顺序存储方式保存, 数据结构定义如下:

```
typedef struct {  
    int SqBiTNode[MAX_SIZE];  
    int ElemNum;  
}SqBiTree;  
//MAX_SIZE为已定义常量  
//保存二叉树节点值的数组  
//实际占用的数组元素个数
```

T 中不存在的结点在数组 $SqBiTNode$ 中用 -1 表示。例如, 对于下图所示的两棵非空二叉树 T_1 和 T_2 ,



T_1 的存储结果如下:

T1.SqBiTNode	40	25	60	-1	30	-1	80	-1	-1	27		
T1.ElemNum=10												

T_2 的存储结果如下:

T2.SqBiTNode	40	50	60	-1	30	-1	-1	-1	-1	-1	35	
T2.ElemNum=11												

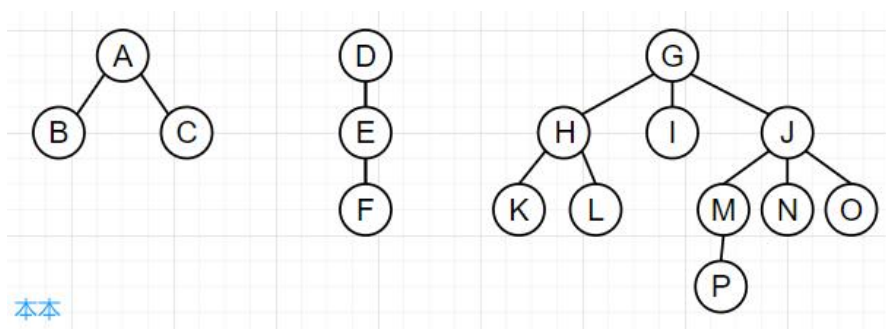
请设计一个尽可能高效的算法, 判定一棵采用这种方式存储的二叉树是否为二叉搜索树,若是,则返回 $true$,否则,返回 $false$ 。要求:

- 1) 给出算法的基本设计思想。
- 2) 根据设计思想, 采用 C 或 $C++$ 语言描述算法, 关键之处给出注释。

5.4.4 树、森林应用题(原书 P177)

(01) 给定一棵树的先根遍历序列和后根遍历序列,能否唯一确定一棵树? 若能,请举例说明; 若不能,请给出反例。

(02) 将下面一个由 3 棵树组成的森林转换为二叉树。



(03) 已知某二叉树的先序序列和中序序列分别为 $ABDEHCFIMGJKL$ 和 $DBHEAIMFCGKLJ$,请画出这棵二叉树,并画出二叉树对应的森林。

(04) 编程求以孩子兄弟表示法存储的森林的叶结点数。

(05) 以孩子兄弟链表为存储结构, 请设计递归算法求树的深度。

5.5.3 树与二叉树的应用应用题(原书 P190)

(01) 设给定权集 $w = \{5, 7, 2, 3, 6, 8, 9\}$, 试构造关于 w 的一棵哈夫曼树, 并求其加权路径长度 WPL 。

(02) 【2012 统考真题】设有 6 个有序表 A, B, C, D, E, F , 分别含有 10, 35, 40, 50, 60 和 200 个数据元素, 各表中的元素按升序排列。要求通过 5 次两两合并, 将 6 个表最终合并为 1 个升序表, 并使最坏情况下比较的总次数达到最小。请回答下列问题:

- 1) 给出完整的合并过程, 并求出最坏情况下比较的总次数。
- 2) 根据你的合并过程, 描述 $n(n \geq 2)$ 个不等长升序表的合并策略, 并说明理由。

(03) 【2020 统考真题】若任意一个字符的编码都不是其他字符编码的前缀,则称这种编码具有前缀特性。现有某字符集 (字符个数 ≥ 2) 的不等长编码, 每个字符的编码均为二进制的 0、1 序列, 最长为 L 位, 且具有前缀特性。请回答下列问题:

- 1) 哪种数据结构适宜保存上述具有前缀特性的不等长编码?
- 2) 基于你所设计的数据结构, 简述从 0/1 串到字符串的译码过程。
- 3) 简述判定某字符集的不等长编码是否具有前缀特性的过程。

6.1.2 图的基本概念应用题(原书 P203)

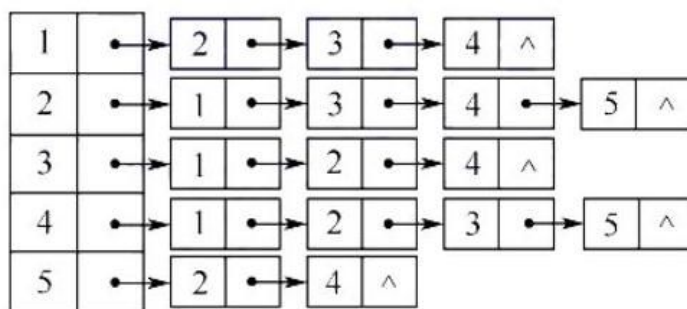
(01) 图 G 是一个非连通无向图, 共有 28 条边, 该图至少有多少个顶点?

6.2.6 图的存储及基本操作应用题(原书 P213)

(01) 已知带权有向图 G 的邻接矩阵如下图所示, 请画出该带权有向图 G 。

0	15	2	12	∞	∞	∞
∞	0	∞	∞	6	∞	∞
∞	∞	0	∞	8	4	∞
∞	∞	∞	0	∞	∞	3
∞	∞	∞	∞	0	∞	9
∞	∞	∞	5	∞	0	10
∞	4	∞	∞	∞	∞	0

(02) 设图 $G = (V, E)$ 以邻接表存储, 如下图所示。画出其邻接矩阵存储及图 G 。



(03) 对 n 个顶点的无向图和有向图, 分别采用邻接矩阵和邻接表表示时, 试问:

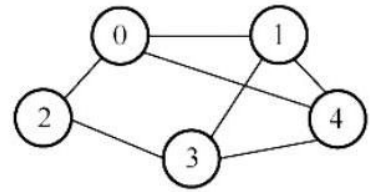
- 1) 如何判别图中有多少条边?
- 2) 如何判别任意两个顶点 i 和 j 是否有边相连?
- 3) 任意一个顶点的度是多少?

(04) 如何对无环有向图中的顶点重新编号,使得该图的邻接矩阵中所有的 1 都集中到对角线以上?

(05) 写出从图的邻接表表示转换成邻接矩阵表示的算法。

(06) 【2015 统考真题】已知含有 5 个顶点的图 G 如下图所示。

请回答下列问题：



- 1) 写出图 G 的邻接矩阵 A (行、列下标从 0 开始)。
- 2) 求 A^2 , 矩阵 A^2 中位于 0 行 3 列元素值的含义是什么?
- 3) 若已知具有 $n(n \geq 2)$ 个顶点的图的邻接矩阵为 B , 则 $B^m (2 \leq m \leq n)$ 中非零元素的含义是什么?

(07) 【2021 统考真题】已知无向连通图 G 由顶点集 V 和边集 E 组成, $|E| > 0$, 当 G 中度为奇数的顶点个数为不大于 2 的偶数时, G 存在包含所有边且长度为 $|E|$ 的路径 (称为 EL 路径)。设图 G 采用邻接矩阵存储, 类型定义如下:

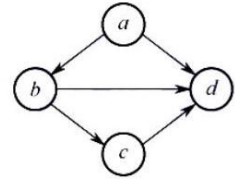
```
typedef struct{                                //图的定义
    int numVertices, numEdges;                //图中实际的顶点数和边数
    char VerticesList[MAXV];                  //顶点表.MAXV为已定义常量
    int Edge[MAXV][MAXV];                     //邻接矩阵
}MGraph;
```

请设计算法 $int IsExistEL(MGraph G)$, 判断 G 是否存在 EL 路径, 若存在, 则返回 1, 否则返回 0。要求:

- 1) 给出算法的基本设计思想。
- 2) 根据设计思想, 采用 C 或 $C++$ 语言描述算法, 关键之处给出注释。
- 3) 说明你所设计算法的时间复杂度和空间复杂度。

(08) 【2023 统考真题】已知有向图 G 采用邻接矩阵存储, 类型定义如下:

```
typedef struct{           //图的定义
    int numVertices, numEdges; //图中实际的顶点数和边数
    char VerticesList[MAXV]; //顶点表.MAXV为已定义常量
    int Edge[MAXV][MAXV]; //邻接矩阵
}MGraph;
```



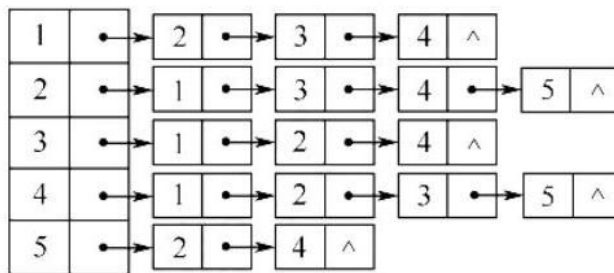
将图中出度大于入度的顶点称为 K 顶点。例如, 在上图中, 顶点 a 和 b 为 K 顶点。

请设计算法 `int printVertices(MGraph G)`, 对给定的任意非空有向图 G , 输出 G 中所有的 K 顶点, 并返回 K 顶点的个数。要求:

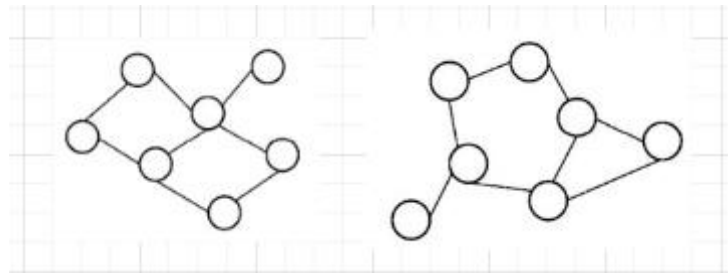
- 1) 给出算法的基本设计思想。
- 2) 根据设计思想, 采用 C 或 $C++$ 语言描述算法, 关键之处给出注释。

6.3.4 图的遍历应用题(原书 P225)

(01) 图 $G = (V, E)$ 以邻接表存储, 如下图所示, 试画出图 G 的深度优先生成树和广度优先生成树 (假设从结点 1 开始遍历)。



(02) 给定一个连通无向图, 采用邻接表存储, 将图的所有顶点分别染成红色或蓝色, 若存在一种染色方法使图中每条边的两个顶点的颜色都不相同, 则称这个图能被二分。



- 1) 判断上面两个无向图是否能被二分, 若能被二分, 则请标出每个顶点的颜色。
- 2) 请设计一种算法用来判断图是否能被二分, 仅用语言描述算法的思想即可。
- 3) 给出你设计的算法的时间复杂度和空间复杂度。

(03) 试设计一个算法, 判断一个无向图 G 是否为一棵树。若是一棵树, 则算法返回 *true*, 否则返回 *false*。

(04) 分别采用基于深度优先遍历和广度优先遍历算法判别以邻接表方式存储的有向图中是否存在由顶点 v_i 到顶点 v_j 的路径 ($i \neq j$)。注意, 算法中涉及的图的基本操作必须在此存储结构上实现。

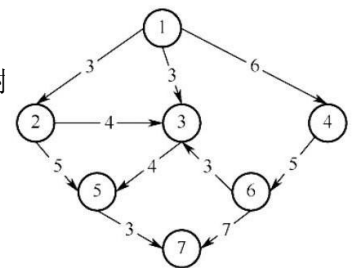
(05) 假设图用邻接表表示, 设计一个算法, 输出从顶点 V_i 到顶点 V_j 的所有简单路径。

6.4.6 图的应用应用题(原书 P248)

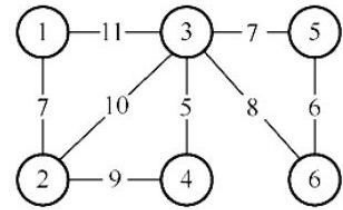
(01) 下面是一种称为“破圈法”的求解最小生成树的方法: 所谓“破圈法”, 是指“任取一圈, 去掉圈上权最大的边”, 反复执行这一步骤, 直到没有圈为止。试判断这种方法是否正确。若正确, 说明理由; 若不正确, 举出反例 (注: 圈就是回路)。

(02) 已知有向图如下图所示。

- 1) 写出该图的邻接矩阵表示并据此给出从顶点 1 出发的深度优先遍历序列。
- 2) 求该有向图的强连通分量的数目。
- 3) 给出该图的任意两个拓扑序列。
- 4) 若将该图视为无向图, 分别用 *Prim* 算法和 *Kruskal* 算法求最小生成树

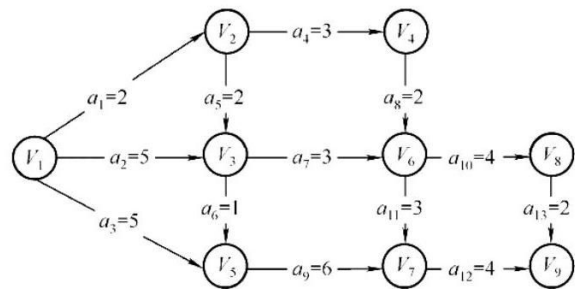


(03) 对下图所示的无向图,按照 *Dijkstra* 算法,写出从顶点 1 到其他各个顶点的最短路径和最短路径长度 (顺序不能颠倒)。



(04) 下图所示为一个用 *AOE* 网表示的工程。

- 1) 画出此图的邻接表表示。
- 2) 完成此工程至少需要多少时间?
- 3) 指出关键路径。
- 4) 哪些活动加速可以缩短完成工程所需的时间?



(05) 下表给出了某工程各工序之间的优先关系和各工序所需的时间 (其中“—”表示无先驱工序), 请完成以下各题:

- 1) 画出相应的 *AOE* 网。
- 2) 列出各事件的最早发生时间和最迟发生时间。
- 3) 求出关键路径并指明完成该工程所需的最短时间。

工序代号	A	B	C	D	E	F	G	H
所需时间	3	2	2	3	4	3	2	1
先驱工序	—	—	A	A	B	A	C、E	D

(06) 一连通无向图, 边非负权值, 问用 *Dijkstra* 最短路径算法能否给出一棵生成树, 该树是否一定是最小生成树? 说明理由。

(07) 试编写利用 *DFS* 实现有向无环图拓扑排序的算法。

(08) 【2009 统考真题】带权图 (权值非负, 表示边连接的两顶点间的距离) 的最短路径问题是找出从初始顶点到目标顶点之间的一条最短路径。假设从初始顶点到目标顶点之间存在路径, 现有一种解决该问题的方法:

- ① 设最短路径初始时仅包含初始顶点, 令当前顶点 u 为初始顶点。
- ② 选择离 u 最近且尚未在最短路径中的一个顶点 v , 加入最短路径, 修改当前顶点 $u = v$ 。
- ③ 重复步骤②, 直到 u 是目标顶点时为止。

请问上述方法能否求得最短路径? 若该方法可行, 请证明; 否则, 请举例说明。

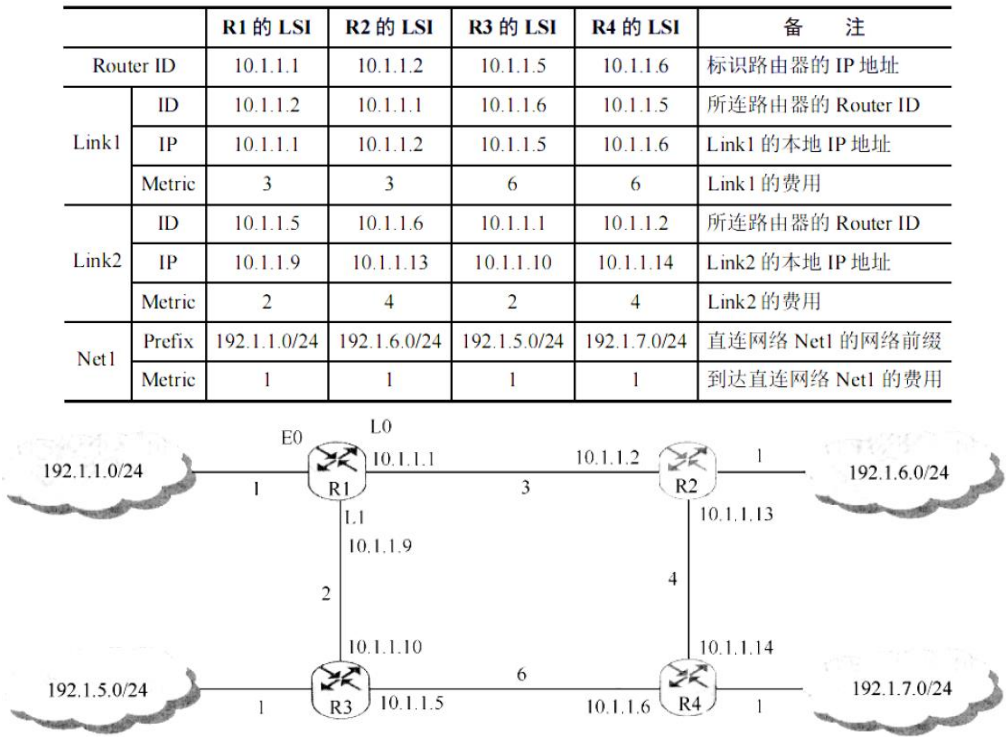
(09) 【2011 统考真题】已知有 6 个顶点 (顶点编号为 0~5) 的有向带权图 G , 其邻接矩阵 A 为上三角矩阵, 按行为主序 (行优先) 保存在如下的一维数组中。

4	6	∞	∞	∞	5	∞	∞	∞	4	3	∞	∞	3	3
---	---	----------	----------	----------	---	----------	----------	----------	---	---	----------	----------	---	---

要求:

- 1) 写出图 G 的邻接矩阵 A 。
- 2) 画出有向带权图 G 。
- 3) 求图 G 的关键路径, 并计算该关键路径的长度。

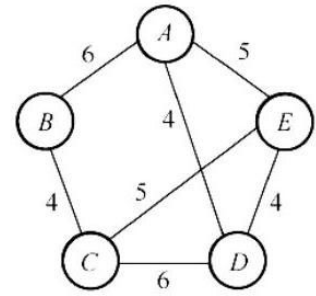
(10) 【2014 统考真题】某网络中的路由器运行 *OSPF* 路由协议,下表是路由器 *R1* 维护的主要链路状态信息 (*LSI*),*R1* 构造的网络拓扑图 (见下图) 是根据题下表及 *R1* 的接口名构造出来的网络拓扑。



- 请回答下列问题。
- 1) 本题中的网络可抽象为数据结构中的哪种逻辑结构？
 - 2) 针对表中的内容, 设计合理的链式存储结构, 以保存表中的链路状态信息 (*LSI*)。要求给出链式存储结构的数据类型定义,并画出对应表的链式存储结构示意图 (示意图中可仅以 *ID* 标识结点)。
 - 3) 按照 *Dijkstra* 算法的策略, 依次给出 *R1* 到达子网 192.1.x.x 的最短路径及费用。

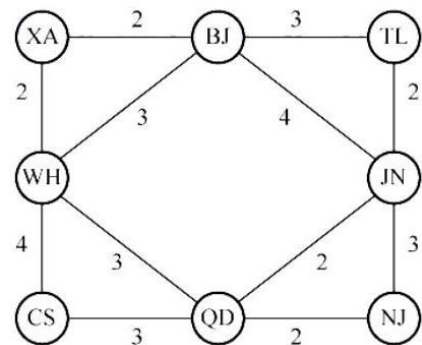
(11) 【2017 统考真题】使用 Prim 算法求带权连通图的最小 (代价) 生成树 (MST)。请回答下列问题:

- 1) 对下列图 G , 从顶点 A 开始求 G 的 MST, 依次给出按算法选出的边。
- 2) 图 G 的 MST 是唯一的吗?
- 3) 对任意的带权连通图, 满足什么条件时, 其 MST 是唯一的?



(12) 【2018 统考真题】拟建设一个光通信骨干网络连通 BJ 、 CS 、 XA 、 QD 、 JN 、 NJ 、 TL 和 WH 等 8 个城市, 下图中无向边上的权值表示两个城市之间备选光缆的铺设费用。请回答下列问题:

- 1) 仅从铺设费用角度出发, 给出所有可能的最经济的光缆铺设方案 (用带权图表示), 并计算相应方案的总费用。
- 2) 该图可采用图的哪种存储结构? 给出求解问题 1) 所用的算法名称。
- 3) 假设每个城市采用一个路由器按 1) 中得到的最经济方案组网, 主机 $H1$ 直接连接 TL 的路由器, 主机 $H2$ 直接连接 BJ 的路由器。若 $H1$ 向 $H2$ 发送一个 $TTL = 5$ 的 IP 分组, 则 $H2$ 是否可以收到该 IP 分组?



(13) 【2024 统考真题】2023 年 10 月 26 日, 神舟十七号载人飞船发射取得圆满成功, 再次彰显了中国航天事业的辉煌成就。载人航天工程是包含众多子工程的复杂系统工程, 为了保证工程的有序开展, 需要明确各子工程的前导子工程, 以协调各子工程的实施。该问题可以简化、抽象为有向图的拓扑序列问题。已知有向图 G 采用邻接矩阵存储, 类型定义如下。

```
typedef struct                                //图的类型定义
{
    int numVertices, numEdges; //图的顶点数和有向边数
    char VerticesList[MAXV]; //顶点表, MAXV为已定义常量
    int Edge[MAXV][MAXV]; //邻接矩阵
}MGraph;
```

请设计算法: `int uniquely(MGraph G)`, 判定 G 是否存在唯一的拓扑序列, 若是, 则返回 1, 否则返回 0。要求如下。

- 1) 给出算法的基本设计思想。
- 2) 根据设计思想, 采用 C 或 C++ 语言描述算法, 关键之处给出注释。

7.2.4 顺序查找和折半查找应用题(原书 P276)

(01) 若对有 n 个元素的有序顺序表和无序顺序表进行顺序查找, 试就下列三种情况分别讨论两者在相等查找概率时的平均查找长度是否相同。

- 1) 查找失败。
- 2) 查找成功, 且表中只有一个关键字等于给定值 k 的元素。
- 3) 查找成功, 且表中有若干关键字等于给定值 k 的元素, 要求一次查找能找出所有元素。

(02) 有序顺序表中的元素依次为 017, 094, 154, 170, 275, 503, 509, 512, 553, 612, 677, 765, 897, 908。

- 1) 试画出对其进行折半查找的判定树。
- 2) 若查找 275 或 684 的元素, 将依次与表中的哪些元素比较?
- 3) 计算查找成功的平均查找长度和查找不成功的平均查找长度。

(03) 已知一个有序顺序表 $A[0 \dots 8n-1]$ 的表长为 $8n$, 并且表中没有关键字相同的数据元素。假设按下述方法查找一个关键字值等于给定值 X 的数据元素: 首先在 $A[7], A[15], A[23], \dots, A[8k-1], \dots, A[8n-1]$ 中进行顺序查找, 若查找成功, 则算法报告成功位置并返回; 若不成功, 则当 $A[8k-1] < X < A[8 \times (k+1)-1]$ 时, 可确定一个缩小的查找范围 $A[8k] \sim A[8 \times (k+1)-2]$, 然后可在这个范围内执行折半查找。特殊情况: 若 $X > A[8n-1]$ 的关键字, 则查找失败。

- 1) 画出描述上述查找过程的判定树。
- 2) 计算相等查找概率下查找成功的平均查找长度。

(04) 写出折半查找的递归算法。初始调用时, low 为 1, $high$ 为 $ST.length$ 。

(05) 线性表中各结点的检索概率不等时,可用如下策略提高顺序检索的效率: 若找到指定的结点,则将该结点和其前驱结点 (若存在) 交换,使得经常被检索的结点尽量位于表的前端。试设计在顺序结构和链式结构的线性表上实现上述策略的顺序检索算法。

(06) 已知一个 n 阶矩阵 A 和一个目标值 k 。该矩阵无重复元素, 每行从左到右升序排列, 每列从上到下升序排列。请设计一个在时间上尽可能高效的算法, 判断矩阵中是否存在目

标值 k 。例如, 矩阵为 $\begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$, 目标值为 8, 判断存在。要求:

- 1) 给出算法的基本设计思想。
- 2) 根据设计思想, 采用 C 或 C++ 语言描述算法, 关键之处给出注释。
- 3) 说明你的算法的时间复杂度和空间复杂度。

(07) 【2013 统考真题】设包含 4 个数据元素的集合 $S = \{'do', 'for', 'repeat', 'while'\}$, 各元素的查找概率依次为 $p_1 = 0.35, p_2 = 0.15, p_3 = 0.15, p_4 = 0.35$ 。将 S 保存在一个长度为 4 的顺序表中, 采用折半查找法, 查找成功时的平均查找长度为 2.2。

1) 若采用顺序存储结构保存 S , 且要求平均查找长度更短, 则元素应如何排列? 应使用何种查找方法? 查找成功时的平均查找长度是多少?

2) 若采用链式存储结构保存 S , 且要求平均查找长度更短, 则元素应如何排列? 应使用何种查找方法? 查找成功时的平均查找长度是多少?

7.3.4 树形查找应用题(原书 P300)

(01) 一棵二叉排序树按先序遍历得到的序列为 (50,38,30,45,40,48,70,60,75,80), 试画出该二叉排序树, 并求出等概率下查找成功和查找失败的平均查找长度。

(02) 按照序列 (40,72,38,35,67,51,90,8,55,21) 建立一棵二叉排序树, 画出该树, 并求出在等概率的情况下, 查找成功的平均查找长度。

(03) 依次把结点 (34,23,15,98,115,28,107) 插入初始状态为空的平衡二叉排序树, 使得在每次插入后保持该树仍然是平衡二叉树。请依次画出每次插入后所形成的平衡二叉排序树。

(04) 给定一个关键字集合 $\{25, 18, 34, 9, 14, 27, 42, 51, 38\}$, 假定查找各关键字的概率相同, 请画出其最佳二叉排序树。

(05) 试编写一个算法, 判断给定的二叉树是否是二叉排序树。

(06) 设计一个算法,求出指定结点在给定二叉排序树中的层次。

(07) 利用二叉树遍历的思想编写一个判断二叉树是否是平衡二叉树的算法。

(08) 设计一个算法,求出给定二叉排序树中最小和最大的关键字。

(09) 设计一个算法,从大到小输出二叉排序树中所有值不小于 k 的关键字。

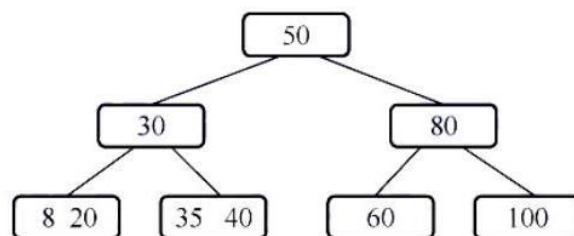
(10) 编写一个递归算法, 在一棵有 n 个结点的、随机建立起来的二叉排序树上查找第 k ($1 \leq k \leq n$) 小的元素, 并返回指向该结点的指针。要求算法的平均时间复杂度为 $O(\log_2 n)$ 。二叉排序树的每个结点中除 *data*、*lchild*、*rchild* 等数据成员外, 增加一个 *count* 成员, 保存以该结点为根的子树上的结点个数。

7.4.3B 树和 B+树应用题(原书 P317)

(01) 给定一组关键字 {20,30,50,52,60,68,70}, 给出创建一棵 3 阶 B 树的过程。

(02) 对如下图所示的 3 阶 B 树,依次执行下列操作,画出各步操作的结果。

- 1) 插入 90 2) 插入 25 3) 插入 45 4) 删除 60 5) 删除 80



(03) 利用 B 树做文件索引时,若假设磁盘页块的大小是 $4000B$ (实际应是 2 的次幂,此处是为了计算方便),指示磁盘地址的指针需要 $5B$ 。现有 20000000 个记录构成的文件,每个记录为 $200B$, 其中包括关键字 $5B$ 。

试问在这个采用 B 树作索引的文件中, B 树的阶数应为多少? 假定文件数据部分未按关键字有序排列,则索引部分需要占用多少磁盘页块?

7.5.5 散列(Hash)表应用题(原书 P329)

(01) 若要在散列表中删除一个记录,应如何操作? 为什么? 按照处理冲突的方法为开放地址法和拉链法分别说明。

(02) 假定把关键字 key 散列到有 n 个表项 (从 0 到 $n-1$ 编址) 的散列表中。对于下面的每个函数 $H(key)$ (key 为整数), 这些函数能够当作散列函数吗? 若能, 它是一个好的散列函数吗? 说明理由。设函数 $random(n)$ 返回一个 0 到 $n-1$ 之间的随机整数 (包括 0 与 $n-1$ 在内)。

1) $H(key) = key/n$ 。

2) $H(key) = 1$ 。

3) $H(key) = (key + random(n)) \% n$ 。

4) $H(key) = key \% p(n)$; 其中 $p(n)$ 是不大于 n 的最大素数。

(03) 使用散列函数 $H(key) = key \% 11$, 把一个整数值转换成散列表下标, 散列表的长度为 11, 现在要把数据 $\{1, 13, 12, 34, 38, 33, 27, 22\}$ 依次插入散列表。

1) 使用线性探测法来构造散列表。

2) 使用链地址法构造散列表。

试针对这两种情况, 分别确定查找成功所需的平均查找长度, 及查找不成功所需的平均查找长度。

(04) 已知一组关键字为 $\{26, 36, 41, 38, 44, 15, 68, 12, 6, 51, 25\}$, 用链地址法解决冲突, 假设装填因子 $\alpha = 0.73$, 散列函数的形式为 $H(key) = key \% P$, P 为不大于表长的最大素数, 请回答以下问题:

- 1) 构造出散列函数。
- 2) 分别计算出等概率情况下查找成功和查找失败的平均查找长度 (查找失败的计算中只将与关键字的比较次数计算在内即可)。

(05) 设散列表为 $HT[0 \cdots 12]$, 即表的大小为 $m = 13$ 。现采用双散列法解决冲突, 散列函数和再散列函数分别为:

$$H_0(key) = key \% 13 \quad \text{注: \% 是取模运算 (= mod)}$$

$$H_i = (H_{i-1} + REV(key + 1) \% 11 + 1) \% 13; \quad i = 1, 2, 3, \dots, m - 1$$

其中, 函数 $REV(x)$ 表示颠倒十进制数 x 的各位, 如 $REV(37) = 73, REV(7) = 7$ 等。若插入的关键码序列为 $(2, 8, 31, 20, 19, 18, 53, 27)$, 请回答:

- 1) 画出插入这 8 个关键码后的散列表。
- 2) 计算查找成功的平均查找长度 ASL 。

(06) 【2010 统考真题】将关键字序列 (7,8,30,11,18,9,14) 散列存储到散列表中。散列表的存储空间是一个下标从 0 开始的一维数组, 散列函数为 $H(key) = (key \times 3) \bmod 7$, 处理冲突采用线性探测再散列法, 要求装填 (载) 因子为 0.7。

- 1) 请画出所构造的散列表。
- 2) 分别计算等概率情况下, 查找成功和查找不成功的平均查找长度。

(07) 【2024 统考真题】将关键字序列 20,3,11,18,9,14,7 依次存储到初始为空、长度为 11 的散列表 HT 中, 散列函数 $H(key) = (key \times 3) \% 11$ 。 $H(key)$ 计算出的初始散列地址为 H_0 , 发生冲突时探查地址序列是 H_1, H_2, H_3, \dots , 其中, $H_k = (H_0 + k^2) \% 11, k = 1, 2, 3, \dots$ 。请回答下列问题:

- 1) 画出所构造的 HT , 并计算 HT 的装填因子。
- 2) 给出在 HT 中查找关键字 14 的关键字比较序列。
- 3) 在 HT 中查找关键字 8, 确认查找失败时的散列地址是多少?

8.2.4 插入排序应用题(原书 P344)

(01) 给出关键字序列 $\{4, 5, 1, 2, 6, 3\}$ 的直接插入排序过程。

(02) 给出关键字序列 $\{50, 26, 38, 80, 70, 90, 8, 30, 40, 20\}$ 的希尔排序过程 (取增量序列为 $d = \{5, 3, 1\}$, 排序结果为从小到大排列)。

8.3.3 交换排序应用题(原书 P352)

(01) 已知线性表按顺序存储, 且每个元素都是不相同的整数型元素, 设计把所有奇数移动到所有偶数前边的算法 (要求时间最短, 辅助空间最小)。

(02) 试编写一个算法, 使之能够在数组 $L[1 \cdots n]$ 中找出第 k 小的元素 (从小到大排序后处于第 k 个位置的元素)。

(03) 荷兰国旗问题: 设有一个仅由红、白、蓝三种颜色的条块组成的条块序列, 存储在一个顺序表中, 请编写一个时间复杂度为 $O(n)$ 的算法, 使得这些条块按红、白、蓝的顺序排好, 即排成荷兰国旗图案。请完成以下两个的算法实现:

```
typedef enum{RED, WHITE, BLUE}color;      // 设置枚举数组
```

```
void Flag_Arrange(colora[], intn) {...}
```

(04) 【2016 统考真题】已知由 $n(n \geq 2)$ 个正整数构成的集合 $A = \{a_k \mid 0 \leq k < n\}$, 将其划分为两个不相交的子集 A_1 和 A_2 , 元素个数分别是 n_1 和 n_2 , A_1 和 A_2 中的元素之和分别为 S_1 和 S_2 。设计一个尽可能高效的划分算法, 满足 $|n_1 - n_2|$ 最小且 $|S_1 - S_2|$ 最大。要求:

- 1) 给出算法的基本设计思想。
- 2) 根据设计思想, 采用 C 或 C++ 语言描述算法, 关键之处给出注释。
- 3) 说明所设计算法的平均时间复杂度和空间复杂度。

8.4.3 选择排序应用题(原书 P363)

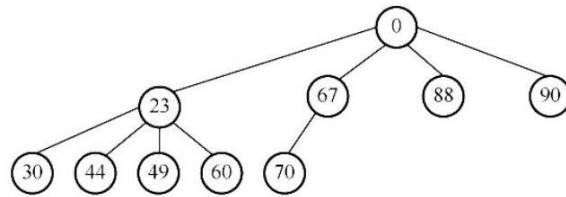
(01) 指出堆和二叉排序树的区别?

(02) 画出一棵二叉树,使得它既满足大根堆的要求又满足二叉排序树的要求。

(03) 若只想得到一个序列中第 $k(k \geq 5)$ 个最小元素之前的部分的排序序列, 则最好采用什么排序算法?

(04) 通常使用的堆也称二叉堆, 因为它用完全二叉树来实现的, 树中结点最多只有两个孩子。同理可以有 m 叉堆, 即用完全 m 叉树来实现的堆。

1) 下图是一个 m 叉小根堆, 问 m 值是多少? 向这个堆插入一个元素 65 后, 堆中的元素如何变化? 再删除堆顶元素呢? 请画出变化后的树形。



2) 从 0 开始对完全 4 叉树中的结点从左到右、从上到下进行编号。若给定一个结点 k , 其父结点的编号是多少?(若存在), 其第 i ($i=1,2,3,4$) 个孩子的编号是多少?

3) 在 m 叉堆中进行插入和删除操作的时间复杂度是多少?

(05) 编写一个算法,在基于单链表表示的待排序关键字序列上进行简单选择排序。

(06) 试设计一个算法,判断一个数据序列是否构成一个小根堆。

(07) 优先队列 (*PriorityQueue*) 是一种数据结构,它类似于普通队列,但每个元素都有一个优先级。元素在入队时会根据其优先级来排序,而不按照先入先出的顺序来排序。每次从优先队列中出队时,出队的是优先级最高的元素,而不是最早进入队列的元素。

队列中的元素的数据结构的定义如下:

```
typedef struct{
    int value;        //元素的值
    int priority;     //元素的优先级, priority越大, 优先级越高
}PriorityQueueElement;
```

请设计一个优先队列, 要求满足: ①初始时队列为空; ②入队时, 不允许增加队列的占用空间; ③出队后, 出队元素所占用的空间可重复使用, 即整个队列所占用的空间不变; ④入队操作和出队操作的时间复杂度始终保持为 $O(\log_2 n)$ 。请回答:

- 1) 该队列是应选择链式存储结构, 还是选择顺序存储结构?
- 2) 给出优先队列的数据结构的定义。
- 3) 用伪代码给出入队操作和出队操作的基本过程 (关键之处可用文字描述)。

(08) 【2022 统考真题】现有 $n(n > 100000)$ 个数保存在一维数组 M 中, 需要查找 M 中最小的 10 个数。请回答下列问题。

- 1) 设计一个完成上述查找任务的算法, 要求平均情况下的比较次数尽可能少, 简述其算法思想 (不需要编程实现)。
- 2) 说明你所设计的算法平均情况下的时间复杂度和空间复杂度。

8.5.4 归并排序、基数排序与计数排序应用题(原书 P376)

(01) 已知序列 {503, 87, 512, 61, 908, 170, 897, 275, 653, 462}, 采用非递归的二路归并排序算法对该序列做升序排序时需要几趟排序? 给出每一趟的结果。

(02) 设待排序的关键字序列为 {12, 2, 16, 30, 28, 10, 16*, 20, 6, 18}, 试写出使用最低位优先 (*LSD*) 基数排序算法每趟排序后的结果。

(03) 【2021 统考真题】已知某排序算法如下：

```
void cmpCountSort(int a[], int b[], int n){
    int i, j, *count;
    count = (int *)malloc(sizeof(int)*n);
                                                    //C++语言: count=new int[n];

    for(i=0;i<n;i++)    count[i]=0;
    for(i=0;i<n-1;i++)
        for(j=i+1;j<n;j++)
            if(a[i] < a[j])    count[j]++;
            else                count[i]++;
    for(i=0; i<n; i++)    b[count[i]]=a[i];
    free(count);          //C++语言: delete count;
}
```

请回答下列问题。

- 1) 若有 $\text{inta[]} = \{25, -10, 25, 10, 11, 19\}$, $b[6]$; 则调用 $\text{cmpCountSort}(a, b, 6)$ 后数组 b 中的内容是什么?
- 2) 若 a 中含有 n 个元素, 则算法执行过程中, 元素之间的比较次数是多少?
- 3) 该算法是稳定的吗? 若是, 阐述理由; 否则, 修改为稳定排序算法。

8.6.3 各种内部排序算法的比较及应用应用题(原书 P384)

(01) 设关键字序列为 $\{3, 7, 6, 9, 7, 1, 4, 5, 20\}$, 对其进行排序的最小交换次数是多少?

(02) 设顺序表用数组 $A[]$ 表示, 表中元素存储在数组下标 $1 \sim m + n$ 的范围内, 前 m 个元素递增有序, 后 n 个元素递增有序, 设计一个算法, 使得整个顺序表有序。

- 1) 给出算法的基本设计思想。
- 2) 根据设计思想, 采用 $C/C++$ 描述算法, 关键之处给出注释。
- 3) 说明你所设计算法的时间复杂度与空间复杂度。

(03) 设有一个数组中存放了一个无序的关键序列 K_1, K_2, \dots, K_n 。现要求将 K_n 放在将元素排序后的正确位置上, 试编写实现该功能的算法, 要求比较关键字的次数不超过 n 。

8.7.6 外部排序应用题(原书 P393)

(01) 若某个文件经内部排序得到 80 个初始归并段,试问:

- 1) 若使用多路平衡归并执行 3 趟完成排序, 则应取得的归并路数至少应为多少?
- 2) 若操作系统要求一个程序同时可用的输入 / 输出文件的总数不超过 15 个, 则按多路归并至少需要几趟可以完成排序? 若限定这个趟数, 可取的最低路数是多少?

(02) 假设文件有 4500 个记录, 在磁盘上每个块可放 75 个记录。计算机中用于排序的内存区可容纳 450 个记录。试问:

- 1) 可以建立多少个初始归并段? 每个初始归并段有多少记录? 存放于多少个块中?
- 2) 应采用几路归并? 请写出归并过程及每趟需要读 / 写磁盘的块数。

(03) 设初始归并段为 $(10,15,31), (9,20), (22,34,37), (6,15,42), (12,37), (84,95)$ 。试利用败者树进行 m 路归并, 手工执行选择最小的 5 个关键字的过程。

(04) 给出 12 个初始归并段, 其长度分别为 30, 44, 8, 6, 3, 20, 60, 18, 9, 62, 68, 85。现要做 4 路外归并排序, 试画出表示归并过程的最佳归并树, 并计算该归并树的带权路径长度 WPL 。

(05) 【2023 统考真题】对含有 $n(n>0)$ 个记录的文件进行外部排序, 采用置换 - 选择排序生成初始归并段时需要使用一个工作区, 工作区中能保存 m 个记录。请回答:

1) 若文件中含有 19 个记录, 其关键字依次是 51, 94, 37, 92, 14, 63, 15, 99, 48, 56, 23, 60, 31, 17, 43, 8, 90, 166, 100, 则当 $m=4$ 时, 可生成几个初始归并段? 各是什么?

2) 对任意的 $m(n \gg m > 0)$, 生成的第一个初始归并段的长度最大值和最小值分别是多少?