



中国科学院—中国科学技术大学

2005 年招收攻读硕士学位研究生入学考试试题

试题名称: 计算机软件基础

第一部分: 编译原理 (30 分, 请将答案写在答题纸上, 写在试卷上不给分)

一、(5 分) 下面是用正规式表示的变量声明:

$(\text{int} \mid \text{float}) \text{id} (, \text{id})^*$;

请改用上下文无关文法表示, 也就是写一个上下文无关文法, 它和该正规式等价。

二、(10 分) 在 C 语言中, $3++$ 和 $(\text{id} + \text{id})++$ 这样的表达式被编译时, 编译器都会报告如下的错误:

invalid lvalue in increment

现有如下简化的 C 语言表达式文法:

$E \rightarrow E + E \mid (E) \mid E ++ \mid \text{id} \mid \text{num}$

请你写一个语法制导定义或翻译方案, 它检查++的运算对象是否合法。

三、(10 分) 下面是一个 C 语言程序:

```
long f1(i)
long i;
{
    return(i*10);
}

long f2(long i)
{
    return(i*10);
}

main()
{
    printf("f1 = %d, f2 = %d\n", f1(10.0), f2(10.0));
}
```

其中函数 f1 和 f2 仅形式参数的描述方式不一样。该程序在 X86/Linux 机器上的运行结果如下:

f1 = 0, f2 = 100

请解释为什么用同样的实在参数调用这两个函数的结果不一样。

四、(5分) 优化编译器对下面程序的局部变量 i 和 j 不分配空间，为什么？

```
main()
{
    long i, j;
    i = 5;
    j = i * 2;
    printf("%d\n", i+j);
}
```

第二部分：操作系统（40分，请将答案写在答题纸上，写在试卷上不给分）

五、填空（6分，每空1分）

1. 虚拟存储管理方案否定了实存储方式的_____和_____等两个特性；
2. 文件目录的最基本功能是_____；
3. 在 UNIX 系统中，i 结点主要对应进程映像三部分中的_____；
4. CPU 的执行状态有_____和_____两种。

六、简答（14分，每题7分）

1. 在引入 cache 的虚拟分页式存储管理方案中，若给出一逻辑地址，其对应的物理地址如何转换得到？
2. 在 UNIX 中，实现文件名到外存盘块的映射采用的方法是什么？

七、问答题（8分）

Unix 操作系统使用 copy-on-write 方式来实现 fork() 系统调用以便创建一个新进程。但在创建进程时并不能完全使用 copy-on-write 方式，因为进程的某些部分必须在创建时即时 copy，请问进程的哪些部分必须在创建时即时 copy，哪些部分可以采用 copy-on-write 方式而不用即时 Copy，这样做有什么好处？

八、分析与设计（12分）

设有一个公有信箱和若干私有信箱。公有信箱可由所有人在其中存取信件；而私有信箱只能由特定拥有者使用，在其中存取信件。显然，信箱是有容量限制的。

1. 分析使用者在收发信件时可能存在的同步与互斥关系；（5分）
2. 选定一种信号量机制，设计出相应的同步机制管理程序；（4分）
3. 分析你的程序是否会出现死锁和饥饿现象；若有，可以采用什么样的解决方法？（3分）

第三部分: 数据结构 (80 分, 请将答案写在答题纸上, 写在试卷上不给分)

九、选择题 (10 分, 每题 2 分)

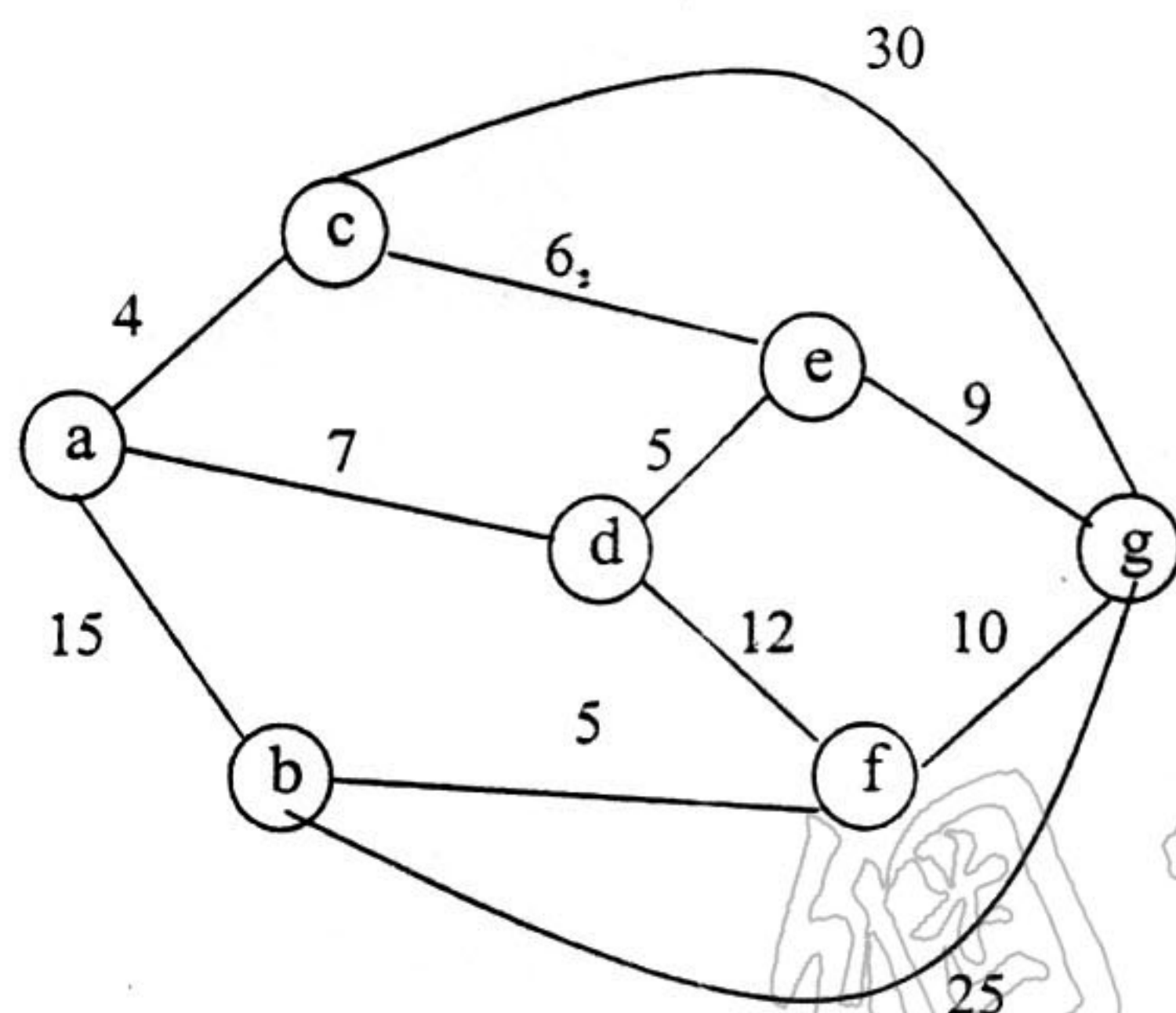
1. 用单循环链表表示队列, 下述说法中正确的是_____。
 - a) 必须设头指针和尾指针才能使入队、出队都方便;
 - b) 可以只设一个头指针, 就可以使入队、出队都方便;
 - c) 可以只设一个尾指针, 就可以使入队、出队都方便;
 - d) 无论如何, 上述三种方法都只可能使入队方便。
2. 设输入序列为 1、2、3、4, 则下述序列中_____不可能是出栈序列。
 - a) 1、2、3、4
 - b) 4、3、2、1
 - c) 1、3、4、2
 - d) 4、1、2、3、
3. 已知二叉树的_____序列, 不能唯一确定一颗二叉树。
 - a) 前序遍历和中序遍历
 - b) 前序遍历和后序遍历
 - c) 中序遍历和后序遍历
 - d) a), b) 和 c)
4. 设有向图 G 用邻接表存储, 则求每个顶点入度的时间复杂度为_____。
 - a) $O(n)$
 - b) $O(n+e)$
 - c) $O(n^2)$
 - d) $O(n \cdot e)$其中: n 为图中的顶点数, e 为图中弧的个数。
5. 在一棵 m 阶 B-树中, 若在某结点中删除一个关键字而导致结点合并, 则该结点中原有_____个关键字。
 - a) $\lceil m/2 \rceil - 1$
 - b) $\lceil m/2 \rceil$
 - c) $m-1$
 - d) m

十、问答题 (20 分, 每题 5 分)

1. 假设顺序表的长度为 n , 在顺序表中删除第 i 个元素的概率为 p_i , 其中 $p_i = \frac{2(n-i)}{n(n-1)}$, 问在该顺序表上删除一个元素时所需移动元素的平均次数是多少?
2. 已知先序遍历和中序遍历某棵二叉树的结点序列分别为: ACKIJHLDBEGF 和 IKJCHLAEGBDF, 请给出这棵二叉树。
3. 已知一棵 3 阶 B-树中有 14 个关键字, 问该 B-树深度至多为多少? (包括叶结点在内)
4. 已知某序列有 7 个记录, 其关键字分别为 40, 35, 76, 11, 27, 51, 13, 现对该序列进行快速排序, 请给出一趟排序之后的关键字序列。

十一、计算题（30 分，每题 10 分）

1. 已知某个字符串只可能出现 8 种字符(A-H)，其中各个字符出现的次数分别为：
A 5, B 25, C 3, D 6, E 10, F 11, G 36, H 4，现要对这个字符串用
0, 1 二个数字进行前缀编码，要求使整个字符串编码后的总长度最短。请给出编
码的方法和各个字符的编码。（需给出相应的树）
2. 已知有如下所示的无向图，画出该图的邻接表，并求该图的一棵最小生成树。



3. 已知有如下的算法（伪代码），其中，数组 $A[1..n]$ 是输入数组，它的每一个元素都是介于 0 到 k 之间的整数；数组 $B[1..n]$ 是输出数组；数组 $C[0..k]$ 为辅助存储空间。请给出当输入数组 $A[1..12] = (6, 1, 3, 1, 2, 4, 5, 6, 2, 5, 3, 5)$ 而且 $k = 6$ 时，算法的运行结果。

```

Sample—Algorithm ( A, B, k)
{
1   for i = 0 to k do
2       C[i] = 0
3   for j = 0 to length(A) do      /* length(A) = n */
4       C[A[j]] = C[A[j]] + 1
5   writeln C                      /* 打印数组 C 的所有元素 */
6   for i = 1 to k do
7       C[i] = C[i] + C[i-1]
8   writeln C
9   for j = length(A) downto 1 do
10      { B[C[A[j]]] = A[j]
11        C[A[j]] = C[A[j]] - 1
12      }
12  writeln B
}
    
```


十二、算法题 (20 分, 每题 10 分)

1. 一个三叉小根堆定义如下: 以数组 $A[1..n]$ 作为存储结构的 n 个元素的序列 $\{k_1, k_2, \dots, k_n\}$ 当前仅当满足如下关系时, 称之为三叉小根堆。

$$\begin{cases} A[i] \leq A[3i-1], & \text{如果 } 3i-1 \leq n \\ A[i] \leq A[3i], & \text{如果 } 3i \leq n \\ A[i] \leq A[3i+1], & \text{如果 } 3i+1 \leq n \end{cases}$$

请写一算法, 要求将任意的一个无序序列 $A[1..n]$ 调整为三叉小根堆。

2. 已知图 G 是一个 DAG 图, $A_{n \times n}$ 是 G 的邻接矩阵。请写一算法, 该算法对 G 中顶点重新编号, 使得新的邻接矩阵为一个上三角阵, 输出每个结点的旧序号和新序号之间的对应关系 (不要求输出新的邻接矩阵); 并给出算法的时间复杂度。