国科大计算机考研全套视频和资料，真题、考点、典型题、命题规律独家视频讲解！
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

中科院软件与计算所2003年硕士研究生入学考试试题
考试科目：软件基础（含编译原理，操作系统和数据结构）

第一部分 编译（40'）

1.（10分）叙述下面的正规式描述的语言，并画出接受该语言的最简DFA的状态转换图。

(1 | 01)* 0*

2.（10分）某语言有两种语句：

$S \rightarrow$ 过程调用语句 | 下标变量赋值语句

过程调用语句的形式是：**id(id, id, …, id)**，即过程名加置于圆括号中的变量表。

下标变量赋值语句的形式是：**id(id, id, …, id)** := **id(id, id, …, id)**，赋值号两边都是数组名加置于圆括号中的变量表。

(a) 请你完成过程调用语句和下标变量赋值语句的文法设计，得到一个以语句S为开始符号的LR(1)文法。不得超过6个产生式，不需要给出你的文法是LR(1)文法的证明。

(b) 如果想在LR分析的同时完成语义分析和中间代码生成，基于你的文法有什么困难？

3.（10分）

(a) 为下面的算术表达式文法写一个语法制导的翻译方案，它将每个子表达式E的符号（即值大于零还是小于零）记录在属性$E.sign$中（属性值分别用POS或NEG表示）。你可以假定所有的整数都不为零，这样就不用担心零的符号。

$E \rightarrow E*E \mid +E \mid -E \mid unsigned\_integer$

(b) 为上面的表达式产生栈机器代码。代码执行后，表达式的值留在栈上。你自己设计所需的栈机器指令，并写清楚指令的含义。

4.（10分）在C语言的教材上，称&为地址运算符，&a为变量a的地址。但是教材上没有说明表达式&a的类型是什么。另外，教材上说，数组名代表数组的首地址，但是也没有说明这个值的类型。它们所带来的一个问题是，如果a是一个数组名，那么表达式a和&a的值都是数组a的首地址，但是它们的使用是有区别的，初学时很难掌握。

下面我们给出4个C文件，请你根据编译报错信息和程序运行结果，写出表达式a和&a的类型表达式。若你能掌握它们的类型，那么它们的区别就清楚了，你也就会正确使用它们了。

（1）文件1：

```
typedef  int  A[10][20];
A  a;

A  *fun()
{
    return(a);
}
```

该函数在Linux上用gcc编译时，报告的类型错误如下：

第6行：warning: return from incompatible pointer type

（2）文件2：

国科大计算机考研全套视频和资料，真题、考点、典型题、命题规律独家视频讲解！
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

```
typedef   int   A[10][20];
A   a;

A   *fun()
{
     return(&a);
}
```

该函数在 Linux 上用 gcc 编译时，没有类型方面的错误。

（3）文件 3：

```
typedef   int   A[10][20];
typedef   int   B[20];
A   a;

B   *fun()
{
     return(a);
}
```

该函数在 Linux 上用 gcc 编译时，没有类型方面的错误。

（4）文件 4：

```
typedef   int   A[10][20];
A   a;

fun()
{
     printf("%d,%d,%d\n", a, a+1, &a+1);
}

main()
{
      fun();
}
```

该程序的运行结果是：

134518112, 134518192, 134518912

第二部分 操作系统（40’）

五. 1、操作系统内核有强内核和微内核，unix 是前者，windowsNT 是后者，简介微内核比强内核的优点。（4’）

2、若只有进程控制，其独立性表现在？引入线程后，独立性有何改变（4’）

3、请求调页存储系统确定页面大小的标准（4’）

六、 1.死锁的证明

在 m 个同类资源，n 个进程共享它，每次进程只能获得或释放至多一个资源，问会不会发

国科大计算机考研全套视频和资料，真题、考点、典型题、命题规律独家视频讲解！
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

生死锁，若：
(1)、设每个进程所需资源数为 ri 1<=ri<=m (6')
2、windows NT 页面大小为 4KB，采用两级页表机构，为提高 设了 32K 或 64K 的 Cache，试叙述 windows NT 地址变换过程的页面调度策略。（10'）
3、假设有一种新磁盘技术，两者即磁盘与内存访问时间在同一数量级上，作下面哪些修改以采用更快的磁盘访问速度。 （12'）
（1） 进程调度（4'） （2）内存管理（4'） （3）磁盘驱动程序（4'）

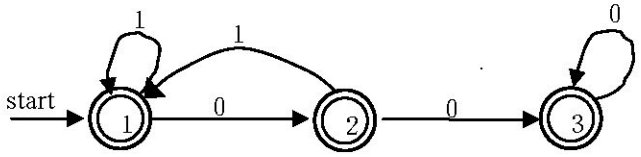
第三部分 数据结构（70 分）
七. 选择（5×2'）
八.简答（10×2'）

九、（5×5'分）
1、广义表，设 H 表示 Get head ，T 表示 Get Tail 从下表中分解出原子 a,请给出 H、T 操作序列。
L=((( )),(b,c),((b,(c,a)),(c,d)),((e),d))
2、串序列 T="abcabcabca"模式串 w="abca"用 kmp 算法，求 next[1:10]
3、一无向图，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。
4、判断向一无环图增加一边是否会使图中产生环的问题时，应选用什么样的数据结构？（一名话简单回答）在使用这种数据结构时该判断所需时间。
5、设向一棵空平衡二叉树（AVL）中插入关键字序列为[45，24，12，62，70，50，10，38]画出每插入一关键字后该树状态示意图，若在此基础上删除关键字 62，给出删除后的状态图。

十、（15 分）有 n 张扑克牌，存在由记录组成的数组 A（1：n）中，每个记录有三个域，其中，N0 为每张扑克初始序号，一旦给定不改变,Cor 表示每张扑克花色，梅花<方块<红桃<黑桃 ，Val 表子扑克数值 1..13，要将这 n 张由小→大排序，每张只能看一次，低花色比高花色的值小，花色的大小均相同的保持原相对的次序，请写算法，并描述所用附加存储空间结构。

编译原理 2003 部分答案

1．该正规式描述的语言是，所有不含子串 001 的 0 和 1 的串。



2．
(a) $S \rightarrow Call \mid Assign$
$Call \rightarrow \mathbf{id}(\text{id\_list})$

国科大计算机考研全套视频和资料，真题、考点、典型题、命题规律独家视频讲解！
详见：网学天地（www.e-studysky.com）；咨询QQ：2696670126

Assign → **id**(id_list) := **id**(id_list)

id_list → id_list, **id** | **id**

(b) 由于过程参数和下标表达式的中间代码是不一样的，在 **id** 和 id_list, **id** 向 id_list 归约时，不知按哪种方式处理。

3．

(a) $E \rightarrow E_1 * E_2$ {**if** $E_1.sign = E_2.sign$ **then** $E.sign$:= POS **else** $E.sign$:= NEG }

$E \rightarrow +E_1$ { $E.sign := E_1.sign$ }

$E \rightarrow -E_1$ {**if** $E_1.sign$= POS **then** $E.sign$:= NEG **else** $E.sign$:= POS}

$E \rightarrow$ *unsigned_integer* {$E.sign$:= POS}

(b) 指令的解释如下：

PUSH 值： 将值压栈

NEG： 将栈顶值取出，计算其相反数，把结果压入栈

MUL： 将栈顶和次栈顶的值取出，将它们相乘，把结果压栈

产生代码的翻译方案如下：

$E \rightarrow E_1 * E_2$ { emit(MUL)}

$E \rightarrow +E_1$ { }

$E \rightarrow -E_1$ {emit('NEG')}

$E \rightarrow$ *unsigned_integer*{emit('PUSH' *unsigned_integer.lexval*)}

4．对一个t类型的数组a[i_1][i_2]…[i_n]来说，表达式a的类型是：

pointer(array(0.. i_2 –1, … array(0.. i_n –1, t)…))

而表达式&a 的类型是：

pointer(array(0.. i_1 –1, … array(0.. i_n –1, t)…))