



中国科学院 - 中国科学技术大学

2004 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 1 部份, 40 分, 请将所有答案写在答题纸上!!!

一、本题 20 分, 每小题 5 分:

1. 请简述进程和线程之间的异同点。并至少分别给出两个以上的线程相对与进程的优缺点和缺点。
2. 请简述系统调用的过程, 并指出在设计和实现系统调用时需要特别注意的问题。
3. 请简述内部碎片和外部碎片的区别, 并分别给出至少两种在操作系统中能遇到的实例情况。
4. 请以一种典型的操作系统为例, 说明其中进程的动态优先级调度算法的设计方法。

二、本题 10 分, 每小题 5 分:

1. 请比较 LRU 和 LFU 之间的区别, 在用于页面置换时各自需要哪些硬件和数据结构?
2. 假设在一个使用请求调页策略, 包含三个空页框虚拟存储系统中, 下列的页号依次被引用: 1 2 1 3 2 1 4 3 1 1 2 4 1 5 6 2 1。请对于 LRU 和 CLOCK 置换算法, 分别给出页框中的内容变化以及出现的缺页次数。

三、本题 10 分:

现有一个计算机存储系统需要设计, 下图给出了可选购的关键组件。

组件	延迟	最小大小	价格
TLB	10ns	16 个页表项	¥20/项
内存	200ns	16MB	¥2/MB
磁盘	10ms	2GB	¥0.2/MB

经费预算为 2000¥。假定: 页大小固定为 8KB; 系统中需要同时运行 4-5 个应用程序, 每个程序最大大小为 64MB, 工作集为 256KB; TLB 中不包含进程标志符。请讨论以最高执行性能为目标划分预算, 如何选购组件。

第2部份, 30分, 请将所有答案写在答题纸上!!!

四. 本题10分:

下面的文法是 LR(1)文法。请合并该文法 LR(1)项目集规范族中的同心集, 以说明该文法不是 LALR(1)文法。

$S \rightarrow aAd | bBd | aBe | bAe$

$A \rightarrow c$

$B \rightarrow c$

五. 本题10分:

为下面文法写一个语法制导的定义, 用 S 的综合属性 val 给出下面文法中 S 产生的二进制的值。例如, 输入 101.101 时, $S.val := 5.625$ 。(不得修改文法。)

$S \rightarrow L.R | L$

$L \rightarrow LB | B$

$R \rightarrow BR | B$

$B \rightarrow 0 | 1$

六. 本题5分:

在 X86/Linux 机器上, 编译器报告第 15 行有错误:

incompatible types in return

在 C 语言中, 数组和结构都是构造类型, 为什么下面第 2 个函数有类型错误, 而第 1 个函数没有?

```
Typedef int A1[10];
Typedef int A2[10];
A1 a;
Typedef struct {int i;} S1;
Typedef struct {int i;} S2;
S1 s;
```

```
A2 *fun1()
{
    return(&a);
}
```

```
S2 fun2()
{
    return(s);
}
```

七. 本题5分:

在 C 语言中, 若 a 和 b 是相同的结构类型, 那么赋值 $a=b$ 是可以的。但是编译器对这种赋值的实现方式可能和它们的字节数 (size) 有关。下面是两个 C 语言程序及在 X86/Linux 机器上生成的目标代码 (略去了和本题目无关的部分)。扼要叙述这些目标代码中体现出的实现方式上的区别。(在下面汇编程序中给出的注释仅供参考。)

```
struct {long i,j; double m;}a,b;
main()
{
    a=b;
}
```

```
main:
    pushl %ebp
    movl %esp,%ebp
    movl $b,%eax
    movl $a,%edx
    movl (%eax),%ecx
    movl %ecx,(%edx)
    movl 4(%eax),%ecx
    movl %ecx,4(%edx)
    movl 8(%eax),%ecx
    movl %ecx,8(%edx)
    movl 12(%eax),%eax
    movl %eax,12(%edx)
.L1:
    leave
    ret
```

```
struct {long i,j; double m,n;}a,b;
main()
{
    a=b;
}
```

```
main:
    pushl %ebp
    movl %esp,%ebp
    pushl %edi
    pushl %esi
    movl $a,%edi
    movl $b,%esi
    cld
    movl $6,%ecx
    rep
    movsl
.L1:
    leal -8(%ebp),%esp
    popl %esi
    popl %edi
    leave
    ret
```

- 设定方向标志，使地址指针自动增量
- 设定重复次数
- 表示重复下面的指令
- 传送指令

第3部份，共80分，请将所有答案写在答题纸上!!!

八、是非题，请在答题纸上回答“对”或“错”，共12分：

1. 用邻接矩阵作为图的存储结构时，所需存储空间的大小与图的结点数有关，而与边数无关。
2. 在一棵完全二叉树中，度为1的结点数一定不超过1。
3. 若某完全二叉树中，根的关键字为该树中各结点关键字的最大值，则该完全二叉树为一个堆。
4. 在采用线性探测法处理冲突的散列表中所有同义词在表中相邻。
5. 一棵树，若它的左、右子树都是AVL树，则该树也是AVL树。
6. 所有时间复杂度为 $O(n^2)$ 的简单排序算法（如：简单插入排序、冒泡排序、简单选择排序等）均是稳定的排序算法。

九、基本题，共48分，6小题，每题8分：

1. 在顺序表中插入或删除一个元素，所需的元素移动次数是否只与插入或删除的位置 i 有关？为什么？（请回答，并用1或2句话简单叙述理由）。
2. 若某个有向图的邻接矩阵为三角矩阵，该有向图是否一定具有拓扑有序序列？为什么？（请回答，并用1或2句话简单叙述理由）。
3. 含有12个叶子结点的3阶B-树中至少有多少个非叶子结点？（请简单给出求解过程）。
4. 已知有一段文字含有7种不同的字符，这些字符在这段文字中出现的次数分别为：A 4次，B 3次，C 6次，D 5次，E 11次，F 10次，G 1次，现要对这段文字进行二进制编码（前缀码），问这段文字的二进制编码的总长度至少有多少位（二进制位）？（请给出计算过程）。
5. 已知有一个无向图的邻接矩阵如下所示，要求：1) 画出该图；2) 给出该图的一棵最小生成树。

∞	4	1	∞	3	6	∞	∞
4	∞	2	1	∞	∞	2	7
1	2	∞	4	3	∞	∞	∞
∞	1	4	∞	1	5	4	3
3	∞	3	1	∞	3	∞	∞
6	∞	∞	5	3	∞	4	7
∞	2	∞	4	∞	4	∞	2
∞	7	∞	3	∞	7	2	∞

6. 已知先序遍历一棵二叉排序树 T 时，访问 T 中结点所含关键字的序列为 18, 12, 10, 24, 17, 19, 15，请给出这棵二叉排序树（共 7 个结点）。

十、写算法，20 分，每题 10 分：

1. 编写算法统计二叉树中第 k 层的结点总数，其中二叉树采用二叉链表表示， k 作为参数。
2. 写一个非递归算法，从一个具有 n 个元素的数组中同时找出其最大和最小的元素，要求算法所需的附加空间为 $O(1)$ （即常数个附加空间），算法所需的元素间的比较次数在最坏情况下 $\leq 3n/2$ 。

中国科学院 & 中国科学技术大学

2004 年硕士学位研究生入学考试试题参考答案

试题名称： 计算机软件基础

第 1 部份，40 分，请将所有答案写在答题纸上!!!

一、本题 20 分，每小题 5 分：

1. 线程和进程都可以是操作系统的基本调度单位，线程又被称做是轻量级进程。一个进程能够包含多个线程，这些线程共享进程的系统资源，如文件、内存映射等。(1 分) 线程的优点包括：调度单位小，系统资源占用少，提高了处理器的利用率，提高了程序的响应度，线程间通讯更容易 (2 分)；缺点：需要附加的数据结构进行管理；若干系统调用(fork 等)和信号的处理比进程更复杂 (2 分)
2. 软中断，系统陷入内核态。系统调用的处理在核心栈内进行，参数传递可使用寄存器或内存 buffer。(3 分)设计系统调用时需要特别注意参数的安全性检查，避免用户传入非法的参数。(2 分)
3. 外部碎片通常指再也无法被分配出去的较小的存储块；内部碎片指已被分配出去，但永不会被使用的存储块。(1 分) 外部碎片：内存管理中的动态分区分配；文件分配中的连续分配。(2 分)内部碎片：分页，每个进程会有 $\text{page size}/2$ 大小的内部碎片；文件的多个逻辑块保存在一个磁盘块中也会产生内部碎片。(2 分) (也可从于内外存分配中的角度出来回答)
4. 如以 LINUX 为例，首先为进程赋静态优先级 (有一定范围)；然后每隔一段时间 (如 1 个时钟中断周期)，按照某种标准对进程记数 (如进程还能运行的时间 counter)；最后在满足特定条件时 (如所有进程的 counter=0)，动态调整进程的优先级 (如动态优先级=counter>>1+静态优先级)。

二、本题 10 分，每小题 5 分：

- 1 LRU 为最近最久未使用算法，LFU 为最不常用算法，即从程序运行开始被访问次数最少的页面被置换。(2 分)使用时需要的数据结构和硬件支持等不同，如 LRU 需要一个特殊的栈 (保存访问的页) 以及为每个页面设立移动寄存器；而 LFU 只需为每页设置访问计数器。(3 分)

2. LRU，共 11 次缺页。

	1	2	1	3	2	1	4	3	1	1	2	4	1	5	6	2	1
0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2
1		2	2	2	2	2	2	3	3	3	3	4	4	4	6	6	6
2				3	3	3	4	4	4	4	2	2	2	5	5	5	1
	Y	Y		Y			Y	Y			Y	Y		Y	Y	Y	Y

CLOCK，共 9 次缺页

	1	2	1	3	2	1	4	3	1	1	2	4	1	5	6	2	1
0	1	1	1	1	1	1	4	4	4	4	4	4	4	5	5	5	5
1		2	2	2	2	2	2	2	1	1	1	1	1	1	6	6	6
2				3	3	3	3	3	3	3	2	2	2	2	2	2	1
	Y	Y		Y			Y		Y		Y			Y	Y		Y

中国科学院 & 中国科学技术大学

2004 年硕士学位研究生入学考试试题参考答案

三、本题 10 分：

5 个程序需要 $5 \times 64\text{M} = 320\text{M}$ 的磁盘空间。一个磁盘足够保存这些程序，共计 $0.2 \times 2048 = 409.6\text{RMB}$ 。还剩余 $2000 - 409.6 = 1590.4\text{RMB}$ 。将所有 5 个程序都保存在内存中能够省略交换的时间。于是需要 320M 的内存，需要 $320 \times 2 = 640\text{RMB}$ ，还剩下 $1590.4 - 640 = 950.4\text{RMB}$ 。每个进程的工作集大小为 246K ，共计 32 个页面。因此 TLB 必须要能至少容纳 32 个项，从而避免 TLB miss。如此需要 $32 \times 20 = 640\text{RMB}$ 。最后剩余 $950.4 - 640 = 310.4$ ，全部用来购买 TLB， $310.4 / 20 = 15$ 。如果整体的设计思路正确，比如从瓶颈(disk->memory->TLB)开始考虑，加 3 分。

	4 分	2 分	1 分	0 分
TLB	[32, 47]	(47, 64)	>64	<32
内存	[320,]	[128, 320)	[64, 128)	(0, 64)
磁盘	2G	>2G	>2G	<2G

中国科学院 & 中国科学技术大学

2004 年硕士学位研究生入学考试试题参考答案

第 2 部份, 30 分。

四. 本题 10 分:

对活前缀 ac 和 bc 有效的项集分别为 $\{[A \rightarrow c \cdot, d], [B \rightarrow c \cdot, e]\}$ 和 $\{[A \rightarrow c \cdot, e], [B \rightarrow c \cdot, d]\}$, 它们是同心的。合并后变成 $\{[A \rightarrow c \cdot, d/e], [B \rightarrow c \cdot, d/e]\}$, 产生归约-归约冲突。因此该文法不是 LALR(1) 文法。

五. 本题 10 分:

$S \rightarrow L \cdot R$	$S.val := L.val + R.val$
$S \rightarrow L$	$S.val := L.val$
$L \rightarrow L_1 B$	$L.val := L_1.val \times 2 + B.val$
$L \rightarrow B$	$L.val := B.val$
$R \rightarrow B R_1$	$R.val := (R_1.val + B.val) / 2$
$R \rightarrow B$	$R.val := B.val / 2$
$B \rightarrow 0$	$B.val := 0$
$B \rightarrow 1$	$B.val := 1$

六. 本题 5 分:

C 语言对除结构类型以外的所有类型使用结构等价, 而对结构类型使用名字等价。

第 1 个函数能通过结构等价的检查, 而第 2 个函数不能通过名字等价的检查。

七. 本题 5 分:

在结构的字节数较少时, 则为该结构各域分别产生值传送指令。

在结构的字节数较多时, 则根据值传送的源地址 (b 的地址)、目的地址 (a 的地址) 和该结构的长字数, 产生简洁的重复传送指令, 以提高效率。

中国科学院 & 中国科学技术大学

2004 年硕士学位研究生入学考试试题参考答案

第 3 部份, 共 80 分, 请将所有答案写在答题纸上!!!

八、是非题, 请在答题纸上回答“对”或“错”, 共 12 分:

1. 对 2. 对 3. 错 4. 错 5. 错 6. 错

九、基本题, 共 48 分, 6 小题, 每题 8 分:

1. 在顺序表中插入或删除一个元素, 所需的元素移动次数是否只与插入或删除的位置 i 有关? 为什么? (请回答, 并用 1 或 2 句话简单叙述理由)。

答: 不仅与插入或删除的位置 i 有关, 而且与顺序表的长度有关。

2. 若某个有向图的邻接矩阵为三角矩阵, 该有向图是否一定具有拓扑有序序列? 为什么? (请回答, 并用 1 或 2 句话简单叙述理由) 答: 是。

3. 含有 12 个叶子结点的 3 阶 B-树中至少有多少个非叶子结点? (请简单给出求解过程)。

答: 至少有 7 个非叶子结点。

4. 已知有一段文字含有 7 种不同的字符, 这些字符在这段文字中出现的次数分别为: A 4 次, B 3 次, C 6 次, D 5 次, E 11 次, F 10 次, G 1 次, 现要对这段文字进行二进制编码, 问这段文字的二进制编码的总长度至少有多少位 (二进制位)? (请给出计算过程)。

答: 这段文字的二进制编码的总长度至少有 103 个二进制位。

5. 已知有一个无向图的邻接矩阵如下所示, 要求: 1) 画出该图; 2) 给出该图的一棵最小生成树。

∞	4	1	∞	3	6	∞	∞
4	∞	2	1	∞	∞	2	7
1	2	∞	4	3	∞	∞	∞
∞	1	4	∞	1	5	4	3
3	∞	3	1	∞	3	∞	∞
6	∞	∞	5	3	∞	4	7
∞	2	∞	4	∞	4	∞	2
∞	7	∞	3	∞	7	2	∞

答: 最小生成树的边为: (1, 3), (2, 3), (2, 4), (4, 5), (2, 7), (7, 8), (5, 6)。

6. 已知先序遍历一棵二叉排序树 T 时, 访问 T 中结点所含关键字的序列为 18, 12, 10, 24, 17, 19, 15, 请给出这棵二叉排序树 (共 7 个结点)。

答: 该树的中序遍历序列为: 10, 12, 15, 17, 18, 19, 24. 树的根结点为 18, 左子树的根为 12, 右子树的根为 24, 12 的左儿子为 10, 右儿子为 17, 24 的左儿子为 19, 17 的左儿子为 15。

中国科学院 & 中国科学技术大学

2004 年硕士学位研究生入学考试试题参考答案

十、写算法，20 分，每题 10 分：

1. 编写算法统计二叉树中第 k 层的结点总数，其中二叉树采用二叉链表表示， k 作为参数。

```
int countKlevel (BiTree T, int k)
{
    // 初始化队列，队列元素含两项：一为结点指针，另一为该结点的层次
    InitQueue(Q);

    count = 0; // 设置计数器为 0
    if (T != NULL)
    {
        // 将(T, 1)入队，T 指向的结点层次为 1
        EnQueue(Q, T, 1);
        while (!QueueEmpty(Q))
        {
            // 出队，将原队头元素中的结点指针存储到 p，相应的层次存储到 level
            DeQueue(Q, p, level);
            if (level == k) count++;
            else if (level > k) break;
            if (p->lchild != NULL)
                // 将左孩子指针及其层次入队
                EnQueue(Q, p->lchild, level+1);
            if (p->rchild != NULL)
                // 将右孩子指针及其层次入队
                EnQueue(Q, p->rchild, level+1);
        }
    }
    return(count);
}
```

注：该题也可用按层次遍历的方法，可以借用图的 BFS 算法，稍加改动即可。

2. 写一个非递归算法，从一个具有 n 个元素的数组中同时找出其最大和最小的元素，要求算法所需的附加空间为 $O(1)$ （即常数个附加空间），算法所需的元素间的比较次数在最坏情况下 $\leq 3n/2$

答：算法略。

算法思想：第 1 步，先将 n 个元素进行两两比较，得到 $n/2$ 个小元素和 $n/2$ 个大元素，最多需要 $n/2$ 次比较；第 2 步，分别在 $\lceil n/2 \rceil$ 个小元素中找最小值，在 $\lceil n/2 \rceil$ 个大元素中找最大值，共需 $2(\lceil n/2 \rceil - 1)$ 次比较；总共用的比较次数小于 $3n/2$ 。