

2008 年硕士学位研究生入学考试试题

计算机软件基础

所有试题答案写在答题纸上, 答案写在试卷上无效

第 1 部分: 编译原理 共 40 分

一、描述题 (每小题 10 分, 共 40 分)

1. 描述由正规式 $b^*a(bb^*a)^*b^*$ 定义的语言, 并画出接受该语言的最简 DFA。

2. 下面的文法产生代表正二进制数的 0 和 1 的串集:

$$B \rightarrow B0 \mid B1 \mid 1$$

下面的翻译方案计算这种正二进制数的十进制值:

$$\begin{aligned} B &\rightarrow B_1 0 \{B.val := B_1.val \times 2\} \\ &\mid B_1 1 \{B.val := B_1.val \times 2 + 1\} \\ &\mid 1 \{B.val := 1\} \end{aligned}$$

请消除该基础文法的左递归, 再重写一个翻译方案, 它仍然计算这种正二进制数的十进制值。

3. 一个 C 语言的函数如下:

```
func(i) long i; {  
    long j;  
    j = i - 1;  
    func(j);  
}
```

下面左右两边的汇编代码是两个不同版本 GCC 编译器为该函数产生的代码。左边的代码在调用 func 之前将参数压栈, 调用结束后将参数退栈。右边代码对参数传递的处理方式没有实质区别。请叙述右边代码对参数传递的处理方式并推测它带来的优点。

func:		func:	
pushl	%ebp	pushl	%ebp
movl	%esp, %ebp	movl	%esp, %ebp
subl	\$4, %esp	subl	\$8, %esp
movl	8(%ebp), %edx	movl	8(%ebp), %eax
decl	%edx	decl	%eax
movl	%edx, -4(%ebp)	movl	%eax, -4(%ebp)
movl	-4(%ebp), %eax	movl	-4(%ebp), %eax
pushl	%eax	movl	%eax, (%esp)
call	func	call	func
addl	\$4, %esp	leave	
leave		ret	
ret			

4. 在 C 语言中, 如果变量 i 和 j 都是 long 类型, 请写出表达式 $\&i$ 和表达式 $\&i-\&j$ 的类型表达式。为帮助你回答问题, 下面给出一个程序作为提示, 它运行时输出 1。

```
main() {  
    long i, j;  
    printf("%d\n", &i-&j);  
}
```

第 2 部分: 操作系统 共 40 分

二、简答题 (每小题 6 分, 共 24 分)

1. 采用 `fork()` 原语创建一个进程时, 所要完成的工作有哪些?
2. 在虚拟页式存储管理方案中, 影响缺页次数的因素有哪些?
3. 请给出至少 3 种常用的文件物理结构。它们各自的基本思想和优点是什么?
4. 设备驱动程序的目的是为了控制 I/O 传输。其具体工作流程是什么?

三、叙述分析题 (第 1 小题 9 分, 第 2 小题 7 分, 共 16 分)

1. 进程调度算法分析问题

设某操作系统的进程调度算法采用的是动态优先数调度策略, 并且规定: 优先数越大, 其优先级越低。优先数计算公式为: $P_{pri} = P_{cpu}/2 + P_{user} + P_{nice} + NZERO$ 其中: P_{user} 、 $NZERO$ 是优先数基值; P_{cpu} 是每个进程最近一次使用 CPU 的时间; P_{nice} 是用户设置的进程优先数偏移值。

请回答如下问题:

- (1) 对新创建的进程, 分析能否较快得到调度;
- (2) 对正运行的原有进程, 分析是否会被其它进程抢先;
- (3) 对被阻塞的原有进程, 分析能否重新得到机会, 抢占 CPU。

2. 死锁处理问题

哲学家就餐是一种典型的死锁问题。为避免其中死锁现象的发生, 请给出至少 3 种可行的措施。

第3部分: 数据结构 共70分

四、选择题: (每小题2分, 共10分)

1. 已知广义表 $L = ((a), b), (a, (c)), (d, (e)), f)$, 则 $Head[Tail[Head[Tail[L]]]]$ 的运算结果为_____。
a) a b) (a) c) c d) (c) e) 其它
2. 一棵有 n 个结点的完全二叉树至多有 _____ 层结点;
a) $\lfloor \log_2 n \rfloor$ b) $\lfloor \log_2(n+1) \rfloor$ c) $\lceil \log_2 n \rceil$ d) $\lceil \log_2(n+1) \rceil$
3. 一棵有 n 个结点的二叉排序树最多可以有 _____ 层结点;
a) $\lfloor \log_2 n \rfloor + 1$ b) $2\lfloor \log_2 n \rfloor$ c) $\lfloor n/2 \rfloor$ d) n e) 其它
4. 下述排序算法中, _____ 是稳定的排序算法。
a) 简单选择排序 b) Shell 排序 c) 归并排序 d) 堆排序
5. 如果 G 是一个共有 n 个结点的有向完全图, 则该图中共有 _____ 条弧。
a) $\frac{1}{2}n(n-1)$ b) $\frac{1}{2}(n^2-1)$ c) $n(n-1)$ d) n^2-1

五、简答题 (每小题5分, 共20分)

1. 已知入栈序列(次序)为 a_1, a_2, a_3, a_4, a_5 , 请问其合理的出栈序列(次序)共有多少种?
2. 已知2维数组 $A[0:4, 1:5]$ 采用行主序的存储方式, 每个数组元素占用4个存储单元, 其第1个元素 $A[0, 1]$ 的存储地址为1020, 请问 $A[3, 3]$ 的存储地址是什么?
3. 简单插入排序算法的最坏情况时间复杂度为 $O(n^2)$, 而归并排序算法的最坏情况时间复杂度为 $O(n \log n)$, 请问这是否表明排序任意的 n 个元素, 归并排序算法总是比简单插入排序算法快? 请简单加以解释。
4. 一棵有19个关键字的3阶B-树至少有多少非叶结点? (请简单给出求解过程);

六、计算题（每小题 8 分，共 24 分）

1. 已知模式串 $T = "abcabcdabcb"$ ，请给出用 KMP 算法进行串匹配时的 Next 数组的值 $Next[1..12]$ 。
2. 已知有一个有向图的邻接矩阵如下所示，其结点的编号从 1 至 8，请给出由该图中的 1 号结点至其余各个结点的最短距离。

∞	1	2	3	1	5	∞	4
2	∞	4	5	∞	1	3	3
3	6	∞	3	2	4	5	1
8	∞	9	∞	2	1	2	∞
1	9	5	7	∞	1	3	6
∞	5	∞	3	1	∞	5	9
3	3	2	∞	4	2	∞	8
9	∞	1	∞	1	1	2	∞

3. 设 Hash 表地址为 0 到 15，关键字序列为(53, 23, 47, 15, 38, 68, 78, 48, 25, 82)，Hash 函数为 $H(\text{key}) = \lfloor \text{key} / 2 \rfloor \bmod 13$ ，用线性探测法解决冲突，计算在等概率情况下查找成功时关键字的平均比较次数。

七、程序设计题（第 1 小题 6 分，第 2 小题 10 分，共 16 分）

1. 已知二叉树 T 采用二叉链表做为存储结构，每个结点有三个域：Dep, Lchild 和 Rchild；其中 Lchild 和 Rchild 分别为指向其左、右子结点的指针，要求写程序遍历二叉树 T ，修改各个结点的 Dep 域的值(初始值为 0)，使得各个结点的 Dep 域值等于该结点至根结点的距离（即：根结点至该结点的路径所含有的边数）。
2. 已知数组 $A[1..n]$ 中存放着一些实数，有正有负，请写程序找出其最大求和子区间，即：找出下标 i 和 j ， $1 \leq i \leq j \leq n$ ，使得 $\sum_{k=i}^j A[k]$ 的值达到最大。

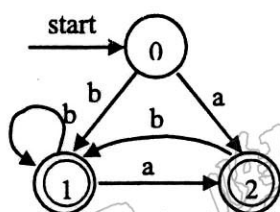
2008 年硕士学位研究生入学考试试题参考答案

(计算机软件基础)

第 1 部分: 编译原理 共 40 分

一、描述题 (每小题 10 分, 共 40 分)

- 1、正规式 $b^*a(bb^*a)^*b^*$ 体现的特点是, 每个 a 的左边都有若干 b , 除非 a 是第一个字母。因此该正规式定义的语言是: 不含子串 aa 的所有 a 和 b 的非空串的集合。最简 DFA 如下:



- 2、消除左递归后的文法如下:

$$B \rightarrow 1 B'$$

$$B' \rightarrow 0 B' \mid 1 B' \mid \epsilon$$

相应的翻译方案如下:

$$B \rightarrow 1 \{B'.i := 1\} B' \{B.val := B'.val\}$$

$$B' \rightarrow 0 \{B'.i := B'.i \times 2\} B'_1 \{B'.val := B'_1.val\}$$

$$\mid 1 \{B'.i := B'.i \times 2 + 1\} B'_1 \{B'.val := B'_1.val\}$$

$$\mid \epsilon \{B'.val := B'.i\}$$

3. 左边的编译器版本: 一般只为局部变量分配空间。调用函数前, 用若干次 `pushl` 指令将参数压栈, 返回后用 `addl $n, %esp` 一次将所有参数退栈 (常数 n 根据调用前做了多少次 `pushl` 来决定)。

右边的编译器版本: 除了为局部变量分配空间外, 同时还为本函数中出现的函数调用的参数分配空间, 并且参数所用空间靠近栈顶。调用函数前, 用 `movl` 指令将参数移入栈顶, 调用结束后无需参数退栈指令。

优点是每次函数调用结束后不需要执行 `addl $n, %esp` 指令, 另外增加优化的可能性。

4. 表达式 $\&i$ 的类型表达式是 `pointer(long)`, 表达式 $\&i - \&j$ 的类型表达式是 `long`。按照 C 语言的规定, 指向同一个类型的两个指针可以相加减, 它们值的差是它们之间的元素个数。

计算机软件基础答案

第2部分：操作系统 共 40 分

二、简答题（每小题 6 分，共 24 分）

1. 采用 fork() 原语创建一个进程时，所要完成的工作有哪些？

参考答案：

- (1) 申请空白 PCB，为新进程分配唯一的数字标识符，并从 PCB 集合中索取一空白 PCB；
- (2) 为新进程分配资源。包括必要的内存，进程需要的其它资源；
- (3) 进一步填充 PCB，其中主要是复制父进程的上下文；
- (4) 将新进程插入 ready 队列。

2. 在虚拟页式存储管理方案中，影响缺页次数的因素有哪些？

参考答案：

- (1) 分配给进程的物理页面数
- (2) 页面本身的大小
- (3) 程序的编制方法
- (4) 页面淘汰算法

3. 请给出至少 3 种常用的文件物理结构。它们各自的基本思想和优点是什么？

参考答案：

顺序文件、链接文件、索引结构、Hash 结构、倒排结构等。

- (1) 顺序结构：一个文件占有若干个连续的物理块。优点：简单、支持顺序存取和随机存取、顺序存取速度快、所需的磁盘寻道次数和寻道时间最少。
- (2) 链接结构：一个文件占有若干个不连续的存储块，各块之间以指针相连。优点：提高了磁盘空间利用率、不存在外部碎片问题、有利于文件插入和删除、有利于文件动态扩充。
- (3) 索引结构：一个文件占有若干个不连续的存储块，这些块的号记录于一个索引块中。优点：能顺序存取和随机存取、支持文件动态增长和插入删除、充分利用外存空间。

4. 设备驱动程序的目的是为了控制 I/O 传输。其具体工作流程是什么？

参考答案：

主要负责接收和分析从设备分配转来的信息，并根据设备分配的结果，结合具体物理设备特性完成以下具体工作

- (1) 预置设备的初始状态；
- (2) 根据请求传输的数据量，组织 I/O 缓冲队列，利用 I/O 缓冲对数据进行加工，包括数据格式处理和编码转换；
- (3) 构造 I/O 程序（在有通道系统中，是通道程序）；
- (4) 启动设备进行 I/O 操作。

计算机软件基础答案

三、叙述分析题 (第 1 小题 9 分, 第 2 小题 7 分, 共 16 分)

1. 进程调度算法分析问题

参考答案:

- 对于新创建的进程: 由于 $P_{cpu}/2=0$, 所以其优先数会较小, 因此会比较快地得到调度;
- 对于正运行的原有进程: 随着进程执行, 其 P_{cpu} 的值不断增加, 因而可能会被其它进程抢先;
- 对被阻塞的原有进程, 由于公式中 P_{cpu} 被 2 所除, 因此每当计算优先数时, $P_{cpu}/2$ 项会衰减; 因此当其优先数小于当前进程和其它进程时, 又会重新抢占 CPU。

2. 死锁处理问题

参考答案:

- (1) 最多允许 4 个哲学家同时坐在桌子周围;
- (2) 仅当一个哲学家左右两边的筷子都可用时, 才允许他拿筷子;
- (3) 给所有哲学家编号, 奇数号的哲学家必须首先拿左边的筷子, 偶数号的哲学家则反之;
- (4) 把哲学家分为三种状态, 思考, 饥饿, 进食, 并且一次拿到两只筷子, 否则不拿。

计算机软件基础答案

第3部分: 数据结构 共70分

四、选择题: (每小题2分, 共10分)

1. d) 2. d) 3. d) 4. c) 5. c)

五、简答题 (每小题5分, 共20分)

1. 答: 合理的出栈序列共有42种;
2. 答: $A[3, 3]$ 的存储地址是1088.
3. 答: 不是。因为简单插入排序算法的运行时间与数据的分布有关, 在最理想的情况(元素从小到大有序)其运行时间为 $O(n)$, 优于相同情况下归并排序算法的运行时间。
4. 答: 一棵有19个关键字的3阶B-树至少有11个非叶结点。

六、计算题 (每小题8分, 共24分)

1. 答: $Next[1..12] = "0, 1, 1, 1, 2, 3, 4, 1, 2, 3, 4, 5"$
2. 答: 1至2距离为1; 1至3距离为2; 1至4距离为3;
1至5距离为1; 1至6距离为2; 1至7距离为4; 1至8距离为3
3. 在等概率情况下查找成功时关键字的平均比较次数为 $14/10=1.4$

七、程序设计题 (第1小题6分, 第2小题10分, 共16分)

1. 采用二叉树先序遍历算法, 将深度做为子过程的参数, 子结点的深度是父结点的深度加1。
2. 可以用分治法。
也可以设置临时变量 u 代表目前已知的最大区间和的值, 用 u_i 、 u_j 表示与 u 相关的区间, 采用设置标记的方法, 计算目前的和 $temp$, 用一次循环即可以求解。