

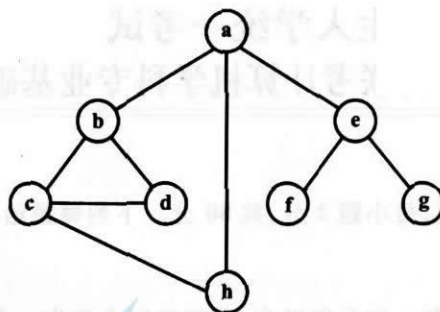
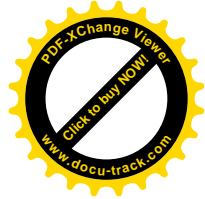
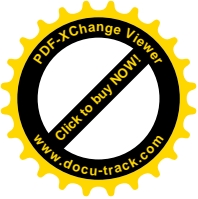
2013 年全国硕士研究生入学统一考试

计算机科学与技术学科联考计算机学科专业基础综合试题

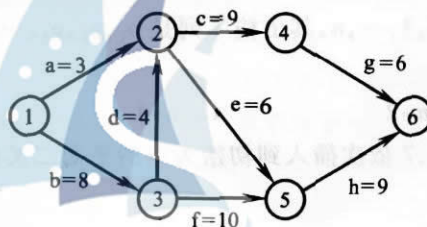
一、单项选择题:1~40 小题,每小题 2 分,共 80 分。下列每题给出的四个选项中,只有一个选项符合试题要求。

- 已知两个长度分别为 m 和 n 的升序链表,若将它们合并为一个长度为 $m+n$ 的降序链表,则最坏情况下的时间复杂度是
A. $O(n)$ B. $O(m \times n)$ C. $O(\min(m, n))$ D. $O(\max(m, n))$
- 一个栈的入栈序列为 $1, 2, 3, \dots, n$, 其出栈序列是 $p_1, p_2, p_3, \dots, p_n$ 。若 $p_2 = 3$, 则 p_3 可能取值的个数是
A. $n-3$ B. $n-2$ C. $n-1$ D. 无法确定
- 若将关键字 $1, 2, 3, 4, 5, 6, 7$ 依次插入到初始为空的平衡二叉树 T 中, 则 T 中平衡因子为 0 的分支结点的个数是
A. 0 B. 1 C. 2 D. 3
- 已知三叉树 T 中 6 个叶结点的权分别是 $2, 3, 4, 5, 6, 7$, T 的带权(外部)路径长度最小是
A. 27 B. 46 C. 54 D. 56
- 若 X 是后序线索二叉树中的叶结点, 且 X 存在左兄弟结点 Y , 则 X 的右线索指向的是
A. X 的父结点 B. 以 Y 为根的子树的最左下结点
C. X 的左兄弟结点 Y D. 以 Y 为根的子树的最右下结点
- 在任意一棵非空二叉排序树 T_1 中, 删除某结点 v 之后形成二叉排序树 T_2 , 再将 v 插入 T_2 形成二叉排序树 T_3 。下列关于 T_1 与 T_3 的叙述中, 正确的是
I. 若 v 是 T_1 的叶结点, 则 T_1 与 T_3 不同
II. 若 v 是 T_1 的叶结点, 则 T_1 与 T_3 相同
III. 若 v 不是 T_1 的叶结点, 则 T_1 与 T_3 不同
IV. 若 v 不是 T_1 的叶结点, 则 T_1 与 T_3 相同
A. 仅 I、III B. 仅 I、IV C. 仅 II、III D. 仅 II、IV
- 设图的邻接矩阵 A 如下所示。各顶点的度依次是
$$A = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

A. 1, 2, 1, 2 B. 2, 2, 1, 1 C. 3, 4, 2, 3 D. 4, 4, 2, 2
- 若对如下无向图进行遍历, 则下列选项中, 不是广度优先遍历序列的是
A. h, c, a, b, d, e, g, f B. e, a, f, g, b, h, c, d
C. d, b, c, a, h, e, f, g D. a, b, c, d, h, e, f, g



9. 下列 AOE 网表示一项包含 8 个活动的工程。通过同时加快若干活动的进度可以缩短整个工程的工期。下列选项中,加快其进度就可以缩短工程工期的是



- A. c 和 e B. d 和 c C. f 和 d D. f 和 h
10. 在一棵高度为 2 的 5 阶 B 树中,所含关键字的个数最少是
A. 5 B. 7 C. 8 D. 14
11. 对给定的关键字序列 110,119,007,911,114,120,122 进行基数排序,则第 2 趟分配收集后得到的关键字序列是
A. 007,110,119,114,911,120,122 B. 007,110,119,114,911,122,120
C. 007,110,911,114,119,120,122 D. 110,120,911,122,114,007,119
12. 某计算机主频为 1.2 GHz,其指令分为 4 类,它们在基准程序中所占比例及 CPI 如下表所示。

指令类型	所占比例	CPI
A	50%	2
B	20%	3
C	10%	4
D	20%	5

该机的 MIPS 数是

- A. 100 B. 200 C. 400 D. 600
13. 某数采用 IEEE 754 单精度浮点数格式表示为 C640 0000H,则该数的值是
A. -1.5×2^{13} B. -1.5×2^{12} C. -0.5×2^{13} D. -0.5×2^{12}
14. 某字长为 8 位的计算机中,已知整型变量 x 、 y 的机器数分别为 $[x]_{\text{补}} = 1\ 1110100$, $[y]_{\text{补}} = 1\ 0110000$ 。若整型变量 $z = 2 * x + y/2$,则 z 的机器数为
A. 1 1000000 B. 0 0100100 C. 1 0101010 D. 溢出

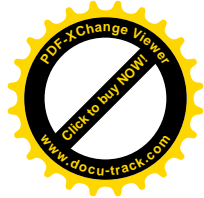
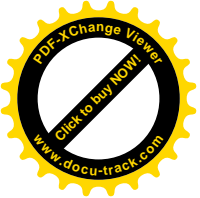


二、综合应用题:41~47 小题,共 70 分。

41. (13 分)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。
42. (10 分) 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。请回答:
- (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?
43. (9 分) 某 32 位计算机, CPU 主频为 800 MHz, Cache 命中时的 CPI 为 4, Cache 块大小为 32 字节; 主存采用 8 体交叉存储方式, 每个体的存储字长为 32 位、存储周期为 40 ns; 存储器总线宽度为 32 位, 总线时钟频率为 200 MHz, 支持突发传送总线事务。每次突发传送总线事务的过程包括: 送首地址和命令、存储器准备数据、传送数据。每次突发传送 32 字节, 传送地址或 32 位数据均需要一个总线时钟周期。请回答下列问题, 要求给出理由或计算过程。
- (1) CPU 和总线的时钟周期各为多少? 总线的带宽(即最大数据传输率)为多少?
 - (2) Cache 缺失时, 需要用几个读突发传送总线事务来完成一个主存块的读取?
 - (3) 存储器总线完成一次读突发传送总线事务所需的时间是多少?
 - (4) 若程序 BP 执行过程中, 共执行了 100 条指令, 平均每条指令需进行 1.2 次访存, Cache 缺失率为 5%, 不考虑替换等开销, 则 BP 的 CPU 执行时间是多少?
44. (14 分) 某计算机采用 16 位定长指令字格式, 其 CPU 中有一个标志寄存器, 其中包含进位/借位标志 CF、零标志 ZF 和符号标志 NF。假定为该机设计了条件转移指令, 其格式如下:

15	11	10	9	8	7	0
0 0 0 0 0		C	Z	N	OFFSET	

其中, 00000 为操作码 OP; C、Z 和 N 分别为 CF、ZF 和 NF 的对应检测位, 某检测位为 1 时表示需检测对应标志, 需检测的标志位中只要有一个为 1 就转移, 否则不转移, 例如, 若 $C=1, Z=0, N=1$, 则需检测 CF 和 NF 的值, 当 $CF=1$ 或 $NF=1$ 时发生转移; OFFSET 是相对偏移量, 用补码表示。转移执行时, 转移目标地址为 $(PC)+2+2 \times \text{OFFSET}$; 顺序执行时, 下条指令地址为 $(PC)+2$ 。请回答下列问题。



计算机学科专业基础综合试题参考答案及解析 (2013 年)

一、单项选择题

1. D

解析: m 、 n 是两个升序链表, 长度分别为 m 和 n 。在合并过程中, 最坏的情况是两个链表中的元素依次进行比较, 比较的次数最少是 m 和 n 中的最小值。

2. C

解析: 除了 3 本身以外, 其他的值均可以取到, 因此可能取值的个数为 $n-1$ 。

3. D

解析: 利用 7 个关键字构建平衡二叉树 T , 平衡因子为 0 的分支结点个数为 3, 构建的平衡二叉树如下图所示。



4. B

解析: 利用二叉树的 6 个叶子结点的权构建最小带权生成树, 最小的带权路径长度为 $(2+3) \times 3 + (4+5) \times 2 + (6+7) \times 1 = 46$ 。

5. A

解析: 根据后续线索二叉树的定义, X 结点为叶子结点且有左兄弟, 那么这个结点为右孩子结点, 利用后续遍历的方式可知 X 结点的后继是其父结点, 即其右线索指向的是父结点。

6. C

解析: 在一棵二叉排序树中删除一个结点后再将此结点插入到二叉排序树中, 如果删除的结点是叶子结点, 那么在插入结点后, 后来的二叉排序树与删除结点之前相同。如果删除的结点不是叶子结点, 那么再插入这个结点后, 后来的二叉树可能发生变化, 不完全相同。

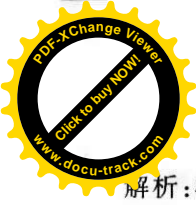
7. C

解析: 各顶点的度是矩阵中此结点对应的横行和纵列非零元素之和。

8. D

解析: D 选项是深度优先遍历不是广度优先遍历的顺序。

9. C



解析:根据 AOE 网的定义可知,关键路径上的活动时间同时减少,可以缩短工期。

10. A

解析:一棵高度为 2 的 5 阶 B 树,根结点只有到达 5 个关键字的时候才能产生分裂,成为高度为 2 的 B 树。

11. C

解析:基数排序的第 1 趟排序是按照个位数字来排序的,第 2 趟排序是按照十位数字的大小进行排序的,答案是 C 选项。

12. C

解析:基准程序的 $CPI = 2 \times 0.5 + 3 \times 0.2 + 4 \times 0.1 + 5 \times 0.2 = 3$ 。计算机的主频为 1.2 GHz,为 1 200 MHz,该机器的 MIPS 为 $1\,200/3 = 400$ 。

13. A

解析:IEEE 754 单精度浮点数格式为 C640 0000H,二进制格式为 1100 0110 0100 0000 0000 0000 0000 0000,转换为标准的格式为:

S	阶码	尾数
1	1000 1100	100 0000 0000 0000 0000 0000

因此,浮点数的值为 -1.5×2^{13} 。

14. A

解析:将 x 左移一位, y 右移一位,两个数的补码相加的机器数为 1 1000000,答案选择 A。

15. C

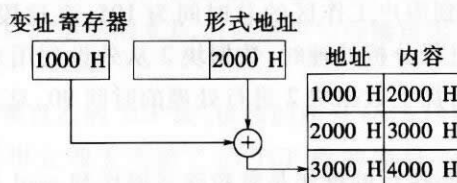
解析:设校验位的位数为 k ,数据位的位数为 n ,应满足下述关系: $2^k \geq n + k + 1$ 。 $n = 8$,当 $k = 4$ 时, $2^4 (=16) > 8 + 4 + 1 (=13)$,符合要求,校验位至少是 4 位。

16. A

解析:虚拟地址为 03FF F180H,其中页号为 03FFFH,页内地址为 180H,根据题目中给出的页表项可知页标记为 03FFFH 所对应的页框号为 0153H,页框号与页内地址之和即为物理地址 015 3180 H。

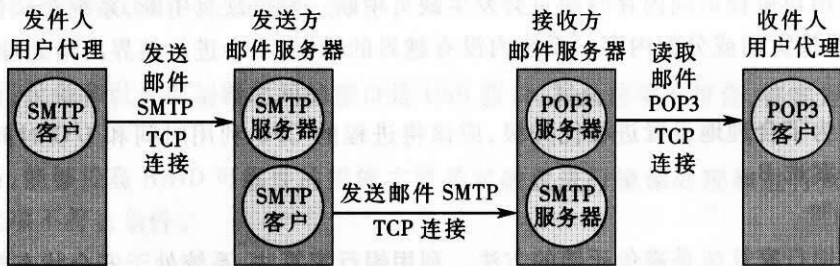
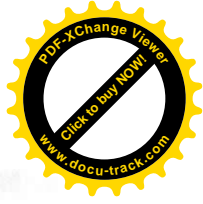
17. D

解析:根据变址寻址的主要方法,变址寄存器的内容与形式地址的内容相加之后,得到操作数的实际地址,根据实际地址访问内存,获取操作数 4000H。



18. C

解析:采用 4 级流水执行 100 条指令,在执行过程中共用 $4 + (100 - 1) = 103$ 个时钟周期。CPU 的主频是 1.03 GHz,也就是说每秒钟有 1.03 G 个时钟周期。流水线的吞吐率为



二、综合应用题

41. 【答案要点】

(1) 给出算法的基本设计思想:(4 分)

算法的策略是从前向后扫描数组元素,标记出一个可能成为主元素的元素 Num 。然后重新计数,确认 Num 是否是主元素。

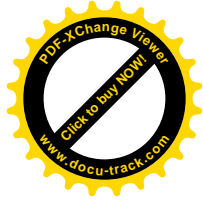
算法可分为以下两步:

- ① 选取候选的主元素:依次扫描所给数组中的每个整数,将第一个遇到的整数 Num 保存到 c 中,记录 Num 的出现次数为 1;若遇到的下一个整数仍等于 Num ,则计数加 1,否则计数减 1;当计数减到 0 时,将遇到的下一个整数保存到 c 中,计数重新记为 1,开始新一轮计数,即从当前位置开始重复上述过程,直到扫描完全部数组元素。
- ② 判断 c 中元素是否是真正的主元素:再次扫描该数组,统计 c 中元素出现的次数,若大于 $n/2$,则为主元素;否则,序列中不存在主元素。

(2) 算法实现:(7 分)

```
int Majority( int A[ ],int n )
```

```
{
    int i,c,count=1;           // c 用来保存候选主元素,count 用来计数
    c = A[0];                  // 设置 A[0]为候选主元素
    for ( i=1;i<n;i++)         // 查找候选主元素
        if ( A[i] == c )
            count++;            // 对 A 中的候选主元素计数
        else
            if ( count > 0 )    // 处理不是候选主元素的情况
                count--;
            else                // 更换候选主元素,重新计数
                { c = A[i];
                  count = 1;
                }
    if ( count>0 )
        for ( i=count=0;i<n;i++) // 统计候选主元素的实际出现次数
            if ( A[i] == c )
                count++;
}
```

```
if ( count > n/2 ) return c;    // 确认候选主元素
else return -1;                // 不存在主元素
}
```

【(1)、(2)的评分说明】

- ① 若考生设计的算法满足题目的功能要求且正确,则(1)、(2)根据所实现算法的效率给分,细则见下表:

时间复杂度	空间复杂度	(1)得分	(2)得分	说明
$O(n)$	$O(1)$	4	7	
$O(n)$	$O(n)$	4	6	如采用计数排序思想,见表后 Majority1 程序
$O(n\log_2 n)$	其他	3	6	如采用其他排序的思想
$\geq O(n^2)$	其他	3	5	其他方法

```
int Majority1( int A[],int n )// 采用计数排序思想,时间: $O(n)$ ,空间: $O(n)$ 
{
    int k, *p,max;
    p = (int *) malloc( sizeof(int) * n);    // 申请辅助计数数组
    for ( k=0; k < n; k++) p[k] = 0;        // 计数数组清 0
    max = 0;
    for ( k=0; k<n; k++)
    {p[ A[k] ] ++;                          // 计数器+1
      if ( p[ A[k] ] > p[ max ] )max = A[k]; // 记录出现次数最多的元素
    }
    if ( p[ max ] > n/2 ) return max;
    else return -1;
}
```

- ② 若在算法的基本设计思想描述中因文字表达没有非常清晰反映出算法思路,但在算法实现中能够清晰看出算法思想且正确的,可参照①的标准给分。
- ③ 若算法的基本设计思想描述或算法实现中部分正确,可参照①中各种情况的相应给分标准酌情给分。
- ④ 参考答案中只给出了使用 C 语言的版本,使用 C++或 Java 语言的答案视同使用 C 语言。

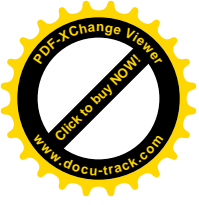
(3) 说明算法复杂性:(2分)

参考答案中实现的程序的时间复杂度为 $O(n)$,空间复杂度为 $O(1)$ 。

【评分说明】若考生所估计的时间复杂度与空间复杂度与考生所实现的算法一致,可各给 1 分。

42. 【答案要点】

- (1) 采用顺序存储结构,数据元素按其查找概率降序排列。(2分)



采用顺序查找方法。(1分)

查找成功时的平均查找长度 = $0.35 \times 1 + 0.35 \times 2 + 0.15 \times 3 + 0.15 \times 4 = 2.1$ 。
(2分)

(2)

【答案一】

采用链式存储结构,数据元素按其查找概率降序排列,构成单链表。(2分)

采用顺序查找方法。(1分)

查找成功时的平均查找长度 = $0.35 \times 1 + 0.35 \times 2 + 0.15 \times 3 + 0.15 \times 4 = 2.1$ 。(2分)

【答案二】

采用二叉链表存储结构,构造二叉排序树,元素存储方式见下图。(2分)



采用二叉排序树的查找方法。(1分)

查找成功时的平均查找长度 = $0.15 \times 1 + 0.35 \times 2 + 0.35 \times 2 + 0.15 \times 3 = 2.0$ 。(2分)

【(1)、(2)的评分说明】

- ① 若考生以实际元素表示“降序排列”,同样给分。
- ② 若考生正确求出与其查找方法对应的查找成功时的平均查找长度,给2分;若计算过程正确,但结果错误,给1分。
- ③ 若考生给出其他更高效的查找方法且正确,可参照评分标准给分。

43. 【答案要点】

(1) CPU 的时钟周期为: $1/800 \text{ MHz} = 1.25 \text{ ns}$ 。(1分)

总线的时钟周期为: $1/200 \text{ MHz} = 5 \text{ ns}$ 。(1分)

总线带宽为: $4 \text{ B} \times 200 \text{ MHz} = 800 \text{ MB/s}$ 或 $4 \text{ B}/5 \text{ ns} = 800 \text{ MB/s}$ 。(1分)

(2) Cache 块大小是 32 B,因此 Cache 缺失时需要一个读突发传送总线事务读取一个主存块。(1分)

(3) 一次读突发传送总线事务包括一次地址传送和 32 B 数据传送:用 1 个总线时钟周期传输地址;每隔 $40 \text{ ns}/8 = 5 \text{ ns}$ 启动一个体工作(各进行 1 次存取),第一个体读数据花费 40 ns,之后数据存取与数据传输重叠;用 8 个总线时钟周期传输数据。读突发传送总线事务时间: $5 \text{ ns} + 40 \text{ ns} + 8 \times 5 \text{ ns} = 85 \text{ ns}$ 。(2分)

(4) BP 的 CPU 执行时间包括 Cache 命中时的指令执行时间和 Cache 缺失时带来的额外开销。命中时的指令执行时间: $100 \times 4 \times 1.25 \text{ ns} = 500 \text{ ns}$ 。(1分) 指令执行过程中 Cache 缺失时的额外开销: $1.2 \times 100 \times 5\% \times 85 \text{ ns} = 510 \text{ ns}$ 。BP 的 CPU 执行时间: $500 \text{ ns} + 510 \text{ ns} = 1010 \text{ ns}$ 。(2分)

【评分说明】

- ① 执行时间采用如下公式计算时,可酌情给分。