

2023 年全国硕士研究生招生考试

计算机学科专业基础试题

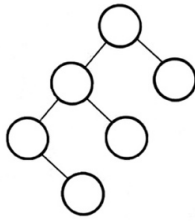
参考时间：180分钟

一、单项选择题：1~40 小题，每小题 2 分，共 80 分。下列每小题给出的四个选项中，只有一个选项是最符合题目要求的。

- 下列对顺序存储的有序表（长度为 n ）实现给定操作的算法中，平均时间复杂度为 $O(1)$ 的是
 - 查找包含指定值元素的算法
 - 插入包含指定值元素的算法
 - 删除第 i ($1 \leq i \leq n$) 个元素的算法
 - 获取第 i ($1 \leq i \leq n$) 个元素的算法
- 现有非空双向链表 L ，其结点结构为：

prev	data	next
------	------	------

，prev 是指向直接前驱结点的指针，next 是指向直接后继结点的指针。若要在 L 中指针 p 所指向的结点（非尾结点）之后插入指针 s 指向的新结点，则在执行了语句序列“ $s \rightarrow next = p \rightarrow next$; $p \rightarrow next = s$;”后，下列语句序列中还需要执行的是
 - $s \rightarrow next \rightarrow prev = p$; $s \rightarrow prev = p$;
 - $p \rightarrow next \rightarrow prev = s$; $s \rightarrow prev = p$;
 - $s \rightarrow prev = s \rightarrow next \rightarrow prev$; $s \rightarrow next \rightarrow prev = s$;
 - $p \rightarrow next \rightarrow prev = s \rightarrow prev$; $s \rightarrow next \rightarrow prev = p$;
- 若采用三元组表存储结构存储稀疏矩阵 M ，则除三元组表外，下列数据中还需要保存的是
 - M 的行数
 - M 中包含非零元素的行数
 - M 的列数
 - M 中包含非零元素的列数
 - I、III
 - I、IV
 - II、IV
 - I、II、III、IV
- 在由 6 个字符组成的字符集 S 中，各字符出现的频次分别为 3,4,5,6,8,10，为 S 构造的哈夫曼编码的加权平均长度为
 - 2.4
 - 2.5
 - 2.67
 - 2.75
- 已知一棵二叉树的树形如下图所示，若其后序遍历序列为 f, d, b, e, c, a ，则其先（前）序遍历序列是



A. a, e, d, f, b, c

B. a, c, e, b, d, f

C. c, a, b, e, f, d

D. d, f, e, b, a, c

6. 已知无向连通图 G 中各边的权值均为 1。下列算法中，一定能够求出图 G 中从某顶点到其余各顶点最短路径的是

I. 普里姆(Prim)算法

II. 克鲁斯卡尔(Kruskal)算法

III. 图的广度优先搜索算法

A. 仅I

B. 仅III

C. I、II

D. I、II、III

7. 下列关于非空 B 树的叙述中，正确的是

I. 插入操作可能增加树的高度

II. 删除操作一定会导致叶结点的变化

III. 查找某关键字总是要查找到叶结点

IV. 插入的新关键字最终位于叶结点中

A. 仅I

B. I、II

C. III、IV

D. I、II、IV

8. 对含 600 个元素的有序顺序表进行折半查找，关键字间的比较次数最多是

A. 9

B. 10

C. 30

D. 300

9. 现有长度为 5、初始为空的散列表 HT ，散列函数 $H(k) = (k + 4) \% 5$ ，用线性探查再散列法解决冲突。若将关键字序列 2022, 12, 25 依次插入 HT 中，然后删除关键字 25，则 HT 中查找失败的平均查找长度为

A. 1

B. 1.6

C. 1.8

D. 2.2

10. 下列排序算法中，不稳定的的是

I. 希尔排序

II. 归并排序

III. 快速排序

IV. 堆排序

V. 基数排序

A. I、II

B. II、V

C. I、III、IV

D. III、IV、V

11. 使用快速排序算法对数据进行升序排序，若经过一次划分后得到的数据序列是 68, 11, 70, 23, 80, 77, 48, 81, 93, 88，则该次划分的枢轴是

A. 11

B. 70

C. 80

D. 81

12. 若机器 M 的主频为 1.5 GHz，在 M 上执行程序 P 的指令条数为 5×10^5 ， P 的平均 CPI 为 1.2，则 P 在 M 上的指令执行速度和用户 CPU 时间分别为

寄存器编号。

I1	add s2, s1, s0	// $R[s2] \leftarrow R[s1] + R[s0]$
I2	load s3, 0(s2)	// $R[s3] \leftarrow M[R[s2] + 0]$
I3	beq t2, s3, L1	// if $R[t2] = R[s3]$ jump to L1
I4	addi t2, t2, 20	// $R[t2] \leftarrow R[t2] + 20$
I5	L1: ...	

若采用转发（旁路）技术处理数据冒险，采用硬件阻塞方式处理控制冒险，则在指令 I1~I4 执行过程中，发生流水线阻塞的指令有

- A. 仅 I3 B. I2、I4 C. I3、I4 D. I2、I3、I4
20. 某存储器总线宽度为 64 b，总线时钟频率为 1 GHz，在总线上传输一个数据或地址需要一个时钟周期，不支持突发传送方式。若通过该总线连接 CPU 和主存，主存每次准备一个 64 b 数据需要 6 ns，主存块大小为 32 B，则读取一个主存块需要的时间是
- A. 8 ns B. 11 ns C. 26 ns D. 32 ns
21. 下列关于硬件和异常/中断关系的叙述中，错误的是
- A. CPU 在执行一条指令过程中检测异常事件
B. CPU 在执行完一条指令时检测中断请求信号
C. 开中断时 CPU 检测到中断请求后就进行中断响应
D. 外部设备通过中断控制器向 CPU 发中断结束信号
22. 下列关于 I/O 控制方式的叙述中，错误的是
- A. 查询方式下，通过 CPU 执行查询程序进行 I/O 操作
B. 中断方式下，通过 CPU 执行中断服务程序进行 I/O 操作
C. DMA 方式下，通过 CPU 执行 DMA 传送程序进行 I/O 操作
D. 对于 SSD、网络适配器等高速设备，采用 DMA 方式输入/输出
23. 与宏内核操作系统相比，下列特征中，微内核操作系统具有的是
- I. 较好的性能 II. 较高的可靠性
III. 较高的安全性 IV. 较强的可扩展性
- A. II、IV B. I、II、III C. I、III、IV D. II、III、IV
24. 在操作系统内核中，中断向量表适合采用的数据结构是
- A. 数组 B. 队列 C. 单向链表 D. 双向链表
25. 某系统采用页式存储管理，用位图管理空闲页框。若页大小为 4 KB，物理内存大小为 16 GB，则位图所占空间的大小是

A. 128 B B. 128 KB C. 512 KB D. 4 MB

26. 下列操作完成时，导致 CPU 从内核态转为用户态的是

- A. 阻塞进程 B. 执行 CPU 调度
C. 唤醒进程 D. 执行系统调用

27. 下列由当前线程引起的事件或执行的操作中，可能导致该线程由执行态变为就绪态的是

- A. 键盘输入 B. 缺页异常
C. 主动出让 CPU D. 执行信号量的 *wait()* 操作

28. 对于采用虚拟内存管理方式的系统，下列关于进程虚拟地址空间的叙述中，错误的是

- A. 每个进程都有自己独立的虚拟地址空间
B. C 语言中 *malloc()* 函数返回的是虚拟地址
C. 进程对数据段和代码段可以有不同的访问权限
D. 虚拟地址空间的大小由内存和硬盘的大小决定

29. 进程 P1、P2 和 P3 进入就绪队列的时刻、优先级（值越大优先权越高）及 CPU 执行时间如下表所示。

进程名	进入就绪队列的时刻	优先级	CPU 执行时间
P1	0 ms	1	60 ms
P2	20 ms	10	42 ms
P3	30 ms	100	13 ms

若系统采用基于优先权的抢占式 CPU 调度算法，从 0 ms 时刻开始进行调度，则 P1、P2 和 P3 的平均周转时间为

- A. 60 ms B. 61 ms C. 70 ms D. 71 ms

30. 进程 R 和 S 共享数据 *data*，若 *data* 在 R 和 S 中所在页的页号分别为 p1 和 p2，两个页所对应的页框号分别为 f1 和 f2，则下列叙述中，正确的是

- A. p1 和 p2 一定相等，f1 和 f2 一定相等
B. p1 和 p2 一定相等，f1 和 f2 不一定相等
C. p1 和 p2 不一定相等，f1 和 f2 一定相等
D. p1 和 p2 不一定相等，f1 和 f2 不一定相等

31. 若文件 F 仅被进程 P 打开并访问，则当进程 P 关闭 F 时，下列操作中，文件系统需要完成的是

- A. 删除目录中文件 F 的目录项

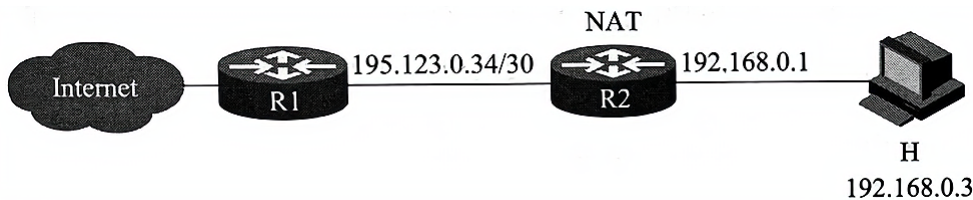
37. 若甲向乙发送数据时采用 CRC 校验，生成多项式为 $G(x) = X^4 + X + 1$ （即 $G = 10011$ ），则乙接收到下列比特串时，可以断定其在传输过程中未发生错误的是

- A. 1 0111 0000

B. 1 0111 0100
- C. 1 0111 1000

D. 1 0111 1100

38. 某网络拓扑如下图所示，其中路由器 R2 实现 NAT 功能。若主机 H 向 Internet 发送 1 个 IP 分组，则经过 R2 转发后，该 IP 分组的源 IP 地址是



- A. 195.123.0.33

B. 195.123.0.35
- C. 192.168.0.1

D. 192.168.0.3

39. 主机 168.16.84.24/20 所在子网的最小可分配 IP 地址和最大可分配 IP 地址分别是

- A. 168.16.80.1, 168.16.84.254

B. 168.16.80.1, 168.16.95.254
- C. 168.16.84.1, 168.16.84.254

D. 168.16.84.1, 168.16.95.254

40. 下列关于 IPv4 和 IPv6 的叙述中，正确的是

- I. IPv6 地址空间是 IPv4 地址空间的 96 倍
- II. IPv4 首部和 IPv6 基本首部的长度均可变
- III. IPv4 向 IPv6 过渡可以采用双协议栈和隧道技术
- IV. IPv6 首部的 Hop Limit 字段等价于 IPv4 首部的 TTL 字段

- A. I、II

B. I、IV

C. II、III

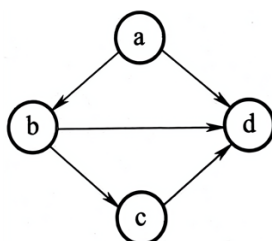
D. III、IV

二、综合应用题：41~47 小题，共 70 分。

41. (13 分)已知有向图 G 采用邻接矩阵存储，类型定义如下：

```
typedef struct                                // 图的类型定义
{
    int numVertices, numEdges;                // 图的顶点数和有向边数
    char VerticesList[ MAXV ];                // 顶点表，MAXV 为已定义常量
    int Edge[ MAXV ][ MAXV ];                // 邻接矩阵
} MGraph;
```

将图中出度大于入度的顶点称为 K 顶点。例如题 41 图中，顶点 a 和 b 为 K 顶点。



题 41 图

请设计算法：int printVertices (MGraph G)，对给定的任意非空有向图 G ，输出 G 中所有的 K 顶点，并返回 K 顶点的个数。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 $C++$ 语言描述算法，关键之处给出注释。

42. (10 分)对含有 $n(n > 0)$ 个记录的文件进行外部排序，采用置换-选择排序生成初始归并段时需要使用一个工作区，工作区中能保存 m 个记录。请回答下列问题。

(1) 若文件中含有 19 个记录，其关键字依次是 51,94,37,92,14,63,15,99,48,56,23,60,31,17,43,8,90,166,100，则当 $m = 4$ 时，可生成几个初始归并段？各是什么？

(2) 对任意的 $m(n \gg m > 0)$ ，生成的第一个初始归并段的长度最大值和最小值分别是多少？

43. (14 分)已知计算机 M 字长为 32 位，按字节编址，采用请求调页策略的虚拟存储管理方式，虚拟地址为 32 位，页大小为 4 KB；数据 Cache 采用 4 路组相联映射方式，数据区大小为 8 KB，主存块大小为 32 B。现有 C 语言程序段如下：

```
int a[24][64];  
...  
for(i = 0; i < 24; i++)  
    for(j = 0; j < 64; j++) a[i][j] = 10;
```

已知二维数组 a 按行优先存放，在虚拟地址空间中分配的起始地址为 0042 2000H，sizeof(int)=4，假定在 M 上执行上述程序段之前数组 a 不在主存，且在该程序段执行过程中不会发生页面置换。请回答下列问题：

- (1) 数组 a 分布在几个页面中？对于数组 a 的访问，会发生几次缺页异常？页故障地址各是什么？
- (2) 不考虑变量 i 和 j ，该程序段的数据访问是否具有时间局部性？为什么？
- (3) 计算机 M 的虚拟地址(A31~A0)中哪几位用作块内地址？哪几位用作 Cache 组号？ $a[1][0]$ 的虚拟地址是多少？其所在主存块对应的 Cache 组号是多少？
- (4) 数组 a 占用多少主存块？假设上述程序段执行过程中数组 a 的访问不会和其他数据发生 Cache 访问冲突，则数组 a 的 Cache 命中率是多少？若将循环中 i 和 j 的次序按如下方式调换：

```

for(j = 0; j < 64; j++)
    for(i = 0; i < 24; i++) a[i][j] = 10;

```

则数组 a 的 Cache 命中率又是多少？

44. (9分)题 43 中的 C 程序段在计算机 M 上的部分机器级代码如下，每个机器级代码行中依次包含指令序号、虚拟地址、机器指令和汇编指令。

for (i = 0; i < 24; i++)			
1	00401072	C7 45 F8 00 00 00 00	mov [ebp - 8], 0
2	00401079	EB 09	jmp 00401084h
3	0040107B	8B 55 F8	mov eax , [ebp - 8]

7	00401088	7D 32	jge 004010bch
for (j = 0; j < 64; j++)			
8	0040108A	C7 45 FC 00 00 00 00	mov [ebp - 4] , 0

a[i][j] = 10;			

19	004010AE	C7 84 82 00 20 42 00 0A 00	mov [ecx + edx * 4 + 00422000h] , 0Ah
		00 00	
20

请回答下列问题。

- (1) 第 20 条指令的虚拟地址是多少？
- (2) 已知第 2 条 `jmp` 和第 7 条 `jge` 都是跳转指令，其操作码分别是 `EBH` 和 `7DH`，跳转目标地址分别为 `0040 1084H`、`0040 10BCH`，这两条指令都采用什么寻址方式？给出第 2 条指令 `jmp` 的跳转目标地址计算过程。

- (3) 已知第 19 条 mov 指令的功能为“ $a[i][j] \leftarrow 10$ ”，其中 ecx 和 edx 为寄存器名，0042 2000H 是数组 a 的首地址，指令中源操作数采用什么寻址方式？已知 edx 中存放的是变量 j，ecx 中存放的是什么？根据该指令的机器码判断计算机 M 采用的是大端还是小端方式。
- (4) 第一次执行第 19 条指令时，取指令过程中是否会发生缺页异常？为什么？

45. (7 分)现要求学生使用 swap 指令和布尔型变量 lock 实现临界区互斥。lock 为线程间共享的变量，lock 的值为 TRUE 时线程不能进入临界区，为 FALSE 时线程能够进入临界区。某同学编写的实现临界区互斥的伪代码如题 45(a)图所示。

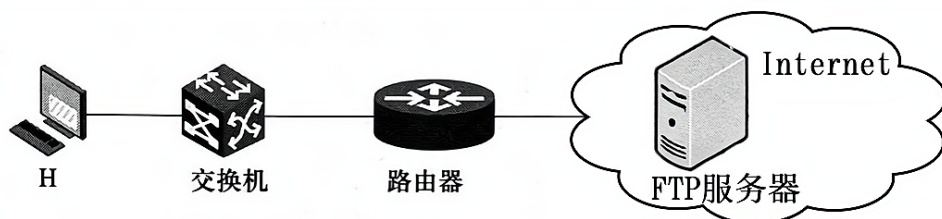
某同学编写的伪代码		newSwap()的代码
<pre> bool lock = FALSE; // 共享变量 bool key = TRUE; if (key == TRUE) swap key, lock; // 交换 key 和 lock 的值 临界区; lock = TRUE; 退出区 </pre>		<pre> void newSwap(bool *a,bool *b) { bool temp = *a; *a = *b; *b = temp; } </pre>

请回答下列问题。

- (1) 题 45(a)图的伪代码中哪些语句存在错误？将其改为正确的语句(不增加语句的条数)。
- (2) 题 45(b)图给出了交换两个变量值的函数 newSwap() 的代码，是否可以用函数调用语句“newSwap (&key, &lock)”代替指令“swap key, lock”以实现临界区互斥？为什么？
46. (8 分)进程 P 通过执行系统调用从键盘接收一个字符的输入，已知此过程中与进程 P 相关的操作包括：①将进程 P 插入就绪队列；②将进程 P 插入阻塞队列；③将字符从键盘控制器读入系统缓冲区；④启动键盘中断处理程序；⑤进程 P 从系统调用返回；⑥用户在键盘上输入字符。以上编号①~⑥仅用于标记操作，与操作的先后顺序无关。请回答下列问题。

- (1) 按照正确的操作顺序，操作①的前一个和后一个操作分别是上述操作中的哪一个？操作⑥的后一个操作是上述操作中的哪一个？
- (2) 在上述哪个操作之后 CPU 一定从进程 P 切换到其他进程？在上述哪个操作之后 CPU 调度程序才能选中进程 P 执行？
- (3) 完成上述哪个操作的代码属于键盘驱动程序？
- (4) 键盘中断处理程序执行时，进程 P 处于什么状态？CPU 处于内核态还是用户态？

47. (9 分)某网络拓扑如题 47 图所示，主机 H 登录 FTP 服务器后，向服务器上传一个大小为 18 000 B 的文件 F。假设 H 为传输 F 建立数据连接时，选择的初始序号为 100，MSS=1 000 B，拥塞控制初始阈值为 4 MSS，RTT = 10 ms，忽略 TCP 段的传输时延；在 F 的传输过程中，H 均以 MSS 段向服务器发送数据，且未发生差错、丢包和乱序现象。



请回答下列问题。

- (1) FTP 的控制连接是持久的还是非持久的？FTP 的数据连接是持久的还是非持久的？H 登录 FTP 服务器时，建立的 TCP 连接是控制连接还是数据连接？
- (2) H 通过数据连接发送 F 时，F 的第 1 个字节的序号是多少？在断开数据连接过程中，FTP 服务器发送的第二次挥手 ACK 段的确认序号是多少？
- (3) H 通过数据连接发送 F 的过程中，当 H 收到确认序号为 2101 的确认段时，H 的拥塞窗口调整为多少？收到确认序号为 7101 的确认段时，H 的拥塞窗口调整为多少？
- (4) H 从请求建立数据连接开始，到确认 F 已被服务器全部接收为止，至少需要多长时间？期间应用层数据平均发送速率是多少？