

2020 年全国硕士研究生招生考试 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。

1、将一个 10×10 对称矩阵 M 的上三角部分的元素 $m_{ij}(1 \leq i \leq j \leq 10)$ 按列优先存入 C 语言的一位数组 N 中，元素 $m_{7,2}$ 在 N 中的下标是：

A、15 B、16 C、22 D、23

参考答案：C

答案解析：「数据结构」考研复习攻略出处：<http://www.noobdream.com/Major/article/187/>

按上三角存储， $m_{7,2}$ 对应的是 $m_{2,7}$ ，在它之前有：

第 1 列：1 第 2 列：2

.....

第 6 列：6 第 7 列：1

前面一共 $1+2+3+4+5+6+1$ 个元素，共 22 个元素，数组下标从 0 开始，故下标为 $m_{2,7}$ 的数组下标为 22。

2、对空栈 S 进行 Push 和 pop 操作，入栈序列 a,b,c,d,e 经过 Push,Push,Pop,Push,Pop,Push,Push,Pop 操作后得到的出栈序列是：

A、b,a,c B、b,a,e C、b,c,a D、b,c,e

参考答案：D

答案解析：「数据结构」考研复习攻略出处：<http://www.noobdream.com/Major/article/183/>

操作	执行该操作后的栈(左侧为栈底)	出栈元素
Push	a	
Push	ab	
Pop	a	b
Push	ac	
Pop	a	c
Push	ad	
Push	ade	
Pop	ad	e

3、对与任意一棵高度为 5 且有 10 个节点的二叉树，若采用顺序存储结构保存，每个结点占 1 个存储单元（仅存放结点的数据信息），则存放该二叉树需要的存储单元数量至少是：

A、31 B、16 C、15 D、10

参考答案：A

答案解析：「数据结构」考研复习攻略出处：<http://www.noobdream.com/Major/article/205/>

由于题目明确说明只存储结点数据信息，所以采用顺序存储时要用数组的下标保存结点的父子关系，所以对于这棵二叉树存储的结果就是存储了一棵五层的满二叉树，五层的满二叉树结点个数为 $1+2+4+8+16=31$ ，所以至少需要 31 个存储单元。

4、已知森林 F 及与之对应的二叉树 T ，若 F 的先根遍历序列是 a,b,c,d,e,f ，后根遍历序列是 b,a,d,f,e,c 则 T 的后遍历序列是：

A、b,a,d,f,e,c

B、b,d,f,e,c,a

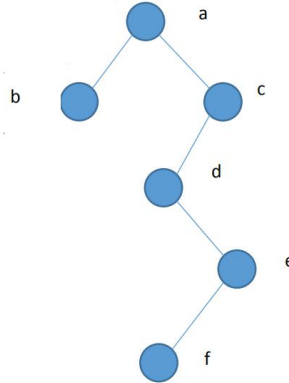
C、b,f,e,d,c,a

D、f,e,d,c,b,a

参考答案: C

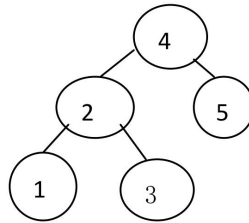
答案解析: 「数据结构」考研复习攻略出处: <http://www.noobdream.com/Major/article/210/>

森林的先根遍历对应它自己转化后二叉树的先序遍历, 森林的后根遍历对应它自己转化后二叉树的中序遍历, 所以先根和后根可以唯一确定森林转化后的二叉树, 如下:



后序遍历为: b, f, e, d, c, a

5、下列给定的关键字输入序列中, 不能生成如下二叉排序树的是:



A、4, 5, 2, 1, 3

B、4, 5, 1, 2, 3

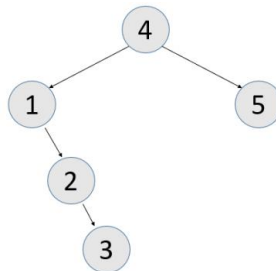
C、4, 2, 5, 3, 1

D、4, 2, 1, 3, 5

参考答案: B

答案解析: 「数据结构」考研复习攻略出处: <http://www.noobdream.com/Major/article/212/>

在 4, 5, 1, 2, 3 中由于 1 先插入, 所以 1 会成为 4 的左孩子, 2 会成为 1 的右孩子, 如下图。



6、修改递归方式实现的图的深度优先搜索 (DFS) 算法, 将输出 (访问) 定点信息的语句移到退出递归前 (即执行输出语句后立刻退出递归)。采用修改后的算法遍历有向无环图 G, 若输出结果中包含 G 中的全部顶点, 则输出的顶点序列是 G 的:

A、拓扑有序序列

B、逆拓扑有序序列

C、广度优先搜索序列

D、深度优先搜索序列

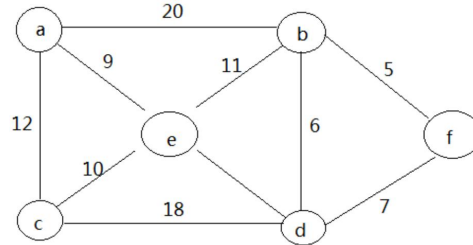
参考答案: B

答案解析: 「数据结构」考研复习攻略出处: <http://www.noobdream.com/Major/article/223/>

DFS 是一个递归算法, 在遍历的过程中, 先访问的点被压入栈底。拓扑有序是指如果

点 U 到点 V 有一条弧,则在拓扑序列中 U 一定在 V 之前, 深度优先算法搜索路径恰恰是一条弧, 栈的输出是从最后一个被访问点开始输出, 最后一个输出的点是第一个被访问的点, 所以是逆拓扑有序序列。

7、已知无向图 G 如下所示, 使用克鲁斯卡尔 (Kruskal) 算法求图 G 的最小生成树, 加入到最小生成树中的边依次是:



- A、(b,f)(b,d)(a,e)(c,e)(b,e)
- B、(b,f)(b,d)(b,e)(a,e)(e,c)
- C、(a,e)(b,e)(c,e)(b,d)(b,f)
- D、(a,e)(c,e)(b,e)(b,f)(b,d)

参考答案: A

答案解析: 「数据结构」考研复习攻略出处: <http://www.noobdream.com/Major/article/211/>
先将所有边按权值排序, 然后依次取权值最小的边但不能在图中形成环, 此时取得权值序列为 5, 6, 此时 7 不能取因为形成了环, 接下来去 9, 10, 11, 按权值对应的边。

8、若使 AOE 网估算工程进度则下列叙述中正确的是:

- A、关键路径是从原点到汇点边数最多的一条路径;
- B、关键路径是从原点到汇点路径长度最长的路径;
- C、增加任一关键活动的时间不会延长工程的工期;
- D、缩短任一关键活动的时间将会缩短工程的工期。

参考答案: B

答案解析: 「数据结构」考研复习攻略出处: <http://www.noobdream.com/Major/article/224/>
A 改为权值之和最大的路径, B 的最长就是指权值之和最大, C 增加关键活动一定会增加工期, D 减小一关键活动不一定会缩短工期

9、下列关于大根堆 (至少含 2 个元素) 的叙述中正确的是:

- I. 可以将堆看成一颗完全二叉树; II、可采用顺序存储方式保存堆;
 - III、可以将堆看成一棵二叉排序树; IV、堆中的次大值一定在根的下一层。
- A、I II III B、II III IV C、I II IV D、I II III IV

参考答案: C

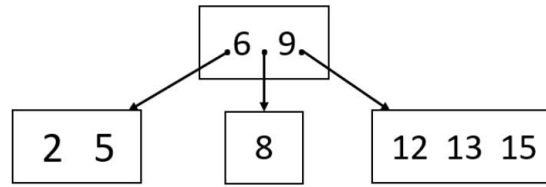
答案解析: 「数据结构」考研复习攻略出处: <http://www.noobdream.com/Major/article/241/>
III 错误, 因为堆只要求根大于左右子树, 并不要求左右子树有序。

10、依次将关键字 5, 6, 9, 13, 8, 2, 12, 15 插入初始为空的 4 阶 B 树后, 根节点中包含的关键字是:

- A、8 B、6, 9 C、8, 13 D、9, 12

参考答案: B

答案解析: 「数据结构」考研复习攻略出处: <http://www.noobdream.com/Major/article/241/>
根据 B 树的性质, 最后生成的 B 树如下图所示:



11、对大部分元素已有序的数组进行排序时，直接插入排序比简单选择排序效率更高，其原因是：

- I、直接插入排序过程中元素之间的比较次数更少；
- II、直接插入排序过程中所需要的辅助空间更少；
- III、直接插入排序过程中元素的移动次数更少。

A、I B、III C、I,II D、I,II,III

参考答案：A

答案解析：「数据结构」考研复习攻略出处：<http://www.noobdream.com/Major/article/238/>

直接插入排序在有序数组上的比较次数为 $n-1$ ，简单选择排序的比较次数为 $1+2+\dots+n-1=n(n-1)/2$ 。II，辅助空间都是 $O(1)$ ，没差别，III，因为本身已经有序，移动次数均为 0。

12、下列给出的部件中其位数（宽度）一定与机器字长相同的是：

- I、ALU II、指令寄存器 III、通用寄存器 IV、浮点寄存器

A、I,II B、I,III C、II,III D、II,III,IV

参考答案：B

答案解析：机器字长：是指计算机进行一次整数运算所能处理的二进制的位数，通常与 CPU 的寄存器位数，加法器有关。

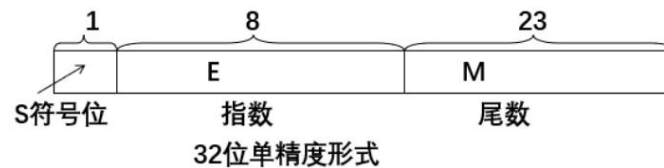
13、已知带符号整数用补码表示，float 型数据用 IEEE 754 标准表示，假定变量 x 的类型只能是 int 或 float。当 x 的机器数为 C800 0000H 时，x 的值可能是：

- A、 -7×2^{27} ； B、 -2^{16} ； C、 2^{17} D、 25×2^{27} ；

参考答案：A

答案解析：C800 0000H = 1100 1000 0000 0000 0000 0000 0000 0000H

若为 float 转化为 1 100 1000 0 000 0000 0000 0000 0000 0000H = -2^{17}



若为 int，则对 C800 0000H 求补，得 B800 0000H，即 -7×2^{27}

14、在按字节编址，采用小端方式的 32 位计算机中，按边界对齐方式为以下 C 语言结构型变量 a 分配存储空间。

```
1. Struct record{
2.     short x1;
3.     int x2;
4. } a;
```

若 a 的首地址为 2020 FE00H，a 的成员变量 x2 的机器数为 1234 0000H，则其中 34H 所在存储单元的地址是：

- A、2020 FE03H； B、2020 FE04H； C、2020 FE05H； D、2020 FE06H；

参考答案：D

答案解析：根据按边界对齐和小端方式的定义（王道书第 2 章和第 4 章的常见问题分别给出了大小端存放和按边界对其的定义及举例），给出变量 a 的存放方式如下：

x1 (LSB)	x2 (MSB)	null	null
00H	00H	34H	12H

首地址为 2020 FE00H, 按字节编址, 则 34H 所在单元地址为 2020 FE00H +6= 2020 FE06H。

15、下列关于 TLB 和 Cache 的叙述中错误的是:

- A、命中率与程序局部性有关; B、缺失后都需要去访问主存;
C、缺失处理都可以由硬件实现; D、都由 DRAM 存储器组成。

参考答案: D

答案解析: Cache 由 SRAM 组成。TLB 通常用相联存储器组成, 也可以由 SRAM。DRAM 需要不断的刷新, 性能较低, 肯定不能选

16、某计算机采用 16 位定长指令字格式, 操作码位数和寻址方式位数固定, 指令系统有 48 条指令, 支持直接、间接、立即、相对 4 种寻址方式, 单地址指令中直接寻址方式可寻址范围是:

- A、0~225; B、0~1023; C、-128~127; D、-512~511;

参考答案: A

答案解析: 48 条指令需要 6 位操作码字段, 4 种寻址方式需要 2 位寻址特征位, 故寻址范围为 0~255。注意, 主存地址不能为负数。

17、下列给出的处理器类型中理想情况下 CPI 为 1 的是:

- I、单周期 CPU; II、多周期 CPU; III、基本流水线 CPU; IV 超标量流水线 CPU
A、I, II; B、II,III; C、II,IV; D、III,IV;

参考答案: B

答案解析: II 肯定错误, IV 是通过增加功能部件实现的并行。

在理想情况下, I 单周期 CPU, 令指令周期=时钟周期; III 基本流水线 CPU, 让每个时钟周期流出一条指令 (执行完一条指令)。

18、下列关于“自陷”(Trap, 也称陷阱)的叙述中错误的是:

- A、自陷是通过陷阱指令预先设定的一类外部中断事件;
B、自陷可用于实现程序调试时的断点设置和单步跟踪;
C、自陷发生后 CPU 将转去执行操作系统内核相应程序;
D、自陷处理完成后返回到陷阱指令的下一条指令执行。

参考答案: A

答案解析: 自陷是属于内中断。

19、QPI 总线是一种点对点全工双周同步串行总线, 总线上的设备可同时接收和发送信息, 每个方向可同时传输 20 位信息 (16 位数据+4 位校验位), 每个 QPI 数据包有 80 位信息, 分 2 个时钟周期传送, 每个时钟周期传递 2 次, 因此 QPI 总线带宽为每秒传送次数*2B*2。若 QPI 时钟频率为 2.4GHz, 则总线带宽为:

- A、4.8 B、9.6 C、19.2 D、38.4 (单位 GB/s)

参考答案: C

答案解析: 2.4G(频率) * 2(每个时钟周期传递两次) * 2(全双工) * 2B/S= 19.2GB/s

20、下列事件中属于外部中断事件的是:

- I、访存时缺页; II、定时器延时; III、网络数据包到达 选项暂无
A、I II B、I III C、II III D、I II III

参考答案: C

答案解析: II 外部中断, 描述的是时钟中断; III 外部中断, 外部事件。

21、外部中断包括不可屏蔽中断 (NMI) 和可屏蔽中断, 下列关于外部中断的叙述中错误的是:

- A、CPU 处于关中断状态时也能响应 NMI 请求;
- B、一旦可屏蔽中断请求信号有效, CPU 将立即响应;
- C、不可屏蔽中断的优先级比可屏蔽中断的优先级高;
- D、可通过中断屏蔽字改变可屏蔽中断的处理优先级。

参考答案: B

答案解析: A: [关中断] 在中断服务过程中为了保护中断现场不被新的中断所打断, 在保护现场的过程中, CPU 不应响应更高级中断源的中断请求, 但不可屏蔽中断通过 NMI 控制, 不受中断标志位的影响, 即使在关中断的情况下也会被响应。 B: CPU 相应终端必须满足以下三个条件: ① 中断源有中断请求② CPU 允许中断及开中断③ 一条指令执行完毕且没有更紧迫的任务 C: 一般来说, 在中断判优过程中, 不可屏蔽中断优于可屏蔽中断 D: 正确

22、若设备采用周期挪用 DMA 方式进行输入输出, 每次 DMA 传送的数据块大小为 512 字节, 相应的 I/O 接口中有一个 32 位数数据缓冲寄存器, 对于数据输入过程, 下列叙述中错误的是:

- A、每准备好 32 位数据, DMA 控制器就发出一次总线请求;
- B、相对于 CPU, DMA 控制器的总线使用权的优先级更高;
- C、在整个数据块的传送过程中, CPU 不可以访问主存储器;
- D、数据块传送结束时, 会产生“DMA 传送结束”的中断请求。

参考答案: C

答案解析: DMA 与主存相互并行, DMA 工作期间不影响 CPU 的工作。

23、若多个进程共享同一个文件 F, 则下列叙述中正确的是:

- A、个进程只能用“读”方式打开文件 F;
- B、在系统打开文件表中仅有一个表项包含 F 的属性;
- C、各进程的用户打开文件表中关于 F 的表项内容相同;
- D、进程关闭 F 时系统删除 F 在系统打开文件表中的表项。

参考答案: B

答案解析: 既可以是读的方式, 也可以是写的方式, A 错误。系统打开文件表整个系统只有一张, 同一个文件打开多次只需要改变引用计数, 不需要对应多项, B 正确。用户进程的打开文件表关于同一个文件不一定相同, C 错误。进程关闭文件时, 文件的引用计数减少 1, 引用计数变为 0 时才删除, D 错误。

24、下列选项中支持文件长度可变, 随机访问的磁盘存储空间分配方式是:

- A、索引分配; B、链接分配; C、连续分配; D、动态分区分配。

参考答案: A

答案解析: 链接分配不能支持随机访问, B 错误。连续分配不支持可变文件长度, C 错误。动态分区分配是内存管理方式非磁盘空间管理方式, D 错误。

25、下列与中断相关的操作中, 由操作系统完成的是:

- I、保存被中断程序的中断点; II、提供中断服务; III、初始化中断向量表; IV、保存中断屏蔽字;
- A、I,II; B:I,II,IV; C III,IV; D II,III,IV.

参考答案: D

答案解析: 中断的保存硬件和软件分别都要保存部分寄存器内容, 硬件保存程序计数器 PC, 操作系统保存程序状态字 PSW, 不仅仅由操作系统单独完成, I 错误。

26、下列与进程调度有关的因素中在设计多级反馈队列调度算法时需要考虑的是:

- I 就绪队列的数量; II 就绪队列的优先级; III 各就绪队列的调度算法; IV 进程在就绪队列间的迁移条件;
- A、I,II; B、III,IV; C、II,III,IV; D I,II,III,IV

参考答案: D

答案解析: 多级反馈队列调度需要综合考虑优先级数量、优先级之间的转换规则等, I, II, III, IV 均正确。

27、某系统中有 A,B 两类资源各 6 个, t 时刻资源分配及需求情况如下表所示

进程	A 已分配数量	B 已分配数量	A 需求总量	B 需求总量
P1	2	3	4	4
P2	2	1	3	1
P3	1	2	3	4

t 时刻安全检测结果是:

- A、存在安全序列 P1,P2,P3;
B、存在安全序列 P2,P1,P3; C、存在安全序列 P2,P3,P1;
D、不存在安全序列。

参考答案: B

答案解析: 此道题作出需求矩阵 $NEED = MAX - ALLOCATED$ 即可, 如下图。

$$\begin{array}{cc} A & B \\ 4 & 4 \\ 3 & 1 \\ 3 & 4 \end{array} - \begin{array}{cc} A & B \\ 2 & 3 \\ 2 & 1 \\ 1 & 2 \end{array} = \begin{array}{cc} A & B \\ 2 & 1 \\ 1 & 0 \\ 2 & 2 \end{array}$$

同时, 由 allocated 矩阵得知当前 available 为 (1, 0)。由需求矩阵可知, 初始只能满足 p2 需求。P2 释放资源后 available 变为 (4, 1)。此时仅能满足 p1 需求, p1 释放后可以满足 p3。故得到顺序 p2→p1→p3。B 正确。

28、下列因素影响请求分页系统有效(平均)访存时间的是: I、缺页率; II、磁盘读写时间; III、内存访问时间; IV 执行缺页处理程序的 CPU 时间;
A、II,III; B、I,IV; C、I,III,IV; D、I,II,III,IV。

参考答案: D

答案解析: I 影响缺页中断发生的频率; II, IV 影响缺页中断的处理时间; II 影响访问慢表和访问目标物理地址的时间, 故 I, II, III, IV 均正确。

29、下列关于父进程与子进程的叙述中错误的是:

- A、父进程与子进程可以并发执行;
B、父进程与子进程共享虚拟地址空间;
C、父进程与子进程有不同的进程控制块;
D、父进程与子进程不能同时使用同一临界资源。

参考答案: B

答案解析: 父进程可以和子进程共享一部分共享资源, 但是不和子进程共享虚拟地址空间, 在创建子进程时, 会为子进程分配空闲的进程描述符、唯一标识的 pid 等, B 错误。

30、对于具备设备独立性的系统下列叙述中错误的是:

- A、可以使用文件名访问物理设备;
B、用户程序使用逻辑设备与物理设备之间的映射关系;
D、更换物理设备后必须修改访问该设备的应用程序。(缺一个选项)

参考答案: D

答案解析: 设备可以看作特殊文件, A 正确。B 为知识点, 正确。访问设备的驱动程序与具体设备无关, D 错误。

31、某文件系统的目录由文件名和索引节点号构成。若每个目录项长度为 64 字节, 其中 4 个字节存放索引节点号, 60 个字节存放文件名。文件名由小写英文字母构成, 则该文件系统能创建的文件数量的上限为:

- A、 2^{26} ; B、 2^{32} ; C、 2^{60} ; D、 2^{64} ;

参考答案: B

答案解析：最多创建文件个数=最多索引节点个数。由题，索引节点占 4 个字节，对应 32 位，最多可以表示 2^{32} 个文件，B 正确。

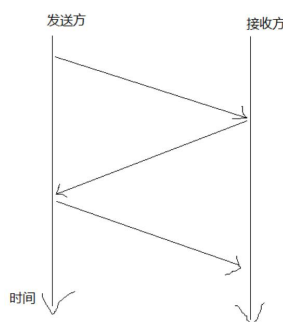
32、下列准则中实现临界区互斥机制必须遵循的是：

I、两个进程不能同时进入临界区； II、允许进城访问空闲的临界资源；
III、进程等待进入临界区的时间是有限的； IV、不能进入临界区的执行态进程立即放弃 CPU。 A、I,IV； B、II,III； C、I,II,III； D、I,III,IV；

参考答案：C

答案解析：I, II, III 分别符合互斥、空闲让进、有限等待的原则，不能立即进入临界区的进程可以选择等待部分时间，IV 错误。故 C 正确。（个人认为 IV 也对，四个原则：II 是空闲让进，I 是忙则等待，III 是有限等待，IV 是让权等待。不排除是回忆版真题有误）

33、下图描述的协议要素是



I、语法； II、语义； III、时序

A、仅 I； B、仅 II； C 仅 III； D、I,II 和 III；

参考答案：B

答案解析：协议由语法、语义和同步(时序)三部分组成。语法规定了传输数据的格式；语义规定了所要完成的功能，即需要发出何种控制信息、完成何种动作以及做出何种应答；同步规定了执行各个操作的条件、时序关系等，即时间实现顺序的详细说明。该图显然描述的是各操作的执行顺序，并未描述传输数据的格式或所要完成的功能。

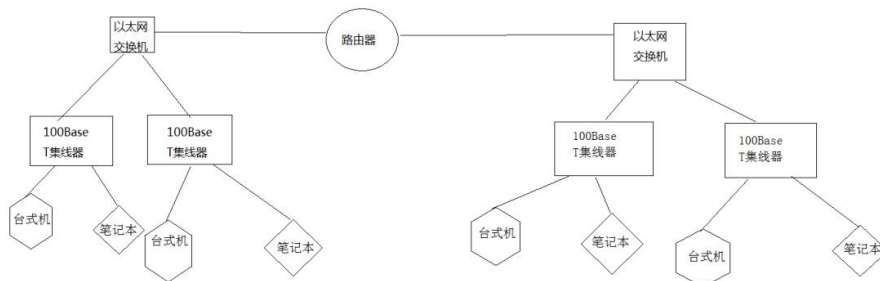
34、下列关于虚电路网络的叙述中错误的是：

A、可以确保数据分组传输顺序； B、需要为每条虚电路预分配带宽；
C、建立虚电路时需要进行路由选择； D、依据虚电路号（VCID）进行数据分组转发。

参考答案：B

答案解析：虚电路服务需要有建立连接过程，每个分组使用短的虚电路号，属于同一条虚电路的分组按照同一路由进行转发，分组到达终点的顺序与发顺顺序相同，可以保证有序传输，不需要为每条虚电路预分配带宽。

35、下图所示的网络冲突域和广播域的个数分别是：



A、2, 2; B、2, 4; C、4, 2; D、4, 4;

参考答案: C

答案解析: 网络层设备路由器可以隔离广播域和冲突域, 链路层设备普通交换机只能隔离冲突域, 物理层设备集线器、中继器既不能隔离冲突域也不能隔离广播域。题中共有 2 个广播域, 4 个冲突域。

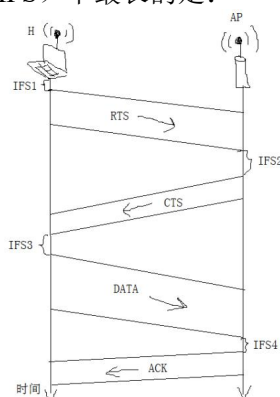
36、假设主机采用停-等协议向主机乙发送数据帧, 数据帧长与确认帧长均为 1000B。数据传输速率是 10kbps, 单项传播延时是 200ms。则甲的最大信道利用率:

A、80%; B、66.7%; C、44.4%; D、40%

参考答案: D

答案解析: 发送数据帧和确认帧的时间分别为 800ms, 800ms。发送周期为 $T=800+200+800+200=2000\text{ms}$ 。采用停止-等待协议, 信道利用率为 $800/2000=40\%$ 。

37、某 IEEE 802.11 无线局域网中主机 H 与 AP 之间发送或接收 CSMA/CA 帧的过程如下图所示, 在 H 或 AP 发送帧前所等待的帧间间隔时间 (IFS) 中最长的是:



A、IFS1; B、IFS2; C、IFS3; D、IFS4;

参考答案: A

答案解析: 为了尽量避免碰撞, 802.11 规定, 所有的站在完成发送后, 必须再等待一段很短的时间(继续监听)才能发送下一帧。这段时间通称为帧间间隔 IFS (InterFrame Space)。帧间间隔的长短取决于该站要发送的帧的类型。IEEE 802.11 推荐使用 3 种帧间间隔 (IFS), 以便提供基于优先级的访问控制。DIFS (分布式协调 IFS): 最长的 IFS, 优先级最低, 用于异步帧竞争访问的时延。PIFS (点协调 IFS): 中等长度的 IFS, 优先级居中, 在 PCF 操作中使用。SIFS (短 IFS): 最短的 IFS, 优先级最高, 用于需要立即响应的操作。网络中的控制帧以及对所接收数据的确认帧都采用 SIFS 作为发送之前的等待时延。当结点要发送数据帧时, 载波监听到信道空闲时, 需等待 DIFS 后发送 RTS 预约信道, IFS1 对应的帧间间隔 DIFS, 时间最长, 图中 IFS2, IFS3, IFS4 对应 SIFS。

38、若主机甲与主机乙已建立一条 TCP 连接, 最大段长 (MSS) 为 1KB, 往返时间 (RTT) 为 2ms, 则在不出现拥塞的前提下, 拥塞窗口从 8kB 增长到 20KB 所需的最长时间是:

A、4ms; B、8ms; C、24ms; D、48ms;

参考答案: C

答案解析: 在不出现拥塞的前提下, 拥塞窗口从 8KB 增长到 20KB 所需的最长时间 (由于慢开始门限可以根据需求设置 所以这里面为了求最长时间 可以假定在慢开始门限小于等于 8KB, 这样由 8KB-20KB 的过程中都是加法增大), 考虑拥塞窗口达到 8KB 时, 以后的每个轮次拥塞窗口逐次加 1, 需 $12 \times 2 = 24\text{ms}$ 后达到 20KB 大小。也有同学反馈: 题目是从 8k-32k, 那么答案就选 D。

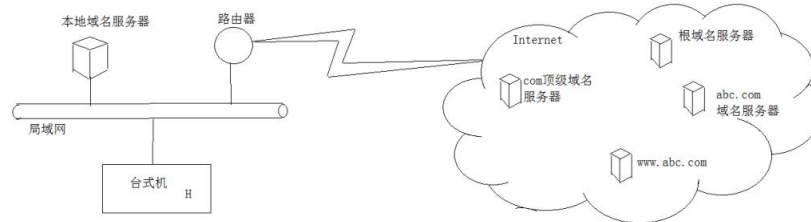
39、若主机甲与主机乙建立 TCP 连接时发送的 SYN 段中的序号为 1000, 在断开连接时, 甲发送给乙的 FIN 段中的序号为 5001, 则在无任何重传的情况下, 甲向乙已经发送的应用层数据的字节数为:

A、4002; B、4001; C、4000; D、3999;

参考答案: C

答案解析: 主机甲与主机乙建立 TCP 连接时发送的 SYN 段中的序号为 1000, 则在数据数据传输阶段所用序号起始为 1001, 在断开连接时, 甲发送给乙的 FIN 段中的序号为 5001, 在无任何重传的情况下, 甲向乙已经发送的应用层数据的字节数为 $5001-1001=4000$ 。

39、假设下图所示网络中的本地域名服务器只提供递归查询服务, 其他域名的服务器均只提供迭代查询服务; 局域网内主机访问 Internet 上各服务器的往返时间 (RTT) 均为 10ms, 忽略其他各种时延, 若主机 H 通过超链接 <http://www.abc.com/index.html>, 请求浏览纯文本 Web 页 index.html, 则从点击超链接开始到浏览器接收到 index.html 页面为止, 所需最短、最长时间分别是:



A、10ms, 40ms; B、10ms, 50ms; C、20ms, 40ms; D、20ms, 50ms;

参考答案: D

答案解析: 忽略各种时延情况下, 最短时间, 即本地域名服务器存在域名与 IP 地址映射关系, 仅需主机向本地域名服务器递归查询一次 10ms, 传送数据 10ms, 最短时间共需 20ms; 最长时间即本地域名服务器不存在域名与 IP 地址映射关系, 需向本地域名服务器递归查询一次后, 迭代查询各级域名服务器 3 次, 需 40ms, 传送数据 10ms, 最长时间共需 50ms。

二、综合应用题: 41~47 小题, 共 70 分。

41、定义三元组 (a, b, c) (a, b, c 均为正数) 的距离 $D=|a-b|+|b-c|+|c-a|$. 给定 3 个非空整数集合 S_1, S_2, S_3 , 按升序分别存储在 3 个数组中。请设计一个尽可能高效的算法, 计算并输出所有可能的三元组 (a, b, c) ($a \in S_1, b \in S_2, c \in S_3$) 中的最小距离。例如 $S_1=\{-1, 0, 9\}$, $S_2=\{-25, -10, 10, 11\}$, $S_3=\{2, 9, 17, 30, 41\}$ 。则最小距离为 2, 相应的三元组为 $(9, 10, 9)$, 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;
- (3) 说明你所设计算法的时间复杂度和空间复杂度。

参考答案:

提示: 很多同学第一反应就是暴力枚举三个集合, 然后有同学就会想优化, 比如查找的时候使用二分来加速等等。但是很多同学没有注意到这个题本质是求三个最接近的数, 而且没有集合都是有序的, 所以只需要枚举少部分可能产生最小值的情况即可。

(1) 由于三个集合已经拍好序了, 我们可以从三个集合最小的数开始, 尽量让三个集合的数的值接近甚至相等, 所以我们需要把最小值所在的集合往后遍历, 重复这个过程即可, 同学们可以自己组织语言去详细描述这个算法过程, 不理解的同学看下面的代码。

(2) 完整代码如下

```
1. #include <stdio.h>
2. const int maxn = 1005;
3. int s1[maxn], s2[maxn], s3[maxn];
4. int len_s1, len_s2, len_s3; //三个集合的长度
5. int _abs(int x, int y) { //求两个数的距离
6.     if (x - y < 0) return y - x;
```

```

7.     else return x - y;
8. }
9. int main() {
10.    scanf("%d%d%d", &len_s1, &len_s2, &len_s3); //输入 3 个集合的分别的元素个数
11.    for (int i = 1; i <= len_s1; i++) {
12.        scanf("%d", &s1[i]); //输入集合 1 的元素
13.    }
14.    for (int i = 1; i <= len_s2; i++) {
15.        scanf("%d", &s2[i]); //输入集合 2 的元素
16.    }
17.    for (int i = 1; i <= len_s3; i++) {
18.        scanf("%d", &s3[i]); //输入集合 3 的元素
19.    }
20.    int ans = 0x3f3f3f3f; //无穷大
21.    while (len_s1 >= 1 && len_s2 >= 1 && len_s3 >= 1) {
22.        int d = 0;
23.        d += _abs(s1[len_s1], s2[len_s2]); //求两个元素的距离
24.        d += _abs(s1[len_s1], s3[len_s3]);
25.        d += _abs(s2[len_s2], s3[len_s3]);
26.        if (d < ans) ans = d; //更新更小的答案
27.        if (s1[len_s1] >= s2[len_s2] && s1[len_s1] >= s3[len_s3]) len_s1--;
28.        if (s2[len_s2] >= s1[len_s1] && s2[len_s2] >= s3[len_s3]) len_s2--;
29.        if (s3[len_s3] >= s1[len_s1] && s3[len_s3] >= s2[len_s2]) len_s3--;
30.    }
31.    printf("%d\n", ans); //输出答案
32.    return 0;
33. }
34. /*
35. 输入数据
36. 3 4 5
37. -1 0 9
38. -25 -10 10 11
39. 2 9 17 30 41
40. 输出答案: 2
41. */

```

(3) 时间复杂度

将三个序列的长度分标记为 L_1, L_2, L_3 , 时间复杂度为 $O(L_1 + L_2 + L_3) = O(\max(L_1, L_2, L_3))$, 空间复杂度为 $O(1)$

42、若任一个字符的编码都不是其他字符编码的前缀, 则称这种编码具有前缀特性。现有某字符集(字符个数 ≥ 2)的不等长编码, 每个字符的编码均为二进制的 0, 1 序列, 最长为 L 位, 且具有前缀特性。请回答下列问题:

- (1) 哪种数据结构适宜保存上述具有前缀特性的不等长编码?
- (2) 基于你所设计的数据结构, 简述从 0/1 串到字符串的译码过程;
- (3) 简述判定某字符集的不等长编码是否具有前缀特性的过程。

题目解析:

- (1) 哈夫曼树。(书上哈夫曼树应用的例子就是前缀编码)
- (2) 译码过程: 从前到后依次取 0/1 串的每一位, 如果为 0 就进入左子树, 如果为 1 就进入右子树, 直到整个 0/1 串遍历结束, 遇到叶子节点, 该叶子节点代表的字符串就是该 0/1 串的字符串, 如果没有遇到叶子节点或者遇到了叶子节点但 0/1 串未遍历结束, 则该 0/1 串所代表的字符串不在该字符集中。
- (3) 若某字符集的不等长编码不能确定一颗哈夫曼树, 则说明该字符集的不等长编码不具有前缀特性

43、有实现 $x*y$ 的两个 C 语言函数如下:

```
1. unsigned umul (unsigned x, unsigned y) {
2.     return x*y;
3. }
4. int imul(int x, int y) {
5.     return x*y;
6. }
```

假定某计算机 M 中 ALU 只能进行加减运算和逻辑运算。请回答:

- (1) 若 M 的指令系统中没有乘法指令, 但有加法、减法和位移等指令, 则在 M 上也能实现上述两个函数中的乘法运算, 为什么?
- (2) 若 M 的指令系统中有乘法指令, 则基于 ALU、位移器、寄存器以及相应控制逻辑实现乘法指令时, 控制逻辑的作用是什么?
- (3) 针对以下 3 种情况: (a) 没有乘法指令; (b) 有使用 ALU 和位移器实现的乘法指令; (c) 有使用阵列乘法器实现的乘法指令, 函数 `umul()` 在哪种情况下执行时间最长? 哪种情况下执行的时间最短? 说明理由
- (4) n 位整数乘法指令可保存 $2n$ 位乘积, 当仅取低 n 位作为乘积时, 其结果可能会发生溢出。当 $n=32$, $x=2^{31}-1$, $y=2$ 时, 带符号整数乘法指令和无符号整数乘法指令得到的 $x*y$ 的 $2n$ 位乘积分别是什么(用十六进制表示)? 此时函数 `umul()` 和 `imul()` 的返回结果是否溢出? 对于无符号整数乘法运算, 当仅取乘积的低 n 位作为乘法结果时, 如何用 $2n$ 位乘积进行溢出判断?

题目解析:

- (1) 乘法运算也可以通过加法操作和移位操作实现 $x*y$ 可视为 y 个 x 或 x 个 y 相加的结果
- (2) 实现相加和移位的控制
- (3) 最长: a 最短: c

a) 情况下执行时间最长, 需要利用其他指令来实现乘法功能

b) 情况下使用了 ALU 与位移器, 由多次相加及位移操作串行实现乘法操作

c) 情况使用阵列乘法器做并行乘法运算, 时间显然最快

(4) 带符号整数指令乘法: $7FFF\ FFFFH * 2 = 0000\ 0000\ FFFF\ FFFE H$

无符号整数指令乘法: $7FFF\ FFFFH * 2 = 0000\ 0000\ FFFF\ FFFE H$

`umul()` 返回 $FFFF\ FFFE H$ 未溢出

`imul()` 返回 $FFFF\ FFFE H$ 有溢出(结果成了负数) 高 n 位全 0 则未产生溢出, 否则产生溢出

44、假定主存地址为 32 位，按字节编址，指令 Cache 和数据 Cache 与主存之间均采用 8 路组相联映射方式，直写（Write Through）写策略和 LRU 替换算法，主存块大小为 64B，数据区容量各为 32KB。开始时 Cache 均为空，请回答下列问题：

(1) Cache 每一行中标记（Tag）、LRU 位各占几位？是否有修改位？

(2) 有如下 C 语言程序段：

```
1. for(k=0; k<1024; k++)
2.     S[k]=2*s[k];
```

若数组 S 及其变量 k 均为 int 型，int 型数据占 4B，变量 k 分配在寄存器中，数组 s 在主存中的起始地址为 0080 00C0H，则该程序段执行过程中，访问数组 S 的数据 Cache 缺失次数为多少？

(3) 若 CPU 最先开始的访问操作是读取主存单元 0001 003H（或者 0001 0003H？好像记错了）中的指令，简要说明从 Cache 中访问该指令的过程，包括 Cache 缺失处理过程。

题目解析：

(1) 主存地址为 32 位，其中 6 位为块内地址

标记：32 - 6 - 6 = 20 位

LRU：3 位

直写写策略无修改位

(2) 在该程序的执行过程中，产生一次缺失时会更新 Cache 中的 16 块，之后的 15 次访问均命中 Cache，故平均每 16 次访问 Cache 就会产生一次缺失，K 的范围为 1024，即共 1024 次访问 Cache，Cache 缺失次数=1024/16=64 次

(3) CPU 访指的大致过程：

CPU 发出读请求时，若访存地址在 Cache 中命中，就将此地址转换成 Cache 地址，直接对 Cache 进行读操作，与主存无关；若 Cache 不命中，则仍需访问主存，并把此字所在的一块一次性从主存调入 Cache，若此时 Cache 已满，就根据 LRU 替换算法，用这个块替换 Cache 中的一块信息。

45、现有 5 个操作 A、B、C、D 和 E，操作 C 必须在 A 和 B 完成后执行，操作 E 必须在 C 和 D 完成后执行，请使用信号量的 wait()，signal()，操作（P、V 操作）描述上述操作之间的同步关系，并说明所用信号量及其初值。

题目解析：

本题要求实现操作的先后顺序，属同步问题。

分别设置信号量 ABCDE 对应 5 个操作，初值都为 0。5 个进程的动作可描述为：

```
1. A(){
2.     完成动作 A;
3.     V(A);
4. }
5. B(){
6.     完成动作 B;
7.     V(B);
8. }
9. C(){
10.    //C 必须在 A、B 都完成后才能完成
```



```

11.    P(A);
12.    P(B);
13.    完成动作 C;
14.    V(C);
15. }
16. D(){
17.    完成动作 D;
18.    V(D);
19. }
20. E(){
21.    //E 必须在完成 C、D 之后执行
22.    P(C);
23.    P(D)
24.    完成动作 E;
25.    V(E);
26. }

```

46、某 32 位系统采用基于二级页表的请求分页存储管理方式，按字节编址，页目录项和页表项长度均为 4 字节，虚拟地址结构如下：

页目录号（10 位）	页号（10 位）	页内偏移量（12 位）
------------	----------	-------------

某 C 程序中数组 $a[1024][1024]$ 的起始虚拟地址为 1080 0000H，数组元素占 4 字节，该程序运行时，其进程的页目录起始物理地址为 0020 1000H，请回答下列问题：

（1）数组元素 $a[1][2]$ 的虚拟地址是什么？对应的页目录号和页号分别是什么？对应的页目录项的物理地址是什么？若该目录项中存放的页框号为 00301H，则 $a[1][2]$ 所在页对应的页表项的物理地址是什么？

（2）数组 a 在虚拟地址空间中所占区域是否必须连续？在物理地址空间中所占区域是否必须连续？

（3）已知数组 a 按行优先方式存放，若对数组 a 分别按行遍历和按列遍历，则哪一种遍历方式的局部性更好？

题目解析：

（1）①页面大小 $= 2^{12} \text{B} = 4096 \text{B} = 4 \text{KB}$ 。每个数组元素 4B，每个页面可以存放 $4 \text{KB} / 4 \text{B} = 1024$ 个数组元素，刚好是数组的一行。1080 000H 的虚页号为 1080H，因此 $a[0]$ 行存放在虚页号为 1080H 的页面中， $a[1]$ 行存放在页号为 1081H 的页面中。 $a[1][2]$ 的虚拟地址为 $1081000 \text{H} + 4 \times 2 = 1081008 \text{H}$ 。

②转换为二进制 0001000010000001 000000001000，根据虚拟地址结构可知，对应的页目录号为 4，页号为 129。

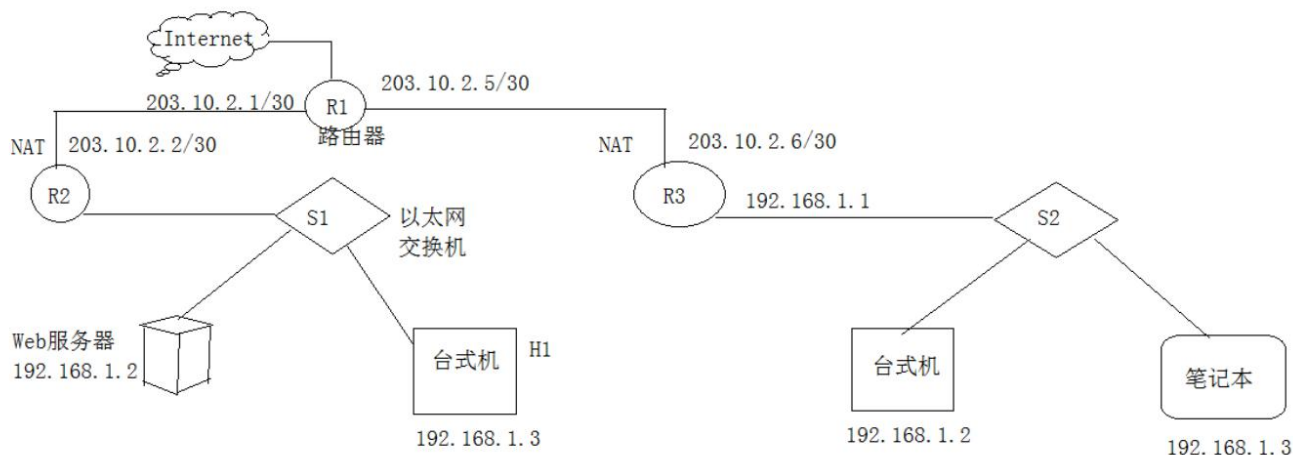
③进程的页目录表起始地址为 0020 1000H，每个目录项长 4B，因此 4 号页目录项的物理地址是 $00201000 \text{H} + 4 \times 4 = 00201010 \text{H}$ 。

④页目录项存放的页框号为 00301H，该页框的起始地址为 00301000H，因此 $a[1][2]$ 所在页的页号为 129，每个页表项 4B，因此对应的页表项物理地址是 $00301000 \text{H} + 129 \times 4 = 00301000 \text{H} + 204 \text{H} = 00301204 \text{H}$ 。

（2）数组 a 在虚拟地址空间中所占区域必须连续，在物理地址空间中所占区域可以不连续。

(3) 按行遍历的局部性更好, 因为“按行有限方式存放”意味着同一行内的所有数据元素都存放在同一个页面中(页面大小 4096B 刚好可以存放一整行的数据元素)。遍历同一行的所有元素访问的都是同一个页面, 同一列的各个元素都存放在不同的页面中。

47、某校园网有两个局域网, 通过路由器 R1、R2 和 R3 互联后接入 Internet, S1 和 S2 为以太网交换机, 局域网采用静态 IP 地址配置, 路由器部分接口以及各主机的 IP 地址如图所示:



假设 NAT 转换表结构为:

外网		内网	
IP地址	端口号	IP地址	端口号

请回答下列问题:

- (1) 为使 H2 和 H3 能够访问 Web 服务器(使用默认端口号), 需要进行什么配置?
- (2) 若 H2 主动访问 Web 服务器时, 将 HTTP 请求报文封装到 IP 数据报 P 中发送, 则 H2 发送 P 的源 IP 地址和目的 IP 地址分别是? 经过 R3 转发后, P 的源 IP 地址和目的 IP 地址分别是? 经过 R2 转发后, P 的源 IP 地址和目的 IP 地址分别是?

题目解析:

(1) 路由器 R2 开启 NAT 服务, 当路由器 R2 从 WAN 口收到来自 H2 或 H3 发送过来的数据根据 NAT 转换表发送给 WEB 服务器对应端口。R2 的 NAT 转换表可设置如下:

外网		内网	
IP 地址	端口号	IP 地址	端口号
203.10.2.6	默认端口号	192.168.1.2	80
203.10.2.6	默认端口号	192.168.1.3	80

(2)

H2 发送 P 的源 IP 地址: 192.168.1.2 目的 IP 地址 203.10.2.2

R3 转发后 P 的源 IP 地址: 203.10.2.6 目的 IP 地址: 203.10.2.2

R2 转发后 P 的源 IP 地址: 203.10.2.6 目的 IP 地址: 192.168.1.2