新时代大学计算机通识教程

习题解答

桂小林 编著

人民邮电出版社
2021.12.03

目　录

第 1 章 信息与社会.....3

第 2 章 计算系统与平台.....5

第 3 章 程序设计与问题求解.....7

第 4 章 计算机网络与网络安全.....12

第 5 章 物联网技术及应用.....18

第 6 章 大数据分析与人工智能.....22

第 1 章 信息与社会

一、选择题（参考答案：C、B、C、D、C）

1. 在计算机的数值表示中，1 个字节等于（　　）。
   A. 1 位　　B. 2 位　　C. 8 位　　D. 16 位
2. 在计算机的数值表示中，1K 字节等于（　　）。
   A. 1000 字节　　B. 1024 字节　　C. 1000 位　　D. 1024 位
3. 一个 16 进制数 76F 转换为二级制数为（　　）。
   A. 11111011111　　B.0111011011111　　C. 111011011111　　D. 以上都不是
4. 将某个人的姓名和身份证号分布到网络上，将违反（　　）。
   A. 社会道德　　B. 隐私安全　　C. 社会法律　　D. 以上都是
5. 关于十进制数与二进制数的关系，下面表述准确的是（　　）。
   A. 任意一个十进制数都可以用一个固定长度（如 16 位）的二进制表示
   B. 任意一个二进制数都可以用一个固定长度（如 16 位）的十进制表示
   C. 有些十进制数不能用一个固定长度的二进制数表示
   D. 以上表述都不准确

二、问答题

1. 什么是信息？在计算机系统内，为什么信息需要采用二级制进行表示？

答：信息通常是指音讯或消息，即通信系统传输和处理的对象，泛指人类社会传播的一切内容。在计算机中，由于数据以器件的物理状态表示，容易寻找或制造具有两种不同状态的电子元件（如电子开关的接通与断开、晶体管的导通与截止等）。所以，计算机内部一律采用二进制来表示数据。二进制的两种不同状态刚好实现了逻辑值的真与假。

2. 语音编码有哪几种方式？各有何优缺点？

答：语音编码就是对模拟的语音信号进行编码，将模拟信号转化成数字信号。语音编码的基本方式可分为波形编码、参量编码（音源编码）和混合编码。不同语言编码方法的优缺点参考书中的表 1-6。

3. 调研国内与信息安全相关的主要法律。

答：为了有效保护网络与信息安全，打击网络与信息犯罪，中国陆续制定了相关法律。主要包括：

1994 年发布施行的、2011 年修订的《中华人民共和国计算机信息系统安全保护条例》；
2000 年发布施行的、2021 年修订的《互联网信息服务管理办法》；
2002 年发布起施行的《中华人民共和国计算机软件保护条例》；
2017 年发布起施行的《中华人民共和国网络安全法》。

4. 为什么一台计算机不能表示实数坐标系中的所有数值？

答：由于计算机的位数有限。例如，一个 32 位的计算机表示的数据个数是有限的，最多 2^32 个。所以，在有限位数的计算机系统中，能够表示的数据也是有有限的。而实数坐标系中的所有数值是无限的。因此，一台计算机不能表示实数坐标系中的所有数值。

5. 一台计算机能够表示的数值多少主要由什么决定？

答：由计算机的字长决定。

三、计算题

1. 计算十进制数 90.75 的二进制数、八进制数和十六进制数表示。

答：对应的二进制数为 1011010.11B，八进制数为 132.6Q，十六进制数为 5A.CH。

2. 计算二进制数 11000100011.011 的十进制数、八进制数和十六进制数表示。

答：每四位一组,不足补 0，即 0110 0010 0011.0110，求出十六进制数为：623.6H

每三位一组,不足补 0，即 011 000 100 011.011，求出十六进制数为：3043.3Q

采用加权方法，求出十进制数为：1571.375

3. 已知英文数字“1”的 ASCII 字符为 31H，计算英文数字“5”和“8”的 ASCII 编码。

答：因为“1”的 ASCII 字符为 31H，英文数字“5”和“8”的 ASCII 分别为 35H 和 38H。

4. 已知英文数字“A”的 ASCII 字符为 41H，计算英文数字“B”和“Z”的 ASCII 编码。

答：因为英文数字“A”的 ASCII 字符为 41H，英文数字“B”和“Z”的 ASCII 编码分别为 41H+1H=42H，41H+25=41H+19H=5AH。

5. 已知英文数字“a”的 ASCII 字符为 61H，计算英文数字“c”和“r”的 ASCII 编码。

答：因为英文数字“a”的 ASCII 字符为 61H，英文数字“c”和“r”的 ASCII 编码分别为 61H+2H=63H，61H+14=61H+0EH=6FH。

四、综合应用题

1. 利用网络手段，查询中文字符“林”的 UTF-8 编码。

答：利用 http://www.mytju.com/classcode/tools/encode_utf8.asp，得到 E6 9E 97。

2. 利用网络手段，查询汉字“西”和“安”的机内码。

答：利用 http://www.mytju.com/classcode/tools/encode_utf8.asp，“西”的机内码为 CEF7，“安”的机内码为 B0B2。

3. 给定英文字符“B”的 8 列、12 行点阵如下图所示，给出该点阵字符的字形编码序列。

答：00H,00H,F0H,48H,48H,70H,48H,48H,48H,F0H,00H,00H

4. 给定中文字符“汉”的 16 列、12 行点阵如下图所示，给出该点阵字符的字形编码序列。

答：按照行列顺序，该点阵字符的字形编码序列如下。

04H,04H;　第 1 行
02H,FEH;　第 2 行
00H,44H;　第 3 行
09H,44H;　第 4 行
05H,48H;　第 5 行
02H,28H;　第 6 行
02H,28H;　第 7 行
0CH,10H;　第 8 行
04H,28H;　第 9 行
04H,44H;　第 10 行
05H,82H;　第 11 行
00H,00H。　第 12 行

第 2 章 计算系统与平台

一、选择题（参考答案：A、C、C、B、C、C）

1. 8086CPU 从存储器中预取指令，它们采用的存取原则为（ ）。

A．先进先出　B．先进后出　C．随情况不同而不同　D．随机

2. 8086 微处理器执行取指令操作时，段地址由 CS 寄存器提供，段内偏移地址由下列哪一个寄存器提供。

A. BX　B. BP　C. IP　D. SP

3. 世界上第一台电子计算机（ENIAC）于（ ）年在美国诞生。

A. 1944　B. 1945　C. 1946　D. 1947

4. 微型机与其他计算机都采用了（ ）的工作原理

A. 存储数据　B. 存储程序　C. 存储文件　D. 存储图形

5. 一个完整的计算机系统由（ ）组成

A．主机、键盘和显示器　B．主机及外部设备

C．硬件系统和软件系统　D．操作系统及应用软件

6. 操作系统是现代计算机系统不可缺少的组成部分。操作系统负责管理计算机的（ ）。

A.程序　B.功能　C.资源　D.进程

二、问答题

1. 什么是计算平台？计算平台包括哪两大类？

答：计算平台是指提供计算能力的计算机系统及其支撑网络，由硬件、软件及网络等组成。计算平台包括单计算机系统、多计算机系统两大类。

2. 简要说明图灵机模型的工作思想。

答：图灵的基本思想是用机器来模拟人们用纸笔进行数学运算的过程。

3. 简述冯诺依曼体系的计算机的工作原理。

答：冯•诺依曼体系的计算机工作原理为：采用二进制形式表示计算机的指令和数据；将程序（由一系列指令组成）和数据存放在存储器中，并让计算机自动地执行程序。即采用“存储程序和程序控制”的思想进行工作。

4. 说明源码、补码、反码表示方法的优缺点，为什么计算机最终采用了补码表示？

答：原码是一种最直接、方便的编码方案，但是它的符号位不能直接参加加减运算，必须单独处理。反码的符号位可以和数值位一起参加运算，但符号位一旦有进位，结果就会发生偏差，因此要进行修正，增加了电路实现的复杂性。补码运算时可以把符号位与数值位一起处理，不需要事先判断参加运算数据的符号位，电路实现上更加简单，因此，现代计算机的运算器一般都采用补码形式进行加减法运算。

5. 简述 8086 微处理器的结构和功能。

答：8086 主要分为两个独立的功能模块：总线接口部件 BIU（Bus Interface Unit）和执行部件 EU（Execution Unit），它们可以并行工作。BIU 通过 AB、DB、CB 三总线与外部存储器和输入/输出端口交换数据；执行部件负责执行指令。

6. 简述微型计算机系统的组成及其软、硬件结构。

答：一个计算机系统由硬件和软件两部分构成。硬件是指计算机系统中的实体部分，它由电子的、磁性的、机械的、光的元器件组成，包括运算器、控制器、存储器、输入设备和输出设备五大部分组成。软件是指在计算机硬件上运行的各种程序和有关文档的总称，包括系统软件、应用软件和软件开发环境（程序设计软件）三大类。

7. 系统总线分为哪三类？每种系统总线的作用是什么？

答：系统总线由数据总线（DB）、地址总线（AB）和控制总线（CB）构成。控制总线上传送一个部件对另一个部件的控制信号，起到定时、控制作用，目标是保证计算机同步、协调工作。地址总线上传送总线主控（如 CPU）发出的地址信号，向存储器、外设发送地址。数据总线在存储器、外设与 CPU 间传送数据，数据总线是双向的，这些数据可以包括地址信息。

8. 在 8086 CPU 硬件复位后，程序从何处开始执行？

答：FFFF0H

9. 什么是多核处理器？

答：多核处理器是指在一枚处理器中集成两个或多个完整的计算引擎(内核)，多个引擎协同工作。

10. 在 8086 中，如何计算存储器的地址？

答：段地址*16+偏移地址

三、填空题

1. 已知 8086 的段地址 CS=8000H，IP=9000H，则该 8086 输出在 AB 上的地址为 89000H。

2. 十进制数“-127”的八位二进制补码可表示为 1 000 0001、八位二进制反码可表示为 1 000 0000 。

3. 十进制数“128”的八位二进制原码可表示为 无法表示 ，八位二进制补码可表示为无法表示。

4. 两个二进制数 10101110B 和 11110000B 的逻辑“与”的结果为 1010000 ,逻辑“异或”的结果为 01011110。

5. 已知某显示器的分辨率为 1280×1024，色彩为 256 位，则该显示器存储一屏幕的图像所需要的存储空间为 40M 字节。【计算方法：1280*1024*256/8 = 41943040 字节= 40960KB = 40MB 】

四、综合应用题

1. 已知 X=+77H,Y=-77H，采用八位补码运算时，设计一种方案判断 X-Y 是否存在溢出？

答：采用单符号位进行判断。[X]补=[+77H]补= 0 111 0111B， [-Y]补=[77H]补= 0 111 0111B，

[X-Y]补 = [X]补 + [-Y]补 = 1 110 1110B。溢出。

也可以采用双符号判定。[X]补=[+77H]补= 00 111 0111B， [-Y]补=[77H]补= 00 111 0111B，

[X-Y]补 = [X]补 + [-Y]补 = 01 110 1110B。正溢出。

2. 已知一个浮点数为 2.0*10^80，试设计一种方案，利用一个 16 位计算机系统来表示和存储这个浮点数，并计算这个浮点数在存储后的误差是多少。

答：由于给定的浮点数超过了定点数表示的范围（2^16），所以，只能使用浮点数进行表示。将 2.0*10^80 转化为 0.2*10^81。尾数表示 0.2 时，如果用 16 位二进制原码表示为 0.0011001100110011，则存在误差。指数如果采用 10 为阶，则 81 可以精确表示；如果采用 2 为阶，则 10^81 转换为 2^270 则存在误差。

3. 在 Linux 操作系统上，使用 KVM 虚拟化技术，配置一个 windows 操作系统。

答：略。

第 3 章 程序设计与问题求解

一、选择题（参考答案：AD、D、C、A、D、C）

1. 关于问题与问题求解，下列说法正确的是（　　）。
   A.问题求解是人们为寻求问题答案而进行的一系列思维活动。
   B.问题是客观存在的，提出问题与发现问题与人对事情的好奇心和求知欲无关。
   C.所有问题都是有科学研究价值的。
   D.人类进行问题求解的一般思维过程可分为问题分析、提出假设和检验假设。
2. 关于查找和排序，下列叙述正确的是（）。
   A.排序只能对数字进行。　　B.冒泡排序属于选择排序。
   C.所有的排序类算法，其效率与对象的原始状态无关。
   D.折半查找只能在有序数据序列中进行。
3. 在 Python 中，实现多分支选择结构的较好方法是（　）。
   A．if　B．if-else　C．if-elif-else　D．if 嵌套
4. 下列四组选项中（每组 3 个），均是合法的用户标识符的选项是（　）。
   A. float　ly897　_S　　B. for　int　x^2
   C. 5W　P_0　in　　D. a_123　abc　True
5. 下面不是合法的字符串是（　　）
   A. 'abc'　B. "ABC"　C. 'He said:"OK" '　D. 'can't
6. 下面正确的函数定义是（　）。

```
A. define fun ( x, y):
       z = x + y
B. def fun (x, y)
       z = x + y
C. def fun (x, y):
     z = x + y
    return z
D. def fun (x = 2, 2):
       z = x + y
       return z
```

二、问答题

1. 什么是指令？什么是程序？二者有何关联？

答：指令是指示计算机执行某种操作的命令，它由一串二进制数码组成。程序是一组为完成某种功能而按一定顺序（通常由算法确定）编排的指令序列，是人与计算机之间传递信息的媒介。二者有何关联是：程序是由指令构成的。

2. 程序设计语言可以分为哪几类？各有何优缺点？

答：程序设计语言可以分为：机器语言、汇编语言和高级语言。机器语言难于记忆但效率高；机器语言较易记忆效率较高；高级语言方便记忆但效率不高。

3. 比较编译型程序设计语言与解释型程序语言的优缺点。

答：解释类程序设计语言的执行方式类似于我们日常生活中的“同声翻译”。源代码由相应语言的解释器一边“翻译”，一边执行，效率较低，但修改、调试、纠错更加容易。

编译类程序设计语言的将程序源代码“翻译”成目标代码，可脱离其语言环境独立执行，防拷贝强、效率高，但修改、调试困难。

4. 如何对问题进行合理抽象和描述？

答：采用自然语言、流程图、实体-关系图（即 E-R 图）对问题进行精确描述。

5. 什么是计算思维？计算思维的核心思想是什么？

答：计算思维就是利用计算机相关技术解决领域应用问题的一种思维方法，其核心思想就

是问题抽象、自动化求解和智能优化。计算思维的核心思想是运用计算机科学的基础概念进行问题求解、系统设计。

6. 什么是代码复用？有哪几种代码复用方法？

答：代码复用是指将软件代码进行多次调用或使用的过程。代码复用包括目标代码和源代码的复用。源代码复用包括自定义函数和构件库两种。

7. 什么是递归算法？递归算法深度是否存在边界？如何确定这一边界？

答：递归算法是一种重复地将复杂问题分解为同类子问题，直到子问题依次获得求解的一种计算方法。递归算法的深度存在边界，该边界由计算机系统的堆栈大小决定。

8. 简述贪心算法的使用场合及其优缺点。

答：贪心算法是指在对问题求解时，总是做出在当前看来是最好的选择。因此，贪心算法适合于“只顾眼前利益”情况、或只需要获得局部最优解的场合。

9. 对插入排序、选择排序、冒泡排序、快速排序 4 种算法进行性能比较。

答：参见教材的表 3-15.

三、问题求解与编程题

1. 编程实现：求 2-500 之间的素数，并且每行显示 9 个素数。

答：参考教材中的程序 3-10 进行编程。

```
def isPrime(n):     #验证是否是素数的函数
     if n <= 1:
          return 0
     for i in range(2, n):
          if n % i == 0:
               return 0   #不是素数
     return 1             #是素数
#主程序
k = 0   # 控制分行
for n in range(2,501):  #遍历对象 2 至 500（含）
     if (isPrime(n)):   #调研函数进行枚举并验证
         print("%5d"%n, end=",")     #输出素数
         k = k+1
     if  k==9:         #没 9 个素数占用 1 行
         print()
         k= 0
```

运行结果为：

```
    2,    3,    5,    7,   11,   13,   17,   19,   23,
   29,   31,   37,   41,   43,   47,   53,   59,   61,
   67,   71,   73,   79,   83,   89,   97,  101,  103,
  107,  109,  113,  127,  131,  137,  139,  149,  151,
  157,  163,  167,  173,  179,  181,  191,  193,  197,
  199,  211,  223,  227,  229,  233,  239,  241,  251,
  257,  263,  269,  271,  277,  281,  283,  293,  307,
  311,  313,  317,  331,  337,  347,  349,  353,  359,
  367,  373,  379,  383,  389,  397,  401,  409,  419,
  421,  431,  433,  439,  443,  449,  457,  461,  463,
```

467， 479， 487， 491， 499，

2. 编程实现：从键盘输入 5 个一位自然数，将这 5 个自然数按照输入的先后顺序转换为一个 5 位整数。

答：将输入的字符序列转换为 10 进制整数。

```
hexdig =0
for i in range(5):
    print("请输入第", i, "个一位自然数", end=" ")
    dig = int(input("#"))
    hexdig = hexdig*10 + dig
print("转换的结果为：", hexdig)
```

运行结果为：

```
请输入第 0 个一位自然数#1
请输入第 1 个一位自然数#3
请输入第 2 个一位自然数#5
请输入第 3 个一位自然数#8
请输入第 4 个一位自然数#9
转换的结果为： 13589
```

3. 编程实现：给定 10 个 100 以内的自然数，使用列表结构进行存储，使用冒泡排序算法找出其中最小者和最大者。

答：参考教材中的程序 3-20 进行编程。

```
dig =[2,14,35,69,76,15,23,87,82,69]   # 列表 dig 初始化要排序的数列
def bubbleSort(dig):
     for i in range(1, len(dig)):
        for j in range(0, len(dig)-i):
            if dig[j] > dig[j+1]:
                 temp= dig[j]; dig[j]= dig[j+1]; dig[j+1]= temp #数据交换
     return dig
#主程序
dig1 = bubbleSort(dig)  # 将排序结果放在列表 dig1 中
print("max=", dig1[9], ", min=", dig1[0])    #显示排序结果
```

运行结果为：

```
max= 87 , min= 2
```

4. 问题求解：某天晚上，某某在家中被害，侦查过程中发现 A、B、C、D 四人到过现场，警察在询问他们时，A 说：“我没有杀人”；B 说：“C 是杀人凶手”；C 说：“D 是杀人者”；D 说：“C 在冤枉好人”。警察经过判断，4 人中有 3 人说的是真话，1 人说的是假话，4 人中只有一人是凶手，那么，凶手到底是谁呢？请对上述问题进行分析和编程求解。

答：具体求解方法如下：

（1）问题分析：

用 0 表示不是凶手，用 1 表示凶手，则每个人的取值范围就是 0 或 1，根据四人的回答，可以构成的关系表示式如下：A 没有杀人，A==0；C 是杀人凶手，C==1；D 是杀人者，D==1；C 在冤枉好人，D==0。

结合四位嫌疑人的回答，警察判断的逻辑表达式如下：

条件 1：四人中三人说的是真话，即(A==0)+(C==1)+(D==1)+(D==0)=3

条件 2：四人中只有一人是凶手，即 A+B+C+D=1

（2）问题求解：

基于上述逻辑进行推理的过程虽然可以通过人工进行，其搜索空间较大。我们可以使用 Python 程序在每个人的取值范围为 0 和 1 的所有可能中进行搜索，当条件 1 和条件 2 同时满足时，即为凶手。

（3）编程实现：

```
sum= 0
for A in range(0,2):
    for B in range(0,2):
        for C in range(0,2):
            for D in range(0,2):
                sum = sum + 1
                if ((A==0)+(C==1)+(D==1)+(D==0)==3) and (A+B+C+D==1):
                    print("A=", A, "; B=", B, "; C=", C, "; D=", D )
```

运行结果为：A= 0 ; B= 0 ; C= 1 ; D= 0。所以 C 是凶手。

四、实验题

装修公司新进了一块长为 a、宽为 b 的大木板。根据装修要求，需要将其裁剪成若干块相同的正方形木板，使得在尽量不浪费木板的情况下，裁剪的正方形木板的长度 c 尽可能大。提示：（1）可以将该问题转化为 c 为 a 和 b 的最大公约数问题；（2）可以用递归算法和迭代算法分别实现，并比较二者的执行效率。

解：

（1）问题分析

该问题要求不浪费布料，所以 c 应该同时被 a 和 b 整除，而且 c 尽可能大。显然，c 为 a 和 b 的最大公约数，表示为 gcd(a,b)。根据最大公约数定理，gcd(a,b)的计算公式如下（约定 a 大于等于 b）：

gcd(a,b) = gcd (b, a mod b)

上述迭代方法称之为辗转相除法，也叫做“欧几里得算法”。该方法是希腊数学家欧几里得在其著作《几何原本》提出的。

辗转相除法表明，a 和 b 的最大公约数等于 b 和 a 模 b（或者说 a 除以 b 的余数）之结果的最大公约数。

（2）问题求解

显然，根据上述欧几里得算法的公式，可以容易设计出自顶向下的递归算法；也可以设计出迭代算法。

其中，迭代算法的基本思路为：a 和 b 分别表示两个正整数，其中 a≥b，t 表示 a 除以 b 的余数，即 t=a mod b 的结果；如果 t 不等于 0，则迭代计算，即 a 更新为 b 的值，b 更新为余数 t，重复上述过程；如果 t 等于 0，则迭代过程终止，b 是所求的结果。

例如，求 8251 和 6105 的最大公因数的迭代过程如下。

第 1 次迭代：8251＝6105×1＋2146，a=8251，b = 6105，t = 2146。

第 2 次迭代：6105＝2146×2＋1813，a=6105，b = 2146，t = 1813。

第 3 次迭代：2146＝1813×1＋333，a=2146，b = 1813，t = 333。

第 4 次迭代：1813＝333×5＋148，a=1813，b = 333，t = 148。

第 5 次迭代：333＝148×2＋37，a=333，b = 148，t = 37。
第 6 次迭代：148＝37×4＋0，a=148，b = 37，t = 0。
算法终止，37 为 8251 与 6105 的最大公因数。
例如，求 6643 和 5913 的最大公因数的迭代过程如下。
第 1 次迭代：6643＝5913×1＋730，a= 6643，b= 5913，t= 730。
第 2 次迭代：5913＝730×8＋73，a= 5913，b= 730，t= 73。
第 3 次迭代：37＝37×1＋0，a= 37，b= 37，t = 0。
算法终止，73 为 6643 和 5913 的最大公因数。

上述迭代过程虽然要求 a 和 b 为整数，但实际上，如果输入的 a 和 b 为小数，也可求解。只需要将 a、b 同时乘以 10 的若干次幂转换为整数即可。

（3）编程实现

根据上述的迭代关系，本问题迭代算法的实现程序如下。

```
def gcd(a, b):
   while(a % b !=  0):
       t = a%b
       a = b  #迭代变量
       b = t  #迭代变量
   return b
# 主程序
a = int(input("请输入 a="))
b = int(input("请输入 b="))
print(a, "和", b, "的最大公约数为", gcd(a, b))
```

运行结果为：
请输入 a=8251
请输入 b=6105
8251 和 6105 的最大公约数为 37
请输入 a=6643
请输入 b=5913
6643 和 5913 的最大公约数为 73

第 4 章 计算机网络与网络安全

一、选择题（参考答案：A、C、C、B、C；C、A、B、C、D；B、C、D、A、C）

1. 局域网的英文缩写为（　　）。
   A、LAN　B、WAN　C、ISDN　D、MAN
2. 计算机网络中广域网和局域网的分类是以（　　）来划分的。
   A、信息交换方式　B、网络使用者
   C、网络连接距离　D、传输控制方法
3. OSI（开放系统互联）参考模型的最底层是（　　）。
   A、传输层　B、网络层　C、物理层　D、应用层
4. 在 Internet 中，用来进行数据传输控制的协议是（　　）。
   A. IP　B. TCP　C. HTTP　D. FTP
5. Internet 的域名中，顶级域名为 gov 代表（　　）。
   A. 教育机构　B. 商业机构　C. 政府部门　D. 军事部门
6. 在 Web 服务网址中，http 代表（　　）。
   A. 主机　B. 地址　C. 协议　D. TCP/IP
7. 超文本的含义是（　　）。
   A. 文本中可含有图像　B. 文本中可含有声音
   C. 文本中有超级链接　D. 文本中有二进制字符
8. 用 Internet 访问某主机可以通过（　　）。
   A. 地理位置　B. IP 地址　C. 域名　D. 从属单位名
9. 在 Internet 电子邮件系统中（　）。
   A.　发送邮件和接收邮件都使用 SMTP 协议。
   B.　发送邮件使用 POP3 协议，接收邮件使用 SMTP 协议。
   C.　接收邮件使用 POP3 协议，发送邮件使用 SMTP 协议。
   D.　发送邮件和接收邮件都使用 POP3 协议。
10. 下列 IP 地址中，能够直接分配给主机的是_____。
   A.　192.168.0.1　B.　127.11.10.101　C.　224.10.10.10　D.　202.117.48.255
11. Ethernet 采用的介质访问控制方式为（　）
   A.CSMA　B.CSMA/CD　C.CDMA　D.CSMA/CA
12. 在 OSI 参考模型中能实现路由选择、拥塞控制与互连功能的层是（　　）。
   A. 传输层　B. 应用层　C. 网络层　D. 物理层
13. 在下面给出的协议中，（　　）属于 TCP/IP 的应用层协议。
   A. TCP 和 FTP　B. IP 和 UDP
   C. RARP 和 DNS　D. FTP 和 SMTP
14. 在下面对数据链路层的功能特性描述中，不正确的是（　　）。
   A. 通过交换与路由，找到数据通过网络的最有效的路径
   B. 数据链路层的主要任务是提供一种可靠的通过物理介质传输数据的方法
   C. 将数据分解成帧，并按顺序传输帧，并处理接收端发回的确认帧
   D. 以太网数据链路层分为 LLC 和 MAC 子层，在 MAC 子层使用 CSMA/CD 的协议
15. 网络层、数据链路层和物理层传输的数据单位分别是（　）。
   A. 报文、帧、比特　B. 包、报文、比特
   C. 包、帧、比特　D. 数据块、分组、比特

二、问答题

1．简述 OSI 分层模型中各层的主要功能？

答：应用层提供网络服务与最终用户的接口；表示层提供数据表示、加解密和解压缩等功能；会话层建立、管理和终止网络会话（即通信连接）；传输层定义传输数据的协议端口号，以及流量控制和差错校验功能；网络层进行逻辑地址寻址并实现不同网络之间的路径选择；数据链路层建立逻辑连接、进行硬件地址寻址、差错校验等；物理层建立、维护、断开物理连接。

2．局域网与广域网相比主要特点是什么？

答：局域网（LAN）分布在相对较小的区域，通常不超过 10km，如同一房间内的若干计算机，同一楼内的若干计算机；同一校园、厂区内的若干计算机等。广域网（WAN）是指网络中的计算机跨区域分布，地理范围能够覆盖 10km 甚至 1000km 以上的范围，比如同一个省、同一个国家或同一个洲甚至跨越几个洲等。广域网也称为互联网（Internet），通常是由多个局域网或城域网组成。

3．IEEE 802 体系结构的特点是什么？

答：按照 IEEE 802 标准，局域网体系结构分为三层，即物理层、媒体链路控制层（MAC 子层）和逻辑链路控制层（LLC 子层）。不同类型的局域网通常使用不同的介质访问控制协议，如以太网、令牌环、令牌总线等。它们所遵循的都是 IEEE 制定的以 802 开头的标准。

4.IP 路由器的主要功能是什么？

答：IP 网络中的路由选择是由路由器完成的。路由器通过执行一定的路由协议，为 IP 数据报寻找一条到达目的主机或网络的最佳路由，并转发该数据报，实现路由选择。I

5．防火墙技术有什么作用？防火墙能防病毒吗？

答：防火墙是一种用来加强网络之间访问控制的特殊网络设备，常常被安装在受保护的内部网络连接到 Internet 的点上，它对传输的数据包和连接方式按照一定的安全策略对其进行检查，来决定网络之间的通信是否被允许。防火墙能够依据一定规则阻止病毒传播。

6．有哪些防御攻击的安全措施？

答：病毒检测与防护、杀毒软件、防火墙等。

7．什么是计算机病毒？

答：计算机病毒（computer virus）是一种人为制造的、能够进行自我复制的、具有对计算机资源的破坏作用的一组程序或指令的集合。

8．访问控制技术有哪些，各有什么特点？

答：访问控制就是在身份认证的基础上，依据授权对提出的资源访问请求加以控制。访问控制可以限制对关键资源的访问，防止非法用户的侵入或合法用户的不慎操作所造成的破坏。访问控制按照访问手段可分为自主访问控制和强制访问两类。自主访问控制(DAC)由客体自主地来确定各个主体对它的直接访问权限。强制访问控制(MAC)的每一个数据对象被标以一定的密级，每一个用户也被授予某一个级别的许可证。对于任意一个对象，只有具有合法许可证的用户才可以存取。

三、综合计算题

1．在以太网中，位串 01110111110011111101 需要在数据链路层上被发送，请问经过位填充后实际被发送出去的什么？

答：在以太网中，由于采用了零比特插入技术，在传输前需要加入前导标识 01111110，所以数据中每 5 个 1 需要插入一个 0。即，位串 01110111110011111101 就变成了位串 0111011111<u>0</u>0011111<u>0</u>101。

2. 若要将一个 B 类的网络 202.117.0.0 划分为 14 个子网，请计算出每个子网的子网掩码，

以及在每个子网中主机 IP 地址的范围是多少？

答：该网络的子网掩码是 255.255.240.0。划分的 14 个子网的掩码地址分别为：

202.117.16.0、202.117.32.0、202.117.48.0、202.117.64.0、202.117.80.0、202.117.96.0、202.117.112.0 、 202.117.128.0 、 202.117.144.0 、 202.117.160.0 、 202.117.176.0 、202.117.192.0、202.117.208.0、202.117.224.0。

每个子网的主机 IP 地址范围分别为：

202.117.16.1--202.117.31.254、202.117.32.1--202.117.47.254、
202.117.48.1--202.117.63.254、202.117.64.1--202.117.79.254、
202.117.80.1--202.117.95.254、202.117.96.1--202.117.111.254、
202.117.112.1--202.117.127.254、202.117.128.1--202.117.143.254、
202.117.144.1--202.117.159.254、202.117.160.1--202.117.175.254、
202.117.176.1--202.117.191.254、202.117.192.1--202.117.207.254、
202.117.208.1--202.117.223.254、202.117.224.1--202.117.239.254

3. 若要将一个 B 类的网络 202.117.0.0 划分子网，其中包括 3 个能容纳 16000 台主机的子网，7 个能容纳 2000 台主机的子网，8 个能容纳 254 台主机的子网，请写出每个子网的子网掩码和主机 IP 地址的范围。

答：16000 台主机，最少需要 14 位地址进行寻址。2000 台主机，最少需要 11 位地址进行寻址。254 台主机，最少需要 8 位地址寻址。

3 个能容纳 16000 台主机的子网掩码和主机 IP 地址的范围如下：

202.117.0.0/18；　　202.117.0.1--202.117.63.254
202.117.64.0/18；　202.117.64.1--202.117.127.254
202.117.128.0/18；　　202.117.128.1--202.117.191.254

7 个能容纳 2000 台主机的子网掩码和主机 IP 地址的范围如下：

202.117.192.0/18---　202.117.192.0/21　　202.117.192.1--202.117.199.254
202.117.200.0/21　　202.117.200.1--202.117.207.254
202.117.208.0/21　　202.117.208.1--202.117.215.254
202.117.216.0/21　　202.117.216.1--202.117.223.254
202.117.224.0/21　　202.117.224.1--202.117.231.254
202.117.232.0/21　　202.117.232.1--202.117.239.254
202.117.240.0/21　　202.117.240.1--202.117.247.254

8 个能容纳 254 台主机的子网掩码和主机 IP 地址的范围如下：

202.117.248.0/21----202.117.248.0/24
202.117.248.1--202.117.248.254
202.117.249.0/24　202.117.249.1--202.117.249.254
202.117.250.0/24　202.117.250.1--202.117.250.254
202.117.251.0/24　202.117.251.1--202.117.251.254
202.117.252.0/24　202.117.252.1--202.117.252.254
202.117.253.0/24　202.117.253.1--202.117.253.254
202.117.254.0/24　202.117.254.1--202.117.254.254
202.117.255.0/24　202.117.255.1--202.117.255.254

4. 对于一个从 192.168.0.0 开始的超网，假设能够容纳 4000 台主机，请写出该超网的子网掩码以及所需使用的每一个 C 类的网络地址。

答：该超网的子网掩码为 192.168.0.0/20。

C 类的网络地址为：

192.168.00.0/24、192.168.01.0/24、192.168.02.0/24、192.168.03.0/24、192.168.04.0/24、192.168.05.0/24、192.168.06.0/24、192.168.07.0/24、192.168.08.0/24、192.168.09.0/24、192.168.10.0/24、192.168.11.0/24、192.168.12.0/24、192.168.13.0/24、192.168.14.0/24、192.168.15.0/24

四、实验题

1. 在 Windows10 上配置网络防火墙。

答：Windows10 操作系统上自带网络防火墙功能，用户可以自行配置，实现网络防火墙功能，也可以使用系统默认的防火墙功能。

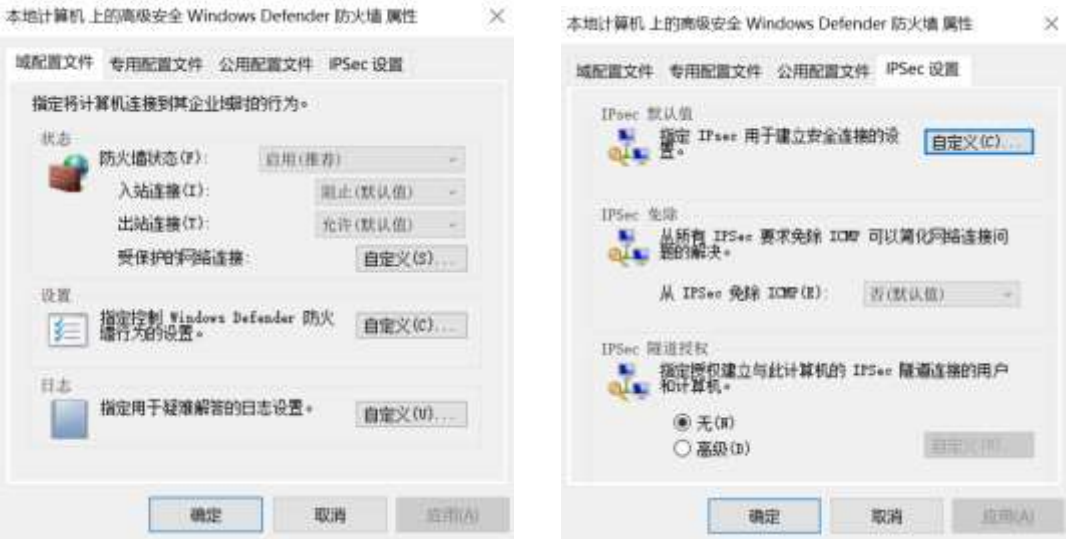
下面以 Windows 10 操作系统为例，介绍 Windows Defender 防火墙中的 IPSec 配置过程。

（1）选择 “开始→控制面板→系统安全→Windows Defender 防火墙”，打开 Windows 防 Defender 防火墙。如图 1 所示。

图 1　Windows10 的防火墙界面

（2）点击图 1 中左侧的 “高级设置”，出现图 2 所示的界面，并在界面中点击“操作”中的“属性”，出现图 3(a)所示界面。选择“IPSec 设置”选项卡，出现图 3(b)所示界面。

图 2 Windows10 的防火墙高级设置界面



（a） （b）

图 3 高级安全 Windows 防火墙属性的界面

在图 3(b)所示的界面中，包括三个与 IPSec 相关的选项，各选项的功能说什么如下：

（a）IPSec 默认值

可以配置 IPSec 用来帮助保护网络流量的密钥交换、数据保护和身份验证方法。单击“自定义”按钮可以显示“自定义 IPsec 设置”对话框，如图 5-15 所示。用户一般选择默认设置即可；如果需要自定义，则可选择“高级”。

当具有活动安全规格时，IPSec 将使用该项设置规则建立安全连接，如果没有对密钥交换（主模式）、数据保护（快速模式）和身份验证方法进行指定，则建立连接时将会使用组策略对象（GPO）中优先级较高的任意设置，顺序如下：最高优先级组策略对象（GPO）→本地定义的策略设置→IPSec 设置的默认值（比如身份验证算法默认是 Kerberos V5 等更多默认可以直接点击图 4 所示下面窗口的“什么是默认值”帮助文件）。

图 4 自定义 IPSec 设置界面

图 5 自定义 IPSec 隧道授权界面

（b）IPSec 免除

此选项设置确定包含 Internet 控制消息协议（ICMP）消息的流量包是否受到 IPsec 保护 ICMP 通常由网络疑难解答工具和过程使用。注意，此设置仅从高级安全 Windows 防火墙的 IPsec 部分免除 ICMP，若要确保允许 ICMP 数据包通过 Windows 防火墙，还必须创建并启用入站规则，另外，如果在“网络和共享中心”中启用了文件和打印机共享，则高级安全 Windows 防火墙会自动启用允许常用 ICMP 数据包类型的防火墙规则。可能也会启用与 ICMP 不相关的网络功能，如果只希望启用 ICMP，则在 Windows 防火墙中创建并启用规则，以允许入站 ICMP 网络数据包。

（c）IPSec 隧道授权

只在以下情况下使用此选项：具有创建从远程计算机到本地计算机的 IPsec 隧道模式连接的连接安全规则，并希望指定用户和计算机，以允许或拒绝其通过隧道访问本地计算机。

具体操作流程为：在图 3(b)界面下面“IPSec 隧道授权”部分选择“高级”，然后单击“自定义”可以显示“自定义 IPsec 隧道授权”对话框可以为需要授权的计算机或用户进行隧道规则授权。IPSec 隧道授权设置界面如图 5 所示。

2. 使用 Foxmail 配置一个客户端邮件系统。

答：首先，需要下载 foxmail 软件，如 https://www.foxmail.com/

然后，按照软件要求，进行安装。

其次，安装完成后进行配置。

第 5 章 物联网技术及应用

一、选择题（参考答案：B、A、B、A、B；A、C、B）

1. 基于 RFID 技术的唯一编码方案，即产品电子编码（EPC）最早提出的是（　）。
   A、斯坦福大学　B、麻省理工学院　C、哈佛大学　D、西安交通大学
2. 2009 年 8 月 7 日，温家宝总理在无锡微纳传感网工程技术研发中心视察并发表重要讲话，提出了（　）。
   A、感知中国　B、物联中国　C、中国制造 2025　D、工业 4.0
3. RFID 系统中，无源标签的能耗来自（　）。
   A、光照　B、磁场　C、电池　D、振动
4. 目前流行的智能手机的计步功能，主要通过（　）传感器实现。
   A、加速度　B、温度　C、光　D、声音
5. 利用支付宝进行地铁支付，其技术实现主要是基于（　）。
   A、一维码　B、二维码　C、RFID　D、图像
6. 2008 年 8 月，美国麻省理工学院的三名学生宣布成功破解了波士顿地铁资费卡。主要原因是（　）。
   A、密码过于简单　B、物理保护措施不力　C、机器故障　D、内部泄密
7. 不属于电子钱包的是（　）。
   A、微信零钱　B、支付宝花呗　C、银行信用卡　D、京东白条
8. 关于空间单位描述准确的是（　）
   A、GPS 可以进行室内定位　B、北斗可以进行室外定位
   C、WiFi 只能室内定位　D、蜂窝只能室外定位

二、简答题

1. 什么是物联网？

答：物联网是通过使用 RFID、传感器、红外感应器、全球定位系统、激光扫描器等信息采集设备，按约定的协议，把任何物品与互联网连接起来，进行信息交换和通信，以实现智能化识别、定位、跟踪、监控和管理的一种网络（或系统）。

2. 物联网的四个主要特征是什么？简述每个特征的含义。

答：尽管对物联网概念还有其他一些不同的描述，但内涵基本相同。经过近十年的快速发展，物联网展现出了与互联网、无线传感网不同的特征。物联网主要特征包括：全面感知、可靠传递、智能处理和广泛应用四个方面。

（1）全面感知：利用电子标签、条形码、传感器、红外感应器等多元感知装置或设备，随时随地获取物品信息。

（2）可靠传递：通过各类网络通信设施，将各种传感器的信息利用可靠传输到数据中心。

（3）智能处理：根据感知数据进行智能化识别、定位、跟踪、监控和管理。

（4）广泛应用：在零售、物流、交通和工业等领域进行广泛应用。

3. 简述第 4 次工业革命和中国制造 2025 的含义。

答：2011 年 4 月的汉诺威工业博览会上，德国政府正式提出了工业 4.0 （Industry 4.0）战略。工业 4.0 的核心就是物联网，其目标就是实现虚拟生产和与现实生产环境的有效融合，提高企业生产率。中国制造 2025 是中国政府实施制造强国战略的第一个十年行动纲领。围

绕实现制造强国的战略目标，《中国制造 2025》提出，坚持“创新驱动、质量为先、绿色发展、结构优化、人才为本”的基本方针，坚持“市场主导、政府引导，立足当前、着眼长远，整体推进、重点突破，自主发展、开放合作”的基本原则，通过“三步走”实现制造强国的战略目标。即，第一步：到 2025 年迈入制造强国行列；第二步：到 2035 年中国制造业整体达到世界制造强国阵营中等水平；第三步：到新中国成立一百年时，综合实力进入世界制造强国前列。

4. 简述传感器的主要分类方法。

答：传感器一般是根据物理学、化学、生物学等特性、规律和效应设计而成的，其种类繁多，往往同一种被测量可以用不同类型的传感器来测量，而同一原理的传感器又可测量多种物理量，因此传感器有许多种分类方法。具体包括：（1）根据被测对象划分，常见的有温度传感器、湿度传感器、压力传感器、位移传感器、加速度传感器；（2）按照工作原理可以分为电学式、磁学式、谐振式、化学式等传感器。

5. 简述一维条形码和二维条形码的主要区别

答：一维条形码技术在商业、交通运输、医疗卫生、快递仓储等行业得到了广泛应用。但是，一维条形码存在非常多的缺陷。首先，其表征的信息量有限，每英寸只能存储十几个字符信息。其二，一维条形码只能表达字母和数字，而不能表达汉字和图像。其三，一维条形码不具备纠错功能，比较容易受外界污染的干扰。二维条形码的诞生解决了一维条形码不能解决的问题。二维条形码是用某种特定的几何图形按一定规律相应元素位置上用“点”表示二进制“1”,用“空”表示二进制“0”，由“点”和“空”的排列组成的代码。二维码是一种比一维码更高级的条形码格式。一维码只能在一个方向（一般是水平方向）上表达信息，而二维码在水平和垂直方向都可以存储信息。一维码只能由数字和字母组成，而二维码能存储汉字、数字和图片等信息，因此二维码的应用领域要广得多。

6. 请调研生活中使用二维码的场合，思考为什么利用二维码能够在景区进行自助导游？

答：随着现在旅游业的不断发展，智慧旅游应运而生。通过手机二维码扫描自助语音导览系统，就可以免费（或收费）收听语音导游讲解，更加清晰的了解到景区的风景内容。利用二维码进行导游，首先要构建一个 Web 服务网站，将每个讲解内容的语音或视频放在网站的一个页面上；然后将该该页面的链接生成一个二维码，张贴在景区的相关位置供游客扫描；游客用手机扫描后会自动链接到预先设置网页上。从而实现自助导游功能。

7. 简述北斗卫星定位的基本原理，说明其主要应用。

答：北斗卫星导航系统由空间段、地面段和用户段三部分组成，可在全球范围内全天候、全天时为各类用户提供高精度、高可靠定位、导航、授时服务，并且具备短报文通信能力，已经初步具备区域导航、定位和授时能力，定位精度为分米、厘米级别，测速精度 0.2 米/秒，授时精度 10 纳秒。北斗卫星定位系统由 35 颗卫星组成，包括 5 颗静止轨道卫星、27 颗中地球轨道卫星、3 颗倾斜同步轨道卫星。5 颗静止轨道卫星定点位置为东经 58.75°、80°、110.5°、140°、160°，中地球轨道卫星运行在 3 个轨道面上，轨道面之间为相隔 120°均匀分布。由于北斗卫星分布在离地面 2 万多千米的高空上，以固定的周期环绕地球运行，使得在任意时刻，在地面上的任意一点都可以同时观测到 4 颗以上的卫星。

8. 什么是 RFID 技术？RFID 系统的基本组成部分有哪些？

答：RFID 技术是指利用电磁信号和空间耦合（电感或电磁耦合）的传输特性实现对象信息

的无接触传递，从而实现对静止或移动物体的非接触自动识别技术。通常，RFID 系统由射频电子标签、读写器和数据管理系统组成。

9. 简述 RFID 的识读原理，给出一种安全的 RFID 识读方法。

答：RFID 读写器利用天线在周围形成电磁场，发射特定的询问信号，当电子标签感应到这个信号后，就会给出应答信号，应答信号中含有电子标签携带的数据信息。读写器在读取数据后对其进行梳理，最后将数据返回给后台系统，进行相应操作处理。为了阻止非授权的 RFID 读写器访问非授权的电子标签，多种基于 RFID 安全认证的识读协议相继提出。在这些安全认证协议中，比较流行的是基于 Hash 运算的安全认证协议，它对消息的加密通过 Hash 算法实现。Hash Lock 协议是一种经典的隐私增强的 RFID 识读协议。该协议是 MIT 的 Sarma 等人提出的，不直接使用真正的节点 ID，取而代之的是一种短暂性节点，即临时节点 ID。这样做的好处是，保护了真实的节点 ID。

三、综合应用题

1. 已知《物联网技术导论》的国际标准书号 978-7-302-51064-？，给出其校验码计算方法。

答：13 位书号校验码的计算方法与 EAN-13 的校验码计算方法相同。具体计算过程如下：

（1）先求偶数位的和，然后乘以 3。即 N1 =(7+7+0+5+0+4)*3 = 23*3 = 69

（2）再求奇数位的和。即 N2 = 9+8+3+2+1+6 = 29

（3）计算 N1 与 N2 之和除以 10，得余数 M。即 M=(69+29)/10= 8

（4）计算 10-M 的校验码。即 10-8 = 2

2. 已知《计算机学报》的国际标准期刊号 ISSN 0254-4164，给出校验码的计算方法。

答：ISSN 校验码的计算方法为：前 7 位数字依次以 8～2 加权后求和、再以 11 为模数进行计算得到余数。若余数为 0，则校验码为 0；否则校验码为 11 减余数，余数如果为 10，则用“X”表示。

ISSN 0254-4164，给出校验码的计算方法为：0*8+2*7+5*6+4*5 + 4*4+1*3+6*2 =95。

95 mod 11 = 7,因为 7 不为 0，则校验码为 11-7=4。

3. 利用 Python 程序实现 EAN-13 的校验码计算。

答：

```
#EAN 校验码计算
EANList=[9,7,8,7,3,0,2,5,1,0,6,4,2]
N1=0; N2=0
for i in range(6):
    N1 +=  EANList[2*i+1]  #对应算法中的偶数位
    N2 +=  EANList[2*i]     #对应算法中的奇数位
M = (3*N1 + N2) % 10
CheckBit = 10 - M
print(CheckBit)
```

运行结果为：2

4. 利用网络平台生成字符串“西安交通大学”的二维码。

答：网络上的二维码生成工具或平台很多，如，https://tool.oschina.net/qr/。

四、实验题

1. 利用 Python 实现 EAN-13 的二进制序列编码，并以 6966090118201 为例进行测试。

答：相关程序参考教材的程序 5-2。

2. 利用 Python 的 qrcode 库生成网络链接“https://gr.xjtu.edu.cn/web/xlgui”的二维码。

答：相关程序参考教材的程序 5-5。

```
import qrcode    # 导入 qrcode 库
qr =qrcode.QRCode(version=10,    #设置版本号
error_correction=qrcode.constants.ERROR_CORRECT_L,     #设置容错级别
box_size=20, border = 10 )  #设置二维码图大小、图的边界
img = qr.add_data('https://gr.xjtu.edu.cn/web/xlgui') # 添加二维码数据
qr.make()                #将数据编译成 qrcode 数组
img=qr.make_image()    #生成二维码图片存储在 img 对象中
img.save('qr2.png')      #将 img 存储在硬盘上当前目录下的 qr2.png 文件中
img.show()             #显示二维码
```

程序运行结果如下：

第 6 章 大数据分析与人工智能

一、选择题（参考答案：D、C、A、B、A；B、A、A、C、D）

1. 下列选项中，不属于大数据的特征是（ ）。
   A. 海量 B. 高速 C. 多样 D. 实时
2. SQL 语句中，创建基本表的命令是（ ）。
   A. ALTER B. GRANT C. CREATE D. DELETE
3. 当图片的分辨率为 1024*768，色彩为 16 位时，则该图片占用的存储空间为（ ）。
   A. 1536KB B. 1536MB C. 12288KB D. 以上都不是。
4. 下面属于结构化的数据为（ ）。
   A. 图片 B. 姓名 C. 视频 D. 音频
5. 从关系模式中找出满足给定条件的那些元组称为（ ）。
   A. 选择 B. 投影 C. 连接 D. 查询
6. 从关系模式中挑选若干属性组成新的关系称为（ ）。
   A. 选择 B. 投影 C. 连接 D. 查询
7. 对称加密算法 DES 是（ ）的英文缩写。
   A. Data Encryption Standard B. Data Encode System
   C. Data Encryption System D. Data Encode Standard
8. 如果恺撒置换密码的密钥 Key=4，设明文为 YES，则密文是（ ）。
   A. BHV B. CIW C. DJX D. AGU
9. RSA 的公开密钥(n,e)和秘密密钥(n,d)中的 e 和 d 必须满足（ ）。
   A. 互质 B. 都是质数 C. e*d ≡ 1 mod n D. e*d ≡ n-1
10. 下面不是隐私保护的主要方法的是（ ）。
   A. 匿名 B. 假名 C. 加密 D. 副本

二、问答题

1. 什么是大数据？简要说明物联网大数据的 5V 特征。

答：对“大数据”概念认识可谓“仁者见仁，智者见智”。根据麦肯锡全球研究所的定义，大数据（big data）是一种规模大到在获取、存储、管理、分析方面大大超出了传统数据库软件工具能力范围的数据集合。根据 IBM 公司的观点，大数据是指所涉及的资料量规模巨大到无法透过目前主流软件工具，在合理时间内达到撷取、管理、处理、并整理成为帮助企业经营决策更积极目的的资讯。

物联网大数据正在呈现出 5V 特征，即 Volume（海量）、Velocity（高速）、Variety（多样）、Value（低价值密度）、Veracity（真实性）。

2. 什么是关系数据库？

答：数据的组织结构如果支持关系模型的特性，则该数据库为关系数据库。

3. 什么是云存储？举例说明两种典型的云存储方式。

答：云存储是在云计算概念上延伸和发展出来的一个新的概念，是指通过集群应用、网格技术或分布式文件系统等功能，将网络中大量各种不同类型的存储设备通过应用软件集合起来协同工作，共同对外提供数据存储和业务访问功能的一个系统。

典型的云存储模式包括云盘存储，电子邮件存储等。

4. 什么是数据预处理？预处理包括哪几个过程？

答：不管是通过什么方式获取数据，在进行存储和分析之前，一般都需要进行预处理，取其精华，去其糟粕，目标是为减少存储空间、提高存储与服务效率。数据预处理一般包括数据清洗、数据集成、数据转换和数据归约等。

5. 简要说明分类和聚类的主要区别和联系。

答：分类是一种在分类数已知的情况下的数据分析方法。它使用类标签已知的样本建立一个分类函数或分类模型（也常常称作分类器）。聚类是一种根据数据对象的相似度等指标进行数据分析方法。聚类中，数据对象分成由类似的对象组成的多个类。

6. 简要说明数据安全三原则的含义。

答：数据安全需求随着应用对象不同而是不同，需要有一个统一的数据安全标准。这个标准就是数据安全三原则，即数据机密性（confidentiality）、完整性（integrity）和可用性（availability）三原则（简称 CIA 原则）。数据机密性是指通过加密，保护数据免遭泄漏，防止信息被未授权用户获取，包括防分析。数据完整性是指数据的精确性和可靠性。数据可用性是指期望的数据或资源的使用能力，即保证数据资源能够提供既定的功能，无论何时何地，只要需要即可使用，而不因系统故障或误操作等使用资源丢失或妨碍对资源的使用。

7. 什么是明文？什么是密文？什么是加密？什么是解密？

答：加密是保证数据安全的主要手段。加密之前的信息是原始信息，称为明文（plaintext）；加密之后的信息，看起来是一串无意义的乱码，称为密文（ciphertext）。把明文伪装成密文的过程称为加密（encryption）；将密文还原为明文的过程称为解密（decryption）。

8. 什么是对称加密？什么是非对称加密？二者的主要不同是什么？

答：加密密钥和解密密钥相同的加密技术，称为对称加密；加密密钥和解密密钥不相同的加密技术，称为非对称加密。在对称加密算法中，加密和解密使用同一套规则，密钥 在传递过程中可能泄露。在非对称加密中，加密密钥（公钥）和解密密钥（私钥）不同，公钥在网络上传递，私钥只有自己拥有，不在网络上传递，这样即使知道了公钥也无法解密。

9. 在人工智能的相关技术中，机器学习、深度学习、联邦学习有何关联和区别？

答：机器学习就是对计算机的一部分数据进行学习，然后对另一部分数据进行预测或者判断，换句话说，就是让机器去分析数据找规律，并通过找到的规律对新的数据进行处理。深度学习就是一种为了让层数较多的多层神经网络可以训练，能够运行起来、并演化出来一系列的新的结构和新的方法的过程。联邦学习是一个机器学习的框架，能有效帮助多个机构在满足用户隐私保护、数据安全和政府法规的要求下，进行数据使用和机器学习建模。

三、综合应用题

1. 通过腾讯问卷或问卷星设计一个包含 5 道考试题的问卷，每题 20 分。

答：略。

2. 利用电子表格对下述数据进行分析，给出拟合曲线和拟合函数。

X	1	2	3	4	5	6	7	8	9
Y	2.5	3	4	6	7	9	20	30	40

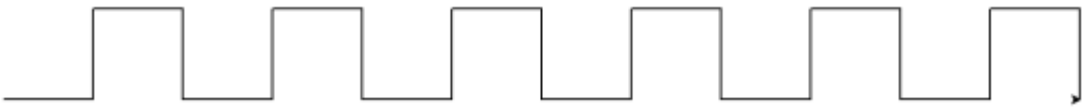
答：略。

3. 利用 Python 的 turtle 库绘制一个方波图形。

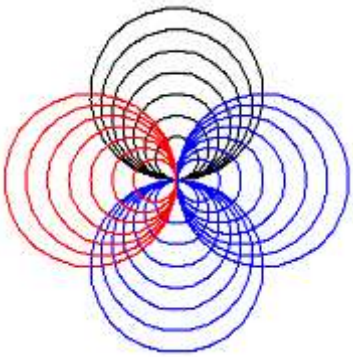
答：具体程序如下。

```
from turtle import *
penup()
goto(-800,0)  #画笔起点
pensize(2)    #笔画大小
speed(0)          #画笔速度
for i in range(6):  #绘制 6 个方波
     pendown()
     forward(100)    #画线功能
     left(90)
     forward(100)    #画线功能
     right(90)
     forward(100)    #画线功能
     right(90)
     forward(100)    #画线功能
     left(90)
```

程序运行结果如下：



4. 利用 Python 的 turtle 库绘制下图所示的多同切圆图形。



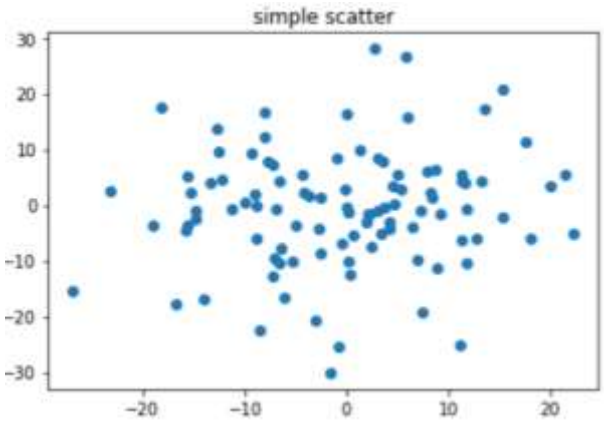
具体程序如下。

```
import turtle
def drawcir(cor):
     turtle.color(cor)
     turtle.pensize(2)
     turtle.circle(20)
```

```
    turtle.circle(30)
    turtle.circle(40)
    turtle.circle(50)
    turtle.circle(60)
    turtle.circle(70)
    turtle.circle(80)
#主程序
drawcir('black')
turtle.left(180)
drawcir('blue')
turtle.left(90)
drawcir('blue')
turtle.left(180)
drawcir('red')
turtle.penup()
turtle.goto(320,0)
turtle.pendown()
drawcir('black')
turtle.left(120)
drawcir('gold')
turtle.left(120)
drawcir('blue')
turtle.hideturtle()
```

5. 利用 Python 的 matplotlib 库绘制一个如下所示的散点图形。

答：



具体程序如下。

```
import matplotlib.pyplot as plt
import numpy as np
fig,ax=plt.subplots() #建立对象
ax.plot(10*np.random.randn(100),10*np.random.randn(100),'o')#设置对象
ax.set_title("simple scatter")
plt.show()
```

6. 利用 Python 的 matplotlib 库绘制一个简单的柱状图。

答：

X	1	2	3	4	5	6	7	8	9
Y	87	56	23	11	29	39	121	108	33

（1）导入模块并且起别名 import matplotlib.pyplot as plt。

（2）设置柱状图的标题，X 轴，Y 轴的显示信息。

（3）设置中文汉字正常显示否则可能出现中文乱码。

（4）设置柱状图的样式，显示颜色，风格等等。

（5）使用 plt.show()显示柱状图。

具体程序如下。

```
import matplotlib.pyplot as plt
X = [1,2,3,4,5,6,7,8,9]
Y = [87,56,23,11,29,39,121,108,33]
# 显示中文
plt.rcParams['font.sans-serif'] = [u'SimHei']
plt.rcParams['axes.unicode_minus'] = False
plt.bar (X, Y,0.4 color='r' )
plt.xlable('X 轴')
plt.ylable('Y 轴')
plt.title ('柱状图')
plt.show()
```

7. 设 26 个英文字母 a～z 的编号依次为 0～25。已知仿射变换为 c=(7m+5) mod 26，其中 m 是明文的编码，c 是密文的编码。试对明文“computer”进行加密，得到相应的密文。

答：首先计算 computer 的字母编号。

字母	a	b	c	d	e	f	g	h	i	j
序号	0	1	2	3	4	5	6	7	8	9
字母	k	l	m	n	o	p	q	r	s	t
序号	10	11	12	13	14	15	16	17	18	19
字母	u	v	w	x	y	z				
序号	20	21	22	23	24	25				

根据上表和仿射变换公式可得下表，可得密文为“tzlgpihu”。

明文字母	c	o	m	p	u	t	e	r
变换前序号	2	14	12	15	20	19	4	17
变换后序号	19	103	89	110	145	138	33	124

模 26 运算后	19	25	11	6	15	8	7	20
密文字母	t	z	l	g	p	i	h	u

8. 在使用 RSA 的公钥体制中，已截获发给某用户的密文为 c=10，该用户的公钥 pk=5，n=35，那么明文 m 等于多少？

答：首先需要求出解密密钥 sk。

已知 N =35，则只有一种可能，即 p*q=5*7=35；(p-1)*(q-1)=4*6=24。

根据公式 sk×pk ≡ 1 (mod (p-1)(q-1))，又 pk=5，所以

5*sk≡ 1 (mod 24)，即 5*sk mod 24 = 1。

由此推测 5*5=25。即 25 除以 24 刚好余 1。所以 sk=5。

密文 $c = m^e \bmod N = m^5 \bmod 35 = 10$。

明文 $m = c^d \bmod N = 10^5 \bmod 35 = 5$。

9. 利用 RSA 算法运算，如果 p=11，q=13，pk=103，对明文 3 进行加密。求 sk 及密文。

答：

答：N = p*q=11*13=143；(p-1)*(q-1)=10*12=120

根据公式 d×e ≡ 1 (mod (p-1)(q-1)) 又 pk=103，所以

103*sk≡ 1 (mod 120)。即 103*sk mod 120 = 1。

103x7=721。721 除以 120 刚好余 1。所以 sk=7。

密文 $c = m^{pk} \bmod N = 3^{103} \bmod 143 = 16$。

四、实验题

1. 使用 turtle 库实现字符“汉”的显示。

答：首先，按照第 1 章习题中的“汉”字字模，参考教材中的程序 6-4 进行实现。

具体程序如下。

```
from turtle import *
zm = [
0x04, 0x04, 0x02, 0xFE, 0x00, 0x4, 0x09, 0x44,
0x05, 0x48, 0x02, 0x28, 0x02, 0x28, 0x0C,0x10,
0x04, 0x28, 0x04, 0x44, 0x05,0x82, 0x00, 0x00
      ]
def draw2Byte(zmdig, size):      #画一行的函数（16 位）
     x = 0x8000
     for i in range(16):          #从高到低按位获取数值
          dig = zmdig // x       #取整
          zmdig = zmdig%x    #模 x 运算
          if dig ==0 :           #点的数值为 0
               penup()               #抬笔
          else:          #点的数值为 1
               pendown()         #落笔
          forward(size)          #前行
          penup()        #抬笔
          x = x//2               #右移
#主程序
```

```
penup()
size = 10   # 字体大小
line = 100      # 字的行位置
goto(-100,line) #画笔起点
pensize(size)   #笔画大小
speed(0)        #画笔速度
for i in range(0,24,2):
     zmdig16 = zm[i]*2**8 + zm[i+1]  #两个字节显示一行
     draw2Byte(zmdig16,size)
     line = line - size + 1  # 缩小行间距
     goto(-100,line)
```

程序输出结果如下：

2. 使用 turtle 库函数，利用人工智能技术，设计一个馋吃蛇游戏。

答：略。