**必修一**

**一、数据、信息、知识、智慧**

1. **数据**是现实世界客观事物的符号记录，是信息的载体，是计算机加工的对象。表现形式主要有数字、文字、图形、声音和视频等。

直接获取：传统方法有调查法、访谈法、观察法、实验法等；

现代化方法有传感器、探测器等设备及相关技术方法；

间接获取：查阅文献资料、关注媒体资源、浏览权威网站、利用搜索引擎在网络中进行检索、从相关数据公司购买等。

1. **信息**是经过加工处理的、具有意义的数据。信息是对客观世界中各种事物的运动状态和变化的反映，是客观事物之间相互联系和相互作用的表征。数据是信息的符号表示；信息是数据的内涵，是对数据的语义解释。

信息的特征： 普遍性，传递性， 依附性，可处理性，时效性，共享性，真伪性，价值相对性。

1. **知识**是信息经过加工提炼后形成的抽象产物。
2. **智慧**是指为了达到预定目标而运用知识解决问题的创新思维能力。
3. 数据、信息、知识和智慧是逐渐递进的概念，前者是后者的基础和前提，后者是前者的抽象与升华。从数据到信息的转化，主要理解数据间的相互关系；从信息到知识的转化，主要理解信息间的相互联系,发现和抽象其规律；从知识到智慧的转化，主要理解和应用知识原理与法则解决现实世界的问题。

**二、数据编码**

1.编码：信息从一种形式或格式转换为另一种形式或格式的过程。

2. 模拟信号与数字信号

用传感器直接获得的信号一般为模拟信号，如声音、温度、压强等。模拟信号的值是随时间连续变化的，波形光滑。模拟信号的值称为模拟数据。

数字信号随时间的变化是非连续的**，**数字信号可以由模拟信号转换得来，数字信号的值即数字数据，直接用计算机所能理解的二进制表示。

对模拟信号进行**采样、量化、编码**后转为数字数据，方便用计算机存储和处理。**采样：时间的离散化、数字化；量化：幅度的离散化、数字化；编码：数据的格式化、文件化。**

3. 图形图像编码：

数字图像有位图图像和矢量图形两种。

矢量图是根据几何特性来绘制图形，只能靠软件生成。它的特点是任意放大后图像不会失真，因为矢量图形与分辨率是无关的，文件大小一般比位图要小。常见的矢量图格式有：\*.ai（Illustrator软件）、\*.cdr (CorelDraw软件)、\*.dwg（AutoCAD软件）等。

位图最小单位为光栅点(或称像素)，由纵横排列的像素点组成，因而也叫做点阵图(或像素图)。它与矢量图的区别在于可以包含丰富的颜色，放大后出现马赛克现象。

常见的位图格式有：\*.bmp、\*.jpg/\*.jpeg、\*.png、\*.psd等。

位图图像的数据量(单位:B)=图像分辨率×图像量化位数÷8 （图像量化位数，也叫图像深度，是指图像中每个像素点记录颜色所用二进制数的位数。）

4. 声音的数字化

数字化音频文件的大小：采样频率(Hz)×量化位数(bit)×声道数×时长(s)÷8

常见的音频文件存储格式：\*.wav 、\*.wma 、\*.mp3、\*.amr等

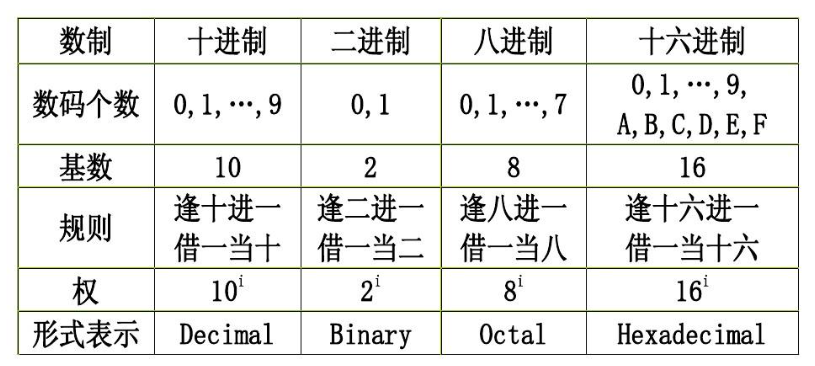
常见的音频文件编辑软件：Adobe Audition、GoldWave等

5.视频编码

常见视频类型静态的图像连续播放就形成视频，常见的视频类型有\*.avi、\*.mp4、\*.wmv、\*mpeg、\*.flv 等。

6. 进制与数制转换

整数转换成R进制数(R≥2)采用“除R取余法”，把R进制非负整数转换成十进制数采用“按权展开求和法”。



7.文本数据的编码

（1） ASCII码( American Standard Code for Information Interchange, 美国信息交换标准码)用7位(标准ASCII码)或8位(扩展ASCII码)二进制数表示一个字符。标准ASCII码字符集中定义了大小写英文字母、标点符号和数字等字符和符号，共128个。计算机内部用一个字节来存放一个ASCII码字符，最高位用0表示。例如，“A” 的ASCII码为01000001，即65。

（2）中文编码

国标码：计算机要处理汉字，必须对汉字进行编码，我国先后制定了多个汉字编码方案，如处理简体中文的GB2312、GBK，用于繁体中文的BIG5（大五码）。它们均使用两个字节表示一个汉字，故又称为双字节码。

机内码：在计算机中为了能同时存储中文和英文，在汉字国标码的基础上稍作修改变成了机内码，用于存储汉字，具有唯一性，即一个汉字只有一个机内码。

输入码：也称汉字外码，是为将汉字输入到计算机中设计的代码。常见的各种输入法，如五笔输入法、搜狗拼音输入法、语音输入法等都属于输入码。

输出码（字形码）：因汉字是象形文字，为了将汉字显示和打印出来，需要一种编码将汉字的字形数字化，即输出码。

（3）Unicode码

是为了解决传统字符编码方案的局限性而产生的。Unicode字符集，是全球可以共享的编码字符集，涵盖了世界上主要文字的字符，其中包括简繁体汉字，共计74686个汉字，通常用两个字节存放编码。

**三、数据的计算**

1. 计算：人类分析问题采用的方法。计算可以看作是一种获得信息的过程。
2. WPS电子表格软件使用：

数据的排序

数据的筛选

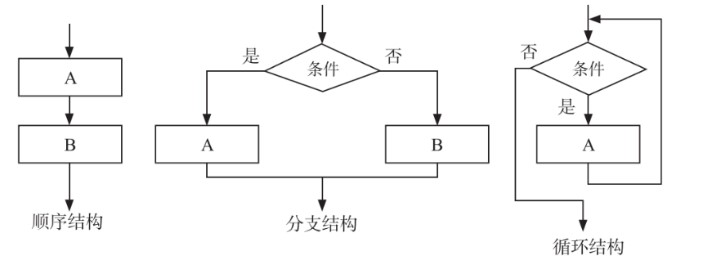
公式函数的应用

数据的分类汇总（先按照分类字段排序，再依据汇总方式计算选定汇总项）

数据的可视化表达：基本图表类型有柱状图、饼图和折线图等。

1. 常见的计算方式：人工方式、面向大众的软件和编程方式。

**四、计算机解决问题的过程**



1. 算法的特性 ：有穷性，确定性，一个算法可以有0个或多个输入，一个算法必须要有输出，可行性。
2. 程序设计语言是人们编制计算机程序所使用的语言。

|  |  |
| --- | --- |
| 类型 | 特点 |
| 机器语言 | 由二进制的0、1代码指令构成，能被计算机直接识别 |
| 汇编语言 | 符号化的机器语言，采用英文助记符代替机器指令，依赖于特定的计算机系统。属于低级语言 |
| 高级语言 | 更接近于自然语言，不依赖于特定的计算机系统，通用性好．可读性强。有Java、C/C＋＋以及Python等 |

高级语言的翻译程序有两种类型：编译程序和解释程序

1. 算法的评价 ：时间复杂度（定性描述一个算法的运行时间） 、空间复杂度（定性描述一个算法或程序运行所需要的存储空间大小）。

4.常见的算法

**解析法**：又称为分析法，是应用解析式去求解数学模型的方法。

**枚举法**：基本原理是根据已知条件，在给定的范围内对所有可能的答案按一定的顺序逐一枚举和验证，从中找到符合要求的答案。 又称穷举法。是一种最为直接，实现最为简单，同时又最为耗时的解决问题的算法思想。

枚举算法的两大要素：枚举对象，判断条件。

枚举解决问题的一般结构：循环＋判断。其优势在于正确性容易证明。

常见程序框架: 　　　　for 枚举对象 in 枚举范围:

　　　　　 　 if 满足判定条件:

　　　　　　　 　 输出解

**二分查找**：又称折半查找，要求数列有序排序，以递增为例，要查找的元素先与数列中点位置的元素进行比较，如果正好是，则查找成功，如果比中心位置的元素小，则将待查序列缩小为左半部分，否则为右半部分，每一次比较后都可以将查找区间缩小一半。

**递归**：直接或间接地调用自身的方法称为递归。面对大规模复杂问题的求解时，递归的基本思想就是把规模较大的问题层层转化为规模较小的同类问题求解。递归主要有两个组成要素，一个是递推关系，一个是边界条件。在计算机编程中，我们经常将递归和自定义函数联系起来.

递归的经典应用：斐波那契数列和汉诺塔**。**

F(n)=

**迭代**（概念性）：迭代通常是为了接近并达成所需的目标或结果，每一次对过程的重复称为一次“迭代”，而每一次迭代得到的结果会被用来作为下一次迭代的初始值。

迭代算法思想：迭代法也称辗转法，每一次对过程的重复称为一次“迭代”，而每一次迭代得到的结果用来作为下一次迭代的初始值，通常是为了**接近并到达**所需的目标或结果。

使用迭代算法解决问题，有三个关键步骤：(1) 确定迭代变量；(2) 建立迭代关系式；(3) 对迭代过程进行控制。

**数值计算**

1、绘制函数图像：一般要用到numpy和matplotlib两个模块

**①**numpy是一个科学计算包，其中包括很多数学函数，如三角函数、矩阵计算方法等。通过该模块中的arange函数可以创建一个等差数列**。**

例：arange(0，4\*numpy.pi，0.1)可以实现0—4π之间每隔0.1取一个值。其中numpy.pi表示π。

**②**matplotlib模块是最出色的绘图库，调用matplotlib.pyplot时，坐标系可以根据数值范围自动生成。matplotlib的绘图原理很简单，利用plot画线函数就可以在直角平面内轻松地将(x，y)坐标点连接成平滑曲线。

**非数值计算** 分治策略

1. 非数值运算更多探讨“算法”问题

2. 分治策略：将一个难以直接解决的大问题，分割成一些较小的同类问题，各个击破。分治策略的典型运用：二分查找与递归。 可用“分”“治”“合”三个字概括。

3 查找的基本算法有：顺序查找、二分查找、分块查找、哈希查找等。

**五、程序设计语言**

**（一）Python基础**

1. 常量：在程序运行过程中值始终保持不变的量。所有的常数都是常量，例如π。常量可以是数字、字符串、布尔值、空值None(和 False 不同，它不表示 0，也不表示空字符串，而表示没有值，也就是空值)。

2. 变量：在程序运行过程中值允许改变的量。

变量的命名规则：

(1) 必须以字母或下划线开头，后面可以是字母、数字或下划线。

(2) 区分大小写。

(3) 不能使用保留字，如if、for、False、True、and等。

(4) 不能包含空格。

3. 运算符与表达式

常用的算术运算符如下表所示：

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 算术运算符 | + | - | \* | \*\* | / | // | % |
| 功能 | 加 | 减 | 乘 | 幂 | 除 | 整除 | 取模 |
| 算术表达式 | a+b | a-b | a\*b | a\*\*b | a/b | a//b | a%b |

表达式由一个或多个操作数通过运算符组合而成。

4. 赋值语句

基本格式：变量＝表达式

赋值运算符：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 关系运算符 | += | -= | \*= | /= | %= |
| 操作 | x+=y即x=x+y | x-=y即x=x-y | X\*=y即 x=x\*y | x/=y即 x=x/y | x%=y即x=x%y |

5. 输入语句：input()，表示从键盘输入数据（得到的数据类型是字符串型）。

6. 输出语句：print()，表示在屏幕上输出数据。如有多个输出项，可以使用英文逗号分隔。

7. 数据类型的转换函数：

(1) float()：根据一个字符串或整数创建一个新的浮点数(小数)

(2) int()：根据一个字符串或浮点数创建一个新的整数

(3) str()：根据一个数(也可以是任何其他类型)创建一个新的字符串

(4) len()：获取字符个数

(5) ord()：将字符类型的参数转换为对应的 ASCII 数值或者 Unicode 数值

(6) chr()：将整数类型的参数转换为对应的字符（常和ord()配对使用）

8.列表类型

列表：是一个有序的集合，它可以包含多个数据元素，列表中的元素允许是不同的数据类型。

(1) 创建列表：列表名＝[　 ，　 ，　 ，　 ，　 ]

(2) 访问列表中的元素：列表名[索引号]

正负索引：从左往右第一个元素索引号为0；从右往左第一个元素索引号为-1。

(3) 列表末尾增加元素：**列表名.append()**

(4移除列表中元素（默认最后一个元素）：**列表名. pop()**

(5) 列表中**插入**元素：**列表名.insert( ，)**

(6) 删除列表中的某个元素：del 列表名[索引号]

(7) 修改列表中的元素：列表名[索引号]=

(8) 反向列表中的元素（**逆序**）：**列表名.reverse()**

(9) 列表**排序**：**列表名.sort()**

**（二）分支**

|  |  |  |
| --- | --- | --- |
| 分支类型 | 基本格式 | 流程图 |
| 单分支 | if条件：语句或语句组1 |  |
| 双分支 | if条件：语句或语句组1  else：语句或语句组2 |  |
| 多分支 | if条件1：语句或语句组1  elif条件2：  　语句或语句组2  ……  elif条件N：  　语句或语句组N  else：语句或语句组N＋1 |  |

常用的关系运算符如下表所示:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 关系运算符 | == | > | >= | < | <= | != |
| 含义 | 等于 | 大于 | 大于等于 | 小于 | 小于等于 | 不等于 |
| 关系表达式 | a==b | a>b | a>=b | a<b | a<=b | a!=b |

**（三）循环**

1. 计数循环

**for 循环变量 in 列表或range() ：** range()函数：生成数字序列

**循环体**

range()函数使用格式：range(起,终,步长)，默认从0开始，左闭右开，步长为1时可以省略。

2. 条件循环

**while 关系表达式：**

**语句或语句组**

当关系表达式一直为真时，条件循环将会一直执行下去，一般我们称之为“死循环”。

3. 循环的提前结束

如果想提前结束循环(在不满足结束条件的情况下结束循环)，可以使用**break或continue**语句。

4. 循环嵌套

Python语言允许在一个循环体里面嵌套另一个循环。

当2个(甚至多个)循环结构相互嵌套时，位于外层的循环结构常简称为“外层循环”或“外循环”，位于内层的循环结构常简称为“内层循环”或“内循环”。break 只有中断当前所在循环。

**（四）函数**

1. 函数是一个能完成某一独立功能的子程序，也就是程序模块。

函数是对复杂问题的一种“自顶向下，逐步求精”思想的体现。

2. 系统函数 ：函数库里的标准函数,是程序设计语言或操作系统提供给用户的一系列已经编写好的程序。如print()函数和input()函数，运算类函数abs()、int()、float()、sqrt()、round()、sum()、str()、ord()等，函数无需导入，直接使用。

3. 自定义函数

**def 函数名(形式参数)：** 当有多个参数时，各参数之间用逗号隔开。

**函数体(语句或语句组)**

**return 返回值**

其中参数用来向函数传递值，当有多个参数时，各参数之间用逗号隔开，调用时要注意参数顺序；函数执行完成后，由return语句将表达式的值返回给调用者，结束函数。

**函数调用格式：函数名(参数)**

4. ****Python标准库：****Python自带的标准库。Python标准库无需安装，只需要先通过import方法导入便可使用其中的方法。time、math、OS、random等。

主程序中导入库：import　发布的程序库名，如：import math

程序体中调用：库名．函数名()，例如time.time()、random.randint(0,1)等。

**（五）\*数据与结构**

1. 数据类型

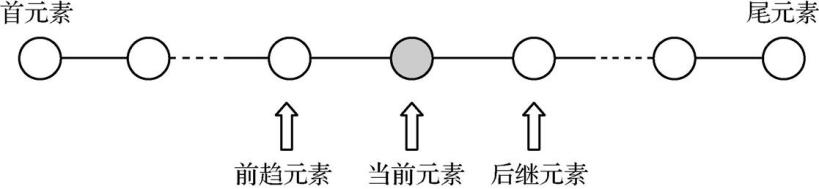
**简单数据类型**指不能分解为更小的数据类型。如**整型int、浮点数float、字符串str、布尔型bool**。

**复合数据类型**指由简单数据类型或者复合数据类型组成。如**列表list、字典dict**等**。**

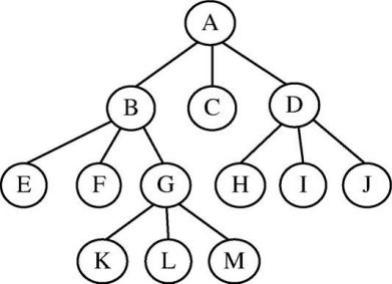
2. 数据结构：是存在特定关系的数据元素的集合

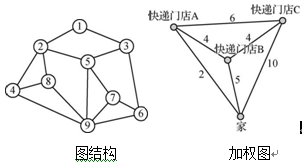
**数据结构**也称为逻辑结构，**主要有集合结构、线性结构、树结构和图结构**(又称网状结构)四种类型。

3.线性数据结构：又称线性表。在线性数据结构中，除首元素没有前趋元素、尾元素没有后继元素外，其他元素都只有一个前趋元素和一个后继元素。线性表中数据元素之间是一对一的关系。



4. 队列：是一种有限制的线性结构，他的数据元素只能在一端依次添加(进队)，在另一端依次删除(出队)。其特点是“先进先出”。

5. 树结构：树结构是一种具有层次关系的非线性结构。树是由n(n>＝0)个节点组成的有限集合。



树结构

1. 图结构：是由一组节点(称为顶点)和一组节点间的连线(称为边或弧)构成的一种数据结构。图结构中的每个顶点都可以与其他顶点有边相连**。**

|  |  |  |  |
| --- | --- | --- | --- |
| 结构类型 | 结构特点 | 数据(节点)之间的关系 | 生活中相应结构应用举例 |
| 队列(线性) | 线性结构 | 一对一 | 排队付款、过马路、医院就诊电子牌的队列…… |
| 树 | 非线性结构 | 一对多 | 行政区划、书的目录结构、磁盘文件存储结构…… |
| 图 | 非线性结构 | 多对多 | 铁路运输图、高速公路网、全国航运图…… |

**（六）图形用户界面**

1. 图形用户界面：又称图形用户接口(GraphicalUserInterface,**GUI**)，是指采用图形方式显示的计算机操作用户界面。在Python语言中,tkinter是实现图形用户界面的模块。可以快速创建GUI应用程序。

操作步骤：

(1) 导入**tkinter**模块；（2)创建交互对象：使用相应代码进行创建。如矩形容器(Frame)、文本框(Entry)、命令按钮(Button)、文本标签(Label)等。

参考代码如下：

　　from tkinter import\*　# 导入模块

　　root=Tk()　# 创建一个窗口

　　root.title("旅游线路")　# 设置窗口标题

　　root.geometry(id:2147492647;FounderCES300x120id:2147492654;FounderCES)# 设置窗口大小

root.resizable(0,0)# 禁止调整窗口大小

# 在窗口上建一个“添加线路”按钮

Button(root,text="添加线路",relief="solid",width=10,height=2).pack()

2. 对象：用来建立计算机世界的实体与客观世界的实体相一致的映射关系。问题的解决方案可以用一些彼此交互的对象来形象化表示。

在面向过程的程序设计中，我们关注的是各种类型的数据及其操作；而在面向对象的程序设计中，将数据及其操作隐藏细节后封装为类，函数调用式的计算演变为对象之间的交互。

**（七）文件**

1. 文件：在解决问题的过程中，我们通常将数据组织在一起形成队列、树和图等逻辑结构,将其保存在内存中，程序退出后相关数据也会丢失。如果需要更持久地存储数据，可以将数据以文件形式存储到外存储设备中。文件是存储在外存中相关数据的集合。

2. 文本文件的读写：在Python语言中，open函数可以将文本文件(.txt)打开，用write方法将内存中的数据存储到外存储器的文件中，操作结束时一定要用close方法关闭文件。

### 六、数据处理与可视化

**（一）.数据采集**

数据采集，即根据需求采用适当的方法和工具获取所需要的数据。数据采集的主要环节包括明确数据需求，确定数据来源，选择采集方法和实施数据采集。

直接数据采集 、间接数据采集 、系统日志采集法 、系统日志采集数据、网络数据采集法（网络爬虫、API） 、其他数据采集法 （如感知设备数据采集，它是指通过传感器、摄像头和其他智能终端自动采集信号、图片或音视频来获取数据。 ）

**（二）数据整理**

数据整理的目的是对数据进行校验和标准化。数据整理就是通过去重、补漏和勘误等方法，删除重复数据、补全缺失数据和校正错误数据，并对数据进行统一性和标准化处理，以确保数据是相关和准确的。

删除重复数据 、补全缺失数据 、校正错误数据

**（三）数据分析**

1. 数据分析：用恰当的统计分析方法对收集来的大量数据进行分析，提取有用信息，并形成结论的过程。

2. 数据分析的一般过程：根据分析的目标提出假设，然后收集有关数据，选用恰当的分析方法进行分析，验证假设是否正确，继而得出相应的结论。

3. 数据分析的基本方法

（1） 对比分析： 对比分析是将两个或两个以上的数据进行比较，分析它们的差异，从而揭示这些数据所代表的事物发展变化情况和规律。

对比分为横向对比和纵向对比。横向对比指的是类似的事物或者同类的事物之间进行比较。纵向对比指的是相同事物的不同时期进行比较。

（2） 平均分析： 平均分析是运用计算平均值的方法，来反映总体在一定时间、地点条件下某一数量特征的一般水平。平均指标中最常用的是算术平均数，如平均身高、平均工资和平均降水量等。

平均分析和对比分析常结合使用。

（3）结构分析法：也称构成分析法，是将各个部分与总体进行对比，是分析事物内部的结构和部分与整体之间关系的方法。

结构分析法的基本表现形式就是计算结构指标，结构指标即各个部分相对于总体所占的百分比。

（4）关联分析法：就是分析并发现存在于大量数据之间的关联性或相关性，从而描述一个事物中某些属性同时出现的规律和模式，如购物车案例（啤酒与尿布）

（5） 聚类分析法 :聚类分析是一种探索性的分析，在分类的过程中，人们不必事先给出一个分类的标准，聚类分析能够从样本数据出发，自动进行分类。聚类分析的算法有很多，其中K-平均（K-Means）算法是一种经典的自上而下的聚类分析方法。

（6）数据分类法 :数据分类通常的做法是，基于样本数据先训练构建分类函数或者分类模型，该分类器具有将待分类数据项映射到某一特点类别的功能。

贝叶斯分类技术在众多分类技术中占有重要地位。贝叶斯分类技术通过对已分类的样本子集进行训练，学习归纳出分类函数，利用训练得到的分类器实现对未分类数据的分类。

4. 数据可视化表达：以图形、图像和动画等方式更加直观生动地呈现数据及数据分析结果，揭示数据之间的关系、趋势和规律等的表达方式。

图表是最常用的数据可视化表达方式之一。基本图表类型,如柱状图、饼图和折线图等。

有关趋势的分析：折线图、分组柱形图、堆叠柱形图、柱线混合图

有关比例的分析：百分比圆环图、饼图、圈图、堆叠面积图

有关关系的分析：逻辑关系（散点图、雷达图、网络关系图、词云）、空间关系（数据地图、动态热力图、3D动态显示图）

5. 数据分析报告

项目研究结果的展示，也是数据分析结论的有效承载形式。通过报告不仅是把数据分析的起因、过程、结果及建议完整的展现出来，更为决策者提供科学、严谨的决策依据。

在数据分析报告中，首先需要明确数据分析的目的和背景，阐述目前存在的问题及通过分析希望解决的问题；其次需要描述数据来源和数据分析的思路、方法和模型；最后需要重点呈现数据分析的过程、结论和建议。

6**.** 大数据

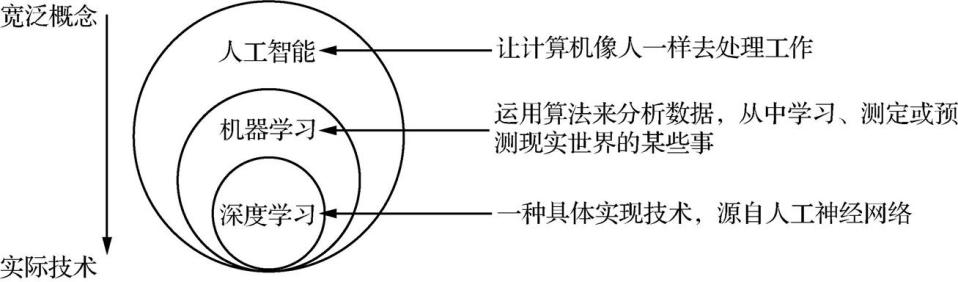
以容量大、类型多、存取速度快、应用价值高为主要特征的数据集合，它正快速发展为对数量巨大、来源分散、格式多样的数据进行采集、存储和关联分析，从中发现新知识、创造新价值、提升新能力的新一代信息技术和服务业态。

大数据的“大”是一个相对概念，没有具体标准，如果一定要定一个标准，那么10TB－100TB通常称为大数据的门槛。

大数据分析的一般流程：

大数据采集 分布式存储 分布式处理 分析 数据可视化

**七、人工智能**



1. 人工智能(AI)：是指利用计算机或计算机控制机器模拟、延伸和扩展人的智能，感知环境，获取知识并使用知识进行分析、辨别的技术与应用。

2. 人工智能分类：弱人工智能和强人工智能，目前处于弱人工智能时代。强人工智能是指能思考，有知觉，有自我意识的人类级别的智能机器。

3. 常见技术：

* **搜索：分盲目搜索和启发式搜索。**盲目搜索，按预定控制策略进行搜索，不考虑问题本身特性。启发式搜索，在搜索过程中加入评估函数等启发信息，不断自动调整搜索方向，加速求解过程。
* **机器学习**，是当前人工智能的核心技术之一，目的是使计算机模拟或实现人类的学习行为，获取新的知识或技能，重新组织已有的知识结构，从而不断完善自身的性能。
* **神经网络**，是人工智能领域的研究热点之一，是一种模拟动物神经网络行为特征，进行并行信息处理的算法模拟。它通过自身训练、学习某种规则，最终在给定输入时得到最接近期望值的输出结果。
* **智能传感器**，具有信息处理功能的传感器，智能传感器带有微处理器，具备采集、处理、交换信息等功能，是传感器集成化与微处理机相结合的产物。是人工智能的神经末梢，用于全面感知外界环境。例如，智能隐形眼镜、葡萄糖手表等。发展趋势为高敏度、高精度、高可靠性、微型化和集成化等。

4. 人工智能的关键技术：主要包括机器学习、知识图谱、自然语言处理、计算机视觉、人机交互、生物特征识别、虚拟现实与增强现实等。

5. 应用领域：当前人工智能的主要应用于自动驾驶、车牌识别、语音识别、人脸识别及智能机器人等。

**必修二**

**一、信息社会**

信息社会四个方面的主要特征：数字生活、信息经济、网络社会、在线政府。

物质、信息和能量是信息社会的三大资源。

信息技术简称IT，一切与信息的获取、加工、表达、交流、管理和评价等有关的技术都可以称之为信息技术。计算机与网络通信技术是信息技术的核心技术。

**二、信息系统的组成和功能**

信息系统额组成

1. 信息系统的定义

信息系统(Information System)是由计算机硬件和软件、网络和通信设备、信息资源、信息用户和规章制度组成的以处理信息流为目的的人造系统。

2. 信息系统的基本组成

信息系统的基本组成：人、硬件、软件、数据和通信方式。

(1) 信息系统以人为主导

(2) 信息系统以计算机技术和网络技术为基础

信息系统的运行依托于计算机硬件、软件、数据库系统和网络技术等。

* 计算机硬件是信息系统的物质基础，数据的输入、输出、存储、计算等都要依靠计算机硬件设备的支持。由程序和文档构成的软件则起到对计算机资源进行管理和组织的作用。
* 数据库系统是提供存储、访问、维护数据服务的信息处理系统。
* 计算机网络则是实现资源共享和信息交换的重要基础。

(3) 信息系统处理的是数据

3. 信息系统的分类：简单系统、复杂系统和随机系统**。**

4. 信息系统对社会发展的影响

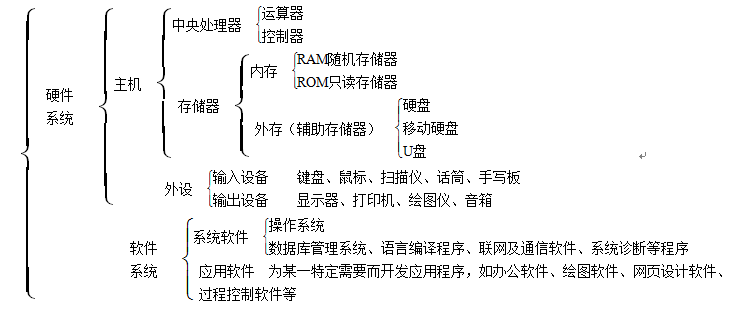
信息系统是管理各种信息的重要手段，它给人们带来了诸多便利，促进了社会的发展。但是，随着信息系统应用的日益普及，人们对它的依赖性增强、安全威胁增加、安全事件频出，信息系统的局限性也逐渐地显现出来。

信息系统的功能

* **事务处理系统**(Transaction Processing System, TPS)：能在数据(信息)发生处将它们记录下来，保存到数据库中供其他信息系统使用，实现预定的业务流程，提高事务处理效率并保证其正确性。显著特点是具有在线处理功能，能够远程提交或取消事务。
* **管理信息系统**(Management Information System, MIS)：是一个以人为主导，利用计算机硬件、软件、网络通信设备以及其他办公设备，进行信息的收集、传输、加工、储存、更新、拓展和维护的系统。
* **决策支持系统**(Decision Support System, DSS)：是以管理科学、运筹学、控制论和行为科学为基础，以计算机技术、仿真技术和信息技术为手段，针对决策问题支持决策活动的具有智能作用的人机系统。该系统能够为决策者提供所需的数据、信息和背景资料，帮助明确决策目标和进行问题的识别，建立或修改决策模型，提供各种备选方案，并且对各种方案进行评价和优选，通过人机交互功能进行分析、比较和判断，为正确的决策提供必要的支持。
* **专家系统**(Expert System, ES)：是一种在特定领域内具有专家水平解决问题能力的程序系统。它能够有效地运用专家多年积累的经验和专门知识，通过模拟专家的思维过程，解决原本需要专家才能解决的问题。

信息系统的局限性：对外部环境有依赖，本身有安全隐患，可能加剧数字鸿沟。

1. **信息系统的集成**

**1.计算机系统的组成** 

* 冯·诺依曼计算机体系结构的基本工作原理是存储程序和程序控制。
* 衡量计算机速度的性能指标有主频、运算速度、存储周期、字长等
* 常见操作系统：Windows、Linux、macOS、iOS、Android、HarmonyOS等

1. **计算机系统互联**

计算机网络在信息系统中主要起到数据传输、资源共享、分布式处理三个方面的作用。

* 带宽：用来表示网络通信线路传送数据的能力。单位时间内从网络中的某一个点到另一个点所能通过的“最大数据量”。带宽的单位是比特每秒(b/s)。
* 速率单位为(b/s比特每秒，或bps)，通常描述的网速100M，实际的速率100Mbps，如下载100MB的文件需要时间为：100MB\*8/100Mbps。
* 网络协议：在计算机网络中进行数据交换而建立的规则、标准或约定的集合

TCP/IP协议(传输控制协议/网络协议)：网络中最基本的通讯协议。TCP/IP协议不是单一协议，而是一个协议组。使用TCP/IP协议，需要详细设置IP地址、子网掩码、默认网关、DNS服务器等参数**。**

**FTP协议：文件传输协议**

地址格式：ftp：//用户名：密码@FTP服务器IP地址或域名

**HTTP协议：超文本传输协议**

地址格式：http：//服务器IP地址或域名

* **IP地址格式：32位的二进制数组成，有网络地址和主机地址两部分。采用点分十进制表示法**(a.b.c.d)，其中a，b，c，d都是**0~255之间的十进制整数**。其中A、B、C三类地址由InternetNIC在全球范围内统一分配，D、E类为特殊地址**。**

|  |  |  |  |
| --- | --- | --- | --- |
| 类别 | 最大网络数 | IP地址范围 | 单个网段**最大主机数** |
| A类 | 126(27－2) | 1.0.0.1—126.255.255.254 | 16777214 |
| B类 | 16384(214) | 128.0.0.1—191.255.255.254 | 65534 |
| C类 | 2097152(221) | 192.0.0.1—223.255.255.254 | 254 |

* 公有地址：由国际互联网信息中心(INIC)负责，向其申请，通过IP地址可以访问互联网
* 私有地址：属于非注册地址，专门供组织机构内部使用。以下是留用的内部私有地址

A类：10.0.0.0—10.255.255.255

B类：172.16.0.0—172.31.255.255

C类：192.168.0.0—192.168.255.255

* **子网掩码：是32位二进制**，是一种用来指明一个IP地址的哪些位标识的是主机所在的子网。子网掩码不能单独存在，**它必须结合IP地址一起使用。子网掩码的作用是将某个IP地址划分成网络地址和主机地址两部分。**

|  |  |  |
| --- | --- | --- |
| 类别 | 子网掩码的二进制数值 | 子网掩码的十进制数值 |
| A类 | 11111111 00000000 00000000 00000000 | 255.0.0.0 |
| B类 | 11111111 11111111 00000000 00000000 | 255.255.0.0 |
| C类 | 11111111 11111111 11111111 00000000 | 255.255.255.0 |

* **网关**：从一个网络向另一个网络发送信息，也必须经过一道“关口”，这道关口就是网关。比如网络设备路由器**。**
* **DNS服务器**：DNS(Domain Name System，**域名系统服务器主要用于域名解析**，**进行域名(domain name)和与之相对应的IP地址 (IP address)转换的服务器。**
* 域名：一串用点分隔的字符组成的互联网上某一台计算机或计算机组的名称。国内域名注册归中国互联网络信息中心(CNNIC)管理。

　.com　商业机构　.net　网络组织　.org　非营利性组织　.gov　政府机构　.edu　教育机构

**3.无线局域网的组建**



无线网络名称（SSID）

1.常见网络设备：路由器、交换机、网卡、集线器、中继器、网桥等。

(1) 网卡：分有线网卡和无线网卡。

(2) 交换机：具备自动寻址能力和数据交换作用的设备**。**连接在同一个交换机上的各个设备处于同一个网络中**。**

(3) 路由器：网络中进行网间连接的关键设备，它的每个端口都可以连接一个网络。路由器的主要工作是为经过路由器的每个数据包寻找一条最佳传输路径，并将数据有效地传送到目的站点。

(4) 网络连接：通过传输介质把网络设备连接起来,构成网络。

2. 无线网络：指应用无线通信技术将计算机设备互联起来，构成可以互相通信和实现资源共享的网络体系。

3. 常用的传输介质：有线传输介质、无线传输介质。无线传输介质：微波、红外线和激光等；有线传输介质：双绞线、同轴电缆、光纤等。

4、无线网络接入方式：移动通信网络接入、Wi-Fi（Wireless Fidelity）接入、蓝牙技术（Bluetooth Technology）、红外数据传输（IrDA）、ZigBee（是一种新兴的短距离、低功率、低速率无线接入技术）、．可见光无线通信技术（简称Li-Fi）、近场通信技术NFC。

**4.基于物联网的信息系统**

1. **物联网的概念**：通过二维码识读设备、射频识别装置、红外线感应器、全球定位系统和激光扫描器等信息传感设备，按约定的协议，将任何物品与互联网相连接，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。

2. 物联网的体系结构：**感知层、网络层与应用层**

(1) 感知层

感知层是物联网的基础，是联系现实物理世界与虚拟信息世界的纽带。

感知层设备的主要分类：自动感知设备、人工生成信息设备

(2) 网络层

网络层也称“传输层”，将来自感知层的各类信息通过基础承载网络传输到应用层，相当于人的神经系统。

目前，物联网的网络层都是基于现有的通信网和互联网建立的。

(3) 应用层

物联网的应用层可进一步分为管理服务层与行业应用层。

3. 支持物联网的共性技术：信息安全、网络管理、对象名称解析服务、服务质量。

4. **物联网应用的相关技术**：传感器技术、RFID技术、嵌入式系统技术、移动通信技术、智能数据处理技术、智能控制技术、位置服务技术、云计算技术、信息安全技术**。**

(1) **传感器**：

按功能分：光敏传感器、声敏传感器、气敏传感器、化学传感器、压敏传感器、热敏传感器**。**

常见传感器：位移传感器、形变传感器、重力传感器、加速度传感器、湿度传感器、温度传感器、超声波测距传感器等。

(2) **RFID技术**：即Radio Frequency Identification，**射频识别技术**，通过无线射频方式进行非接触双向数据通信，利用无线射频方式对记录媒体(电子标签或射频卡)进行读写，从而达到识别目标和数据交换的目的。

5. 大数据与物联网的关系：物联网推进大数据，大数据助力物联网。物联网中可以通过大量传感器采集实时数据，数据经过分析、处理，结合程序可以进行智能预测和控制。

6.物联网的应用领域有智能家居、智能交通、智能医疗、智能电网、智能物流、农业物联网、智能电力、智能安防、智能汽车、智能建筑、智能水务、智能商业、智能工业等。

**5、远程控制系统的组建**

**6、小型信息系统的组建**

### 四、信息系统的软件与应用

**（一）信息系统的设计**

* 信息系统生命周期，是信息系统的产生直到报废或停止使用的生命周期。信息系统生命周期内有问题定义、可行性分析、概要设计、系统设计、编码、调试和测试、验收与运行、维护、报废阶段。
* 瀑布模型：开发过程是通过设计一系列阶段顺序展开的，从系统需求分析开始直到产品发布和维护，每个阶段都会产生循环反馈。信息系统开发从一个阶段流到下一阶段，当一个阶段出现差错，会被下面的各个阶段放大。

该模型特点：① 必须等前一阶段工程完成后，才能开始后一阶段的工作；② 前一阶段的输出文件是后一阶段的输入文档。只有前一阶段正确，后一阶段才能得到正确结果。相邻两个阶段具有因果关系。

#### （二）信息系统的体系结构及工作过程

客户机/服务器结构（Client/Server结构，简称C/S结构）

浏览器/服务器结构（Browser/Server结构，简称B/S结构）

对等网络结构（P2P）

#### （三）信息系统的软件及其作用

##### 操作系统（OS）

操作系统（Operating System,简称OS）是管理和控制计算机硬件与软件资源的计算机程序，是直接运行在“裸机”上的最基本的系统软件，任何其他软件都必须在才做系统的支持下才能运行。操作系统是用户和计算机的接口，同时也是计算机硬件和其他软件的接口。

目前，服务器操作系统主要有Unix、Linux、Windows，桌面客户端操作系统有Macintosh、Unix、Linux、Windows等，移动终端操作系统有iOS、Android（以Linux为基础）等。

* **数据库系统**

1. 概念**：**数据库是以一定的组织方式存储在计算机中的相互关联的数据集合。

2. **数据表**：数据是以二维表的形式组织存储的。表中的**一列称为一个字段**，每个字段都有一个字段名；表中的**一行称为一条记录**，保存有关实体属性的数据；可以唯一确定一条记录的字段称为主键。可以对数据表进行插入、删除和更新等操作。

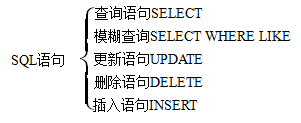
3. **数据库管理系统**(Database Management System，DBMS)，是定义、创建、维护数据库的一种工具，它允许用户管理和控制数据库中的数据。常见的数据库管理系统有Sybase、ORACLE、MySQL、Access等。

4. 表间关系及查询：表间关系就像绳子将两个“表”联结在一起，可以利用“查询”面板中的“查询设计”选项按照提示创建选择查询,得到查询结果。

查询准则：如果只是简单地查找某个字段为特定值的记录，只要在该字段对应的“条件”栏中输入该值即可；如果不仅是查找某个特定值，在准则中可以使用表达式。在表达式中可通过操作符设置查询范围，主要的操作符包括AND、OR、BETWEEN…AND、LIKE等。

5. **数据库系统是由数据库DB、数据库管理系统DBS(及其应用开发工具)、数据库应用系统DBAS、数据库管理员ADMIN和用户USER组成的存储、管理、处理和维护数据的系统。**

**6.**结构化查询语言简称SQL，是一种数据库查询和程序设计语言，用于存取数据以及查询、更新和管理关系数据库系统。



**五、****信息系统的安全风险防范**

**1.信息系统安全风险**

**2.加密与解密**

1. 密码的概念：指用来核对用户ID以验证用户就是本人的一组字符。

2. 密码破解技术中最基本的方法就是暴力破解，也叫密码穷举

3. 创建安全密码的一般技巧

●使用长度不少于8个字符的密码。密码长度越长越不容易被破解。

●在可能的情况下，尽量使用字母、数字和特殊字符(如 ，#)相结合的密码。

●不要使用电话号码、身份证号或生日等信息作为密码。

●不要使用整个用户ID或用户ID的一部分作为密码。

●不要使用字典中能找到的词语作为密码，即使是字母次序颠倒过来的常用词语也不可以。

为保障数据安全，还可以采取限制登录尝试次数、使用验证码、使用用户的具有排他性的自然特征(如指纹、人脸等)作为身份验证依据等措施。

4. 数字签名技术

数字签名技术是将摘要信息用发送者的私钥加密，与原文一起传送给接收者。**接**收者只有用发送者的公钥才能解密被加密的摘要信息，数字签名是个加密的过程，数字签名验证是个解密的过程。

5. 加密的概念

加密：就是将原始信息(数据)隐匿起来，使之在缺少特殊信息(数据)时不可读。

解密：将密文还原成明文的过程称为解密(或解码)。

明文：加密前的原始信息(数据)称为明文。

密文：加密后的信息(数据)称为密文。

7. 加法密码

明文中的所有字母都在字母表上向后(或向前)按照一个固定数目进行偏移后被替换成密文，这种密码称为加法密码，又称为移位密码。

8. 恺撒密码

恺撒密码作为一种最古老的对称加密体制，在古罗马的时候就已经很流行，它是加法密码的代表。

恺撒密码只是简单地将明文中的每一个字母用字母表中该字母后的第3个字母替换。例如，将明文中的a用d替换，b用e替换，……，z用c替换。

**3.信息系统安全技术**

计算机病毒是在计算机程序中插入的破坏计算机功能或者数据的一个程序、一段可执行代码，具有传播性、隐蔽性、传染性、潜伏性、可触发性、表现性和破坏性等特性。

计算机病毒防护常识：

（1）安装杀毒软件，定期进行全盘病毒、木马扫描，经常更新杀毒软件。

（2）安装防火墙或者安装自带防火墙的杀毒软件。

（3）经常进行操作系统和重要软件的更新，尤其要注意安装安全补丁以修复漏洞。

（4）不随意接受、打开陌生人发来的电子邮件或通过QQ传递的文件或网址。

（5）使用移动存储器前，先查杀病毒。

（6）不要轻易打开电子邮件附件中的文档文件。

（7）禁用浏览器中不必要的加载项。

（8）经常查看系统的启动项，去除不必要的启动服务。

**4.信息系统安全管理**

|  |  |
| --- | --- |
| 阶段 | 信息系统安全管理措施 |
| 事先防御 | 网络隔离、访问控制、加密传输 |
| 实时监测 | 病毒监控、入侵检测、系统/用户行为监控 |
| 事后响应 | 报警、急救、取证、问责、修复、加固 |

**5.信息社会的伦理与道德**

**6.信息社会的法律与法规**

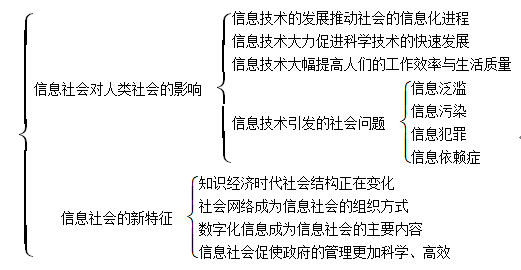
* Cookie：是网站在用户访问时植入计算机中的非常小的文本文件，这些文件含有识别用户信息的数据。当该用户再次访问的时候，网站会自动呈现与上次用户浏览内容相关的内容或产品广告等。许多网站的隐私政策都会声明是如何使用cookie的。
* “第三方”是个法律术语。 在访问网站时，你是第一方，网站是第二方，其他的人或组织是第三方。
* 信息公开的方式：

主动公开：要依法向公众公开相关的政府或其他组织的相关信息。对于主动公开的信息，一般会声明公开范围、公开方式和公开期限。

依申请公开：指行政机关根据公民、法人或者其他组织的申请，依照法律规定和本机关的职权，向申请人公开政府信息的行政行为。

救济方式：指当事人之间出于各种原因发生争议或纠纷后，当事人一方或双方（多方）为解决争议或纠纷而采取的各种行为的形式。

* 知识产权主要有两类：一类是著作权（也称为版权、文学产权），另一类是工业产权（也称为产业产权）。

**7.信息社会的未来发展**