

# 上海师范大学天华学院

SHANGHAI NORMAL UNIVERSITY TIANHUA COLLEGE



---

## 《人工智能专业》

## 2022 年春季招生

---

## 考试大纲



上海师范大学天华学院 人工智能专业

2022 年 1 月

## 大纲简介：

结合天华学院“专、通、雅”协调发展与“16条学生能力与素质培养标准”的要求，通过《信息科技》课程考试，了解学生掌握信息技术的基本概念、原理和应用，进而了解学生的学习能力和学习效果，以确保学生入学后能顺利进行后续专业课程的学习。要求学生学习信息技术基本知识点，掌握信息技术相关的具体操作技能，具有较高的信息素养和创新能力。

该大纲选用《高中信息科技（第一册）》，华东师范大学出版社，并在此基础上进行适当调整。

1. **考核形式：**闭卷考试

2. **考试题型与分值：**

- 单选题（40分）
- 判断题（40分）
- 简答题（30分）
- 论述题（30分）
- 试卷总计（共140分）

3. **考试时间：**120分钟

## 第一章 信息与信息技术

### 1.1 信息与信息处理

#### 1.1.1 信息

**信息**是能够用来消除不确定性的东西。对于信息接收者而言，只有当他获得信息之后，这种不确定性才能减少或消除。香农关于信息的定义揭示了信息的作用，也提供了一种衡量信息的方法，即使用消除不确定性的多少来衡量信息。

**信息与物质、能量**共同构成了人类社会赖以生存和发展的三大资源。

#### 1.1.2 信息处理

**信息处理**：获取原始信息，对它进行加工处理，使之成为有用信息的过程。其主要包括对信息的获取、存储、加工、表达和传输等。

**信息的获取和存储**：一般来说，在获取信息之前要先进行需求分析，即明确需要获取什么样的信息，它主要表现在时间范围、地域范围和内容范围，即“获取什么时间、什么地点、什么样的信息”，然后根据需求分析的结果，选择信息收集的途径和工具。信息的获取是信息加工处理的基础，人们通过各种途径获取信息，并对这些信息进行有效处理和评价，以便更好地利用它们解决各种问题。

**光盘、U 盘（即 USB 闪存盘）、云盘（网络存储工具）**是保存信息最常用的存储设备。通常，利用数据库系统来对大容量的信息进行管理的做法非常普遍，这样一方面可以安全有效地保存信息，另一方面也便于信息的查找与应用。

**存储容量**：是指存储器可以容纳的二进制信息量。一个二进制“位”称为一个比特（即 1 bit），每八个比特组成一个字节（即 1Byte）。例如：1 KB（千字节，Kilobyte）=1024 B，1 MB（兆字节，Megabyte）=1024 KB，1 GB（吉字节，Gigabyte）= 1024 MB，1 TB（太字节，Terabyte）=1024 GB。

表 1.1 常见信息加工形式

信息类型	信息加工形式
文本信息	文字录入及编辑，版面、正文样式、文字效果及页面设置等
表格信息	制作表格，利用数据进行数值分析和图表分析等
多媒体信息	声音、图像及视频影像的编辑合成等

**信息的加工：**包括计算、转换、重组和表示等处理过程。掌握信息处理的基本技能可以帮助人们快速、准确、及时地加工信息。表 1.1 列出了常见的一些信息加工形式。此外，诸如科学计算、文字或语音识别、数据管理、网上检索等也属于信息加工的范畴。

**信息表达方式：**有文字、声音、图形、影像、形体动作等，也可采用多媒体技术以及虚拟现实、增强现实等先进的可视化呈现技术。一般应根据不同的接收对象、环境，选择合适的信息表达方式。

为了便于信息交流，信息表达要规范化，必须遵守一定的标准，否则可能导致信息交流的失败。文件格式就是信息表达规范化的一个例子，例如 MP3、JPEG、MPEG 标准等。

### 1.1.3 人与信息处理工具

**智能手机**就是现代人们经常使用的典型的信息处理工具。手机最基本的功能是通话，人们可以通过它与远在千里之外的亲人进行语音交流。手机通话的基本工作方式是：一方面手机接收来自话筒的语音信息，经过处理转换成数据信号，用一定波段的电磁波发射出去；另一方面，手机也会反复检测一定波段的电磁波信息，一旦识别到是给本手机的信息时，就接收并转换成语音信息，并通过听筒播放。在此过程中，对信息转换的规则通常是预先存储在手机中的。

**信息传输**的目的都是为了完成从发送方到接收方的信息传递。图 1.1 所示为信息通信模型。



图 1.1 信息通信模型

- (1) **信源和信宿**分别为信息的发送方和接收方，可以是人或设备。
- (2) **发送器**的作用是将信源发出的信息按照一定的规则转换成便于传输的某种信号。
- (3) **信道**是指信息的传输通道，如广播、电视、计算机网络等。
- (4) **接收器**负责按照一定的规则把接收到的信号还原成发送的信息。

## 1.2 信息与数字化

为了更好地处理各种各样的信息，人们常常将各种信息用数据的形式表示出来，通常意义上的**数字、文字、声音、图片、动画和影像都是数据**。

**信息的数字化：**把各种形式的信息转化为二进制数形式的过程。计算机内部采用二进制，它处理和存储信息的最小单位是“位”，即一个二进制数字。要使用计算机进行信息处理，首先要把待处理的信息用二进制数来表示。

### 1.2.1 数制

**计数制：**是指用一组固定的符号和规则来表示数值的方法。

**进位计数制：**是指按进位的方法进行计数。我们日常习惯使用的十进制数，由 0,1,2,3, …, 9 这十个数码组成，运算时遵照“逢十进一”的规则，由低位向高位进位的方法进行计数。二进制数，由 0 和 1 这两个数码组成，运算时遵照“逢二进一”的规则。

### 1.2.2 信息编码

**编码：**是对原始信息符号按一定的规则所进行的变换。方便使用是信息编码的主要目的。生活中有许多信息编码的例子，例如：我们每个人都有学号，每一张公民身份证上都有一个身份证号马路上行驶的车辆都有车牌号，电影院里的座位都有座位号等等。

**编码规则：**首先，编码要尽量做到不重复；其次，编码要有规律；此外，编码的位数要尽量短，以方便记忆、保存和交流。

在计算机中，信息是以 0 和 1 的方式编码的，它的奇妙之处在于，一串由二进制数组成的数据，可以用来表示数值，也可以用来表示某个字符、标点符号或汉字，甚至还可以用来表示图像或声音。

**字符编码：**我们经常使用的数字、英文和汉字，它们的二进制编码称为字符编码，并且都有相应的国家标准。目前计算机中普遍采用的英文字符编码是 ASCII 码 (American Standard Code for Information Interchange, 美国信息交换标准码)，它已被国际标准化组织 (ISO) 定为国际标准。ASCII 码用 7 位二进制数表示一个英文字符，共可表示  $2^7 = 128$  种不同字符。其中，控制字符 34 个，阿拉伯数字 10 个，大、小写英文字母各 26 个，各种标点符号和运算符 32 个。在计算机中，实际用 8 位二进制数存储一个英文字符，最高位为“0”。

**汉字编码：**汉字在计算机内部也是以二进制方式存储的。汉字的编码包括输入码、内码和字形码。

**输入码：**为方便汉字输入而形成的汉字编码称为输入码。不同的输入方法，形成了不同的汉字输入码，同一个汉字对应的输入码不唯一。常见的输入码有以下几类：区位码、音码（如全拼、简拼、双拼等）、形码（如五笔字型）、音形码（如智能 ABC）。

**内码：**在计算机内部用来进行汉字的存储、处理的编码叫机内码或内码。内码是唯一的，也就是说每个汉字对应一个内码。由于汉字数量众多，所以每个汉字内码可用两个字节、三个字节甚至四个字节来表示。

**字形码：**为显示和打印输出汉字而形成的汉字编码称为字形码。计算机通过汉字内码在字库中找出汉字的字形码，实现其转换。目前，汉字字库主要用点阵、矢量和曲线轮廓三种方式表示汉字字形码。

**计算机中图像**的表示：矢量图和位图。矢量图使用直线和曲线等属性来描述图形，它们都是通过数学公式计算获得的。矢量图的数据量较小，其最大的优点是无论放大、缩小或旋转等都不会失真；最大的缺点是难以表现色彩层次丰富的逼真图像效果。

**位图**：基本思想是把一幅图像按照行、列分割成许许多多的点，每一个点称为一个像素，对每个像素上显示的颜色信息逐一编码后，整幅图像就可以用这些已编码像素的集合来表示。位图格式一般有 BMP、GIF、TIF、JPG 等。BMP 是微软推出的位图格式；GIF 压缩比较高，文件较小，普遍用于网页制作；TIF 是一种非失真的压缩格式，通常用于较专业的领域，如图书出版、海报印刷等，它支持多种编码方式，常见的有 LZW 编码等；JPG 是一种成熟的彩色静止图像的编码技术，其压缩比很高，失真不明显。

**声音编码**：声音是一种以声波形式传递的连续变化的模拟信号。声波由各种频率的正弦波合成，波的振幅反映了声音响度的强弱，波的频率反映了声音音调的高低。通常，对声音的编码的常用方法是按照一定的时间间隔采集声波的振幅，并将其转化为二进制数序列，即通过采样和量化两个步骤实现。

**采样**：就是采集模拟信号的样本，每秒钟的采样样本数称为采样频率。采样的过程就是按一定采样频率得一个模拟量值序列。

**量化**：就是把采集到的模拟量值序列转换为一个二进制数序列。这个二进制数序列就是对声音的编码。采样频率越高、量化级越多，声音还原的质量越好，存储声音的数据量也就越大。例如，CD 音频格式采用的采样频率为 44.1kHz，量化的值用 16 位二进制数表示，立体声双声道（每次生成两个声波数据），每秒数据量的计算公式为： $44100 \times 16 \times 2 \div 8 = 176400$  字节。

### 1.2.3 信息的压缩

**信息压缩**：很多信息，尤其是图像和音、视频等多媒体信息，经数字化转换后生成的数据量较大。这些生成的数据中包含了许多冗余信息，在不损失有用信息的前提下，按照一定的编码规则对数据进行重新组织，以去除数据冗余。信息压缩的目的，是为了更少地占用存储空间和更大地缩短传输时间。常用的压缩方法分为无损压缩和有损压缩两种。

**无损压缩**：是指对压缩后的数据进行还原，还原后的数据与原来的数据完全相同。常见的无损压缩软件有 WinZip 和 WinRAR 等，压缩后所生成的文件称为压缩包，“体积”可能只有原来的几分之一甚至更小。压缩包中的数据必须用压缩软件还原，即恢复到原来的形式后才能使用，这个过程称作解压缩。

**有损压缩：**是指在压缩过程中会损失一定的信息，压缩后的数据无法还原成原始数据。有损压缩被广泛应用于动画、声音和图像文件中。如果采用有损压缩，还原后的数据肯定没有原来的精确，质量也有所损失，但不会导致人们对原始数据表达的信息产生误解。常用的有损压缩格式有：**MP3**（声音信息压缩格式）、**JPG**（静态图像信息压缩格式）、**MPEG**（视频信息压缩格式）等等。

### 1.3 信息技术及发展

**数字城市：**是综合运用**GIS**（地理信息系统）、遥感、遥测、宽带网络、多媒体及虚拟仿真等技术，对城市的基础设施、功能机制进行信息自动采集、动态监测管理和辅助决策服务的技术系统；它具有城市地理、资源、生态环境、人口、经济、社会等复杂系统的数字化、网络化、虚拟仿真、优化决策支持和可视化表现等强大功能。“数字城市”不是一个纯技术的概念，它同时也意味着城市管理和规划体制的一次大变革。

#### 1.3.1 信息技术

**信息技术：**主要是指利用计算机技术和现代通信技术，实现获取信息、加工信息、传递信息、使用信息等的相关技术。广义地说，凡是能扩展人的信息功能的技术，都是信息技术。它主要包括以下几方面技术。

##### （1）感测技术

感测技术就是获取信息的技术，主要是对信息进行提取、识别或检测并能通过一定的计算方式显示计量结果。比如：电子温度计、电子血压仪、燃气报警器等都采用了感测技术的工作原理；而目前人们常用的识别系统，如人脸或指纹识别系统、车牌识别系统、**OCR**等，也是感测技术的一种应用。感测技术与通信技术相结合而产生的遥感遥测技术，如雷达技术、全球卫星定位系统(**GPS**)等，使人感知信息的能力得到进一步的加强。

##### （2）通信技术

通信技术就是传递信息的技术。它的主要功能是实现信息从空间一点到另一点的快速、可靠和安全转移。信息传递技术实现了信息资源的更大范围使用，如广播技术就是一种传递信息的技术。

##### （3）计算技术

计算技术就是加工信息的技术，它包括对信息的编码、运算、判断等。其目的是为了信息更有效更可靠，并且可以在对信息进行处理的基础上提炼知识、发现规律，为决策提供依据，如分布式计算技术等。

##### （4）控制技术

控制技术是在对获取的信息进行加工和逻辑判断的基础上作出决策并对操作对象实施控制的技术。控制技术能根据指令自动对外部事物的运动状态和方式实施控制，目前广泛用于自动化领域。

## 1.4 信息安全与社会责任

### 1.4.1 信息安全

**信息安全**是为了保护信息的正常使用、存储和传输，也包括使用存储、和传输这些信息的系统，以防止对信息无意或有意的损害、破坏、窃取、泄露、未授权的修改以及误用。简单地说，信息安全主要是确保信息的完整性、保密性、可用性和可控性。

(1) **信息的完整性**是指信息在存储或传输过程中保持不被修改、不被破坏和不丢失；

(2) **信息的保密性**是指信息不泄露给非授权的个人或实体并供其利用；

(3) **信息的可用性**是指信息可被合法用户访问并按要求使用；

(4) **信息的可控性**是指对信息的内容及传播具有控制能力，任何信息在一定范围内都应该是可控的。

**影响信息安全的因素**主要有：(1) **威胁和攻击**。威胁与攻击的发生可能是自然因素，也可能是人为因素造成的。自然因素是指一些不可预测的因素，例如火灾、地震等自然灾害。另外还有一种说法是“不可抗力”因素。除了自然灾害，还包括战争。计算机病毒和黑客攻击都是一种人为的潜在危害，可能引发信息泄露、信息敲诈甚至是蓄意破坏造成系统瘫痪。(2) **知识产权的损害**。例如非法使用或复制软件产品或其他信息资源。(3) **服务质量差**。因为信息的安全使用依赖于很多支持系统的成功运转，这些支持系统包括电力供应、数据网络、设备供应商以及相关的服务人员，无论在信息的存储与传输中的哪个环节出了问题，都会影响信息的正常使用，其造成的原因可能是人为过失，也可能是管理机制问题。(4) **误操作**。一般是信息使用者因失误执行了删除、改写等具有破坏性的操作命令，从而导致了信息受损。

**信息安全措施**主要有以下几方面：(1) 信息安全的法律法规和道德教育，规范人们安全使用信息和相关设备的行为。(2) 制定信息安全规划与制度，使信息安全的政策、标准能够有效地实施。(3) 采用合理的技术手段，例如系统升级设置防火墙、信息加密、数据备份等，以防止信息的泄露和破坏。

### 1.4.2 知识产权

**知识产权**：是一种无形财产权，是从事智力创造性活动取得成果后依法享有的权利，通常分为两部分，即“**工业产权**”和“**版权**”（著作权）。



知识产权包括对下列各项知识财产的权利：文学、艺术和科学作品；表演艺术家的表演及唱片和广播节目；人类一切活动领域的发明；科学发现；工业品外观设计；商标、服务标记以及商业名称和标志；在工业、科学、文学或艺术领域内由于智力活动而产生的一切其他权利。在信息技术领域中，计算机及所有周边设备、通信设备等硬件的设计制造和计算机软件的原创新成果都受到法律的保护。作为使用者必须具备保护知识产权的意识，自觉购买和使用正版软件，抵制盗版或未经授权的软件。未经软件版权人授权，不得对其软件进行复制、修改和传播，更不能进行商业销售活动。

**自媒体(We Media)：**是一个普通市民通过数字科技与全球知识体系相联，提供并分享他们真实看法、自身新闻的途径。直白地说，自媒体是指传播者通过互联网平台，将自主采集或把关过滤的内容传递给他人的个性化传播渠道。

## 第二章 信息处理工具——计算机

计算机系统由**硬件系统**和**软件系统**组成。

### 2.1 计算机硬件系统

#### 2.1.1 计算机的硬件组成及工作原理

ENIAC 是 1946 年在美国诞生的世界上第一台电子计算机。有学者针对 ENIAC 提出了计算机设计制造的建议，其中主要的建议是：计算机应采用二进制运算；程序可以像数据那样存放在存储器中，由计算机自动控制执行。以后设计制造的计算机都采纳了这一建议，我们把遵循这种结构的计算机称为**程序存储控制结构计算机**。

具有程序存储控制结构的计算机由**运算器、控制器、存储器、输入设备和输出设备**等五个逻辑部件组成；在计算机内部采用**二进制编码**的形式表示数据和指令。

**运算器**负责对数据进行加工处理完成各种算术和逻辑运算。**控制器**的作用是协调并控制计算机的各部件按指令序列执行指定的操作。运算器和控制器是计算机的核心，合称为**中央处理器(CPU)**。存储器是具有记忆功能的器件用于存放数据或程序。存储器通常分为**内存储器(内存)**和**外存储器(外存)**两类。输入设备负责程序或数据输入输出设备则用于输出运算结果；输入、输出设备又称为**I/O 设备或外部设备**。I/O 设备需要通过 I/O 接口才能将外部程序或数据输入计算机以及将计算机内部的数据输出。

为了实现 CPU 存储器 I/O 接口之间的连接，计算机的硬件系统一般采用**总线结构**。计算机依靠总线把硬件系统的各个部件连接起来，各类信息通过总线进行传导。依据传递信息内容的不同，总线可分为**数据总线、地址总线和控制总线**三种。

**程序执行的过程：**程序存储控制结构的计算机以存储程序原理为基础，操作者要预先把指挥计算机如何进行操作的程序和原始数据存放到计算机内存中。计算机运行程序时，CPU从给定的内存单元读取一条指令代码，分析该指令的功能，然后执行相应的操作。当一条指令执行完毕以后 CPU 就会读取并执行下一条指令。计算机就是这样不停地执行指令，完成预定的工作。

## 2.1.2 微型计算机基本配置

### (1) 中央处理器（CPU）

CPU 是整台计算机的核心部件，它包括**运算器**和**控制器**两大部分，采用**超大规模集成电路**制成。它直接安插在主板的 CPU 插槽上，由于它高速运行会产生很大的热量，所以通常会在 CPU 芯片上方安装散热片和风扇。CPU 性能的好坏对计算机的总体性能起着非常重要的作用。衡量 CPU 性能的指标主要有：主频、字长和高速缓存(cache)。

**主频**即 CPU 的时钟频率，它在很大程度上决定了计算机的运算速度。例如，AMD 公司生产的 Opteron256 其主频为 3GHz。**字长**指的是 CPU 一次可以处理的二进制位数，字长越长，其运算精度越高。**高速缓存**用来存放 CPU 频繁使用访问的程序和数据。它的存取速度远远高于内存，以接近 CPU 的速度与之交换信息。从概念上讲，它属于存储器，但由于它与 CPU 封装在一块芯片中，所以它的容量也作为反映 CPU 性能的重要指标。

### (2) 存储器

存储器通常分为**内存储器**和**外存储器**两类。

**内存储器**又称主存储器，简称“主存”或“内存”，包括只读存储器(ROM)和随机存取存储器(RAM)两类。

① **ROM 只读存储器**是一种只能读出而不能写入信息的存储器。当计算机断电或关机时，其内部存放的信息不会丢失。ROM 中通常存放计算机系统引导程序、开机检测程序和系统初始化程序等对计算机运行十分重要的信息。

② **RAM 随机存取存储器**是一种既能写入又能读出信息的存储器。当机器断电或关机时 RAM 内部存放的信息就会丢失。计算机中的内存一般指的就是 RAM。

CPU 运行所用到的指令和数据都存放在内存中，CPU 频繁地读写内存中的信息，对内存的速度要求也很高。内存的速度用存取时间衡量，存取时间指启动一次存储器操作到完成该操作所需的时间，以纳秒(ns)为单位。在实际芯片产品中，内存的速度往往用内存主频来表示**内存主频**代表该内存所能达到的最高工作频率，它是以 MHz(兆赫)为单位来计量的。内存主频在一定程度上代表着内存所能达到的速度，以 DDR3-2400MHz 内存为例：2400MHz

就是该内存的主频。那么 DDR3-2400MHz 就比 DDR3-1600MHz 的内存访问速度快。另外，**存储容量**也是衡量内存性能的一个重要指标。多块内存芯片被连接在一块电路板上，这就是内存条。内存条直接安插在主板的内存插槽上，并且有多种规格，常见的有 4GB、8GB、16GB 等。

**外存储器**又称辅助存储器，简称“外存”，一般用来存放需要永久保存或暂时不用的程序和数据。外存储器的存储容量大，但它的存取速度相比内存慢很多，断电后内部信息依然保存，如硬盘、光盘、U 盘等。

① **硬盘**是最常用的外存储器，它通过**磁头与涂有磁存储介质的碟片**之间的相对运动来读写数据。衡量**硬盘的性能指标**包括存储容量、缓存容量和转速等。

② **光盘**是利用激光读写数据的存储介质，一般光盘分为只读和可读写两种。只读光盘中的内容是无法改变的。光盘的使用寿命比硬盘长，但是读取速度比硬盘慢。光盘的读和写需要使用光盘驱动器。

③ **U 盘**是常用的移动存储设备，不需要物理驱动器，直接通过 USB 接口与计算机连接，是一种即插即用设备。U 盘体积小重量轻适合随身携带用于移动办公及文件交换。

### (3) 输入、输出设备

**常用的输入设备**有键盘、鼠标、扫描仪、摄像头等；**常用的输出设备**有显示器、打印机、音箱、绘图仪等。

输入、输出设备种类繁多，即使同一类型的设备往往也有不同的规格型号，规格型号不同，性能也不同。**键盘**有 83 键、101 键等规格，可向计算机输入数字、字母和符号。101 键的键盘右边有数字小键盘，可专门用于输入数字，特别适合经常处理大量数字输入与运算的人员使用。屏幕尺寸大小、分辨率等都是反映显示器性能的指标，显示器的清晰度主要是与它的分辨率有关。一台**分辨率为 1600x1200 像素**的显示器的水平和垂直方向分别有 1600 和 1200 个像素。

相对于 CPU 输入、输出设备都是慢速设备而且输入、输出设备中的数据在表示形式上与计算机内部的形式不一致，这就需要用到一种转换设备。输入、输出接口就是用于实现输入输出设备和计算机内部之间数据交换的电路，也称为**输入输出适配器**。输入输出接口电路大部分都安装在计算机的主板上，通过总线与 CPU 或内存储器连接。

**显卡**是连接显示器的接口，**声卡**是连接音箱的接口，**网卡**是连接网络设备的接口，**USB 接口**可以连接 U 盘、键盘、鼠标等大量外部设备，**HDMI 接口**是音视频传输的接口。这些

接口插在主板的插槽中或者直接集成在主板上。此外，在主板上还有空余的插槽可以用来扩充计算机和外设的连接接口。

**通用串行总线(universal serial bus,简称 USB)**，是连接计算机系统与外部设备的一种串口总线。USB 接口的优点是可以方便连接不同类型的 USB 设备，如 USB 键盘、USB 鼠标、USB 硬盘等。USB 接口的传输速度也要比传统的接口快很多，USB1.0 版本最早于 1996 年推出，传输速度是 **1875KB/s**，USB2.0 版本的传输速度为 **60MB/s**，而 USB3.0 版本的传输速度更是达到了 **640MB/s**。

## 2.2 计算机软件系统

**计算机软件**是为解决计算机管理和各种应用问题所编写的各种程序以及与此相关的文档资料。没有安装任何软件的计算机称为“裸机”。计算机系统只有在“裸机”的基础上，配备一定的软件后，才能发挥作用。

### 2.2.1 计算机指令和程序

#### (1) 计算机指令

让计算机执行某个操作的命令称为指令。二进制代码是计算机内部唯一能识别的符号。所以计算机能识别的指令也只能采用 0 和 1 编码的二进制代码来表示，这种二进制代码的集合称为机器语言。**指令由操作码和操作数组成。**

**机器语言和汇编语言**都是针对特定的计算机系统，不同类型的计算机具有各自的指令集。为了能够满足常见的操作，计算机的指令集中通常包含**算术和逻辑运算、传送、控制、输入和输出**等各种指令。

#### (2) 计算机程序

根据解决某一问题或完成某一任务的具体步骤，使用者将一条条指令进行有序的排列。这一指令序列称为程序。

程序员往往采用一种程序设计语言完成程序的设计，常用的程序设计语言分为三类：**机器语言、汇编语言和高级语言**。

采用机器语言编写的程序所占内存空间较小，执行速度快，缺点是指令难记，编写的程序难以阅读且通用性差。汇编语言与机器语言相比，易读易记。

使用汇编语言编写的程序，必须翻译成机器语言后才能被计算机执行，其通用性仍然很差。高级语言接近人类的自然语言，易学易记，可读性强，修改也很方便。

采用高级语言编写程序，编写的程序通用性强；使用高级语言编写的程序需要经过编译或解释才能被计算机识别，执行速度慢，占用内存空间大。

## 2.2.2 计算机软件的分类

计算机软件通常分为**系统软件**和**应用软件**。

### (1) 系统软件

**系统软件**是计算机设计者或生产厂商提供的使用和管理计算机的软件,它通常包括操作系统、语言处理程序、数据库管理系统等。

**操作系统**(operating system, 简称 OS)是直接运行在“裸机”上的最基本的系统软件。它对计算机全部软硬件资源进行控制和管理。主要包括:**处理器管理**负责 CPU 的分配和控制; **存储器管理**负责内存的分配与回收; **IO 设备管理**负责 I/O 设备的分配与操作; **文件管理**负责文件的读写增删、共享和安全保护; 另外, 操作系统还作为用户和计算机硬件系统之间的接口。

其他软件必须在操作系统的支持下才能运行。目前常用的操作系统有 **Windows、macOS、Unix、Linux、Android 和 iOS** 等。为适应网络环境和管理功能的需要, 有些操作系统还提供了适用于网络管理的服务器版操作系统。

**Unix** 是 20 世纪 60 年代末开发的一款多用户多任务的分时操作系统。经过几十年的发展, Unix 已经成为目前国际上使用广泛、极具影响力的操作系统之一。**很多其他的操作系统都是它的变体,比如 HP-UX、Solaris、AIX 以及著名的 Linux 等。**

**语言处理程序**(language processor)即各种程序设计语言的翻译程序, 如汇编程序( assembler)编译程序( compiler)解释程序( interpreter)等。它们的作用都是把源程序翻译成二进制代码表示的机器语言, 使计算机能够识别和执行。常见的程序设计语言有 **Pascal、C、C++、Visual Basic、Java 和易语言**等。目前比较流行的是可视化的面向对象的语言如 Visual Basic、Visual C++、Delphi 和易语言等。另外, 还有**编程工具**(programming tool), 它们是编写、调试和运行程序的辅助工具, 如文本编辑程序等。

**数据库管理系统**(database management system)是用于建立、使用和维护数据库的软件。数据库是长期储存在计算机中的有组织的、可共享的数据集合。按照数据模型分, 数据库有**关系数据库**和**面向对象数据库**等几类。常见的数据库管理系统有 SQL Server、Access、Oracle 和 DB2 等。

### (2) 应用软件

**应用软件**是为完成某种应用或解决某类问题而编制的专用程序。由于应用面广, 应用软件很难分类, 这里列举几种与我们关系比较密切的应用软件。

#### ① 办公软件

常用的办公软件包括文字处理软件、报表制作软件、演示文稿制作软件等。常见的有 WPS 套件(包括 WPS 文字、WPS 表格、WPS 演示等)和 Office 套件(包括 Word、Excel、PowerPoint 等)。

## ② 多媒体制作和播放应用软件

多媒体制作和播放应用软件包括图形/图像处理软件动画制作软件、音/视频编辑软件以及音/视频播放软件等。这类软件比较多,如图像处理软件 Photoshop, 视频编辑软件 Premiere, 动画制作软件 Flash, 视频播放软件 RealPlayer 和 Media Player 等。多媒体应用软件的工具用来帮助多媒体应用开发人员提高开发工作效率。它们大体上都是一些应用程序生成器, 能将各种媒体素材按照超链接的方式组织起来, 形成多媒体应用系统。

## ③ 管理信息系统

管理信息系统(management information system, MIS)是一个由人、计算机等组成的, 能进行管理信息的收集、传递、存储加工、维护和使用的系统。现在几乎每个行业、每种业务都有自己的 MIS 如人事管理信息系统、财务管理信息系统、学生管理信息系统、教学管理信息系统、商场管理信息系统、物流管理信息系统、企业资源管理信息系统等。

## ④ 生产过程控制系统

大型的现代化工厂都利用计算机(或计算机网络)系统实时控制产品的生产。这类系统的计算机安装了许多特殊的 I/O 设备, 它们能够实时采集生产过程中的数据, 了解生产情况, 并及时发出控制指令进行操纵、指挥管理和决策。如炼钢控制系统、车钢控制系统、乙烯生产控制系统、卫星发射系统等。

对于有一些软件产品, 其属于系统软件还是应用软件的界线会比较模糊。一般来说, Office 套件属于应用软件, 然而其中的 Access 则是数据库管理系统。同样, Windows 作为操作系统属于系统软件, 但在安装时也会集成一些如画图、计算器等应用软件。还有一类软件, 是介于系统软件和应用软件之间的, 如软件开发环境等。

## 第三章 计算机网络与因特网

### 3.1 计算机网络简介

#### 3.1.1 计算机网络的基本概念

**计算机网络:** 利用各种通信线路, 把地理上分散的、彼此独立的多台计算机连接起来, 遵守某种约定进行通信, 实现资源的共享及相互协同工作的系统。

**计算机网络的三要素:** (1) 计算机设备; (2) 通信线路及连接设备; (3) 网络协议。

**计算机网络的功能：**（1）数据通信；（2）资源共享；（3）分布式处理。

**计算机网络的分类：**按照网络的规模和覆盖范围：（1）局域网(LAN, Local Area Network)；（2）广域网(WAN, Wide Area Network)。按照网络拓扑结构：（1）总线型；（2）星型；（3）环型；（4）网状型；（5）树状型。

**网络传输速率：**每秒网络传输线路能够传送的二进制位个数，单位为 bps(bit per second) 比特率。K、M、G 和 T 分别表示千、兆、吉和太(1Kbps=2<sup>10</sup>bps=1024bps, 1Mbps=1024Kbps, 1Gbps=1024Mbps, 1Tbps=1024Gbps)。

### 3.1.2 网络连接组件

**网络连接组件：**网络适配器(网卡)、传输线(网线)、集线器、交换机、路由器等。

#### (1) 集线器

集线器采用共享带宽的工作方式，它本身并不“认识”目的计算机地址，所以只能采用广播形式发送数据，这就很容易造成网络堵塞。一般来说，接收数据的只有一台计算机，而绝大部分数据流量都是无效的，这就造成了整个网络数据传输效率相当低。另外，由于发送的数据在同一网络上，每台计算机都能侦听到，这显然也是不够安全的。所以，目前很少使用独立的集线器产品。

#### (2) 交换机

交换机采用独享带宽的工作方式，其内部通常存有一张各端口与所连接计算机的地址的对照表，当交换机的某个端口接收到从某台计算机发来的数据后，会根据数据提供的目的地址信息自动查找表中对应的端口，然后将数据从该端口转发出去，使数据迅速传送到目的计算机上。交换机具备可以在同一时刻进行多个端口数据传输的能力。通过交换机的过滤和转发，可以减少误送或错送数据情况的出现，避免共享冲突，不易产生网络堵塞。

#### (3) 路由器

路由器连接多个网络的网络设备，它能对不同网络之间的数据信息进行“翻译”，以使它们能够相互“读懂”对方的数据，从而构成一个更大的网络。它与上述集线器和交换机不同，它不是应用于同一网络的设备，而是应用于不同网络之间的设备，属于网际设备。

路由器最主要的作用就是**路由选择**，也就是接收从网络上计算机发来的数据，然后根据这个数据中包含的源地址信息和目的地址信息，对照路由器自己内部的路由表(该表记录着目的计算机地址与端口的对应关系)，然后把该数据直接转发到目的计算机中去。在该数据的传输过程中，路由器会分析判断发出请求的源地址信息及接收请求的目的地址信息，从而找出一条最佳、最快捷的通信路径。

**网络传输介质：**①**有线传输介质：**双绞线、光纤（传输速率更高、抗干扰能力更强，适合远距离传输）、同轴电缆（可传输数字信号和模拟信号，比双绞线传输速率更高）。②**无线传输介质：**电磁波。

### 3.1.3 网络协议

**网络协议：**计算机在网络中实现通信时必须遵守的规则和约定。**常见网络协议：**万维网的超文本传输协议(HTTP 协议)，电子邮件发送和接收协议(SMTP 和 POP3 协议)，而传输控制协议和网际协议(TCP/IP 协议)是因特网中最核心的一套协议。

**计算机网络的工作方式：**（1）客户机/服务器模式（Client Server, C/S 模式）；（2）浏览器/服务器（Browser/Server, B/S 模式）。

## 3.2 因特网基础

### 3.2.1 因特网

**因特网：**是一个全球性的计算机数据通信网络，又称国际互联网（英文名为 Internet），是连接全世界成千上万台计算机的网络的集合，是当今最大的国际性计算机互联网络。因特网也被称为是“网络的网络”。

整个因特网是由非营利性的民间组织**因特网协会(Internet Society ,ISOC)**来管理的。这个协会下设了很多机构，主要职责是研究因特网存在的问题及其解决方法，开发因特网的标准和协议，推动互联网的建设和发展，受理因特网地址和域名的注册等。

**因特网的发展：**世界上第一个计算机网络(阿帕网 ARPANET)：美国国防部高级研究计划署开发的世界上第一个运营的网络。20 世纪 70 年代末，为了满足网络用户开发各种网络应用的需要，首次使用 TCP/IP 协议成功地进行了网络连接的试验。20 世纪 80 年代，美国建成了为所有科研单位服务的**网络国家科学基金会网络（简称 NSF 网络）**。这个网络以五个**超级计算机中心**为核心，通过高速通信线路连成一个**主干网**，并使用 TCP/IP 协议。到 1991 年，连入 NSF 网的网络数达到 3000 个。1988 年，NSF 网络正式改称为因特网。1993 年，NSF 取消了对商业应用的限制。从此，因特网开始突飞猛进地发展起来。1995 年，NSF 退出因特网的管理。随着网络应用的研究成果不断涌现，1971 年，诞生了**基于分布式网络发送消息的电子邮件程序**。1978 年，由**第一个计算机公告牌系统 CBBS(computerized bulletin board system )**演变成现在的 BBS 系统。1990 年，**万维网(world wide web ,缩写为 WWW, 简称 Web)**计划被提了出来，同时相应的**超文本标记语言(hypertext markup language ,HTML)、超文本传输协议(hypertext transfer protocol, HTTP)**和**统一资源定位器(uniform resource locator ,URL)**等基础技术陆续被开发出来。直到 1993 年，历史上第一个



支持多媒体的网页浏览器 Mosaic 诞生，这才证明了万维网真正的实用价值。于是，万维网上的通信量爆炸性地增长，极大地推动了因特网的发展。20 世纪 90 年代，随着万维网上的资源呈指数级增长，人们为了有效获取信息，提出了各种信息搜索方式。**搜索引擎**为用户获取信息提供帮助，搜索引擎的开发不仅为人们在因特网上信息搜索带来了极大的便利，也为其他信息产业的发展提供了巨大商机。

### 3.2.2 TCP/IP 协议

因特网是一个分组交换网络，这意味着从一台计算机向另一台计算机发送数据时，这些数据必须被划分成较小的分组。在所有的分组全部到达目的计算机后，就将这些分组重新组合成原来的完整信息数据。负责把数据分割成分组，对这些分组进行因特网选路并在接收端重组分组的协议有两个：（1）**网际协议 IP(Internet protocol)**：负责分组的选路；（2）**传输控制协议 TCP(transmission control protocol)**：负责把原始数据划分成分组，并在接收数据的计算机上重组为原始数据。IP 协议和 TCP 协议是因特网传输协议中最核心的两个协议，因特网中的计算机都要安装 TCP/IP 协议。

### 3.2.3 因特网的地址和域名

因特网上的**每台计算机**都需向 IP 地址分配管理机构提出申请，**请求分配唯一的 IP 地址**。所以一个 IP 地址实际上反映的是一台**计算机（或路由器）与因特网的一个连接**。

当前的 **IP 地址**采用国际统一标准的 **IPv4 协议**，是一个 **32 位的二进制数**。由于人们阅读和使用二进制数十分不便，因而因特网定义了一种 IP 地址标准写法。它规定按 8 位（即一个字节）为一组，把 IP 地址的 32 位分成四组，组与组之间用圆点作分隔，每组的值用十进制数表示。例如，**IP 地址 10000010 00000001 00010000 00000001 可写成 130.1.16.1**。

**IP 地址包含两大部分：**（1）**网络地址**；（2）**主机地址**。

**域名：**为了方便人们记忆与使用网络地址，可使用含有一些具体特征的名字来标识网络主机，每个域名有若干个子域组成，子域与子域之间用圆点相隔，最右边的是最高层域名，由右向左层次逐级降低，最左边的子域是主机名，一般格式为：**计算机名.网络名.机构名.域名**。域名有时可以根据组织机构来划分，也可以根据国家 and 地区代码来划分，常见的域名如表 3.1 所示。

表 3.1 常见域名表

机构域		地理域	
域名	类型	域名	类型
com	商业机构	cn	中国
edu	教育机构或设施	uk	英国
gov	非军事性的政府机构	bj	北京
int	国际性机构	sh	上海
mil	军事机构或设施	js	江苏
net	网络组织或机构	hk	中国香港
org	非盈利性组织机构	tw	中国台湾

例如：上海师范大学天华学院的域名为：sthu.edu.cn

如何从域名找到相应的 IP 地址呢？域名与 IP 地址的转换是由**域名系统（domain name system, DNS）**来实现的。

### 3.2.4 因特网的接入方式

**因特网服务提供商(Internet service provider ,ISP)**：向用户提供因特网接入、信息发布服务等业务的通信运营商，例如：中国移动、中国联通、中国电信。ISP 具备了与因特网主干网络相连的高速通信线路，拥有 IP 地址、域名等网络资源。ISP 的主要工作就是配置用户与因特网相连所需的设备，并建立通信连接，为用户提供信息服务。

**光纤接入**：目前使用最广泛的一种因特网接入方式。现在比较常见的有，光纤线路直接入户，然后通过光纤调制解调器与用户终端通信；或者是光纤先接入大楼，然后再通过一定的网络设备分别转接到每个用户终端。光纤网络的传输距离远，传输速率通常也很高，目前最高可达每秒千兆。光纤接入与有线通接入的上、下行传输速率都可以根据实际需要来设置。

**无线接入**：用户采用无线通信手段连接因特网的接入技术。无线接入也可以分为**移动接入**和**固定接入**两大类。各种无线接入方式中，目前比较普遍采用的有**WiFi(wireless fidelity)**接入和**蜂窝移动通信接入**。固定接入方式主要有：（1）微波接入；（2）卫星接入。

**移动接入方式**：智能手机和一部分便携式计算机可以使用移动接入方式，大多采用**无线应用协议(wireless application protocol, WAP)**。

**WiFi** 是一种允许电子设备连接到一个无线局域网(WLAN)的技术，它可以实现将有线网络信号转换成无线网络信号。现在几乎所有的智能手机、平板电脑和笔记本电脑都支持WiFi上网。WiFi是当今使用最广的一种无线网络传输技术。

**蜂窝移动通信**是采用蜂窝无线组网方式，在终端和网络设备之间通过无线通道彼此连接起来，从而实现终端的移动通信。蜂窝移动通信技术从20世纪70年代至今，经历了好几代的发展和改良，例如第四代蜂窝移动通信技术（简称4G）是集第三代蜂窝移动通信技术（简称3G）与WLAN于一体，能够快速传输数据和高质量音频、视频及图像等的移动通信技术，其在下载速率及覆盖范围上都有着明显的优势。目前5G技术已得到大规模商业应用。

**第五代移动通信技术**（5th Generation Mobile Communication Technology，简称5G）是具有高速率、低时延和大连接数特点的新一代宽带移动通信技术，是实现人机物互联的网络基础设施。

为了彻底解决因特网的地址危机，IETF(Internet工程任务组)早在20世纪90年代中期就提出了下一代的**IPv6因特网协议**。IPv6每个地址占**128位**，地址空间大于 $3.4 \times 10^{38}$ 。

### 3.3 因特网的应用

#### 3.3.1 电子邮件

电子邮件(electronic mail, E-mail)是因特网上一种重要的信息服务方式。收发电子邮件需要由**电子邮件服务器、电子邮件协议和电子邮箱**配合来完成。电子邮件地址的格式为：**用户名@域名**。其中，“用户名”是用户自己定义的字符组合，“域名”是提供电子邮件服务的**服务商名称**，如小陈的电子邮件地址为 xiaochen\_edu@sohu.com,那么搜狐邮件服务器就相当于邮局，xiaochen\_edu@sohu.com 就相当于小陈的信箱。

传统信件的收发是通过邮递员的分检和传送实现的，而电子邮件则是通过两个应用协议来实现的。一个是**简单邮件传输协议(simple mail transfer protocol,SMTP)**,它的功能是把电子邮件先送至发送人的电子邮件服务器中，然后该服务器再将邮件发送至接收人的电子邮件服务器中。另一个是**电子邮局协议第3版(post office protocol 3,POP3)**,它的功能是帮助接收人利用邮件管理软件（如QQ邮箱、163邮箱等），将邮件从他的电子邮件服务器接收到他的计算机中。

#### 3.3.2 万维网

万维网是由许多信息资源相互链接而成的一个世界范围的信息网，其信息资源可以供用户进行访问。万维网服务器中的各种资源是以网页的形式整合后，采用超文本传输协议传送，最后通过浏览器的解释重新展现的。

**超文本传输协议(hypertext transfer protocol, HTTP):**是用于从万维网服务器传输超文本到本地浏览器的传送协议，是 TCP/IP 协议族中的一种数据通信协议。该协议的主要功能是传送浏览器的网页请求和服务器返回的网页。例如，当我们输入网站的域名或相应的 IP 地址访问一个网站时，浏览器通过 HTTP 协议与万维网服务器建立连接，并将地址栏中的内容作为访问请求传送给万维网服务器。万维网服务器接收到该请求后，便将我们所需的数据传送到我们的计算机中，然后断开与计算机的连接。数据经浏览器解释后以网页的形式展现，我们就可以访问所需资源了。

**统一资源定位器:**万维网上的每一个网页、每一种资源都有具体的地址，这个地址是用统一资源定位器(uniform resource locator, URL)表示的。不管这个资源存放在世界的哪个角落，只要知道它的 URL，马上就可以找到它。

**URL 的常见形式:** 协议名://域名/路径/文件说明，其中“协议名”是指访问该资源所采用的协议；“域名”是指该资源所在的主机，也可以换作对应的 IP 地址。例如：上海师范大学天华学院人工智能学院的 URL 为：http://ai.sthu.edu.cn，其中，“协议名”是 http（超文本传输协议），“域名”是 ai.sthu.edu.cn。