

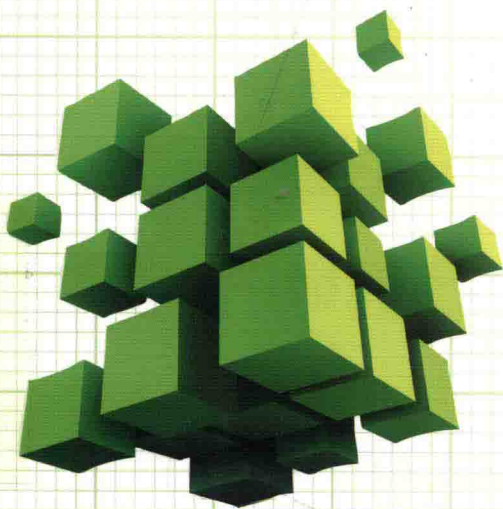
应用型高等院校改革创新示范教材

APPLICATION-ORIENTED INNOVATIVE TEXTS

# Python

## 程序设计与应用

主 编 张广渊



中国水利水电出版社  
www.waterpub.com.cn

应用型高等院校改革创新示范教材

# Python 程序设计与应用

主 编 张广渊



中国水利水电出版社  
[www.waterpub.com.cn](http://www.waterpub.com.cn)

·北京·

## 内 容 提 要

本书是一本基础性强、可读性好、适合入门的 Python 语言教学教材。读者通过本书的学习，能够快速掌握 Python 语言的入门知识，并通过项目实践学习，了解项目的基本开发流程和常用第三方库的使用。

本书内容分为两大部分：第一部分是 Python 语言编程基础知识，共 10 章，第 1 章是绪论，第 2 章是 Python 编程环境搭建，第 3 章是 Python 程序设计入门，第 4 章是程序控制，第 5 章是列表、元组、字典、集合，第 6 章是函数，第 7 章是模块，第 8 章是面向对象程序设计，第 9 章是文件，第 10 章是异常处理；第二部分是项目实践，共有 4 个小项目，分别是根据函数绘制曲线，分析微信好友数据，网页数据下载与处理，Django+MySQL Web 开发。

本书既面向计算机软件编程零基础和刚入门的读者，也适用于具备一定基础、开始学习第三方库的使用和初级项目实施的读者，还可作为参加全国计算机等级考试 Python 语言考试的参考用书。

## 图书在版编目 (C I P) 数据

Python 程序设计与应用 / 张广渊主编. — 北京 :  
中国水利水电出版社, 2019.3  
应用型高等院校改革创新示范教材  
ISBN 978-7-5170-7480-9

I. ①P… II. ①张… III. ①软件工具—程序设计—  
高等学校—教材 IV. ①TP311.561

中国版本图书馆 CIP 数据核字 (2019) 第 031731 号

策划编辑：石永峰 责任编辑：张玉玲 加工编辑：赵佳琦 封面设计：李 佳

书 名	应用型高等院校改革创新示范教材 Python 程序设计与应用 Python CHENGXU SHEJI YU YINGYONG
作 者	主 编 张广渊
出版发行	中国水利水电出版社 (北京市海淀区玉渊潭南路 1 号 D 座 100038) 网址: www.waterpub.com.cn E-mail: mchannel@263.net (万水) sales@waterpub.com.cn
经 售	电话: (010) 68367658 (营销中心)、82562819 (万水) 全国各地新华书店和相关出版物销售网点
排 版	北京万水电子信息有限公司
印 刷	三河市航远印刷有限公司
规 格	184mm×260mm 16 开本 13.75 印张 335 千字
版 次	2019 年 3 月第 1 版 2019 年 3 月第 1 次印刷
印 数	0001—3000 册
定 价	38.00 元

凡购买我社图书，如有缺页、倒页、脱页的，本社营销中心负责调换  
版权所有·侵权必究

# 前 言

Python 语言已成为目前最受欢迎的计算机语言之一，其生态链应用也越来越广泛，从绘制图形到词频统计，从图像处理到 Web 开发，从科学计算到嵌入式开发，越来越多的场景可以使用 Python 语言来解决。尤其是人工智能时代的到来，调用方便、科学计算功能强大的 Python 依旧在 AI 领域拥有最强大的竞争力，可以说，Python 语言就是人工智能领域的 BASIC 语言。

本书的编写旨在推动 Python 语言教学作为应用型本科大学相关专业的公共基础课程来进行教授。本书内容分为两大部分：第一部分是 Python 语言编程基础知识，通过浅显易懂的语言结合丰富的配图，使得编程初学者能够快速入门掌握 Python 语言；第二部分是项目实践，通过设计项目实践环节，初步展现解决问题的过程和方法，使初学者能够使用编程思维解决问题，并通过项目实施了解简单项目的开发流程，掌握一些常用第三方库的使用。因此，本书既面向计算机软件编程零基础和刚入门的读者，也适用于具备一定基础、开始学习第三方库的使用和初级项目实施的读者。

本书依据全国计算机等级考试 Python 语言考试大纲在内容上做了对应的编排，覆盖了大纲所要求掌握的内容范围，在基础知识部分各章的后面都附有大量的习题可供练习，可作为参加全国计算机等级考试 Python 语言考试的参考用书。

本书第一部分基础知识共分 10 章。第 1 章阐述了计算机的基本概念、软硬件的发展历史和现状，并对 Python 语言的基本概念和特点作了介绍；第 2 章介绍了 Python 编程环境的搭建，主要包括 Anaconda、IDLE 和 PyCharm 三种常用环境的搭建和使用；第 3 章介绍了 Python 语言基本内容，主要包括标识符、基本运算、赋值、字符串操作和内置函数的介绍；第 4 章主要介绍选择和循环等程序控制结构；第 5 章介绍了 Python 语言常用的四种结构，包括列表、元组、字典和集合；第 6 章介绍了函数的使用、参数的传递以及全局变量和局部变量的使用；第 7 章介绍了模块、包和第三方库的引入和使用；第 8 章对面向对象程序设计方法进行了描述；第 9 章介绍了文件的基本操作，并对 CSV 文件和 Excel 文件在 Python 中的读写调用进行了详细叙述；第 10 章专门围绕异常处理进行了详细描述。

第二部分项目实践分 4 个项目进行练习。项目一主要是围绕如何绘制图形以及如何项目进行发布的任务来进行 turtle、matplotlib 和项目打包发布介绍；项目二设计实现对微信好友数据进行分析 and 绘制图表的任务；项目三设计了一个网页爬虫任务，通过项目实施实现对 urllib、etree、xpath、time、Pandas、jieba、wordCloud 等第三方库的介绍和使用；项目四基于 Django 框架和 MySQL 数据库，实现了一个简单的基于动态数据展示的 Web 应用。

本书第 1、10 章由张广渊编写，第 2、9 章、项目 1 和项目 2 由倪燃编写，第 3、4 章由吴昌平编写，第 5、8 章和项目 3 由李风云编写，第 6、7 章和项目 4 由朱振方编写，全书由张广渊统稿。

本书内容基于 Python 3 编写，所有源程序代码均在 Python 3 编程环境下运行通过。Python 计算生态和资源可从 <https://github.com/vinta/awesome-Python> 获得。

由于编者水平有限，在本书编写过程中难免出现错误和疏漏，恳请广大读者批评指正。

编 者

2018 年 12 月

# 目 录

前言

## 第一部分 基础知识

第 1 章 绪论	1	练习二	40
1.1 计算机发展历史	1	第 3 章 Python 程序设计入门	41
1.1.1 计算机的历史	1	3.1 标识符、关键字和数据类型	41
1.1.2 个人计算机的历史	2	3.1.1 标识符、关键字	41
1.2 软硬件基本知识	3	3.1.2 标准数据类型	43
1.2.1 冯·诺依曼结构	3	3.2 常量与变量	43
1.2.2 计算机软件发展	7	3.2.1 常量	43
1.3 常用编程语言介绍	9	3.2.2 变量	45
1.3.1 机器语言	9	3.3 运算符与表达式	46
1.3.2 汇编语言	10	3.3.1 算术运算符	46
1.3.3 高级程序设计语言	10	3.3.2 关系运算符	47
1.4 Python 语言简介	11	3.3.3 逻辑运算符	47
1.4.1 Python 语言的历史	11	3.3.4 位运算符	48
1.4.2 Python 语言的特点	12	3.3.5 赋值运算符	48
1.4.3 Python 语言应用	12	3.3.6 成员运算符	49
练习一	13	3.3.7 标识(身份)运算符	49
第 2 章 Python 编程环境搭建	15	3.3.8 运算符优先级	49
2.1 Anaconda 安装与 Python 编程界面	15	3.4 源程序书写风格	50
2.1.1 Anaconda 安装过程	15	3.4.1 注释	50
2.1.2 Anaconda 中的 Python 编辑器 Spyder	19	3.4.2 缩进	50
2.2 Anaconda 其他常用功能	24	3.4.3 语句换行	50
2.2.1 常用应用程序及 pip 模块安装命令	24	3.4.4 必要的空格与空行	51
2.2.2 常用科学包介绍	25	3.5 赋值语句	51
2.3 了解 IDLE 开发环境	30	3.6 基本输入/输出	51
2.3.1 启动 IDLE	31	3.6.1 Python 基本输入	51
2.3.2 使用 IDLE 编写 Python 程序	32	3.6.2 Python 基本输出	52
2.4 使用 PyCharm 编辑器	33	3.7 字符串操作	56
2.4.1 下载安装	34	3.7.1 字符串的存储方式	56
2.4.2 新建项目	34	3.7.2 字符串切片	56
2.4.3 配置 PyCharm	36	3.7.3 字符串运算符	57

3.8 Python 内置函数	58	6.1.1 一个最简单的函数	97
3.8.1 Python 数学函数	58	6.1.2 函数的语法结构	97
3.8.2 字符串函数	59	6.1.3 函数的调用	98
3.8.3 类型判断和类型间转换	61	6.1.4 return [表达式]	98
练习三	62	6.2 函数调用过程中的参数传递	99
第4章 程序控制	66	6.2.1 必选参数	99
4.1 选择结构	66	6.2.2 关键字参数	100
4.2 循环结构	70	6.2.3 默认参数	100
4.2.1 while 语句	70	6.2.4 不定长参数	101
4.2.2 for 语句	71	6.3 匿名函数	102
4.2.3 break 语句、continue 语句和 pass 语句	73	6.3.1 匿名函数的定义	102
4.2.4 循环嵌套	74	6.3.2 匿名函数的特征	102
练习四	75	6.3.3 匿名函数使用	103
第5章 列表、元组、字典、集合	78	6.4 局部变量和全局变量	104
5.1 列表	78	6.4.1 变量的作用域	104
5.1.1 列表的创建与删除	79	6.4.2 局部变量和全局变量	104
5.1.2 列表元素的访问	79	6.4.3 global 和 nonlocal 关键字	105
5.1.3 用列表的基本操作实现案例任务	81	6.5 系统内置函数介绍	106
5.1.4 列表的高级操作	82	练习六	109
5.1.5 用列表的高级操作实现案例任务	82	第7章 模块	112
5.2 元组	83	7.1 模块概述	112
5.2.1 元组的创建与删除	83	7.2 模块的导入	113
5.2.2 访问元组的元素	84	7.2.1 直接使用 import 导入模块	113
5.2.3 元组的高级操作	84	7.2.2 通过 sys 模块导入自定义模块 的 path	115
5.2.4 用元组实现案例任务	85	7.2.3 通过.pth 文件找到自定义模块	115
5.3 字典	86	7.3 第三方库的使用	116
5.3.1 字典的创建和删除	86	7.4 包	116
5.3.2 字典元素的基本操作	87	7.4.1 引入包的原因	116
5.3.3 用字典实现案例任务	88	7.4.2 导入和使用包	117
5.3.4 字典的高级操作	90	练习七	118
5.4 集合	90	第8章 面向对象程序设计	119
5.4.1 集合的基本操作	91	8.1 面向对象的程序设计思想	119
5.4.2 用集合实现案例任务	91	8.2 类的定义和使用	120
5.4.3 集合运算	92	8.2.1 创建类和对象	120
5.5 列表、元组、字典、集合的应用小结	93	8.2.2 数据成员	121
练习五	94	8.2.3 方法	122
第6章 函数	97	8.3 类的继承与方法重载	124
6.1 函数的基本结构	97	8.3.1 派生类的定义和使用	125

8.3.2 方法重载	126	9.2.3 使用 CSV 模块读写文件	144
8.4 面向对象程序设计应用小结	127	9.3 其他类型文件的读取与写入	145
练习八	128	9.3.1 xlrd 模块读取 Excel 文件	145
<b>第 9 章 文件</b>	<b>129</b>	9.3.2 xlwt 模块写 Excel 文件	146
9.1 文件的基本操作	129	9.4 使用 OS 模块处理文件和目录	146
9.1.1 文件与文件对象	129	9.4.1 os 模块常用方法	147
9.1.2 打开和关闭文件	130	9.4.2 遍历目录及子目录	147
9.1.3 文件写入	132	9.5 文件使用综合实例	148
9.1.4 顺序读取写入文件的内容	133	练习九	150
9.1.5 随机读取文件内容的方法	136	<b>第 10 章 异常处理</b>	<b>152</b>
9.1.6 综合实例——使用文件来保存游戏	138	10.1 什么是异常?	152
9.2 读取存储 CSV 格式	142	10.2 如何处理异常?	154
9.2.1 CSV 文件格式	142	10.3 处理多个异常	155
9.2.2 常规 CSV 文件存取过程	142	练习十	156

## 第二部分 项目实践

<b>项目 1 根据函数绘制曲线</b>	<b>158</b>	13.1 使用 urllib 爬取数据	179
11.1 主要问题	158	13.2 相关知识	180
11.1.1 turtle 模块绘制坐标系	158	13.2.1 网络请求与响应	180
11.1.2 使用 NumPy 模块获得散点数列	160	13.2.2 字符编码与解码	181
11.1.3 2D 绘图模块 turtle 动画展示曲线	160	13.2.3 HTML	182
11.2 项目实施	161	13.3 用 etree 和 xpath 提取数据	183
11.2.1 程序流程	161	13.4 将提取数据的数据存到列表中	186
11.2.2 数据代码	162	13.5 将列表中的数据存为 CSV 文件和 TXT 文件	186
11.2.3 运行结果	164	13.5.1 用 time 库获取日期时间	187
11.3 项目扩展	165	13.5.2 使用 Pandas 库实现数据处理	187
11.3.1 使用 matplotlib 显示静态曲线	165	13.6 分词数据和词云数据分析	189
11.3.2 打包成 exe 文件	166	13.6.1 使用 jieba 库实现分词	190
<b>项目 2 分析微信好友数据</b>	<b>168</b>	13.6.2 使用 wordcloud 库生成词云	190
12.1 项目准备	168	13.6.3 使用 matplotlib 库实现词云 的可视化	191
12.1.1 微信模块 wxpy	168	<b>项目 4 Django+MySQL Web 开发</b>	<b>193</b>
12.1.2 绘图模块 matplotlib.pyplot	171	14.1 概述	193
12.2 项目实施	174	14.2 Django 简介	195
12.2.1 程序流程	174	14.2.1 框架介绍	195
12.2.2 数据处理	175	14.2.2 架构设计	195
12.2.3 全部代码	176	14.2.3 工作机制	196
12.2.4 运行结果	177		
<b>项目 3 网页数据下载与处理</b>	<b>179</b>		

14.2.4	部署	197	14.5.3	项目启动	200
14.2.5	文档	197	14.5.4	Web 工程添加页面	201
14.3	MVC 框架介绍	197	14.5.5	动态数据显示	203
14.4	Python Django 的安装	198	14.5.6	数据库准备	204
14.5	使用 PyCharm 和 Django 创建简单的 Web 服务器	198	14.5.7	连接数据库	207
14.5.1	软件安装	199	附录	全国计算机等级考试二级 Python 语言 程序设计考试大纲 (2018 年版)	211
14.5.2	PyCharm 新建 Django 工程	199			

# 第一部分 基础知识

## 第 1 章 绪论



### 本章导读

本章从计算机发展历史开始介绍计算机的软硬件基本组成,使同学们对计算机硬件和软件编程建立初步的概念。本章还对 Python 语言入门涉及到的知识进行了介绍。



### 本章要点

- 计算机发展历史
- 冯·诺依曼结构
- 操作系统与应用软件
- Python 语言介绍

## 1.1 计算机发展历史

### 1.1.1 计算机的历史

#### 1. 机械计算机

最早的计算工具是我国的算盘,距今已有 2000 多年的历史。人们按照一定的规则上下拨动算珠,可以进行不同类型的计算。

在 8 至 9 世纪,阿拉伯数字传入欧洲。在 15 世纪,John Napier 发明了奈氏骨牌,可以进行数值乘法。

第一台机械式计算器于 1642 年由 Blaise Pascal 实现,可以进行加法运算。1694 年由 Gottfried Wilhem von Leibniz 实现的计算器除了能做加法,还能通过连续相加和移位实现乘法运算。

第一台商用机械计算器由 Charles Xavier Thomas 实现,可以做加减乘除的运算。

真正被认为是现代计算机前身的是 Babbage 在 1833 年对分析机的设想。该机可并行对 50 位十进制数字的数值进行操作,并可存储 1000 个这样的数。它包括输入设备、控制单元、处理机、存储器以及输出设备等现代计算机所具备的器件。可惜由于加工的问题,该机最终没能实现。

Babbage 还提出了穿孔卡片的设想。穿孔卡片最终在 1890 年实现。穿孔卡片不仅仅提供了输入输出的方法,而且还是一种存储数字的存储器。它在计算机发展史上占有重要的地位,成为许多计算机公司的开端。著名的 IBM 公司的前身就是从穿孔卡片开始的。

## 2. 电子计算机

1942年,物理学家 John V. Atanasoff 发明了第一台电子计算机。该计算机使用了现代数字开关技术。

第二次世界大战极大地推动了计算机的发展。在这种背景下,1946年,John P. Eckert 和 John W. Mauchly 及同事建造了第一台电子计算机,即 ENIAC。该机使用了约 18000 个真空管,占地  $167\text{m}^2$ ,耗电达 180kW,采用卡片打孔作为 I/O,可运算 10 位数字,每秒可进行 300 次乘法运算。但运行不同的程序,ENIAC 要重新接线或设置开关。

对计算机进行编程的思想是 1945 年由数学家 John V. Neumann 提出,其思想被称为“存储程序技术”。第一代可编程电子计算机出现在 1947 年,这些计算机使用了 RAM (随机存储器) 存储程序和数据。

### 1.1.2 个人计算机的历史

#### 1. 第 1 阶段

1971—1973 年是 4 位和 8 位低档微处理器时代,通常称为第 1 代,其典型产品是 Intel 4004 和 Intel 8008 微处理器和分别由它们组成的 MCS-4 和 MCS-8 微机。Intel 4004 是一种 4 位微处理器,由于功能有限,主要用于计算器、电动打字机、照相机、台秤、电视机等家用电器上,使这些电器设备具有智能化,从而提高它们的性能。Intel 8008 是世界上第一种 8 位的微处理器,存储器采用 PMOS 工艺,用于简单的控制场合。

#### 2. 第 2 阶段

1971—1977 年是 8 位中高档微处理器时代,通常称为第 2 代,其典型产品是 Intel 8080/8085、Motorola 公司的 M6800、Zilog 公司的 Z80 等。它们的特点是采用 NMOS 工艺,集成度提高约 4 倍,运算速度提高约 10~15 倍(基本指令执行时间  $1\sim 2\mu\text{s}$ ),指令系统比较完善,具有典型的计算机体系结构和中断、DMA 等控制功能。它们均采用 NMOS 工艺,集成度约 9000 只晶体管,平均指令执行时间为  $1\mu\text{s}\sim 2\mu\text{s}$ ,采用汇编语言、BASIC、Fortran 编程,使用单用户操作系统。

#### 3. 第 3 阶段

1978—1984 年是 16 位微处理器时代,通常称为第 3 代,其典型产品是 Intel 公司的 8086/8088、Motorola 公司的 M68000、Zilog 公司的 Z8000 等微处理器。其特点是采用 HMOS 工艺,集成度(20000~70000 晶体管/片)和运算速度(基本指令执行时间是  $0.5\mu\text{s}$ )都比第 2 代提高了一个数量级。指令系统更加丰富、完善,采用多级中断、多种寻址方式、段式存储机构、硬件乘除部件,并配置了软件系统。1981 年 IBM 公司推出的个人计算机采用 8088CPU。

1982 年,英特尔公司在 8086 的基础上,研制出了 80286 微处理器。1984 年,IBM 公司推出了以 80286 处理器为核心组成的 16 位增强型个人计算机 IBM PC/AT。由于 IBM 公司在发展个人计算机时采用了技术开放的策略,从而使个人计算机风靡世界。

#### 4. 第 4 阶段

1985—1992 年是 32 位微处理器时代,又称为第 4 代,其典型产品是 Intel 公司的 80386/80486, Motorola 公司的 M69030/68040 等。其特点是采用 HMOS 或 CMOS 工艺,集成度高达 100 万个晶体管/片,具有 32 位地址线和 32 位数据总线。每秒钟可完成 600 万条指令 (Million Instructions Per Second, MIPS)。

1989年,大家耳熟能详的80486芯片由英特尔推出。这款经过四年开发、耗费3亿美元资金投入的芯片的伟大之处在于它首次突破了100万个晶体管的界限,集成了120万个晶体管,使用了 $1\mu\text{m}$ 的制造工艺。80486的时钟频率从25MHz逐步提高到33MHz、40MHz、50MHz。

### 5. 第5阶段

1993—2005年是奔腾(Pentium)系列微处理器时代,通常称为第5代,其典型产品是Intel公司的奔腾系列芯片及与之兼容的AMD的K6系列微处理器芯片。内部采用了超标量指令流水线结构,并具有相互独立的指令和数据高速缓存。随着MMX(Multi Media eXtended)微处理器的出现,使微机的发展在网络化、多媒体化和智能化等方面跨上了更高的台阶。

多能奔腾是继Pentium之后英特尔又一个成功的产品,其生命力也相当顽强。多能奔腾在原Pentium的基础上进行了重大的改进,增加了片内16KB数据缓存和16KB指令缓存,4路写缓存以及分支预测单元和返回堆栈技术。特别是新增加的57条MMX多媒体指令,使得多能奔腾即使在运行非MMX优化的程序时,也比同主频的Pentium CPU要快得多。

1997年推出的Pentium II处理器结合了Intel MMX技术,能以极高的效率处理影片、音效以及绘图资料,首次采用Single Edge Contact(S.E.C)匣型封装,内建了高速快取记忆体。这款芯片让计算机使用者获取、编辑以及通过互联网和亲友分享数码相片,编辑与新增文字、音乐或制作家庭电影特效,使用视频电话以及在网上传送影片。Intel Pentium II处理器晶体管数目为750万个。

1999年推出的Pentium III处理器加入70个新指令,加入的SIMD扩展指令集称为MMX,能大幅提升视频、3D、音乐等应用的性能,还能大幅提升互联网的使用经验,让使用者能浏览逼真的线上博物馆与商店,以及下载高品质影片。Intel首次导入 $0.25\mu\text{m}$ 技术,Intel Pentium III晶体管数目约为950万个。

2000年推出的Pentium 4处理器内建了4200万个晶体管,采用 $0.18\mu\text{m}$ 的电路。Pentium 4初期推出版本的速度就高达1.5GHz,晶体管数目约为4200万个。

### 6. 第6阶段

2005年至今是酷睿(Core)系列微处理器时代,通常称为第6代。“酷睿”是一款领先节能的新型微架构,设计的出发点是提供卓然出众的性能和能效,提高每瓦特性能,也就是所谓的能效比。早期的酷睿是基于笔记本处理器的。

2010年6月,Intel再次发布革命性的处理器——第2代酷睿i3/i5/i7。Intel在酷睿架构后推出了Tick-Tock战略,每两年升级一次工艺,每两年升级一次架构,每年都有新一代处理器问世。2018年,第8代酷睿i3/i5/i7已占据市场主流,已有第9代酷睿处理器曝光。

## 1.2 软硬件基本知识

### 1.2.1 冯·诺依曼结构

PC(Personal Computer,个人计算机)是由硬件系统和软件系统两大部分组成的。

硬件(Hardware)是指实际的物理设备,包括计算机的主机及其外部设备。

软件(Software)是指实现算法的程序及其文档,包括计算机本身运行所需的系统软件(System Software)和用户完成特定任务所需的应用软件(Application Software)等。

## 1. PC 的硬件结构

冯·诺伊曼 (Von Neumann) 提出的 EDVAC 方案明确提出电子计算机由运算器、控制器、存储器、输入设备和输出设备五个部分构成, 并描述了这五部分的功能和相互关系, 提出了两个重要思想: 存储程序和使用二进制。

一个复杂的运算, 总可以分解成一系列简单的基本操作步骤。用计算机解题, 首先应确定这个问题的算法, 明确运算过程, 然后用一系列计算机可直接执行的基本操作即指令组成的程序来实现这个过程。“存储程序”的思想是: 把程序和所需数据事先以一定顺序存储在计算机的存储器中, 运行时从存储器中逐一取出程序中一条条指令, 并实现其基本操作, 最后达到解题的目的。存储程序摆脱了 ENIAC 那种烦琐费时的线路连接方法, 使计算机的应用通用化, 运行真正自动化。

(1) 二进制。“二进制”的思想是指在计算机中, 指令和数据都以二进制形式表示。计算机采用二进制, 而不采用十进制, 是因为二进制数据有以下一些主要特点:

1) 二进制数容易表示。二进制数只含有两个数字 0 和 1, 因此可用大量存在的具有两个不同稳定物理状态的元件来表示, 例如: 可用指示灯的不亮和亮、继电器的断开和接通、晶体管的截止和导通、磁性元件的反向和正向剩磁、脉冲电位的低和高等, 来分别表示二进制数字的 0 和 1。计算机中采用具有两个稳定状态的电子或磁性元件表示二进制数, 这比十进制的每一位要用具有 10 个不同的稳定状态的元件来表示, 实现起来要容易得多。同时, 由于表示二进制数的元件的状态数少, 故数据传送不易出错, 工作稳定可靠。

2) 二进制数的运算规则简单。二进制数的加法和乘法的运算规则都比十进制数简单得多, 这使得计算机中的运算部件的结构也相应比较简单。它的加法规则和乘法规则都只有 4 条:

$$0+0=0 \quad 0+1=1 \quad 1+0=1 \quad 1+1=10 \quad 0 \times 0=0 \quad 0 \times 1=0 \quad 1 \times 0=0 \quad 1 \times 1=1$$

实际上, 由于二进制数的运算规则简单, 它的算术运算都可通过移位、相加这两种简单的操作来实现, 这就使得计算机的运算部件结构简单、运行可靠。

3) 二进制可进行逻辑计算。因二进制数的两个数字 0 和 1 与逻辑代数的逻辑变量取值一样, 故可采用二进制数进行逻辑运算。这样就可以运用逻辑代数作为工具来分析和设计计算机中的逻辑电路, 使得逻辑代数成为计算机设计的数学基础。

这一思想冲破了 ENIAC 中仍采用的传统十进制的影响, 使得计算机的存储容量大大增加, 结构大为简化, 运算速度大大提高。

为了和十进制区别, Python 在书面表示时, 二进制采用 0b 或 0B 作为前缀, 八进制采用 0o 或 0O 作为前缀, 十六进制采用 0x 或 0X 作为前缀。

(2) 冯·诺伊曼结构。“存储程序和程序控制”的思想为冯·诺伊曼机的基本结构, 即计算机硬件系统由运算器、控制器、存储器、输入设备和输出设备组成。到目前为止, 几乎所有计算机的结构都是冯·诺伊曼结构。

1) 运算器。运算器 (arithmetic unit) 的主要功能是完成对数据的算术运算、逻辑运算和逻辑判断等操作。在控制器的控制下, 它对取自存储器或寄存在其内部寄存器的数据进行算术或逻辑运算, 其结果暂存内部寄存器或送到存储器。

2) 控制器。控制器 (control unit) 的主要作用是控制各部件的工作, 使计算机自动地执

行程序。它按存储顺序取出指令，并对指令进行分析，然后向各部件发出相应的控制信号，使这些部件协调动作，完成指令所规定的操作。这样逐一执行一系列指令，使计算机能够按照这一系列指令组成的程序的要求自动运行。

控制器和运算器合在一起成为中央处理器（Central Processing Unit, CPU），它是计算机的核心部件。

3) 存储器。存储器（memory）是用来存储程序和数据部件。用户先通过输入设备把程序和数据存储在存储器中。运行时，控制器从存储器逐一取出指令加以分析，发出控制命令以完成指定的操作；根据控制命令，从存储器取出数据送到运算器中运算或把运算器中的结果送到存储器保存。由此可见，既可从存储器进行“读”，又可对存储器进行“写”。

衡量存储器的指标有三：一是存储容量，二是存储速度，三是价格。人们从未感到过存储容量已经够用，存储速度已经够快。因此存储器总是不断发展的，即其容量越来越大，速度越来越快，价格越来越低，体积越来越小，耗电越来越省，寿命越来越长。存储器的种类很多，有快有慢，有贵有贱。于是人们在进行存储系统的设计时，大都采用多种类型的存储器建立一个存储层次体系。因为速度快的存储器价格贵，容量就不能做得很大；而价格便宜的存储器可以把容量做得很大，但它的存储速度却比较慢。因此，设计人员必须在容量、速度、价格三者之间进行权衡。

典型的存储层次包括高速缓冲存储器（cache）、主存储器（main memory）、辅助存储器（auxiliary storage）、海量存储器（mass storage），存取速度由高自低依次排列。高速缓冲存储器采用速度很高的半导体静态存储器，有的微型计算机甚至把它和微处理器做在一起。主存储器的速度也比较高，常采用半导体动态存储器。辅助存储器则主要用硬磁盘或软磁盘，而海量存储器采用容量极大的磁带或光盘。粗略地分，存储器则有内存和外存两个层次。

运算器、控制器、内存储器合起来称为计算机的主机。

4) 输入设备和输出设备。输入（input）设备能把程序、数据、图形、图像、声音、控制现场的模拟量等信息，通过输入接口转换成计算机可以接收的电信号。常用的输入设备有键盘、鼠标器、操纵杆、卡片输入机、纸带输入机、光笔、语音识别装置、数字化仪、扫描仪、条形码阅读器、光学字符阅读机（Optical Character Reader, OCR）、调制解调器（modem）及各种模/数（A/D）转换器等。输出（output）设备能把计算机运行结果或过程，通过输出接口转换成人们所要求的直观形式或控制现场能接受的形式。常见的输出设备有显示器、打印机、绘图仪、卡片穿孔机、纸带穿孔机、语音合成装置、缩微胶卷输出设备、调制解调器及各种数/模（D/A）转换器等。

## 2. PC 系统的类型

PC 系统的类型可以从软件和硬件两个角度划分，我们在这里主要讨论从硬件方面划分的类型。

当 CPU 读取数据时，数据通过外部数据总线连接进入 CPU。不同的 CPU 有不同的数据总线宽度，主板为了连接数据源和 CPU，就必须设计与 CPU 总线宽度相匹配的总线。例如，如果 CPU 的数据总线宽度为 32 位，则主板的总线宽度也应该是 32 位，这样系统可以在单个周期内移动 32 位有效数据进出处理机。CPU 的数据总线宽度是主板和存储器系统设计的主要影响因素，它指明了在一个周期内有多少位能移入或移出芯片。486 系统具有 32 位 CPU 总线。Pentium 系列的 CPU 都具有 64 位的数据总线，即支持 Pentium 系列 CPU 的主板总线宽度也是

64 位。不能将 64 位的 CPU 用在 32 位的主板上。

因此，我们可以将 PC 系统按硬件分为 8 位、16 位、32 位和 64 位等类型。

### 3. PC 的系统部件

现代 PC 既简单又复杂。说它简单，是因为这些年来系统所用的许多元器件被集成为越来越少的组件；说它复杂，是因为现代 PC 系统中的每一个部件都完成了比以前系统中相同部件多得多的功能。

下面是组装一个现代计算机系统所需要的基本组件：

(1) 主板。主板是系统的核心。所有的其他部件都通过主板相互连接。可用的主板有几种不同的外形和形状参数。主板上通常包含下面部件：CPU 插座；芯片组；存储器插槽；ROM BIOS；超级 I/O 芯片；总线槽；时钟；电池等。

芯片组中包含最重要的电路。芯片组控制着 CPU 总线、存储器、PCI 总线、ISA 总线、系统资源等。它体现了主板的主要特性和规范。

芯片组决定了系统能够支持什么样的 CPU，可以安装什么样的存储器，机器所能支持的最高频率是多少以及系统采用何种外部总线。在主板上还有一个 ROM 存储芯片，用来存放 BIOS 信息，其中包括 POST 自检程序、自举装载程序和部分硬件的基本驱动程序等。

(2) CPU。CPU（中央处理单元）也称为中央处理器，可以说它是整个计算机的大脑，所有的指令和程序都是在这里执行的，它是在很小的硅片上集成了几百万个晶体管。计算机的性能和执行指令的速度很大程度上取决于它。

(3) 存储器。这里的存储器指的是系统存储器，通常称为 RAM（随机存储器），用来保存 CPU 所需要的数据和指令。由于 RAM 本身的特性，在断电后所有的内容都会被清除。在重新打开电源后，需要重新装入程序和指令供 CPU 使用。计算机一开始装入的程序称为初始化程序，是存放在 ROM（只读存储器）中，ROM 不同于 RAM，它在断电后内容不会消失。

(4) 显示适配器。现在的计算机都采用视频系统来显示和用户的交互信息，包括文本、图像和多媒体信息等。显示适配器（通常被称为显示卡或显卡）用来控制这些在屏幕上显示的信息。显示适配器有四个基本组成部件：显示主芯片、显示存储器（称为显存或视频 RAM）、数/模（D/A）转换器以及视频 BIOS。显示主芯片通过向显示存储器中写入数据来控制屏幕上显示的内容。

(5) 显示器。显示器就是 PC 机和用户直接交互的屏幕，一般采用独立部件与机箱分离实现，也有的便携系统或原装 PC 将显示器与机箱整合在一起。

(6) 机箱。机箱（又称机壳或机架）是组成 PC 的基础，所有 PC 的其他部件都安装在机箱内。我们可根据不同的应用选择不同的机箱设计。

(7) 电源。由于 PC 的各部件一般都要求直流供电，因此在 PC 机中我们需要一个电源转换器把 220V 交流电（有的国家市电标准是 110V AC）转换成各部件所要求的 3.3V、5V 或 12V 直流电。电源的质量对整个系统的稳定性有很大的影响。

(8) 软盘驱动器/硬盘驱动器：软盘驱动器和硬盘驱动器都采用磁存储原理存储数据。软盘驱动器由于速度慢、容量小，已经被淘汰。硬盘驱动器的容量和读写速度要比软盘驱动器大和快。在现代 PC 中，操作系统以及绝大部分的应用软件都存放在硬盘驱动器中供系统使用。硬盘驱动器不同的型号有不同的容量和其他特性。

(9) CD-ROM。CD-ROM 和 DVD-ROM 是大容量可更换介质的存储器。采用特殊的设

备(一般采用刻录机)可以实现对盘片的写和重写,由于设备价格和介质价格的不断降低,在现在的 PC 中都有配置,而且在主板的 BIOS 中也提供了对 CD-ROM 的启动支持。

(10) 声卡和音箱。在 PC 中加入声卡和音箱后就可以代替 PC 扬声器实现 PC 对声音的回放和处理。目前大部分的主板都集成有声卡,也有单独的声卡可供用户选择。高质量的声卡可提供高效的声音回放、精美的声音合成功能。

要使 PC 发出声音,需要为声卡连接音箱,一般建议采用有源音箱实现较高质量的声音回放。音箱是一个能将模拟电子脉冲信号转换成机械性的振动,并通过空气的振动再形成人耳可以听到的声音的装置。

(11) 键盘。键盘目前是 PC 上的默认输入设备,用户通过键盘与 PC 实现信息的交互。键盘根据不同的应用有不同的布局、尺寸和外形,有些键盘还有专用的功能键和符号键。而且由于键盘采用标准接口与主机相连接,用户可以自由选择适合自己的键盘。

(12) 鼠标。鼠标是另一种通用的输入设备,它是随着采用图形用户接口(GUI)的操作系统的出现而出现的。目前市场上也有其他类型的定位设备,但由于鼠标的廉价、方便,因此应用最为广泛。标准鼠标由两个按键进行输入,也有的鼠标由三个按键和滚轮来实现更多的功能。

(13) 其他可选硬件。计算机系统除了键盘和鼠标这些常用设备外,根据需要也可选取以下的可选硬件,以便扩充其输入/输出功能。

1) 扫描仪。扫描仪是一种主要用于图像输入的设备。其内部主要包括灯管、光电耦合器、机械传动结构、主控单元等。它工作时,主要是利用主控单元完成复杂的图像扫描和图像预处理运算,然后通过接口卡完成计算机和扫描仪之间的数据传输。常用的扫描仪有三种,即手持式扫描仪、台式扫描仪以及滚筒式扫描仪。

2) 打印机等设备。常用的打印机有三种,即针式打印机、喷墨打印机、激光打印机,它们的主要任务是在纸或者其他介质上输出可以永久保存的图形和文字。

## 1.2.2 计算机软件发展

面对只有硬件的计算机(称为裸机),人们只能用机器语言编制解题程序,这给计算机的应用带来了极大不便。如果要方便灵活地使用计算机解决实际问题,还必须给它配备各种软件。如果把计算机的硬件比作人的身躯,则计算机的软件就是人的灵魂。计算机软件是相对于计算机硬件而言的。一般地说,软件是所有程序及数据结构的总称。1983年,美国电气和电子工程学会明确地给软件下了这样一个定义:软件是计算机程序、方法、规则、相关的文档以及在计算机上运行它时所必需的数据。这一定义深刻阐明了软件的实质,也充分表明了软件与程序的区别。随着计算机硬件制造工艺技术的发展,硬件成本已不断下降,而软件成本在整个计算机系统中所占比重越来越大。

软件通常分为系统软件和应用软件两大类。系统软件是计算机设计制造者提供的使用和管理计算机的软件,包括操作系统、语言处理系统、数据库管理系统和常用服务程序等。应用软件是用户利用计算机提供的各种系统软件而开发的解决各种实际问题的软件。

### 1. 微机上的典型操作系统

在所有软件中,操作系统(Operating System, OS)是最基本、最重要的,它是对“裸机”功能的第一次扩充,所有其他软件都须通过操作系统对硬件功能进行再扩充,它们须在操作系统的统一管理和支持下运行,可以说操作系统是硬件与其他软件的接口。用户不是直接使用计

算机硬件资源，而是通过操作系统提供的命令去方便地操作计算机，所以它又是用户与计算机之间的接口。

操作系统的作用是控制和管理计算机的硬件资源（如中央处理器、存储器、打印机及调制解调器、鼠标等）和软件资源，从而提高计算机的利用率，方便用户使用计算机。

在微型计算机系统中，常用的操作系统有：

PC-DOS(Personal Computer Disk Operating System)最初是美国 Microsoft 公司专门为 IBM 的 PC 系列微型计算机开发并授权给 IBM 公司使用的一种单用户操作系统，1993 年 9 月协议失效后，IBM 不断推出高版本的 PC-DOS，而 Microsoft 则不断推出高版本的 MS-DOS，直到 MS-DOS6.22。

UNIX 是美国贝尔实验室开发的通用多用户多任务交互式的操作系统，其源代码的 90% 以上以及绝大部分系统程序都是用 C 语言编写的，因而具有很好的可移植性。UNIX 最初适用于小型计算机，20 世纪 80 年代后，在大量有影响的大、中、小型机上都可以运行 UNIX，特别是以微机为硬件环境的分时系统进入市场后，UNIX 得到极大推广，各种 16 位、32 位微机竞相移植 UNIX，因而被公认为 32 位微机的主要操作系统。

Windows 是 Microsoft（微软）公司推出的一种运行在 286 以上机型上的图形窗口式操作系统。1985 年 11 月，Windows 1.0 问世。1992 年 4 月，Microsoft 推出了 Windows 3.1。Windows 3.1 增加了 TrueType 字体，用户界面更加友好，并首次采用了对象链接与嵌入(Object Linking and Embedding, OLE) 技术。从 1983 年提出设想，到 1994 年 Windows 3.11 问世，经历了十几年的风风雨雨，最终成为用户乐于使用的个人计算机操作系统，并主宰了个人计算机操作系统产品的市场。1995 年 8 月，Microsoft 推出最新的 Windows 95，它是一个真正的 32 位操作系统软件，正是因为 Windows 95 的推出，全世界计算机产业的发展更加迅速。现在还有 32/64 位的 Windows XP、Windows 7、Windows 8、Windows 10 等可供用户选择使用。

## 2. 语言处理系统

语言处理系统由各种程序设计语言处理程序（亦称编译程序）组成，它位于操作系统的外层。语言处理程序分三类：汇编程序、解释程序和编译程序。

(1) 汇编程序。汇编程序(assembly)是将汇编语言编写的源程序翻译加工成机器语言表示的目标程序的一种软件。汇编程序一般还提供查错、修改等功能，并对源程序中出现的伪指令等作相应的处理。

(2) 解释程序。解释程序(interpreter)将高级程序设计语言编写的源程序按动态的运行顺序逐句进行翻译并执行，即每翻译一句就产生一系列完成该语句功能的机器指令并立即执行这一系列机器指令，如此进行直至源程序运行结束。在这一过程中，若出现错误，则系统会显示出错信息，待修正后才能继续下去。解释程序的这种工作方式，便于实现人机会话。

(3) 编译程序。编译程序(compiler)是能将用高级程序设计语言编写的源程序翻译成用汇编语言或机器语言表示的目标程序。

编译程序把源程序翻译成目标程序一般经过词法分析、语法分析、中间代码生成、代码优化和目标代码生成五个阶段。

从以上所述可见，汇编程序和编译程序都产生目标程序，而解释程序不产生目标程序，解释程序可提供人机对话的工作方式，使得用户对源程序的调试、修改和扩充比较方便，但程

序执行的速度比较慢；编译程序对源程序进行编译产生目标程序，将来执行的速度较快，但对源程序修改后必须重新编译。

### 3. 数据库系统

随着计算机在信息管理领域中日益广泛深入的应用，不仅产生了数据库系统（Data Base System），随之还出现了各种数据库管理系统（Data Base Management System, DBMS）。

数据库系统是一门综合的软件技术，它研究如何有效地组织数据和方便地处理数据。数据库系统是一个记载和维护数据信息的系统，它由数据、硬件、软件 and 用户四部分构成。

数据是数据库系统的重要资源。在系统中，一般把它组织成一些数据库存储，它具有冗余小、可共享等特点。所谓数据库，可理解成按一定的方式组织起来的操作数据的集合。

硬件是数据库系统的物质基础，包括存储系统中数据的存储设备以及有关的控制设备，如硬盘、光盘等。

软件是数据库管理系统。它是用户和物理数据库之间的接口，能把数据库的物理细节屏蔽起来，向用户提供一个使用方便灵活的友善的工作界面。

数据库管理系统是一组软件的集合，用来定义数据库，帮助和控制用户为增加、删除、修改和检索数据时对数据库的访问和使用，提供数据独立性、完整性和安全性的保障。

### 4. 应用软件

在计算机硬件和系统软件的支持下，面向具体问题和具体用户的软件称为应用软件。应用软件是一些具有一定功能并满足一定要求的应用程序的组合。目前，一些应用软件有的已逐步标准化、模块化，形成了解决某类典型问题的应用程序组合，即软件包（package），如财务管理软件包、统计软件包、运筹学软件包等。

随着计算机应用的日益广泛深入，各种应用软件的数量不断增加，质量日趋完善，使用更加灵活方便，通用性越来越强，人们只要略加学习一些基础知识和基本操作方法，就可以利用这些应用软件进行数据处理、文字处理、辅助设计等。

必须指出，如同软硬件的界限在不断变化一样，系统软件和应用软件之间也不存在明显的界限。

## 1.3 常用编程语言介绍

### 1.3.1 机器语言

计算机指令系统中的指令由 0 和 1 组成并且能被机器直接理解执行，它们被称为机器指令。机器指令的集合就是该计算机的机器语言，即计算机可以直接接收、理解的语言。机器语言能利用机器指令精确地描述算法，并且所编程序占用内存空间少、执行速度快。

但是我们已经看到，用机器语言编写程序是一件十分烦琐的工作，不仅要记住用 0、1 代码表示的各条指令的不同功能，而且这种全部由 0、1 代码组成的程序，直观性很差，容易出错，阅读检查和修改调试都较困难。

不仅如此，由于不同类型的计算机的指令系统不同，机器语言也不同。因此，机器语言是一种面向机器的语言。