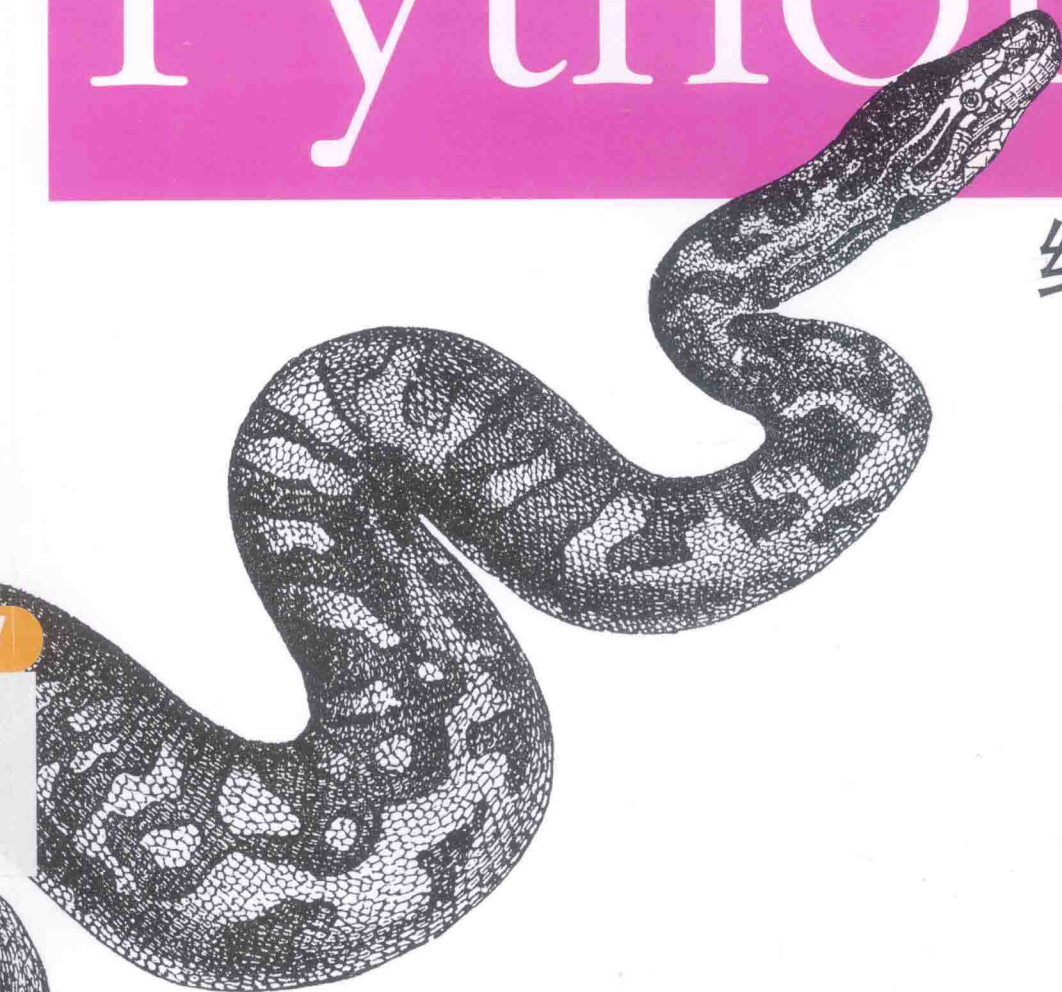


Programming Python

第4版(上册)
涵盖 Python 3.x

Python

编程



O'REILLY®

中国电力出版社

Mark Lutz 著

邹晓 瞿乔 任发科 等译

第四版

Python编程

Mark Lutz 著
邹晓 瞿乔 任发科 等译

O'REILLY®

Beijing • Cambridge • Farnham • Köln • Sebastopol • Tokyo

O'Reilly Media, Inc. 授权中国电力出版社出版

中国电力出版社

图书在版编目 (CIP) 数据

Python编程: 第4版/ (美) 卢茨 (Lutz, M.) 著; 邹晓等译. —北京: 中国电力出版社, 2015.1

书名原文: Programming Python, Fourth Edition

ISBN 978-7-5123-5530-9

I. ①P… II. ①卢… ②邹… III. ①软件工具—程序设计 IV. ①TP311.56

中国版本图书馆CIP数据核字 (2014) 第024404号

北京市版权局著作权合同登记

图字: 01-2012-7973号

©2011 by O'Reilly Media, Inc.

Simplified Chinese Edition, jointly published by O'Reilly Media, Inc. and China Electric Power Press, 2014.
Authorized translation of the English edition, 2011 O'Reilly Media, Inc., the owner of all rights to publish and sell the same.

All rights reserved including the rights of reproduction in whole or in part in any form.

英文原版由O'Reilly Media, Inc. 出版2011。

简体中文版由中国电力出版社出版2014。英文原版的翻译得到O'Reilly Media, Inc.的授权。此简体中文版的出版和销售得到出版权和销售权的所有者——O'Reilly Media, Inc.的许可。

版权所有, 未得书面许可, 本书的任何部分和全部不得以任何形式重制。

封面设计/ Karen Montgomery, 张健
出版发行/ 中国电力出版社 (www.cepp.sgcc.com.cn)
地 址/ 北京市东城区北京站西街19号 (邮政编码100005)
经 销/ 全国新华书店
印 刷/ 北京丰源印刷厂
开 本/ 787毫米×980毫米 16开本 95印张 1796千字
版 次/ 2015年1月第一版 2015年1月第一次印刷
印 数/ 0001-3000册
定 价/ 198.00元 (上下册)

敬告读者

本书封底贴有防伪标签, 刮开涂层可查询真伪
本书如有印装质量问题, 我社发行部负责退换

版权专有 翻印必究

O'Reilly Media, Inc. 介绍

O'Reilly Media通过图书、杂志、在线服务、调查研究和会议等方式传播创新知识。自1978年开始，O'Reilly一直都是前沿发展的见证者和推动者。超级极客们正在开创着未来，而我们关注真正重要的技术趋势——通过放大那些“细微的信号”来刺激社会对新科技的应用。作为技术社区中活跃的参与者，O'Reilly的发展充满了对创新的倡导、创造和发扬光大。

O'Reilly为软件开发人员带来革命性的“动物书”；创建第一个商业网站（GNN）；组织了影响深远的开放源代码峰会，以至于开源软件运动以此命名；创立了Make杂志，从而成为DIY革命的主要先锋；公司一如既往地通过多种形式缔结信息与人的纽带。O'Reilly的会议和峰会集聚了众多超级极客和高瞻远瞩的商业领袖，共同描绘出开创新产业的革命性思想。作为技术人士获取信息的选择，O'Reilly现在还将先锋专家的知识传递给普通的计算机用户。无论是通过书籍出版，在线服务或者面授课程，每一项O'Reilly的产品都反映了公司不可动摇的理念——信息是激发创新的力量。

业界评论

“O'Reilly Radar博客有口皆碑。”

——Wired

“O'Reilly凭借一系列（真希望当初我也想到了）非凡想法建立了数百万美元的业务。”

——Business 2.0

“O'Reilly Conference是聚集关键思想领袖的绝对典范。”

——CRN

“一本O'Reilly的书就代表一个有用、有前途、需要学习的主题。”

——Irish Times

“Tim是位特立独行的商人，他不光放眼于最长远、最广阔的视野并且切实地按照Yogi Berra的建议去做了：‘如果你在路上遇到岔路口，走小路（岔路）。’回顾过去Tim似乎每一次都选择了小路，而且有几次都是一闪即逝的机会，尽管大路也不错。”

——Linux Journal

译者序

Python语言自其诞生之日起，就以其“优雅”、“明确”、“简单”的设计哲学，吸引着来自世界各地的开发人员。经由他们的使用和推介，至今，Python已渗透到计算机的诸多领域。

本书是Python应用的手册指南，它涵盖了Python编程的方方面面，从系统管理到图形界面编程，从文本处理到网络编程，从数据库到语言扩展……这些主题的探讨中，作者提供了大量的实际代码，通过对这些代码的研读，对读者的Python编程水平以及实际项目问题的解决都将不无裨益。

本书主要由邹晓、瞿乔、任发科翻译，参与翻译的其他人员还有张超、栾瑞鹏、张君兰、刘甜、马亮、官毅、潘婧、杨洁、颜秦进。由于译者水平所限，翻译难免有疏漏之处，愿广大读者给予批评指正。

——常新居士

目录

上册

前言	1
----------	---

第一部分 序幕

第1章 先睹为快	19
“Python编程：长话短说”	19
任务	20
第1步：表示记录	20
第2步：持久存储记录	30
第3步：走进OOP	41
第4步：增加控制台交互	51
第5步：增加GUI	54
第6步：增加Web界面	65
演示总结	81

第二部分 系统编程

第2章 系统工具	85
“os.path”的知识	85
系统编程概述	87

介绍sys模块.....	98
介绍os模块.....	101
第3章 脚本运行上下文.....	115
我要增加一个变量.....	115
当前工作路径.....	115
命令行参数.....	118
shell环境变量.....	121
标准流.....	124
第4章 文件和目录工具.....	143
“五个简单步骤擦除你的硬盘!”.....	143
文件工具.....	143
目录工具.....	171
第5章 并行系统工具.....	184
“指挥猴子做事”.....	184
进程分支.....	185
线程.....	192
程序退出.....	218
进程间通信.....	227
multiprocessing模块.....	247
启动程序的其他方法.....	261
一个跨平台的程序启动框架.....	267
其他系统工具.....	271
第6章 完整的系统程序.....	273
“愤怒的Grep命令”.....	273
小游戏：找到最大的Python文件.....	274
分割与合并文件.....	284
生成重定向网页.....	294
一个回归测试脚本.....	298
复制目录树.....	306

比较目录树	309
搜索目录树	320
访问器：目录遍历“++”	330
播放媒体文件	343

第三部分 GUI编程

第7章 图形化用户界面 355

“嗨，我们正瞧着你！”	355
Python GUI 开发之选	357
tkinter概览	362
循序渐进的GUI学习曲线	366
tkinter编码的替代之选	370
添加按钮和回调函数	376
添加用户定义的回调函数	379
添加多个组件	391
用类实现组件的自定义设置	397
用类复用GUI部件	400
本章小结	406
Python/tkinter对Tcl/Tk的转换	408

第8章 tkinter之旅：第一部分 410

“组件、工具和GUI，太棒了！”	410
配置组件外观	411
顶层窗口	414
对话框	420
绑定事件	436
Message和Entry组件	442
Checkbutton、Radiobutton和Scale	450
三种运行GUI代码的方法	463
图像	475
使用PIL查看和处理图像	481

第9章 tkinter之旅：第二部分..... 496

“如今的菜单，到处都是垃圾内容” 496

菜单 496

Listbox和Scrollbar 510

Text 516

Canvas 537

网格 550

时间工具、线程和动画 567

结束组件之旅 580

第10章 GUI编码技巧 581

“搭建一个更好的用户界面” 581

GuiMixin：普通工具的混合类 582

GuiMaker：自动化生成菜单和工具栏 586

ShellGui：命令行工具的GUI 596

GuiStreams：重定向流到组件 606

动态重载回调处理器 610

封装顶层窗口接口 612

GUI、线程和队列 616

添加GUI到非GUI代码的其他方法 627

PyDemos和PyGadgets启动器 641

第11章 完整的GUI程序 651

“Python、开源和Camaro” 651

PyEdit：一个文本编辑器程序/对象 654

PyPhoto：一个图像查看器和调整大小的工具 694

PyView：一个图像和说明的幻灯片显示工具 704

PyDraw：绘制和移动图像 714

PyClock：一个模拟/数字时钟组件 723

PyToe：一个三连棋游戏组件 737

小结 742

下册

第四部分 互联网编程

第12章 网络脚本 745

“接收、登录和退出”	745
Python互联网发展方案	750
探索互联网	754
套接字编程	760
处理多个客户端	775
使套接字看起来像文件和数据流	798
一个简单的Python文件服务器	811

第13章 客户端编程 823

“与我连接！”	823
FTP：通过网络传输文件	824
用ftplib传输文件	824
用ftplib传输目录	843
用ftplib传输目录树	861
处理Internet电子邮件	868
POP：抓取电子邮件	870
SMTP：发送电子邮件	878
email：解析和撰写邮件内容	889
一个基于控制台的电子邮件客户端	914
mailtools工具包	923
NNTP：访问新闻组	957
HTTP：访问网站	959
再次探讨urllib包	962
客户端编程的其他选择	967

第14章 PyMailGUI客户端 969

“卢克，快用源码！”	969
------------------	-----

PyMailGUI的主要改动	975
演示PyMailGUI	984
PyMailGUI的实现	1022
改进建议	1073
第15章 服务器端编程	1082
“哦，我们织了怎样一张复杂的网啊”	1082
什么是服务器端CGI脚本？	1083
运行服务器端示例	1086
沿着CGI学习曲线继续攀登	1091
在CGI脚本里保存状态信息	1128
Hello World选择器	1136
重构代码，使其更容易维护	1145
关于HTML和URL转义的更多讨论	1154
向客户端和服务器传输文件	1161
第16章 PyMailCGI服务器	1179
“停留芝加哥期间的任务”	1179
PyMailCGI网站	1180
根页面	1188
通过SMTP发送邮件	1191
处理收到的邮件	1213
实用工具模块	1222
Web 脚本编程要考虑的平衡	1237
第五部分 工具和技巧	1247
第17章 数据库和持久化	1247
“给我来一份持久化，但不要加pickle”	1247
Python中可选的持久化方案	1247
DBM文件	1249
Pickle对象	1253

Shelve文件	1259
面向对象的数据库ZODB	1268
SQL数据库接口	1271
ORM：对象关系映射	1296
PyForm：一个持久化对象的查看器（扩展阅读）	1297
第18章 数据结构	1299
“玫瑰是红色的，紫罗兰是蓝色的；列表是可变的，而集合是个傻瓜”	1299
堆栈实现	1300
集合的实现	1312
实现内置类型的子类	1322
二叉搜索树	1323
检索图	1328
队列的排列组合	1333
反序和排序	1335
PyTree：一个通用的树对象浏览器	1339
第19章 文本和语言	1342
“快看杰克做功课。功课，杰克，功课”	1342
Python中文本处理的策略	1342
字符串方法工具	1343
正则表达式模式匹配	1352
XML和HTML解析	1366
高级语言工具	1375
定制语言解析器	1377
PyCalc：一个计算器程序/对象	1393
第20章 Python/C集成	1417
“迷失在C中”	1417
在C中扩展Python：概述	1420
一个简单的C扩展模块	1421
SWIG集成代码生成器	1424
封装C环境调用	1429

用SWIG封装C++类	1435
其他扩展工具	1445
在C中嵌入Python：概述	1447
基本嵌入技巧	1450
注册回调处理器对象	1462
在C中使用Python类	1466
其他集成相关话题	1469

第六部分 尾声

第21章 结语：Python与开发周期	1475
“全书至此完结，以下是生活的意义”	1476
“我们编程的方式有些问题”	1476
“Gilligan因素”	1476
做正确的事	1477
走进Python	1478
在沉没的泰坦尼克号上	1483
那么“Python：续集”打算讲点什么？	1486
最后的总结	1486

前言

“现在，来点完全不同的……”

本书探索了如何将Python编程语言应用于常见的应用领域和实际的可扩展性任务。它会告诉你一旦掌握Python的基本原理，就可以利用这门语言来做些什么。

本书假定你对于书中涵盖的各应用领域，例如GUI、互联网、数据库、系统编程等认识尚浅，因此本书将以教程的形式，从头介绍每个主题。在此过程中，它会侧重于常用工具和程序库的介绍，而不是语言的基础知识。本书最终将成为一个参考资源，以供读者深入理解Python在真实世界，以及实际编程工作中所扮演的角色。

作为另一个主题，本书探讨了Python作为软件开发工具角色的方方面面，而正是这一角色使多数人认为Python超越了传统意义上的“脚本”语言。实际上，书中的许多例子专门为此做了规模上的调整。其中，我们会逐步开发出数千行代码的电子邮件客户端。这样的大规模编程一向是颇具挑战性的工作，但你将会发现，当使用Python完成这些程序时也相当快，而且更为简单。

本书第四版已经作了更新，介绍了Python 3.X的语言、程序库和实践。特别地，书中的程序示例使用了Python 3.1，这也是本书写作时最新的Python版本，并且在出版前，其主要程序示例成功通过了Python 3.2 alpha 3版的测试。本版也经过重新组织，不但简化了之前版本的一些内容，而且涉及到新近出现的工具与主题。

由于本版的读者既包括初来乍到的新手，也不乏之前版本的老朋友，因此在深入代码之前，我想借这个前言来阐述一下本书的目的和范围。

关于本书

本书以教程的形式，介绍了Python在常见应用领域和任务中的使用。它将教你如何应用Python进行系统管理、GUI编程和Web开发，并且探索了Python在网络、数据库、前端、脚本层、文字处理等方面的作用。虽然本书从始至终都在使用Python，但本书的重点在于Python在现实世界任务中的应用，而不是语言的基本知识。

本书的生态系统

根据其范围，本书被设计为Python系列书籍两卷中的第二部分，并对第三部分加以补充。最重要的是，本书是核心语言书籍《Python学习手册》在应用编程方面的后续，其主题是阅读本书的正式前提。这三本书的相互关系如下：

- 《Python学习手册》(*Learning Python*)深入地探讨了Python编程的基本知识。它集中研究Python语言的核心，并且其主题是阅读本书的前提条件。
- 《Python编程》(*Programming Python*)，也就是本书，涉及了Python在现实世界编程任务中的应用。本书关注程序库和工具，并认为你已经了解了Python的基础知识。
- 《Python Pocket Reference》为这里没有详尽列出的细节提供了快速参考。它虽然讲得不多，但它可以让你快速查找细节。

在某种意义上，本书对于应用编程就如同《Python学习手册》对于语言核心，也是一个循序渐进的教程，因此本书对读者的背景几乎没有任何要求，所有的主题都将从头介绍。比如，通过研究本书涵盖的Web基础知识，将可以构建简单的网站，并且当需要更进一步时，也可以理解更高级的框架和工具。GUI编程也是同样循序渐进地去讲解，从入门到精通。

另外，本书被设计成由快速参考书籍《Python Pocket Reference》进行补充，《Python Pocket Reference》提供了一些精巧的细节，并且作为查找细末枝节的资源。它仅仅用作参考，因此很大程度上没有例子和叙述，但它为《Python学习手册》的基础知识和《Python编程》的应用知识做了增补。由于本书第四版目前包含了Python 2.X和3.X的工具，因此它也为读者在Python的两个版本的产品过渡提供了资源（稍后将更详细的讨论这一点）^{注1}。

注1：披露：我是这里提到的三本书的作者，这使得我可以随心所欲地控制它们的范围从而避免重叠。但这也意味着，作为作者，我要尽量避免评论其他Python书籍，其中一些非常好，并且涉及我的书中没有探讨的主题。请从网上查找其他的Python资源。这三本书是我13年Python教学的反映，其来源于1995年<这里要插一个色彩斑驳的老照片>所著的、最初的那本《Python编程》。

本书没有提及的部分……

由于我刚刚提到的相关书籍在范围上的划分，本书遵循以下两个明确的约定：

- 本书不涉及Python语言的基础知识。
- 本书不打算成为一本语言参考书。

前一个约定表明，核心语言是《Python学习手册》的专属领域。如果你完全是Python语言的新手，我建议你在研读本书之前先看看那本书，就像本书所要求的那样。当然，书中的例子也会呈现一些语言方面的技巧，而更大的例子将说明众多核心概念是如何汇聚成实际程序的。比如说，面向对象编程（OOP）常常在我们所编写的更大的程序中才会得到最佳的体现。本书认为你已经了解了足够多的Python基础知识，以便能够理解书中的示例代码。我们这里主要关注的是程序库和工具。如果我们用到的基本代码对你来说不甚明朗，请查阅其他资源。

后一个约定反映了人们多年来对本书的一个常见误解（实际上，如果我们在1995年更有先见之明的话，这本书其实最好改成《应用Python》）。我想竭尽所能让大家清楚：这绝对不是参考书，这是一本教程。虽然你可以使用索引和目录来搜索细节，但本书的设计初衷不是这样的。实际上，《Python Pocket Reference》提供了对这类细节的快速参考，一旦你开始撰写那些非同寻常的代码，你就会发现其用处。还有许多其他的参考资源，这包括其他书籍和Python自己的参考手册。这里，我们的目标是一个循序渐进的教程，该教程将教你如何将Python应用于常见任务，但不打算面面俱到地记录所有的细节。

关于第四版

如果这是你第一次看到本书，也许对最新的变化没什么兴趣，尽可以跳过该部分。然而，对于之前版本的读者，本书的第四版主要有三方面的改变：

- 本书已更新到Python 3.X（仅包含）。
- 本书进行了精简以突出重点并为新主题准备空间。
- 本书关注了Python世界出现的新主题和新工具。

在这些变化中，第一条可能是最重要的改变，这个版本使用了Python 3.X，以及对应版本的标准库和其用户的惯用法。为了更全面地解释本版的这些改变，我要补充更多的细节。

本版的具体变化

由于本书之前的版本受众广泛，因此这里要快速描述一下本版中最突出的变化：

精简已有内容来容纳新主题

本书之前的版本已经达到1600页，根本没有空间来容纳新的Python主题（Python 3.X面向Unicode就意味着许多新内容）。幸运的是，Python世界最近的变化使我们这次可以削减一些目前不怎么重要的东西，以便为新内容腾出地方。

当然，这个过程中并没有牺牲内容的深度，本书还是一如既往的丰富殷实。总的来说，这次更新的一个主要目标就是厚度，因此，我下面提及的其他改变和删减在某种程度上都是为了帮助容纳新内容。

只涵盖Python 3.X

本书的例子和叙述所使用的均是Python 3.X版本。因此，除了Python 3.X和2.X共同的部分，本书已不再支持Python 2.X。虽然这个共同部分的交集很大，以至于Python 2.X的读者也可以使用本书，但现在本书仅面向Python 3.X。

这正是本版书的厚度没有较之以前增加的主要原因。由于Python 3.X与2.X并不兼容，而且Python 3.X被认为是Python语言的未来。所以通过将范围限制在Python 3.X上，我们可以避免在两个版本的不同处重复介绍。因此，对于本书这样投注大量内容在更新的例子上，可是又只能罗列一个版本的情况下，这种版本上的限制就尤为重要。

对于仍然介于2.X和3.X的读者，我会在稍后更多地谈及Python 3.X的变化。这里描述的3.X最重大的变化大概是PyEdit和PyMailGUI中新的国际化支持。虽然Python 2.X也有Unicode，但Python 3.X的重新设计迫使这样的系统重新回顾它们之前所使用的ASCII方法。

收录新的库与工具

自上一版图书面世后，出现了很多新的库和工具，并且越来越受欢迎。本书会对这些库和工具进行介绍，包括新标准库工具，例如子进程（第2章和第3章）和多重处理（第5章），以及新的第三方Web框架和ORM数据库工具包。这些新的库与工具涵盖的内容并不广泛（许多流行的第三方扩展是带有版权的复杂系统，最好由专门的书籍对其进行介绍），但在这本书里，我们至少对这些工具进行了总结性介绍。

例如，Python 3.1中的新tkinter.ttk的Tk主题组件集将在本书的第7章中介绍，但只是简单介绍。作为一项规则，本书往往会在不经意的时候顺便提到这些扩展，不仅仅只是展示代码而缺乏足够的解释。

更精简的前言

我删去了所有使用并且运行程序例子的命令。请参考例子中的README文件获取使用细节说明。而且，以前版本中的致谢部分有些冗余，本书不必再复制这些内容，所以这个版本更精简了。书的内容介绍也被删除了，请参见内容表对本书的结构进行概览。