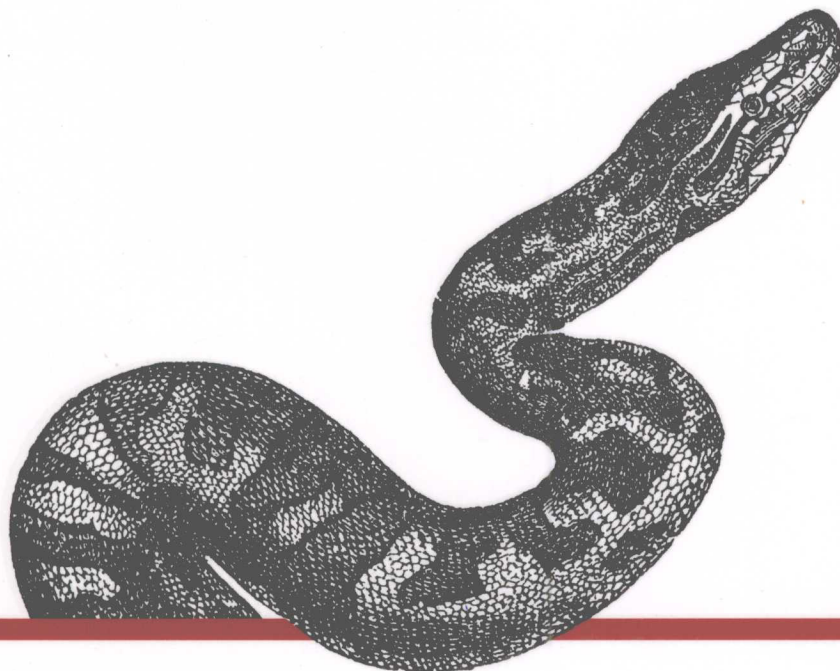


Python in a Nutshell



PYTHON

技术手册

(第2版)

O'REILLY®

[美] Alex Martelli 著

程胜 杨萍 译

 人民邮电出版社
POSTS & TELECOM PRESS

Python 3.x 対応



PYTHON

技術手帳

(第2版)

O'REILLY

1000 以上の Python 3.x
コード例を収録

Python 3.x 対応
第2版

O'REILLY®

PYTHON 技术手册 (第 2 版)

[美] Alex Martelli 著

程胜 杨萍 译

人 民 邮 电 出 版 社

北 京

图书在版编目 (CIP) 数据

Python技术手册 : 第2版 / (美) 马特利
(Martelli, A.) 著 ; 程胜, 杨萍译. — 北京 : 人民邮
电出版社, 2010. 6
ISBN 978-7-115-22583-2

I. ①P… II. ①马… ②程… ③杨… III. ①软件工
具—程序设计—技术手册 IV. ①TP311.56-62

中国版本图书馆CIP数据核字(2010)第056707号

版权声明

Copyright©2006 by O'Reilly Media, Inc.

Simplified Chinese Edition, jointly published by O'Reilly Media, Inc. and Posts & Telecom Press, 2010. Authorized translation of the English edition, 2006 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.

本书中文简体版由 O'Reilly Media, Inc. 授权人民邮电出版社出版。未经出版者书面许可, 对本书的任何部分不得以任何方式复制或抄袭。

版权所有, 侵权必究。

Python 技术手册 (第 2 版)

- ◆ 著 [美] Alex Martelli
- 译 程 胜 杨 萍
- 责任编辑 刘映欣
- ◆ 人民邮电出版社出版发行 北京市崇文区夕照寺街 14 号
邮编 100061 电子函件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
北京鑫正大印刷有限公司印刷
- ◆ 开本: 787×1000 1/16
印张: 39.5
字数: 829 千字 2010 年 6 月第 1 版
印数: 1-3 500 册 2010 年 6 月北京第 1 次印刷

著作权合同登记号 图字: 01-2009-1823 号

ISBN 978-7-115-22583-2

定价: 89.00 元

读者服务热线: (010) 67132705 印装质量热线: (010) 67129223

反盗版热线: (010) 67171154

内 容 提 要

本书是一本全面介绍有关 Python 语言和 Python 程序开发专业知识的参考手册。书中详细介绍了 Python 开发工具的安装和使用、Python 语言的语法结构、Python 内置对象、库和模块以及 Python 与其他语言的扩展和嵌入，并专门介绍了有关 Python 网络和 Web 编程的内容和实例。本书列举了 Python 对象和模块中提供的所有类型、方法和函数，并辅以适当的示例，系统地展示了 Python 包含的功能及其使用方法。

本书的内容面向 Python 语言的各层次用户，可以为不同层次的读者提供相应的信息。针对从其他编程语言转到使用 Python 语言的程序员，本书还重点介绍了 Python 的跨平台功能。

O'Reilly Media, Inc. 介绍

为了满足读者对网络和软件技术知识的迫切需求，世界著名计算机图书出版机构 O'Reilly Media, Inc. 授权人民邮电出版社，翻译出版一批该公司久负盛名的英文经典技术专著。

O'Reilly Media, Inc. 是世界上在 UNIX、X、Internet 和其他开放系统图书领域具有领导地位的出版公司，同时也是联机出版的先锋。

从最畅销的 *The Whole Internet User's Guide & Catalog*（被纽约公共图书馆评为 20 世纪最重要的 50 本书之一）到 GNN（最早的 Internet 门户和商业网站），再到 WebSite（第一个桌面 PC 的 Web 服务器软件），O'Reilly Media, Inc. 一直处于 Internet 发展的最前沿。

许多书店的反馈表明，O'Reilly Media, Inc. 是最稳定的计算机图书出版商——每一本书都一版再版。与大多数计算机图书出版商相比，O'Reilly Media, Inc. 具有深厚的计算机专业背景，这使得 O'Reilly Media, Inc. 形成了一个非常不同于其他出版商的出版方针。O'Reilly Media, Inc. 所有的编辑人员以前都是程序员，或者是顶尖级的技术专家。O'Reilly Media, Inc. 还有许多固定的作者群体——他们本身是相关领域的技术专家、咨询专家，而现在编写著作，O'Reilly Media, Inc. 依靠他们及时地推出图书。因为 O'Reilly Media, Inc. 紧密地与计算机业界联系着，所以 O'Reilly Media, Inc. 知道市场上真正需要什么图书。

前言

Python 编程语言可以很好地协调一些看起来似乎很明显的矛盾：Python 编程语言格式优雅并注重实效、简单而且功能强大、非常高层但是并不妨碍用户对底层的比特 (bit) 和字节 (Byte) 的处理，Python 编程语言适合于编程新手，对 Python 专家也非常适用。

本书的目标读者是已经对 Python 有一些了解的程序员，以及刚开始从其他编程语言转到使用 Python 的有经验的程序员。本书是一本有关 Python 的参考指南，包括 Python 本身、Python 庞大的标准库中最常用的部分，以及一些最流行和最有用的第三方模块和软件包，这些第三方模块和软件包涵盖了广泛的应用程序开发领域，包括 Web 和网络编程、图形用户界面 (GUI)、XML 处理、数据库交互操作以及高速数值计算等。本书将主要关注 Python 的跨平台功能，还将介绍如何扩展 Python 模块，以及如何将 Python 嵌入使用 C 或 Java™ 编写的其他应用程序中的一些基础知识。

本书的组织结构

本书分为以下 5 个部分。

第 1 部分 “Python 入门指南”

第 1 章 “Python 简介”

本章介绍了 Python 语言及其实现的一般特性，并介绍在哪里可以获得有关 Python 的帮助和信息。

第 2 章 “安装”

本章介绍了如何获得 Python 以及在开发者的计算机上安装 Python。

第 3 章 “Python 解释器”

本章介绍了 Python 解释器程序、解释器的命令行选项，以及如何使用 Python 解释器在交互式会话中运行 Python 程序。本章还提到了一些特别适合于编辑 Python 源代码的文本编辑器和可以全面检查 Python 源代码的辅助程序，还介绍了一些已经发展得很完善的集成开发环境，包括 IDLE，IDLE 是标准 Python 发布版本附带的免费集成开发环境。

第 2 部分 “核心 Python 语言和内置对象”

第 4 章 “Python 语言”

本章介绍了 Python 的语法、内置数据类型、表达式、语句，以及如何编写和调用 Python 函数。

第 5 章 “面向对象的 Python”

本章介绍了 Python 中的面向对象编程功能。

第 6 章 “异常”

本章介绍了如何处理 Python 程序中的错误和异常情况。

第 7 章 “模块”

本章介绍了如何使用 Python 软件将代码组合到模块和包中，如何定义和导入模块，以及如何安装以标准 Python 方式打包的第三方 Python 扩展模块。

第 8 章 “核心内置”

本章介绍了 Python 的内置数据类型和内置函数，以及标准 Python 库中最基础的模块（粗略地讲，也就是在其他一些编程语言中被内置到语言本身以提供各种功能的模块）。

第 9 章 “字符串和正则表达式”

本章介绍了 Python 强大的字符串处理功能，包括 Unicode 字符串和正则表达式。

第 3 部分 “Python 库和扩展模块”

第 10 章 “文件和文本操作”

本章介绍了如何使用内置 Python 文件对象、来自 Python 标准库中的一些模块和用于富文本（Rich text）I/O 的平台相关扩展模块对文件和文本进行处理。本章还介绍了有关国际化和本地化的问题，以及如何使用 Python 定义文字模式的交互式命令会话这样的特殊任务。

第 11 章 “持久化和数据库”

本章介绍了 Python 的序列化机制和持久化机制，以及 Python 与 DBM 数据库、Berkeley 数据库和关系数据库（基于 SQL）之间的接口。

第 12 章 “时间操作”

本章介绍了在 Python 中如何使用标准库和常用扩展模块来处理时间和日期。

第 13 章 “控制执行”

本章介绍了如何在 Python 中完成高级执行控制，包括对动态生成的代码的执行控制和对垃圾收集操作的控制。本章还介绍了 Python 的一些内部类型，以及与注册“清理”

函数有关的特殊问题，“清理”函数是在程序终止时执行的函数。

第 14 章 “线程和进程”

本章介绍了 Python 的并发执行功能，包括如何在一个进程中运行多个线程，以及如何在单台计算机上运行多个进程。本章还介绍了如何访问进程的环境，以及如何通过内存映射机制访问文件。

第 15 章 “数值处理”

本章介绍了 Python 标准库模块和第三方扩展包中的数值计算功能。本章特别介绍了如何使用十进制浮点型数字，而不是默认的二进制浮点型数字。还介绍了如何获得和使用伪随机数和真正的随机数。

第 16 章 “数组处理”

本章介绍了用于执行数组处理的内置对象和扩展包，主要针对传统的 Numeric 第三方扩展，并提到了其他一些最近开发的数组处理解决方案。

第 17 章 “Tkinter GUI”

本章介绍了如何使用标准 Python 发布版本附带的 Tkinter 包开发 Python 图形用户界面，并简要提到了其他一些可选的 Python GUI 框架。

第 18 章 “测试、调试和最优化”

本章介绍了如何使用 Python 工具和方法来确保开发者的程序的正确性（也就是说，程序完成了开发者想要其实现的功能），查找并纠正程序中的错误，以及检查并增强程序的性能。本章还介绍了“警告”的概念和用来处理“警告”的 Python 库模块。

第 4 部分 “网络和 Web 编程”

第 19 章 “客户端网络协议模块”

本章介绍了 Python 标准库中用来帮助程序员编写网络客户端程序的一些模块，并专门通过从客户端处理各种网络协议和处理 URL 来介绍这些模块。

第 20 章 “套接字和服务器端网络协议模块”

本章介绍了 Python 与底层网络机制（套接字）的接口，用来帮助程序员编写网络服务器程序的标准 Python 库模块，以及如何使用标准模块和强大的 Twisted 扩展模块进行异步（事件驱动）网络编程。

第 21 章 “CGI 脚本和其他解决方案”

本章介绍了 CGI 编程的基础知识，如何使用标准 Python 库模块在 Python 中实现 CGI 编程，以及如何使用“cookie”处理 HTTP 服务器端编程中的会话状态。本章还提到了许多 CGI 编程的可选方案，这些可选方案可以使用流行的 Python 扩展模块来实现服务

器端 Web 编程。

第 22 章 “MIME 和网络编码方式”

本章介绍了如何在 Python 中处理电子邮件，以及如何处理其他网络数据结构和编码的文档。

第 23 章 “结构化文本：HTML”

本章介绍了可以用来处理和生成 HTML 文档的 Python 库模块。

第 24 章 “结构化文本：XML”

本章介绍了可以用来处理、修改和生成 XML 文档的 Python 库模块和流行的扩展模块。

第 5 部分 “扩展和嵌入”

第 25 章 “扩展和嵌入经典 Python”

本章介绍了如何使用 C 和其他经典编译语言来编写 Python 扩展模块、如何在用其他语言编写的应用程序中嵌入 Python，扩展 Python 以及访问现有 C、C++ 和 Fortran 库的其他可选方法。

第 26 章 “扩展和嵌入 Jython”

本章介绍了如何使用 Python 的 Jython 实现中的 Java 类，以及如何在 Java 编写的应用程序中嵌入 Jython。

第 27 章 “发布扩展和程序”

本章介绍了几个可以帮助开发者对 Python 扩展、模块和应用程序进行打包并进行发布的工具。

本书使用的惯例

本书使用了以下这些惯例。

引用惯例

在函数/方法引用输入参数中，只要有可能，每个可选参数都将被显示为一个使用 Python 语法 `name=value` 表示的默认值。内置函数不需要接受命名参数，因此参数的名称并不是很重要。有些可选参数可以按照这些参数的出现与否得到最好的诠释，而不是通过其默认值。在这些情况下，本书通过在方括号 ([]) 中包含这个参数以表示该参数是可选的。在有多于一个的可选参数时，将形成嵌套的方括号。

如何使用本书中的代码示例

本书的目的是帮助读者完成自己的 Python 开发工作。一般情况下，读者有可能会在自

己的程序和文档中使用本书中的代码。对此，不需要联系出版社以获得许可，除非打算利用本书中的主要代码进行二次商业开发。例如，使用本书中的一部分代码段编写程序是不要求获得许可的。但是，如果要销售或发布一张包含来自 O'Reilly 图书中的示例的 CD-ROM，则要求获得许可。如果要将本书中的大量示例代码复制到自己开发的产品资料中，则必须获得许可。

我们感谢，但并不要求读者在使用本书的代码示例时注明引用源。引用源通常包括标题、作者、出版社和 ISBN。例如，*Python in a Nutshell, Second Edition*, by Alex Martelli. Copyright 2006 O'Reilly Media, Inc., 0-596-10046-9。

如何联系我们

作者已经尽自己最大努力测试并验证了本书中的所有内容，但是读者可能还是会发现有些功能被更改了（甚至会在书中找到一些错误！）。请将发现的任何错误告知出版社，还可以告诉出版社对本书将来版本的建议，通信地址是：

美国：

O'Reilly Media, Inc.
1005 Gravenstein Highway North
Sebastopol, CA 95472

中国：

100080 北京市西城区西直门成铭大厦 C 座 807 室
奥莱利技术咨询（北京）有限公司

出版社为本书提供了一个网页，网页上列出了本书的勘误表、示例和所有其他信息。可以通过下面的链接访问这个网页：

<http://www.oreilly.com/catalog/pythonian2>

要想询问技术问题或者对本书进行评论，请发送电子邮件到：

bookquestions@oreilly.com
info@mail.oreilly.com.cn

要想获得有关 O'Reilly 的图书、会议、资源中心和 O'Reilly Network 的更多信息，请访问 O'Reilly 的网站：

<http://www.oreilly.com>

致谢

衷心感谢曾经在这本书上帮助过我的所有人，包括本书第 1 版和现在的第 2 版。许多 Python 的初学者、专业人士和专家曾经阅读过本书的一部分初稿，并提供了一些反馈

意见，这些意见可以帮助我让这本书更加清楚、精确和正确，并具备更好的可读性。除了要感谢所有这些提供了大量高质量的反馈意见和其他帮助的人之外，我还必须单独感谢我在 Google 公司的同事们，特别是 Neal Norwitz 和 Mohsin Ahmed。

本书的第 1 版得到了来自于 Python 各个领域的专家（Aahz 是有关线程的专家、Itamar Shtull-Trauring 是有关 Twisted 的专家、Mike Orr 是有关 Cheetah 的专家、Eric Jones 和 Paul Dubois 是有关 Numeric 的专家，而 Tim Peters 是有关线程、测试和最优化的专家）、一个非常优秀的技术审校小组（包括 Fred Drake、Magnus Lie Hetland、Steve Holden 和 Sue Giller）以及本书的编辑 Paula Ferguson 不可或缺的帮助。本书的第 2 版也从编辑 Jonathan Gennick 和 Mary O'Brien、技术审校 Ryan Alexander、Jeffery Collins 和 Mary Gardiner 的辛勤工作中受益匪浅。我要对 O'Reilly 工具小组的优秀员工们致以特殊的感谢，他们（不管是直接和亲自的帮助，还是通过他们开发出来的那些非常有帮助的工具）帮助我解决了几个非常困难的技术问题。

一直以来，我非常思念我的家庭：我的孩子 Flavia 和 Lucio，我的姐姐 Elisabetta 和我的父亲 Lanfranco，他们已经回到了我的祖国意大利，而我在 Google 的工作让我来到了美国的加利福尼亚。

但是，在我内心深处，无比感谢，甚至不止是感谢的一个极其了不起的人，就是我的妻子 Anna Martelli Ravenscroft，我的 *Python Cookbook* 一书第 2 版的合作者，Python 软件基金会的会员之一，以及所有作者都可能会梦寐以求的最严格、最出色的技术审校。除了对本书数不清的直接贡献之外，在过去的一年中，Anna 还非常神奇地设法为我创造了足够多的和睦、安静和自由时间（尽管我同时还承担了 Google 公司的 Uber Tech Lead 这个非常奇妙和具有挑战性的工作），是她使得这本书成为可能。实际上，这也是一本属于她的书，她至少付出了和我同样多的努力。

目录

第 1 部分 Python 入门指南

第 1 章 Python 简介	2
1.1 Python 语言	2
1.2 Python 标准库和扩展模块	4
1.3 Python 的实现	4
1.4 Python 的开发和版本	7
1.5 Python 的资源	9
第 2 章 安装	13
2.1 从源代码安装 Python	13
2.2 从二进制文件安装 Python	18
2.3 安装 Jython	19
2.4 安装 IronPython	20
第 3 章 Python 解释器	21
3.1 Python 程序	21
3.2 Python 开发环境	25
3.3 运行 Python 程序	28
3.4 Jython 解释器	29
3.5 IronPython 解释器	29

第 2 部分 核心 Python 语言和内置对象

第 4 章 Python 语言	32
4.1 词法结构	32
4.2 数据类型	37
4.3 变量和其他引用	44

4.4	表达式和运算符	48
4.5	数值运算	50
4.6	序列运算	52
4.7	集合运算	57
4.8	字典运算	59
4.9	print 语句	61
4.10	控制流语句	61
4.11	函数	69
第 5 章	面向对象的 Python	81
5.1	类和实例	81
5.2	特殊方法	104
5.3	装饰器	114
5.4	元类	115
第 6 章	异常	119
6.1	try 语句	119
6.2	异常传播	125
6.3	raise 语句	126
6.4	异常对象	127
6.5	自定义异常类	131
6.6	错误检查策略	133
第 7 章	模块	139
7.1	模块对象	139
7.2	模块加载	144
7.3	包	149
7.4	发布工具 (distutils)	150
第 8 章	核心内置	153
8.1	内置类型	153
8.2	内置函数	157
8.3	sys 模块	164
8.4	copy 模块	167
8.5	collections 模块	168
8.6	functional 模块	171
8.7	bisect 模块	171
8.8	heapq 模块	172

8.9	UserDict 模块	173
8.10	optparse 模块	173
8.11	itertools 模块	177
第 9 章	字符串和正则表达式	180
9.1	字符串对象的方法	180
9.2	string 模块	183
9.3	字符串格式化	186
9.4	pprint 模块	189
9.5	repr 模块	190
9.6	Unicode	190
9.7	正则表达式和 re 模块	193
第 3 部分	Python 库和扩展模块	
第 10 章	文件和文本操作	206
10.1	其他与处理文件有关的章节	206
10.2	本章的组织结构	206
10.3	文件对象	207
10.4	文件 I/O 的辅助模块	214
10.5	StringIO 和 cStringIO 模块	218
10.6	压缩文件	219
10.7	os 模块	226
10.8	文件系统操作	227
10.9	文本输入和输出	239
10.10	富文本 I/O	242
10.11	交互式命令会话	247
10.12	国际化	250
第 11 章	持久化和数据库	258
11.1	序列化	258
11.2	DBM 模块	266
11.3	Berkeley DB 接口	269
11.4	Python 数据库 API (DBAPI) 2.0	272
第 12 章	时间操作	281
12.1	time 模块	281

12.2	datetime 模块	285
12.3	pytz 模块	289
12.4	dateutil 模块	290
12.5	sched 模块	292
12.6	calendar 模块	293
12.7	mx.DateTime 模块	294
第 13 章	控制执行	303
13.1	动态执行和 exec 语句	303
13.2	内部类型	307
13.3	垃圾收集	308
13.4	终止函数	312
13.5	站点和用户自定义	313
第 14 章	线程和进程	315
14.1	Python 中的线程	315
14.2	thread 模块	316
14.3	Queue 模块	317
14.4	threading 模块	319
14.5	线程程序架构	324
14.6	进程环境	327
14.7	运行其他程序	328
14.8	mmap 模块	333
第 15 章	数值处理	337
15.1	math 和 cmath 模块	337
15.2	operator 模块	339
15.3	随机数和伪随机数	341
15.4	decimal 模块	343
15.5	gmpy 模块	344
第 16 章	数组处理	345
16.1	array 模块	345
16.2	数值数组计算的扩展包	347
16.3	Numeric 包	348
16.4	数组对象	348
16.5	通用函数 (ufuncs)	366
16.6	辅助 Numeric 模块	371

第 17 章 Tkinter GUI	373
17.1 Tkinter 基础知识	374
17.2 部件基础知识	377
17.3 常用的简单部件	383
17.4 容器部件	388
17.5 菜单部件	390
17.6 文本部件	393
17.7 画布部件	401
17.8 布局管理	407
17.9 Tkinter 事件	410
第 18 章 测试、调试和最优化	415
18.1 测试	415
18.2 调试	425
18.3 warnings 模块	433
18.4 最优化	436
 第 4 部分 网络和Web 编程	
第 19 章 客户端网络协议模块	452
19.1 URL 访问	452
19.2 Email 协议	460
19.3 HTTP 和 FTP	461
19.4 网络新闻	465
19.5 Telnet	468
19.6 分布式计算	469
19.7 其他协议	471
第 20 章 套接字和服务端网络协议模块	472
20.1 socket 模块	472
20.2 SocketServer 模块	479
20.3 事件驱动套接字程序	483
第 21 章 CGI 脚本和其他解决方案	494
21.1 Python 中的 CGI	495
21.2 Cookie	501
21.3 其他服务器端方案	505