

开发人员专业技术丛书

*Python Web
Development with Django*

Django Web开发指南

Jeff Forcier

(美) Paul Bissex 著
Wesley Chun

徐旭铭 等译



机械工业出版社
China Machine Press

本书讲述如何用Python框架Django构建出强大的Web解决方案，本书讲解了使用新的Django 1.0版的各种主要特性所需要的技术、工具以及概念。

全书分为12章和6个附录，内容包括，Django Python实战，Django速成：构建一个Blog，起始，定义和使用模型，URL、HTTP机制和视图，模板和表单处理，Photo Gallery，内容管理系统，Liveblog，Pastebin，高级Django编程，高级Django部署。附录内容包括命令行基础，安装运行Django，实用Django开发工具，发现、评估、使用Django应用程序，在Google App Engine上使用Django，参与Django项目。

本书适用于Python框架Django初学者，Django Web开发技术人员。

Simplified Chinese edition copyright © 2009 by Pearson Education Asia Limited and China Machine Press.

Original English language title: *Python Web Development with Django* (ISBN 978-0-13-235613-8) by Jeff Forcier, Paul Bissex, Wesley Chun, Copyright © 2009.

All rights reserved.

Published by arrangement with the original publisher, Pearson Education, Inc., publishing as Addison-Wesley.

本书封面贴有Pearson Education（培生教育出版集团）激光防伪标签，无标签者不得销售。

版权所有，侵权必究。

本书法律顾问 北京市展达律师事务所

本书版权登记号：图字：01-2009-2370

图书在版编目（CIP）数据

Django Web开发指南 / (美) 杰佛 (Jeff, F.), (美) 鲍尔 (Paul, B.), (美) 陈仲才 (Wesley, C.) 著；徐旭铭等译. —北京：机械工业出版社，2009.5

(开发人员专业技术丛书)

书名原文：Python Web Development with Django

ISBN 978-7-111-27028-7

I . D… II . ① 杰… ② 鲍… ③ 陈… ④ 徐… III . 软件工具—程序设计—指南
IV . TP311.56-62

中国版本图书馆CIP数据核字（2009）第065987号

机械工业出版社（北京市西城区百万庄大街22号 邮政编码 100037）

责任编辑：李东震

北京京北印刷有限公司印刷

2009年5月第1版第1次印刷

186mm×240mm · 18.25印张

标准书号：ISBN 978-7-111-27028-7

定价：49.00元

凡购本书，如有倒页、脱页、缺页，由本社发行部调换

本社购书热线：(010) 68326294

计算机行业真是一个很有意思的行业，它每天都发明无数新名词却又拒绝它们进入生产生活领域。一项技术往往需要十数年的成长才会被业界慢慢接受，而其中的大部分还未来不及成熟就已在沙滩上了。Java当初依赖于快十年才火起来，C++即使借了C的光也是多年媳妇熬成婆。Python比Java发明的更早，早期和Perl抢饭吃，虽有Zope/Pione这样出色的作品，依旧难逃出局的命运。随着Web 2.0的改革号角吹响，开发人员们开始意识到，轻型的框架才是可持续发展的硬道理。一时间，小到只有一个文件的web.py，大到像Quixote等能支撑豆瓣(douban.com)这样大型应用的平台百花齐放，连Python的创始人Guido van Rossum都忍不住跑来凑热闹，扬言要挑一个高手的功用。那么最后到底是谁入了Guido的法眼呢？那就更要隆重推出本书的主角——Django了！能被Guido看中那就说明它靠谱进了Google，这不，Google的新概念云计算的产品之一——Google App Engine (GAE) 已率先支持了一个尊贵过的Django框架，本书会向你介绍如何在GAE上运行Django的程序。

相比其他Python Web框架，Django更为一体化，它安装简单且灵活多变，这很符合Python 开箱即用的特点。选择Django，你无需安装其他组件就能编写序运行了，同时它的灵活性体现出来了。开发者可以在掌握了这些基础后，把各个组件换成自己顺手的工具，这对快速开发来说真的很Web 2.0来说至关重要。如果你还在Web开发的入门徘徊不决的话，不妨来试试 Django，或者你可以访问CPUG的wiki来看看Python主流框架从而加以比较。

http://wiki.woodpecker.org.cn/moin/PyWebFrameworks
http://wiki.woodpecker.org.cn/moin/PyWebAppFrameworks
http://wiki.woodpecker.org.cn/moin/PyWebFrameVs
最后，感谢机械工业出版社华章公司的陈冀康老师和本书的编辑，没有你们就没有这本书中文版的出版。

解卦序

关于本书

欢迎来到DjangоЯ的世界，很高兴能和你一起进行这趟旅程。你会发现有了这个强大的Web框架，做每件事都变得便捷起来——从设计开发新应用到不用大力调弄地修改代码就能为现有的代码提供新的特性和功能。

欢迎使用Django

市面上已经有了些讲解Django的书籍，但是本书的特别之处在于它着重介绍的三个方面：无论你的背景是什么都能读懂它。同时，你还能完整地了解这个框架以及它的能力。Django基础、各种示例以及Django的进阶内容。我们希望写出一本关于这个主题最完整的教程，图P1根据你对Python和Django的理解程度给出了不同的起点。当然，我们的建议是最重要的一章，Django Python 实战

章节指引

第一部分向Django以及Python的新用户介绍了基本的内容，不过即使是经验的读者，我们也会推荐您读一下第3章，“起始”。第一章，Django Python 实战

这一章为那些不了解Python的读者做了一下介绍。本章不仅展示了基本语法，还进一步深入介绍了Python内存模型、数据类型，特别是在Django里大量使用的结构。

第二章，Django展示了这个框架的强大能力。

这一章是为那些希望通过Python介绍，直接在15~20分钟完成一个Django应用的读者准备的。本章完美地展示了这个框架的强大能力。

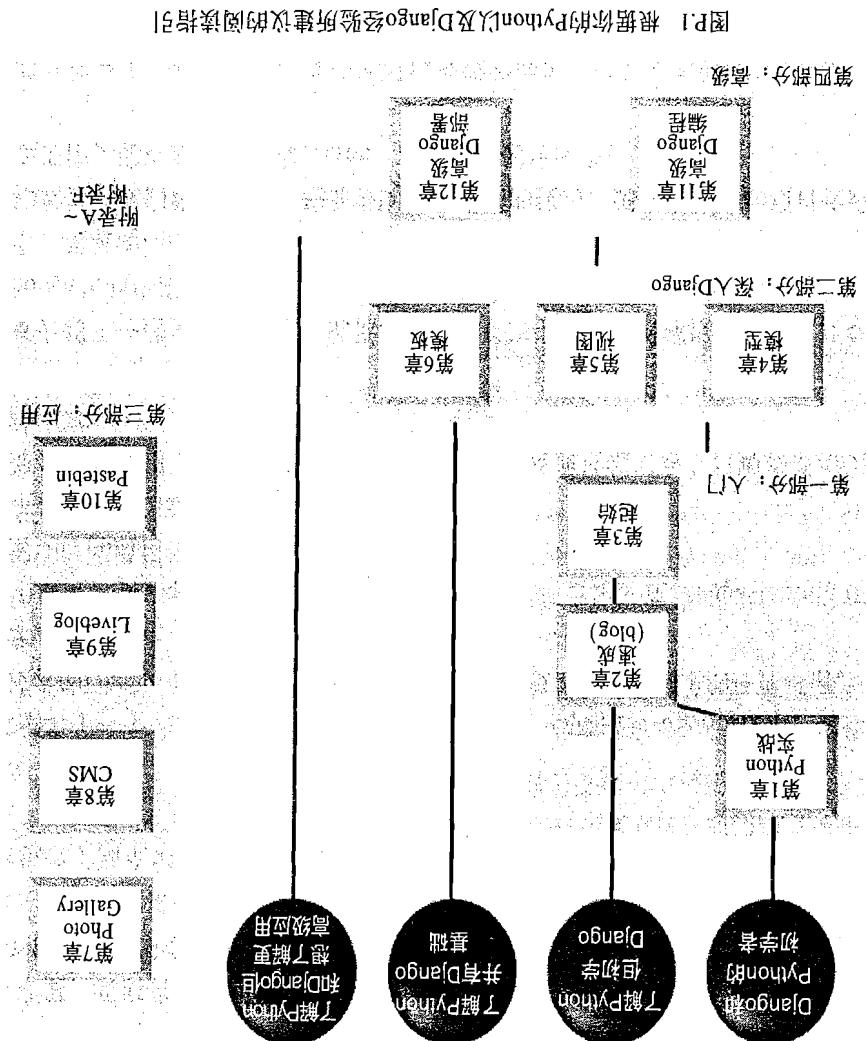
对那些更没耐性的读者，这一章介绍了开发Web应用基础的方法方面（对新手和老鸟都有助于其他Web应用框架的魅力所在）。

第二部分：深入Django

第二部分覆盖了框架所有组成部分。为第三部分中的例子打下了基础。

前言

- 第4章，定义和使用模型
- 在第4章里，我们将学习如何定义和使用数据模型（model），以及Django里对象关系映射（ORM）的基础，包括从简单的值到复杂的对象的关系。
- 第5章，URL、HTTP机制和视图
- 这一章详细介绍了Django处理URL和HTTP协议的方法，这包括中间件，以及如何使用Django简便快捷的通用视图（view）等。另外还介绍了怎样自定义或者重写每一个视图。
- 第6章，模板和表单处理
- 第6章介绍了框架最后一个主要的部分，Django的模板语言和它处理表单的机制。这包含了向用户显示数据以及从用户那里获取数据。
- 第7部分：Django应用实例
- 在第7部分里，我们将创建4个不同的应用，每一个都展示了Django开发中不同的部分和组件。它们将介绍一些新构想，并且扩展在第一部分和第二部分里讲解的概念。



图P.1 根据你的Python以及Django经验所建议的阅读指引

第7章，Photo Gallery

在第7章里，我们将学习如何在你的URL结构上应用“不要重复自己”(DRY)的原则，以及创建一个包含图片预览的简单的照片浏览应用程序。

第8章，内容管理系统

第8章包含了两个创建CMS类系统的方法，还介绍了一些用户贡献的第三方Django应用的使用方法。

第9章，Liveblog

第9章编写了一个“liveblog”——这个站点使用了一些高级JavaScript技术。它作为一个应用了Ajax技术的Django项目，展示了使用任何Ajax工具包都是非常简单的一件事情。

第10章，Pastebin

在第10章里，我们将通过学习创建一个几乎没有任何自定义逻辑的pastebin（在线着色工具）来了解Django通用视图的力量。

第四部分：高级Django技术和特性

第四部分主要介绍了自定义Django的admin应用以及通过编写命令行脚本来和Django应用交互等高阶内容。

第11章，高级Django编程

第11章介绍了一些关于改进代码的话题，比如RSS生成、扩展模板语言，以及如何更好地使用Django的admin应用。

第12章，高级Django部署

在第12章里，我们将学习到一些部署Django应用的技巧，怎样在Django项目代码之外和应用程序一起工作，比如命令行脚本、cron、测试和数据导入等。

附录

附录部分集合了一些和本书有关但不足以成章的内容。学习基本的UNIX命令、Django安装和部署策略、开发工具等。

附录A，命令行基础

附录A介绍了基本的UNIX命令。如果你从来没接触过的话，这个附录将对你非常有用。

附录B，安装运行Django

在附录B里，我们将学习安装所有运行Django所需的必要组件，包括各种可能的数据库和Web服务器，以及一些特定的部署策略的小技巧。

附录C，实用Django开发工具

附录C简要介绍了一些你可能熟悉也可能不熟悉的开发工具，比如代码控制、编辑器等。

附录D，发现、评估、使用Django应用程序

出色的程序员写代码，但是伟大的程序员重用代码！在附录D中，我们来分享一些关于到哪里以及如何寻找可重用的Django应用的经验。

附录E，在Google App Engine上使用Django

附录E从一个独特的视角出发，考察了Google新的App Engine如何使用Django，以及学习

了大量的辅助资料，本书也在很多地方有所提及。[我们的网站](http://withdjangocom)<http://withdjangocom> 包含了大量辅助的资料，本书也在很多地方有所提及。

本书资源

在撰写这本书和示例应用期间，我们用到了所有三个主要的平台——Mac OS X、Linux 和 Windows。另外，我们还测试了所有主要的浏览器（虽然可能没有在书里的图中显示出来），包括 Firefox、Safari、Opera 和 Internet Explorer。

在本书中，我们使用粗体来介绍新的和重要的术语，用斜体来表示强调，<http://links/表记> URL，用等宽字体来描述 Python 变量名和命令行命令等。如果有多行的代码和命令，則把它们包含在等宽字体的块里，像这样：

```
>>> print "This is Python!"  
This is Python!
```

编写体例

附录 F 对比了如何向 Django 项目贡献代码并且成为社区的一份子。

附录 E，[参见与 Django 项目](#) 如何把你的 Django 应用放到 App Engine 框架上去运行。

致謝

虽然我的名字可能是排在作者列表里的头一个，但是如果说没有另外两位作者的努力和贡献，这本书是不可能完成的。Paul和Wesley是两位出色的绅士学者，能和他们共事感觉很棒。说到绅士学者，Django的核心团队也一样想得起这个名字。最开始的四大金刚——Adrian Holovaty、Jacob Kaplan-Moss、Simon Wilson和Wilson Miner——为Django奠定了（而且继续奠定了）一份坚实的基础，在这之上，Malcolm Tredinnick、George Bauer、Luke Plant、Russell Keith-Magee和Robert Wittemans进一步扩展了这个基础。要知道我不是一个容易被感动的人，但是这里的每一位都给了我很大的灵感。

Jeff Forcier
New York, NY
2008年8月

感谢所有围绕在 Django、Python 和其他 Web 应用开发的开源漏洞赏金社区。成百上千热心的开发者和维护者们的辛勤努力让全世界都能免费用上强大的软件。我的合作者们让我受益匪浅，他们不仅给这项任务带来了必要的知识和技能，还为之做出了很大的贡献。尽管我们分散在美洲的不同角落，我还是有幸见到了 Jeff 和 Wes。感谢西马摩咨询塞州开发组，感谢和你们之间许多非常有趣且 geeky 的讨论，感谢你们对本书的极大热情。

感谢 Hallmark Institute of Photography 的主席 George J. Rosa III，给了我极大的信任和支持，让我选择最好的工具（当然，包括 Djangogeo）来最好地完成任务。

在 2008 年夏天经历了一场严重的车祸后，我得到了很多来自家庭、朋友和社区的关心和支持。每一份祝福、每一张卡片、每一分善款和每一顿饭都是很重要的。非常感谢你们。

最后，感谢我亲爱的妻子 Kathleen，感谢你的支持、照顾和爱。

2008年9月
Northampton, MA
Paul Bissell

编写我的第二本书实在是一个非常棒的体验。首先我想要向我的两位出色的合作者致敬，十分高兴能和他们一起工作。他们有汲取他人的Python技巧和介绍Django的经验的能力。我很高兴能合作完成这本出色的Django教程，并且期待能和他们有机会在写作或是教学上再次合作。能编写这样一本完全由开源项目组成的书是非常令人欣慰的，开发人员每天都在使用相同的工具开发改变社会的软件产品。

我要特别感谢Debra Williams Cauley从我第一天加入这个项目开始就帮助我们管理整个进程。虽然在这过程中我们发生了很大的人事变化，但是她一直都让我们专注于写作上面。就像她也赞同的一样，我们的理念是要为社区写出一本“合适的书”，而不仅仅是一本迎合市场需要的Django教程。感谢我们所有技术评论，Michael Thurston（策划编辑），Joe Blaylock和Antonio Cangiano，以及所有在这本书草稿阶段就在Rough Cuts上留下反馈的人，因为你们，这本书才会更加完美。我还要感谢Matt Brown，Django Helper for Google App Engine的首席维护，在审阅附录E中给予的协助，以及Eric Walstad和Eric Evenson最后的终审和评注。

最后，如果没有我们家里团结一致的支持，我是不可能完成这本书的。

Wesley Chun
Silicon Valley, CA
2008年8月

一句话，你想要的就是Python和Django。我们编写这本书的目的就是要帮助你尽量简化快流物的工作，而其组件又是松散耦合的，可以让你在需要的时候随时替换。

你需要一个背后有依赖的、助人为乐的用户和开发者区的框架。这个框架能作为一个整体生成CSV文件、生成饼状图、科学计算或者是图像处理等。

你需要用一门真正的编程语言来编写代码，它必须是干净强大、成熟的语言，需要有大量的文档支持。你希望它有一个漂亮的标库以及大量的高品质的第三方库可以满足各种需要，比如生成CSV文件、生成饼状图、科学计算或者是图像处理等。

你希望通过自己的测试面又不要你自己来维护的框架就好了。

你拿着这本书的原因就是希望你能找到更好的做法。要是有一个强大灵活、实现优雅、经历

更好的做法

务？你的测试覆盖率有多少？

那次升职操作系统、Web服务器、编程语言的时候会不会把它弄挂了？能不能在未来引入修改时不需要你敲动鼠标？它能不能支撑复杂的功能如会话管理、本地化，又或者数据仓库事不过，只要你（或者你的团队、公司或者客户）还在使用它，你就能保证它能正常工作。

恭喜！你已经完成了一个Web框架。

的代码库越来越大、越来越复杂，但是它也变得非常强大。

对绝大多数项目来说你不再需要直接调用框架代码，只需要改动一下配置文件就可以了。虽然你有了教训以后，你回过头来把基础设施和最好的add-on从各个项目里拿出重新组合在一起。这绝对是没有任何办法维护的。

但是，事情可没那么简单。如果客户端需要的特性不在你的库里怎么办，没关系，加快推进就工作时间。

一套自己的库来提供这些功能，或者说，从你最新的、最伟大的“创造”中把这些库提取出来。然后，如果你要开始一个新项目的话，你第一件要做的事情就是安装你的库。这能大大节约你的时间。

Web框架的前世今生

验证码是多么痛苦的一件事情。你得创建数据库结构，把数据导出导进数据库，处理URL，验证用户名输入，提供编辑工具，还得关心安全性可用性……

只要你知道一点构建动态网站是怎么回事的话，那么就一定能体会到不断重复地发明某些特性的痛苦，可能你每次都需要重新实现这些特性实在是太浪费生命了。所以，你决定要重新开发

如果你是一名Web工程师，那么Django有可能彻底改变你的生活。至少它改变了我们。

速地在实际环境下学习使用Django。

冲出堪萨斯，走向世界

Django原本是Adrian Holovaty和Simon Willison为堪萨斯州劳伦斯的一家家族媒体World Online编写的。当时是需要它能够迅速开发出数据库驱动的新闻系统。

在Web领域里证明了自己的实力后，Django在2005年7月作为一个开源项目公开发布。搞笑的是，虽然这个时候很多人都觉得Python已经有太多的Web框架了，Django还是迅速获得了大量追随者。今天，Django已经不仅仅是Python框架里的领头羊了，就是与其他所有Web框架里相比也毫不逊色。

当然，World Online还在大量使用Django，它的一些核心工程师还在那里工作。但是自从Django项目开源以来，全世界的公司和组织都已经把它用在了无数大大小小的项目里了。这里有一份不完整的列表：

- The Washington Post
- The Lawrence Journal-World
- Google
- EveryBlock
- Newsvine
- Curse Gaming
- Tabblo
- Pownce

虽然这里还有成千上百的Django网站没有列出，但是Django的传播和成长已经势不可挡，未来会有无数受欢迎的站点采用Django开发。我们希望你也是其中之一。

用Python和Django更好地进行Web开发

Web开发通常是一件很烦琐的工作。你必须要从测试起就面对浏览器的不兼容性、疯狂的爬虫、宽带和服务器的限制以及总体框架等一系列挑战。

虽然我们相信本书将Django的基础讲解得非常完备了，但是我们还要讨论很多其他难点——你80%的时间都会花在这些20%的工作上。我们和许多Django工程师一起工作讨论，听取他们的意见和建议并且帮助他们解决问题。在编写本书的过程中，我们都提醒自己注意这些问题和难点。

要是我们觉得Django和Python不好的话，我们也不会费这么大劲来为它们写一本书。但是当遇到有限制或者有陷阱的时候，我们还是会如实相告的。我们的目标是帮助你完成任务。

序言	1
前言	1
致谢	1
引言	1
第二部分 深入Django	1
第1章 Django Python实践	1
1.1 Python技术就是Django技术	1
1.2 入门：Python交互解释器	2
1.3 Python基础	3
1.4 Python标准类型	5
1.5 读写控制	19
1.6 异常处理	21
1.7 文件	23
1.8 国数	24
1.9 面向对象编程	33
1.10 正则表达式	35
1.11 常见错误	36
1.12 代码风格	41
1.13 总结	43
1.14 第2章 Django速成：构建一个Blog	44
第2章 编写Django应用	46
2.1 创建项目	44
2.2 运行开发服务器	46
2.3 创建Blog应用	47
2.4 设计你的Model	48
2.5 设置数据库	48
2.6 配置自动admin应用	51
2.7 使用admin	52
2.8 建立ImageField	55
2.9 最后的打包	57
第3章 起始	61
3.1 动态网站基础	61
3.2 逻辑模型、视图和模板	63
3.3 Django架构总览	64
3.4 Django的核心理念	66
3.5 总结	68
第4章 定义和使用模型	69
4.1 定义模型	69
4.2 使用模型	80
4.3 总结	91
第5章 URL、HTTP机制和视图	92
5.1 URL	92
5.2 HTTP建模：请求、响应和中间件	96
5.3 视图与逻辑	100
5.4 总结	105
第6章 模板和表单处理	106
6.1 模板	106
6.2 表单	112
6.3 总结	123
第7章 Photo Gallery	125
7.1 模型	126
7.2 集成文件上传	127
7.3 安装PIL	128
7.4 处理ImageField	128
7.5 构建自定义File字段	130
第三部分 Django应用实例	130

目 录

译者序
前言
致谢
引言

第一部分 入门

第1章 Django Python实践	1
1.1 Python技术就是Django技术	1
1.2 入门：Python交互解释器	2
1.3 Python基础	3
1.4 Python标准类型	5
1.5 读写控制	19
1.6 异常处理	21
1.7 文件	23
1.8 国数	24
1.9 面向对象编程	33
1.10 正则表达式	35
1.11 常见错误	36
1.12 代码风格	41
1.13 总结	43
1.14 第2章 Django速成：构建一个Blog	44
第2章 编写Django应用	46
2.1 创建项目	44
2.2 运行开发服务器	46
2.3 创建Blog应用	47
2.4 设计你的Model	48
2.5 设置数据库	48
2.6 配置自动admin应用	51
2.7 使用admin	52
2.8 建立ImageField	55
2.9 最后的打包	57
第3章 起始	61
3.1 动态网站基础	61
3.2 逻辑模型、视图和模板	63
3.3 Django架构总览	64
3.4 Django的核心理念	66
3.5 总结	68
第4章 定义和使用模型	69
4.1 定义模型	69
4.2 使用模型	80
4.3 总结	91
第5章 URL、HTTP机制和视图	92
5.1 URL	92
5.2 HTTP建模：请求、响应和中间件	96
5.3 视图与逻辑	100
5.4 总结	105
第6章 模板和表单处理	106
6.1 模板	106
6.2 表单	112
6.3 总结	123
第7章 Photo Gallery	125
7.1 模型	126
7.2 集成文件上传	127
7.3 安装PIL	128
7.4 处理ImageField	128
7.5 构建自定义File字段	130
第三部分 Django应用实例	130

7.6 使用ThumbnailImageField	134
7.7 设置DRY URL	135
7.8 Item应用的URL布局	137
7.9 用模板把它们都串在一起	138
7.10 总结	143
第8章 内容管理系统	144
8.1 什么是CMS	144
8.2 Flatpages	144
8.3 超越Flatpages：一个简单的自定义 CMS	147
8.4 改进建议	162
8.5 总结	163
第9章 Liveblog	164
9.1 究竟什么是Ajax	164
9.2 设计应用程序	165
9.3 应用程序布局	166
9.4 加入Ajax	169
9.5 总结	176
第10章 Pastebin	177
10.1 定义模型	177
10.2 创建模板	179
10.3 设计URL	180
10.4 试运行一下	181
10.5 限制最近Paste显示的数量	184
10.6 语法高亮	185
10.7 通过Cron Job清除	186
10.8 总结	187

第四部分 高级Django技术和特性

第11章 高级Django编程	189
11.1 自定义Admin	189
11.2 使用聚合	193
11.3 生成下载文件	195
11.4 用自定义Manager来增强Django ORM	200
11.5 扩展模板系统	202
11.6 总结	211
第12章 高级Django部署	212
12.1 编写工具脚本	212
12.2 自定义Django codebase	214
12.3 缓存	215
12.4 测试Django应用	223
12.5 总结	229
附录	
附录A 命令行基础	231
附录B 安装运行Django	240
附录C 实用Django开发工具	254
附录D 发现、评估、使用Django应用 程序	261
附录E 在Google App Engine上使用 Django	264
附录F 参与Django项目	273
后记	275

Python 2.x还是3.x

在编写本书的时候，Python 正在从2.x 向新一代的版本3.0 迈移。3.x 版本家族不保证对

第一章里，我们要向你介绍一些我们觉得如果要成为一名出色的Django 工程师就必须知道的Python 技术。我们主要关注的是那些作为Django 工程师一定要知道的Python 概念，而不是再讲一份通用的Python 教程。所以基本上这一章里没有什么 Django 版的。

Django 提供了一个高级的框架，用它只需要很少的几行代码就能完成一个Web 应用。这种轻盈、强大、灵活的框架让你在设计方案时无需太多考虑。Django 是以Python 写成，Python 是一门面向对象的应用程序开发语言，它同时兼具系统语言（如C/C++ 和Java）的强大和脚本语言（如Ruby 和Visual Basic）的灵活迅速。这让Python 的用户创建解决问题的各种难题非常难。

1.1 Python 技术就是Django 技术

这一章要介绍的Python，主要是语言中和Django 开发有关的核心特性和技巧。光有基本的Python 技术是无法高效地开发Django 的，你还需要知道一些更深层次的东西，这样当遇到这些特定的特性、这方面的知识和Django 的需求时，就不会无所适从。对Python 新人或者编程新人来说，先阅读一些其他语言的Python 知识以及本章的内容会帮助你快速入门——选择哪种方式完全取决于你自己的节奏来。

第1章 Django Python 实战

第3章 起始

第2章 Django速成：构建一个Blog

第1章 Django Python 实战

C2341291Y



清华大学图书馆

但是对稍后我们讨论到的复杂对象来说，这个差异就会非常明显，这是因为Python会让你肯定在没有使用print时对象应该如何行为。

```
10
11 print 10
12
13 >>> print 10
14
15
16 >>> 10
```

虽然后面会详细讲解变量和循环，这里我们先来体验一下稍微复杂一点的Python是什么样的。这里有以下几个for循环。

请注意Hello World例子中的不同之处。当直接把一个对象丢给解释器的时候，它会用引号告诉你这是个字符串。而在使用print语句时，引号则不会显示出来，因为你告诉了解释器去显示字符串的内容（当然是不包括引号的）。下面这个例子就和字符串的情况有点不同（对字符串来说显示结果是一样的）。

print是非常に有用的一个命令。它不仅向你的用户提供了相关的应用程序信息，还是一个无可替代的调试工具。虽然也有不显示调试用print功能，“打印”变量值的方法，就像我们前面那样，但是这通常会产生和print不同的结果。

```
>>> print 'Hello World'
Hello World
>>> Hello World
Hello World
>>> Hello World!
```

过程中直接把这些例子输入到解释器里去。像是这样：

在本书中，你会看到很多用Python shell的交互操作：>>>打头的代码段。你可以在阅读有十多年编程经验的老鸟，也是每天都要不开Python shell的！

就能轻易启动它，或者也可以直接从命令行或者是你操作系统的应用程序菜单里启动。这样，你就能在一个切身的体会，对Python以及接下来的Django会更有信心。就算是Python高手，哪怕在阅读本章的过程中，我们建议你启动Python解释器来直接感受一下。大多数Python IDE再运行源文件了，你可以直接用它测试一小段代码。它不但会检查代码的正确性，还能让你在把这段新代码加入到源文件之前进行各种不同的操作，比如查看数据结构、修改值等。

1.2 Jupyter：Python交互解释器

Jupyter只有在绝大多数用户都准备好的时候才会进行升级。

Django的核心团队现在还不打算立刻迁移至3.0上去（和绝大多数面向框架的项目一样，这样的转换破坏性非常大，一定要非常小心），所以这里我们只是提一下这个转换而已。

Jupyter的推出让大家能够更方便地使用这一转模工具：他们会提供可靠的2.x-10-3.x转模工具，并且给予足够的时间保证大家都能顺利完成。

旧版本的兼容性，所以用2.x编写的代码完全有可能无法在3.x下工作。但是Python的核心开发团队正在努力使这一转模尽可能的顺利：他们会提供可靠的应用程序的2.x-10-3.x转模工具，并且给予足够的时间保证大家都能顺利完成。

Python的矩阵运算并学习对(#)表示的。如果它是第一行里第一个字符，那么这整个一行就

蘇玉

在这一节里我们要介绍Python的基本的各个方面，包括矩阵、变量、运算符和基本Python类型。下面几节里会更详细地介绍Python的主要类型。绝大多数Python（和Django）代码都是围绕儿子里会更详细地介绍Python的主要类型。绝大多数Python（和Django）代码都是围绕着以.py为结尾的文本文件，这是告诉系统统一个Python文件的标准做法。你还会看到相关的扩展名诸如.pyo，这不会影响你的使用，现在可以暂时先忽略它们。

1.3 Python基础

要在解釋器里使用Django，你需要在Python解釋器里調試Django應用或框架的話那就太方便了。但是如果你只是按照常規启动解釋器并且试着导入Django模块的话，你只能得到一个Django_SETTINGS_MODULE没有设置的错误。为了方便起见，Django提供了一个manage.py shell命令，它能运行一些必要的环境变量来避免这个问题。

如果你安装了IPython的话，manage.py shell会默认使用它。如果你安装了IPython，但是又希望使用标准Python解释器的话，可以运行manage.py shell plain。虽然在后面的例子里我们是继续使用默認的解释器，不过我们还是强烈推荐用IPython。

Python语言非常重要的一个方面就是没有区分函数的花括号{}。我们用对齐来帮助你发现其实这种方
式时间以后，你就会发现其实这种方式也没有看上去那么糟糕。

字符串或者tab都可以）。如果你习惯于其他语言的话，这可能需要一点时间来适应。但是过一
关手解器器最后一点要说的就是：当你熟练使用解器器了以后，你应该考虑试试另一个类
似的工具IPython。要是你喜欢交互解释器的话，IPython的能力完全是在另一个数量级上的
哦！它提供了系统shell访问、命令行记数、自动对齐、命令历史等许多特性。你可以在
<http://ipython.scipy.org/doc/html/install.html>找到更多关于IPython的信息。它不是随Python一起发布，而是一个第
三方应用。

你已经知道 Python 的代码段是用对齐而不是花括号来区分开的了。我们前面提到这样就能

```
print 'Python and %s are number %d, %s (Django, foo)
if show_output and foo == 1:
show_output = True
```

将它们组合起来。Python 还支持取反的布尔运算符 not。下面是一些例子：

标准的比较运算符，如 <、>=、==、!= 等也是支持的，你还可以分别用布尔运算符 and 和 or。不过这不包括其他语言里有的自增/自减运算符（++ 和 --）。

如 +、-、* 等，以及它们的相对的赋值运算符 +=、-=、*= 等。就是说 $x = x + 1$ 和 $x += 1$ 是一样的。说到运算符，基本上和你熟悉的其他编程语言所支持的运算符一样。这包括算术运算符，

运算符

像只鸭子的话，我们就可以把它当作一只鸭子。

不管它是不是有额外的属性。这叫做 “duck-typing”，即，只要它走起来像只鸭子，叫起来也像只鸭子的话（比如，它有一个字符串应该使用哪种方法），它就可以被认为那个类型，现的像个类型的话。不过，只要这个变量能够表示什么地摊定变量在任何给定的时刻指向的是什么类型的对象。不过，只要这个变量能够表示上去。注意，除非有其他变量也引用了 foo 前面引出的那个字符串（这是完全可能的！），在这个例子中，变量 foo 首先被映射到一个字符串对象 bar 上，接着再映射到另一个整数对象 bar 上去了。

```
I
>>> foo
>>> foo = 1
>>> bar,
      bar,
>>> foo
>>> foo = 'bar',
      bar,
```

所以，变量的值可以在任何时候发生改变，像这样：

Python 的变量不像其他语言那样需要先为之声明一个特定的类型。Python 是一门 “动态类型”的语言。我们可以把变量想象成指向一个匿名对象的名字，这个对象才真正地保存着实际的值。所以，变量的值可以在任何时候发生改变，像这样：

变量和赋值

用它们也很简单，这样配置选项也变得可用了起来。

settings.py 配置文件——注释去掉了不必要的常选项，或者是不用干嘛认值的值，要重新启用它们不仅可以用来自解释器附近的代码，还可以组织代码运行。一个很好的例子就是 Django 的注释。

```
# This entire line is a comment
foo = 1          # short comment: assign int 1 to 'foo'
foo = 1          # short comment: assign int 1 to 'foo'
```

都是注释。# 还可以出现在行的中间，这代表从井字符开始的地方起，同一行中剩下的部分都属于注释。例如：

都是注释。# 还可以出现在行的中间，这代表从井字符开始的地方起，同一行中剩下的部分都

Python有两个主要的数据类型：int（整数）和float（浮点数）。根据KISS原则，Python只

数字

布尔值和数字一样是字符串。下面来看Python中的数字。

在章最后那个例子里的None。这个Python的特殊值和其他语言里的NULL或者void是一样的意思。None在做布尔判断时总是为False。

所决定是不是要执行代码时就要依赖于这些值来判定对象的布尔值。

在Python的if和while语句里做条件判断都是False，但是这个非空的列表却有一个True值。在Python的if和while语句里做条件判断都是False，但是最后一个例子有点绕：虽然列表的两个上面的例子里所有bool的输出都是有意义的。就是最后一个例子有点绕：

```
True
>>> bool(None, 0)
False
>>> bool('')
False
>>> bool(0.0)
True
>>> bool(-1.23)
False
>>> download_complete = False
>>> download_complete = True
True
>>> bool(download_complete)
```

值，可以被当作变量值来显式地赋值。

你可以用bool函数来决定任何Python对象的布尔值，甚至True和False它们自己也是会读的。是False，而所有非零的数值则为True。类似地，空的容器即为False，非空的容器就为True。

的数据值是什么，所有Python值都可以表示为布尔值。举个来说，任何等号两侧的数值都被认为和大多数语言一样，Python可以表示两种布尔值：真（True）和假（False）。无论其本身或着说数据结构。在我们开始介绍主要数据类型之前，首先要注意到的是所有Python对象都有内在的布尔值。

对象的布尔值

现在我们来介绍为Django程序员会用到的标准类型。它们可以是标量（scalar）或者字符串（literal），（比如数字和字符串）。也可以是用来自很多个Python对象组织在一起的“容器”，或者说是数据结构。在我们开始介绍主要数据类型之前，首先要注意到的是所有Python对象都有内在的布尔值。

1.4 Python标准类型

同样的，Python也避免使用很多符号。不仅是那些花括号，Python没有结尾的分号（;），没有货币符号（\$），在条件语句里也不需要小括号（()）。你可能注意到了一些“@”符号和大量的下画线（_），但是这就是仅有的符号了。Python的作者相信只有干净和容易阅读的代码才能避免混乱。

非常容易地辨别其他代码段是属于哪里的，更进一步来说，我们可以消除所谓的“else悬挂”（dangling-else）问题，因为一个else子句只能属于其中一个if，绝对不会产生。

Python中字符串类型的长整数。像1L, -42L, 9999999999999999L等。
④ Python曾经还有一个整数类型long, 但是它现在已经被合并到int里来了。在旧有的代码和文档里你还能看到

右移位(<>和>>)，以及对它们进行赋值运算，如a=, <<=等。

最后，Python可以对整数进行二进制位操作，与(&)、或(|)、异或(^)、取反(~)和

```
0.0
>>> 1.0 // 2.0      # Floor division (// operator)
0
>>> 1 // 2          # Floor division (// operator)
0.5
>>> 1.0 / 2.0       # True division (float operands)
0
>>> 1 / 2           # Floor division (int operands)
```

回整数结果：

Python还提供了显式的“向下取整的除法floor division”，不管操作数的类型是什么，一律都返回截断结果(即向下取整floor division)，面对浮点数来说则是“真正的除法true division”。如果两个操作数都是整数的话，除法运算符/代表“经典除法classic division”，就是说它和指数运算(**)。

数字支持大多数常用的算术运算符：加(+)、减(-)、乘(*)、除(/和//)、取模(%)
数字运算符

类型	描述	举例
int	带符号整数(没有大小限制)	-1, 0, 0xE8C6, 0377, 42
float	双精度浮点数	1.25, 4.3e+2, -5, -9.3e, 0.375
complex	复数(实部+虚部)	2+2j, 3-j, -10.3e+5-60j

表1.1 Python内置数字类型

表1.1总结了这些数字类型，并且给出了更多的例子。
首先，它的值范围比较小，但是更精确。Python还为科学计算提供了一个内置的复数类型。
嗯……最后一个例子有点奇怪。浮点数的范围很大，但是，它们在表示有理数的时候不是
非常精确。反而还有一个浮点数类型Decimal(这不是一个内置类型，必须通过decimal模块从
访问)，它的值范围比较小，但是更精确。Python还为科学计算提供了一个内置的复数类型。

```
1.1000000000000001
>>> 1.1
1.1
>>> -13
-13
>>> -9 - 4
-9 - 4
3.75
>>> 1.25 + 2.5
1.25 + 2.5
```

这些整数和浮点数的例子，以及在解释器里的使用方法：
有一种整数类型int，而不是像其他语言那样提供了好几种整数类型。除了十进制整数还可以
表示为十六进制数和八进制数。浮点数就是和其他语言一样的双精度的浮点实数。下面就是一

到表开始，然后再介绍元组。表1-2列出了我们要讨论的类型并且给出了几个例子。不过我们还是需要告诉它的含义和作用。至于你应该早就知道字符串是什么了，我们将从以下几个地方——当然它的作用是完全不同的。虽然它不是你应该用数据结构的第一选择，没什么特别的地方了。基本上它就是一个“身有残疾的只读列表”，除此之外它也没有特别的类型是元组tuple。基本上它就是一个“身有残疾的只读列表”，除此之外它也没有特别的类型是元组tuple。

序列是Python类型里很典型的可以迭代的类型：即这是一类你可以让你每次获取一个元素位置的对像。例如，my_list[2]就会返回到表中第三个元素（下标从0开始）。你都是用for循环而不是next方法），它还支持随机访问——即你可以直接要求序列中某个元素。其内部会里的元素直到耗尽为止。Python序列不但支撑这种方法的迭代（虽然99%的情况代表背后的的基本思想是你可以用next方法持续地要求下一个对象，而它则不停地“读出”这个对象的值）。迭代通过下标顺序访问。Python的序列类型基本上也是一样的意思，但是它可以包含不同类型的对象，同时其长度还可以长遍伸缩。在这一节里，我们将讨论两个十分常用的Python类型：列表list([1,2,3]) 和字符串string("Python")。他们都属于序列sequence这种数据结构的一种。

很多编程语言都将数据组织为数据结构，数组通常是很长的，由一组相似的对象组合而成，

序列和迭代

现在我们来看看字符串和Python重要的容器类型。
2006) 中的“数字”章节，或是查阅任何完整的参考书籍，也可以在网上搜索Python的文档。要知道更多关于这些函数的信息，可以参考《Core Python Programming》(Prentice Hall,

```
1A.
>>> chr(65)
97
>>> ord('A')
(65, 97)
>>> divmod(15, 6)
(2, 3)
10.0
>>> float(10)
1.2
>>> round(1.15, 1)
45
>>> int(45.67)
123
>>> int(123L)
```

位，而abs函数则可以取得一个数的绝对值。下面是一些使用这些函数的例子：
Python还提供了很多内置函数用以操作数字，像round函数可以将浮点数精确到指定的某一位。

12.0。complex和bool函数也是一样的道理。
以这里用“factory”。即，int(12.34)会创建一个值为12的整数对象，而float(12)则会返回一个现有序列的类型。实际上你是在原来那个对象的基础上返回了一个新的对象（所谓“改变”）。这是在Python里我们不用这些个术语，因为你并没有真的“转换”或者“cast”，但是在Python里我们不用这些个术语，因为你并没有真的“改变”这个现有序列的类型。实际上你是在原来那个对象的基础上返回了一个新的对象（所谓“conversion”或者“casting”），但是你并不用这些个术语，因为你并没有真的“转换”或者“cast”，有的读者把这叫

内置数字工厂函数

获取更多关于对象复制的内容。
② 这里所谓的“拷贝”，是指引用的拷贝，而不是对象本身。更准确的说法是“浅拷贝”。参见下面章节可以了解更多。

例子都是作用于字符串之上，但是切片语法对列表和所有其他序列类型也都是有效的。

```
>>> s = 'Python'
>>> s[0]
'P'
>>> s[2:4]
'yh'
>>> s[-1]
'n'
>>> s[3:-1]
'Pytho'
>>> s[3:]
'Python'
>>> s[:1]
'P'
>>> s[:4]
'Pyth'
>>> s[4:4]
''
>>> s[2:4]
'yh'
>>> s[1:4]
'yth'
>>> s[1:4]
'yth'
>>> s = 'Python'
```

为起始，到第二个下标为止（但是不包括）的子集。

你还可以一次索引序列中的多个元素，在Python中称之为切片（slicing）。切片是用一组用冒号（：）隔开的索引来表示的，比如[*im*]。当要求切片的时候，解释器会取以第一个下标！
data[len(data) - 1]来获取数组中最后一个元素吧？在上面代码里的最后一个例子中，我们只需要一个简单的-1就可以了。

```
data[len(data) - 1]
```

Python还提供非常灵活的负索引！我们一定都写过类似这样的代码data[-len(data) - 1]或是一

```
s[-1]
```

言不同，Python的字符串可以同时被当作是离散的对象和一组单独的字母。

刚才我们提到序列是可以直接索引的，接下来是一些操作字符串的例子。和许多其他语言

序列切片

```
str
'django', 'in', '%s is number %d' % ('Python', 1), 'hey there'
list
[123, 'foo', 3.14159, [], [x.upper() for x in words]
tuple
(456, 2.71828), (), ('need a comma even with just 1 item,')
```

表1.2 序列类型举例

其他序列操作符

在上一节我们看到了用`[]`和`[:]`进行切片操作的方法。还有其他一些可以作用于序列之上的操作包括了连结`(+)`、复制`(*)`以及检查是否是成员`(in 和 not in)`。和前面一样，这里用字符串来举例，当然这些也是能用于其他序列之上的。

```
>>> 'Python and' + 'Django are cool!'
'Python andDjango are cool!'
>>> 'Python and' + ' ' + 'Django are cool!'
'Python and Django are cool!'
>>> '-' * 40
'-----'
>>> 'an' in 'Django'
True
>>> 'xyz' not in 'Django'
True
```

连结的其他形式

我们建议你避免在序列上使用`+`操作符。当你还是新手的时候，它确实能很方便地将字符串组合在一起。但是这样做的效率不高。（细节的原因需要解释Python下C的实现手法，这不是本书的内容——你只要相信我们就好了。）

例如拿字符串来说，你可以用稍后字符串一节中要讨论的字符串格式化操作符`(%)`来取代`'foo'+'bar'`，像这样`'%s%s' % ('foo', 'bar')`。另一种办法用`join`方法，特别是当要把一组字符串拼在一起的时候非常方便，例如`"join(['foo', 'bar'])`。对于列表来说，`extend`方法也可以把另一个列表中的内容加进来（相比`list1 += list2`，`list1.extend(list2)`要好得多）。

列表

Python类型里很像其他语言中数组的一种类型是列表`list`。列表是可变的，可以改变大小的序列，它能够保存任何数据类型。下面的例子里我们展示了如何创建一个列表，以及你能进行的操作。

```
>>> book = ['Python', 'Development', 8]          # 1) 创建列表
>>> book.append(2008)                          # 2) 附加对象
>>> book.insert(1, 'Web')                      # 3) 插入对象
>>> book
['Python', 'Web', 'Development', 8, 2008]
>>> book[:3]                                  # 4) 切片头三个元素
['Python', 'Web', 'Development']
>>> 'Django' in book                         # 5) 对象属于列表吗?
False
>>> book.remove(8)                            # 6) 显式移除对象
>>> book.pop(-1)                             # 7) 通过索引移除对象
2008
>>> book
['Python', 'Web', 'Development']
>>> book * 2                                  # 8) 重复 / 复制
```

列表推导式

列表推导式 (list comprehension) 是一个由逻辑代码组成的结构 (construct)，这是从另一个语言 Haskell 借鉴来的)，它构造了一个包含由那段逻辑代码所产生的值或对象的列表。例如，有一个包含整数 0~9 的列表。如果我们想要对其中的每一个数字加一，并且返回的结果也要

一个排序好的或者是倒序的拷贝。

2.4f 以上的版本提供了内置函数 sorted 和 reversed，它们分别接受一个列表作为参数并且返回一个排好序的或者是一个排好序的拷贝。Python

当然，有时候我们需要获得一个排好序的拷贝而不是直接在给定序列上排序。Python 大写的字符串 (修改过的) 拷贝。稍后会详细介绍可改变性。

Python 的新手可能会觉得 sort 不返回一个排好序的列表的行为相当奇怪，所以使用的時候一定要小心。另一方面，之前看到的字符串方法 upper 都返回了一个字符串 (包含了所有字母都是大写的字符串)。这是因为和列表不同，字符串是不可改变的，所以upper 方法只能返回一个 (修改过的) 字符串。

列表的内置函数如 sort、append 和 insert 等都是直接对对象进行修改而且没有返回值。

不过要是你把文件或者是类实例放进去的话，那结果就一定是定义的了。

对所有数值进行排序 (从小到大)，然后对字符串字典序排序。这个例子还算有点意义，字符串和数字) 怎么能相互比较呢? Python 所用的算法是“尽量猜测”来获取“正确的行为”：先

对一个混合类型的列表进行“排序”的结果其实是未定义的。没有关系的对象 (比如字符串

```
[8, 2008, 'Development', 'Python', 'Web']  
>>> book  
>>> book.sort()  
# 注意：这是直接排序……，没有返回值!
```

我们先重设前面例子中的列表，并对其进行排序，然后再讨论其细节。

列表的方法

如你所见，列表是一种非常灵活的对象。下面我们就深入讨论一下它的方法。

9. 用另一个列表将其展开列表。

8. 显示复制操作符 * 的用法。

7. 根据位置 (即下标) 移除 (并返回) 一个元素。

6.无论元素的位置，从列表中移除它。

5. 成员检查 (元素是否属于列表)。

4. 截取头三个元素的一个切片。

3. 在第二个位置上插入一个字符串 (下标为1)。

2. 在列表尾部添加另一个整数。

1. 创建一个初具有两个字符串和一个整数的列表。

我们来逐条解释一下上面的例子：

```
[Python, 'Web', 'Development', 'With', 'Django']  
>>> book  
>>> book.append(['With', 'Django']) # 9) 合并到当前列表  
>>> book.extend(['Web', 'Development', 'Python', 'Web', 'Development'])
```

Python 的另一个序列类型是字符串，虽然它们都是用单引号或者双引号括起来的（“this is a string”或是“this is a string”），但是你可以把它们想象成一个字符串组。另外，和列表不同的是，字符串是不能修改的，其大小也不能改变。任何试图改变字符串长度或是修改内容的行

序

生成器表达式是Python 2.4以后才支持的特性，所以如果你还在使用Python 2.3的话，就没办法使用它了，目前它在Python程序员之间的名气也还不是很高。不过，要是你的输入序列可能会变得很大时，最好考虑采用生成器表达式而不是列表表达式。

```
<>>> even_numbers = (x for x in range(100000) if x % 2 == 0)
<>>> even_numbers
<>>> <generator object at 0x...>
<>>> len(even_numbers)
100000
```

在上一个例子中，我们用到表推导式在10个元素的列表中查找偶数，但如果列表中有一万个甚至上十万个数字时要怎么办？要是列表中不是简单的整数，而是复杂庞大的数据结构时又该怎么办？在这种情况下，生成器表达式节省内存的特点就能帮助我们解决问题了。我们只要稍微修改一下语法，将方括号替换成小括号就能把列表推导式变成生成器表达式（`genexp`）了。

Python还支持另一种和列表推导式类似的数据结构，叫做生成器表达式（generator expression）。除了它有一种称之为“惰性计算”（lazy evaluation）的特点，它和列表推导式的用途基本一致。它工作的方式是每次处理一个对象，而不是一口气处理和构造遍整个数据结构，这样做的潜在优点是可以节省大量的内存（虽然有时候要付出一点点性能上的代价）。

生威器美送夫

第二个例子展示了如何使用最后的`if`子句来过滤元素。它不包含任何做出修改的逻辑代码，自身也是一个合法的表达式（结果就是`x`的值）。在过滤序列时列表推导式非常有用。

```
>>> even_numbers = [x for x in range(10) if x % 2 == 0]
[0, 2, 4, 6, 8]
>>> even_numbers
```

阅读理解排列表推导式推导法是先从里面的for循环开始，向右查看是否有if条件（上面一个子序列）上的值外。

和列表一样，列表推导式也采用方括号表示，并且用到了一个简写版的Python for循环表达式。虽然我们还没有讲到循环（马上就会介绍），但是你还是应该很容易理解这个推导式。第一部分是一个生成器表达式元素的表达式，第二部分是一个输入表达式 (input expression)，会产生或

```
>>> data = [x + 1 for x in range(10)]
>>> data
[1, 2, 3, 4, 5, 6, 7, 8, 9, 10]
```

算第一个倒霉的遇害者吗？有了倒霉推算式（或者简写为“histcomps”），我们可以这样：

为实际上只是创建了一个新的修改过的字符串而已。不过对普通使用来说，这些操作都是透明的，只有在处理内存问题时才会显得重要。

字符串一样，字符串也有很多方法，这里要再次重申，因为字符串是不可改变的，字符串和列表一样，字符串也有很多方法，这里要再次重申，因为字符串是不可改变的，字符串没有一个能修改现有的对象，它们只是返回了一个修改过的拷贝而已。在编写本书时，字符串已经有超过37个方法了！我们主要介绍一些可能会在Django里用到的方法。下面是一些例子：

```

>>> s = 'Django is cool'
# 1
>>> words = s.split()
# 2
>>> words
['Django', 'is', 'cool']
# 3
>>> ', '.join(words)
'Django is cool'
# 4
>>> ', '.join(words)
'Django::is::cool'
# 5
>>> ' '.join(words)
'Django is cool'
# 6
>>> s.upper()
'DJANGO IS COOL'
# 7
>>> s.upper().lower()
'django is cool'
# 8
>>> s.title()
'Django Is Cool'
# 9
>>> s.capitalize()
'Django Is Cool'
# 10
>>> s.count('o')
2
# 11
>>> s.find('go')
4
# 12
>>> s.startswith('Python')
False
# 13
>>> s.replace('Django', 'Python')
'Python is cool'

```

我们来小结一下上面的例子：

- 创建字符串。
- 用空字符串把字符串分隔到列表里去。
- #2的操作。 (把字符串组合成一个字符串并以空格分隔。)
- 和#3一样，但是以一对冒号分隔。(将所有子串合在一起。)
- 创建字符串。
- 用空字符串分隔到列表里去。
- #2的操作。 (把字符串分隔成一个字符串并以空格分隔。)
- 创建字符串 (并且丢弃) 一个新字符串，在原有版本上全部大写。(另外还可以参考lower方法。)
- 展示了如何串联函数来确认一个字符串是不是全部大写。(另外还可以参考lower方法。)
- 让字符串的每个单词都以大写开头，剩下的部分都是小写。

```

    >>> mystring = u"This is Unicode!"
    >>> print mystring
    This is Unicode!
    但是打印和转换raw或unicode字符串的操作却不会打印出字符。
    >>> mystring = u"This is Unicode!"
    >>> print mystring
    This is Unicode!
    >>> mystring = u"This is Unicode!"
    Unicode字符串。这些指字符串（designator）在编写代码和在解释器里显示字符串时都用的着。
    Python字符串可以在引号前加一个标记符：r代表这是一个raw字符串，而u则代表这是一个
    字符串指字符串

```

字符串方法	描述
count	字符串中子串出现的次数
find	查找子串（类似的还有index, rfind, rindex）
join	用一种分隔符合并子串
replace	替换字符串
split	将字符串分成子串（类似的还有splitlines）
strip	移除字符串首尾的空白符（另请参见rstrip, lstrip）
startswith	字符串是不是以子串开头（另请参见endswith）
title	大写每个单词的第一个字母（另请参见capitalize, swapcase）
upper	大写整个字符串（类似的有lower）
isupper	字符串是大写的么？（类似的还有islower等）

表1.3 常用的Python字符串方法

Python的文章来获取更多关于字符串方法的内容。

总结在本书出现过的办法。和你推測的一样，字符串还有很多其他方法，我们推荐你查阅[第13章](#)。除了isupper方法，字符串还有很多其他的以is-开头的方法，例如isalnum、isalpha等。

例子#7展示了只要你明確知道每一个方法调用所返回的结果是什么类型，就可以将困难串联起来使用。因为我们知道upper方法返回的是字符串，所以完全可以调用到类别的方法。

例子#7展示了只要你知道字符串是一个列表的话，那么你就可以调用到类别的方法。

如果字符串全部去掉（或者strip方法也行，它会移除字符串首尾的空白符）。

和split方法类似，还有一个splitlines方法，它会特别地去查找换行符（而不是空白符）。

如果有一个包含这些字符串的字符串，比如从文件中读取数行的时候，你可以用rstrip方法将所有结尾的空白符掉（或者strip方法也行，它会移除字符串首尾的空白符）。

14. 简单的查找替换。
13. 检查字符串是不是以给定子串开头。在这里为否。（另请参见endswith方法）。
12. 和#10一样，但是没有找到匹配，所以返回-1。
11. 找出字符串go在字符串中的起始下标（索引值为4）。(还可以看考index方法。)
10. 计算子串o在字符串里出现的次数。
9. 只大写第一个单词的开头，其他部分小写。

尾符来保证执行的话，使用三引号就可以了，像下面这样的XML代码段。

```
尾符来保证执行的话，使用三引号就可以了，像下面这样的XML代码段。
上面的例子还展示了三引号的使用，这是Python独有的特性。你可以在字符串中插入
一些数据必须和转换说明一一对应。你可以参考Python文档以获得所有可用的格式化转换说明。
在上面的例子中，我们有一个字符串（其格式化字符串的转换说明是%）和一个整数
```

```
there
hi
>>> print hi
hi\there,
>>> hi
here...
>>> hi = ...hi
Django is number 1,
>>> %s is number %d, %s[6], 1)
```

的转换说明 (directive)。这里是一些例子。

在本章之前的例子中，已经见过字符串格式化操作符 (%) 了。它通过“格式化字符串”(format string) 来为将不同的输入类型打印到字符串做准备，这个格式化字符串包含了某些特殊字符或字符串操作符和三引号字符串。

因为通常普通的Python字符串只含很少的一些字符（包括Western字母加上一些特殊符号），它们不足以显示很多非英语言的字符。Unicode是一种新型的字符串，拥有大量的编码方式，所以不存在这个限制。Django (在编写本书时) 已经花费了相当多的精力来保证框架的每个部分都能支持Unicode。因此在开发Django的过程中，你会遇到很多Unicode字符串对象。

在使用正则表达式的Python代码中，raw字符串是十分常见的。拿Django来说，就是在你的URL的配置规则中，Django根据匹配URL请求和你提供的正则表达式的结果来控制分派。用raw字符串来编写这些规则能让他们看上去更清晰一些。而且为了统一性，不管一个正则表达式是否用到了反斜杠，通常一律都使用raw字符串。

Django中的raw字符串

另一个使用raw字符串的地方是正则表达式，因为它大量使用了诸如反斜杠 (\) 这样的特殊字符。在正则表达式一节中，我们会用raw字符串来表示一个正则表达式，`r"\w+@\w.\w+"`，这比普通的（需要转义符的）字符串要容易阅读多了，`"\w+@\w+\.\w+"`。

```
filename = r'C:\temp\newfolder\robots.txt'.
```

一个新行，不过有时候你不需要发生类似这样的转换，比如一个DOS的文件名：“raw” 指示告诉解释器不要转换字符串中的任何特殊字符。例如，特殊字符串通常代表

```
This is Unicode!
>>> str(mystring)
This is Unicode!
```

```

xml = '''
<?xml version="1.0"?>
<Request version=".1f">
    <Header>
        <APIName>PWDapp</APIName>
        <APIPassword>youllneverguess</APIPassword>
    </Header>
    <Data>
        <Payload>%s</Payload>
        <Timestamp>%s</Timestamp>
    </Data>
</Request>
'''
```

最后，注意上面例子中虽然用到格式化转换说明，但是却没有格式化操作符和相对应的元组。这是因为字符串格式化操作符就是一个操作符而已。你完全可以先在代码里定义好格式化字符串，稍后再用操作符和元组填充之。

```

import time      # use time.ctime() for timestamp
VERSION = 1.2    # set application version number

# [...]

def sendXML(data):  # define some sendXML() function
    'sendXML() - transmit XML data'

# [...]
payload = 'super top-secret information'
sendXML(xml % (VERSION, payload, time.ctime()))
```

元组

元组 (tuple) 是前面讨论的列表的近亲。表面上两者的一个明显区别在于列表用方括号表示而元组用小括号，不过更深层次的区别就又要论及Python的对象模型了。虽然列表允许并且提供了方法来改变它的值，但元组是不可改变的，即它不允许你改变它的值——这也是它们没有方法的部分原因。

乍看之下，新学Python的程序员会很奇怪为什么要有这么一个单独的数据类型，或者说，干嘛不直接提供一个“只读”的列表呢？看起来，这个理由满充分的，但实际上，元组提供的功能远比一个“只读”列表要多的多。它们的主要作用是作为参数传递给函数调用、或是从函数调用那里获得参数时，保护其内容不被外部接口修改。

这不是说它们对程序员就没有用了。虽然它们在前台上的用武之地不大，但是在后台却是使用得相当频繁的。例如在Django的配置文件中就会见到大量的元组。虽然没有方法，但是它们依然进行常用的序列操作，也可以用在内置函数里。

Django中容易犯的和元组相关的错误

在典型的Django应用里你经常会看到元组的使用，对Python的新手来说，这或许是最晦涩

第二个别字看懂有点问题，怎么回事呢？记住，元组是由逗号决定的，而不是小括号。所以序列类型了——特别是单个元素的元组要求在最后“必须”跟一个逗号的时候。下面我们就来看一看能不能理解下面的这些例子：

```
>>> a = ("one", "two")
('one', 'two')
>>> b = ("just-one")
('just-one',
one)
>>> c = ("just-one",)
('just-one',)
>>> d = "just-one"
'just-one'
>>> e = ([0])
[0]
>>> f = ([0])
[[0]]
```

很多Django的配置都是用元组来指定的——admin选项、URLconf规则，以及很多在settings.py中的设置。所以Django的有些部分和其他部分比较起来更容易告诉你错误的内容。如果你在设置admin选项时用了字符串而不是元组，那么你会得到一个很有用的错误信息，例如告诉你“admin.list_display”必须是一个列表或是元组。另一方面，如果你的ADMINS或MANAGERS设置缺少了结尾的逗号，你会发现服务器会试图给你名字里的每个字母发送错误邮件的email！因为这是一个在Django的新手中非常容易犯的错误，我们会在下面的“常见错误”一节中再次强调。

与数字一样，所有序列类型都有一个特殊的工厂函数来复杂创建所需求类型的实例：list、tuple和list，另外还有一个为Unicode字符串准备的unicode的工厂函数。通常str负责为一个对象提供可读的或者说“可打印的”(printable)字符串表示。在Python里的另外一个函数repr也提供相似的功能，但是它产生的是一个“可计算的”(evaluable)字符串表示。这通常意味着，它是一个将Python对象作为字符串的纯表示(pure representation)，如果对这个字符串进行eval语句的评估该可以把它再转成对象。

内置函数len能告诉你序列里有多少个元素。max和min分别返回序列中“最大”和“最小”的对象。any和all告诉你序列中是否“任意”或“全部”元素都是True。

在其他语言里的for循环环需要自己计数，而籍由range函数的帮助，Python中的for循环则更具有迭代的风格。不过，内置函数enumerate则把这两种风格结合了起来。它返回了一个特殊的迭代器，既能访问序列的下标，同时又能访问其相对应的元素。

在表1.4中我们列出了所有这些常用的序列函数。

表1.4 内置序列函数和工厂函数

函数	描述 ^①
str	(可打印的)字符串表示(参见repr, unicode)
list	列表表示
tuple	元组表示
len	对象的势
max	序列中“最大的”对象(参见min)
range	给定范围里可迭代的数字(参见enumerate, xrange)
sorted	返回排好序的列表(参见reversed)
sum	序列值加和(数字)
any	是不是有元素为True?(参见all)
zip ^②	返回N个元组的迭代器,其中每个元组包含了N个序列里对应的元素

①虽然这里很多函数的描述都标为“序列”。但是实际上这些函数可以用在所有“可迭代”对象上。这里说的可迭代对象是指任何类似序列的数据结构,只要它可以迭代就行,比如序列、迭代器、生成器、字典的键值和文件里的行等。

②这里的描述不是很清楚,有点拗口。zip函数接受多个序列,然后将每个序列对应位置的元素组织在一起组成一个新的元组,最后返回由这些元组组成的列表。它的长度由参数序列中最短的长度决定。举个简单的例子:

```
>>> zip ([1, 2, 3], ('a', 'b', 'c'))
[(1, 'a'), (2, 'b'), (3, 'c')]
```

映射类型: 字典

字典是Python里唯一的映射类型。字典是可变的、无序的、大小可变的键值映射,有时候也称为散列表(hashes)或是关联数组。虽然其语法和序列很相似,但你是用键而不是下标来访问元素值,而且字典用花括号({})而不是方括号(列表)或是小括号(元组)来定义的。

字典是目前讲到的Python语言里最重要的数据结构。它是大多数Python对象的幕后黑手。不管对象是什么类型或怎么使用的,绝大多数情况下在幕后都有一个字典在管理着它的属性。闲话少说,我们来看看究竟什么是字典以及它能干什么。

字典的操作、方法和映射函数

下面是一些如何使用字典的例子。我们来讨论一下这些代码的作用以及给出这些操作符、方法和函数的用法。

```
>>> book = { 'title': 'Python Web Development', 'year': 2008 }
>>> book
{'year': 2008, 'title': 'Python Web Development'}
>>> 'year' in book
True
>>> 'pub' in book
False
>>> book.get('pub', 'N/A')           # where book['pub'] would get an error
'N/A'
>>> book['pub'] = 'Addison Wesley'
>>> book.get('pub', 'N/A')           # no error for book['pub'] now
'Addison Wesley'
>>> for key in book:
```

1.5/小结了常用的字典方法。
的那个替换掉。最后，我们用遍用的内置函数len获取字典的长度，即字典里多少键-值对。表
第一行用del命令移除了一个键-值对。随后我们将title键赋予了另一个字符串，把之前

2

>>> len(d)

```
{'year': 2008, 'title': 'Python Web Development with Django'}
>>> d
>>> d['title'] = 'Python Web Development with Django'
>>> del d['pub']
```

我们来演示如何用del关键字来移除键-值对，让字典回到之前开始的状态：
这种情况，所以如果你把另一个对象用一个已经在存在的键赋值的话，它会覆盖前一个值。下面
这是当你试图用一个已经在存在的键在散列表里添加另一个值时会发生的事情。在Python里没有
如果你早就知道散列表的话，那么你一定对“键冲突”(key collision) 这个概念很熟悉了。

```
{'year': 2008, 'pub': 'Addison Wesley', 'title': 'Python Web Development'}
```

>>> d

>>> Addison Wesley,

```
>>> d.setdefault('pub', Addison Wesley)
```

>>> d

```
>>> d = {'title': 'Python Web Development', 'year': 2008}
```

会产生错误了。

另一个更加强大的方法是setdefault。它的作用和get一样，而且要是键不存在的话，它还会
用默认值自动创建键-值对，这样当后面再调用“dict.get(key)”或“dict[key]”时，它们就不
会生产错误了。

后来get方法更加安全，因为它总是返回一个值而不会发生错误。

设置这样一个默认值的话就会返回None，你也可以用方括号语法，就像从字典里获取一个元
组一样，d['pub']。这里的区别是pub不是字典的一个成员，所以d['pub']会返回一个错误。相比
设置这回第4步，如果传递给get的键不存在于字典的话，它就会返回第二个参数（如果你没有
回到底4步，因为字典的键不存在于字典的话，它就会返回None）。

我们先来看看最后一段代码。我们用了一个for循环来迭代字典里的所有键。这是非常典型的

手法。你还可以确认我们刚刚说过的特点：字典的键是无序的（如果我们关心顺序的话，我们
就会用序列！），而且正是这种无序让字典在查找元素值的时候非常地高效。

7. 迭代整个字典并显示每一对键-值。

6. 再次使用get方法，不过这次成功地获取了值。

5. 加入一个新的键-值对。

4. 使用get方法获取指定键的值（在这里获取的是默认值）。

3. 检查字典的时候含有某个键（两次，一次为真，一次为假）。

2. 显示这个对象。

1. 创建一个初学者字典，它包含有一个字符串和一个整数。它们的键都是字符串。

那么上面的代码都做了些什么？我们来总结一下：

```
title : Python Web Development
pub : Addison Wesley
year : 2008
...
print key, ':', book[key]
```

```

if data[0] == 'y':
    print "You typed '%s'" % data[0]
else:
    print "You typed '%s'" % data[0]

```

Python里是怎么工作的了。

和Bourne家族的脚本语言 (sh, ksh和bash) 是一样的。只要举一个简单的例子你就会明白它在Python和其他语言一样也提供了if和else语句。Python的“else-if”实际上只是写成elif，这就表或者元组重复的执行一段代码) 等逻辑，以变量形式出现的数据本身并没有太多的含义。

不在它们之上加语系附件判断 (根据条件选择代码执行的“路径”) 和循环 (根据某种形式的列表或者元组重复的执行一段代码) 等逻辑，以变量形式出现的数据本身并没有太多的含义。

在了解Python变量的基础之后，我们现在来看看除了对它们赋值以外还能做什么。如果

条件判断

你会发现在这些标准类型里，列表和字典 (“dict”) 是应用程序里用的最频繁的数据结构。元组和字典主要是用在函数调用之间交换参数和返回值。字符串和数字也时有用到。Python还有很多数据类型可供选用，但是这里介绍的都是在编写Django应用过程中会经常碰到的类型。

1.5 编程控制

Django中“类似字典的”数据类型

面，有的地方我们不仅需要字典这样的键-值行为 (key-value behavior)，还需要一些字典没有提供的功能。一个最显著的例子就是HttpRequest对象中保存GET和POST参数的QueryObject对象。由于为同一个给定的参数 (字典的键) 提交多个值在这里是合法的，但是Python的字典却无法提供这个功能，我们需要一个特殊的数据结构。如果你有一个应用程序员处理HTTP请求这种有点怪异的特性的话，请参见Django文档关于Request和Response对象的章节。

字典方法	描述
keys	所有键 (参见iterkeys)
values	所有值 (参见itervalues)
items	所有键-值对 (参见iteritems)
get	获取给定的键或默认值 (参见setdefault, fromkeys)
pop	从字典里移除并返回值 (参见clear, popitem)
update	用另一个字典更新字典

表1.5 常用的Python字典方法

```

STATUSES_CHOICES = enumrate(("solid", "sguishi", "lqid"))
用到choices关键字的地方——关于model field参数的细节请参见第4章。其用法如下：

在编写Django代码时一个使用enumrate非常方便的地方就是model的意义，特别是在
在Django的Model中使用enumrate

```

```

2 3.14
1 abc
0 123
...
... print i, value
>>> for i, value in enumrate(data):
>>> data = (123, 'abc', 3.14)
举个例子，enumrate是一个能让你同时迭代和计数的内置函数（for循环自身没办法计数）：

什么时候用for循环是需要经验和技术的。
如在调试的时候，你是没办法在列表推导式里使用print语句的。选择什么时候用列表推导式，很多时候都既可以（而且最好！）转换成列表推导式。不过有时候也需要简单的for循环，例如
多简单的循环都可以（而且最好！）转换成列表推导式。不过有时候也需要简单的for循环，例如
如果你还记得列表推导式的话，你会发现其实这个例子就能转换成一个推导式。事实上很

```

```

print line
if error, in line:
for line in open('/tmp/some_file.txt'):
    ... print this
    ... #打印变量这种东西。
但是Python这里，for则更像是shell脚本里的foreach循环，它能让你专注于问题本身而不用关心
for。在其他语言里，for只是作为一个可以计数的循环而存在，和它的搭档while差不多。
不过老实说，在Python里你不会经常需要这样使用while循环，因为我们有更加强大的循环
机制for。

```

```

4
3
2
1
0
...
... i += 1
... print i
>>> while i < 5:
>>> i = 0
语句为假：

```

和许多高级语言一样，Python也有while循环。while会重复执行同一段代码直到它的条件

循环

```

print 'Invalid key entered!'
else:

```

记住虽然大多数错误都会导致异常，但一个异常不一定代表一个警告，有时它们可以作为一个信号通知上层的调用函数，异常处理可以处理从简单的单个异常到一系列多个不同的异常。你可以看到这里我们处理了一个异常（except语句的代码）：

```
# attempt to open file, return on error
try:
    f = open(filename, 'r')
except IOError, e:
    f = None
    print "ERROR: you provide invalid data", e

# now we can use f as usual
for line in f:
    print line.strip()

# finally, close the file
f.close()
```

当然你也可以为多个异常创建对应的多个处理语句：

```
try:
    f = open(filename, 'r')
except IOError, e:
    print "ERROR: you provide invalid data", e
except ValueError, e:
    print "ValueError", e
```

和C++、Java等现代语言一样，Python也提供了解算表达式。如同在本章开头中的例子，它想程序员提供了一种在运行时发现错误，进行恢复处理，然后继续执行的能力。Python try-except语句和其他语言里的try-catch结构是很相似的。

1.6 算法基础

```

class Icecream(models.Model):
    flavor = models.CharField(max_length=50)
    status = models.IntegerField(choices=STATUS_CHOICES)
    models.IntegerField(choices=CHOICES)
    class Meta:
        ordering = ('-status',)

def __str__(self):
    return self.flavor

```

最后一个except语句利用了Exception是(几乎)所有异常的根类(root class)这一事实，所以如果有异常没有被它之前的任何语句捕获的话，这个异常会在最后一个语句里被捕获。

Python还提供一个try-finally语句。我们不关心捕提到的是什么错误，无论错误是不是发生，这些代码“必须”运行，比如关闭文件、释放锁，把数据写进磁盘还会连接池等。例如：

```
try:
    do_something()
finally:
    log("ERROR: data retrieval accessing a non-existent element")
```

当没有异常发生的时候，finally中的代码会在try完成之后立即运行。如果发生任何错误，那么finally的代码还是会执行，但是它不会消除异常，异常依然还会在调用链里寻找能处理它的语句。

```
try:
    get_mutex()
    do_something()
except (IndexError, KeyError, AttributeError), e:
    log("ERROR: data retrieval accessing a non-existent element")
finally:
    free_mutex()
```

自Python 2.5起，try-finally可以和except一起使用了。(之前的版本不支持这种写法。)

到目前为止，我们只讨论了如何捕捉异常，那么怎么才能抛出异常呢？答案是使用raise语句。假设你在自己的库里创建了一个API调用要求传入一个大于0的正整数。在内置函数instance的帮助下(检测对象的类型)，你代码看起来应该是这样的：

```
# check if integer
if not isinstance(must_be_positive_int, int):
    raise TypeError("must_be_positive_int must pass in an integer!")

# check if positive
if must_be_positive_int < 1:
    raise ValueError("must_be_positive_int must be greater than zero!")

# normal processing here
```

表1.6列出了最常用和你在学习Python过程中很有可能会碰到的异常。

读写：

```

>>> f.close()
bar
foo
...
...
print line.rstrip()
for line in f:
    f = open('test.txt', 'r')
    f.close()
    f.write('bar\n')
    f.write('foo\n')
    f = open('test.txt', 'w')

```

在之前的代码里我们已经见过几个使用内置函数open的例子了。它的作用是打开文件以供

1.7 文件

和任何其他的Python程序一样，Django也大量使用了异常。不过大多数情况下这些都是在内部处理，不属于日常使用Django的范围。但是有些异常却是需要你在Django应用里直接使用的。例如，抛出一个Http404的异常会通知Django转去处理HTTP 404 “Not Found”的错误。这种在发生错误时能立即处理Http404错误的能力（而不是小心翼翼地通过传回一个错误的标志）在创建Web应用里是非常好用的。

Django里的异常

当前异常的完整列表可以在官方文档 <http://docs.python.org/lib/module-exceptions.html> 的 exceptions模块一节中找到。

异常	描述
AssertionError	试图访问一个对象没有的属性，比如foo.x，但是foo没有属性x
AttributeError	试图访问一个对象没有的属性，比如foo.x，但是foo没有属性x
EOFError	输入/输出异常；基本上是无法打开文件
ImportError	无法引入模块或者包；基本上是路径访问错误
IndentationError	语法错误；代码没有正确对齐
IndexError	下标索引超出序列边界，比如当x只有三个元素，却试图访问x[5]
KeyError	试图访问字典不存在的键
KeyboardInterrupt	Ctrl+C被按下了
NameError	使用一个还未被赋予对象的变量
SyntaxError	Python代码非法，代码不能编译
TypeError	传入对象类型与要求的不符
UnboundLocalError	试图访问一个还未被设置的局部变量，基本上是由于你有一个同名的全局变量，导致你以为正在访问它
ValueError	传入一个调用者不期望的值，即传值的类型是正确的

表1.6 常见的Python异常

```
import httpplib
def check_web_server(host, port, path):
    h = httpplib.HTTPConnection(host, port)
    h.request('GET', path)
```

现在我们来看一点更实际的例子。

调用函数则更加简单：给出函数名和一对小括号，然后在括号里放人所需的参数就可以了。

```
123
>>> foo(123)
...
...
>>> def foo(x):
    print x
```

的语句要用一对括号即可：

声明函数的方法是用def关键字，然后是函数名以及小括号里的参数列表。（如果无参数

声明和调用函数

- 类处理器
- (函数签名里的) 变长参数
- (函数调用里的) 参数容器
- 匿名函数和lambda
- 函数是first-class的对象
- (函数签名里的) 默认参数
- (函数调用里的) 关键字参数
- 声明和调用函数

要讨论一些更深入的用法，包括（但不限于）：

在Python里创建函数非常简单。在这一章里我们已经看过一些函数的声明了。这一节我们

1.8 函数

Programming》的“文件和I/O”一章里找到介绍。

最后，另外这里没有论及的一些不太常用的文件方法都可以轻易地在《Core Python

们会组合并为一行。

我们一定要调用字符串的strip方法来去掉它们。（否则输出会多出一个空行来，因为print会自动加上回车。）类似地，通过write或writeln方法发送给文件的字符串也需要断开得，否则它们

断开得（根据你的操作系统，可以是in、tin、或it）在读入行时是保留的，这就是为什么要

个简单的for循环在大多数情况下就已经足够了。

一个文件对象本身就是一個迭代器，所以通常没有必要直接使用read或readlines 方法。一

个字符串列表用正确的顺序写回到文件里去。

本文件来说，readlines方法能将文件里所有行读入到一个列表里，而writeln方法则可以将一

个字符串列表用正确的顺序写回到文件里。对文件来说，除了write方法以外，还有一个read方法可以把整个文件的内容读入到一个字符串里。

```

resp = h.getresponse()
print 'HTTP Response:'
print '    status =', resp.status
print '    reason =', resp.reason
print 'HTTP Headers:'
for hdr in resp.getheaders():
    print '    %s: %s' % (hdr,

```

这个函数的作用是什么？它接受一个主机名或者IP地址（host），服务器的端口号（port），以及一个路径名（path），然后尝试和运行在指定主机和端口号上的Web服务器连接。如果成功的话，它会向给定的路径提交一个GET请求。若执行下面的这段代码检查Python网站的服务器，你会得到的输出如下：

```

>>> check_web_server('www.python.org', 80, '/')
HTTP Response:
status = 200
reason = OK
HTTP Headers:
content-length: 16793
accept-ranges: bytes
server: Apache/2.2.3 (Debian) DAV/2 SVN/1.4.2 mod_ssl/2.2.3 OpenSSL/0.9.8c
last-modified: Sun, 27 Apr 2008 00:28:02 GMT
etag: "6008a-4199-df35c880"
date: Sun, 27 Apr 2008 08:51:34 GMT
content-type: text/html

```

(函数调用里的) 关键字参数

除了这样“常规的”调用约定，Python还允许你指定关键字参数（keyword argument），这可以让函数更加清晰，更加容易使用，同时，也消除了记忆固定参数顺序的需要。关键字参数通过“键=值”这样的形式加以指定，下面是对之前例子里的一点修改（输出被省略了）：

```
>>> check_web_server(port=80, path='/', host='www.python.org')
```

根据给出的关键字，Python就会在执行程序时将相应的对象正确的赋值给那些变量。

(函数签名里的) 默认参数

Python函数的另一个特点是可以为参数提供默认值，这样向它们传递参数就不是必需的了。很多函数都有一些对所有调用值都一样的变量，在这种情况下定义默认值就可以让函数好用一点。

参数的默认值是在函数签名里用等于号直接指定的。拿刚才检查网站的代码为例，大多数网站都是运行在80端口上，并且检查“网站是否运行”通常只要测试首页就行了。所以我们可以加入这样的默认值：

```
def check_web_server(host, port=80, path='/'):
```

这里，不要和关键字参数弄混淆了。关键字参数只能用于“函数调用”，而默认参数则是用于“函数声明”。所有必须提供的参数一定要出现在任何可选参数之前，不能混在一起或者颠倒顺序。

```
def check_web_server(host, port=80, path):      # INVALID
```

将列表和字典作为默认参数

这里我们要提醒你注意一个Python初学者常犯的错误。还记得之前我们讨论的关于可变和不可变的变量么？列表和字典是可变的，而字符串和整数是不可变的。因为这样，将列表和字典作为默认参数可能会非常危险，特别是当它们持续地穿过多个函数调用时，像这样：

```
>>> def func(arg=[]):
...     arg.append(1)
...     print arg
...
>>> func()
[1]
>>> func()
[1, 1]
>>> func()
[1, 1, 1]
```

这种可变对象的特性不是非常明显，这也是为什么我们要在这里特别提到的原因。一不小心你就会发现有些函数会引入非常奇怪的行为。

函数是First-Class对象

在Python里你可以把函数（和方法）当作和其他对象一样的使用，例如，把它们存放在容器里，赋值给其他变量，作为参数传递给函数等。唯一的不同是你可以这个函数对象，就是说当附上括号和参数时，你就把它当作了一个函数来对待。在深入讨论之前，我们先要理解对象引用的概念。

引用

当执行def语句的时候，你实际上是在创建一个函数对象并将其赋值或绑定到当前名字空间里的一个名字上，但这可以只是它许多引用（或别名alias）里的第一个。每次当你向另一个函数调用传递一个函数对象，将其放入一个容器，或是赋值给一个局部变量时，你就为那个对象创建了一个额外的引用，或者说别名。

作为一个例子，下面的代码在全局名字空间里创建了不是一个，而是两个变量，因为函数一经定义，它就和其他变量没什么两样了：

```
>>> foo = 42
>>> def bar():
...     print "bar"
...
```

和其他Python对象一样，函数可以有任意多的引用。下面是一些例子来使之更容易理解。首先是bar的正常用法：

这有点不太好理解。如果我们将其default设置为`datetime.date.today()`的话——注意这里的括号——函数将会在model类义的时候被调用，这不是我们想要的，所以我们传递的是函数对象。`Django`会意识到这一点，并且在创建实例的时候调用函数，为我们生成需要的值。

```
date = models.DateField(default=datetime.date.today)
entry = models.TextField()
class DiaryEntry(models.Model):
```

import datetime

在默认情况下就能收到创建时间，你可以在它被调用时传入一个标准库的函数对象。例如，如果你希望`DateField`另一个利用函数对象的地方是在model field里的默认参数。例如，如果你希望`DateField`在这里代码里，`listview`作为一个函数对象被直接使用，而不是一个包含了函数名的字符串。

```
)
```

```
url(r'^list/$', ListView),
urlpatterns = patterns('',
```

```
from myproject.myapp.views import ListView
from django.conf.urls.defaults import *
```

就是在URLconf文件里设置`Django`view。

`Django`充分利用了Python函数对象可以像其他值一样被传递这一特性。最常见的例子是当我们要用函数或者对象的名字（比如`bar`）和实际调用或执行它（比如`bar()`）区分开来了。

注意只有当我们想要调用函数的时候才会加上小括号。在把它当作变量或者对象传给方法的时候，你只需要用函数的名字（就像上面我们创建`function_list`做的那样）就行了。这样引用函数对象或者对象的名字（比如`bar`）和实际调用或执行它（比如`bar()`）区别开来了。

```
bar
bar
...
function()
...
>>> for function in function_list:
    >>> function()
>>> function_list = [bar, baz]
```

且传入参数就可以了。例如：

要使用一个存放容器里的函数对象，你只需要和其他对象一样引用它，然后跟上小括号并

```
bar
bar
...
>>> bar()
>>> bar()
>>> bar()
```

接着，我们将`bar`赋值给另一个名字`baz`，如此之前的函数`bar`现在也可以通过`baz`来引用了。

Django里用到lambda函数的地方不太少，但是有一个地方看起来特别的适合：即所谓“认证装饰器”（authentication decorator），它的作用是确认用户有足够的权限访问某些类地展示了Python函数First-Class的特性。

第三个例子其实不太现实（我们需要重写使用这个函数时），这时，第二个例子就比较恰当。（例如当程序员意识到要重写使用这个函数时），这时，第二个例子就比较恰当。通常这是解决问题的好办法。不过，也有很多lambda最终“成长”为一个普通的函数这三个语句最大的不同之处在于可读性和重用性。第一个例子更简洁同时目的相当明确，

```
get_last_name = lambda person: person.last_name
sorted(list_of_people, key=get_last_name)
```

甚至我们还可以这样写：

```
sorted(list_of_people, key=lambda person: person.last_name)
```

```
def get_last_name(person):
    return person.last_name
```

当然下面这样的写法也是等价的：

这是正确的原因为key期望的是一个函数对象，而lambda返回的正好是一个匿名函数。

```
sorted(list_of_people, key=lambda person: person.last_name)
```

排序的话，则可以这样：

lambda的一个常见用途就是为sorted这样的函数工具提供一个函数对象，它在众多参数里直接受为排序的数据。例如有一个代表人物的复杂对象列表，我们需要按照他们的姓氏属性来有一个key参数，这个key必须是一个函数，可以用在列表里的每个对象的元素上生成一个

函数对象，或者你可以选择将它保存为一个变量，或是保存为一个回调函数以便稍后执行。

lambda的语法如下：lambda args:表达式。在执行的时候，lambda返回一个可以立即使用的

使用lambda

一个动作，而不是返回或生成一个值。

结果不是对象的代码被称为“语句”。例如if或者print语句，for和while循环等。它们执行向对象。例如42、1+2、int(123)、range(10)等。

Python代码由表达式和语句组成，并由Python解释器负责执行。它们的主要区别是“表达式”是一个值，它的结果一定是一个Python对象。当Python解释器计算它时，其结果可以是任何对象。例如42、1+2、int(123)、range(10)等。

表达式 vs. 语句

这里我们要弄清楚“表达式”expression和“语句”statement之间的区别以避免任何混淆。同，故而没有名字，所以也称为匿名函数。它们通常只是一行表达式，所以一般都是用完就扔。一个表达式组成，这个表达式代表了函数的“返回值”。这样的函数和普通函数声明的方式不是一个函数是Python提供的另一个函数编程的特性。创建的方法是使用lambda关键字，它由

匿名函数

```
host_info = {'host': 'www.python.org', 'port': 80, 'path': '/'}

建立一个具有和('www.python.org', 80, '/')相似内容的字典。
```

这样的写法就明显要干嘛得多，而双引号对字典的用法也是类似的。现在我们来创建一个字典：

```
check_web_server(*host_info)

将其“打开”。所以下面的例子和上面的代码是等价的：
```

然而这种写法既不可扩展（要是函数有超过一打的参数怎么办？）也不好看。这时单引号就可以解决我们的问题，因为当调用函数时，表达式在计算一个带星号前缀的元组或列表时会将它们拆开。

```
check_web_server(host_info[0], host_info[1], host_info[2])

这时调用就会变成：
```

```
host_info = ('www.python.org', 80, '/')
# http://www.python.org/
```

三元组里的语法怎么办？例如：

```
用check_web_server(127.0.0.1, 8000, '/admin')即可调用这个函数。那要是这些信息在一个
```

```
def check_web_server(host, port, path):
    我们来重新看一下之前的例子check_web_server函数。下面是函数的签名：
```

```
函数调用里的*和**。
星号则代表附近有字典“出没”。我们先从函数调用开始。
```

一般说来，无论是函数调用还是声明，单个星号表示有元组（或是列表）“出现”，而两个星号则代表附近有字典“出没”。我们在从函数调用开始。

```
这一步我们要讨论Python里*和**的特殊含义，它们都和函数有关但是函数调用时和函数声明时却有着不同行为。在开始之前，我们先要对所有C/C++程序员讲解，这里的星号和指针可没有任何关系！
```

```
*args和**kwargs
```

以①开头的行就是一个“函数装饰器”（function decorator），在本章稍后你会学到这个概念。装饰器通过“包装”函数（例如这里的vote函数）来改变它们的行为。这里的参数为参数，并返回一个布尔值（True或False）。因为这里的测试非常简单（我们只是简单地返回了User对象的一个特定的属性），只需一行代码就可以了。

```
"""
Process a user's vote
"""

def vote(request):
    user_passes_test(lambda u: u.is_allowed_to_vote)

def process_a_user's_vote():
    user_passes_test(u: u.is_allowed_to_vote)

    这样的函数可以用一般的def foo(): 结构来定义，但是这不影响我们理解它的作用。
```

就返回True，反之则返回False。

页面。一种办法是将一个代表已登录用户的User对象传递给一个函数，如果允许用户访问

也就等于：

于是函数的调用就变成了：

check_web_server(**host_info)

这告诉函数在打开字典时，每个键是参数的名字，同时对应的值是函数调用的参数。即，

你可以同时使用这些技术，这和手动的通过下标或是关键字参数调用函数是一样的。

函数签名里的*和**虽然看样子相似但作用却完全不同：它们让Python得以支持变长参数，

有时也称为“varargs”。即函数可以接受任何数量的参数。

当定义一个有三个参数的函数时（没有默认值的参数），调用者必须传入正好三个参数。

默认参数虽然引入了一些灵活性，但函数依然受限于所定义参数的最大数目。

如果需要更大的灵活性，则可以用星号表示的元组来定义一个变长参数，就好像传入一个“购物袋”一样包含了所有元素。现在我们来创建一个这样的“daily sales total”函数：

def daily_sales_total(*all_sales):

 total = 0.0

 for each_sale in all_sales:

 total += float(each_sale)

 return total

def daily_sales_total(*args):

 dailly_sales_total(*args)

这样你向这个函数传递多少参数，它都能够处理。all_sales就是一个包含了所有参数的元

组（这也是为什么我们可以在这个函数定义里选择all_sales的原因）。

不管你怎么向这个函数传递多少参数，它都能够处理。all_sales就是每一个包含了所有参数的元

组（这也是为什么我们可以在这个函数定义里选择all_sales的原因）。

你还可以把普通参数和变长参数混在一起使用，这时vararg就会捕捉所有剩下的参数，例如

你现在在这个假设的check_web_server函数定义就能接受额外的参数了。

def check_web_server(host, port, path, *args, **kwargs):

 # 代码省略

 # 代码省略

学习Python函数和函数式编程中最后一个可能也是最难懂的概念就是装饰器（decorator）。在这里，Python的装饰器指的是一种让你能改变或者说“装饰”函数行为的机制，它能让函数执行一些和原本设计不同，或是在原有基础上额外的操作。装饰器也是也可以说是对函数的一个包装。这些额外的任务包括写日志、计时、过滤等。

Python里一个被包裹或被装饰的函数（对象）通常会被重新赋值给原来的名字，这样被包裹的函数能和普通的版本保持兼容——因为使用装饰器就是在原有的功能上再“加盖”额外的功能。

卷之三

Django 的数据模型 API 查询经常需要包含关键字符串。例如：

```
    bob_stories = Story.objects.filter(title__contains="bob", subtitle__contains="bob", byline__contains="bob")
```

这样，你就可以很方便地创建字典了：

```
    bob_stories = {title: bob, subtitle: bob, byline: bob}
```

这样，你就可以很方便地创建字典了：

```
    bob_stories = Story.objects.filter(**boobargs)
```

这样，你就可以很方便地创建字典了：

```
    bob_stories = dict((f + '_contains', bob) for f in ('title', 'subtitle', 'text', 'byline'))
```

这样，你就可以很方便地创建字典了：

```
    bob_stories = Story.objects.filter(**boobargs)
```

有助于组织动态获取的数据对象的过滤参数（例如从搜索表单里得到的选项）。

这样的函数可以按照 `f()`, `f(a, b, c)`, `f(a, b, foo=c, bar=d)` 等任何方式调用，它可以接受任何形式的输入。当然了，函数内部如何处理 `args` 以及 `kwargs` 的内容则取决于它的作用。

实际上，还有一种全部用类参数组成的所谓“通用”Python方法签名（“universal” Python method signature）。
def f(*args, **kwargs):

下面这个简单的例子记录了函数调用的发生：

```
def log(func):
    def wrappedFunc():
        print "*** %s() called" % func.__name__
        return func()
    return wrappedFunc

@log
def foo():
    print "inside foo()"
```

现在我们来看看执行代码后所生成的结果：

```
>>> foo()
*** foo() called
inside foo()
```

在本章稍早的地方，我们看见过这样一个接受一个参数的装饰器。

```
@user_passes_test(lambda u: u.is_allowed_to_vote)
```

在这个例子里，实际上是调用了一个函数，然后返回一个事实上的装饰器——user_passes_test自己并不是一个装饰器，它只不过是一个接受了参数的函数，然后用这些参数返回一个可用的装饰器。它的语法是这样的：

```
@decomaker(deco_args)
def foo():
    pass
```

这个下面的代码是等价的，注意这里Python表达式是如何串联在一起的：

```
foo = decomaker(deco_args)(foo)
```

这里的“装饰器生成器”（decorator-maker，或者叫decomaker）接受参数deco_args并且返回一个可以包装foo函数的装饰器。

最后一个例子展示了如何同时应用多个装饰器：

```
@deco1(deco_args)
@deco2
def foo():
    pass
```

这里我们不再多做深入讨论了，不过根据现有的代码，你可以认定它和下面的代码是等价的：

```
foo = deco1(deco_args)(deco2(foo))
```

你或许会想，“为什么要用装饰器呢”？实际上，包裹函数的概念在Python里不算什么新的东西，同样接受对象，修改并重新赋值给原来的变量也不是新的概念。不同的是装饰器可以让你用一个简单的@符号完成这一切动作。

如果你需要一份更完整更易读的装饰器教程，请访问Kent John的“Python Decorators”教

实例化

如果你有使用其他语言来创建实例，不过在Python里你只要像调用函数一样调用类的构造函数——也没有new和free关键字。稍后我们会更详细地讨论这一点。

如果name是一个属性的话，你必须用self.name作为完整的名字来引用它。也就是说self开头以及点-属性(dotted-attribute)形式出现的变量都是一个实例对象。如果self指向的是类的一个特定的实例对象，那么实例就是它的具体实现，一个在执行过程中创建出来让你操作的真类是一份图纸，那么实例就是它的一份具体实现，(在其他语言里用this来代表这个self的概念)。如果

变量self指向的是类的一个特定的实例(在其他语言里用this来代表这个self的概念)。这样，就是每个方法都必须带上一个额外的self参数——Python在这方面非常明确。

version这样类的静态成员(Static class member)的创建是通过在类定义里赋值完成的。这样的成员就是属于整个类的变量，可以在所有实例之间共享。方法(method)的意义和函数一样，就是每个方法都必须带上一个额外的self参数——Python在这方面非常明确。

```
class AddressBookEntry(object):
    version = 0.1

    def __init__(self, name, phone):
        self.name = name
        self.phone = phone

    def update_phone(self, phone):
        self.phone = phone

    self.phone = phone
```

我们选用的第一个建模现实世界里的问题是创建一份通讯录。要在Python里创建一个类，你需要提供class关键字，这个新的类的名字，以及一个或多个基类(base class)。例如，如果你不想从任何现成类继承的话，只需要Python的根(root)类或类型object作为你的基类。要创建一个Car(轿车)或是Truck(卡车)类的话，你可以用Vehicle(车辆)作为一个基类。就像这里我们创建的“通信录”类，AddressBookEntry，就行了，就像这里我们创建的“通信录”类，AddressBookEntry：

首先声明，这一节不是一份面向对象编程(OOP)的教程。我们只不过是准备介绍一下如何在Python里创建和使用类。如果你不知道什么是OOP的话，使用Python是学习它最简单的办法之一——当然最好你还是先读一下其他高级教程，不过我们不做强求。OOP的主要目标是在代码和现实问题之间提供一个合理的映射关系，并且鼓励代码的重用和共享。我们还会展示一下Python独有的特性。

类的定义

程，<http://personalpages.tds.net/~ken37/kk/00001.html>。

1.9 面向对象编程

所以名字用的也是`__init__`。在实例化一个对象时，你需要传入`__init__`所需的参数。在Python在创建对象时会用你提供的参数自动调用`__init__`，并最终返回将一个新创建的对象返回给你。

我们说的是调用方法，不过其实也完全可以说是函数调用。虽然你在方法声明里被要求提供`self`参数，但是在（以常规的方式）调用方法时却不用这么麻烦，Python会自动为你传入`self`对象。下面的例子创建了两个`AddressBookEntry`类的实例：

```
john = AddressBookEntry('John Doe', '408-555-1212')
jane = AddressBookEntry('Jane Doe', '650-555-1212')
```

Python会为每个实例调用`__init__`，随后返回这个对象。仔细看，这里并没有在调用时给出`self`，只有`name`和`phone number`。Python会为你传入`self`的。

现在你就可以直接访问属性了，比如`john.name`、`john.phone`、`jane.name`和`jane.phone`。如同下面的例子，基本上对实例属性的访问没有什么限制：

```
>>> john = AddressBookEntry('John', '408-555-1212')
>>> john.phone
'408-555-1212'
>>> john.update_phone('510-555-1212')
>>> john.phone
'510-555-1212'
```

再次强调一下，注意这里的`update_phone`方法实际上有两个参数，`self`和`newphone`。但是我们只需要提供新的`phone number`就可以了，Python会自动将指向`john`的实例对象作为`self`传递进来。

Python还支持动态的实例属性，即那些没有在类定义里声明的属性，可以“凭空”创造出来。

```
>>> john.tattoo = 'Mom'
```

这是一个非常好用的特性，完全展示了Python的灵活性。你可以随时随地创建任意数目的属性。

继承

创建一个子类和创建一个类是很相似的，你只需要提供一个或多个基类（不是对象）就行了。继续刚才的例子，现在我们来创建一个员工通信录条目类。

```
class EmployeeAddressBookEntry(AddressBookEntry):
    def __init__(self, name, phone, id, social):
        AddressBookEntry.__init__(self, name, phone)
        self.empid = id
        self.ssn = social
```

注意这里我们没有给`self.name`和`self.phone`分配对应的`name`和`phone number`——这是因为在`AddressBookEntry.__init__`调用里已经为我们处理了。当你按这种方式重写某个基类方法的时候，你必须显式地调用它（原来的方法），就像上面这个例子那样。注意这一次我们必须传入`self`参数，因为我们指向的是`AddressBookEntry`类而非实例。

Django使用了一种叫做正则表达式 (regular expression) 的字符串匹配机制来定义网站的 URL。如果没有正则表达式 (也常称为 “regex”) 的话，我们将不得不定义每一个可能的 URL，在对 /index/ 或 /blog/posts/new/ 的时候可能还好，但是遇上稍微动态一点的 URL，像 /blog/posts/2008/05/21/ 这样的 URL 就很麻烦了。

1.10 正则表达式

这里定义了一个从django.db.models.Model继承而来的新类 BlogPost，它有三个自定义的变量。继承 Model 还赋予了 BlogPost 其他方法和属性，包括了允许你从数据库里查询 BlogPost 对象，创建新的对象，以及访问相关的关系。

```
class BlogPost(models.Model):
    ordering = ('-timestamp',)
    class Meta:
        ordering = '-timestamp'
        get_latest_by = 'timestamp'
        verbose_name = 'Blog Post'
        verbose_name_plural = 'Blog Posts'

    body = models.TextField(max_length=150)
    title = models.CharField(max_length=150)
    class BlogPost(models.Model):
        from django import db
        from django.db import models
        timestamp = models.DateTimeField()
        class Meta:
            ordering = ('-timestamp',)
```

Django 的数据 model (绝大多数 Django 应用的核心) 是从 Django 的内置类 django.db.models.Model 继承而来的子类。Django 的 Model 类提供了大量的强大特性，我们将在第 4 章里做深入介绍。现在，我们先用第 2 章里的一段代码来做一个例子：

```
class MyModel(models.Model):
    pass
    class Meta:
        pass
```

就像用来创建装饰器的“嵌套函数” (inner function) 一样，你还可以创建嵌套类 (inner class)，即在类的内部定义的类，例如：

```
class MyModel(models.Model):
    inner_class = inner_class()
```

除了那些被重写的基本类方法外，子类还会继承所有一切，所以这里 EmployeeAddress 和 BookEntry 也同样拥有 name 和 phone 属性，以及 update_phone 方法。和绝大部分 Python 编序和框架一样，Django 在其内部以及用户关心的外部特证里也都大量使用了继承。

个只有一个元素的元组？又或者，为什么在面向对象的Python代码里到处都是self？
我们在这一节里要讨论的是—些Python新手经常容易犯的错误，比如：怎么做才能创建一

1.11 常见错误

re.search则稍微宽容一点，所以我们能得到一个非空的结果。
因为“foo”只能匹配“seafood”的一部分，所以调用re.match的结果为空（None）。但是

```
>>> re.match(r'foo', 'bar')
None
>>> m = re.match(r'foo', 'seafood')
<_sre.SRE_Match object at ...>
>>> print m
<_sre.SRE_Match object at ...>
>>> m.group()
'foo'
...
>>> if m is not None: print m.group()
>>> m = re.match(r'foo', 'seafood')
>>> print m
None
>>> if m is not None: print m.group()
>>> m = re.match(r'foo', 'seafood')
>>> print re
```

而匹配表示整个字符串都需要符合模式的描述。例如：
这里我们要区分“查找”和“匹配”的概念。查找是在目标字符串里搜寻任何匹配的模式，

查找和匹配

一样，正则表达式最好用raw字符串来表达，这样就能避免大量使用反斜杠这样的转义符。
当它失败的时候，你得到的是None。注意这里用到的raw字符串标志r——和本章之前提到的
search函数会成功时返回一个Match对象，调用它的group方法就可以得到匹配的字符串。

```
>>> m = re.match(r'bar', 'seafood')
None
>>> print m
<_sre.SRE_Match object at ...>
>>> m.group()
'bar'
>>> if m is not None: print m.group()
>>> m = re.match(r'foo', 'seafood')
>>> print m
<_sre.SRE_Match object at ...>
>>> m.group()
'foo'
...
>>> if m is not None: print m.group()
>>> m = re.match(r'foo', 'seafood')
>>> print m
None
>>> if m is not None: print m.group()
>>> m = re.match(r'foo', 'seafood')
>>> print re
```

re.match返回一个匹配对象，随后可以用这个对象的group或groups方法获取匹配的模式。
Python的正则表达式可以通过re模块来访问，这是在查找函数中使用非常频繁的一个组件。

re模块

们在Python里的用途。
的余下部分会假设你已经熟悉了regex的使用（至少也是读过一点教程了），现在我们来看看它
不花太多时间在这个话题上了。你可以到www.djangoproject.com的网站上找到一些很好的资源。本节
很多书籍和在线教程都介绍了究竟什么是正则表达式以及如何编写它们，所以这里我们就

单元素的元组

如果这样(1)。)好像这样(1)。

初学者虽然能明白(和(123, 'xyz', 3.14)都是元组，但是却不太能意识到(1)其实不是一个元组。Python实际上是在多处重载了小括号。比如当用在表达式里的时候，小括号的作用是分组。如果你想要一个单元素的元组，在那个元素后面必须跟一个不太漂亮不过一定要有的尾巴(逗号)。

我们已经见过很多种导入模块和它们属性的方法了：

```
import random
print random.choice(range(10))
```

和

```
from random import choice
print choice(range(10))
```

第一种方法是将模块的名字设置为一个包含的名字空间里的全局变量，这样你就可以好像是访问全局性那样访问choice函数。而在第二个例子中，我们是直接把choice引入到全局名字空间里来(而非模块的名字)。因此不再需要把这个属性当成是模块的成员了。实际上我们也只有了这个属性而已。

Python新手之间经常有一种误解，以为第二种方法只导入了一个函数，而没有导入整个模块。这是不对的。整个模块其实已经被导入了，但是只有那个函数的引用被保存了起来。所以from-import这种语法并不能带来性能上的差异，也没有节省什么内存。

新手经常会想的一个问题是他们有两个模块m.py和n.py都导入了foo.py模块。当导入n时，foo岂不是会被人两次?简单的来说，没错，是这样的，但是和你想的有点不一样。

不能重复导入一个模块

Python有导入模块(import)和加载模块(loading)之分。一个模块可以被导入任意多次，但是它只会被加载一次。就是说，当Python碰到一个已经加载过的模块又被导入时，它会跳过加载的过程，所以你无需担心额外消耗内存的问题。

Package是Python在文件系统上发布一组模块的一种方式，它使用常见的dot-ed-attribute方式来访问子模块，就像它们是另一个对象的属性而已一样。或者说，如果你发布的软件产品里有几个子模块的话，基本上是不可能把它们都放在同一个目录下的。当然这样不行，但你真的想这么做?为什么不利用文件系统来用一种更符合逻辑更直观的方式组织模块呢?

```
Phone/
    __init__.py
    util.py
    Voicedata/
        util.py
```

例如。假设有一个电话程序。我们可以组织这样的一个目录结构：

在这一节里你会了解到，可读性在Python编程里是非常重要的一个部分。这也解释

五

初学者经常会问的一个问题是，到底Python是“传引用”的，还是“传值”的？这个问题的答案没办法用简单的是否来回答，所以只能说是“看情况”——有的对象在传入函数时是一个引用，而有些则是被复制过来，即传值。而判断的数据就是看对象的可改变性（mutability），而这一点又取决于对象的类型。由于这种双重行为，Python程序员通常不用“传引用”或是“传值”这种说法，取而代之的是对象是可变的（mutable）还是不可变的（immutable）。

可读文件

这看样子有点笨拙。在实际工作中，基于效率考量你可能会挑一点边缘，当然更多情况下还是为了可以少打几个字母啦！Python和Django都是以简单易用著称，尽量遵循DRY的原则。下面的例子可能更加贴近现实一点：

```
import Phone.Mobile.Analog as pma
```

```
pma.dial()
```

Phone类是顶层目录，或者叫package。在它之下有子package，不过实际上它们就是包含在其他Python模块的子目录里了。你可以看到每个子目录下都有一个init_.py文件。这告诉Python解释器这些目录下的文件应该被当作是一个package而不是普通文件。它们一般都是空文件，当然也可以在使用任何package代码之前做一些初始化的工作。

假設現在我們要訪問analog cell phone的dial函數。我們可以執行：

```
import Phone.Mobile.Analog
Phone.Mobile.Analog.dial()
```

```
__init__.py
posts.py
isach.py
Fax/
__init__.py
g3.py
mobile/
__init__.py
Analog.py
Digital.py
__init__.py
```

而不是浅拷贝。如果确实需要后面这种行为，即“深拷贝”（deep copying），你必须使用`copy.deepcopy()`方法。

其实都是一个指向内存中列表对象的引用，在任何一个列表表中修改它都会影响到另一个列表，而列表是可变的，所以把`mylist2`得到的只是另一个列表的副本，而列表是可变的，所以把`mylist2`得到的只是另一个列表的副本。但是，`mylist2`的第三个对象是一个新的整数和字符串对象，所以把`mylist2`进一步完全正确。但是，`mylist2`的第三个对象是一个

`mylist`中前两个对象是不可变的（一个是整数和一个字符串），所以`mylist2`得到了两个全

```
[2, 'a', ['biz', 'bar']]  
>>> print mylist2  
[1, 'a', ['biz', 'bar']]  
>>> print mylist
```

两个列表的新值。

你在这里希望对嵌套列表`mylist`的修改不会影响到`mylist2`，但实际上不是这样的！看看这

```
>>> mylist[2][0] = 'biz'  
>>> mylist2[0] = 2  
>>> mylist2 = list(mylist)  
>>> mylist = [1, 'a', ['foo', 'bar']]
```

为什么这一点如此重要呢？请看下面这个嵌套列表的例子：

我们在内部有一个对象，即在内存中只有一份对象，而有两份引用。我们提到了不可变对象是传值的，而可变对象是传引用的。不管是向函数传递参数或是任何形式的对象复制来说，这都是一样的：不可变对象（比如整数）被真正复制，而可变对象只是复制了一个对象的引用，即在内存中只有一份对象，而有两份引用。

现在我们来看一个Python新手很容易犯的错误，可改变性和复制对象。在本书开始的时候，

复制对象和可改变性

一份修改过的拷贝。

假设通过`newlist = list(mylist)`或`newlist = mylist[:]`，随后在新的拷贝之上调用sort方法来获取对象作为输入，并且返回一个排序序或者倒置的拷贝。这在不需要修改原对象或者是希望节省几行代码的情况下非常有用。如果你一定要用Python 2.3的话，你需要手动复制这个列表（一行代码的情况）。

Python新手彻底搞晕了！很多其他列表方法，如`reverse()`、`extend()`、以及字典方法`update()`（向字典添加新的键值对）等都是直接就地修改对象。

一个常用的例子是`list.sort()`，它就是直接在列表上排序而不是返回它，这很容易把很多对象的拷贝（即这类函数返回的都是`None`）。

调用可变对象的方法是一个需要小心谨慎对待的区域。如果你调用的函数有任何修改后

可改变性如何影响方法调用

会变得非常有用。

不是。这种能够提供“份列表”对象的能力在需要不可变性的场合下（例如字典里所有键）

了为什么Python要提供两种“列表”类型，列表`list`和元组`tuple`，其中列表是可变的而元组

须导入copy模块中的copy.deepcopy。请在使用前仔细阅读相关文档——这种类型的拷贝通常都是递归的（如果有循环引用的话就会产生问题），而且不是所有对象都是可以深拷贝的。

构造函数v.s. 初始化程序

虽然Python是面向对象的语言，但是它和传统OOP语言的一个区别之处在于它没有显式的构造函数的概念。Python里没有new关键字因为你并没有真的实例化你的类。相反，Python会为你创建实例并调用初始化程序——这是在你的对象被创造出来之后，但是在Python将它返回给你之前调用的第一个方法。它的名字是__init__。

要实例化一个类，或者说要创建一个对象，你可以像调用函数一样调用这个类。

```
>>> class MyClass(object):
...     pass
...
>>> m = MyClass()
```

此外，由于Python会为你自动调用__init__，如果它需要参数的话，你必须在“调用”类的时候提供给它。

```
>>> from time import ctime
>>> class MyClass(object):
...     def __init__(self, date):
...         print "instance created at:", date
...
>>> m = MyClass(ctime())
instance created at: Wed Aug 1 00:59:14 2007
```

类似的，Python程序员通常不需要实现析构函数来销毁对象，只需让它离开作用域就行了（这时它们会被自动垃圾回收）。不过，你可以在Python对象里定义一个__del__方法来当作析构函数，而且你可以用del语句来显式地销毁一个对象（例如，del my_object）。

动态实例属性

Python新手另一个可能会搞混的东西（特别是来自其他面向对象语言背景的程序员）是实例的属性可以动态分配，即使是在类定义已经完成甚至已经创建实例以后。例如这里的AddressBook类：

```
>>> class AddressBook(object):
...     def __init__(self, name, phone):
...         self.name = name
...         self.phone = phone
...
>>> john = AddressBook('John Doe', '415-555-1212')
>>> jane = AddressBook('Jane Doe', '408-555-1212')
>>> print john.name
John Doe
>>> john.tattoo = 'Mom'
>>> print john.tattoo
Mom
```

代码的文档是非常に有用の、Python不仅允许你为代码创建文档，还让你可以在运行时访问它。模块、类，以及函数和方法都可以创建docstring。比如下面这个简单的例子foo.py：

创建字符串（即“docstring”）

这里的主要原因是可读性，另外当你要在里面添加额外代码时无需更多编辑。

```
return
if __finished__:
    return
else:
    print("我们还是推荐你把它分开放到多行里去，像这样：")
    print("但是从屏幕上看起来一切正常的话，很有可能")
    print("虽然这样的代码是完全合法的：")
    print("不要像标题一样把一组代码写在同一行里")
```

转換成空格后，你就可以发现哪里的对齐有问题了。

如果Python解释器向你报错说代码有错误，但是从屏幕上看起来一切正常的话，很有可能就是代码的什么地方混进了tab因此编辑器“错乱地”显示了你的代码。在最浅的地方所有tab都被转换成空格后，你就可以发现问题里的对齐有问题了。

无论在什么平台上开发，你的代码总是有可能会被移除或是复制到另一个不同架构的机器上，或是运行在一个不同的操作系統上。由于制表符（tab）在不同平台上处理的方式都不同（例如在Win32上是四个空格，而到了POSIX或UNIX系统上就变成了八个），所以最好还是干脆避免使用tab。

使用空格而非Tab

在一个用空格区分行代码的语言里，再加上各种各样的用户类型，如果只用一个或是两个空格来对齐Python代码的话，编辑起来实在是太伤眼睛了。同样用八空格的话，代码又太容易折行。Guido在最早的文章中就一直建议使用四个空格是一个完美的折中。

四空格对齐

Python里有很多“正确的编码风格”的元素、推荐和建议，但是总的来说就是要保持代码的味道要够“Pythonic”。遵循Python编码风格，DRY（不要重复自己），易读的代码，鼓励优雅的方案和代码重用等设计哲学的编程系统就算是Pythonic的标准，Django也不例外。在这里有限的篇幅里，我们只能提供一些基本的指导。剩下的部分则来自Python、Django的经验，在它们高度重视的社区。Python有一份官方的风格指导，你可以在PEP8里找到它们（<http://www.python.org/dev/peps/pep-0008>）。

1.12 代码风格

注意这里的self tattoo没有在类、方法声明和类方法的任何地方出现过——甚至在__init__里也没有！我们能够在运行时动态的创建属性。这就是动态实例属性（dynamic instance attribute）。

```
foo.py:
#!/usr/bin/env python
"""foo.py -- sample module demonstrating documentation strings"""

class Foo(object):
    """Foo() - empty class ... to be developed"""

def bar(x):
    """bar(x) - function docstring for bar, prints out its arg 'x'"""
    print x
```

这里说的“可以在运行时访问”指的是：如果你启动解释器并导入模块时，你可以通过每个模块、类和函数的`__doc__`属性来访问它的docstring。

```
>>> import foo
>>> foo.__doc__
'foo.py -- sample module demonstrating documentation strings'
>>> foo.Foo.__doc__
'Foo() - empty class ... to be developed'
>>> foo.bar.__doc__
'bar(x) - function docstring for bar, prints out its arg "x"'
```

此外，你还可以用内置函数`help`来“漂亮地”打印docstring。这里我们只给出一个模块的例子（你可以自行实验类和函数的docstring）。

```
>>> help(foo)
Help on module foo:
NAME
    foo - foo.py -- sample module demonstrating documentation strings

FILE
    c:\python25\lib\site-packages\foo.py

CLASSES
    __builtin__.object
        Foo

        class Foo(__builtin__.object)
            | Foo() - empty class ... to be developed
            |
            | Data descriptors defined here:
            |
            |     __dict__
            |         dictionary for instance variables (if defined)
            |
            |     __weakref__
            |         list of weak references to the object (if defined)

FUNCTIONS
    bar(x)
        bar(x) - function docstring for bar, prints out its arg "x"
```

这一章的内容非常丰富，我们希望能让读者所有Django的程序员来到Python的世界。虽然不能没有我们预想的那么全面，但是这毕竟是关于Django的一本书，介绍Python不是我们的主要任务。不过我们还是尽量把Django所需要的Python相关知识浓缩在这里。

本章的前三个部分基本上都是关于Python语言的教材。最后一部分是一些“教性的技能”：常见的错误，编码风格指导等。希望我们已经展示了足够的内容来帮助你读写一些基本的Python代码，这样你就能顺利地进入Django的世界了。

在下一章里，我们准备手把手带你完成一个简单的blog应用来感受一下 Django。虽然这个blog不像外界商业的系统那样完整，但是它应该能让你感到Django应用开发的乐趣，而且在这过程中你还能练习一下在本章里学到的Python技巧。

1.13 总结

Python标准库里的大多数软件都有`docstring`，所以要是你需要内置函数、内置类型的方法，或是任何模块和包属性的帮助信息的话，只要你导入了相关属性的必要模块，就可以随时寻求帮助，如`help(open)`, `help("strip")`, `help(time,ctime)`等。

现在我们来看看这个命令为你创建的目录里有什么内容。在UNIX上它看起来是这样的：
示例】来启动这个窗口。另外，你会看到C:\WINDOWS\system32这样的提示符，而不是\$符号。
在Windows上，你需要先打开一个DOS命令窗口。可以通过【开始→程序→附件→命令提示符】来启动这个窗口。

djangos-admin.py startproject mysite

为blog项目创建一个项目目录的djangos-admin.py命令是：

变量量里以保证它可以从此命令执行执行。
在UNIX上，它的默认安装路径为/usr/bin，在Win32上，它的位置是Python安装目录下的Scripts文件夹，比如C:\Python25\Scripts。无论是哪一个平台，你都要保证djangos-admin.py在PATH环境变量量里以保证它可以从此命令执行执行。
在Win32上，djangos-admin.py的命令用来帮助创建这样项目的目录。
组织Django代码最简单的方式是使用Django的“项目”(project)：一个包含了组成单个网站的所有文件的目录。Django提供了一个叫djangos-admin.py的命令来帮助创建这样项目的目录。

2.1 创建项目

且重复那些会导致错误的步骤还能帮你加深印象！
作者在学习Django时就用过这个办法，这样比无助地盯着错误信息看上几个小时要有效率，而且能帮助你理解出错的原因。如果还是没发现哪里不对，那就抛掉整个项目重新来过。
如果你确实打算在电脑上实际操作，同时遇到和你在这里看到的结果不同的情况时，请停
下来重新检查刚才的步骤，然后复习之前两到三步。看看是不是有些看上去不重要的地方，
或是哪些不能理解的步骤被遗漏了。如果还是没发现哪里不对，那就抛掉整个项目重新来过。
如果你确实打算在电脑上实际操作，同时遇到和你在这里看到的结果不同的情况时，停
下来重新检查刚才的步骤，然后复习之前两到三步。看看是不是有些看上去不重要的地方，
或是哪些不能理解的步骤被遗漏了。如果还是没发现哪里不对，那就抛掉整个项目重新来过。
我们建议你在阅读的同时实际动手构建这个blog。如果没有其他现代Web框架开发的经验
没有这个耐性)的话，简单读一下本章也是好的。特别是如果你有其他现代Web框架开发的经验
就更好了，因为很多基本的概念都是相似的。
如果你确实打算在电脑上实际操作，同时遇到和你在这里看到的结果不同的情况时，停
下来重新检查刚才的步骤，然后复习之前两到三步。看看是不是有些看上去不重要的地方，
或是哪些不能理解的步骤被遗漏了。如果还是没发现哪里不对，那就抛掉整个项目重新来过。
在Win32的命令行窗口里，你可以输入echo %PYTHONPATH%。你可以在第1章里读到更多关
于Mac OS X、FreeBSD等类UNIX的系统上，你可以用echo \$PYTHONPATH命令来查看其内容。
所以，首先打开终端，进入cd命令进入你的PYTHONPATH环境变量所指向的目录。在Linux、
本章所有工作都是在你选择的shell (bash、tcsh、zsh和Cygwin等) 下以命令行方式完成。

本章假设你已经安装了Django。如果没有的话，请参见附录B。

注意

Django自称是“最适合开发有限期的完美Web框架”。所以现在我们来挑战一下，看看到底
能在多短的时间里用Django完成一个简单的blog。(关于追求完美的部分我们稍后再细说。)

第2章 Django速成：构建一个Blog

Python而已。

startproject命令创建的所有文件都是Python的源码文件。这里没有XML、ini文件，或是任何时髦的配置语法。Django追求的是尽可能地保持“纯Python”这一理念。这让你在拥有诸多灵活性的同时无需在框架里引入任何复杂性。例如，如果你想让你的settings文件从其他文件导入设置，或是计算一个值而避免硬编码的话都可以执行无阻——因为它就是

注意

• URLs.py文件在Django里叫URLconf，它是一个将URL模式映射到你应用程序上的配置文件。URLconf是Django里非常强大的一个特性。

• urls.py文件在Django里叫URLconf，它是一个将URL模式映射到你应用程序上的配置文件。URLconf是Django里非常强大的一个特性。

• settings.py文件包含了项目的默认设置。包括数据库信息、调试标志以及其他一些重要的变量。你项目里安装的任何应用都可以访问这个文件。稍后在本章进行过程中我们会更多地看到它是什么可以执行的。我们马上就会用到它。

• manage.py文件是一个同这个Django项目一起工作的工具。你可以从此它在目录列表中的权

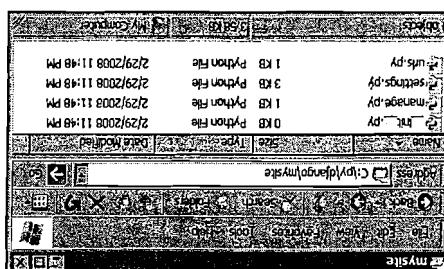
限了_init_.py，startproject命令还创建了以下三个文件：

子包的内容。)

如果精通Python的话，你一定知道_init_.py会把这个项目目录变成一个Python的包 (package) —— 相关Python模块的一个集合。这让我们可以用Python的“点记号”(dot-notation) 来指定项目中的某个部分，比如mysite.urls。(你可以阅读第1章，以获取更多关

注意

图2.1 Win32上的mysite文件夹



如果你是在Win32平台上进行开发，那么打开资源管理器窗口你就可以看到图2.1这样的文件夹，这里我们所创建的目录是C:\Pydjangomy，项目的內容将会存放在这里。

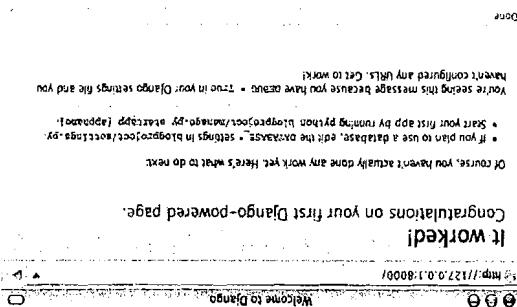
```
$ cd mysite
$ ls -l
total 24
-rw-r--r-- 1 pbsx pbsx 0 jun 26 18:51 __init__.py
-rw-r--r-- 1 pbsx pbsx 546 jun 26 18:51 manage.py
-rw-r--r-- 1 pbsx pbsx 2925 jun 26 18:51 settings.py
-rw-r--r-- 1 pbsx pbsx 227 jun 26 18:51 urls.py
```

log从左到右有4个部分：时间戳、请求、HTTP状态码，以及字节数。（你看到的字节数或

07/DEC/2007 10:26:37] "GET / HTTP/1.1" 404 2049

同时，如果你查看终端的话，你会看到dev服务器记录下了你的GET请求。

图2.2 Django初启的“It worked!”页面



在浏览器里输入这个链接，你会看到Django的“It worked!”页面，如图2.2所示。

在Win32平台上你应该能看到类似如下的输出，只是要退出时要按下的快捷键是Ctrl+Break。
Django version 1.0, using settings 'mysite.settings',
Development server is running at <http://127.0.0.1:8000/>,
quit the server with CONTROL-C.
0 errors found.
Validating models...
而不是Ctrl+C。

在Win32平台上你应该能看到类似如下的输出，只是要退出时要按下的快捷键是Ctrl+Break

./manage.py runserver # or "manage.py runserver" on Win32
settings文件。启动dev的命令如下：
manage.py工具，这是一个简单的包裹脚本，能直接告诉django-admin.py去读入项目特定的
运行开发服务器（或“dev”）简单到只需要一个命令就行了。这里我们要用到项目里的
• 知道如何为admin应用编写并显示静态的媒体文件，所以你就可以直接使用它。
设置来说都是很有必要的。
• 它会自动检测到你对Python源码的修改并且重新加载那些模块。相比每次修改代码后都
很方便了。
在一台新的服务器或是没有服务器环境的开发机上，或是就想实验证一下的话，那就非常
• 不需要安装Apache、Lighttpd，或是其他任何实际生产所需的Web服务器软件——如果你
中最好用的就是Django的内置Web服务器了。这个服务器不是用来部署公共站点，而是用来做
快速开发的。其优点在于：

到这里，你还没有构建完blog应用，不过Django为你提供了一些可以就地使用的方便。其

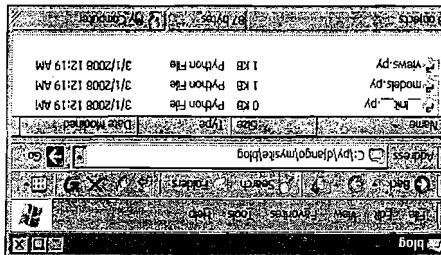
2.2 运行开发服务器

要告诉Django这个app是项目的一部分，你需要去编辑settings.py文件（也称之为“配置文件”）。打开配置文件并在文件尾部找到INSTALLED_APPS元组。把你的app以模块的形式添加到元组里，就像这样（注意结尾的逗号）：

```
mysite.blog,
```

和项目一样，app也是一个包。现在models.py和views.py还没有真正的代码，它们只是类文件而已。实际上对我们这个简单的blog来说，我们根本用不着views.py文件。

图2.3 Win32上的mysiteblog文件夹



```
-rw-r--r-- 1 ppx ppx 0 jun 26 20:33 __init__.py
-rw-r--r-- 1 ppx ppx 57 jun 26 20:33 models.py
-rw-r--r-- 1 ppx ppx 1 ppx ppx 26 jun 26 20:33 views.py
total 16
```

```
$ ls -l blog/
```

首先是UNIX下的，接着是Windows资源管理器里的截图（图2.3）。这样就完成项目的创建了。现在我们在项目的目录下有了一个blog目录。这是它的内容，manage.py来创建这个blog app。

```
./manage.py startapp blog # or ".\manage.py startapp blog" on win32
```

有了项目以后，就可以在它下面创建应用（按Django的说法是“app”）了。我们再次使用

2.3 创建Blog应用

如果你成功启动服务器后，我们就可以来设置你的第一个Django应用了。

如果服务器不工作的話，請重新檢查每一個步驟。不要怕麻煩！有時候直接關掉整個項目，然後重新開始可能會比勞心勞力地檢查每一個文件，每一行代碼要來的簡單。

提示

“worked!”页面只是Django委婉地告诉你这一点的方式。
你会发现有所不同。) 这里状态码404 (“Not Found”) 是因为还没有为项目定义任何URL。而 “It

2.4 设计你的Model

现在我们来到了这个基于Django的blog应用的核心部分：models.py文件。这是我们定义blog数据结构的地方。根据DRY原则，Django会尽量利用你提供给应用程序的model信息。我们先来创建一个基本的model，看看Django用这个信息为我们做了点什么。

用你最习惯的编辑器（如果它能支持Python模式就最好）打开models.py文件。你会看到这样的占位文本：

```
from django.db import models

# Create your models here.
```

删掉注释，加入下列代码：

```
class BlogPost(models.Model):
    title = models.CharField(max_length=150)
    body = models.TextField()
    timestamp = models.DateTimeField()
```

这是一个完整的model，代表了一个有3个变量的“BlogPost”对象。（严格说来应该有4个，Django会默认为每个model自动加上一个自增的、唯一的id变量。）

这个新建的BlogPost类是django.db.models.Model的一个子类。这是Django为数据model准备的标准基类，它是Django强大的对象关系映射（ORM）系统的核心。此外，每一个变量都和普通的类属性一样被定义为一个特定变量类（field class）的实例。这些变量类也是在django.db.models里定义，它们的种类非常多，从BooleanField到XMLField应有尽有，可不止这里看到的区区三个。

2.5 设置数据库

如果你没有安装运行数据库服务器的话，我们推荐使用SQLite，这是最快最简单的办法。它的速度很快，被广泛接受，并且将它的数据库作为一个文件存放在文件系统上。访问控制就是简单的文件权限。更多关于在Django里使用数据库的信息请参阅附录B。

如果你安装了数据库服务器（PostgreSQL、MySQL、Oracle和MSSQL），并且希望用它们而非SQLite的话，你需要用相应的数据库管理工具为Django项目创建一个新的数据库。在这里我们把数据库取名为“djangodb”，不过你可以选用任何喜欢的名字。

当你有了一个（空的）数据库以后，接下来只需要告诉Django如何使用它即可。这就需要用到项目的settings.py文件。

使用数据库服务器

很多人都选用PostgreSQL或MySQL这样的关系数据库和Django配合使用。这里有6个相关的设置（虽然你可能只要两个就够了）：DATABASE_ENGINE、DATABASE_NAME、DATABASE_HOST、DATABASE_PORT、DATABASE_USER和DATABASE_PASSWORD。它

们的作用从名字上就能看的出来。你只需在相应的位置填入正确的值就可以了。例如，为MySQL的设定看上去是这样的：

```
DATABASE_ENGINE = "mysql"
DATABASE_NAME = "djangodb"
DATABASE_HOST = "localhost"
DATABASE_USER = "paul"
DATABASE_PASSWORD = "pony" # secret!
```

注意

我们没有指定DATABASE_PORT是因为只有当你的数据库服务器运行在非标准的端口上时才需要那么做。比如，MySQL服务器的默认端口为3306。除非你修改了这个设置，否则根本不用指定DATABASE_PORT。

要知道创建新数据库和（数据库服务器要求的）数据库用户的细节，请参阅附录B。

使用SQLite

SQLite非常适合测试，甚至可以部署在没有大量并发写入的情况下。因为SQLite使用本地文件系统作为存储介质并且用原生的文件系统权限来做访问控制，像主机、端口、用户或密码这种信息一律统统不需要。因此Django只要知道以下的两个设置就能使用你的SQLite数据库了。

```
DATABASE_ENGINE = "sqlite3"
DATABASE_NAME = "/var/db/django.db"
```

注意

当SQLite配合Apache这样真正的Web服务器一起使用时，你需要确认拥有Web服务器进程的账号也拥有对数据库文件以及包含数据库文件目录的写权限。在我们这里用dev服务器做开发的情况下，这通常不是一个大问题，因为运行dev服务器的用户（你）同时也拥有项目的文件和目录。

SQLite在Win32平台上也是非常受欢迎的选择之一，因为它是随同Python一同免费发布的。假设我们已经为项目（和应用程序）创建了C:\py\django目录，现在再来创建一个db目录。

```
DATABASE_ENGINE = 'sqlite3'
DATABASE_NAME = r'C:\py\django\db\django.db'
```

如果你不太熟悉Python的话，你会注意到这个例子和前一个例子的不同之处，之前我们在sqlite3上用的是双引号，而在Win32的版本里，我们用的是单引号。这个和平台是没有关系的——Python没有字符串类型，所以单引号和双引号都是被等同对待。只要保证你用同样的引号引用字符串就可以了！

你还应该注意到了文件夹名字前的小写“r”。如果你没有读过第一章，那么你应该知道它的意思是这个对象是一个raw字符串，或者说它把字符串里所有字符都逐字保留下来，不对任何特殊字符组合做转义。例如，\n通常代表一个新行，但是在raw字符串里，它（字面上）代

表了两个字符：一个反斜杠和一个小写n。所以raw字符串的作用（特别是对这里的DOS文件路径来说）就是告诉Python不要转义任何特殊字符（如果有的话）。

创建表

现在你可以告诉Django用你提供的连接信息去连接数据库并且设置应用程序所需的表。命令很简单：

```
./manage.py syncdb      # or ".\manage.py syncdb" on win32
```

在Django设置数据库的过程中你会看到类似这样的输出：

```
Creating table auth_message
Creating table auth_group
Creating table auth_user
Creating table auth_permission
Creating table django_content_type
Creating table django_session
Creating table django_site
Creating table blog_blogpost
```

当你执行syncdb命令时，Django会查找INSTALLED_APPS里的每一个models.py文件，并为找到的每一个model都创建一张数据库表。（在稍后我们碰到华丽的多对多关系时会有例外，但是对这个例子来说是正确的。如果你用的是SQLite，你还可以看到django.db数据库会在你指定的位置上被创建出来。）

INSTALLED_APPS里的其他默认条目也都拥有model。manage.py syncdb的输出就确认了这一点，你可以看到Django为每个app都创建了一个或多个表。

但这还不是syncdb命令输出的全部。你还会被问到一些和django.contrib.auth app有关的问题。

```
You just installed Django's auth system, which means you don't have any superusers defined.
Would you like to create one now? (yes/no): yes
Username (Leave blank to use 'pbx'):
E-mail address: pb@e-scribe.com
Password:
Password (again):
Superuser created successfully.
Installing index for auth.Message model
Installing index for auth.Permission model
```

现在你在auth系统里建立了一个超级用户（就是你自己）。这在等一会我们加入Django的自动admin应用时就会变得很方便。

最后，这一过程还包裹了一些和特性fixture有关的代码，在第4章里我们在讨论这个。这让你在一个新建立的应用里预先载入数据。我们在这里没有用到这个特性，所以Django会跳过它。

```
Loading 'initial_data' fixtures...
No fixtures found.
```

删除第二行开头的 '#' 号（也可以同时删除最后一行）并保存文件。这样你就告诉 Django 去加载默认的 admin站点，这是被用于 contrib/admin 应用程序的一个特殊对象。

```
# ./manage.py syncdb
Creating table django_admin_log
Installing index for admin_LogEntry model
Installing index for admin_LogEntry model
Loading initial_data fixtures...
No fixtures found.
```

自动化的后台应用框架，admin模块上是Django“量级上的明珠”。任何对Web应用创建简单单的“CRUD”（Create, Read, Update, Delete）接口感到厌倦的人来说，这绝对是个好消息。我们在第11.1节里深入地介绍了admin。现在我们只要拿来自用就OK了。

由于这不是Django的必要组件，你必须在settings.py文件里指明你要使用它——就和指定blog app一样。打开settings.py并在INSTALLED_APPS元组里的django.contrib.auth下面添加这一行。

```
django.contrib.auth,
```

每次往项目里添加新的应用后，你都要运行一下syncdb命令确保它所必需的表已经在数据库里创建了。在这里可以看到在INSTALLED_APPS里添加admin app并运行syncdb命令会往我们

2.6 没置自功能admin应用

到此数据集的初始化就完成了，下次你再对这个项目运行 `syndb命令`（只要你在添加「应用或是model时就要求执行」时，就不会再看到那么多输出了，因为它不需要再次设置表或者提示你创建超级用户了。

图2.5 admin主页

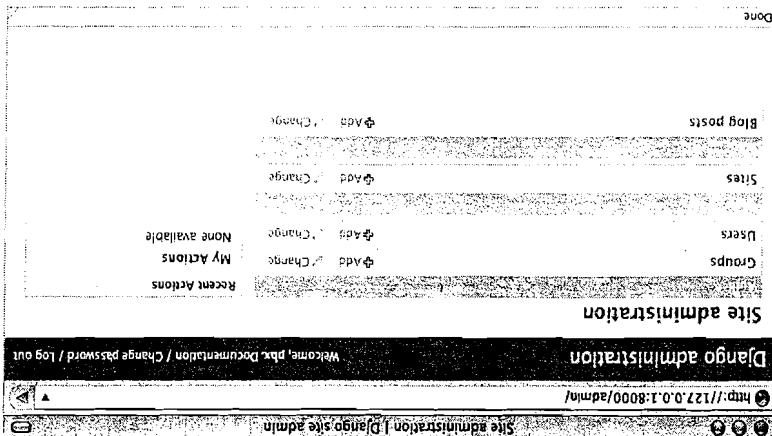


图2.4 admin登录页面

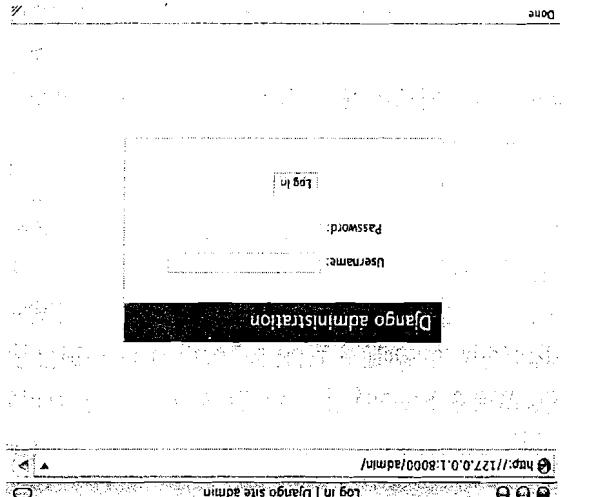


图2.5所示。

输入你之前创建的“超级用户”名字和密码。登录后，你就可以看到admin的主页了，如图2.4所示。再次输入用户名和密码，点击“Log in”，你将进入admin登录窗口，如图2.5所示。

到这里我们已经在Django里设置了一个名为admin的app，并向其注册了一个名为User的model，现在是时候试试它了。在后面的章节里你还会看到admin的高级使用，特别是本书第三部分和第四部分。

2.7 使用admin

这里admin的简单使用只不过冰山一角，它还可以通过为给定model建立一个特殊的Admin类来指定许多不同和admin有关的选项，然后向那个类注册model。这个我们在稍后再讲。

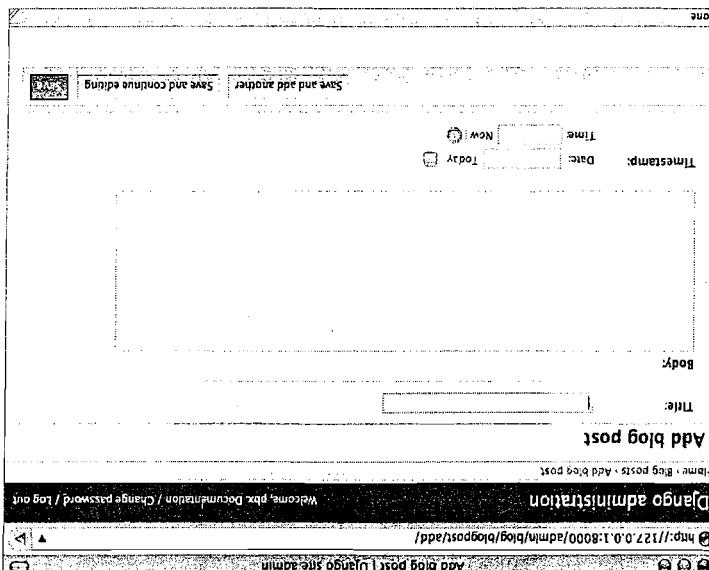
输出不会有任何变化——你一定不喜欢看见所有条目都被称为“BlogPost object”，如图2.8。你会看见页面上加入了另一个BlogPost。如果你刷新页面或是离开应用程序再回来的话，现在点击右上角的“Add Blog Post +”按钮添加另一篇不同内容的帖子。当你回到列表视图里，你会看到如何为你的对象的默认标签指定一个特定变量，或是特别定制的字符串。

为什么帖子有“BlogPost object”这么难看的名字？Django的设计是希望你能灵活地处理任意类型的内容，所以它不会去猜测对于给定的内容什么变量才是最适合的。在第三部分的例子中，你会看到如何为你对对象的默认标签指定一个特定变量，或是特别定制的字符串。

完成之后，点击Save按钮。你会收到一条确认消息（“The blog post ‘BlogPost object was added successfully.’”）和一个列出你所有Blog帖子的列表——目前只有一篇而已，如图2.7所示。

给你的帖子取一个名字然后往里填一点内容。至于时间戳，你可以点击Today和Now的快进链接来获取当前的日期和时间。你还可以点击日期或时间标志来方便地选择时间。

图2.6 通过admin添加新内容



有没有内容的Blog？点击Blog Posts右侧的Add按钮。Admin会显示一个表单让你添加新的帖子，如图2.6所示。

三个最常见的“我的app没有显示在admin里”的原因是：1) 它没向admin.site.register注册你的model类；2) models.py里有错误；以及3) 它记在settings.py中的INSTALLED_APPS里添加app。

如果没有的话，请重新检查之前的步骤。
本书稍后会解释这个界面，现在只需确认你的应用模型Blog和截图一样出现在屏幕上。

所示。你不是第一个这么想的人，“总有办法让它看起来顺眼一点吧。”

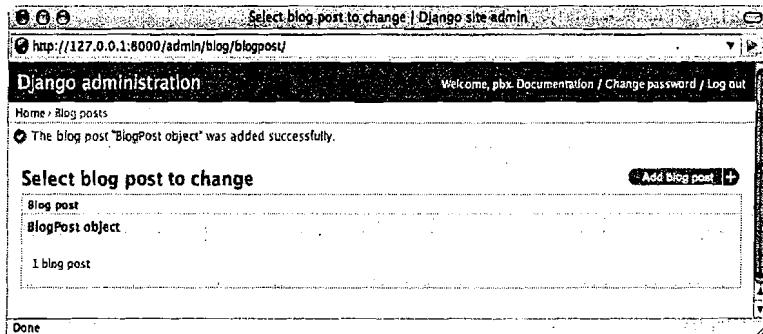


图2.7 成功地保存你的第一篇blog



图2.8 不太有用的小节页面

不过我们不需要等到那个时候来清理admin视图里的这些列表显示。之前我们通过很少的配置就激活了admin工具，即向admin app注册我们的model。现在只要额外的两行代码以及对注册调用的一些修改，我们就可以让列表看起来更漂亮更有用。更新你的mysite/blog/models.py文件，添加一个BlogPostAdmin类，并将它加到注册代码那一行里，于是现在models.py看起来是这样的：

```
from django.db import models
from django.contrib import admin

class BlogPost(models.Model):
    title = models.CharField(max_length=150)
    body = models.TextField()
    timestamp = models.DateTimeField()

class BlogPostAdmin(admin.ModelAdmin):
    list_display = ('title', 'timestamp')

admin.site.register(BlogPost, BlogPostAdmin)
```

开发服务器会注意到你的修改并自动重新加载model文件。如果你去看一下命令行shell，

```
<h2>{{ post.title }}</h2>
<p>{{ post.timestamp }}</p>
<p>{{ post.body }}</p>
```

模板：

Django的模板语言相当简单，我们直接来看代码。这是一个简单的显示单个blog帖子的

创建模板

。参数。

- 一个URL模式，它用来把收到的请求和你的视图函数匹配，有时也会向视图传递一些
- 一个视图（view）函数，它负责获取要显示的信息，通常都是从数据库里取得。
- 一个模板（template），模板负责将传递过来的信息显示出来（用一种类似Python字典的字典来说，一个页面具有三个类型的组件：

完成我们应用的数据库部分和admin部分后，现在来看看面向公众的页面部分。从Django

2.8 建立Blog的公共部分

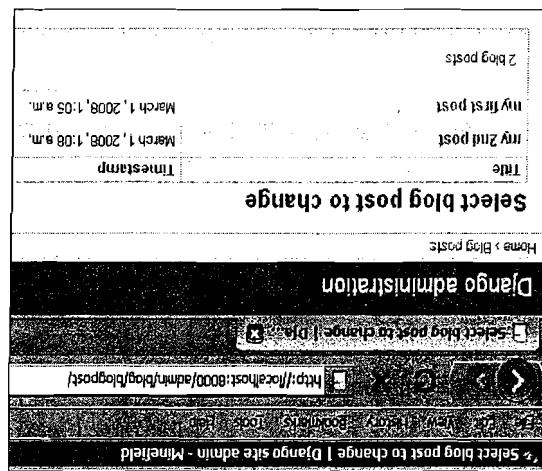
就像我们多次提到的，在本章第三和第四部分里会更详细地展示这些内容。

Admin还有很多有用的特性只需一到两行代码就可以激活：搜索、自定义排序、过滤等。

点一下Title会按照升序排列标题，再点一下则变成降序排列。

试着点一下Title和Timestamp列的标题——每一个都影响了你的条目是如何排序的。例如，

图2.9 这样就好多了



BlogPostAdmin类里list_display变量来生成输出的，如图2.9所示。

刷新一下页面，现在你就可以看到一个更有意义的页面了，它是根据你添加到
就能看到这些结果。

第4章) 的一个简单例子, 载取数据库里所有BlogPost对象。Django对 ForeignKey的全部力量。这一行只是使用ORM (对象关系映射, 详见第3章和第6行: 当我们把BlogPost类作为django.db.models.Model的一个子类时, 我们就获得了可以通过URLconf接受其他参数, 你将来会大量用到这个特性。

第5行: 每个Django视图函数都将django.http.HttpRequest对象作为它的第一个参数。它还可以通过import那几行 (它们载入了我们需要的函数和类), 我们来逐行解释一下这个视图函数。

```
先略过import那几行 (它们载入了我们需要的函数和类), 我们来逐行解释一下这个视图
returndjango.shortcuts.render(request, 'archive.html')
    c = Context({'posts': posts})
    t = loader.get_template('archive.html')
    posts = BlogPost.objects.all()
def archive(request):
    from mysite.blog.models import BlogPost
    from django.http import HttpResponseRedirect
    from django.template import Context, loader
    context = Context({'posts': posts})
    template = loader.get_template('archive.html')
    response = template.render(context)
    return response
```

现在我们来编写一个从数据库读取所有blog帖子的视图函数, 并用我们的模板将它们显示出来。打开blog/views.py文件并输入:

创建一个视图函数

```
mysite/blog/templates/archive.html
<h2>{{ post.title }}</h2>
<p>{{ post.timestamp }}</p>
<p>{{ post.body }}</p>
</div>
</div>
</div>
```

模板本身的名字是随意取的 (以foo.html也没问题), 但是templates目录的名字则是强制的。Django在默认情况下会在搜索模板时逐个查看你安装的应用程序下的每一个templates目录。

原来的3行没用, 我们只是简单地增加了一个叫做for的块标签 (block tag), 用它将模板渲染到序列中的每个元素上。其语法和Python的循环语法规是一致的。注意和变量标签不同, 其标签是包含在% ... %里的。

```
&lt;for post in posts &gt;
    &lt;div&gt;
        &lt;h2&gt;{{ post.title }}&lt;/h2&gt;
        &lt;p>{{ post.timestamp }}</p>
        &lt;p>{{ post.body }}</p>
    &lt;/div&gt;
&lt;/for &gt;
```

现在我们将稍微改进一下这个模板, 通过Django的for模板标签让它能显示多篇blog帖子。

这些是变量标签 (variable tag), 用于显示给模板的数据。在变量标签里, 你可以用Python风格的dot-ed-notatoin (点记号) 来访问传递给模板的对象的属性。例如, 这里假设你传递了一个叫“post”的BlogPost对象。这三行模板代码分别从BlogPost对象的title、timestamp和body变量里读取了相应的值。

这就是一个HTML (虽然Django模板可以用于任何形式的输出) 加上一些大括号里的特殊

- 第7行：这里我们只需告诉Django模板的名字就能创建模板对象t。因为我们把它保存在app下的templates目录里，Django无需更多指示就能找到它。
- 第8行：Django模板渲染的数据是由一个字典类的对象context提供的，这里的context c只有一对键和值。
- 第9行：每个Django视图函数都会返回一个django.http.HttpResponse对象。最简单的就是给其构造函数传递一个字符串。这里模板的render方法返回的正是一个字符串。

创建一个URL模式

我们的页面还差一步就可以工作了——和任何网页一样，它还需要一个URL。

当然我们可以直接在mysite/urls.py里创建所需的URL模式，但是那样做只会在项目和app之间制造混乱的耦合。Blog app还可以用在别的地方，所以最好是它能为自己的URL负责。这需要两个简单的步骤。

第一步和激活admin很相似。在mysite/urls.py里有一行被注释的示例几乎就是我们需要的代码。把它改成这样：

```
url(r'^blog/', include('mysite.blog.urls')),
```

这会捕捉任何以blog/开始的请求，并把它们传递给一个你马上要新建的URLconf。

第二步是在blog应用程序包里定义URL。创建一个包含如下内容的新文件，mysite/blog/urls.py：

```
from django.conf.urls.defaults import *
from mysite.blog.views import archive

urlpatterns = patterns('',
    url(r'^$', archive),
)
```

它看起来和基本的URLconf很像。其中的关键是第5行，注意URL请求里和根URLconf匹配的blog/已经被去掉了——这样blog应用程序就变得可以重用了，它不用关心自己是被挂接到blog/下，或是news/下，还是what/i/had/for/lunch/下。第5行里的正则表达式可以匹配任何URL，比如/blog/。

视图函数archive是在模式元组第二部分里提供的。（注意我们传递的不是函数的名字，而是一个first-class的函数对象。当然用字符串也行，你在后面会看到。）

现在来看看效果吧！开发服务器还在运行中么？如果没有，执行manage.py runserver来启动它，然后在浏览器里输入http://127.0.0.1:8000/blog/。你可以看到一个简单朴素的页面，显示了所有你输入的blog帖子，有标题、发布时间和帖子本身。

2.9 最后的润色

你有好几种方式来继续改进这个基本的blog引擎。我们来讨论几个关键的概念，把这个项

毫不夸张地说，我们的模板实在是太平庸了。毕竟这是一本关于Web编程而不是Web设计的书，美术方面的东西你就算自己处理吧，但是模板毕竟是模板系统里另一个能让你减少工作量的特性，特别是当你的页面风格快速成长的时候。

我们这个模板现在是自给自足的。但是如果我们站有一个blog，一个相册和一个链接页面，并且我们需要把这些都基于同一个基础风格的话该怎么办？经验告诉我们我们要用复制粘贴的办法做出三个几乎完全一样的模板肯定不行的。在Django里正确的做法是创建一个叫base.html的模板，然后在这之上扩展出其他特定模板来。在mysite/blog/templates目录里，创建一个叫base.html的模板，其内容如下：

```
<html>
  <head>
    <title>MySite Example</title>
    <meta name="description" content="A simple blog example."/>
    <meta name="author" content="by Carl" />
  </head>
  <body>
    <div id="content">
      <h1>{{ post.title }}</h1>
      <p>{{ post.content }}</p>
      <ul>
        <li>{{ post }}</li>
      </ul>
    </div>
  </body>
</html>
```

这里的是`% extends ... %`标签告诉Django去查找一个叫base.html的标签，并将这个模板里的命令块的内容填入到那个模板里相对应的块里去。现在你应该可以看到类似图2.10的页面了（当然了，你的blog内容比我这个应该更加精彩）。

模板的精确定位
目标得再漂亮一点。

一个很酷的特性：过滤器 (filter) 来把它弄得人性化一点。

虽然时间戳很好用，但是它的ISO8601格式却有点怪异。我们现在用Django模板系统里另

通过模板过滤器格式化时间戳

出现在前面。

在逗号后面加上title，并且你有两个相同发布时间的帖子“A”和“B”的话，“A”就会带小括号的字符串。Django在这里要的是一个元组，你可以非序性任意数目的变量。如果你

千万别忘了小括号里结尾的那个逗号！它代表这是一个单元素的元组，而不是一个

注意

话则是按升序排列。)

现在看一下blog的首页(/blog/)。最新的帖子应该出现在页面最上方了。字符串“-

ordering = ('-timestamp',)

class Meta:

设置model默认排序的方法是给它定一个Meta类，然后设置ordering属性。

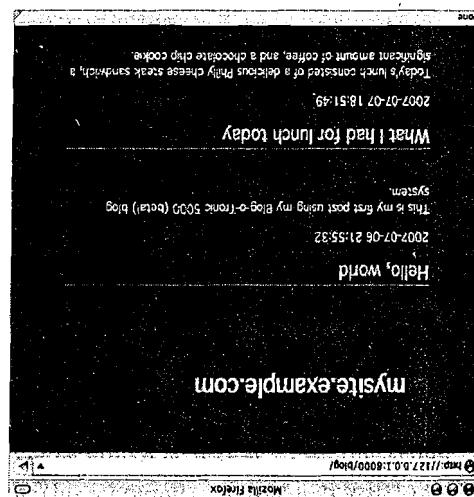
部分都会采用这个排序结果。

如果在model里设置我们想要的排序方式，Django里任何访问数据的代码里的BlogPost.objects.all()上添加排序功能。在这里修改model会比较好，因为基本上帖子都是按时间倒序排列的。如果在model里设置我们想要的排序方式，Django里任何访问数据的

实面上，有好几种方法可供选择。我们可以在model里加入一个默认的顺序，或者在视图单。

按日期排序

图2.10 模板带点风格的blog



本的“live blog”。

- 继续构建示例应用了：一个内容管理系统（CMS）、一个Pastebin、一个相册和一个基于Ajax的小章的经验教训总结。第4章将带你进入框架的细节，你会找到很多之前例子中碰到的关于的设计师哲学，以及简要地复习一下Web开发原则，它们不仅仅是对Django本身，而且也是对书中讨论过的“通用视图”来避免在views.py里编写任何代码（参见第10章里的Pastebin示例）。
- 添加一个搜索功能，让用户能迅速定位包含特定词语的blog帖子（参见第8章里的CMS例子）。
 - 发布关于最新帖子的Atom或RSS Feed（参见第11章）。

概念：

下面是一些可以透过Django内建的特性就能添加到blog上的功能，这能让你有个基本的留在其所属的应用程序内部。

- URLconf系统，在给予你URL设计极大的灵活性的同时还能将应用程序特定的URL部分保留在模板过滤器，可以在不解释应用逻辑商业逻辑的前提下改善表示的数据显示（比如日期）。
- 模板系统，可以用来生成HTML、CSS、JavaScript以及任何文本输出格式。
- 自动化的admin应用组件，提供了完整的内页编辑特性，甚至非技术背景的用户也能使用。
- 数据模型的创建采用纯Python方式完成，你不用编写维护任何SQL代码或XML描述文件。
- 内置的Web服务器能让你的工作自给自足，同时它能在代码被修改时自动重新加载你的代码。

我们就可以给这个blog引擎不停地加入新特性（很多人就是这么干的！），但是作为体验好几个Django优雅、简洁的特性：

Django能力的话，希望你已经看的差不多了。在构建这个基本的blog app过程中你已经看到了

2.10 总结

空格——Django的模板引擎对空格敏感。

这个格式化字符串会返回“Friday, July 6th”风格的日期。注意这里不要在冒号两侧留有

```
<p>{{ post.timestamp|date:"I, F jS" }}</p>
```

说明。例如，如果你要显示星期几，但是不要显示年份的话，代码就变成下面这个样子了：
如果date过滤器默认以风格式不喜欢的话，你可以传递一个strftime风格的格式化字符串为它的参数。不过这里它用的不是Python里time模块的转换代码，而是PHP的date函数的格式化它的参数。

用于变量之上。刷新blog首页。现在你可以看到日期的显示更加友好了（“July 7”）。

只要你这样用一个竖杠，或“管道”符接在变量名后面（大括号内部）就能把过滤器应

```
<p>{{ post.timestamp|date }}</p>
```

由于这是一个表示层的(presentation)细节而非数据结构或是商业逻辑的细节，最适合它的位置应该是模板。打开archive.html文件并修改“post.timestamp”一行。

与任何大型软件项目一样，Django包含了大量的概念、特性和工具，而且它所要解决的问题（即Web开发）的范围也是相得的。在开始学习使用Django的细节之前，你先得理解这些问题是什么，以及像Django这样的框架是怎么解决它们的。

这一章将会逐一介绍这些基本概念，首先是对于Web的一个总体认识，然后是Web框架模型及其组成部分的解释，最后是Django的创造者对软件开发的一些普遍的看法和理念。上一章里的一些概念虽然已经有些老套了——但即使是再有经验的Web工程师也可以温故知新嘛。有时候之所以不识庐山真面目，主要是因为我们的思想都被某种特定的语言或工具束缚住了，要是看问题的时候站的高一点，就能看得更清楚些了。

这里有一个重要的提示：如果你对Web开发相当熟悉的话，这里的某些概念已经有些老套了。Web开发理论上来讲还是很简单的。用户向Web服务器请求一个文档，Web服务器随即获取生成这个文档，服务器再把结果返回给用户的浏览器；最后浏览器将这个文档渲染出来。细看HTTP（超文本传输协议）封装了Web页面服务的整个过程，它是Web的基石。由于这个协议是面向客户端/服务器端通信之用，所以它主要就是由请求（request，客户端发送到服务器端）和响应（response，服务器端到客户端）两个部分组成。而两者之间的服务器上发生的事情都不属于HTTP关心的范畴，完全是由服务器软件决定的（看下面）。

请求封装了过程的第一部分——客户端向服务器端要求一个给定的文档。请求的核心就是URL（指向所需求文档的“路径”），当然它可以通过一系列方法进一步参数化，让单个的地址或URL展现多种行为。

响应主要是由一个正文（body，通常是以Web页面的文本）和相应的包头（header）组成。包头里是关于所需求的数据的额外信息，例如最后更新的时间，在本地可以缓存多久，内容的类型等等。另外，响应里非HTML的内容可以是纯文本，文档（PDF，Word，Excel等），声音片段等。

Django将请求和响应表示成相对简单的Python对象，用属性来表示其数据，以及用方法来进行更复杂的操作。

3.1 动态网站基础

一章！

本章探讨这些基础知识能让你在设计和实现时做出更好的选择。所以，请不要跳过这一章！

Web开发理论上来讲还是很简单的。用户向Web服务器请求一个文档，Web服务器随即获取生成这个文档，服务器再把结果返回给用户的浏览器；最后浏览器将这个文档渲染出来。细看HTTP（超文本传输协议）封装了Web页面服务的整个过程，它是Web的基石。由于这个协议是面向客户端/服务器端通信之用，所以它主要就是由请求（request，客户端发送到服务器端）和响应（response，服务器端到客户端）两个部分组成。而两者之间的服务器上发生的事情都不属于HTTP关心的范畴，完全是由服务器软件决定的（看下面）。

请求封装了过程的第一部分——客户端向服务器端要求一个给定的文档。请求的核心就是URL（指向所需求文档的“路径”），当然它可以通过一系列方法进一步参数化，让单个的地址或URL展现多种行为。

响应主要是由一个正文（body，通常是以Web页面的文本）和相应的包头（header）组成。包头里是关于所需求的数据的额外信息，例如最后更新的时间，在本地可以缓存多久，内容的类型等等。另外，响应里非HTML的内容可以是纯文本，文档（PDF，Word，Excel等），声音片段等。

Django将请求和响应表示成相对简单的Python对象，用属性来表示其数据，以及用方法来进行更复杂的操作。

粗看起來，Web的作用就是传输数据或内容共享（这里的內容可以是任何东西——blog帖子，金融数据，电子书等）。在早期的Web里，内容都是手工编写的HTML文件，一般都是保存在数据库的文件系统上。这叫做静态（static）内容，因为同样的URL请求返回的信息总是一样的。之前描述的“路径”在这里相对简单，一般都没有参数，毕竟它只不过是指向服务器文件系统上的静态内容而已。不过今时不同往日，现在大多数的内容都是动态的了，因为一个给定URL可以根据参数的不同返回完全不一样的数据。

大部分这种动态的特性是通过将数据保存在数据库里的，数据库里，数据不再是简单的文本字符串，你可以创建有多个组成部分的数据，并将它们连在一起来表达其中的相互关系。SQL（结构化查询语言）是用来定义和查询数据库的语言，通常被进一步抽象为一个ORM（对象关系映射），它可以把数据库里的数据映射为面对对象语言里的代码对象。

SQL数据库按照表（table）的形式来组织，每张表由行（row，例如，条目、对象）和列（column，属性、变量）组成，基本上和数据表格很相似。Django提供了一个强大的ORM机制，Python的类就代表了表，对象代表了其中的每一行，而对象的属性则代表了列。

Web开发里的最后一个部分就是如何表示或者格式化用户所请求的信息，经由HTTP返回的信息，以及SQL数据库查询所返回的信息。通常，结果都是表示成HTML（超文本标记语言）或更XML一点的XML，並且配合负责完成浏览器端功能的JavaScript和视觉效果的CSS（层叠样式表）。再新一点的应用还会用JSON（一种“轻型”数据格式）或XML来完成动态内容。

大多数Web框架都提供了模板语言（template language）来处理要显示的数据，它混合了原始的HTML标签以及一些类似编程的语法来循环对对象集合，执行逻辑操作和其他一些动态行为所必需的结构。一个简单的例子就是一份静态HTML文档加上一点点逻辑来显示当前登录用户的用户名，或者在没有用户登录的情况下显示一个“登录”的链接。

有些模板系统希望通过把它们自己的命令实现为HTML属性或是标签来试图完整地兼容 XHTML，所以文档生成的结果可以按照普通HTML那样来解析。其他一些则更贴近常规的编程语言，有时候还会将程序结构用特殊的符号包裹起来以便阅读和解析。Django的模板语言就属于后者。

组合在一起

面对这种。

尽管Web被分成前面说到的三个组成部分，但是有一个关键的地方被忽略掉了：它们之间是如何相互交流的。Web应用程序是怎么根据请求知道要去执行一个SQL查询的，以及它是如何知道要使用哪一个模板来渲染结果的呢？答案是这要看用的是什么工具了：每个Web框架或语言都有不同的方法。不过通常相似的地方总要比不同的地方多，所以虽然下面的两节讲的是Django自己的方式，但是这些概念在其他框架里也可以找到。

3.2 理解模型、视图和模板

就像你刚看到的那样，Web开发经常被分割成几个核心的组件。在这一节里，我们要进一步拓展这些概念，讨论一些编程方法论，然后总览一下Django是如何实现它们的（并在后面的章节里展示细节是实例）。

分层 (MVC)

将（Web或其他的）动态应用程序分层这种思想已经存在很久了，通常都是将其应用在图形化的客户端应用上，学名叫MVC（Model-View-Controller，模型-视图-控制器）范式。即，应用程序被分割成模型（控制数据），视图（定义显示的方法），以及控制器（在两者之间斡旋，并且让用户可以请求和操作数据）。

把一个应用程序以这种风格分成几个部分可以给予程序员足够的灵活性，并且鼓励重用代码。例如，对于一个给定的视图，假设这个模块知道如何图形化数字类型的数据，那么只要中间的胶水代码能把它和数据联系起来，它就可以用在各种不同的数据集上。或者一个特定的数据集可以用多种不同的输出格式来显示，例如刚刚提到的图形化视图，纯文本文件或是可排序的表格等。多个控制器可以根据用户的不同对同一个数据模型做出不同程度的访问控制，或是允许通过GUI应用以及email或是命令行来提交数据。

成功实施MVC架构的关键在于要正确地分割应用程序的不同层次。虽然在某些情况下，在数据模型里存放如何显示它的信息是贪图一些方便，但是却会给将来替换视图带来极大的困难。同样，在图形布局的代码里放置数据库相关的代码会在替换数据库平台的时候让你头痛不已。

Django的办法

Django也遵循了这一分层的原则，不过在做法上却略有不同。首先模型部分保持不变：Django的模型层只负责把数据传入传出数据库。然而Django里的视图却并不是显示数据的最后一步——Django的视图其实更接近MVC里传统意义上的控制器。它们是用来将模型层和表示层（由HTML和Django的模板语言组成，稍后在第6章里会介绍到它）连接在一起的Python函数。按Django开发团队的话来说就是：

我们理解的MVC里，视图的作用是描述将要显示给用户的 data。这不仅仅是数据看上去的外观 (look)，还包括如何表示数据 (present)。视图描述的是你能看哪些数据，而不是怎么看它。这里面的区别很微妙。

换一种说法，Django把表示层一分为二，视图方法定义了要显示模型里的什么数据，而模板则定义了最终信息的显示方式。而框架自己则担当了控制器的角色——它提供了决定什么视图和什么模板一起响应给定请求的机制。

模型

任何应用程序的基本，不管是Web应用，都是它所展现、收集和修改的信息。因此，若将应用程序分层，模型 (model) 将是最底部的一层，它是基础。视图和模板可以根据数据

本章到目前为止，我们已经讲解了组成实际Django系统的主要架构组件以及外围的一些小部件。现在我们可以把它们放到一起，给出一个总览。如图3.1，你可以看到最接近用户的是一些HTTP通信协议。通过URL，他们可以向Django的Web应用发送请求，并且在Web客户端接受响应。

3.3 Django架构总览

虽然HTML是最常见的格式，模板实际上并不一定要它——模板可以生成任何文本格式，甚至CSV，甚至是email消息正文。重要的是它们让一个Django项目能够把数据的表现和决定显示什么数据的视图代码区分开来。

所以大多数Django程序员都使用它的模板语言来渲染HTML页面。基本上，模板就是一些输出动态的经过特殊格式化的HTML文本，支持简单的逻辑结构如循环等。当一个视图要返回一个HTML文件时，它通常会指定一个模板，提供给它所要显示的信息，并在响应里使用模板渲染的结果。

因此在这里引入一个抽象是非常重要的。后退回它，这样的做法也不算是很落伍。只不过在绝大多数情况下，这样做的效率实在太低，而且这回一个HTTP响应的话，那么是正确的——你可以在Python里写出一个字符串当然方法只是返回一个HTTP响应的话，那么是正确的——你可以在Python里写出一个字符串。

我们刚刚说视图的作用是负责显示来自模型的数据。这并非百分之百的正确。如果说视图

模板

Django为这类任务提供了很多方便快捷的帮助函数，但是你也可以选择自己编写一切来完全地控制整个过程，或是大量使用这些快捷方式来进一步快速操作，抑或是两者结合。灵活性和强大都是模板的优点。

或者是在模型里添加这样的新对象，加上一些额外的工作诸如检查应用程序用户验证的状态，并且在这一步通常只有一些简单的任务需要完成，例如显示一个或是一列从模型里取得的对象，HTTP响应对象。在Django的HTTP机制两端之间要执行什么操作完全都在你的控制之下。实际上却很简单：它们是链接到一个或多个定义URL上的Python函数，这些函数都返回一个视图（view）组成Django应用程序里很多（有时候几乎完全是全部）的逻辑。它们的定义其实

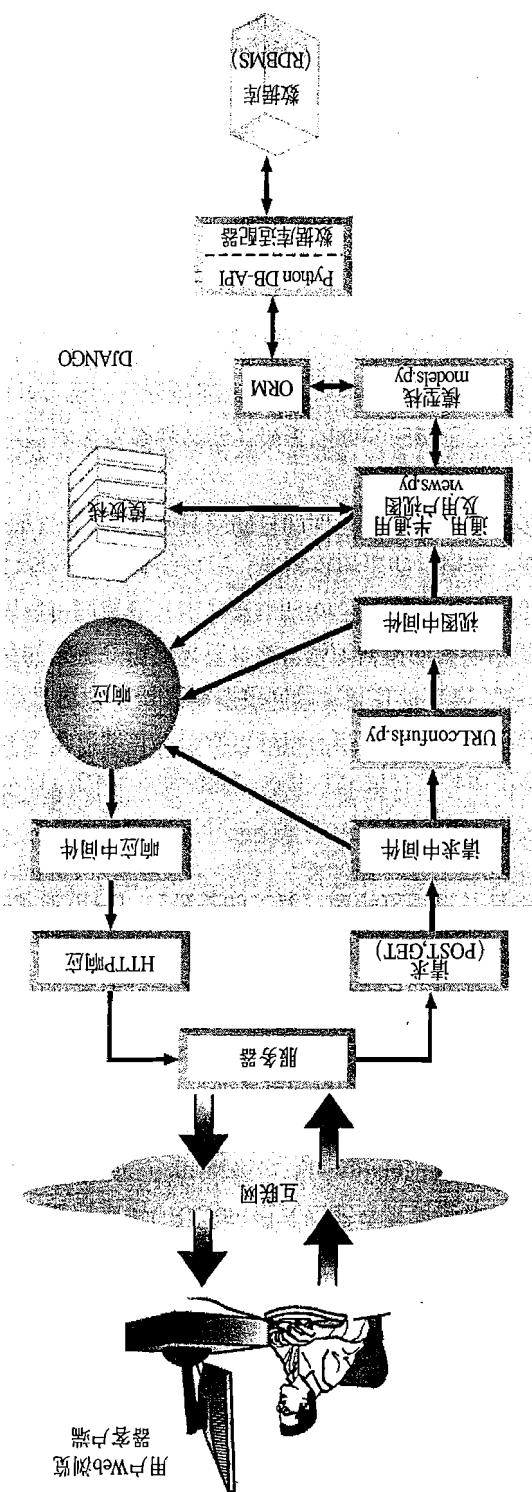
视图

这里的模型产生不良的副作用。在后面设计示例应用的章节里，我们遇到很多这类现实问题。你“只是在修改代码”，但实际上确实正在修改底层的数据存储方式，这经常会对已经存储在数据库里的数据产生不良的副作用。其中一个就是在应用程序部署后修改模型的代码。虽然表面上看模型却不一定总是最高效的。

从设计整个Web应用的角度来说，模型可能是最容易领会却也是最难掌握的部分。在面向对象系统里对一个现实问题进行建模相对来说通常不难，但是对大量的网站来说，最好合实际的模型却不一定总是最高效的。

进出模型的方式以及表现的形式在意转换，但模型却相对稳定得多。

图3.1 Django组件架构图



应，同时客户端可能还用到了Ajax技术，利用JavaScript来处理一些非顺序性的服务器访问。

在另一端（图的底部），你可以看到你的模型（model）和Django ORM管理下的数据库。它们通过Python的DB-API以及作为连接器的数据仓库客户端（通常是以C/C++编写的库）与数据库通信。

最后，两者之间的正是应用程序的心脏，Django。MVC模型在Django的术语里应该是“MTV”。视图（View）是作为控制器通过ORM负责从数据库里创建、更新和删除数据模型。同时根据给定的模板管理最终用户看到的结果。

把所有部分拼起来就是这样，收到的HTTP请求被Web服务器转交给Django，Django在请求的中间件层接受它们。随后根据URLconf模式匹配分派到适合的视图上去，视图会执行所需求的工作的核心部分，用模型（model）和/或模板（template）技需要生成响应。随后响应再通过中间件层进行最后的处理，最后将HTTP响应返回给Web服务器并转发给用户。

3.4 Django的核心理念

Django希望尽量Pythonic一点。使用编程语言的社区通常能对软件项目的设计产生重大影响的因素之一，Django也不例外。使用编程语言的社区通常能对软件项目的设计师产生重大的影响的因素之一，Django也不例外。虽然对这个圈子没有官方的定义，但通常这代表代码所展现的属性应该符合语言的各个方面。

这些属性包括有使用精悍的语法（比如默认的for循环或是更简洁的列表推导式）等。即通常对每种简单任务，只有一种正确的做法（而这里所谓的“做法”通常都已经被结合到语言里了，比如字典的get方法），以及明确自己的目的而不要依赖默认为（例如所有对象方法都要求的self参数）。

如同在第三部分的示范章节里要看到的，很多Django的规定，方法和设计都是或努力接近Pythonic的。因此有Python经验的程序员很容易入门，同时也能够帮助新手养成良好的编程习惯。Django的每一个用户界面都是一个对几乎所有的编程工作来说都非常重要的原则：DRY，不要重复自己。DRY可能是所有编程惯例（programming idiom）最简单的，因为它实在是个常识罢了：如果你要修改在多个地方重复出现的信息，你就平白无故给自己制造了两倍（甚至更多）的工作量。

作为DRY的一个例子，要对一些数据进行一些简单的运算，比如对某个人名下的所有银行账户加和。在设计糟糕的系统里，这种加和运算可能会出现在整个代码里的多个位置，显示用户操作的一部分。每个用户详细信息的页面、或者是显示多个用户综合的页面。在Django ORM这样的页面上，要对一些数据进行一些简单的运算，比如对某个人名下的所有银行账户加和。在设计糟糕的系统里，这种加和运算可能会出现在整个代码里的多个位置，显示用户操作的一部分。每个用户详细信息的页面、或者是显示多个用户综合的页面。在Django ORM这样的

第3章 起始

Django是一个功能完整的Web框架，提供了所有动态Web应用所需要的这一切必要组件：数据库、请求响应、应用框架、应用组件逻辑、模板系统等。同时，它也努力保持开放：你可以根据需要选择或减少Django组件并随时替换成自己的。

Django在编写的时候就考虑到了快速和敏捷开发。工作在一家快节奏的本地报馆里，Django的核心团队需要一组能在极短的时间里实现功能的工具。开源整个框架并没有改变它擅

Django在不同层面上提供了很多快捷方式。这里最明显的就是刚刚提到的通用视图，它由Django在编写时提供的许多组成了。Django在Python视图层上也有许多常见任务提供了快捷的方法，所以当通用视图无法满足程序员的需要时，他们仍然能避免自己动手。这样的快捷方式用数据字典来模板化，换取数据

框架的映射（以及一些HTML模板），就可以在几分钟内完成一个网站了。

Django的许多功能都是先哲时不要关心Django的这种模块化方式。因此，对

Django Web开发的新手来说，最好还是先哲时不要关心Django的这种模块化方式。不过毕竟这样的模块化还是有代价的：有些Django拿出的操作的特性却一定要把整个项目捆在一起才能使用，比如通用视图方法，它能让你轻易地显示、更新和创建数据库记录。因此，对

Django的ORM而转去使用这些工具。相反地，(虽然不太常见) Django的ORM也可以被单独拿出来用于其他项目(甚至非Web应用)，只需要少的几个步骤就能让它工作起来。

Django的ORM面对其他用户比喜好SQLAlchemy或是SQLObject的话，完全可以完全无视数据库层也是一样。如果用户比较喜欢SQLAlchemy或是SQLObject的话，可以完全无视

视图方法并不一定使用Django的模板系统，所以你的视图完全可以让加载Kid或是Cheethah，

例如，有些用户不喜欢Django的模板系统而比较偏爱Kid或Cheethah这样的选择。Django的视图方法这三套纯编写的模板，并将其作为Django的视图的一部分返回。

Django是一个功能完整的Web框架，提供了所有动态Web应用所需要的这一切必要组件：

Django在上面这样的简单情况下应用DRY的难度并不大，它却也是最难在任何时候都严格遵守的戒律之一。很多时候它会和其他(Pythonic等)惯例发生冲突，这时就必须做出妥协。不过，大多数时候这都是值得努力接近的目标，而且随着经验的不断丰富遵守它也会变得容易起来。

的系统里，你只需为Person类创建一个sum_accounts方法就可以轻易遵循DRY原则了，这个方法只需要定义一次，然后就能在上面说到的所有地方使用。

Django与灵活性

这一章涵盖了很[多基础内容](#)：什么是Web开发，Django及类似框架创建网站的方式，以及推动Django自身开发和设计背后的理论。不管你还记得多少，我们希望你从这一章里获得一些什么。

到此，你应该已经了解了开发Web应用的基础背景，以及一个典型的Web框架背后的理论及其组织方式。在第二部分里，我们会详细讲解如何使用Django，讨论它用到的各种类、函数和数据结构，并且向你展示更多代码示例来帮助你理解它们。

3.5 总结

Django和其他大多數現代Web框架（以及其他應用程序開發工具）一樣，依赖于一个强大的數據訪問層，它試圖將底層的關係數據庫和Python的面向對象聯繫起來。這些ORM在開發社區裏仍然是一个爭議很大的話題，有支持的也有反對的。Django在設計之初就決定要使用ORM，我們有4個使用它的理由，特別是Django自己的實現。

本章稍后会介绍到的Django模型对象是定义一组变量的首选方式，而变量通常对应的是一個id为5的Author对象并检查它的author.name值，而不必去写SELECT name FROM authors WHERE id=5这样的SQL查询——比较起来这种类型的數據接口要Pythonic得多。

而且，模型对象能给那个简陋的例子增加许多额外的价值。和很多其他系统一样，Django的数据模型能给那个简陋的例子增加许多额外的价值。很多其他系统一样，Django

为什么使用ORM

admin应用。

Django的数据模型层大量使用了ORM（对象关系映射），理解这个设计背后的原因为什么，以及很多其他应用程序开发工具）一样，依赖于一个强大的数据访问层，它试图将底层的关系数据库和Python的面向对象联繫起来。这些ORM在这个方法优缺点是非常重要的。所以，在一节里我们首先解释一下Django的ORM，然后我们及这个方法优缺点是非常重要的。所以，在一节里我们首先解释一下Django的ORM，然后我们

4.1 定义模型

在第3章中解释过，数据模型通常是在Web应用程序的基础，从这里开始探索Django开发的细节非常合适。虽然这一章有两个主要的部分（定义模型和使用模型），但它们不是独立的小节，而是更多地交织在一起。我们需要在定义模型时就要思考好怎么使用它们，以便生成最有用的类和关系布局。当然，如果没有理解怎么定义和为什么要定义模型的话，你是无法充分利用它们的。

第4章 定义和使用模型

第6章 模板和表单处理

第5章 URL、HTTP机制和视图

第4章 定义和使用模型

第二部分 深入Django

- 你可以定义只读的变量或属性组合，有时候也称之为“数据聚合”（data aggregation）或者“计算属性”（calculated attribute）。例如，一个具有count和cost属性的Order对象可以暴露一个计算两者乘积的total属性。常用的面对对象设计模式因此变得相当容易——比如此外观模式（facades），或是代理模式（delegation）等。
 - Djangos的ORM允许重写内置的数据修改方法，例如保存和删除对象。这样你就可以在数据被保存到数据库之前轻易地对它进行任何操作，或者在删除一条记录之前确保特定的操作语句集成功能（与Django来说，自然就是Python）通常比较简单，可以让你的数据库开发，这张表会越来越长。
 - 可移植性
 - 安全性
 - 表现力
 - 基础和模型的定义并没有直接的联系，然而使用ORM最大的一个好处（和直接编写SQL相比，当然也是最大的不同之处）就是从数据库取记录时使用的查询语法。这类似于直接地在模型类中编写SQL查询语句所导致的问题，例如SQL注入攻击等。ORM还提供了一个智能化的机制来帮助你避免编写错误的查询语句所导致的问题，所以你不必再担心由不合适的或者是错误地组织在一起的代码带来的安全问题。
 - 基础和模型的定义并没有直接的联系，然而使用ORM最大的一个好处（和直接编写SQL相比，当然也是最大的不同之处）就是从数据库取记录时使用的查询语法。这类似于直接地在模型类中编写SQL查询语句所导致的问题，例如SQL注入攻击等。ORM还提供了一个智能化的机制来帮助你避免编写错误的查询语句所导致的问题，所以你不必再担心由不合适的或者是错误地组织在一起的代码带来的安全问题。
- ```

from django.db import models
class Book(models.Model):
 title = models.CharField(max_length=100)

```
- Django丰富的变量类型
- Django的模型拥有多种不同的变量类型，有些和它们在数据库里的实现比较接近，有些则完全方丈模型类。不过这里我们比较性地列出一些最常用的用户变量。首先，我们给出一个基本的Django模型类。

这个例子的作用相当明显：我们为book创建了一个简单的模型，它包含了好几个和数据库相关的概念。不是很复杂（通常这类绘图书分类的工作要考虑到的东西可不容易），作者和页数在这里够用了。而且它完全可以直接使用，把它扔到Django的models.py文件里，你就可以看到，Django用Python的类来表示对象，而对象则通常映射到SQL中的表，对象的属性则是表中的列。这些属性本身也是对象，它们都是Field类的子类。我们来看一些特定的Field类和SQL的列类型很相似，其他的则都提供了某种程度的抽象。我们中的有属性则是表中的列。这些属性本身也是对象，它们都是Field类的子类。前面说到，它们中的有些类是通过常映射到SQL中的表，对象的属性则可以是一样的——作用都是保存文本。区别在于CharField是定长的，而TextField的长度则可以是无限的。具体使用哪一个要看需要，包括数据库的全文搜索能力或是高效存储的需求。

- CharField和TextField：这可能是你会遇到的最常用的变量类型了，这两个类型基本上是一样的——作用都是保存文本。区别在于CharField是定长的，而TextField的长度则可以是无限的。具体使用哪一个要看需要，包括数据库的全文搜索能力或是高效存储的需求。
- BooleanField和NullBooleanField：BooleanField大多适用于你要存储True或False值的场景，但是有时候你还不知道要在储哪个值的时候就需要更大的回旋余地——这时变量可以是空的或null的，所以就出现了NullBooleanField。这个特性体现了有时侯在为数据建模的时候不但要在语义层面考虑问题，同时还需要有技术层面的考量——即不仅要关心数据于验证的更多细节)。
- EmailField、URLField和IPAddressField：这三个变量其实就是CharField加上一点点额外的验证。它和CharField一样存储在数据库里，但是包含了验证码代码来保证它们的值分别是有效的E-mail地址、URL和IP地址。你也可以很轻松地在模型变量上加入验证码来创建出你的账户。它们主要的区别在于它们的值分别是自己的“变量类型”，这些类型和Django的内置类型是平等的。(参见第6章及第7章里关于验证的更多细节)。
- BooleanField和NullBooleanField：BooleanField大多适用于你要存储True或False值的场景，但是有时候你还不知道要在储哪个值的时候就需要更大的回旋余地——这时变量可以是空的或null的，还要关心它的意义。

• FileField：FileField是最复杂的变量之一，因为大多数时候所有和它有关的工作和数据库文件的操作，如它的连接FilePathField很像，但是它允许用户从浏览器上传一个文件，并将它保存在服务器上。它还在模型对象里提供了访问网页URL的方法来接受要上传的文件。这些只是Django的模型定义里常见的变量类型，当新版的Django发行时，新的变量类型也会时不时地被添加进来或是更新。要查看完整最新的模型变量类的列表以及可以进行的操作，可以查阅Django的官方文档。你还会在本书中看到很多这些变量，比如小代码段或是第三部分中的示例应用。

确保表里面每条记录都有一个唯一的标识符的方法既简单又有效。

关系数据库定义里的一个常见的概念就是主键 (primary key)，这个变量保证在整个表中唯一。如果给定一个主键和唯一性。

Django对模型的使用相当普遍，除非你明确规定，否则Django会为你自动生成。所有没有显式指定为主键的模型都会被指定为一个ID属性，其类型为Django的AutoField（一个自增的整数）。AutoField和普通整数一样，而它们在数据库中的类型则是根据你使用的数据库后会有不同。

ID为5的Book对象并且保证它是唯一拥有这个ID的Book时，引用“第5号book”就不会有意义了。

如果想要引用自己的话，也可以把ForeignKey指定为字符串'self'。这在定义层次结构（例如，

```
class Author(models.Model):
 name = models.CharField(max_length=100)
```

```
class Book(models.Model):
 title = models.CharField(max_length=100)
 author = models.ForeignKey("Author")
```

新组织编写上面的例子：

请注意你需要把被引用的类定义在前面，否则Author变量名是无法在Book类的ForeignKey变量里使用的。不过，你可以用字符串来代替，如果类是在同一个文件里定义的，那就只需要重命名使用。另外，你也可以用点记号（比如，`myapp.Author`）。这里在ForeignKey里使用字符串类名，否则的话就需要使用点记号（比如，`myapp.Author`）。这在 ForeignKey里使用字符串重命名的。

```
class Book(models.Model):
 title = models.CharField(max_length=100)
 author = models.ForeignKey("Author")
```

```
class Author(models.Model):
 name = models.CharField(max_length=100)
```

类ForeignKey，其主要参数就是它要引用的模型类，就像下面的例子：

关键的概念相对比较简单，所以从Django对它的实现也比较直观。它们表示成一个Field的子类，外键的实现相对比较简单，所以从Django对它的实现也比较直观。它们表示成一个Field的子类。

然后它是如何将积木一样构建起其他关系类型的。我们先来看看外键的定义，方法（外键），我们有必要再加入一些抽象来表示更复杂的类型。我们先来看看外键的意义，确保关系是在数据库层面上而不是在应用层上定义。然而，由于SQL值提供了一种显式组织数据层开，同时这也是ORM之间相互竞争的区域。Django目前的实现主要是围绕着数据库展开，关系），同时这也正是ORM之间相互竞争的区域。Django目前的实现主要是围绕着数据库展开，定义模型对象之间关系的能力通常是最主要的特点之一（证据就是它的名字——

## 模型之间的关系

保证那个变量的唯一性，不过你不需要把它当作主键。

说到重复，还有一个类似的参数，它可以应用到模型里的任何变量上：`unique=True`。这也能保证永远不会发生重复的情况。

所以指定一个姓名这样的字符串或是其他一些标识符可不是什么好主意，除非你有100%的把握永远不会取到id（这时id会被忽略）成为这张表的主键。这意味着变量的值必须是唯一的，如果你希望有更多的控制主键，只需在模型的某个变量上指定primary\_key=True就行了，

因为这种主键的使用相当普遍，除非你明确规定，否则Django会为你自动生成。所有没有显式指定为主键的模型都会被指定为一个ID属性，其类型为Django的AutoField（一个自增的整

数）。AutoField和普通整数一样，而它们在数据库中的类型则是根据你使用的数据库后会有不同。

因为这种主键的使用相当普遍，除非你明确规定，否则Django会为你自动生成。所有没有显式指

ID为5的Book对象并且保证它是唯一拥有这个ID的Book时，引用“第5号book”就不会有意义了。

一个Container类可以定义一个parent属性来嵌套另一个Container) 等类似场景（比如Employee类可以具有类似supervisor或是hired\_by这样的属性）的时候很常用。

虽然ForeignKey只定义了关系的一端，不过另一端却能根据关系追溯回来。外键从技术上说是一个“多对一”的关系，即可以有多个子对象引用同一个父对象。因此子对象只有一个父对象的引用，而父对象却能访问到一组子对象。拿上面的例子来说，你可以这样使用Book和Author的实例：

```
Pull a book off the shelf - see below in this chapter for details on querying
book = Book.objects.get(title="Moby Dick")
Get the book's author - very simple
author = book.author
Get a set of the books the author has been credited on
books = author.book_set.all()
```

可以看到，这里从Author到Book的“反向关系”是通过Author.book\_set属性来表示的（这是一个manager对象，本章稍后会讲到），它是由ORM自动添加的。你可以通过在ForeignKey里指定related\_name参数来改变它的名字。在上面的例子里，我们如果把author定义成ForeignKey("Author", related\_name="books")的话，最后就是访问author.books而不是author.book\_set了。

### 注意

对简单的对象层次来说，related\_name不是必需的，但是在更复杂的关系里，比如当你有多个ForeignKey的时候就一定要指定了。这时，ORM需要你明确告诉它在ForeignKey的另一端如何区分两个不同的反向关系。如果你忘了的话，Django的数据库管理工具会抛出错误信息警告你！

## 多对多关系

外键通常用来定义一对多（或多对一）的关系——在上面的例子里，一本Book只能有一个Author，而一个Author却可以有很多本Book。不过有时候你需要更大的灵活性。比如之前我们假设每本Book只能有一个Author，那要是一本书同时有好几位作者的话怎么办，就好像本书？

这样的情况光在一段有“多”关系（Author有多本Book）是不够的，要两端都有才行（Book也可以有多个Author）。这就是多对多关系，由于SQL并没有定义这种关系，我们必须通过外键用它能理解的方式实现它。

Django为这种情况提供了第二种关系对象映射变量：ManyToManyField。语法上来讲，它和ForeignKey是一模一样。你在关系的一端定义它，把要关联的类传递进来，ORM就会自动为另一端生成使用这个关系必要的方法和属性（和前面看到的ForeignKey一样，通常就是生成一个\_set manager对象）。不过由于ManyToManyField的特性，在哪一端定义它通常都没有关系，因为这个关系是对称的。

**注意**

如果你打算使用Django的admin，请注意一个多对多关系在admin里只有在定义关系的那一端才会显示关系对象。

**注意**

自引用的ManyToManyField（即一个给定模型为自己定义了一个ManyToManyField）默认就是对称的，因为它假设关系是相对的。然而，这也不是一定正确的，你可以在变量定义里通过指定symmetrical=False来改变这一行为。

根据这一“新发现”，我们必须处理一本书有多名作者的情况，因此修改例子如下：

```
class Author(models.Model):
 name = models.CharField(max_length=100)

class Book(models.Model):
 title = models.CharField(max_length=100)
 authors = models.ManyToManyField(Author)
```

ManyToManyField的用法和外键关系里“多”的那一端差不多：

```
Pull a book off the shelf
book = Book.objects.get(title="Python Web Development Django")
Get the books' authors
authors = Book.author_set.all()
Get all the books the third author has worked on
books = authors[2].book_set.all()
```

ManyToManyField的秘密在于它在背后创建了一张新的表来满足这类关系的查询需要，而这张表用的则是SQL的外键，其中的每一行就代表了两个对象的一个关系，它同时包含了两端的外键。

这张查找表在Django ORM的使用过程中一般是隐藏起来的，不可以单独查询它，只能通过关系的某一端来进行查询。不过，你可以在ManyToManyField上指定一个特殊的选项，through，来指向一个显式的中间模型类。通过through你可以手动地管理这个中间类上的额外变量，同时还能继续方便地管理关系的两端。

下面的代码和之前ManyToManyField示例一模一样，除了它包含一个显式的中间表Authoring，它在关系上添加了一个collaboration\_type变量，并通过through关键字指向这张表。

```
class Author(models.Model):
 name = models.CharField(max_length=100)

class Book(models.Model):
 title = models.CharField(max_length=100)
 authors = models.ManyToManyField(Author, through="Authoring")

class Authoring(models.Model):
 collaboration_type = models.CharField(max_length=100)
 book = models.ForeignKey(Book)
 author = models.ForeignKey(Author)
```

查询Author和Book的方法和之前完全一样，另外还能构造对“authoring”的查询。

如此，Django在创建关系的能力上就显得非常灵活了。  
用一对一关系进行组合。  
除了常见的多对多关系类型外，关系数据库开发中有时候还需要第三种类型，一对  
对一的关系。顾名思义，这个关系的意思就是两端都只有一个关联的对象。  
Django实现这个OneToOneField概念的方式几乎和ForeignKey一样——它接受一个参数，即要关联的类（或者自引用时的“self”）。同样，它还接受一个可选参数related\_name，这样就  
可以在两个相同的类里区分出多个这样的关系来。不同的地方在于，OneToOneField没有在反  
向关系里添加reverse\_manager，而只是增加了一个普通属性而已，因为关系的另一端一定只有  
一种。在Django直接支持模型继承（model inheritance）之前，OneToOneField主要是用来实现  
这种关系类型最常见的用途对象组合或是拥有关系，所以相比现实世界，它更加面向对象  
的一个对象。

模型继承这类关系，而现在则是转向对这个特性的幕后支持。  
关于定义关系的最后一点，ForeignKey和ManyToManyField都可以指定一个  
limit\_choices\_to参数。这个参数接受一个字典，它的键值对是查询的关系名和值（下面会讲到  
这些关键字是什么）。这对指定你定义的关系的可用值来说是一个非常强大的方法。

例如，下面这个版本的Book模型类只能和Smithe的Authors类一起工作：

```
class Author(models.Model):
 name = models.CharField(max_length=100)
 class Meta:
 ordering = ['name']

class Book(models.Model):
 title = models.CharField(max_length=100)
 authors = models.ManyToManyField(Author, limit_choices_to={
 'name__endswith': 'Smith',
 })
 name = models.CharField(max_length=100)
```

在编写本书的时候，Django的ORM里一个相对比较新的特性就是模型继承（model inheritance）。两个模型类之间除了外键以及其他关系之外，还可以和普通的、非ORM的

## 模型继承

注意  
你可以（有时候还是最好是）在类单层上指定这个限制。参见第6章里对ModelChoiceField 和ModelMultipleChoiceField的描述。

这里的关键是把Book的Meta类里的abstract = True设置——它指明了Book是一个抽象基类，只是用来为它实际的模型子类提供属性而存在的。注意SmithBook只是重新定义了authors变量来提供它的limit\_choice\_to选项——因为它继承的不是常用的models.Model而是

```
(1)
 name_endswith': 'Smith',
 authors = models.ManyToManyField(Author, limit_choices_to=()
class SmithBook(Book):
 abstract = True
 class Meta:
 return self.title
 def __unicode__(self):
 authors = models.ManyToManyField(Author)
 num_pages = models.IntegerField()
 genre = models.CharField(max_length=100)
 title = models.CharField(max_length=100)
 class Book(models.Model):
 name = models.CharField(max_length=100)
 class Author(models.Model):
 def __str__(self):
 return self.title
 def __unicode__(self):
 return self.name
 authors = models.ManyToManyField(Author)
 num_pages = models.IntegerField()
 genre = models.CharField(max_length=100)
 title = models.CharField(max_length=100)
 class Book(models.Model):
 name = models.CharField(max_length=100)
 class Author(models.Model):
 def __str__(self):
 return self.title
 def __unicode__(self):
 return self.name
 authors = models.ManyToManyField(Author)
 num_pages = models.IntegerField()
 genre = models.CharField(max_length=100)
 title = models.CharField(max_length=100)
 class Book(models.Model):
 name = models.CharField(max_length=100)
 class Author(models.Model):
 def __str__(self):
 return self.title
 def __unicode__(self):
 return self.name
 authors = models.ManyToManyField(Author)
 num_pages = models.IntegerField()
 genre = models.CharField(max_length=100)
 title = models.CharField(max_length=100)
```

现在我们用抽象基类来重新组织Book和SmithBook的模型层次。

def \_\_str\_\_(self):  
 return self.title

这听起来有点违反DRY的原则了，但实际上在这个场景里是合理的，你可不需要为类创建一张额外的数据表——特别是如果有别的数据表是现有的，或是还有一个应用程序在使用它的时侯。另外这种无遗漏会一个实际的对象层次能表达类定义重写（refactoring of class definitions）的方法其实更整洁。

你重构Python模型的意义，这样变量和方法就可以从基类继承下来。然而在数据库和查询层上并没有基类这个概念，子类在数据库的表现其实还是复制了基类的变量。

抽象基类（abstract base class）这种方法简单来说就是“纯Python的”继承——它允许

### 抽象基类

base class）和多表继承（multi-table inheritance）。

Django目前支持两种不同的继承方式，每种都有自身的优点和缺点：抽象基类（abstract

取决于你和你计划的模型设置。

这里的Book示例太简单，看不出这个特性的优势。但是想象一下，如果是一个有一打甚至更多属性的模型，再加上一些复杂的办法，根据DRY原则，继承这种方式就突然变得很有竞争力了。不过别忘了你还是可以使用Foreignkey或者是OneToOneField的。到底使用哪种技术完全取决于你和你计划的模型设置。

这类区分开来，而不需要重写整个类的定义。

例如，上面的SmithBook类不仅可以在意义上成为一个刚好和Book类一样有两个相同量的类，还可以显式地从Book类继承而来。这里的优点是很明显的——子类能通过添加或是重写变量来和

Python类一样，通过从另一个模型继承来定义模型。（普通类的继承可以在第1章里找到。）

Book，所以它的数据表里已经包含了title、genre和num\_pages列，以及一个多对多到authors的查找表了。在Python类这一层上它还定义了一个\_\_unicode\_\_方法返回title变量的值，这和Book里一样。

换言之，当SmithBook在数据库里被创建，以及被用来创建对象，ORM查询等的时候，它的行为和下面的定义其实是完全一样的：

```
class SmithBook(models.Model):
 title = models.CharField(max_length=100)
 genre = models.CharField(max_length=100)
 num_pages = models.IntegerField()
 authors = models.ManyToManyField(Author, limit_choices_to={
 'name__endswith': 'Smith'
 })

 def __unicode__(self):
 return self.title
```

因为这一行为还延伸到了查询机制以及SmithBook实例的属性，所以下面的查询也是完全合法的：

```
smith_fiction_books = SmithBook.objects.filter(genre='Fiction')
```

不过其实我们的这个例子不是很适合用在抽象基础类里，通常你还是会需要创建普通Book类的。抽象基础类的当然是抽象的——它们不能被直接创建，而是和前面叙述的那样，最好是在模型定义层上提供DRY的支持。下面要讲到的多表继承（multi-table inheritance）则更适合我们这个特定的场景。

关于抽象基础类最后要说的是：在子类的嵌套类Meta会继承，或是和父类里Meta类合并起来（当然除了要把abstract选项重设为False，另外还有一些和数据库相关选项，比如db\_name等）。

此外，如果基类用到了related\_name参数来设置ForeignKey这样的关联变量的话，你需要用一些字符串格式化来避免子类的名字不会发生冲突。不要使用例如“related\_employees”之类的常见字符串，而应该用%(class)s，比如这样“related\_%(class)s”。（关于这种字符串替换的细节请参考第1章）这样，子类的名字就能正确替换，避免发生冲突。

### 多表继承

从定义上来说，多表继承（multi-table inheritance）和抽象基础类的差别不大。还是会用到Python的类继承，但是不再需要abstract = True这个Meta类选项了。在检查模型实例或是查询的时候，多表继承和我们前面看到的也都一样，子类会从父类继承所有属性和方法（除了Meta类里的一些例外，原因刚刚解释过了）。

这里主要的区别在于底层的机制。在这里，父类是拥有自己数据表的完整Django模型，可以正常地实例化，同时还能把自己的属性“借给”子类。其实，这是通过自动在子类和父类之间设置一个OneToOneField，以及幕后的一些小手段把两个对象连在一起实现的，所以子类才能继承父类的属性。

所以换句话说，多表继承其实就是对普通“has-a”关系（或者说，对象组合）的一个方便的包装。因为Django希望尽量的Pythonic，所以如果你需要那个“被隐藏”起来的关系，其实可以通过OneToOneField显式的暴露出来，它会给出父类名字的小写形式并加上一个\_ptr后缀。例如，在下面的例子里，SmithBook会得到一个指向其“父亲”Book实例的book\_ptr属性。

下面的代码是用多表继承实现的Book和SmithBook示例：

```
class Author(models.Model):
 name = models.CharField(max_length=100)
class Book(models.Model):
 title = models.CharField(max_length=100)
 genre = models.CharField(max_length=100)
 num_pages = models.IntegerField()
 authors = models.ManyToManyField(Author)

 def __unicode__(self):
 return self.title

class SmithBook(Book):
 authors = models.ManyToManyField(Author, limit_choices_to={
 'name__endswith': 'Smith'
 })
```

前面提到过，这里唯一的不同就是没有Meta类的abstract选项。在一个空数据库和这个models.py文件上运行manage.py syncdb命令会创建三个主要的表（分别是Author、Book和SmithBook），而抽象基础类只会返回给我们两个表Author和SmithBook。

注意SmithBook实例得到的book\_ptr属性会指向和它们组合的Book实例，而属于（或者一部分，取决于你怎么看待它）SmithBook的Book实例会得到一个smithbook（没有\_ptr后缀）属性。

由于这种形式的继承允许父类拥有自己的实例，Meta的继承有可能会在关系的两端导致问题或冲突。所以，你需要重新定义绝大多数Meta选项，否则两个类都会分享它们（即使ordering和get\_latest\_by没有在子类里定义，它们也会被继承下来的）。这令遵循DRY原则变得稍微困难了一点，不过只要我们努力接近了就好了，哪有那么多十全十美的事情呢。

最后，我们希望你已经弄清楚了为什么这种方式比较适合我们的book模型，我们可以同时实例化普通的Book对象以及SmithBook对象。如果你要用模型继承来映射我们真实世界的关系，很多时候你都会发现多表继承比抽象基础类要好用得多。分辨究竟应该用哪一种（甚至两种都不行）是需要经验累积的。

## Meta嵌套类

模型(model)里定义的变量(fields)和关系(relationships)提供了数据库的布局以及稍后查询模型时要用的变量名——经常你还需要添加\_\_unicode\_\_和get\_absolute\_url这样方法或是重写内置的save或delete方法。然而，模型定义里还有第三个方面，即用来告知Django关于这个模型的各种元数据信息的嵌套类Meta。

顾名思义，Meta类主要处理的是关于模型的各种元数据的使用和显示：比如在一个对象对

如果你正在用urdyang的“admin”应用的话，一定要大量用到`admin.site.register`方法。它们和它们的register函数，可能还有ModelAdmin子类。这些子类允许你定义各种关系在admin应用里当你们想模型交互的时候，模型应该该如何使用的选项。

Admin注册和选项

官方文档。

如同你正在场景里看到的一样，如果没有Meta的帮助，对“人”这个概念建模还是必要的。在排序时必须同时考虑三个变量，要撇开重名，还要在系统利用超过一个“人”的时候告诉它不要用“persons”这种古怪的名字。

这样就行了！Book类在是太简单，大多数Meta类提供的是通用的接口，而且要不是我们真的关心默认排序方法的话，其实连这个Meta类都可以省略了。虽然Meta和Admin都很常用，但它们都是模型定义中可选的部分。Book这个例子太没劲了，现在我们来看一个复杂一点的例子。

```
class Book(models.Model):
 title = models.CharField(max_length=100)
 authors = models.ManyToManyField(Author)
 class Meta:
 ordering = ['title']
 # Alphabetical order
```

多个对象时，它的名字应该怎么表示；在查询数据表时默认的排序顺序是什么；数据表的名字是什么（如果你很在意这个的话）等。此外，多变量唯一性 (multi-field uniqueness constraints) 也是在Meta类里定义，因为这种限制没办法分开放在每个单独的变量声明上定义。我们来给之前 的Book类添加一点元数据。

数据库“同步”  
syncdb这种行为背后的原因除了Django的核心开发团队坚信生产数据（production data）绝不应该轻易交给一个自动化的过程。而且，通常认为只有当程序员足够了解SQL并能手动修改数据库时，才能修改数据库模式（schema）。我们对这一点也表示同意，在使用高级工具时多了解一些底层的技术总是好的。

第2章里提到过，隨每個Django項目創建的manage.py腳本包含了操作數據庫的功能。其中最常用的是manage.py命令是syncdb。不要被這個名字迷惑住，它並不是和有些用戶想像的那樣對整個數據庫進行一次完整的同步。相反，它只是保證所有模型都有對應的數據表，在必要時創建為模型創建新的表——但是不會去修改已經存在的數據表。

用 manage.py 创建和更新数据库

到此我们已经解释了如何定义以及强化模型（model），接下来要详细讨论如何创建模型，通过模型查询数据库，最后是一些关于在底层支撑整个机制的原始SQL的要点。

4.2 便用機器

风格以便符合你整个网站的话，这些功能都是很有用的。最后，如果你发现自己真的在大量使用这些选项，这其实代表了一个信号：你应该考虑舍弃admin转而编写你自己的管理菜单。当然了，先读一下第11.1节，在开始动手之前看看Django的admin究竟有多少潜力可挖。

过链接跳转，这样就能迅速浏览你的信息了。  
• 表单显示：fields, js, save\_on\_top等提供了很多灵活的方法来重写模型里默认表单表现  
形式，以及添加定制的JavaScript includes和CSS类，如果你想自己动手改变admin的外观

- 列表格式化：list\_display、list\_display\_links、list\_filter等类似的选择允许你改变量示在列表视图里的变量（默认是在单独一列中模型实例的字符串表示），以及激活变量查找和

每个选项的具体定义我们就在官方网站文档里了（在第三部分里也有一些admin使用的例子），不过这里列出了两个ModelAdmin选项里最常见的主要类型。

此外，你还可以通过创建 `InlineFormMixin` 子类并在 `ModelAdmin` 子类里引用它们，来为相关的变更操作指定显示出来。另外，把模型层次扩展到内联编辑器上就可以让你在需要的时候把内联表单放在站点里表示出来。

只需向admin注册你的模型类（并且激活Admin应用，参见第2章），它就能为你提供类的列表和菜单，而挂接一个ModelAdmin子类则可以提供更多控制，例如允许你手动挑选进列表视图里显示的变量，或是表单的布局等。

方法来让你进行常见的查询。

这个模型在数据库所有基本查询。Manager是从数据库获取信息的门户，它们包含了好几个对象总是在模型类里，所以除非有特别指定，每个模型类都会展示一个objects属性，它构成查询由模型生成的数据需要使用两个不同但却相似的类：Manager和QuerySet。Manager

## 查询语法

或是查询Django官方文档。

关于更多如何使用这些以及错综复杂的syncdb命令的信息，请参考第三部分里的应用示例，文件，随后把文件重定向给数据库执行（参见附录A来获得管道和重定向的细节）。

syncdb的步骤分解开来一样。你还可以通过重定向把两种方法组合到一个文件里，修改这个文件，当然你也可以把这些命令的输出直接用管道重定向给数据库客户端来执行，这就好像将它们保存到单独的SQL脚本文件里去。

语句打印出来，让程序员有机会验证（例如确保syncdb的行为和程序员期望的一致）或者把它和syncdb不同的是，这些sql\*系列的命令并不更新数据库。相反，它们只是把对应的SQL

| 操作         | 语句                               |
|------------|----------------------------------|
| dumpdata   | 把现有数据库里的数据输出为JSON，XML等格式         |
| loaddata   | 载入初始数据（和sqlcustom类似，但是这里没有原始SQL） |
| sqlcustom  | 显示指定sql文件里的自定义SQL语句              |
| sqlreset   | 显示DROP TABLE语句                   |
| sqlclear   | 显示DROP TABLE语句的组合 (DROPIFCREATE) |
| sqlindexes | 显示对主键创建索引的语句                     |
| sqlall     | 如同上面的sql一样从sql文件中初始化数据库载入语句      |
| sql        | 显示CREATE TABLE语句                 |
| syncdb     | 创建所有应用程序所需要的表                    |

表4.1 manage.py函数

除了syncdb，manage.py还提供了很多数据库相关的函数，syncdb实际上是在构建在这些函数上完成自己的工作。表4.1给出了一些初学者的函数。这里有显示CREATE TABLE语句的sql和sqlall（sqlall还会执行一些初始化数据的载入），创建索引的sqlindexes，清空或删除数据库的sqlreset和sqlclear，以及执行应用程序自己定义initial SQL语句的sqlcustom等。

因此，如果你创建模型，运行syncdb将它载入数据库，然后改变这个模型时，syncdb不会试图去和数据库协调那些改动。它会希望程序员自己去手动修改，或是通过脚本修改，抑或是直接编辑数据库或整个数据库然后重新执行syncdb，这样就能得到最新的数据库模式（schema）。就目前来讲，重要的还是syncdb是把一个模型类转变成数据库需要的方法。

因此，如果你创建模型，运行syncdb将它载入数据库，然后改变这个模型时，syncdb不会试图去和数据库协调那些改动。它会希望程序员自己去手动修改，或是通过脚本修改，抑或是直接编辑数据库或整个数据库然后重新执行syncdb，这样就能得到最新的数据库模式（schema）。

核心项目正在开发中。

快开发的过程。在编写本书的时候，已经有了几个试图解决框架中这一不足的Django相关非

所对应的SQL显然后就是：

```
everyone = Person.objects.all()
```

使用其他Manager的方法也是返回相似的结果，例如all：

```
SELECT * FROM myapp_person WHERE last = "Doe" AND first = "John";
```

而这里对应的SQL就是：

```
John_does = Person.objects.filter(last="Doe", first="John")
```

```
from myproject import myapp.models import Person
```

組合查詢也是可以的，以乙開的Person類型來舉例：

SELECT \* FROM myapp\_book WHERE title like "%tree%"

这样的：

回到这个例子上来，Book.objects.filter()是一个Manager方法，前面讲过了，Manager方法总是返回QuerySet对象。我们只要往上面加一个方法，就可以得到我们想要的。

这里关键字段接受的是模型变量名（比如title）、以下画线和一个句点的说明比如claims，代表“大于”的gt，代表“大于等于”的gte，代表成员测试的in等。每一个都直接（或接近）映射到SQL操作符和关键字上去。官方文档上有全部关于这些操作符的细节。

```
from myProject import Book, myApp, models
```

撰成適合的SQL。我們用Book模型來舉個例子吧。

顾名思义，QuerySet可以被当作一个数据库查询的始发端。它可以被翻译成一个能在数据  
库上执行的SQL字符串。因为绝大多数SQL查询通常都是一组逻辑语句和参数匹配的集合，  
QuerySet完全可以Python层上接受一个相同的版本。QuerySet接受动态的关键字参数然后转

将QuerySet作为数据库查询

感谢Python的动态性、灵活性，以及所謂的“duck typing”特性，多功能的QuerySet对象提供了很多重要的特性，它们可以查询数据集，可以包含数据，还可以组合在一起查询。

如果说Manager为生成查询提供了一个起点，那么QuerySet就是绝大多数查询发生的地方。

中國農業出版社編《中國農業百科全書》農業工程卷編委會編著，中國農業出版社出版。

《中国图书馆报》2013年1月1日第1版

• 第四部分：家庭与社会的互动（家庭对社会的影响）

• *excuse me, I'm sorry to trouble you, but... / I'm afraid I've got the wrong number.* — 来找不在这儿的那位先生是

• 每個應用程式都必須有一個主視窗，這個視窗稱為「視窗」。

Person objEcts. titleer (laste = "Doe"). titleer (title = "John"). titleer (middle = "Quincy")

可直接把它们写成一行。

根据重要的程度，我们一步步地缩小查询结果的范围，直到最后只剩下一个结果（反正不会很多，取决于数据库里有多少个叫John Quincy Doe的人）。因为这些都是Python代码，所以

```
from myproject import Myapp, models import Person
dope_familly = Person.objects.filter(last='Doe')
john_does = dope_familly.filter(first='John')
john_does = john_does.order_by('last', 'first')
print(john_does.query) # John Doe, John Doe
```

易經一譯

QuerySet是懒惰的。只有在必要时它才会去执行数据库查询，比如当被转换成列表或是被前面几节里提到的访问方式访问的时候。这种行为是QuerySet最强大的特性之一。它们不必是一个单独的、一次性的查询，而可以是复杂或是嵌套查询的组合。这是因为QuerySet遵循了很多和Manager一样的方法，例如filter和exclude等。和Manager一样，这些方法会返回新的QuerySet对象——但这时它们进一步被限制在了QuerySet自身的参数范围里了。半个例子会客

组合QuerySet查询

档。很可能就会找到不需要把整个QuerySet装到内存里就能解决问题的办法了。

Django已经尽可能地为ORM提供了很多强大的功能，所以如果你觉得需要把QuerySet转换成列表的时候，最好先停下来浏览一下本书或是官方文档，或者搜索一下Django部件列表的有关信息。

注意如果Query Set里的结果包含很多对象的话，这么做会导致内存或数据库的巨大负担！

有时候，你会发现需要QuerySet做的事情不可能，至少也是不适合用Django ORM现成的功能完成。这时只需要用list函数把QuerySet转换成一个列表就行了，这是一个包含了结果集的真正的列表。虽然这样有时候是必要的，也很好用（比如需要在Python层面上排序时），但是请

OFFSET关键字。

QuerySet很像一个列表。它实现了大部分列表的接口因此你可以执行操作（`for record in queryset:`，`queryset[0]`，`len(queryset)`等操作）。所以如果你熟悉Python的列表、元组或是迭代器的话，访问QuerySet里的包含的数据对你是同样的。而且这些操作都是面向智能的，比如切片和索引操作都会被自动转换成SQL里的`LIMIT`以及`OFFSET`。

将QuerySet作为容器

SQL——这样你就能获得更多的控制权了。

最后，如果你会使用SQL和理解各种查询机制背后的含义的话（无论是结果集还是查询语句），那么在构建ORM查询时会更加得心应手，能比不懂SQL的人物高出更快、更聪明的查詢來。此外，Diango正打算将来允许用户更容易地访问QuerySets对象并调用所生成的查詢來。

力也能改变最终在数据库上执行的SQL。

请注意在可选的Meta模型嵌套类型定义的那些各种和查询相关的选项都会影响到所生成的SQL，比如，`orderBy`会被转换成`ORDER BY`。稍后会看到，`QuerySet`的其他方法以及其它能

```
SELECT * FROM myapp_person;
```

机敏的读者一定觉得有了Person.objects.filter，没有什么事情是做不到的了。但是，如果之前的john\_does的QuerySet不是得自函数调用而是从某个数据结构里取得的话怎么办？这时我们根本不知道我们在处理的是什么查询内容——当然我们也不一定需要知道。

假设给Book模型添加一个due\_date变量，负责显示所有逾期未还的书（即，还书日比今天早的书）。我们可能得到的是图书馆里所有书，或是所有小说书，又或者是所有某个人归还的书等——即任何形式的集合都是有可能的。而我们可以将这样一个集合缩小到只显示我们感兴趣的书，比如逾期不还的书。

```
from myproject.myapp.models import Book
from datetime import datetime
from somewhere import some_function_returning_a_queryset

book_queryset = some_function_returning_a_queryset()
today = datetime.now()
__lt turns into a less-than operator (<) in SQL
overdue_books = book_queryset.filter(due_date__lt=today)
```

除了这种形式的过滤，任何复杂的逻辑都需要使用QuerySet组合，例如所有作者名为Smith的非小说类书籍。

```
nonfiction_smithBook.objects.filter(author_last="Smith").exclude(genre="Fiction")
```

虽然用查询选项里的否定操作也能达到同样的效果，比如（Django ORM曾经支持过的）as\_\_genre\_\_neq等，但是把逻辑放到额外的QuerySet方法里能让结构变得更清楚。还有这种把查询分成多个独立步骤的方式也更加易读。

### 查询结果排序

你应该注意到了QuerySet里有一些Manager对象中没有的方法，因为它们的作用的是修改查询的结果而不是生成新的查询。其中最常用的就是order\_by，它能改变QuerySet默认的排序方法。例如，Person类通常是按照姓氏排序的，我们可以把某个单独QuerySet改成按名字排序，像这样：

```
from myproject.myapp.models import Person

all_sorted_first = Person.objects.all().order_by('first')
```

QuerySet结果的其他部分还是一样的，除了背后SQL层上的ORDER BY子句改成了按我们的要求进行了更新。这就是说我们可以在上面加诸更多东西来组成更复杂的查询，例如查找按名字排序里的前五个人。

```
all_sorted_first_five = Person.objects.all().order_by('first')[5:]
```

甚至可以用前面见过的双下画线语法来排序模型关系。假设Person有一个ForeignKey包含了一个Address，Address里有一个state变量，我们希望按照他所在的州、然后是姓氏，对这些人进行排序。像这样：

QuerySet可以调用 `filter` 和 `order_by` 来组成AND和OR关系（虽然它们和Python里对应的操作符`and`和`or`一样，但它们的优先级不同）。`filter`方法接受一个或多个过滤条件，而`order_by`方法接受一个或多个排序条件。通过将它们组合起来，你可以很容易地对数据库进行复杂的查询。

例如，假设你有一个名为`Person`的模型，它有`name`、`age`、`height`和`weight`等字段。你想查询所有年龄大于18岁且身高大于175厘米的女性。你可以这样写：

```
Person.objects.filter(age__gt=18, height__gt=175).filter(gender='F')
```

或者，如果你想按年龄降序排列，可以这样写：

```
Person.objects.filter(gender='F').order_by('-age')
```

注意，`filter`方法返回的是一个`QuerySet`对象，而不是一个`Person`对象列表。如果想要直接得到一个`Person`对象列表，可以在`filter`方法后面加上`.all()`方法：

```
Person.objects.filter(gender='F').order_by('-age').all()
```

这样，你就可以直接使用`for`循环遍历结果了：

```
for person in Person.objects.filter(gender='F').order_by('-age').all():
 print(person.name, person.age)
```

当然，你也可以使用`get`方法来获取单个对象：

```
person = Person.objects.filter(gender='F').order_by('-age').first()
print(person.name, person.age)
```

除了`filter`方法，还有`exclude`方法，它可以用来排除某些对象。例如，如果你想查询所有年龄小于18岁的女性，可以这样写：

```
Person.objects.filter(gender='F').exclude(age__gt=18)
```

注意，`exclude`方法接受一个或多个过滤条件，与`filter`方法的功能相反。

“W”的人，这就不是我们要的结果了），但是用~Q就能正确表达出这个意图。要是这里用exclude(middle\_startswith="W")的话就不对了（它会排除所有姓Doe，中间名是

```
(Q(last="Smith") & Q(first="John") & ~Q(middle_startswith="W"))
Person.objects.filter(Q(last="Doe"))
```

换取所有姓Doe，名John Smith，但是不取John W.Smith的A。最后，你可以对Q对象使用~元操作符~来取反。虽然QuerySet的exclude方法可能是更直观一点，不过~Q在逻辑很复杂的時候就显示出来优势来。比如在下面这个例子中，只用一行就能

这类手法可以在框架里的任何一部分里。  
如果上面的例子要是用上Python函数式编程工具的话（比如列表推导式，内置的reduce方法，以及operator模块），还能更简洁一点。operator模块提供了和操作符等价的函数，例如 I 对应的or\_，还有和 & 对应的and\_。这样那三行for循环就可以重写成这个样子：  
reduce(or\_, [Q(first=name) for name in first\_names])。不要忘了，Django其实“就是Python”，reduce(or\_, [Q(first=name) for name in first\_names])。不要忘了，Django其实“就是Python”，  
注意

这里，我们创建了一个for循环用！操作符把所有列表里生成的Q对象组合起来。虽然这个例子举的不是很好（这种情况其实用~in查询操作符就可以了），但是它展示了这种通过遍历的方式将Q对象组合起来的能力。

```
specific_does = Person.objects.filter(last="Doe").filter(first_name_keywords)

first_name_keywords = first_name_keywords | Q(first=name)

for name in first_names:
 first_name_keywords = Q() # Empty "query" to build on

first_names = ["John", "Jane", "Jeremy", "Julia"]
```

和QuerySet一样，Q对象也可以组合在一起使用，通过 & 和 | 操作符，你可以返回和操作对应的新的Q对象。比如，你可以像这样用循环来组装出一个巨大的查询。虽然这个例子有点做作（大概不太会有人关心一个特定的名字或是中间名），不过这里主要是为了展示Q的用途。

```
from myproject import Person
from django.db.models import Q
```

```
specific_does = Person.objects.filter(first_name="John") | Q(middle="Quincy")
specific_does = Person.objects.filter(first_name="Doe").exclude(
```

exclude方法的关键字参数，例如：  
或是位操作符 & 和 | 很像，但是千万不要弄混了）。其结果Q对象可以用来作为filter或是

Django会以appname\_modelname这样的格式命名你的数据表。  
这些名字随后会被插入到查询的FROM子句里，一般是用JOIN语句串联起来。在默认情况下，extra的参数里最简单的可能就是tables了，它允许你指定一个包含额外数据库名字的列表。

```
SELECT first, last, (first||last) AS username FROM myapp_person WHERE
extra(first, last, (first||last) AS username, select=(username, "first||last"),
first||last LIKE Jeffrey%"), extra(first, last, (first||last) AS username,
first||last LIKE Jeffrey%"), extra(first, last, (first||last) AS username,
first||last LIKE Jeffrey%"), extra(first, last, (first||last) AS username,
```

里，我们用了同一个SQL结构来搜索和返回一个用PostgreSQL风格的连接符连接的字符串；无法用属性相关的关键字参数如--gt或--contains组成正确的查询语句。在下面的例子被加入到最终的SQL查询as-is里去（或者说几乎是as-is，参见下面的params）。where最适合在where参数接受一个字符串来表示为参数，列表里包含的是原始的SQL WHERE子句，直接

```
person.last, person.age)
print "%s %s is an adult because they are %d years old." % (person.first,
if person.is_adult:
for person in the_folks:
the_folks = Person.objects.all().extra(select=(is_adult, "age > 18"))
SELECT first, last, (age > 18) AS is_adult FROM myapp_person;
```

from myproject.myapp.models import Person

这个例子是用select来添加一个简单的数据库逻辑的属性：

所以也可以用来做优化。

方法会在所有的地方执行）就很方便。而且，有些操作在数据库执行就是要比在Python更快，这时候，以及把这些改动限制在代码中的一部分的时候（而不是应用在模型方法上，因为这些的选择为生成的数据类型添加定制的属性。这在你需要从数据库获取的信息之外再添加一些

select参数接受一个映射到SQL字符串的字典字符串，让你可以在SQL SELECT子句里

| extra参数 | 描述           |
|---------|--------------|
| select  | 修改SELECT语句   |
| where   | 提供额外的WHERE子句 |
| tables  | 提供额外的表       |
| params  | 完全的替换动态参数    |

表4.2 extra接受的参数

一些没有在之前模型里定义的属性。

关于Django查询机制的最后一点（以及引导到下一节，它现在做不到的内容），我们来看看QuerySet的extra方法。这是一个多功能的方法，可以用来修改QuerySet生成的原始SQL的各个部分，它接受四个关键字参数，如表4.2。注意为了更好地展示概念，这一节的例子里会用到

## 用Extra调整SQL

这里是一个tables的例子，简单起见，它用的模型和别的不太一样（回到了Book类这个例子上，还多加了一个author\_last的属性）。

```
from myproject.myapp.models import Book

SELECT * FROM myapp_book, myapp_person WHERE last = author_last
joined = Book.objects.all().extra(tables=['myapp_person'], where=['last = author_last'])
```

最后一个参数就是params。在高级程序语言里进行数据库查询的“最佳实践”之一就是正确地转义或插入动态参数。初学者常犯的一个错误就是简单地用字符串连接来把他们的变量插入到SQL查询里去，但是这样做会给系统带来潜在的安全漏洞和bug。

所以在使用extra时，应该使用params关键字（它接受一个要替换值的列表）来替换where字符串里的%s字符串占位符，像这样：

```
from myproject.myapp.models import Person
from somewhere import unknown_input

Incorrect: will "work", but is open to SQL injection attacks and related problems.
Note that the '%s' is being replaced through normal Python string interpolation.
matches = Person.objects.all().extra(where=['first = %s' % unknown_input()])

Correct: will escape quotes and other special characters; depending on
the database backend. Note that the '%s' is not replaced with normal string
interpolation but is to be filled in with the 'params' argument.
matches = Person.objects.all().extra(where=['first = %s'],
 params=[unknown_input()])
```

## 利用Django没有提供的SQL特性

最后关于Django模型/查询框架的一点就是，ORM并不能完整地覆盖所有可能性。几乎没有哪个ORM敢自称能100%地替换常规的数据库接口。Django也不例外，即便它的工程师们已经并且一直在努力提升它的灵活性。有时候，特别是对有丰富的关系数据库经验的人来说，避开ORM也是相当有必要的。下面的几节就是关于这个话题的内容。

### 定义模式 (schema) 和定制initial SQL

除了标准的表和列，大多数关系型数据库还会提供一些额外的特性，例如视图或统计表(view)，触发器(trigger)，定义在数据行被删除或更新时“联级”(cascade)行为的能力，甚至在SQL层上自定义函数或数据类型等。在本书编写的时候，Django ORM（和几乎所有ORM）在很大程度上都无视了这些东西，但这不是说你就不能用它们了。

Django的模型定义框架里一个才加入不久的特性就是定义initial SQL文件的能力，它们必须存放在应用程序里的sql子目录下并以.sql结尾，比如myproject/myapp/sql/triggers.sql。任何这样的文件都会在你运行manage.py和SQL相关的命令如reset或syncdb，以及被包括到sqlall或sqlreset输出里的时候自动执行。这个特性还有一个自己的manage.py命令，sqlcustom，它（和

```

 "first": "Jane",
 "fileds": [
 "model": "myapp.Person",
 "pk": "2",
],
},
{
 "last": "Doe",
 "middle": "Q",
 "first": "John",
 "fileds": [
 "model": "myapp.Person",
 "pk": "1",
],
}
]

```

这里是一个给Person模型类准备的例子，一个简单的JSON fixture文件：

就会读取这些文件并用其内容创建数据库对象。（参见第9章）

数数据库创建/重设命令，比如manage.py syncdb或reset的时候，Django的序列化模块（参见第9章）会包含了一个initial\_data.json文件（可以是.xml、.yaml，或是其他序列化格式）。每当运行里面包含了initial\_data.json文件（可以是.xml、.yaml，或是其他序列化格式）。Django应用将这个功能的方式和前面介绍的initial SQL很相似，每个Django应用都有一个fixtures目录，来给用户输入的数据分类用的“预装”的数据表，或是在开发期间载入的测试数据。Django fixtures最常见的用法就是在数据库新建或是重建的时候，例如用fixtures添加数据进来，例如用

库无关的（通常是可以读的）表现形式，例如XML、YAML，或是JSON。

保存在纯文本文件里的数据表集合的名字，通常指的不是原始的SQL dump，而是另一种数据之外和数据库相关的Django特性：fixtures。在第2章里我们已经简单地接触过它了，fixtures是

虽然技术上说起来这不算是SQL本身的内容，但我们还是决定在这一节看一下这个在ORM

Fixtures：载入和导出数据

就需要使用QuerySet.extra。

- 自定义函数和数据类型：这些都可以在initial SQL文件里定义，但是要在ORM里引用的

以用通过manage.py sqlall输出功能添加。)

- 触发器和联级：它们都能和普通ORM方法生成的插入或更新操作一起正常使用，而且根据数据库的不同，还可以定义在initial SQL文件里。如果联级限制不能创建的话，也可

通过SQL库访问视图时，任何对视图表的写操作都会导致错误。

- 注意要小心不要执行任何试图修改这样一个个模型的manage.py SQL命令，否则就会发生问题。因为SQL库访问视图时，任何对视图表的写操作都会导致错误。

- 表的布局这种方式来支持它们，然后你就可以通过普通的Django查询API来访问它们了。

- 视图：因为SQL的视图实际上就是只读表，所以你可以用创建Model定义来映射到这些

过使用这个特性来完成。

Django的工具来创建或者重新创建数据库时，它们一定会被包括进来。下面这些功能都可以通过使用initial SQL，你可以将模型定义命令放在Django项目里，并且你知道适当使用

其他sql\*系的命令一样）会打印出它找到的自定义SQL。

如果ORM（包括extra带来的灵活性）怎么样都无法满足你查询的需要的话，你总是可以回

就像你看到的一样，textures能让你用比原始SQL简单一点的语法处理数据。

```
 "model": "testapp.Person",
 "pk": 1,
 "fields": {
 "first": "John",
 "last": "Doe",
 "middle": "Q",
 "fiddle": {
 "model": "testapp.Person",
 "pk": 2,
 "fields": {
 "first": "Jane",
 "last": "Doe",
 "middle": "N",
 "fiddle": {
 "model": "testapp.Person",
 "pk": 3,
 "fields": {
 "first": "Doe",
 "last": "Doe",
 "middle": "D"
 }
 }
 }
 }
 }
]
}
```

除了机密化数据，fixtures通常还被用作是一个比数据库SQL dump工具更中立的（虽然有时比较低效或不太专用）数据库dump格式——例如，你可以在一个Django应用的数据从PostgreSQL里导出来，随后导入到MySQL数据库，如果没有一个中间转换的fixtures步骤的话，这类任务还真不容易。它们的命令是manage.py dumpdata和loaddata。

```
user@example:~/opt/code/myproject$./manage.py syncdb
Installing json fixture 'initial-data' from '/opt/code/myproject/myapp/fixtures'.
Installed 2 object(s) from 1 fixture(s)
```

到底层的数据库接口上执行自定义SQL语句的。Django的ORM也是使用这些模块来和数据库交互的。根据Django数据库设置的不同，这些模块也会随之变化，通常它们都是符合Python DB-API标准的。只需要导入django.db里定义的connection对象，然后就可以获取数据库游标进行查询了。

```
from django.db import connection

cursor = connection.cursor()
cursor.execute("SELECT first, last FROM myapp_person WHERE last='Doe'")
doe_rows = cursor.fetchall()
for row in doe_rows:
 print "%s %s" % (row[0], row[1])
```

#### 注意

要知道更多关于这些模块提供的语法和方法调用的细节，请参阅Python DB-API文档，《Core Python Programming》中的“数据库”一章，或是数据库适配器文档（见withdjango.com）。

## 4.3 总结

本章覆盖了很多内容，我们希望你能从中得到一些有用的想法可以（也应该）带到你自己的数据模型里去。你了解了使用ORM的原因，学习了如何定义简单的Django模型，以及如何定义它们之间复杂的关系，还看到了Django如何使用特殊的嵌套类来指定模型的元数据。此外，希望你也同意QuerySet类作为从数据库获取数据手段的强大与灵活，并且了解了如何在ORM之外操作Django应用程序的数据。

在下面的两章，第5章和第6章里，你会学习到如何在一个Web应用里使用你的模型，比如在控制器逻辑（视图）里设置查询，然后在模板里显示它们。

使用 \$ 字符通常也是基于相同的原因。它确保正则表达式只会匹配URL的最后而不是中间。  
也就是说，如果能匹配它们，而 r\bar/ 则更明确的多，它只会匹配后面那个字符串。  
实际上，^ 的作用主要是消除匹配的二义性。URL /foo/bar/ 和 /bar/ 是不同的，但是正则表  
现处都是，故而省略了），最先要查找的就是分别代表字符串起始和结束的正则表达式字符，^  
显然，这里的正则表达式（参见第1章）才是重中之重。除了缺少打头的斜杠外（因为它  
和 \$。）

```
(r'^archives/(?P<year>\d{4})/(?P<month>\d{2})/(?P<day>\d{2})/$', 'archive'),
(r'^$', 'index'),
urlpatterns = patterns('myproject.myapp.views',
```

- 从 django.conf.urls.defaults import \*
- 举个例子就清楚了，我们用在第2章里的Blog应用程序中的URLconf为例：
- 视图函数对象或字符串；有时还可以带上一个视图函数的字典参数。
  - 一个或多个由正则表达式 (regex) 字符串匹配一个或一组URL组成的Python元组；一个  
数由以下两点组成：
  - 一个打头的前缀字符串，可以为空。

刚刚提到的映射存放在叫做URLconf的Python文件里。这些文件都必须部署出一个  
urlpattern对象，这个对象则应该是Django定义的patterns函数的结果。这个被调用的patterns函  
数由以下两点组成：

Django使用的是一个简单而强大的机制，让你可以通过正则表达式配置映射到Python的视图方  
法上，并且通过相互包含把它们链接在一起。这样的系统不但使用方便，且灵活多变。  
将一个请求的URL和结果的响应联系起来的机制就是任何Web开发框架的关键所在。

## 5.1 URL

这一章稍息介绍Django是如何实现第三章里介绍的HTTP请求-响应架构，然后解释了  
构成控制器逻辑的Python函数，以及一些协助常见任务的内置帮助函数。  
这两个方面是不够的，你还需要控制器逻辑来决定哪个模板渲染什么数据，以及URL分派机  
制来决定对于给定的URL应该执行什么逻辑。  
在这一章里，我们学习了如何定义支撑几乎所有的Web应用的数据模型 (model)。在这一章  
里，我们要向你展示如何通过Django的模板语言和表单显示这些模型。但是，一个Web框架只

# 第5章 URL、HTTP机制和视图

## 注意

如果你注意一下前面例子的第一个元组，就会发现它只包含了 \$。这代表它要匹配的是网站的根URL，（前面说过，Django会去掉开头的斜杠，所以在字符串的开头（/）和结尾 (\$) 之间就只剩下了一个空字符串了。在Django项目里你会经常用到这个来定义首页。

但是如果某个URL项要包含其他URLconf文件的话，就不需要这个 \$ 了，因为一个项要包含其他URL的URL不是最终的URL，只是其中的一部分而已。

## 用url替换元组

这些正则表达式里另一个要注意的地方就是它们可以使用一种叫符号组 (symbolic groups) 的语法来捕捉部分URL的变化。这个特性提供了定义动态URL必要的能力。在上面的例子中，我们有一个blog字符串的部分，可以通过日期来定向它而已。被URL访问的基本信息 (blog帖子) 并没有变化，只是通过日期来访问它而已。就像刚才看到的那样，这里捕获的信息会传递到特定的视图函数里去，然后函数会在数据库查询里或是任何合适的地 方使用它们。

最后，定义完正则表达式以后，接下来要注意的是它所链接的函数以及可选的关键字参数 (这里以字典形式存放)。patterns的第一个参数如果不是空的话就会被加到函数字符串之前。拿刚刚的例子来说，前缀字符串是 "myproject.myapp.views.index" 和 "myproject.myapp.views.archive"。Python模块路径名分别是 "myproject.myapp.views.index" 和 "myproject.myapp.views.archive" 最终完整的URL。

url方法是最近才加入到Django的URL分派机制里的，作用是在替换前面例子里的那些元组出现在所有的URL里保持唯一，你可以在别的地方通过它引用这个特定的URL。我们可以用url重写刚才的例子。

```
from django.conf.urls.defaults import *
```

```
urlpatterns = patterns('myproject.myapp.views',
 url(r'^archives/(\d{4})/(\d{2})/(\d{2})/$', 'archive',
 {'name': 'archive'},),
 url(r'^$', 'index', name='index',),
)
```

## 注意

因为它们都是可选的。你可以一个都不指定，指定其中一个，或者两个都指定。位置参数 (positional argument) 而不是命名参数 (named argument) 会让你更难讲清楚。(或使用元组) 会让设置变得困难许多。

你可以看到，这就是对旧有的元组语法的一个简单的替换。因为它是一个实际的函数而不是一个元组，它迫使以前的语法变成一个约束而已。前两个参数是必需的且没有名字，字符串现在是一个名字的可选参数wargs，以及一个新的可选的名字参数name。

例子的更新版：

上来填充，并且用一个特殊的include函数来将URL分派传递到其他应用程序里去，这里是之前讲过的特性，如认证等。那么这个基础应用程序的URLconf会定义一些小小节让其他应用程

会有多个应用的项目里很常见，比如可以有一个“基础的”应用程序定义首页或者他用于整

上一节里看到的重构思想可以进一步深入下去，将URLconf文件分成多个这样的文件。这在

## 用include来包含其他URL文件

简化。

“catalog”部分还是有点重复。所以下面我们来看看如何通过包含其他URLconf来进一步的精明的读者一定注意到了这不是重构的尽头。URL定义里的“blog”，“guestbook”，和“catalog”部分都是同一个URL定义，而且归功于这些不同的前缀参数，每个URL都有自己不同的映射。当然，一共6个URL定义，而且归功于这些不同的前缀参数，每个URL都有自己不同的映射。

注意这里第二和第三个patterns调用里+操作符的使用。在文件的最后一行，urlpatterns包含了下

```
urlpatterns += patterns('myproject/catalog.views.',
```

```
url(r'', catalog.add/$', 'new_entry'),
```

```
url(r'', catalog.add/$', 'index'),
```

```
urlpatterns += patterns('myproject.guestbook.views.',
```

```
url(r'', blog.topics.(?P<topic_name>(w+)/new/$, 'new_post'),
```

```
url(r'', blog.new/$', 'new_post'),
```

```
url(r'', index),
```

```
urlpatterns = patterns('myproject.blog.views.',
```

```
from django.conf.urls.defaults import *
```

例子，它表示了一个连接有多个应用的顶层URL。

这样的对象连接起来，所以一般都会按照前缀字符串来将它们分隔开来。这里是一个半抽象的部件对象类型，它可以被当作列表或是其他容器类型一样附加元素。这样就可以很方便地将多个是对那些条目数目很大的文件来说。可以这么说的原因是patterns的返回类型是一个Django的内

Django程序员常用的一个技巧就是重构URLconf，把URL打散到多个patterns调用里，至少

## 使用多个patterns对多

细节，请参见第三部分中的应用示例。

最后，要知道更多关于name参数的信息以及如果通过它从代码的其他部分递向引用URL的

也更灵活。

url方法。在本章剩下的部分里，我们会努力用它来做出一个好榜样，因为它比元组方式更强大

还是使用元组多过使用url来设置Django URLconf。不过，我们还是强烈推荐你在线程里使用

在基于元组的语法后面介绍url方法是因为它比较新，即使是你读到这里的時候，外界肯定

```

urls.py

from django.conf.urls.defaults import *

urlpatterns = patterns('myproject.blog.views',
 url(r'^$', 'index'),
 url(r'^blog/', include('myproject.blog.urls')),
)

urlpatterns += patterns('',
 url(r'^guestbook/', include('myproject.guestbook.urls')),
)

urlpatterns += patterns('',
 url(r'^catalog/', include('myproject.catalog.urls')),
)

blog/urls.py

urlpatterns = patterns('myproject.blog.views',
 url(r'^new/$', 'new_post'),
 url(r'^topics/(?P<topic_name>\w+)/new/$', 'new_post'),
)

guestbook/urls.py

urlpatterns += patterns('myproject.guestbook.views',
 url(r'^$', 'index'),
 url(r'^add/$', 'new_entry'),
)

catalog/urls.py

urlpatterns += patterns('myproject.catalog.views',
 url(r'^$', 'index'),
)

```

这个例子比前面的版本要长一点，但是你可以看到对一个现实世界中每部分都包含数个URL的网站来说这种写法好处多多。在其中的每个部分，我们都避免了在URL定义里重复输入“blog”、“guestbook”和“catalog”。特别是现在我们把这个多应用网站里的绝大多数URL都交给了各自的子应用去处理，除了在blog应用下的首页（当然你也可以为它创建一个base或者类似的应用——这完全都是由你安排的）。

URLconf包含（URL including）即使对单个应用也是有价值的——因为使用多个应用还是一个应用根本就没有硬性规定，所以完全可以在一个Django应用里包含几百个URL。当然大多数程序员在这种情况下会迅速开始为它们组织模块，并采用URLconf包含来支持它们。通常来说怎么组织站点是由你决定的，而URLconf机制已经被设置的尽可能的灵活，其中包含（includes）又占了很大的比重。

## 函数对象 v.s. 函数名字符串

在这一节里，我们已经看到了URL所链接到视图函数里的Python模块路径可以用字符串来指定。不过这不是唯一的方法，Django还允许你传入一个可调用的对象（callable object），例如：

```
from django.conf.urls.defaults import *
from myproject.myapp import views

urlpatterns = patterns('', # Don't need the prefix anymore
 url(r'^$', views.index),
 url(r'^blog/', include('myproject.blog.urls')),
)
```

这样一来，它的功能就非常强大了，比如可以使用装饰器函数（decorator function）来包裹通用视图（generic view），甚至创建你自己的可调用对象将更复杂的工作代理给其他视图。第11章会讨论更多关于装饰器以及使用可调用视图的一些其他技巧。

### 注意

有时候在URLconf里把`from myproject.myapp.views import *`覆盖掉来用可调用视图的这种想法很有诱惑力，但是这在混合多个视图模块的时候有可能会产生问题——比如两个不同的视图文件都定义了它们自己的`index`视图。所以比较聪明的做法就是学习上面的例子那样，把每个视图模块导入为它自己的对象（必要时可以使用`from x import y as z`这种语法），这样我们可以保持本地名字空间不被污染。

## 5.2 HTTP建模：请求、响应和中间件

你已经了解了如何设置URL定义以及如何将URL和视图函数联系起来，现在我们来看看这些视图函数的周围都有些什么。第3章里说到，Django把HTTP归结为相对简单的Python请求对象和响应对象。加上URL转发和视图函数，一个请求在你的Web应用程序的流程是这样的：

- Web服务器接到一个HTTP请求。
- Django把Web服务器传过来的请求转换成一个请求对象。
- Django在URLconf里查找正确的视图函数。
- 调用这个视图函数，参数为请求对象以及任何捕捉到的URL参数。
- 然后视图会创建并返回一个响应对象。
- Django将这个响应对象转换成Web服务器可以理解的格式。
- Web服务器将响应发送给客户端。

首先我们看一下请求和响应对象以及它们的组件，然后深入Django中间件，它提供了进入上述过程中多个步骤的接口。之后，下一节的主要内容就是讲解关于视图本身你所需要了解的知识。

## 请求对象

设置完URLconf之后，你就需要定义URL所对应的行为。视图方法的细节稍后会详细介绍，现在我们让你看看视图处理的HTTP请求和响应对象是什么样子的。所有视图函数都接受一个“请求”参数HttpRequest对象，它是一组经过良好封装的属性，代表了从Web服务器传递进来的原始HTTP请求。

### GET和POST字典

Web程序员最常用到的请求数据就是GET和POST数据结构，它们是HttpRequest对象的属性（即request.GET和request.POST）并以Python字典的形式出现。虽然它们结构相同，但是填充的方法却很不一样，这里面的区别比乍看之下要重要的多。它们一起为参数化Web请求提供了一个灵活的方式。

#### 注意

虽然HttpRequest的GET和POST属性看起来和Python的内置dict非常的像，不过其实它们是Django专用的dict子类，QueryDict的实例，这个类型是设计用来模仿在底层HTTP CGI标准里这些数据结构的行为的。所有的键-值对里的值都是列表，即使对只有一个值的键也是一样，以便在HTTP服务器万一真的返回多个值的时候能正确处理它们。不过方便起见一般QueryDict都是和普通字典一样返回一个单值。如果你对多值都感兴趣的话，可以使用QueryDict的方法，比如getlist来获取它们。

GET的参数是作为URL字符串的一部分传递进来的，但是并不属于URL本身，因为它们没有定义另一个资源（或视图），只是改变它们所对应的资源的行为罢了。例如URL /userlist/ 指向了一个列出网站社区上用户的页面，如果用户希望不要一次列出那么大的一个列表，他可以通过GET变量告诉服务器所需的页码：/userlist/?page=2。这样，虽然被访问的还是这个视图，但是程序员可以在GET字典里查找一个page的键值对，然后返回正确的页面，像这样：

```
def userlist(request):
 return paginated_userlist_page(page=request.GET['page'])
```

注意这里的request.GET（包括请求对象里其他字典类属性），若是使用get这样的字典方法会比较好（字典的细节请参考第1章），因为这样一来即使你要查找的参数没有指定的话，程序的逻辑也不会出问题。

POST参数不属于URL的一部分，所以用户是看不到它们的，通常它们是Web页面上由HTML表单生成的。FORM标签的action参数指明了数据要提交给哪一个URL。用户提交表单的时候，URL就会随着由表单字段组成的POST字典一起调用。虽然从技术上说，它们的数据可以通过GET方法提交（很少见，因为没必要弄一个），不过绝大多数的Web表单都是用POST来操作的。

除了GET和POST，HttpRequest还有一个REQUEST字典，它会搜索前面两个字典试图返回一个被请求的键。如果一个给定的键值对可能被任何一个方法发送到视图里，你不知道改用哪一个的时候，这就很方便了。但是根据“尽量显式地表达意思而不要依赖于默认行为”的

Pythonic哲学，大多数有经验的Django程序员都会避免使用这个特性。

### Cookies和会话 (Sessions)

通常来说，大多数cookies都是用来支持一个叫做会话 (session) 的特性。这代表一个网页客户端——大多数网站的认证系统都会用到它，有些商业站点也会用它来跟踪用户的浏览历史。

通常来说，大多数cookies都是用来支持一个叫做会话 (session) 的特性。这代表一个网页客户端 (通常是指浏览器) 而且在cookies里保留一个唯一的会话ID。所以存放关键数据是很不安全的，大多数网站都会选择把信息放在一个服务器端的会话 (session)，然后用这个信息来为那个用户提供定制化的页面。由于cookies在客户端可以被轻易修改，所以存放关键数据是很不安全的，大多数网站都会选择把信息放在一个服务器端的会话 (session) 是独立的，对之而言之后的请求完全不会有概念。而有了会话以后，Web应用就可以绕过这个限制在服务器端保存在数据库里 (例如通知用户提交的表单是否成功保存了他的修改) 并在稍后的响应里把它们发送给用户了。

Django里的会话也是HTTP协议的一个字典类型，`request.session` (注意这里 `session` 是小写，而像其他属性那样是大写——这是因为会话并不是HTTP协议的一部分)。和之前的COOKIES属性一样，Python代码也能读写session。当第一次被传递到代码里来时，它包含的是根据用户名从数据库里读出的会话数据。在写入的时候，它会把你做出的修改保存到数据库里，以供将来读取。

前面讲到的都是请求对象里最常用的基本属性，不过请求对象里还有很多其他变量。`path`: URL里域名后的部分，例如，`/blog/2007/11/04/`。这个通常也是URLconf要处理的字符串。`method`: “GET”或“POST”两者之一，标明了这个请求用的是哪一个HTTP请求方法。`encoding`: 标明了用来解码表单提交数据所需要的编码字符串。`FIELDS`: 这是一个字典类型对象，包含了所有通过文件上传人表单字段上传的文件，其中每一个值又都是另一个字典，里面包含了文件名、内容的类型以及文件的内容。`META`: 这也是一个字典，它包含了所有没有被请求的其他部分处理的HTTP服务器/请求变量量，例如CONTENT\_LENGTH、HTTP\_REFERER、REMOTE\_ADDR和SERVER\_NAME等。

- `user`: Django的认证用户，只有当你的站点激活Django的认证机制时才会出现。
- `raw_post_data`: 请求里包含的POST原始数据。几乎都是推荐直接使用`request.POST`而不是`raw_post_data`，不过若是有任何高级需求的话，也可以直接访问这里。

## 响应对象

到此，你已经读入传递给视图函数的信息了。现在我们来看看负责返回的响应。对我们来说，响应比请求要简单一点——其主要数据就是存放在content属性里的正文（body text）。一般就是一个巨大的HTML字符串，它对HttpResponse非常重要，你有好几种设置它的方法。

最常用的方法就是创建响应对象——HttpResponse接受一个字符串作为构造函数的参数，并将它保存在content里。

```
response = HttpResponse("<html>This is a tiny Web page!</html>")
```

这样你就有了一个完整的响应对象，可以返回给Web服务器并转发给用户浏览器了。不过，有时候需要一点一点的构建响应内容，HttpResponse对象实现了一个有点类似文件的行为来支持这一点，那就是write方法。

```
response = HttpResponse()
response.write("<html>")
response.write("This is a tiny Web page!")
response.write("</html>")
```

这就是说你可以把HttpResponse用在任何接受一个类似文件对象的代码里（比如csv模块里写CSV的工具），由此灵活地生成代码返回给最终用户的信息。

响应对象里的另一个功能就是设置HTTP头，这时你可以把HttpResponse当作是一个字典来用。

```
response = HttpResponse()
response["Content-Type"] = "text/csv"
response["Content-Length"] = 256
```

最后，Django为很多常见的响应类型提供了HttpRequest子类，例如HttpResponseForbidden（即HTTP 403）和HttpResponseServerError（即HTTP 500，或服务器内部错误）。

## 中间件

虽然Django应用程序的基本流程还是比较简单的（接受请求，找到适合的视图函数；返回响应），不过还可以在上面加诸更多层次来变得更灵活强大。其中之一就是中间件（middleware）——这些Python函数可以在上述过程里的多个地方执行来改变整个应用程序的输入（在请求到达视图之前对它进行修改）输出（修改视图创建的响应）。

Django里的中间件组件就是一个Python类，它实现了一个特定的接口，就是说它定义了一些名为process\_request或是process\_view这样的方法。（在下面几节里，我们会讨论其中最常用的一些。）在settings.py文件里的MIDDLEWARE\_CLASSES元组里列出后，Django会内省这些类，并且在适当的时候调用其方法。类在设置文件里罗列的顺序也决定了它们执行的顺序。

Django自带了很多好用的内置中间件，有些可以直接拿来用，有些则需要特定的“contrib.”应用程序支持，比如认证框架。更多信息可以查阅Django官方文档。

### 请求中间件

请求中间件一般用在输入端，它定义为一个实现了process\_request方法的类，像这样：

```
from some_exterior_auth_lib import get_user

class ExteriorAuthMiddleware(object):
 def process_request(self, request):
 token = request.COOKIES.get('auth_token')
 if token is None and not request.path.startswith('/login'):
 return HttpResponseRedirect('/login/')
 request.exterior_user = get_user(token)
```

注意给request.exterior\_user赋值的那一行展示了请求中间件的一个常用用法：给请求对象添加额外属性。在那一行被调用后，process\_request隐含返回None（如果没有显式的return语句，Python的函数一定返回None），然后Django会继续处理其他请求中间件，最后才轮到视图函数本身。

但是，如果对一个有效的认证符号（valid auth token，并且用户当前并没有要视图登录）测试失败的话，我们的中间件会把用户重定向到登录页上去。这也展示了中间件方法的另一种能力。它们可以返回一个立刻发送给请求客户端的HttpResponse（或子类）对象。在这里，因为我们的中间件是一个请求中间件，在那一行之后正常流程里的所有代码（包括原来要调用的视图）全部都会被忽略掉。

### 响应中间件

响应中间件是运行在视图函数返回的HttpResponse对象之上。这样的中间件必须实现process\_response方法，这个方法接受一个request和一个response参数并返回一个HttpResponse或子类的对象。这些就是全部的限制了——你的中间件可以修改得到的响应或是创建一个全新的响应并返回。

响应中间件一个常见的用法就是在响应里插入额外的头信息（header），即可以是全面地（激活HTTP缓存等相关特性），也可以有选择地（把Content-Language设置为当前语言）。

下面是一个简单的例子，它把Web应用输出的所有文本里的“foo”都替换成“bar”：

```
class TextFilterMiddleware(object):
 def process_response(self, request, response):
 response.content = response.content.replace('foo', 'bar')
```

当然我们可选用比较实际的例子，如过滤所有的脏字（这对社区型的网站很有用）等，但为了不让小孩子学坏，还是免了吧！

## 5.3 视图与逻辑

视图（也叫控制器）是所有Django Web应用程序的核心，因为它们提供了几乎所有实际的程序逻辑。在定义和使用模型的时候，我们是数据库管理员；在编写模板的时候，我们是界面设计师；而在编写视图的时候，我们才是真正的软件工程师。

从哪里取得它的参数：  
from django.views.generic.list\_detail import object\_detail  
from django.conf.urls.defaults import \*

结合的参数，共确保视图要渲染和返回的模板存在就可以了。

单！根据应用组件子模块的不同，这些组件将调用不同的类。Django的视图将根据视图函数而已，但是对于被定位的角色来说，它们又都是高质抽象和参数化的函数以达到最大的灵活性而已。

Django里（甚至Web框架里）用的最多的功能大概就是使用预定义的代码来编写组织成大多  
数Web应用的CRUD操作了。CRUD全称是Create（创建）、Read（读取，也叫Retrieve，获取）、  
Update（更新）和Delete（删除），它们是在一个数据库应用里用的最频繁的操作了。显示一列  
数据或是显示单个对象的细节页面？那就是Retrieve。显示一个编辑表单，提交后修改数据  
库？根据应用程序和表单的不同，这可以是Update或Create。Delete的意思就很简单啦。

通用视图

只要它返回的是一个`HttpkResponse`对象，方法里面做什么都可以：我们在那里看到的只是一些API而已。对于任何API，你都可以用已经编写好的代码来实现它，也可以自己从头开始写。我们一个个来看。

```

算上HttpRequest对象，这个URL引用的archive视图函数的名字应该是这样的：
from django.http import HttpResponse

def archive(request, year, month, day, show_private):
 url(r'^archives/(\d+)/(\d+)/(\d+)/$', 'archive',
 {'show_private': True}),
 urlpatterns = patterns('myproject.myapp.views',
 url(r'^archives/(\d+)/$', 'archive',
 {'show_private': False}),
)

```

本派工來說，`Django`挑出圖像應由 `Image` 類管理。對於圖像數的唯一要求就是它必須具有不同的字典數，就構成了視圖函數所要的所有參數。像這樣（和前面的例子稍再加一個可選的字典參數，就構成了視圖函數所要的所有參數。像這樣（和前面的例子稍

虽然视图本身轻量级能占据大部分内存空间，但是 Django框架里有关视图的代码量倒是相当有限。视图代表的是你的商业逻辑，所以对Web应用程序员来说这部分应该需要最少的胶水代码，而大部分工作应交给程序员。同时，Django内置的通用视图（generic view）是Web开发框架里最能为你节省时间的工具之一，我们会和自己定义视图一起详尽地介绍它们以及使用的方法。

- `date-based_*`: 一组基于日期的通用视图，强调了 Django 原本是一个面向出版界框架的之外不需要任何其他商业逻辑。
- `create_update_object_create_update_object`: 对简单对象的创建和更新很有用，你只需要在表单或是模型（分别参见第6章和第4章）里定义表单验证就行了，除此之外不需要任何其他逻辑。
- `list_detail_object_list_detail`: 这两个提供了大多数 web 应用里只读的部分，而且可能用的最多的通用视图，因为显示信息通常不需要什么复杂的逻辑。但是，如果你要执行一些逻辑来准备模板 context 的话，最好还是使用自定义视图吧。
- `simple_direct_to_template`: 对含有动态内容（如简单页面 template），也就是静态 HTML 页面相对，详见第8章）但是没有特定 Python 逻辑的模板很有用，例如首页或是不分页/混合页。
- `template_name`: 允许用户重写视图模板的默认位置，而 extra\_context 则允许用户向模板传递额外的信息。（关于模板和 context 的细节请参见第6章。）Django 的官方文档中可以查到所有通用视图及其参数的细节，这里我们只讨论其中最常用的一些。注意为简洁起见，通用视图都是组成一个顶层的模块层次。

通过视图一般都提供了很多选项。有些是某个视图特有的，但是其他的都是全局的，比如 `template_name` 允许用户重写视图模板的默认位置，而 `extra_context` 则允许用户向模板传递额外的信息（详见第8章）。但是没有特定 Python 逻辑的模板很有用，例如首页或是不分页/混合页相对，详见第8章）但是没有特定 Python 逻辑的模板很有用，例如首页或是不分页/混合页。

Django 的核心团队希望给你尽可能多的余地，哪怕最后的功能有时候乍看起来不太直观。

`Person.objects.filter(is_employee=True)` 而不是 `Person.objects.all()` 等 `object-detail`。

此外，再来一个完整的 `QuerySet` 而不是模型类（这是另一种指定我们需要的兴趣类型的办法），Django 允许我们自己进行过滤。例如，一个只显示员工细节的页面可以这样写：

```
QuerySet可以（也通常）在真正执行数据库查询之前用filter或exclude排除一部分结果。因此，你可以放心地使用object-detail通用视图会过滤掉特定的对象，把查询的结果限制在合理的范围内。
```

传遍 `Person.objects.all()` 上去非常不高效，因为就这样执行的话，`QuerySet` 会将所有一个它可以通過 ID 传递查找对象的 `QuerySet` 作为参数。在这里，我们通过 URL 提供参数，并通过字典参数提供 `queryset`。

在上面的例子里，我们定义了一个正则表达式来匹配 `/people/25/` 这样的 URL，这里 25 代表的是我们要显示的 Person 记录在数据库里的 ID。`object-detail` 通用视图需要一个 `object_id` 参数和一个它可以通過 ID 传递查找对象的 `QuerySet` 作为参数。在这里，我们通过 URL 提供参数，并通过字典参数提供 `queryset`。

```
from myproject import myapp, models import Person
```

最后，我们前面说过，有时候就是没办法使用通用视图，还记得本节开始的地方么，视图函数本身只是一块空白画布，它只是遵循了一个简单的API，你可以在里面填充任何东西。这

自定義視圖

可以看，虽然图像数据接受了一个在URL正则表达式的命名组里定义的last-name参数，但是我们还是把99%的实际工作交给了通用视图去处理。之所以可以这么做是因为通用视图就是普通的Python函数，一样可以指导人和调用。人们很容易陷入一种怪圈，认为框架自身是一种语言，其实就和我们前面多次强调的那样，所有一切都是Python，并且刚才这种技术更显出为什么这种设计思想是出色的。

```

urlls.py

from django.conf.urls import url, patterns
urllibpatterns = patterns('myapp.views',
 url(r'^people/by-Lastname/(?P<Last_name>[w]+/$)', 'last-name-search'),
)

views.py

from django.views.generic.list_detail import object_list
from django.shortcuts import render_to_response
from myapp.models import Person
from myapp.forms import SearchForm

def last_name_search(request):
 if 'q' in request.GET:
 queryset=Person.objects.filter(last_name=request.GET['q'])
 return render_to_response('search_results.html', {'queryset':queryset})
 else:
 return render_to_response('search_form.html', {'form':SearchForm()})

def search(request):
 form=SearchForm(request.GET)
 if form.is_valid():
 last_name=form.cleaned_data['last_name']
 queryset=Person.objects.filter(last_name__contains=last_name)
 return object_list(request, queryset=queryset, extra_context={'form':form})
 else:
 return render_to_response('search_form.html', {'form':form})

```

有时候直接从URLconf文件里调用通用视图还不够用。一般这种情况下，只能重新写一个自定义的视图，但同时，根据需要通用视图还是可以用来帮助你干一些底层的体力活的。

半導用視圖

通用视图也是一把双刃剑。其优点非常明显，它们能在简单视图到中等复杂的视图里，为你节约大量的时间和绝大部分的工作。通过把它们包裹在自定义视图里，还能进一步扩展它们的能力，下面我们就来说到这个话题。但是再强大它也有局限的时候，有时候你还需要自己重新写一个自定义的视图出来不可，哪怕某个通用视图已经能满足90%的要求了。知道什么时候放弃而采用自定义视图是一项很有价值的功能，和软件开发的很多方面一样，只能通过经验慢慢积累。

它们对任何基于日期的数据类型都非常有用。包括面向日期的索引和细节页面加身。

里我们将和你分享一点我们自己的经验，并指点一些框架提供的方便快捷的函数给你。不过总的来说，在这块领域里，只有你自己的技术和经验才能决定下一步怎么走。

### 框架提供的捷径

就像刚才说的，一旦开始编写自定义视图，Django基本上就不再对你做任何限制了。但是，它还是提供了一些捷径给你，其中绝大多数都属于`django.shortcuts`模块。

- `render_to_response`: 这个函数替代了两步到三步的过程，从创建一个Context对象，然后用Template渲染，最后返回包含结果的`HttpResponse`。它接受一个模板的名字，一个可选的context对象（也可以是字典）以及MIME类型，并返回一个`HttpResponse`对象。模板渲染会在第6章里介绍。
- `Http404`: 这个Exception子类会返回一个HTTP 404错误码并且渲染一个顶层的404.html模板（除非你在`settings.py`里重写它）。使用起来和`raise`其他异常没什么两样，所以当你遇到一个404的时候，它就是一个标准的错误，就和你试图把一个字符串加到一个整数上时发生的错误是一样的。这个类定义在`django.http`模块之下。
- `get_object_or_404`和`get_list_or_404`: 这两个函数就是把两个步骤合二为一：获得一个对象或者列表，如果查找失败则抛出`Http404`。它们接受一个klass参数（它很灵活，能够接受一个模型类，一个Manager，或者一个`QuerySet`）和一些数据库查询参数，例如那些要传递给Manager和`QuerySet`的参数，并试图返回所需的对象或者列表。

这里是两个使用这些快捷方式的例子：第一个直接使用`Http404`，而第二个则展示了如何用`get_object_or_404`来一步完成同样的目标——这两个函数在实际中的行为完全相同。先不要管这里的模板路径，在第6章里会详细解释它们的。

这里是手动抛出404异常的方法：

```
from django.shortcuts import render_to_response
from django.http import Http404
from myproject.myapp.models import Person

def person_detail(request, id):
 try:
 person = Person.objects.get(pk=id)
 except Person.DoesNotExist:
 raise Http404

 return render_to_response("person/detail.html", {"person": person})
```

这个例子则是使用`get_object_or_404`，通常你都会用这个方法来替代前面那个：

```
from django.shortcuts import render_to_response, get_object_or_404
from myproject.myapp.models import Person

def person_detail(request, id):
 person = get_object_or_404(Person, pk=id)

 return render_to_response("person/detail.html", {"person": person})
```

其他观察

很多Django程序员发现他们在定义自己的视图函数时总是会用到“args/kwarg”方式。在第1章里就说过，Python函数定义\*args和\*kwarg后就能接受任何位置参数和关键字参数。虽然这是一把双刃剑（出色的文档通常能帮助你明白你的函数签名，不过这里就没有了），但是这确实是一个很实用的技巧，既灵活又快捷。你不用再回到URLconf文件去记住你是怎么给捕获正则表达式参数或是关键字参数命名的，你只需要这样定义函数：

```
Here we can refer to e.g. args[0] or kwargs['object_id']
def myview(*args, **kwargs):
```

然后什么都不用管，直接在需要的时候引用kwargs[“identifier”]就行了。不知不觉地你就会不由自主地使用它，当你要把函数的参数传递给一个代理函数的时候（例如在前面的“半通用”视图里提到的）也会容易得多。

很多Django程序员发现他们在定义自己的视图函数时总是会用到“args/kwarg”方式。在

第1章里就说过，Python函数定义\*args和\*kwarg后就能接受任何位置参数和关键字参数。虽然这是一把双刃剑（出色的文档通常能帮助你明白你的函数签名，不过这里就没有了），但是这确实是一个很实用的技巧，既灵活又快捷。你不用再回到URLconf文件去记住你是怎么给捕获正则表达式参数或是关键字参数命名的，你只需要这样定义函数：

```
Here we can refer to e.g. args[0] or kwargs['object_id']
def myview(*args, **kwargs):
```

然后什么都不用管，直接在需要的时候引用kwargs[“identifier”]就行了。不知不觉地你就会不由自主地使用它，当你要把函数的参数传递给一个代理函数的时候（例如在前面的“半通用”视图里提到的）也会容易得多。

## 5.4 总结

到此我们探索Django基础设施组件的旅程就已经完成一半了。除了在第4章里描述的模型中间件的使用。你还看到了如何把简单的Django视图函数组合在一起，另外还展示了通用视图以及使用。这一章也是本书里这部分的最后一章，第6章描述了框架里第三个主要的部分，即通过模型查询类网页以及用表单和表单验证管理用户的输入。之后，你就会进入到第三部分，看看这些概念是如何在那4个示例应用程序里使用的。

原因。这里是一个context处理器的例子：

另一个给模板context提供数据的方法就是通过context处理器（processor），这是框架里一个类似中间件的部分，你可以在那里定义各种函数，在模板渲染之前来把键值对附加到所有的context上去。这就是说框架这样的特性为什么能保证特定的数据在全局范围内都能够访问到的。另外一个是将context提供数据的方法就是通过context处理器（processor），这是框架里一个context处理器不是那个动态了。

不过如果是这样的话，你应该考虑去用flagepages context应用程序——毕竟一个没有context的模板作为参数传给帮助函数render\_to\_response。技术上来说，你可以用空context来渲染一个模板，或者更常见的（比如自定义视图），你就得在调用模板render方法的时候，自己准备context，或者最常见的时候（比如先准备好（比如在通用视图里），你只要把extra\_context参数附加上去就行了。而其他时候为件事先准备好（比如在通用视图里），你只要把extra\_context参数附加上去就行了。而其他时候第2章和第5章里我们都看过了，每渲染一个模板都需要有一个context。有时候context会出现在渲染的时候表示了一个类的文字的Context对象。

模板作为一个渲染模板的信息成为context——一个模板的context基本上就是一个包含键值对的字典Django里把模板作为一种动态文档，只有在显示的动态信息的时候才能证明自己的存在。

## 理解Contexts

它能帮助你充分发挥Django模板的能力。

最后，请注意虽然模板系统通常被用来生成HTML，而是可以用来生成log文件、E-mail正文、CSV文件等任何文本格式。别忘了这一点，HTML，Django里把模板作为一个渲染模板的信息成为context——一个模板的context基本上就是一个包含键值对的字典。在渲染的的时候表示了一个类的文字的Context对象。

Django的模板语言是设计给前端工程师使用的，他们不一定是最程序员。正因为如此，再加上传输和表现分开的愿意，模板语言绝对不是嵌入式的Python。不过，Django程序员还是会通过可扩展的标签（tags）和过滤器（filter）系统来扩展模板语言的逻辑结构（后面会说到）。

Django的模板语言是设计给前端工程师使用的，他们不一定是最程序员。正因为如此，再加上传输和表现分开的愿意，模板语言绝对不是嵌入式的Python。不过，Django程序员还是会通过可扩展的标签（tags）和过滤器（filter）系统来扩展模板语言的逻辑结构（后面会说到）。

在之前的章节里已经我们已经接触过，模板是一种独立的文本文件，同时包含了静态内容由视图函数本身（通过显示地渲染或者使用render\_to\_response）或者视图的参数（比如通过视图函数标记的逻辑、循环和数据的显示等。使用哪个模板以及渲染什么数据是（比如HTML）和动态标记的逻辑、循环和数据的显示等。使用哪个模板以及渲染什么数据是固定的template\_name参数）决定的。

学习完Django的数据模型和逻辑处理后，我们还剩下最后一个部分：如何显示信息和管理用户输入。我们先介绍一下Django的模板语言和渲染系统，然后在第二部分介绍表单和表单处理。

## 第6章 模板和表单处理

```
<object object at 0x40448>
<>>> print object()
```

正好和HTML的标签格式相同，即文本都包含在< 和 >尖括号里。  
\_unicode\_方法的对象，在模版里是看不到它们的。这是因为Python默认认为一个对象的形式  
他非字符串变量都会被尽量转换成（Unicode）字符串。小心——如果你试图打印没有定义  
在适当你模版里输出context变量的时候，它会像其他调用unicode方法，所以对象以及其

```
</html>
</body>

(% endfor %)
{{ item }}
{% for item in object_list %}

<body>
</head>
<title>{{ title_text }}</title>
<head>
<html>
```

“object\_list”: [“One”, “Two”, “Three”] } ) 的context的例子。  
这里有两个约定用法，都和大括号有关。变量输出用的是双大括号 ({{ variable }})，而其他则  
模块级命令 (block-level command)，通常是以命令 “if” 或者 “for” 等。Django的模版语  
大部分都是松耦合的，所以只要你愿意，完全可以换用另一种模版库。

Django模版语言和其他一些非XML的模版语言 (比如Smarty或是Cheetah) 比  
起来的区别在于，它没有要保持HTML兼容的意思，而是用特殊的字符来把模版变量以及  
语句命令和静态内容 (通常是HTML) 分开。和Django的其他部分一样，模版语言和框架的其  
他部分都是松耦合的，所以只要你愿意，完全可以用另一种模版库。

## 模版语言语法

应用。

这就是它们在元组里指定的顺序是有讲究的，context处理器会按照在设置变量里的顺序逐个  
的位置是一个名为TEMPLATE\_CONTEXT\_PROCESSORS的元组。另一个和中间件相似的地方  
context\_processors.py文件可以放在项目的根目录里，或者是应用程序的目录里。

你可以把context处理器函数在任何地方，不过通常标准化总是有好处的，比如  
context\_processors.py文件可以放在项目的根目录里，或者是应用程序的目录里。  
看上去好像不太有用 (网站指引很少有这么简单的)，不过它展示了context处理器的简单

```
def breadcrumb_processor(request):
 breadcrumbs = request.path.split('/')
 return {
 'breadcrumbs': request.path.split('/')
 }
```

在上面的例子你可能已经注意到了，虽然变量输出和过滤器都很有用，但是真正强大的确实是在标签 (tag) —— 到目前为止我们已经见过它们用来自循环字符串或对象的列表，但是它们还能执行逻辑操作 (如 if %{}, %ifequal {}, %block {}、%include{})。

## 标签

```
</table>
 <tr>
 <td><{ person.name }></td>
 <td><{ person.isAvailable|yesno:"Yes,No" }></td>
 </tr>
<% for person in person_list %>
<tr>
 <th>Name</th>
 <th>Available?</th>
</tr>
</table>
```

虽然大多数过滤器值接受一个字符串输入，不过有些也能进一步地参数化其行为，比如 yesno 这过滤器可以接受任何值（一般是布尔值）并且打印出可读的字符串。

```

 <% for string in string_list %>
 <{ string|lower }>
 <% endfor %>

```

Django 提供了各种各样的过滤器来封装 Web 开发中常见的文本处理工作，例如转义斜杠符号、大写首字、格式化日期、获取列表或元组的长度，以及连接字符串等。这里一个如何使用过滤器来把一个字符串列表转换成小写的例子。

第 11.5 节里你会看到这种方法，过滤器也是 Python 整数。

(1) 因为它们总是接受一个文本输入并返回一个文本输出，所以可以把它们串在一起使用。在什么又是管道的话，请参阅附录 A。过滤器连语法是直接跟 UNIX 管道，它用的也是管道符 |。通过过滤器 (filter) 的机制来转换 context 变量，它和 UNIX 里的管道有点相似——要是你不知道什么是管道的话，[请参阅附录 A](#)。过滤器连语法都是直接跟 UNIX 管道，它用的也是管道符 |。要是你不知道什么是管道的话，[请参阅附录 A](#)。过滤器连语法都是直接跟 UNIX 管道，它用的也是管道符 |。

## 模板过滤器

你可以看到，虽然 Django 模板的语法不是语义正确的 HTML，但是大括号的语法很容易就能让人分辨出哪里是输出语句以及命令的部分，哪里是静态的内容。此外，Django 的开发团队打算的是用模板语言生成文档类型而不仅仅是 HTML，所以他们觉得设计一个针对 XML 输出的模板系统是行不通的。

即使是有经验的 Django 用户有时候也会踩到这个常见的陷阱里去，所以如果你发现要显示的东西没显示出来，第一个要确认的就是那个对象以及那个对象的字符串表示是什么！

扩展这个顶层模板（例如添加第二层的菜单指引），最后，每个单独的站点里的底层模板再将其一个顶层的或全局可见的模板标出页眉/页脚和全局引导，然后在每个子小节的中层模板里，模板的顶部调用，并告知渲染引擎这个模板是从一个更高级的模板继承而来。例如，你可以在

模板继承是通过`% extends %`和`% block %`两个模板标签实现的。`% extends %`必须在

调用和开发变得很困难。

有助于逻辑化的组织模板。而包围，虽然也很有用，却也很容易导致所谓的“include soup”，让`(inherence)`和包含(`inclusion`)来完成代码组合和重用。我们先介绍继承，因为它通常更有模板里有一组很有用的标签，它们可以嵌出当前的模板和其他模板文件交互，通过继承

`blocksAndExtends`

`tagsDjango官方网站`。

`((forloop.counter))`和`((forloop.counter0))`，分别是从1或0开始)等。更多信息和例子请知道这是不是循环里的第一个或是最后一个迭代)执行什么操作，或者是显示循环计数相关的操作，在循环开始和结束的时候`((forloop.first))`或`((forloop.last))`分别返回有`((forloopcontext)`，它提供了各种属性允许你根据使用的属性和所处的循环迭代来执行不同的操作。这种属性提供了各种属性允许你根据使用的属性和所处的循环迭代来执行不同的操作。这种属性

最后关于标签的一点：模块级命令(`block-level`)如`% if %`和`% for %`可以修改它们的

可用(Django官方网站提供了一份出色的列表)就能避免重新发明轮子啦。

希望这个例子展示了Django内建的过滤器和标签的灵活性之处。事先多知道一点有什么

```
(% endfor %)

(% endfor %)
((item))
(% for item in object_list %)

(% if object_list|length>10 %)
```

当然我们还可以用`length`过滤器，它接受一个列表和参数并返回一个布尔值。

```
(% ifequal object_list|length 10 %)

(% endfor %)
((item))
(% for item in object_list %)

```

下面的例子是如何在迭代列表的内容前检查它的长度。

标签从技术上来说是不定型的，可以在标签名后面跟任何形式的输入(细节请参见第11.5节)，但是内置标签和绝大多数用户创建的标签都尽量遵循一定的约定，通常是一列由空格分隔的参数。很多标签参数可以是context变量，而且事实上大多数时候，还可以用过滤器。比如，

`%和% extends %，参见下一节)，等其他一系列任务。`

展示层的模板来提供页面的实际内容。

{% block %}是一个模块级的标签，可以用来定义模板里预备让那些要扩展它的模板去填充的小节。虽然模板的儿子会通常都会使用这些块（block），但其实这不是必需的。它们可以被忽略掉（这样父模板里有什么它就会显示什么）或是进一步代理给更低一层的模板。下面这个简单的例子使用了之前提到的三层网站布局，它的URL有/、/section1/、/section2/、/section1/page1/和)section1/page2/。

我们暂时先忽略根站点的首页，而专注于最底层的“叶子”页面。这里，base.html提供了顶层的封装结构，section的模板提供了页面的标题（指明了用户所处的是站点的哪一部分），而page的模板则提供了simple content。

base.html:

```
<html>
 <head>
 <title>{% block title %}My Web site{% endblock %}</title>
 </head>
 <body>
 <div id="header">
 Section 1
 Section 2
 </div>
 <div id="content">
 {% block content %}{% endblock %}
 </div>
 <div id="footer">
 About The Site
 </div>
 </body>
</html>
```

section1.html:

```
{% extends "base.html" %}

{% block title %}Section 1{% endblock %}
```

section2.html:

```
{% extends "base.html" %}

{% block title %}Section 2{% endblock %}
```

page1.html:

```
{% extends "section1.html" %}

{% block content %}This is Page 1.{% endblock %}
```

```
page2.html:

{% extends "section1.html" %}

{% block content %}<p>This is Page 2.</p>{% endblock %}
```

前面例子里的模板都设置好后，用户在访问)section1/page2/的时候就能在浏览器里看到下面的内容：

```
<html>
 <head>
 <title>Section 2</title>
 </head>
 <body>
 <div id="header">
 Section 1
 Section 2
 </div>
 <div id="content">
 <p>This is Page 2.</p>
 </div>
 <div id="footer">
 About The Site
 </div>
 </body>
</html>
```

模板继承的优点就在于很容易在模板层次里穿梭来找到什么模板生成了任何给定页面里的哪部分HTML。此外，和基于包含的那种方式即在每个子页面里都要包含页眉、页脚、边栏等代码的方式比起来，继承能够节省很多代码量。

#### 包含其他模板

尽管模板继承有这么多优点，不过模板包含依然有其立足之地。有时候你需要重用的一块HTML或者其他文本没法很好的符合继承的需要，比如一个常用的分页元素。Django支持{% include %}包含（用法完全可以猜到），它接受要包含的模板文件的名字，然后把自己替换成这个文件的内容。包含文件自己就可以是一个完整实现的Django模板，它们的内容是遵循包含模板的context来解析的。

除了{% include %}，Django还提供了{% ssi %}标签（这里的ssi代表的Apache家族的SSI，即服务器端包含，Server Side Includes）。{% include %}和{% extends %}引用的是在settings.py里指定的template目录中定义的局部模板文件。而{% ssi %}使用的则是文件系统绝对路径。不过为了安全期间，{% ssi %}被限制在一组特定的目录里，它可以通过settings.py文件里的ALLOWED\_INCLUDE\_ROOTS变量指定。

最后还要注意{% extends %}和{% include %}既可以接受context变量也能接受字符串，这就等于模板可以动态地决定它们要包含的模板或是从什么模板继承了。

虽然模板很适合把信息显示出来，但是把它输入到数据库就是完全另外一回事了，这需要创建并验证HTML表单以及保存提交的信息。Django提供了forms库来把框架里三个主要的组件联系在一起：模型里定义的数据字段，模板里显示的HTML表单标签，还有检验用户输入和显示错误信息的能力。

在编写本书的时候，Django的表单机制还处在一个不断变化的过程，这里我们讨论的库在叫做newforms，它是一个模块化的办法，比它的先驱oldforms要进步的多（虽然现在import django.forms返回的还是这个旧的库）。我们在这里讨论的是newforms，并且引用（如果要使用Form类完成）`from django import newforms as forms`

是和给定模型100%匹配的表单，不过有一个分开的Form类也在一定程度上增加了灵活性。

有了这个单独的实体，你可以隐藏或是省略特定的变量，或是把多个模型里的变量放在一起。当然，有时候你还会希望可以处理和数据绑定在做完全没关系的表单，这些全都可以通过Form类完成。我们先来看一个简单的例子。

```
class PersonForm(forms.Form):
 first = forms.CharField(max_length=100, required=True)
 last = forms.CharField(max_length=100, required=True)
 middle = forms.CharField()

class PersonForm(forms.ModelForm):
 class Meta:
 model = Person
 fields = ('first_name', 'last_name', 'middle_name')
```

虽然这个看起來有点类似于之前的模型类，不过它完全是一个独立的表单，只不过它们正好有着相同的变量而已（除了这些都是forms.Field实例而非models.Field实例）。表单变量接受的参数数和模型里也是非常类似的。

上面的例子定义了一个由三个文本变量组成的表单，在验证的时候（见下面），如果first和last变量没有被填入的话它就会报错。此外它还能让你把这些变量都不超过100个字符长。因此数字字段类型和表单变量类型有很多重叠的地方——如果你想知道在本章的示例之外还有什么区别，[请查看官方文档里表单Field类的详细列表](#)。

为了保持DRY原则，Django允许你使用一种特殊的Form类型ModelForm来为任何模型类或实体取得一个Form子类。一个ModelForm和一个普通的Form基本上一模一样，但是它包含了实现于模型类的表单方法。如果想知道自己模型类的表单方法，可以在模型类或实体里取得一个Form子类。

```
form = PersonForm({'first': 'John', 'middle': 'Quincy', 'last': 'Doe'})
```

时候才会发生。调用对象的save方法了。这个例子只访问了一次数据库而不是两次，这只有当commit=False的不是真的要更新数据库。设置为False后它还是会返回模型对象给你，不过这是你就要自己负责为了提供这类灵活性，save方法接受一个可选的commit参数（默认是True），它可以控制是否将数据已经被转换成Python值了。

（不过还是需要在它到这模型层之前就修改它，但是有时候在表单完成验证后再修改可能会更简单一点POST字典被交给表单之前就修改它；但是有时候在表单完成验证后再修改可能会更简单一点你经常需要在数据从表单提交后到它存储到数据库之前修改它们。有时候你可以直接在

```
Will result in the _unicode_() output for the new Person
print new_person

new_person = form.save()

form = PersonForm({'first': 'John', 'middle': 'Quincy', 'last': 'Doe'})
```

的例子（这个过程的细节会在本章稍后讲到）：

起来，但是现在只要记住save方法能将POST字典到数据库创建（或更新）的数据。继续前面一旦你读到了如何把信息填入到表单并且如何验证它的時候，这个方法的含义就会变得更明显。加入验证成功的話，可以把它们的信息保存为数据库里的一条记录，然后返回Model对象结果。用这种方式生成的表单和手动生成的表单有一个重要的区别，就是它们有一个save方法，用来ModelForm

（required=False）。  
认为ModelForm后它会把你Model类里的变量“复制”到Form变量里来。通常这个过程都是很直观的（一个模型CharField变成了一个表单TextField或是ChoiceField，如果它是实现了choicess的话），除了Django官方上专门列出的一些警告。此外要记住的最重要的事就是变量默认都是required=True的，除非模型把它们设置为blank=True，这时它们才变成可选的变量

这样。这个方法突出了数据定义和数据输入和验证的分离，并且提供了很大的灵活性。

一般你需要为你创建的每一个模型类都至少定义一个ModelForm，哪怕它简单到这个例子

```
class PersonForm(forms.ModelForm):
 class Meta:
 model = Person
 fields = '__all__'

from django import newforms as forms
from myproject.myapp.models import Person
```

个Meta嵌套类（和模型类里的Meta类似），Meta类里有一个必需的属性model，它的值就是所屬的Model类。下面的例子和前面定义的普通Form在功能上是一致的：

的例子里你定义了一个名为Person的表单。ModelForm的Meta类允许你定义两个可选的属性，fields和exclude，这两个是要包含或排除的变量名的列表或是元组（当然了，这两个你每次只能用其中一个！）。例如，下面有時候你会希望自定义表单，而不是作为一个模型的复制品。隐藏特定的变量是一个很常见的需求，稍微少见一点的是排除或是包含所有变量。通常你不需要重新写一个Form子类出来说，因为有几种方式用ModelForm来完成这个任务。

你可以看到，考慮立刻保存还是延迟保存的需求给save方法的使用增加不少复杂度。好在区別于Model

这种复杂度不是必需的，绝大多数情况下你都可以直接save而无需过多担心。好在

```
new_person = form.save(commit=False)
We get a Person object, but the database is untouched.

Update an attribute on our un-saved Person.
new_person.middle = 'Danger'
Update an attribute on our un-saved Person.

After we save to the DB, our Person exists and can be referenced by
the related objects.
So now we save them as well. Don't forget to call this, or your related objects
will mysteriously disappear!
So now we save them as well. Don't forget to call this, or your related objects
will mysteriously disappear!

form = PersonForm(input_including_related_objects)
objects related to the primary one via the ManyToManyField.
These related objects can't be saved at this point, so they are
deferred till later.
These related objects can't be saved at this point, so they are
deferred till later.

new_person = form.save(commit=False)
Update an attribute on our un-saved Person.
new_person.middle = 'Danger'

After we save to the DB, our Person exists and can be referenced by
the related objects.
So now we save them as well. Don't forget to call this, or your related objects
will mysteriously disappear!
```

所以，当一个ModelForm包含了相关对象的信息，并且你又使用了commit=False时，ModelForm会向数据库插入主对象和相关对象的数据。因为关系数据库要求目标记录必须在插入之前就存在，所以不可能在保存相关对象的时候却还想保存你的主要对象。另一个和commit=False有关的场景是当你使用内联编辑的相关对象的时候。这时，一个Django会给你（不是Model对象结果！）添加一个额外的方法save\_m2m，它让你能正确地安排事件的顺序。在这个例子中，假设Person模型有一个自引用的多对多关系。

```
Now we can update the database for real.
Update an attribute on our un-saved Person.
new_person = form.save(commit=False)
Update an attribute on our un-saved Person.

We get a Person object, but the database is untouched.

new_person = form.save(commit=False)
This input to PersonForm would contain "sub-forms" for additional Person
objects related to the primary one via the ManyToManyField.
These objects related to the primary one via the ManyToManyField.
These objects related to the primary one via the ManyToManyField.

new_person = form.save(commit=False)
Update an attribute on our un-saved Person.
new_person.middle = 'Danger'
```

```
from myproject.myapp.models import Person

class PersonForm(forms.ModelForm):
 class Meta:
 model = Person
 exclude = ('middle',)
```

如果一个Person只有first、middle和last三个变量，那下面用fields实现的例子和上面这个功能一模一样：

```
class PersonForm(forms.ModelForm):
 class Meta:
 model = Person
 fields = ('first', 'last')
```

这里要注意的非常重要的一点是，当调用这样的表单的save方法时，它只会试图保存它知道的变量。如果你忽略了一些变量而模型里它们又必须出现的话这就会产生问题了。所以对这样的变量一定要再三确认它们要么被标志为可选的（null=True），要么在default参数里定义了默认值。

除了决定要显示模型里的哪些变量，你还可以重写forms层里的Field子类，它负责验证和显示特定的变量。做法很简单，直接显式地定义它们即可，就像在本章开头的地方看到的那样，它会自动重写从模型里取出来的定义。这对传给表单层的Field（比如max\_length或required）或者是修改类本身（可能是要把TextField显示为CharField，又或者通过把一个CharField变成ChoiceField来添加一些选择等）都很有用。例如，下面的例子就把名字变量的长度收紧了一点：

```
class PersonForm(forms.ModelForm):
 first = forms.CharField(max_length=10)

 class Meta:
 model = Person
```

这里还值得一提的是关系表单变量（relationship form fields），ModelChoiceField和ModelMultipleChoiceField，它们分别对应ForeignKey和ManyToManyField。虽然你可以在模型层变量里指定limit\_choices\_to参数，其实你还可以在表单层变量里自定一个queryset参数（很自然的，它接受的是一个特定的QuerySet对象）。这样，你就可以在模型层重写这个限制（或缺陷）并且自定义ModelForm。就像下面的例子里，我们假设Person模型有一个指向其他Person对象的没有限制的父ForeignKey：

```
A normal, non-limited form (since the Model places no limits on 'parent')
class PersonForm(forms.ModelForm):
 class Meta:
 model = Person

A form for people in the Smith family (whose parents are Smiths)
class SmithChildForm(forms.ModelForm):
 parent = forms.ModelChoiceField(queryset=Person.objects.filter(last='Smith'))
```

```
class Meta:
 model = Person
```

### 继承表单

很多时候，不管是普通的Form还是ModelForm，你都需要利用Python面向对象的特点来避免重复。Form的子类还可以继续被继承，其结果会包含父类里的全部变量。例如：

```
from django import newforms as forms

class PersonForm(forms.Form):
 first = forms.CharField(max_length=100, required=True)
 last = forms.CharField(max_length=100, required=True)
 middle = forms.CharField(max_length=100)

class AgedPersonForm(PersonForm):
 # first, last, middle all inherited
 age = forms.IntegerField()

class EmployeeForm(PersonForm):
 # first, last, middle all inherited
 department = forms.CharField()

class SystemUserForm(EmployeeForm):
 # first, last, middle and department all inherited
 username = forms.CharField(maxlength=8, required=True)
```

还可以“混合”操作，即多重继承。

```
class BookForm(forms.Form):
 title = forms.CharField(max_length=100, required=True)
 author = forms.CharField(max_length=100, required=True)

class InventoryForm(forms.Form):
 location = forms.CharField()
 quantity = forms.IntegerField()

class BookstoreBookForm(BookForm, InventoryForm):
 # Has title, author, location and quantity
 pass
```

在用这种方法继承ModelForm的时候，记住你还可以修改Meta的属性，通常是更新或是往fields或exclude里添加新的值来进一步限制可用的变量。

### 填写表单

在Django的表单库里，任何给定的表单实例要么是绑定的（bound），即和数据有关系的，要么是没绑定的（unbound），即没有数据的。没绑定的空表单主要是用来生成空的HTML表单来让用户填写，因为你没办法验证它们（除非一张空白的表单是所需要的输入，这种情况几乎没有），而且基本上也不会想要把它们的内容保存到数据库里去。绑定的表单则是大多数操作的所在。

使用类似于initial的数据的好处在于它的值可以在类单例键的时候才被调用出来。例如，这

如果用來實例化 PersonForm 的 initialiaze 數是 {first: 'John', last: 'Dee'}，那么实例化上 (last) 的值 "Dee" 就会覆蓋掉类义层上的表单定义里的 "Smith"。

卷五

```

class PersonForm(forms.Form):
 first = forms.CharField(max_length=100, required=True)
 last = forms.CharField(max_length=100, required=True)
 middle = forms.CharField(max_length=100)

 def process_form(self):
 if self.is_valid():
 form = PersonForm(initial={'first': self.cleaned_data['first'],
 'last': self.cleaned_data['last'],
 'middle': self.cleaned_data['middle']})
 if self.request.method == 'POST':
 # Display the form, noting it was submitted.
 return render_to_response('myapp/form.html', {'form': form})
 else:
 # Just display the form, not processing it.
 return render_to_response('myapp/form.html', {'form': form})

```

from djangos shortcuts import render\_to\_response  
from django import template newforms as forms

所以可以在填写表单的时候修改它们任何一个的值。

这里是一个创建表单的例子，它修改之前自定义的PersonForm，这里把姓氏变量填为“Smith”（通过表单定义）并在运行时把名字变量设为“John”（创建表单实例时）。当然，用

你还可以创建一个虽然未绑定，却在模板被打印的时候显示了初始值的表单。这个命名构造函数参数（named constructor argument）initial 是一个字典，和用于绑定的位置参数（positional argument）是一样的。每个表单变量都有这样一个类似的参数，允许它们指定自己的默认值，不过要是有冲突的话，表单级字典会覆盖掉它们。

需要在意的是你可以以增加在表单的数据字段里添加额外的链接对，表单会自动无视那些和它们定义的变量没关系的输入。

```
def process_form(request):
 form = PersonForm(post)
 post = request.POST.copy() # e.g. {Last: 'Doe', First: 'John'}
```

from myProject.myapp.forms import PersonForm

我们可以根据前面和Person关联的ModelForm类生成一个熟悉的表单，把它嵌入到一个可以变成或菜单或理想图画数的开头。Request.POST.copy()在这里不是必需的，但是我们推荐使用它。在万一需要的时候，你可以在保持请求里原有内容不变的同时修改改自己的字典项。

数据通常会随表单填写而变化，从而导致表单的校验逻辑也必须随之变化。如果表单的校验逻辑是硬编码在表单类中，那么每次表单逻辑发生变化时，都需要手动修改表单类的校验逻辑，这不仅效率低下，而且容易出错。因此，我们建议将表单的校验逻辑分离出来，放在一个专门的校验类中，这样可以在表单逻辑发生变化时，只需要修改校验类的逻辑，而不需要修改表单类的逻辑。

卷之三

为简单起见，我们看略了一个视图函数其上关心的东西，丢重复杂的处理。注意这里我们如何根据URL里的值来获取relative对象，然后把relative的对象作为表单里last的初始值传递进来的。换句话说，我们把数据都准备好了，这样孩子会自动填写好父亲的名字——如果你的用户要输入很多数据的话，这个功能或许会很有用。

```
view's URL: /person/<id>/children/add/
add_relative(request, *kargs):
def add_relative(request, *kargs):
 # display the form if nothing was POSTed
 if not request.POST:
 # if not request.POST:
 form = PersonForm(initial={last: relative, pk=kargs['id']})
 else:
 relative = get_object_or_404(Person, pk=kargs['id'])
 form = PersonForm(initial={'last': relative.last})
 return render_to_response('person/form.html', {'form': form})
```

让你得以利用在类库或是模型定义的时侯进行小规模的信息，特别适用于采集到的数据。我们来看看如何在这样一个视图函数里利用这个特性，这个视图函数负责添加新的Person记录，而这个Person记录又需要和另一个这样的记录有关联。假设在Person模型里有一个新的自引用的 ForeignKey，我们来为Person定义一个简单的ModelForm。

```
from django.shortcuts import get_object_or_404, render_to_response
from myproject.myapp.models import Person, PersonForm

class PersonForm(ModelForm):
 class Meta:
 model = Person
```

```
 type="text" name="first" maxlength="100" /></td></tr>
<tr><td><label for="id-first">First:</label></td><td><input id="id-first"
 type="text" name="last" maxlength="100" /></td></tr>
<tr><td><label for="id-last">Last:</label></td><td><input id="id-last"
 type="text" name="middle" maxlength="100" /></td></tr>
<tr><td colspan="2" style="text-align: right; padding-top: 10px;">
 <input type="button" value="Submit" />
</td></tr>
```

我们重拾复习一下这个手动创建的PersonForm：

```
class PersonForm(forms.Form):
 first = forms.CharField(max_length=100, required=True)
 last = forms.CharField(max_length=100, required=True, initial='Smith')
 middle = forms.CharField(max_length=100, required=False)
```

这里是一个例子来帮助你理解这些不同的选项以及它们是如何影响HTML输出的。首先，我们再加一个结尾的字符串。

这样生成的：将变量名首字母大写，把下画线换成空格，要是变量名不是以标点符号开始的话，就是这样的，`<label>标签里的文本默认以字母开头`，在本章的最后一部分会介绍到它。另外，这些标签的name和id属性，以及它们对应的`<label>`标签的，`Form类里`，`都是你自己定义的`。最后，虽然这些显示方法大多是在模版里面，但这并不是强制的——要是你有需要，完全可以在Python代码里调用它们来使用。这些方法打印出表单的整个内部结构，不带`<form>`标签、提交按钮，或`<label>`标签。如果你需要更细致地控制输出，还可以单独地显示每一个表单变量。最后，虽然这些显示方法大多是在Django表单变量都知道在HTML标签里要怎么显示自己，这种行为是通过widgets来改变的，`从任何一部分都可以完全控制`。

東美當霄

当执行验证后，表单对象就会得到这两个新属性之一：一个包含错误信息的字典`errors`，或者是原本绑定到表单上的值的“干净”版的字典`cleaned_data`。这两个属性永远都不会同时出现，因为`cleaned_data`只有在表单验证通过的时候才会生成，而只有验证失败的时候`errors`才会被应用。

```
 return render_to_response('person/form.html', {'form': form})
```

你还可以通过表单的auto\_id构造函数参数来改变id属性和<label>标签的行为：False代表彻底不显示id和label；True的话则会使用变量的属性名（就像上面例子一样）来替换格式化字符串（比如'id\_%s'）里的格式化字符。另外标签里结尾的?符号会被label\_suffix参数重写，这个参数就是一个简单的字符串。

下面展示了如何创建一个关闭了auto\_id和label\_suffix（通过把它设为空）的PersonForm实例：

```
pf = PersonForm(auto_id=False, label_suffix='')
```

以及first变量在表单显示时候的样子：

```
<tr><th>First</th><td><input type="text" name="first" maxlength="100" /></td></tr>
```

最后是相同的设置，但是这次为auto\_id和label\_suffix提供了自定义的字符串值：

```
pf = PersonForm(auto_id='%s_id', label_suffix='?')
```

它的输出如下：

```
<tr><th><label for="first_id">First?</label></th><td><input id="first_id" type="text" name="first" maxlength="100" /></td></tr>
```

显示全部表单

默认情况下，打印表单用的是as\_table方法，它会用<tr>和<td>标签按每行两个变量来打印，为了灵活性它会省略<table>标签。as\_table还有一个兄弟as\_p方法，它使用的是段落标签，而as\_ul则会使用列表项标签（但是会忽略包裹它们自己的<ul>标签）。

### 注意

表单省略“外围”的包裹标签，比如<table></table>是因为如果包含它们，就会难以将整个表单继承到模板的HTML里去。提交按钮也是一样——很多模板设计都要求使用不同的方法来提交表单，比如<input type="button" />或<input type="submit" />，所以Django把这个留给你自己决定了。

当使用这种方式显示表单的时候，验证错误也是自动打印的，如果发生错误：就会根据所使用输出方法的不同，在相应的变量旁边显示一个<ul>标签和若干<li>标签。as\_table和as\_ul会在和变量自身同一个标签里显示错误列表（分别是用<td>和<li>标签），而as\_p则会创建一个新的段落来显示错误列表。不管是哪种情况，错误都会显示在前面的表单元素之前。

你还可以通过继承django.forms.util.ErrorList并将子类作为error\_class参数传递给相应的表单来自定义错误列表的显示方式。而且若是想要改变变量/错误列表的显示顺序，只需要重新排列它们在Form类里出现的顺序就可以了——够简单吧。

逐个显示表单

除了刚才介绍的方法，你还可以更细致地控制组织表单。通过表单的字典键可以访问到每一个单独的Field对象，让你可以随时随地显示它们。有了Python的duck typing，你甚至还可以迭代表单本身。不管你通过什么方式获得它们，每一个变量都有自己的errors属性，这是一个

类似列表的对象，字符串表示和之前显示全部方法里的无序列表是一样的（重写的方法也相同）。

在oldforms里，重写一个表单变量默认HTML表单最简单的办法就是访问变量的data属性，在Django表单的眼里，Widget就是一个个知道如何显示HTML表单元素的对象。Django以相似后用自定义的HTML把它包裹起来——这个技巧在newforms里也可以用。不过Widget出现在老forms里，基本上就削弱了它存在的必要性了。

Widget类的构造函数，字符串表示和之前显示全部方法里的无序列表是一样的（重写的方法也相同）。

在Django表单的眼里，Widget就是一个个知道如何显示HTML表单元素的对象。Django以相似的风格为模型Field类和表单Field类各提供了一个大小适中的默认类库Widget类。每个表单变量都专门配备了一个Widget，这样在模板里渲染表单的时候，它的数据就知道怎么显示出来了。例如，CharField默认使用的是一种Widget的一个子类TextInput，它会被渲染成<input type="text"/>。

通常这些默认的变量Widget已经够用了，你甚至都注意不到它们的存在。不过有时候你需要修改一个变量Widget的属性或是把Widget整个都换掉。前面这个情况一般更常见一点，它提供了一种让你修改变量的HTML属性的方法。下面这个例子修改了文本变量的“size”属性；

```
from django import newforms as forms

class PersonForm(forms.Form):
 first = forms.CharField(max_length=100, required=True)
 last = forms.CharField(max_length=100, required=True)
 middle = forms.CharField(max_length=100, required=False)

 def __init__(self, *args, **kwargs):
 super().__init__(*args, **kwargs)
 self.fields['first'].widget.attrs['size'] = 3
 self.fields['last'].widget.attrs['size'] = 3
 self.fields['middle'].widget.attrs['size'] = 3
```

使用这个表单得到的middle变量就会变成这样：

```
<input id="id_middle" maxlength="100" type="text" name="middle" size="3" />
```

你可以这么修改是因为Widget类（比如TextInput）接受了一个attrs字典，它能直接映射到HTML标签的属性上去。在这个例子，我们不需要限制中间名的实际输入长度（用户还是可以输入长达100个字符），但是我们却希望显示长度比默认的短一些。

重写一个变量的默认Widget需要一个Widget类就可以用整个替换默认的Widget，你只要传递另一个Widget类就足够了，例如，可以用一个Textarea来代替TextInput。利用这一点就可以清楚地分离一个变量的显示（Widget）和它的验证行为（表单变量）。这还意味着如果内置的Widget类不能满足你的需要，你还可以定义自己的Widget类。

虽然重写定义Widget的细节超出了本章的范畴，我们还是可以分享一下如何简单快捷地使用Widget类来节省时间的方法。如果你经常需要用到某个特定Widget的attrs字典的话，你就可以在类来继承这个Widget然后给它定义一个默认的attrs字典。

```
from django import newforms as forms

class LargeTextareaWidget(forms.Textarea):
 def __init__(self, *args, **kwargs):
 kwargs.setdefault('attrs', {}).update({'rows': 40, 'cols': 100})
 super(LargeTextareaWidget, self).__init__(*args, **kwargs)
```

上面例子里对字典耍了一些小聪明。setdefault在给定的键存在的情况下会返回现有的值，若给定的键不存在，那么就返回提供的值。但是，它还修改了字典中来永久保存那个值。它用在这里是确保kwargs关键字参数字典一定拥有attrs字典，不管原来的构造函数参数是什么。然后我们就可以用update来更新attrs字典添加我们需要的默认值了。

最终除了默认拥有40行和100列之外，我们这个新的LargeTextarea widget和普通的Textarea一模一样。然后我们就可以为所有希望显示的比普通文本区域大一点的变量应用这个新的widget了。下面这个例子里，假设我们把和表单相关的自定义类保存在一个本地应用forms.py里。

```
forms.py.

from django import newforms as forms
from myproject.myapp.forms import LargeTextareaWidget

class ContentForm(forms.Form):
 name = forms.CharField()
 markup = forms.ChoiceField(choices=[
 ('markdown', 'Markdown'),
 ('textile', 'Textile')
])
 text = forms.Textarea(widget=LargeTextareaWidget)
```

当然你还可以更进一步。因为Field子类的widget参数只是用来设置它的widget属性，所以我们可以继承这个变量本身来让它总是使用我们的自定义widget。

```
class LargeTextareaWidget(forms.Textarea):
 def __init__(self, *args, **kwargs):
 kwargs.setdefault('attrs', {}).update({'rows': 40, 'cols': 100})
 super(LargeTextareaWidget, self).__init__(*args, **kwargs)

class LargeTextarea(forms.Field):
 widget = LargeTextareaWidget
```

现在我们就可以修改之前的创建表单示例来使用这个自定义的变量了。

```
class ContentForm(forms.Form):
 name = forms.CharField()
 markup = forms.ChoiceField(choices=[
 ('markdown', 'Markdown'),
 ('textile', 'Textile')
])
 text = LargeTextarea()
```

不要忘了，Django就是纯Python。这代表在有需要的时候，你很容易就能像这样把各种类

和对象替换出去。记住这一点有助于你发现其他可以利用自定义的地方。

在这一章里，你学习了Django的模板语法以及如何用context字典来渲染模板，还包括了更复杂的主题诸如模板继承和包含等。此外，你知道了怎样生成表单（单独的或是表示某个特定模型类的）以及验证它们的数据和显示为HTML。最后，我们展示了一些通过widget自定义表单的方法。

这一章是第二部分的结尾，现在你对Django提供的功能应该已经有了一个相当完整的了解。

从模型定义，URL和请求处理，到现在的模板和表单。在下面第三部分的4章里，将要展示使用了这些你已经看到的知识所构建起来的应用示例，并在其中穿插一些新的或是扩展的概念。

### 6.3 总结



很多由内容驱动的网站都有一个常见的特点就是它们不仅允许用户添加文本，同时还可以添加文件——Office文档、视频、PDF文件，当然还有到处都是的图片。在这个应用示例里，我们要向你展示如何在一个简单的Gallery类型的應用程序里使用Django图片上载类ImageField。此外，我们还要创建一个自定义的ImageField子类来自动生成图片预览。最后，我们还为之设计一个动态的根据URL从服务器上读取并显示出来。这个应用已经过简化了，每一个都可以拥有多个Photo，这个层次表示为一个Django项目gallery及其内部的一个items项目（取不出更好的名字了）。我们可以构建这个应用程序来构建一个更常见的图片站点，Item可以变得更像是一个容器或者文件夹，专门用来组织照片。或者你可以把它做成类似于陈列馆的一个应用，每个Item可以拥有额外的属性（比如汽车型号、厂商和年份）来适应这种复杂度。对于我们这个例子，你就可以把Item想象成一本独立的相册。

我们不打算做任何不必要的工作，所以应用程序会尽量使用通用视图，并且所有数据输入都是经由Django的内建admin来完成。这样，它的架构就相当紧凑了：

- 首页展示了每一个“陈列柜”（一组图片预览）。
- 列表页显示所有站点上的Item。
- 每个Item的细节视图都列出所有属于它的Photo（也是预览）。
- 每个Photo的细节视图会显示图片。
- 首页上静态的模板欢迎信息。

我们从定义模型开始，然后一步步地让文件从admin应用里上传。接着详细讨论创建自定义模型类。最后，我们看如何将DRY原则应用到URL上以及创建用来向全世界展示预览和图片的前台模板。

## 第7章 Photo Gallery

第7章 Photo Gallery  
第8章 内容管理系统  
第9章 Liveblog  
第10章 Pastebin

# 第三部分 Django应用实例

**注意**

这个例子假设你用的是Apache+mod\_python，当然你可以修改它以适应其他部署策略。因为一个Gallery需要服务大量的静态内容（比如图片），Django的开发服务器不太适合它。你可以在附录B里找到更多关于Apache配置的信息。

## 7.1 模型

下面是本应用的models.py，除了稍后要做出的一个小改动外，这就是一个完整的文件了。注意get\_absolute\_url方法使用了@permalink装饰器，在本章结束之前会介绍到它。

```
class Item(models.Model):
 name = models.CharField(max_length=250)
 description = models.TextField()

 class Meta:
 ordering = ['name']

 def __unicode__(self):
 return self.name

 @permalink
 def get_absolute_url(self):
 return ('item_detail', None, {'object_id': self.id})

class Photo(models.Model):
 item = models.ForeignKey(Item)
 title = models.CharField(max_length=100)
 image = models.ImageField(upload_to='photos')
 caption = models.CharField(max_length=250, blank=True)

 class Meta:
 ordering = ['title']

 def __unicode__(self):
 return self.title

 @permalink
 def get_absolute_url(self):
 return ('photo_detail', None, {'object_id': self.id})

class PhotoInline(admin.StackedInline):
 model = Photo

class ItemAdmin(admin.ModelAdmin):
 inlines = [PhotoInline]

admin.site.register(Item, ItemAdmin)
admin.site.register(Photo)
```

(就像上面的例子), Web服务器和我们的用户就能访问它们。

最后, 遵章更完臻的应用程序会需要在media目录建立另外一些符号链接 (毕竟你还需要 CSS和JavaScript), 你的Web服务器需要能正确设置来适应它们。例如, 假设你用的是 Apache加mod\_python, 那么你需要在Django把持的URL空间上留出一些空隙, 以便Apache能直接处理你的媒体文件。更多mod\_python配置的信息请参见附录B。

(就像上面的例子), Web服务器和我们的用户就能访问它们。

上面这种情况下只有在你的普通用户也是属于www-data用户组的时候才适用——根据系统设置的不同，你或许会需要使用sudo等类似方法才能设置完成。我们发现在Web服务器域上执行很多系统任务的时候，把自己设置成它用户组的一部分会很方便。只要目录或文件有小组写权限

令行的使用細節請參見附錄A)：

那么每张照片会被保存在 `/var/www/gallery/media/photos/` 里。如果这个目录不存在的话，就要先创建它，另外 web 服务器运行的用户或者用户名需要有对这个目录的写权限。在我们的 Debian 系统上，Apache 是以 `www-data` 用户身份运行的，所以我们要像下面这样设置一番（命

```
MEDIA_ROOT = '/var/www/galaxy/media/'
```

settings.py里是这么设置的吧：

在向以上传文件到图片网站之前，我们需要告诉 Django 哪些文件是图片。fileField 和 ImageField 会将上传的数据保存在 settings.py 里定义的 MEDIA\_ROOT 中的一个子目录下，它会在 upload\_to 变量指定的目录里定义。在我们的模型代码里，我们已经把它设置成了 photo，所以若是

七、准备工作

↓Meta.ordering屬性集。

这里，Item非非富简单，而Photo才是真正的主角——它不仅和分类Item有关系，而且还有标题、图片文件本身和一个可选的说明。这两个对象都向admin应用注册了自己，它们也都有

可以继续来简单测试一下imageField。

PIL安装完成并且重启（或者重新载入）Web服务器后，PIL的错误就应该消失了，我们就能

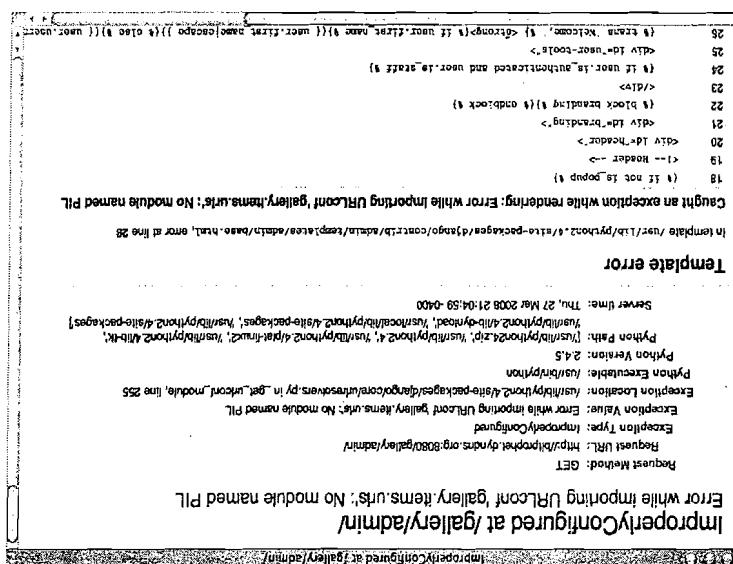
## 7.4 测量imageField

<http://peak.telecommunity.com/DevCenter/EasyInstall>。

不管在什么平台上，安装第三方Python软件包最简单的方法就是Easy Install。它知道怎么处理依赖关系，并且把下载和安装合并到一步完成。你只要运行easy\_install pip就可以完成刚刚描述的所有步骤。更多关于获取和使用Easy Install工具的信息，请访问PEAK开发者中心。

要在基于UNIX的系统上（比如Linux或Mac）安装PIL，从<http://www.pythontutorial.net>上下载源码包，解压后在install目录执行setup.py install即可。在Windows上，则可以从<http://www.pythontutorial.net>上下载安装包，解压后在install目录执行setup.py install即可。

图 7.1 没有安装 PIL 就会看到这个



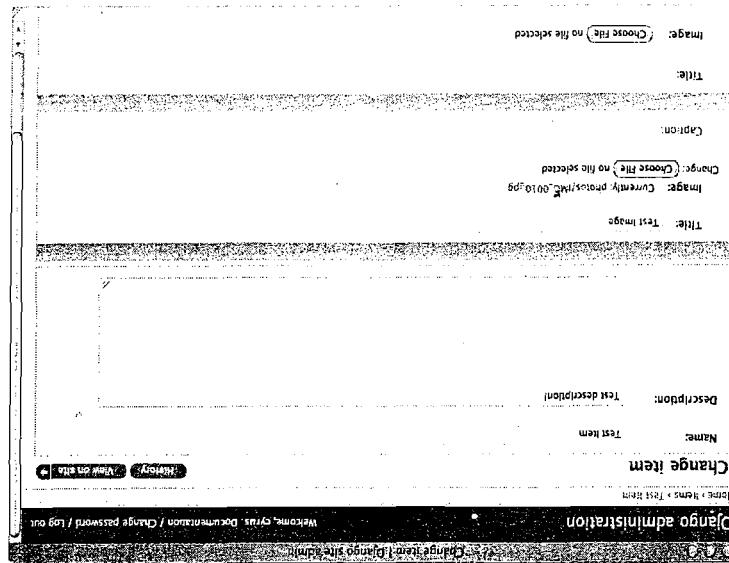
到此（当我们把目光定义为应用加到settings.py并且运行syncdb后），我们几乎就准备好了上传图片了。不过就像马上要看到的，我们还差一口气。要是现在我们就加载admin站点（安装好Django项目并且设置好照片上传文件夹），基本上我们会看到图7.1这样的画面。

卷之三

图7.4所示。

注意第一个Photo里文件选择器上面的Currently:link——点击它就会显示上传的图片，如

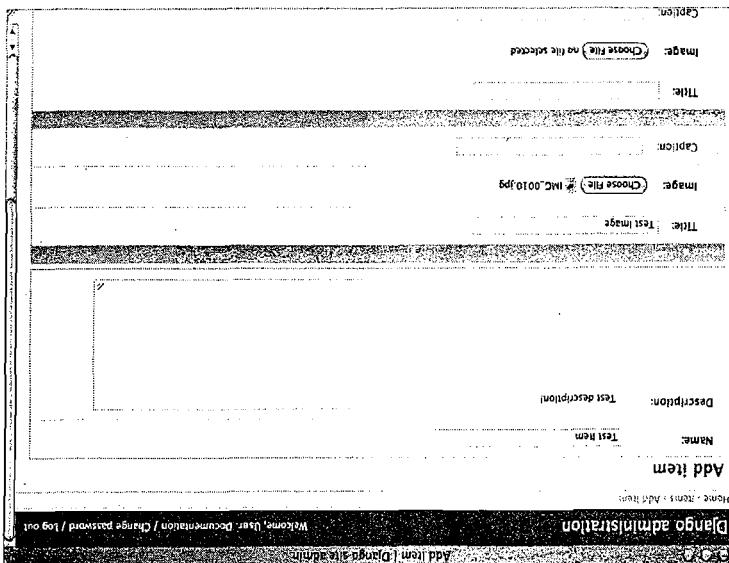
图7.3 文件上传后的admin界面



之上会保存。我们可以通过admin来验证这一点，如图7.3所示。

保存新的Item之后，所选择的图片（在这里是一幅我们一位作者的宠物兔的照片）也会随

图7.2 ImageField



素Item一起编辑。这样我们就可以轻松地在笔item的同时添加图片了，如图7.2所示。

模型的定义后我们发现admin及其实现类都设置好了，所以我们的Photo对象可以随着它们的父

### 不要害怕源码

程序员经常会把他们使用的类库（哪怕是开源的）当成是定义了输入输出行为的黑盒来使用，把它们看的很神秘。虽然有时候这么做是正确的，比如在冗长的低级语言里，这些可能真的是又巨大又笨拙，但是对Python源码来说这种情况却非常见。

写得好的Pythonic库通常都是很容易理解修改的，Django也不例外。我们不会假装整个代码库已经完美重构和注释了，但是大部分都是相当健壮的，程序员（哪怕是初学者）

这更简单也更常见。

由于Django并没有提供生成预览的功能，所以我们准备通过继承ImageField来自己的模型变量，让它能无缝地处理图片预览的生成、删除和显示。Django的官方文档为如何重写ImageField提供了出色的信息——不过在这里，我们只要调整一下现有的行为就可以了，

## 7.5 构建自定义File变量

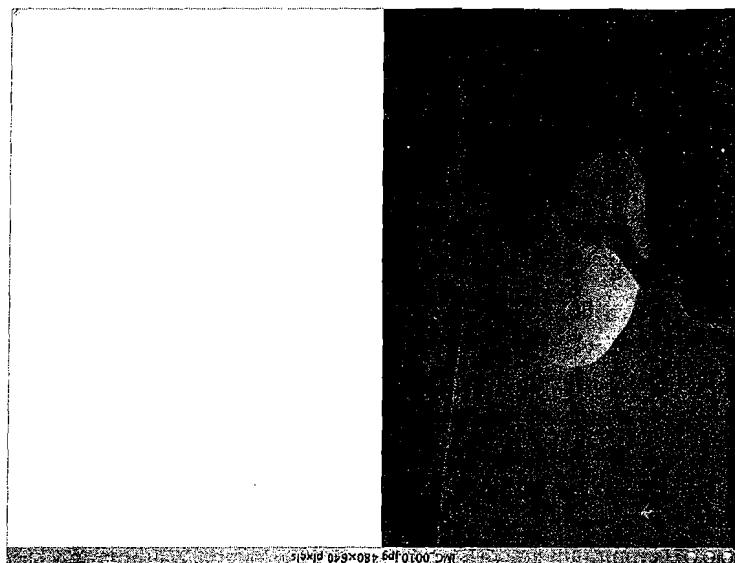
ImageField来生成图片预览。

虽然我们花费了一些篇幅来解釋整个过程，不过你也可以看到，其实这个设置和运行图片上传的逻辑非常直观——到现在为止，我们真正做的也就是定义一个模型，创建一个保存图片的文件夹，还有安装图片类库。现在我们终于要向你展示一点有趣的东西了，即如何扩展的过

```
-rw-r--r-- 1 www-data www-data 140910 2008-03-27 21:26 IMG_0010.jpg
total 144
user@example:/var/www/gallery/media/photos $ ls -l
```

我们也可以在命令行上验证上传的文件。

图7.4 检查当前ImageField的值



要完成这个功能，我们要重写ImageField父类里的4个方法，并增加一个简单的私有方法来帮助重构。这一章里的代码已经有完整的注释（文档在理解新领域的時候是非常有帮助的），但是这里为了阅读方便，我们把它们都删掉了。

我们把原先是ThumbnailImageField（很自然地）起名为ThumbnailImageField，将它保存在Django内部的子类。如果你对Python面对对象的数据方式还不熟悉的話，请參阅第1章，“实現Django Python”。

我们自己上而下地來解釋一下这个文件。

## 初始化

```
from django.db.models.fields.files import ImageField, ImageFile
from PIL import Image
import os
```

每个Python文件（几乎沒有例外的）都由import开始，这个文件也是一样。

ImageField还是很简单的一，这里我们需要的就是父类ImageField和ImageFile，PIL里用來生成的Image类，还有负责处理预览文件的内置os模块。

(续) 也可以通过深入代码了解Django工作机理来获取很多益处。这一章里我们的工作就设置完完整的文档——但是通过阅读django.db.models.ImageField及其父类的源码还是能比较容易地搞清楚它们的。

第7章 Photo Gallery [3]

```

Accepts two additional, optional arguments: thumb_width and thumb_height,
both defaulting to 128 (pixels). Resizing will preserve aspect ratio while
staying inside the requested dimensions; see PIL's Image.thumbnail()
method documentation for details.

"""

attr_class = ThumbnailImageFieldFile

```

这个类就很简单了——我们从ImageField继承了一个新的子类，然后写下一大串docstring。这样，任何人在使用Python的帮助系统，或是自动化文档工具时，都能很容易理解它的作用。

所以实际上这个类只有一行代码，attr\_class是用来更新一个特殊的类，我们的变量用这个类来访问属性。在下一节里我们会详细讨论它。初始化介绍的最后一部分就是\_\_init\_\_：

```

def __init__(self, thumb_width=128, thumb_height=128, *args, **kwargs):
 self.thumb_width = thumb_width
 self.thumb_height = thumb_height
 super(ThumbnailImageField, self).__init__(*args, **kwargs)

```

我们重写的\_\_init\_\_相当简单，只需要把预览文件在待会改变大小时的最大长和宽保存下来就行了。在需要不同预览大小的时候，重用它就很容易了。

### 给变量添加属性

大多数变量都是相对自我约束的，不会去随便改动包含它们的模型对象，但是在我们这个例子里，我们希望能够方便地得到我们提供的额外信息（预览文件的文件名和URL）。方法是去继承ImageField用来管理属性的一个特殊的类，ImageFieldFile。ImageField会用它去查找自身的变量。

例如，用myobject.image.path就可以得到一个叫image的ImageField对象在文件系统上的路径。这里的.path就是ImageFieldFile的一个属性。由于Django会尽量缓存文件数据而把具体文件代理到更低的层次上去，所以这是通过Python的properties完成的。（你可以参阅第1章来复习一下内置函数property的内容。）

下面的代码展示了Django是怎么实现ImageFieldFile.path的：

```

ImageFieldFile.path:

def __get_path(self):
 self._require_file()
 return self.storage.path(self.name)
path = property(__get_path)

```

这段代码取自FieldFile类（它是ImageFieldFile的父类）。回想一下前面的帮助函数\_add\_thumb是怎么转换一个给定文件路径的，你就能猜到我们要怎么把.thumb\_path和.thumb\_url属性加到变量里去了：

```

class ThumbnailImageFieldFile(ImageFieldFile):
 def __get_thumb_path(self):
 return _add_thumb(self.path)
 thumb_path = property(__get_thumb_path)

```

```

 def _get_thumb_url(self):
 return _add_thumb(self.url)
 thumb_url = property(_get_thumb_url)

```

因为`.path`和`.url`的getter函数已经定义好了，而且它们会去安全地处理那些重复性的操作（上面`_get_path`代码里的`self._require_file`调用），所以我们可以省略这些代码。我们只需要`_add_thumb`转换并且把结果用`property`函数附加到需要的属性名上即可。

当`ThumbnailImageFieldFile`在`ThumbnailImageField`上定义以及在`ThumbnailImageField`里由`class`引用之后，我们给变量添加了两个新的属性，你可以在Python代码或是模板里使用它：`myobject.image.thumb_path`和`myobject.image.thumb_url`（这里`myobject`是Django的模型实例，`image`是那个模型上的`ThumbnailImageField`对象）。

继承`ImageFieldFile`然后将子类和`ImageField`的子类连结起来可能不是很常见的手法，绝大部分自定义的模型变量甚至根本用不着深入到这个地步。事实上，作为Django的用户，你可能见不到模型里的这一面。然而它展示了Django核心团队不仅仅在公共API上尽力保持扩展性，而且在框架内部也是如此。

现在我们已经可以访问所需的预览文件的URL和文件系统路径了，接下来就是实际创建（或删除）这个预览文件。

### 保存和删除预览文件

创建预览文件的关键之处，就是`ThumbnailImageFieldFile`（不是`ThumbnailImageField`！）里的`save`方法：

```

 def save(self, name, content, save=True):
 super(ThumbnailImageFieldFile, self).save(name, content, save)
 img = Image.open(self.path)
 img.thumbnail(
 (self.field.thumb_width, self.field.thumb_height),
 Image.ANTIALIAS
)
 img.save(self.thumb_path, 'JPEG')

```

调用父类的`save`方法会帮你正常地处理主要图片文件的保存操作，所以我们的方法只需要做三件事情，即打开原图片文件，创建预览，将预览保存为预览文件名。注意这里的`self.field`让我们能访问属于`File`对象的`Field`，即保存了预览图片大小的地方。利用一开始导入的PIL`Image`类，这里的代码其实非常简单。

最后，当删除原图片时，它们的预览文件也要随之一起删除：

```

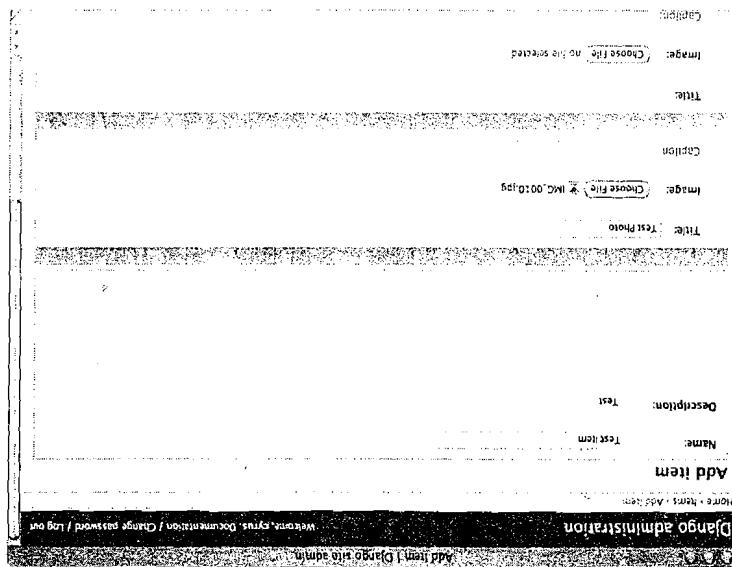
 def delete(self, save=True):
 if os.path.exists(self.thumb_path):
 os.remove(self.thumb_path)
 super(ThumbnailImageFieldFile, self).delete(save)

```

多亏了Python漂亮的语法，这段代码几乎就是在描述自己了。首先获取预览文件名，删除

即使是在文件上传后，看起來也和之前的例子一模一樣，如圖7.6所示。

圖7.5 和前面沒什麼兩樣



單有关的变量，如图7.5所示。

重新加载Web服务器后admin里还没有看到什么显著的变化，因为我们还没有修改任何与表

```
class Photo(models.Model):
 title = models.CharField(max_length=100)
 image = models.ImageField(upload_to='photos')
 caption = models.TextField(max_length=250, blank=True)
 description = models.TextField(max_length=250, blank=True)

 def __str__(self):
 return self.title

 class Meta:
 ordering = ('-title',)
```

定义完ImageField子类后，我们来看看怎么用它。首先在models.py里加入一个新的import：

## 7.6 使用ThumbnailImageField

操作的顺序

在器模型对象被删除时，就会触发ImageField去调用它。  
如果文件不存在，那上面的代码还不能清楚，因为它会先去调用self.path（回忆一下之前的代码）来保证变量的主要文件确实存在！所以我们要等到最后一刻才能删除主文件，以免我们

（如果文件不存在的话，即使不存在，那也算不上什么错误），然后告诉父类去删除它自己的文件（即原图）。如果上面的代码还不够清楚，这里的意思是delete方法会和save一样，当它的

——，以及展示这种方式如何让你可以通过多种方法改变这种行为。  
Django从1.0起引入了一个新的Apache配置说明，PythonOption django.root <root>，它基本取代了我们在这一里介绍的ROOT\_URL功能。不过，我们还是把这部分内容保留下来，作为Django“既是Python”的例子之一

个URL路径，所以一个在 /gallery/ 上的站点里，它需要在所有URL上加上这个字符串前缀。默认情况下，一个Django站点都假设是前面的那种情况——即URL是以域名的根目录开始解析，哪怕Web服务器处理器被挂接在更高的层次上面。因此，一般站点的URL都必须包含全解析路径，所以一个在 /gallery/ 上的站点里，它需要在所有URL上加上这个字符串前缀。都可以正常工作。

到这里为止，我们关注的只是Gallery应用里模型 (model) 的部分。现在我们来看看URL到这个不太寻常的设置方式。由于这个应用开发的方式，我们需要一点背景知识才好理解这个不寻常的设置方式。不过首先我们需要一点背景知识才好理解

它们的模板。在这之前，我们先赶快看一下应用程序里一个比较次要的方面：怎么设置一个大功告成！不过还差一点才能看到这个面貌，因为我们还没有向你展示应用程序里负责显

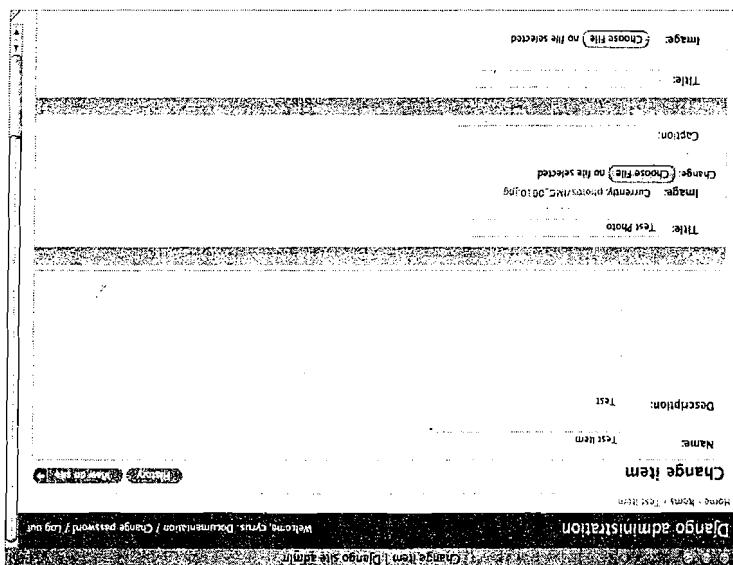
## 7.7 设置DRY URL

简单的方式才好保持URL的DRY呢？

```
-rw-r--r-- 1 www-data www-data 1823 2008-03-30 22:15 IMG_0010.thumb.jpg
-rw-r--r-- 1 www-data www-data 140910 2008-03-30 22:15 IMG_0010.jpg
total 148
user@example:/var/www/gallery/media/photos $ ls -l
```

但是检查一下我们的上传目录，你就能发现我们的劳动果实啦。

图7.6 和之前的页面还是一样



完成了！这个Djangogeo项目经过一系列的调整之后，所有一切就都围绕着ROOT\_URL的值

```
def root_url_processor(request):
 return {'ROOT_URL': ROOT_URL}

from gallery.settings import ROOT_URL

请参阅前面第6章。

比如那些CSS或是JavaScript需要的includes。这只需要一个小小的context处理器就能做到（这个最后，要是模板也能访问ROOT_URL的话就能方便地构建出符合DRY原则的includes URL，


```

```
url(r'^$', include('gallery.urls')),
url(r'^admin/.*$', admin.site.root),
urlpatterns = patterns('',
 from django.conf.urls.defaults import *
 from djangogeo.gallery.urls import *
)
urlpatterns = patterns('',
```

这里是我们“真正的”的根URLconf，称之为real\_urls.py（想不出更好的名字了）：

**注意** Django的URL解析会忽略那个开头的斜杠，所以这里必须要去掉它才能让我们的URL正确LOGIN\_URL) 为了成为正确的绝对路径的形式会要求自带一个开头的斜杠。但是由于我们对ROOT\_URL却发现开头的斜杠，因为用到它的settings.py变量(比如

解卦。

而这个文件对ROOT\_URL及其相关的内容一无所知。这里是根urls.py：

```
url(r'^%s %s$' % (ROOT_URL[1], include('gallery.real_urls'))),
urlpatterns = patterns('',
 from django.conf.urls.defaults import *
 from djangogeo.gallery.urls import *
)
urlpatterns = patterns('',
```

接下來，又由于Django的URL包含系统的工作方式，我们需要用到一个双文件形式的根URLConf设置，这里“普通的”顶层urls.py只用ROOT\_URL，然后去调用“真正的”urls.py，

```
ADMIN_MEDIA_PREFIX = MEDIA_URL + 'admin/'
MEDIA_URL = ROOT_URL + 'media/'
```

```
LOGIN_URL = ROOT_URL + 'login/'
```

这里有登录的URL，媒体文件的URL，以及admin的媒体文件前缀。

因为还有其他一些依赖于URL路径的settings.py变量，所以我们这里就把它们放在一起了，

```
ROOT_URL = 'gallery/'
```

引用它。

所以我们这里就将规矩起来，把这个值保存在settings.py里的一个变量里，然后在必要的地方

名字，此外，包裹`get_absolute_url`的permalink装饰器也要引用它们，像这样：  
面，每一个都有直观的名字。在下一节的模板里，我们将通过`{% url %}`模板标签来引用这些  
如你所见，这个应用程序包含了一个首页，一个所有相册的列表，相册页面，以及相片页

```
name='photo_detail',
},
),
,
template_name: 'photo_detail.html',
queryset: Photo.objects.all(),
kwargs={

url(r'^photos/(?P<object_id>\d+)/$', 'list_detail.object_detail',
),
,
name='item_detail',
},
,
template_name: 'items-detail.html',
queryset: Item.objects.all(),
kwargs={

url(r'^items/(\?P<object_id>\d+)/$', 'list_detail.object_detail',
),
,
name='item_list',
},
,
allow_empty: True,
template_name: 'items-list.html',
queryset: Item.objects.all(),
kwargs={

url(r'^items/$', 'list_detail.object_list',
),
,
name='index',
},
,
extra_context: {'item_list': lambda: Item.objects.all()}
template: 'index.html',
kwargs={

url(r'^$', 'simple.direct-to-template',
),
urlpatterns = patterns('django.views.generic',
from gallery.items import Item,
from django.conf.urls import *
),
,
URlconf，这允许我们给URL取一个唯一的名字。下面是包含在items应用里的urls.py：
```

为了完成URL结构的DRY “属性”，我们要对对象里的`get_absolute_url`方法应用三个  
步骤。第一个也是最重要的部分就是（你在这本书里已经见过的）用`url`函数来定义我们的  
URLconf，这允许我们给URL取一个唯一的名字。下面是包含在`items`应用里的`urls.py`：

## 7.8 Item应用的URL布局

Web服务器的设置把Django挂接到根目录上去）就行了。  
你想把它部署到`http://www.example.com/`上去的话，只需要把ROOT\_URL改成“（并且更新  
了——现在的值是`/gallery/`，表示应用程序的位置是`http://www.example.com/gallery/`。要是

```

class Item(models.Model):
 name = models.CharField(max_length=250)
 description = models.TextField()

 class Meta:
 ordering = ['name']

 def __unicode__(self):
 return self.name

 @permalink
 def get_absolute_url(self):
 return ('item_detail', None, {'object_id': self.id})

```

permalink装饰器要求它包裹的函数返回一个三元组（URL名、一列位置参数和一个命名参数的字典）来重新组成URL。在上面的例子中，item\_detail视图不接受位置参数，但是接受了一个命名参数，而那个参数正是我们在get\_absolute\_url里提供的。

这种设置方式即便在URL结构发生变化的时候也可以让Item.get\_absolute\_url能返回适合的URL，从而达到保持DRY的目的（这里也不是没有代价，比如要是去掉装饰器的话，get\_absolute\_url的行为就会显得很怪异等）。

## 7.9 用模板把它们都串在一起

完成自定义模型变量以及调整好URL设置之后，最后剩下的（因为这里只用到了通用视图）就是模板了。我们采用的是一种简单的继承方式来最大化DRY，首先是基本的结构模板和一些CSS。

```

<html>
 <head>
 <title>Gallery - ${ block title }${ endblock }</title>
 <style type="text/css">
 body { margin: 30px; font-family: sans-serif; background: #fff; }
 h1 { background: #ccf; padding: 20px; }
 h2 { background: #ddf; padding: 10px 20px; }
 h3 { background: #eef; padding: 5px 20px; }
 table { width: 100%; }
 table th { text-align: left; }
 </style>
 </head>
 <body>
 <h1>Gallery</h1>
 ${ block content }${ endblock }
 </body>
</html>

```

接下来是首页。本章的这个例子应用了一点轻型的类似CMS的功能，在admin里显示了一堆“欢迎”用语，这里为简单起见就省略了。我们在这里用的是一段静态的欢迎标语和一列三

个由URLconf控制的特色Item。(它现在显示的是全部内容，不过很容易改成其他形式。)

```
(% extends "base.html" %)

(% block title %)Home(% endblock %)
(% block content %)

<h2>Welcome to the Gallery!</h3>
<p>Here you find pictures of various items. Below are some highlighted
items; use the link at the bottom to see the full listing.</p>

<h3>Showcase</h3>
<table>
 <tr>
 {% for item in item_list|slice:'3' %}
 <td>
 {{ item.name }}

 {% if item.photo_set.count %}

 {% else %}
 No photos (yet)
 {% endif %}

 </td>
 {% endfor %}
 </tr>
</table>
<p>View the full list »</p>

{% endblock %}
```

上面的模板代码渲染出来的页面视图如图7.7所示。

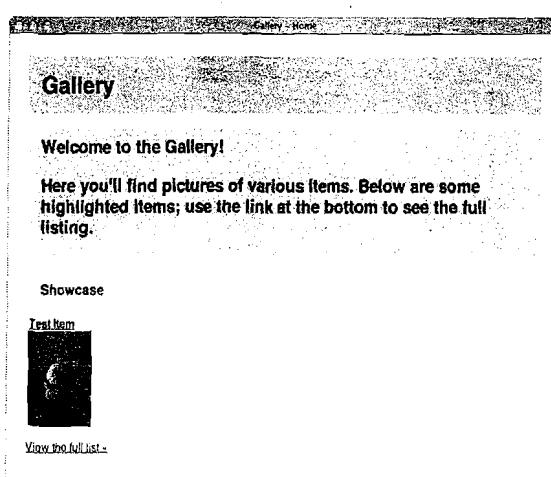


图7.7 Gallery的首页

注意这里分别通过get\_absolute\_url和{% url %}链接到图片细节页面和相册列表的用法，还

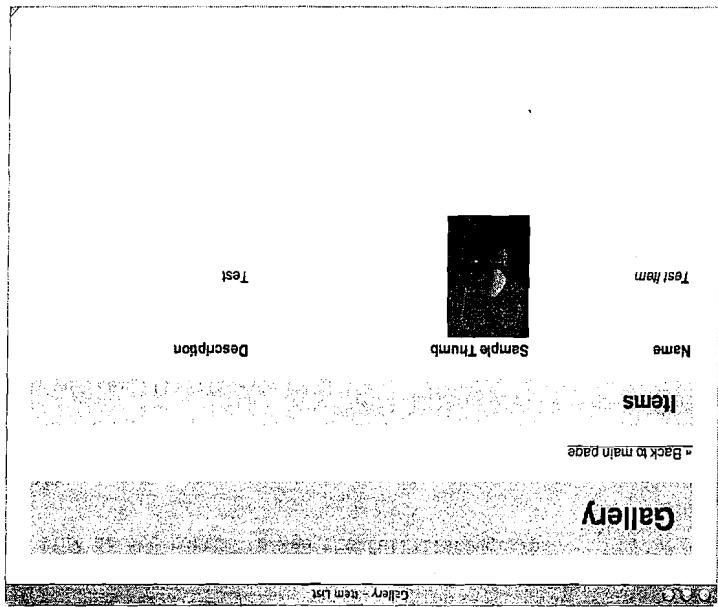
```


 (% if item.photo_set.count %)
 <td>
 <td><i>{{ item.name }}</i></td>
 <td>
 <% for item in object_list %>
 <tr>
 <th>Description</th>
 <th>Sample Thumbnail</th>
 <th>Name</th>
 <tr>
 <td>
 <% if object_list %>
 <h2>Items</h2>
 <p>Back to main page</p>
 (% block content %)
 <ul class="list-group">
 {{ block.title }} Item List <ul class="list-group">
 {{ item }}

 (% endblock content %)
 <% else %>
 <h2>No items found</h2>
 <% endif %>
 </td>
 </tr>
 <% endfor %>
 </td>
 <% endif %>


```

图7.8 Gallery列表页面



一样的，如图7.8所示。

相册列表（items-listing）是一个比首页上的特色列表更完整的版本，它们用的手法其实是过时的Photo模型来选用一张特定的图片作为“代表”来改进——这是应用程序很多可扩展的地方之一。

有每个相册列表上第一幅图片的image.thumb\_url。这个“用哪张预览作为封面”的问题可以通过更新Photo模型来选择一张特定的图片作为“代表”来改进——这是应用程

```


 {% else %}
 (No photos currently uploaded)
 {% endif %}
</td>
<td>{{ item.description }}</td>
</tr>
{% endfor %}
</table>
{% else %}
<p>There are currently no items to display.</p>
{% endif %}

{% endblock %}

```

相册的细节视图 (`items_detail.html`) 和相册列表视图差不多，就是它会列出所有照片而不是封面，如图7.9所示。

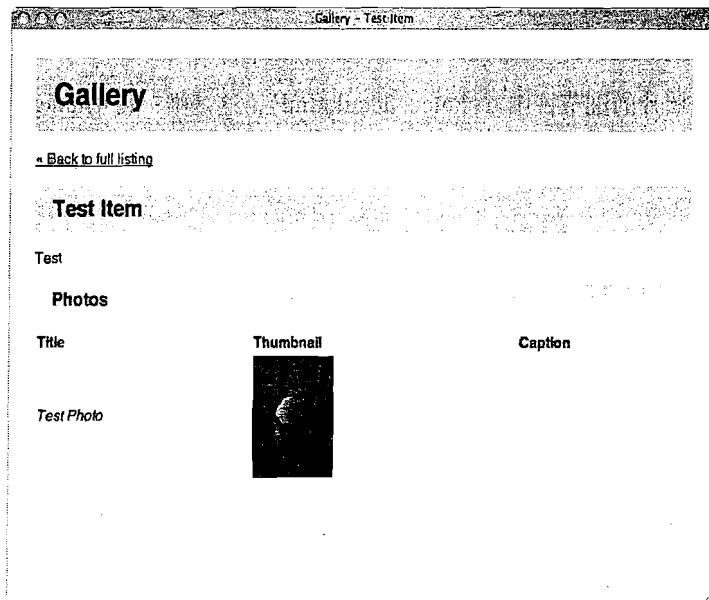


图7.9 相册细节页面

```

{% extends "base.html" %}

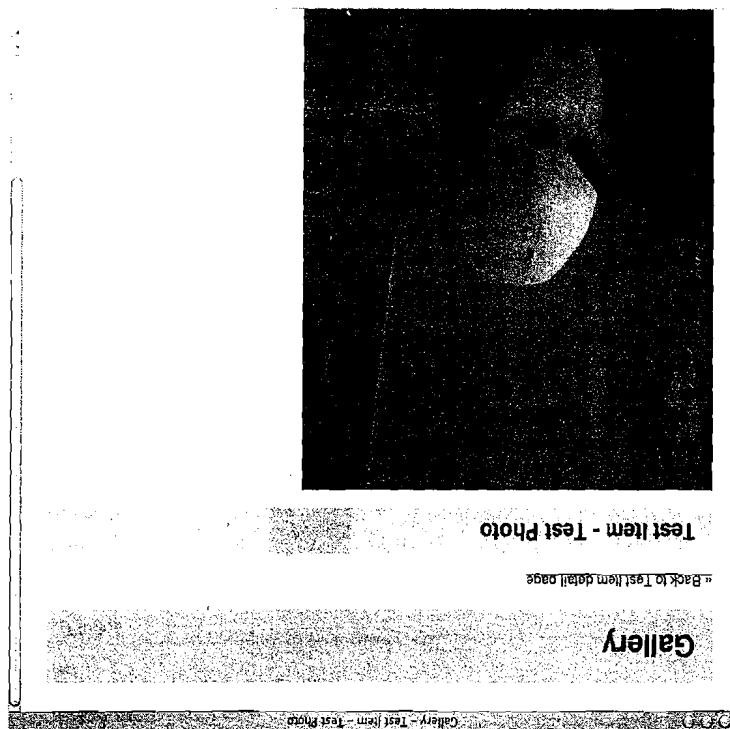
{% block title %}{{ object.name }}{% endblock %}

{% block content %}

<p>« Back to full listing</p>
<h2>{{ object.name }}</h2>
<p>{{ object.description }}</p>

```

图7.10 图片细节视图



7.10所示。

最后即是每张照片的细节视图 (photos\_detail.html)，这是唯一使用image.url的地方，如图

{% endblock %}

```
</table>
({% endfor %}
</tr>
<td>{{ photo.caption }}</td>
</td>

<td><i>{{ photo.title }}</i></td>
</tr>
({% for photo in object.photo_set.all %}

<th>Caption</th>
<th>Thumbnail</th>
<th>Title</th>
</tr>
<table>
<tr>
<td>Photos</td>
```

```
{% extends "base.html" %}

{% block title %}{{ object.item.name }} - {{ object.title }}{% endblock %}

{% block content %}

« Back to {{ object.item.name }} detail page

<h2>{{ object.item.name }} - {{ object.title }}</h2>

{% if object.caption %}<p>{{ object.caption }}</p>{% endif %}

{% endblock %}
```

## 7.10 总结

在风风火火地介绍完后，希望你对整个项目是怎么一同工作的有了一个比较完整的概念。

- 我们定义了模型，然后用admin来展示如何上传图片，包括必要的系统设置。
- 因为需要图片预览的功能，所以定义了一个Django图片变量的子类及其相关的文件类，这里我们重写了几个方法来改变图片的大小，以及提供对预览的访问。
- 我们完全利用了Django的DRY URL特性，包括实现一个“根URL”设置（就像那个在1.0之前刚刚加到Django核心里的一样）来帮助我们保持URL灵活。
- 最后，我们创建了一些简单的模板让用户可以浏览和查看图片。

- Django驱动的最简单的CMS连一行代码也不用写。Django自带的一个叫做“Flatpages”的应用程序最适用于简单的情况。Flatpages最大的优点就在于如果它适合你的话，几乎不用做什么设置，而且也没有代码需要维护。
- 另一个方便的地方就是指向Flatpages页面的URL都在admin里指定了，所以你不用专门去编辑URLconf文件来添加一个新页面。不过别高兴得太早，以下是一些限制：
  - 所有有权访问Flatpages应用的管理员用户都可以修改任何页面，用户不能“拥有”单独的页面。
  - 除了我们讨论的title、content属性和一些专用的变量之外，Flatpages对索的功能相当有限。
  - 由于它是以“contrib”应用的形式提供的，所以没办法轻易地修改它的admin选项，添加你没有办法给某个页面创建时间或是其他的一些数据。

## 8.2 Flatpages

随着开源Django应用社区日渐成熟，我们很高兴地看到一些CMS类型的应用已经变得相当完善，值得推荐给大家使用了，很多系统正在慢慢囊括它自己的用户社区和维护群。说不定能满足你需要的应用早已经出现了。所以先做一个简要的概述是好的（参见附录D），不过如果你需要手动实现的话也不要犹豫哦。

Django新手常问的一个问题是，“有没有一个用Django编写的开源CMS（Content Management System，内容管理系统）？”我们的答案可能不是很让人期待的：你自己写。这一章探讨了几种利用Django编写的方法，首先是利用一个contrib的应用来方便我们创建和发布“普通的”HTML页面，然后再进一步（但还是简单地）完成一个简单的内容管理系统。

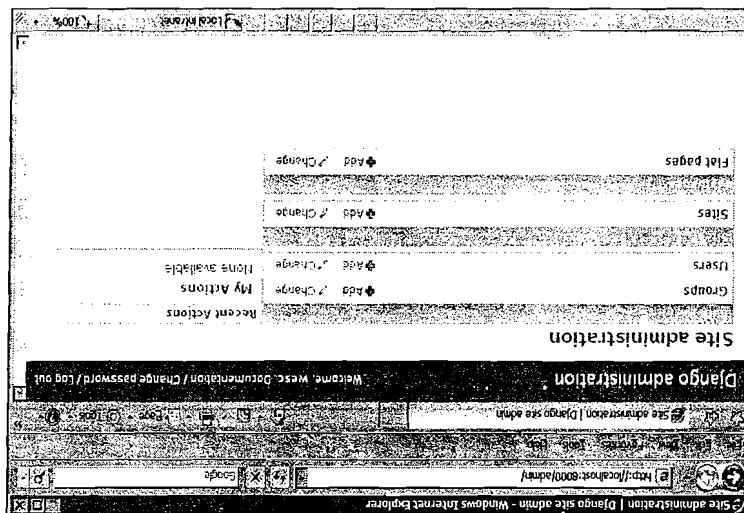
## 第8章 内容管理体系

去正确的错误处理。

下面我们可以直接去看模板的介绍，因为这里除了刚刚输入的URL，你不需要再对它做任何的管理。而且，它会利用一块特殊的Django中间件去拦截404错误，并在Flatpage对象列表里查找出正确的URL。如果找不到的话，Flatpages应用就会出来接手，要是没找到，那么404就会转到的管理。现在创建一个到两个页面，这样在测试的时候你就有内容可看了。

点击Add来创建一个新的Flatpage对象（如图8.2所示）。你只需输入url、title和content。确保这里的URL用斜杠开头。

图8.1 激活Flatpages应用后登录admin页面



完成上述步骤后，登录admin站点，你应该能看到图8.1这样的页面。

6.（重新）启动你使用的Web服务器。

5.更新urls.py，激活默认的admin。

4.运行manage.py syncdb让Django创建必要的数据表。

APPS变量里。

3.把django.contrib.flatpages和django.contrib.admin加入到settings.py里的INSTALLED\_APPS变量里。

2.打开项目的settings.py并更新MIDDLEWARE\_CLASSES设置，添加django.contrib.flatpages.middleware。

1.用django-admin.py创建一个新的Django项目。

以下是如何让Flatpages能跑起来的几个必要步骤：

## 激活Flatpages应用

何设置和使用Flatpages，完成一个更强大的自定义CMS应用。

不过若这些对你都不是问题的话，Flatpages还是相当有用的。我们在下面几节里会探讨如何

新变量或是模型方法。

```

</html>
</body>
<p>{{ flatpage.content }}</p>
<h1>{{ flatpage.title }}</h1>
</body>
</head>
<title>My Dummy Site: {{ flatpage.title }}</title>
</head>
<html>

```

现在在你刚刚创建的目录创建一个最简单的页面模板并保存为default.html:

```

<h1>{{ flatpage.title }}</h1>
<p>{{ flatpage.content }}</p>

```

和你希望的一样，你的模板是由一个叫做flatpage的对象传递过来的。例如：

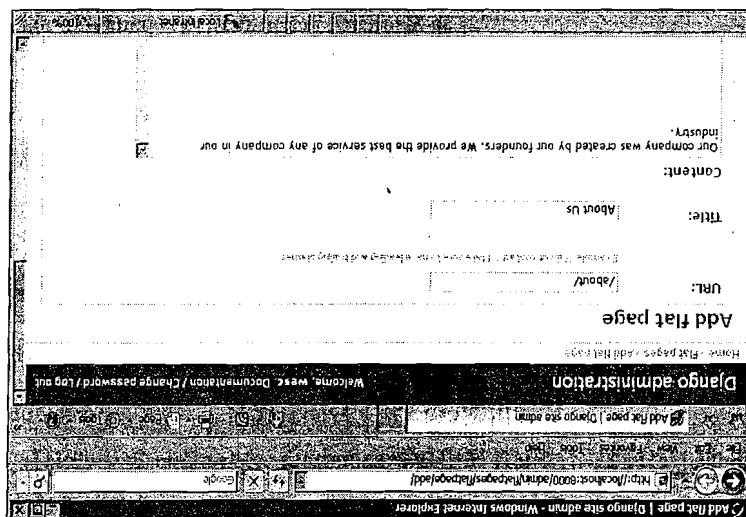
使用app\_directories模板加载器的话，就在INSTALLED\_APPS设置下列出的某个应用库里使用flatpages应用会在所有的模板里查找到一个叫做flatpages/default.html的模板。这就是说你要在项目的“templates”文件夹中建立它。不管是哪种方式，现在先把它建好。

## Flatpage模板

Flatpages的这种404处理方式意味着你可以把它和正常的Django应用一起使用，就是说你可以简单轻松地指定你的flatpages（“关于”或者“法律”等页面）而不用不断地更新URLconf。

### 注意

图8.2 Flatpage的Add界面



先从模型（model）开始。一个应用cms。

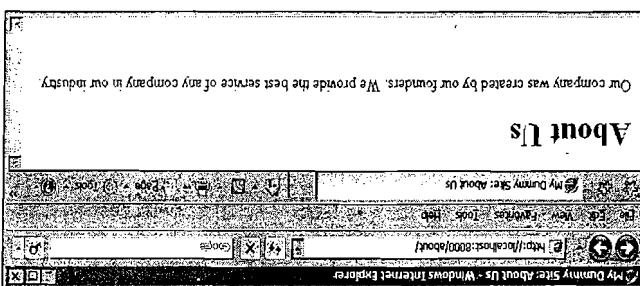
首先，我们要再设置一个Django项目（如果不记得怎么创建项目，数据库等内容的话，请复习一下第2章，“Django集成：构建一个Blog”）。我们给项目起名为cmsproject，它只包含了构建Django应用在一定程度上也就是学习如何尽可能有效地把这些特性组合起来以达成所要的结果。

- 提供基本的工作流——能将一个员工用户和文章联系起来，同时也允许把文章归属为某个产品上的某个部分。
- 维护创建和修改的时间。
- 为文章提供分类功能，并且能按照这些类别来浏览文章。
- 为所有页面提供一个简单的搜索功能。

### 8.3 超越Flatpages：一个简单的自定义CMS

到此你已经体验过了Flatpages的用法以及它适用的方位。下面开始是本章的重点线——一个更强大的自定义CMS应用示例。

图8.3 一个“关于”Flatpage的例子



所示。

且你通过admin创建的Flatpage对你的URL值为/about/的话，你可以在浏览器里输入`http://localhost:8000/about/`。它应该会用default.html模板显示title和content变量的值，如图8.3所示了，现在试着去嵌入你的flatpage页面。例如，如果服务器是运行在你自己的机器上，并且你通过admin创建的Flatpage对你的URL值为/about/的话，你可以在浏览器里输入

测试一下

## 创建模型

下面是我们这个小型CMS的核心模型定义。注意这里引用了两个其他的模型（User和Category），稍后我们就会看到它们的定义以及添加一些必要的import语句。

```
class Story(models.Model):
 """A hunk of content for our site, generally corresponding to a page"""

 STATUS_CHOICES = (
 (1, "Needs Edit"),
 (2, "Needs Approval"),
 (3, "Published"),
 (4, "Archived"),
)

 title = models.CharField(max_length=100)
 slug = models.SlugField()
 category = models.ForeignKey(Category)
 markdown_content = models.TextField()
 html_content = models.TextField(editable=False)
 owner = models.ForeignKey(User)
 status = models.IntegerField(choices=STATUS_CHOICES, default=1)
 created = models.DateTimeField(default=datetime.datetime.now)
 modified = models.DateTimeField(default=datetime.datetime.now)

 class Meta:
 ordering = ['modified']
 verbose_name_plural = "stories"

 @permalink
 def get_absolute_url(self):
 return ("cms-story", (), {'slug': self.slug})

class StoryAdmin(admin.ModelAdmin):
 list_display = ('title', 'owner', 'status', 'created', 'modified')
 search_fields = ('title', 'content')
 list_filter = ('status', 'owner', 'created', 'modified')
 prepopulated_fields = {'slug': ('title',)}

admin.site.register(Story, StoryAdmin)
```

模型类定义里的第一块代码定义了一个包含四个阶段的工作流。当然，你可以给自己的流程添加更多的步骤。

虽然使用Django映射来完成变量选择（就像STATUS\_CHOICES展示的一样）有很多方便的地方，但是在这里它在数据库里最后还是归结为整数。由于稍后要再重新定义“1”代表什么意思就不是那么容易了，所以最好先停下来想一想，确认你的列表是不是合理。当你要根据这个变量对模型实例进行排序的时候就更是如此了，而我们这里正有理由要这么做。

我们还要在公共视图里用这些值来决定访客能看到什么，即，他们可以看到“Published”

和Flapgates应用一样，当你开始构建高级的Django应用时就会发现User模型缺少了很多功能。例如，它组织用户名的方式可能和你的要求有冲突。但是，User特别地好用，算是一个中等完备的解决方案，而且在很多现实世界的应用里也是非常有用的。

除了django.db.models（以及稍后要解释的一个关联的permissions类），我们还需要导入人的有datetimes模块（用于created和modified变量）和来自Django的contrib.auth应用的User模型。此外还有用来自我们的模型向admin应用注册的Django admin模块。

samples

小已經見過它了。

我们对这个模型所做的一个修改（纯粹是为了 admin 的用户），就是在 Metaclass 里指定了一个 verbose\_name\_plural 属性。这能让我们模型不会在 admin 应用里显示 “Stories” 这样错误的一个名字。最后，还有一个生成永久链接的 get\_absolute\_url 方法，在第 7 章 “Photo Gallery” 里我

修改时会自动更新。这个时间戳会显示在文章详细页面上。

• modified: 修改化为当简时间。我们需要一些特殊的数据才能保证在文章被

- created：创建的时间，自动设置为当前时间（利用Python的`datetime`模块）。

- status: 在编辑工作流中的状态。

象的外鍵引用)。

• owner：拥有这个内容的admin用户（或者，从Django的角度来说，这是一个指向user对

不会在 Django 的 admin 应用里的编辑表单上显示出来)。

面的时候就不会有标记翻译的开支了。为了避免混淆，不可以直接编辑这个变量（所以

• `htmlContent`: HTML格式的页面文本。我们在编辑的时候自动渲染它，所以可以在显示页面

- markdown-content: Markdown格式的页面正文（下面会专门介绍Markdown）。

• category: 文章的类别。这是一个指向稍后定义的另一个模型的外键。

• **slug**: 页面在其URL里唯一的名字。这比一个无意义的整数主题要好多了。

- title：我们要在浏览器的标签栏和渲染页面上显示的标题。

定義完STATUS\_CHOICES之後就能夠到變量的定義。

人和其他数据一样在admin里编辑了。

如果不想把这样一个列表硬编码在

目，和/或者应用程序需要所决定的逻辑。

和“Archived”的文章，但是看不到“New

## 完成模型

User对象是直接从Django贡献的“auth”应用里拿来的，那么Category呢？它倒是我们自己的，以下是其模型定义，它应该出现在models.py里Story模型定义的上面。

```
class Category(models.Model):
 """A content category"""
 label = models.CharField(blank=True, max_length=50)
 slug = models.SlugField()

 class Meta:
 verbose_name_plural = "categories"

 def __unicode__(self):
 return self.label

class CategoryAdmin(admin.ModelAdmin):
 prepopulated_fields = {'slug': ('label',)}

admin.site.register(Category, CategoryAdmin)
```

Category模型相当简单，甚至有点琐碎。在Django的应用程序里，你经常会看到这类简单的模型（有时候简单到只定义了一个变量）。不过要是把它变成Story模型里的一个“category”变量的话，那会让事情变得困难起来（给category改名），甚至失去某些功能（比如，给category添加描述属性等）。Django既然提供了方便的途径来创建适合的关系型模型，那么最好还是照做吧。

对于Story，我们还设置了一个verbose\_name\_plural属性，这样admin的用户就不会笑我们连拼写都不会了。

## 控制文章的显示

我们的数据库同时包含了已发布的（之前STATUS\_CHOICES里的3和4）和还未发布的（状态1和2）文章。我们需要一种便捷的方式只把前者显示在站点的公共页面上，而在admin里则显示全部。因为这里涉及的是商业逻辑而非表现风格，所以应该在我们的模型里实现它。

虽然我们也可以通过模板里的{% if ... %}标签来达成，但是这个方案最终会变得异常繁琐和不断重复。（如果你不信的话，我们建议你不妨试试看——说不定在你完成前就受不了这种缺陷了！）根据本书作者的共同经验，永远不要把商业逻辑放到模板里面去，否则时间一长必定搅得一团乱。

我们把这项功能通过自定义Manager加到Story模型里去。关于该技巧更多的细节，请参见第4章，“**定义和使用模型**”里的“**自定义Manager**”一节。在models.py文件的import语句下面加上如下代码：

```
VIEWABLE_STATUS = [3, 4]
```

如果要在Python里使用Markdown，你得先去下载Python-Markdown模块，因为它不是标准库。

## 为什么不用WYSIWYG?

你还可以选择用Textile、RestructuredText等其他轻量型的标记语言。这里我们主要展示的是如何通过重写模型的save方法来“神奇自动”地把Markdown转换成HTML，这样就不需要为每个页面请求执行一次翻译了——我们在前面描述markdown\_content和html\_content变量的时候已经提到了这一点了。

既然基于Web的内容管理系统一般都是针对非技术背景用户的，有心思或许会犯嘀咕，我们这里用Markdown是不是有点 nerd了。确实，若是在Django admin里集成WYSIWYG（所见即所得）的HTML编辑器的话，用户可能会更容易上手。不过除了要多花一点精力实现以外，这个方法的另一个缺点在于WYSIWYG还是不能把Web浏览器变成Microsoft Word，而且很容易出现浏览器不兼容的问题。不过说归说，这类型工具还是能够增加像CMS这样的工具的吸引力和对非技术用户的黏度。要得到WYSIWYG插件的推广和其他的建议，请访问 [wityhdjango.com](http://wityhdjango.com)。

作为模型的最后一部分，我们重写了内置函数save来应用一个轻量型的标记语言，叫做 Markdown，用户在admin里用它来输入文本。Markdown和Wiki风格的语言颇为相似，它提供了创建Web内容更简单的方式。编写Markdown比原生HTML要麻烦得多，但编写这些文件email或是编辑过Wiki页面的人却不会对它感到陌生。

你还可以选择用Textile、RestructuredText等其他轻量型的标记语言。这里我们主要展示的是如何通过重写模型的save方法来“神奇自动”地把Markdown转换成HTML，这样就不需要为每个页面请求执行一次翻译了——我们在前面描述markdown\_content和html\_content变量的时候已经提到了这一点了。

然后我们用常见的objects为名创建一个自定义manager的实例。因为我们在URLconf和视图里的名字除了提醒我们它的作用外并无特殊之处。

就像在第四部分里提到的那样，因为admin\_objects是先定义的manager，所以它就是我们模型的默认manager，admin也会用它——这样就能保证所有阶段的文章都可以被工具修改。这

的过期的文章queryset。

```
admin_objects = models.Manager()
objects = ViewableManager()
```

接着，我们在模型里实例化一个manager对象。在models.py文件的底部（跟在变量和Meta类之后），加上下面两行，注意这里代码要进行适当的对齐，让它属于Story类：

首先我们把VIEWABLE\_STATUS定义一个整数列表，它的值正好对应了可以让公众看到的文章的状态。这是一个模块级别的属性，即将来添加其他方法的时候也能够访问到它。

```
class ViewableManager(models.Manager):
 def get_queryset(self):
 return default_queryset.filter(status__in=VIEWABLE_STATUS)
 def default_queryset = super(ViewableManager, self).get_queryset():
 return default_queryset.filter(status=VIEWABLE_STATUS)
```

(如果你要复习一下super调用的语法，可以参考第1章，“实践Django Python”。)  
保存到数据库里时，模型的save方法会先被调用，将用户输入的Markdown内容翻译成HTML。  
当我们的代码（或者其他任何使用我们模型的应用，比如Django admin）试图把一个对象

```
super().__init__(self).save()
self.modified = datetime.datetime.now()
self.html_content = markdown(self.markdown_content)
def save(self):
```

就行了（注意要保持和模型类里其他代码的区别）。

回到我们的Django应用：要把Markdown内容自动转换为HTML以便保存，我们还要在模型代码里多加上一点代码。这是一个只有三行的函数，把它放在admin.objects赋值的那一行之上  
这里输入的是从Markdown语法写的纯文本，而函数的输出则是合法的HTML。

```
 Bulleted item one
 Bulleted item two
<h2>An alternate H1 style</h2><p>A blockquote</p></blockquote><h2>...
...
* Bulleted item one
* Bulleted item two
* A blockquote

... An alternate H1 style
...
>>> tidy_markdown("""
<p>Click here to buy my book</p>,
<h2>tidy_markdown("Click here to buy my book (<http://whitejango.com/>)")
<h2>Heading Level One</h2>,
<h2>tidy_markdown("# Heading Level One")
<p>Hello</p>,
<h2>tidy_markdown("Hello")
<>
...
return markdown(text).replace("\n", " ")
>>> def tidy_markdown(text):
>>> from markdown import markdown
```

一行。)

如果更清晰明了。（当我们用Markdown输出更多的HTML时，这些换行就能防止输出超很长的  
tidy\_markdown帮助函数，它能去掉Markdown插入到输出里的换行符号（\n），从而使打印结  
果要是喜欢的话可以在Python的解释器里自己试验一番。在这个例子里，我们定义了一个  
不懂Markdown并不会妨碍你理解这个应用程序，不过这里还是为初学者准备了几个例子。

块里要导人的属性名有着相同的名字。

以上的import语句有点绕圈，不过其实这在Python里是很常见的一种手法，即模块和从模  
块里导入的属性名有着相同的名称。

最后，用下面的语句把Markdown函数从markdown模块里导进来：

库的一部分。你可以在`http://www.freewisdom.org/projects/python-markdown/`找到它。安装完

```

info_dict = { 'queryset': Story.objects.all(), 'template_object_name': 'story' }

urlpatterns = patterns('diango.views.generic.list_detail',
 url(r'^([?P<slug>[-\w]+)/$', object_detail, name='cms-story'),
 url(r'^([?P<slug>[-\w]+/$, search/$', category, name='cms-category'),
 url(r'^([?P<category>[-\w]+/$, cms_search/$', name='cms-search'),
 url(r'^([?P<category>[-\w]+/cms-views/$', cms_views, name='cms-views'),
)

```

下面这个文件是上面代码里include调用的应用层urls.py：

```
from django.conf.urls import include, url
from cms.models import Story
from cms import defaults
```

admin 这行还是不一样，另一行则把所有的 URL 配到 CMS 应用的 “cms/” 上去。如果你还需要其他前缀，比如 “stories” 或是 “pages”，当然也是可以在此处指定的。想了解其他员活

```
URL。这是项目目录里的urls.py。
urlpatterns = patterns('',
 url(r'^admin/*$', admin.site.root),
 url(r'^cms/$', include('cmsproject.cms.urls')),
```

重写完save函数后，模型这一块就算是完成了。在讨论视图和模板之前，我们先来看一下

uris.py里的URL模板或  
Python-markdown2。

要了解更多 Markdown 及其语法，请参见官方网站 <http://daringfireball.net/projects/markdown/>。另外 Python-Markdown 自己也有一些非常有用第三方扩展。事实上，本书正是在 WrapedTables “tables”（见 <http://brain-jarress.livejournal.com/5978.html>）的帮助下用 Markdown 编写完或的！如果你有兴趣的话，可以参考另一个 Python 的 Markdown 项目，见 <http://code.google.com/p/>

前的時間點。

因为在HTML的模型变量被标志为editable=False，所以它不会显示在admin的界面里。这让用户的关系更加清晰，并且消除了用户改动并保存而覆盖渲染出来的HTML的可能性，因为这种覆盖可能带来破坏。所有的修改都是用Markdown源码完成，转换成HTML后保存到html\_content变量里，无需任何关注。另外在保存的时候，我们还一并把modified变量更新为当前时间。

录的社区论坛)，每次页面访问去去计算可能是比较好的选择。

纯数据层王义春对能不能喜欢这种内容可以轻易地从另一个字段计算得来的手段。如果这一转换没什么计算成本的话，我们就不需要保存这个渲染出来的HTML。这也是每个不同的项目经常会做的取舍的地方。这里我们假设计算能力是有限的资源，比如一个昂贵的单片机但是内存却不一定很多的站点。而对于更关注数据大小的站点（比如一个有着几千乃至几百万条记录

接着我們就可以  
〔如圖8.8.1〕

(7.8圖略) 由批子

所示的一个小的弹出窗口。

重慶文獻卷之二

#### 图 8.4 admin 页面

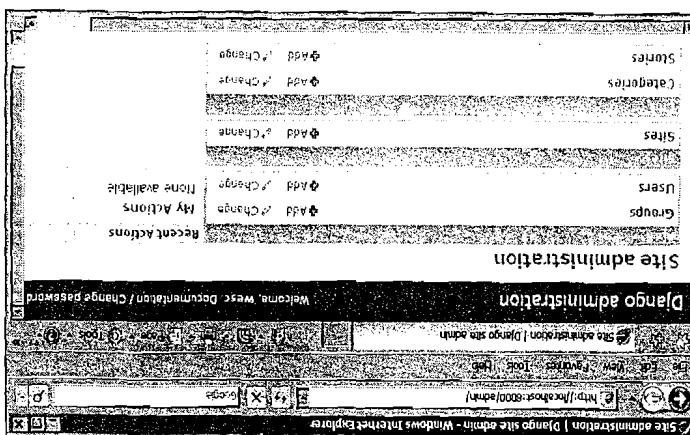


图8-3是点击Add后所显示的Add Story页面。

Admin

对象而不 是默认的名字 object。

这些函数们公用其中一些，`itemplate_object_name`，已经在第4章中提到了。

中国科学院植物研究所植物学国家重点实验室

就像在之前章节里看过的那样，通用视图提供了一个很

数据通常都是相对比较晦涩的，只要适合你的需要就好。

以在这种情况下将图示入述来定义的函数是不行的。不：

三

等人都使用这些视图函数。

饭店之间的协调问题，我们

我们  
那不回的祖國前線

由于这里我们再次利用

配了某个搜索请求的列表。

我们与UML接壤的技术

## 第二章 球形鏡頭

图8.7 加入“Site News”类别

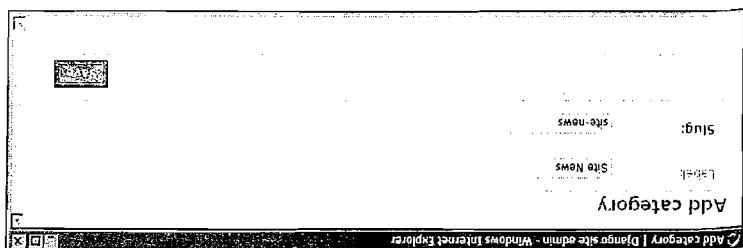


图8.6 在添加文章的同时加入一个类别

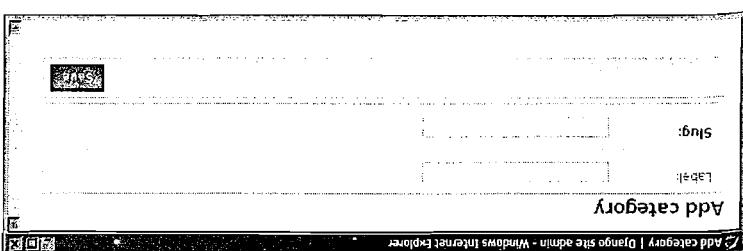
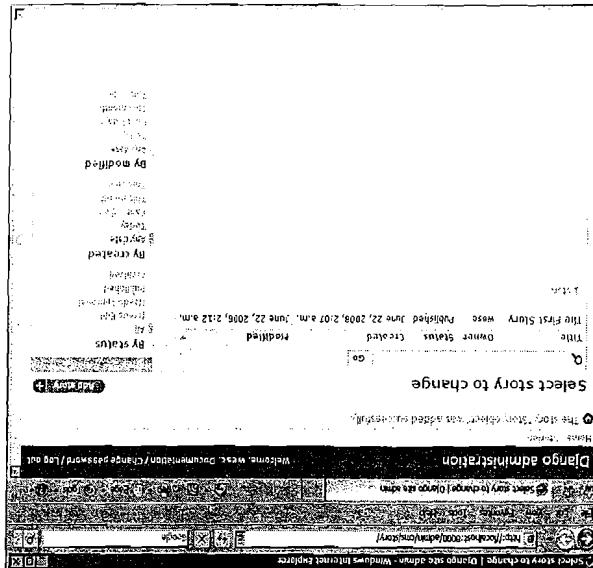


图8.5 在admin里添加一篇文章

试着再随便添加编辑一些文章，确保Published或Archived状态的文章至少各有一篇，这样

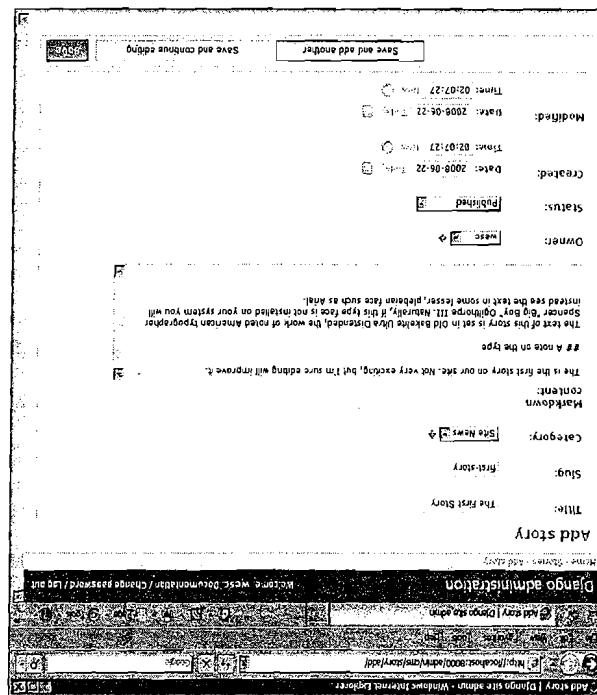
图8.9 在admin里浏览文章的列表；注意右边可选的过滤器



的文章了。

按下保存按钮，你会被重定向到CMS的Story页面（如图8.9），此时应该可以看到刚刚添加

图8.8 完成我们的第一篇文章



我们才能在站点里有东西可看。

### 使用通用视图显示内容

在前面的URLconf里可以看到，我们前台显示的部分基本上用的都是通用视图。不过在分类列表显示里我们还需要一点自定义视图。以下是该应用程序的views.py文件的起始部分。

```
from django.shortcuts import render_to_response, get_object_or_404
from django.db.models import Q
from cms.models import Story, Category

def category(request, slug):
 """Given a category slug, display all items in a category."""
 category = get_object_or_404(Category, slug=slug)
 story_list = Story.objects.filter(category=category)
 heading = "Category: %s" % category.label
 return render_to_response("cms/story_list.html", locals())
```

以上是一个相当简单的视图函数，不过却完成了现有的通用视图处理不了的功能，这也是为什么我们要自己写的原因。接着我们将继续介绍模板，并在稍后再回到第二个自定义视图上来，让它提供一个搜索界面。

### 模板布局

几乎所有的Django项目里都有一个base.html，所有其他的模板都由它扩展而来。这里我们只要扩展两个即可：story\_detail.html和story\_list.html。在cms文件夹里创建这三个文件，并在settings.py文件里为它们设置TEMPLATE\_DIRS。

首先是最简单的base模板：

```
<!DOCTYPE html PUBLIC "-//W3C//DTD HTML 4.01//EN"
"http://www.w3.org/TR/html4/strict.dtd">
<html lang="en">
 <head>
 <title>{% block title %}{% endblock %}</title>
 <style type="text/css" media="screen">
 body { margin: 15px; font-family: Arial; }
 h1, h2 { background: #aaa; padding: 1% 2%; margin: 0; }
 a { text-decoration: none; color: #444; }
 .small { font-size: 75%; color: #777; }
 #header { font-weight: bold; background: #ccc; padding: 1% 2%; }
 #story-body { background: #ccc; padding: 2%; }
 #story-list { background: #ccc; padding: 1% 1% 1% 4%; }
 #story-list li { margin: .5em 0; }
 </style>
 </head>
 <body>
 <div id="header">
 <form action="{% url cms-search %}" method="get">
```

```

 Home •
 <label for="q">Search:</label> <input type="text" name="q">
 </form>
</div>
{* block content *}
{* endblock *}
</body>
</html>

```

稍后我们再解释模板里Django相关的部分。现在，我们需要一个能显示一篇单独文章（story\_detail.html）的模板，就像刚刚所说，它是从base模板扩展而来。

```

{% extends "cms/base.html" %}
{% block title %}{{ story.title }}{% endblock %}
{% block content %}
 <h1>{{ story.title }}</h1>
 <h2>{{ story.category }}</h2>
 <div id="story-body">
 {{ story.html_content|safe }}
 <p class="small">Updated {{ story.modified }}</p>
 </div>
{% endblock %}

```

这在可用的模板里大概算是最简单的一个了——它只需要一个story模板变量。只要传给它一个有title和html\_content属性的对象，它就能够正常工作。

该模板中非常重要的一点就是应用在html\_content变量上的safe过滤器。默认情况下，Django会自动转义模板里所有的HTML以防止用户输入的恶意信息（Web应用里日益严重的安全性问题）。因为我们的Markdown源码都是由可信赖的用户输入的，所以我们有理由相信其内容是“安全的”，因此允许浏览器直接处理HTML，而不用把<em>转义为&lt;em&gt;等。

好几个需要显示多篇文章的视图都要用到我们的列表模板story\_list.html，比如分类列表，搜索结果以及主页。

```

{% extends "cms/base.html" %}
{% block content %}
 {% if heading %}
 <h1>{{ heading }}</h1>
 {% endif %}
 <ul id="story-list">
 {% for story in story_list %}
 {{ story.title }}
 {% endfor %}

{% endblock %}

```

以上就比之前的细节模板稍微复杂了一点。它循环取出story\_list里的每一项，为它们创建<li>元素并且把标题作为链接的文本，而链接则是指向文章的细节页面。

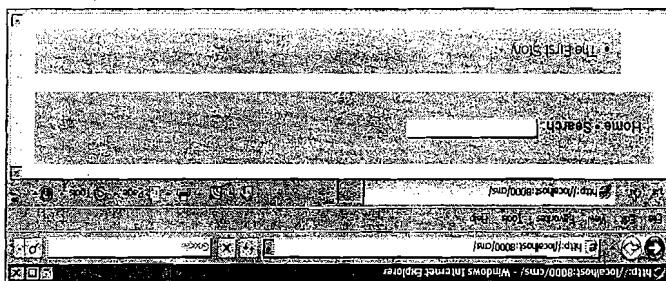
## 添加搜索功能

现在我们来给网站加一个简单的搜索功能，只需要以下几个步骤就能完成。

Site-Search框就运行了(<http://www.google.com/coop/cse/>)，但是如果能对搜索过程以及结果的表

搜索文章内容的功能是必要的。对于一个公共站点来说，你只需要加上一个Google文章的标题即是链接，它们是由之前创建的get\_absolute\_url方法生成的。

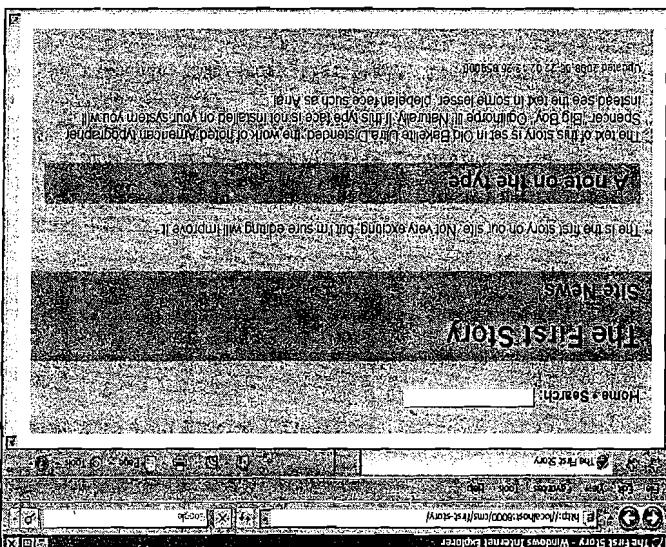
图8.11 列出所有文章的主页



如图8.11所示。

接着，我们访问站点的主页来测试object\_list视图。它的URL为<http://localhost:8000/cms/>。

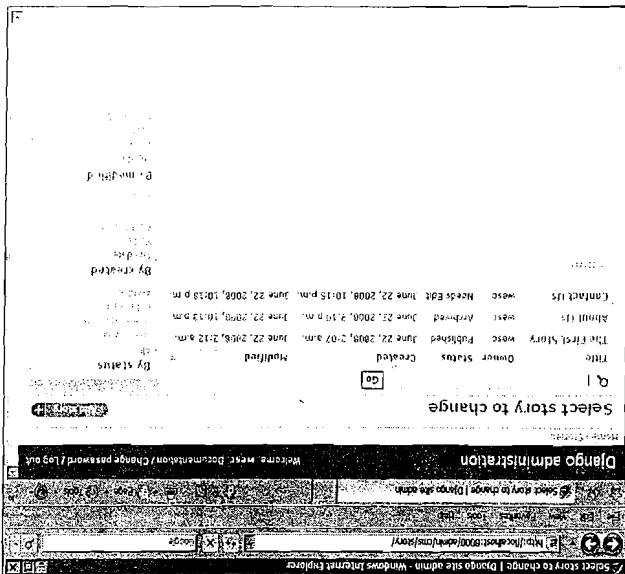
图8.10 我们的第一篇文章的“细节页面”



因为我们用到了slugg变量，所以在浏览器里输入以上URL，结果如图8.10所示。<http://localhost:8000/cms/>

## 显示文章

图8.12 列出全部文章的admin页面



虽然这是一个自定义视图，但是它并不需要一个专用的模板。我们可以直接把story\_list.html模板从重用，只要我们提供它所需要的参数就行了——即在上下文变量story\_list里一个由Story模型对象组成的QuerySet。搜索算法也非常简单，只要能在标题或是Markdown正文里能找到通过表单提交上的文字，那么那个Story就是匹配的。

```
def search(request):
 term = request.GET.get('q')
 if term:
 stories = Story.objects.filter(Q(title__contains=term) | Q(content__contains=term))
 context = {'stories': stories}
 return render_to_response("cms/story_list.html", context)
 else:
 return render_to_response("cms/story_list.html", {})
```

- 编写base.html模板加上一个包含搜索字段的HTML表单，这样每个页面上都能看到它
- 添加一个接受该表单输入，并且查找匹配文章的视图函数
- 至于用显示结果的story\_list.html模板，我们已经创建好了
- 前面的base.html模板，它在头里包含了一个搜索框。要它起作用的话，我们就需要在表单根据文本来的時候专门有一个视图来处理它。

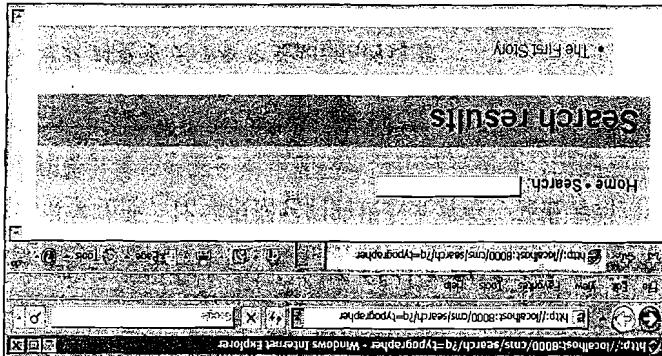
## 注意

在不久的将来，DjangO可能会实现一个更细致的“对象权限”系统，在本书编写的时候，这项新的admin特性还在开发之中。想了解更多信息，请访问withdjangO.com。

## 管理用户

技术上一个用户完全可以修改或删除不属于他的内容，事实上甚至不会限制他们修改拥有权限。我们的系统提供了一个拥有权的概念，每篇文章都和一个特定的DjangO用户名对对象相关联。

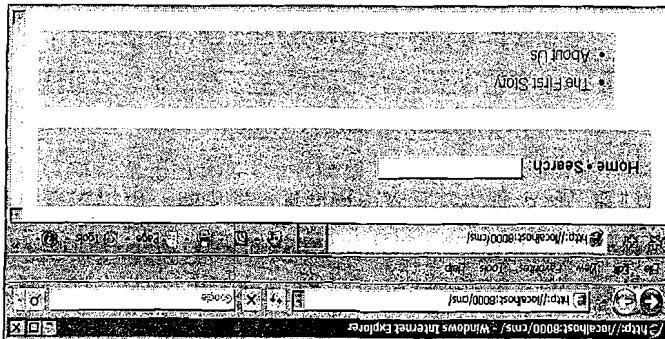
图8.14 搜索结果页面（列出的还是只有可公开访问的文章）



是匹配的，如图8.14。

现在来试一试搜索功能。在搜索框里输入“typographer”，我们可以看到只有第一篇文章

图8.13 显示可公开访问的文章的主要



面没有显示！

但是主页上，我们应该只显示可公开访问的页面及其链接（Published或Archived，这两个都是由VIEWABLE\_STATUS控制的）。当访问主页时，如图8.13所示，注意此时Contact Us页

尽管如此，这种不正规的松散的拥有权管理方式在相互信任的组织内部还是颇为有用的。这就像是在一个办公室环境里，你不会怀疑别人偷了你漂亮的红色订书机或者把你的文件扔进碎纸机里去一样。这里实现拥有权的方便之处在于我们利用了Django内建的用户模型。我们不需要其他任何额外的模型代码。所以我们可以用Django来管理用户。

对于admin的超級用户，你可以用admin来控制谁能编辑用户和用户组，谁又有权限来访问这个简单的内部工作流影响了我们的状态变量以及相关连接：

1. 一个外部的作者或员工提交了页面的内容。这份内容处于草稿形式，并且需要被编辑。
2. 编辑初稿完成，在发布前还需要经审核。
3. 一旦文章被标记为“Published”后，随即便出现在公共网站上。
4. 如果文章过期了，它会被标记为“Archived”。这就是说，它会出现在站点搜索的结果里，但是不会出现在首頁上，“Recent articles”的列表中。

以下这个简单的内部工作流影响了我们的状态变量以及相关连接：

1. 一个外部的作者或员工提交了页面的内容。这份内容处于草稿形式，并且需要被编辑。

2. 编辑初稿完成，在发布前还需要经审核。

3. 一旦文章被标记为“Published”后，随即便出现在公共网站上。

4. 如果文章过期了，它会被标记为“Archived”。这就是说，它会出现在站点搜索的结果里，但是不会出现在首頁上，“Recent articles”的列表中。

注意

你可以在第11章里找到更多关于自定义admin的内容。

注意

注意我们的models.py使用了Django admin里的list\_filter特性来方便地选择属于四个状态下的任何文章。例如，编辑用户可以使用它来选择所有处在Needs Edit状态下的文章，而一名专门负责删除旧资料的实习生可以专注于Archived状态下的文章。

在本章开始的地方我们就提过，CMS架构的种类非常多。你在本章里构建的应用示例根据所需的特性不同，最终可以演变成各种样子。以下是一些建议。

成百上千，把它们全部显示在一頁里就有点吓人了，而且这还可能会拖累网站的性能。同样，如果一个搜索返回了几百个结果，用户也不可能希望一次性能看到它们。好在Django内置了一些对分页的支持，即django.core.paginator模块。想了解更多信息，请查阅Django官方文档。

更重要的是，我们的搜索功能还算能用，但是它和我们熟悉的web搜索引擎比起来还是差了一点。我们在Django内建的全文搜索功能上做了一些改进，但效果并不理想。如果要解决这个问题，最好的办法是使用一个成熟的全文搜索引擎，如Solr或Elasticsearch，它们能够提供更快的搜索速度和更好的用户体验。

## 8.4 改进建议

则是在基于Django的Pastebin。

这一部分还有两个更重要的应用示例：一个使用了Ajax技术来创建一个Liveblog，另一个一个模板层次。

到此，我们希望你已经熟悉了Django应用构建的方式：通过命令行工具创建一个项目和应用，熟悉好模型的定义（包括怎么使用admin），定义URL，使用通用和自定义视图，以及创建Flatpages站点以及更复杂的CMS。

这一章很长，但是你应该看到了如何利用多个Django核心组件和contrib应用来构建简单的

## 8.5 总结

所以我们在本章中讨论的内容是关于如何使用Django来构建一个博客系统。首先，我们介绍了如何使用Django的视图、模型和模板来实现一个基本的博客系统。然后，我们探讨了如何使用Django的中间件来处理请求和响应，以及如何使用Django的管理后台来管理数据。接着，我们学习了如何使用Django的认证系统来实现用户登录和注销功能。最后，我们讨论了如何使用Django的信号和处理器来处理模型的变更事件。

通过本章的学习，你应该能够理解Django的基本框架，并且能够开始着手构建自己的Web应用了。希望你在阅读本章后能够有所收获！

在本章中，我们主要讨论了如何使用Django的视图、模型和模板来实现一个基本的博客系统。首先，我们介绍了如何使用Django的视图、模型和模板来实现一个基本的博客系统。然后，我们探讨了如何使用Django的中间件来处理请求和响应，以及如何使用Django的管理后台来管理数据。接着，我们学习了如何使用Django的认证系统来实现用户登录和注销功能。最后，我们讨论了如何使用Django的信号和处理器来处理模型的变更事件。

通过本章的学习，你应该能够理解Django的基本框架，并且能够开始着手构建自己的Web应用了。希望你在阅读本章后能够有所收获！

在本章中，我们主要讨论了如何使用Django的视图、模型和模板来实现一个基本的博客系统。首先，我们介绍了如何使用Django的视图、模型和模板来实现一个基本的博客系统。然后，我们探讨了如何使用Django的中间件来处理请求和响应，以及如何使用Django的管理后台来管理数据。接着，我们学习了如何使用Django的认证系统来实现用户登录和注销功能。最后，我们讨论了如何使用Django的信号和处理器来处理模型的变更事件。

通过本章的学习，你应该能够理解Django的基本框架，并且能够开始着手构建自己的Web应用了。希望你在阅读本章后能够有所收获！

Ajax的UI部分其实就是漂亮的客户端JavaScript和DOM操作，这些技术是由于近年来越发拥挤的JSON。

注意这些对话的响应部分通常都是HTML或XML（即Ajax里的“X”），又或者是轻型的数据。从实现层面来说，可以把Ajax里的“额外信息”的部分想象成送你的请求，是浏览器和服务器里一个无需重新载入整个页面的普通HTTP对话。稍后我们会介绍这里的细节，现在先“widget”拖放接口特性的个人主页站点。

- 高级“动态的”用户接口——想想Google Maps的地图滚动和放大缩小，或者各种支持器不重载或是重绘整个页面的情况下显示email，收件箱以及表单。
- 无需用户重新载入或者访问他处就能让网页获取额外的信息——想想Gmail是怎么在浏览器当在Web开发里提到“Ajax”的时候，通常指的是一种不但又经常纠缠在一起的行为。

## 9.1 究竟什么是Ajax

注意 和其他应用示例一样，这里我们使用Apache来简化静态文件的处理（对于这个例子来说就是JavaScript）。

不过这也意味着Django并没有把你绑在某个Ajax库上（市面上这种库不计其数），而是把同样的方式提供即时信息，不过通常没有动态更新。

在这一章里，我们要展示一个相对简单的Ajax应用，即“Liveblog”。一个Liveblog是一个列出了系列简短、带时间戳的条目的网页，它能进行自我更新而无需用户干预。要是你对近年来Apple的新闻比较关心的话，你应该已经在各种Mac新闻和谎言的站点（比如macrumorslive.com）上见过这类应用了。同样的概念也可以较小规模地用在普通，静态的blog上，虽然它也用同样的方式提供即时信息，不过通常没有动态更新。

以下的应用示例会向你介绍在Django Web应用里集成Ajax所用的一切知识，同时又不会过多涉及复杂的CS交互或是动画。我们还会在不牵涉具体工具的情况下，指出Django在哪些方面可以和Ajax携手合作。

### 第9章 Liveblog

## 选择一个Ajax库

- 这条“流”按时间顺序显示，即最新的放在最前面。所以最新的信息总是显示在页面顶部。
- 这个“流”由带时间戳的文本段落组成。所以我们模型只需要两个变量。
- 它只跟踪一条连续的消息流。尽量保持简单。
- 应用程序只会有一个网页。不需要其他花哨的东西——我们只是要建立一个一次更新一个事件的liveblog。
- 在深入代码之前，我们要给出一个简单的规格，应用程序要有什幺特性，以及决定使用什幺工具（特别是用哪个Ajax库）来构建它。首先我们先写下一些需求，定义出应用程序要达到的目标。

## 设计应用程序

- Ajax很有用
- Ajax本身是一门完整的编程语言这一事实，网页现在其实已经和传统的GUI动画技术一样好用。当然，因为客户端浏览器只要求了特定部分的数据而非整个页面，另外它创造了一种新的用户体验，因为浏览器窗口不用不停地重绘。这让Web应用感觉起来更接近桌面应用。
- 虽然有时候它们被批评成“华而不实”，但是只要使用得当，高级的动画、拖放功能以及更好的用户体验，一个网页发出请求的能力从各个方面来看都很有用。它能在大量的数据库来讲，一个网页发出请求的能力从各个方面来看都很有用。它能在大量的
- Web 2.0”的特性都可以大大增强用户体验。当你请求带来的减少页面重载的数据和更高的动画特效，都进一步模糊了Web和传统GUI之间的界线。

浏览器和客户端电脑才变得现实起来。若是考虑到样式正确的Web标记元素显示的可能性，所以除了选择工具集，你还要知道那个工具的作用并下载正确的版本。

注意这里media文件夹的结构只是我们自己的约定——Django没有任何规定你需要怎么做。

```
liveproject/
| - urls.py
| - liveupdate
| - templates
| - settings.py
| - __init__.py
| - manage.py
| - views.py
| - urls.py
| - models.py
| - liveupdate
| - __init__.py
|
└── media
```

(输出)：

media文件夹（保存JavaScript的地方），所以我们初给的设置如下所示（来自Unix Tree命令的项目liveproject里。除了应用组件本身，我们在项目范围里还有一个templates文件夹和一个

想起种子开始工作吧！以下的应用示例（我们称之为liveupdate）包含在一个通用的Django

### 9.3 应用程序布局

这里用到的Ajax功能都很简单，上面提到的任何一种框架都可以做到。我们的应用示例采用的是jQuery，不过就作者而言，很大程度上这只是我们随便选择的而已。JavaScript UI项目是Yahoo!近年来出色的成果，专门打包出来供社区使用。

- Yahoo! User Interface (YUI)：(developer.yahoo.com/yui) 这个还在继续进行中的Prototype：(prototypers.org) 源自Ruby on Rails框架，但是已经分离开成为一个独立的类库。

- Mootools：(mootools.net) 它提供了一个极富模块化的下载系统，允许你进行高度自定义的类库设置。

- Mochikit：(mochikit.com) 众多“Pythonic”的JavaScript库之一，它受到Python和Objective-C的很多影响。

- jQuery：(jquery.com) 一个新晋的Ajax库，它提供了一种强大的“串联”语法并可以同时选择和操作多个页面元素。

- Dojo：(dojotoolkit.org) 壮大的Ajax库之一，Dojo吸收了几个较小的库并且提供了多种下载选择。

闲话少说，以下列出了几个最著名的Ajax工具的小结及其获取地址。

看起来还挺麻烦的，不过也是必须的。每次访问页面都需要下载一大块JavaScript对托管的Web服务器来说绝对是数目不小的资源。所以，应该让程序员自行选择并包含应用组件需要的组件。

组织媒体文件，甚至都不一定非要把它们放在项目文件夹里。我们只是出于开发拥有大量 JavaScript、CSS 文件，和图片的大型网站的习惯，在这里专门准备了一个独立的 js 子目录。在该例子中，我们没有使用外部的 CSS 或者图片，但是如果有的话，就还可以放置 img 和 CSS 文件来。把媒体放在 Django 项目文件夹内部可以让它在服务器以及源码控制里的管理简单许多。把 media 文件夹符号链接到 Apache 的 document root 下（要把 Apache 配置成允许 AllowSymlinks）可以在浏览器里面的媒体文件都能正常工作。

我们的 liveupdate 应用里，模型，URL，和视图各有一个文件。根据给定的需求，我们的 liveupdate/urls.py（它应该被 include 在项目的 urls.py 里），它只列出了 Update 对象：

```
from django.conf.urls import url, include
from django.contrib import admin

urlpatterns = patterns('liveupdate.views.generic',
 url(r'^$', 'list_detail.object_list', {'queryset': Update.objects.all()}),
```

现在我们来看看 models.py 文件，它定义了 Update 模型类（包括默认的排序方法）并为它设置管理员：

```
class Update(models.Model):
 title = models.TextField()
 timestamp = models.DateTimeField(auto_now_add=True)
 class Meta:
 ordering = ['-id']

 def __unicode__(self):
 return '[%s %s %s]' % (self.timestamp.strftime('%Y-%m-%d %H:%M:%S'), self.text)

 @admin.register(Update)
```

最后是我们初学的模板 (templates/update\_list.html)，它是用户首次进入站点时看到的“静态的”视图，显示了我们更新列表当时的状态：

```
<html>
<head>
<title>Live Update</title>
</head>
<body>
```

```

<style type="text/css">
 body {
 margin: 30px;
 font-family: sans-serif;
 background: #fff;
 }
 h1 { background: #ccf; padding: 20px; }
 div.update { width: 100%; padding: 5px; }
 div.even { background: #ddd; }
 div.timestamp { float: left; font-weight: bold; }
 div.text { float: left; padding-left: 10px; }
 div.clear { clear: both; height: 1px; }
 </style>

</head>
<body>
 <h1>Welcome to the Live Update!</h1>
 <p>This site will automatically refresh itself every minute with new
 content - please do not reload the page!</p>

 {% if object_list %}
 <div id="update-holder">
 {% for object in object_list %}
 <div class="update" {% cycle even,odd %}"id="{{ object.id }}">
 <div class="timestamp">
 {{ object.timestamp|date:"Y-m-d H:i:s" }}
 </div>
 <div class="text">
 {{ object.text|linebreaksbr }}
 </div>
 <div class="clear"></div>
 </div>
 {% endfor %}
 </div>
 {% else %}
 <p>No updates yet - please check back later!</p>
 {% endif %}
</body>
</html>

```

从逻辑的角度来讲，该模板相当的直白，也没有包含任何JavaScript或JavaScript includes。在下一节里，我们会用jQuery来给模板加入动态的部分。

不过在这之前，我们先测试一下，在给它添加核心功能前先感受一下。激活Admin（如果你不记得怎么做了的话，请参考之前的章节），运行manage.py syncdb，启动Apache，然后访问admin。

你应该可以看到和往常一样的包含了我们Update模型类的admin控件，点击Add然后像图9.1那样填写一些内容。注意因为我们在timestamp变量里指定了auto\_now\_add，所以我们只需要输入文本就好了——这些都是为了更快的更新liveblog。

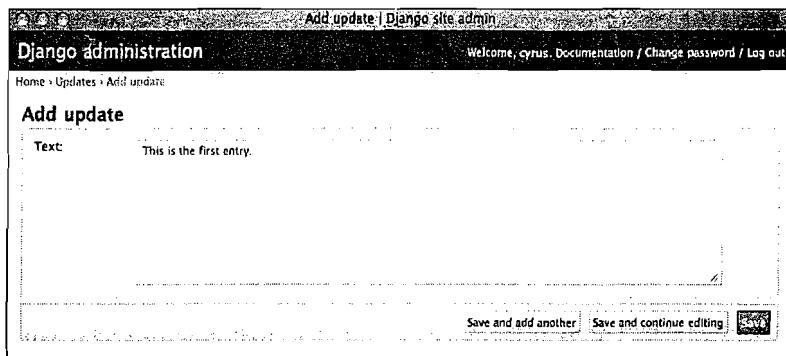


图9.1 添加一个新的Update

添加完之后，admin就会显示新加入的Update，如图9.2所示。

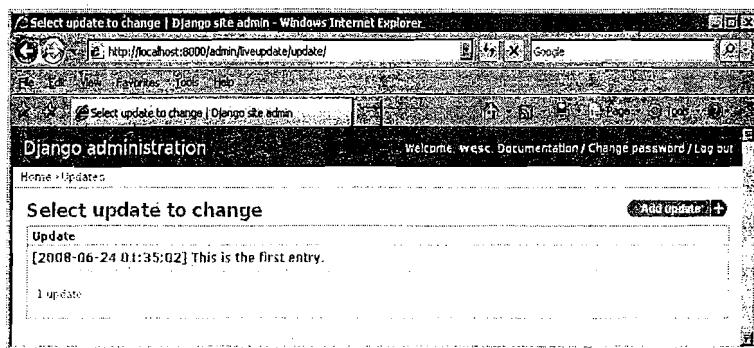


图9.2 登录包含Liveupdate应用的Admin页面

点击前端的根URL，你还可以看到一个“静态”版本的liveblog，如图9.3所示。现在应用程序的基础已经准备好了，我们要开始应用Ajax了。

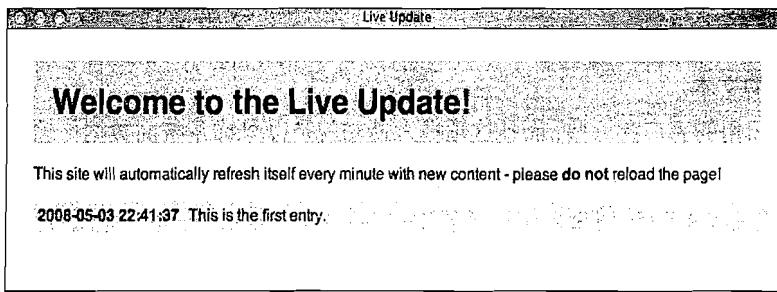


图9.3 只有一条消息的liveblog

## 9.4 加入Ajax

这一节占据了本章剩下的绝大部分篇幅，但是别气馁！在使用Ajax的时候，你需要很多背景知识才能完全理解到底是怎么回事，和往常一样，我们希望能给你一个学习的机会而不是直

## 基本概念

编写代码。

首先看一个使用Ajax需要什么，以及它用什么格式来传递信息，然后是安装和

测试拥有Ajax功能的浏览器，最后展示实际化图标为神奇的代码，包括服务器端和浏览器

操作。告诉你怎么复制粘贴。

实际上来说，网站“实现Ajax”主要包含了以下三个主要任务。

· 导入类库：因为我们要用第三方库来实现功能，所以必须在使用它之前导入模块。

· 定义客户端回调函数：我们用导入的库来编写一个向服务器发出请求的函数，并用

求然后返回HTTP响应的函数，有时也会有一到两个额外的函数（见本章稍后的“创建

知道怎么做”的Django视图）。处理Ajax的视图就是普通的视图（即它也是一个接受HTTP请

求然后返回HTTP响应的函数）。最后，服务器需要知道如何回应你的请求，所以我们必须定义一个

· 定义服务器端逻辑：最后，服务器需要知道如何回应你的请求，所以必须定义一个

视图（见“视图”一节）。

第一步（导入类库）通常就是一个JavaScript包，而那两个函数是简单还是复杂则要取决于

于你逻辑的需要。最典型的，客户端会做比较多的工作，而服务器端经常就是JavaScript代码处理数

据之间的唯一桥梁而已，不过根据你和你用户的特定需求这些部分是可以改变的。另外，如果你

有兴趣使用Ajax库里UI部分的话，不管它是不是和送你请求一起并发执行的，那些代码都应该放在模版这一层上。

## Ajax里的“X”（XML v.s. JSON）

客户端和服务端之间对话理论上可以包含任何JavaScript代码能处理的数据（当然

还得是HTTP可以传输的格式）。但是因为所要的结果一般是要把HTML传输或是添加到网页

上去，所以绝大部分的Ajax对话（就像前面提到的）都是采用XML（ XHTML是它的一种变体）

或者是文本数据格式JSON（JavaScript Object Notation），即一种简单的可以被转换成

而且由于XML和（X）HTML之间紧密的关系，它也非常适合做这项工作，因为JavaScript和其

他Web开发工具都是被设计用来操作这种层次化的数据结构的。此外，你还可以在服务器端就将

数据转化为HTML（例如，用Django的模板和渲染引擎），这样客户端的代码就只需要一个函数

把它放到合适的地方就行了。

按照说XML用的更广泛一些，而且这项技术最明显是以导致系统间数据传输语言而闻名的。

尽管说XML用的更广泛一些，而且这项技术最明显是以导致系统间数据传输语言而闻名的。

JavaScript变量的文本。

客户端和服务端之间对话理论上可以包含任何JavaScript代码能处理的数据（当然

还得是HTTP可以传输的格式）。但是因为所要的结果一般是要把HTML传输或是添加到网页

上去，所以绝大部分的Ajax对话（就像前面提到的）都是采用XML（ XHTML是它的一种变体）

或者是文本数据格式JSON（JavaScript Object Notation），即一种简单的可以被转换成

而且由于XML和（X）HTML之间紧密的关系，它也非常适合做这项工作，因为JavaScript和其

他Web开发工具都是被设计用来操作这种层次化的数据结构的。此外，你还可以在服务器端就将

数据转化为HTML（例如，用Django的模板和渲染引擎），这样客户端的代码就只需要一个函数

把它放到合适的地方就行了。

以下提到了一个优点就是它的语法和Python的数据结构（字符串，字典和列表）惊人的相似。更多

关于JSON语法的信息可以在<http://json.org>找到。

以下是JSON数据结构的一个简单的例子。

关于JSON语法的信息可以在<http://json.org>找到。

227k/s in 0.2s  
100% [=====] 54,075

```
Saving to: ./jquery-1.2.6.min.js
Length: 54075 (53K) [text/x-ec]
HTTP request to jquery.js.googledecode.com[64.233.187.82]:80... connected.
Resolving jquery.js.googledecode.com... 64.233.187.82
2008-05-01 21:52:15 - http://jquery.js.googledecode.com/files/jquery-1.2.6.min.js
https://jquery.js.googlecode.com/files/jquery-1.2.6.min.js
userexample:/opt/code/liveworkspace/media/js/
userexample:/opt/code/liveworkspace $ cd media/js/
userexample:/opt/code/liveworkspace $ wget
https://jquery.js.googlecode.com/files/jquery-1.2.6.min.js
jquery需要在我们的模版里应用，所以它要和自定义的JavaScript一起放在
而在Unix系的主机上，比如Mac OS X和Linux，我们可以用wget或curl命令行工具直接下载文
件，比如（用浏览器的复制功能把下载URL抓下来）：
```

jQuery需要在我们的模版里应用，所以它要和自定义的JavaScript一起放在
而在Unix系的主机上，比如Mac OS X和Linux，我们可以用wget或curl命令行工具直接下载文
件，比如（用浏览器的复制功能把下载URL抓下来）：

1.2.x之前的版本应该也能工作，但是它们缺少getJSON函数，所以需要一条额外的语句来转换
我们采用的是jQuery当简最小化的版本1.2.6，不过任何1.2.x版都能在我们的代码里工作。
而且需要客户端花费不等的CPU时间来解压。

jQuery就是一个单独的类库，而不是其他那样分成多个部分，但是它还是提供了多种
下载方式——最小化的，打包的和压缩的。三个包的功能都是一样的，只不过文件大小不同，

因为我们将已经决定在这个例子使用jQuery了，所以我们要从http://jquery.com上把它下载

## 安装JavaScript

还是网络，这两种格式都有大量的范例，帮助文档和教程。

在这个应用程序里我们采用的是JSON，但是XML仍然是一个很流行的选择。无论是书籍

很多情况下，JSON都可以被直接转换成Python数据结构，这在客户端JavaScript给服务器发送数据时很有用。但是它和Python还是有一点不同的地方，比如JSON的true和false布尔值和Python的就不一样（Python里是首字母大写的），以及JSON的null和Python的None。在这种情况下，使用一个Python/JSON解析器就很有必要了。要知道更多JSON和Python互操作的信息，可以访问上面列出的JSON网站以及这两篇文章http://deron.me/python/json/comparing\_json\_modules/和http://blog.hill-street.net/?p=7。

## JSON变Python

即使你一点也不懂JavaScript，运用你的Python背景也能理解这是一个拥有两个字符串值和一个整数列表值的字典。这样的字典会被转换成JavaScript里的数据结构，随后可以被客户端使用。

```
{"first": "Bob", "last": "Smith", "favorite_numbers": [3, 7, 15]}
```

2008-05-01 21:52:16 (227 KB/s) - `jquery-1.2.6.min.js' saved [54075/54075]

把类库保存到media目录下以后，我们需要在<head>标签里加入下面这行代码，这样就能把它包含到模板里来了（注意 /media/custom 是Apache docroot里一个指向我们liveproject/media目录的符号链接）。

```
<script type="text/javascript" language="javascript"
src="/media/custom/js/jquery-1.2.6.min.js"></script>
```

至此使用jQuery前的准备工作就完成了，我们先设置一个简单的测试，然后才是实现真正的功能。

### 设置和测试jQuery

Ajax库通常都提供了一种简单的方式来访问它的代码，当前浏览器的文档，或当前的DOM对象。jQuery在它几乎所有的功能里都使用了一种相对独特的语法。变量名\$被绑定到一个特殊的可调用对象上，它可以被当作函数一样调用（例如\$(argument)），又可被用来保存特殊的方法（比如\$.get(argument)）。

例如，\$(document)返回一个通常可以和普通的JavaScript document变量相比较的对象，但是jQuery还给它加了点料。其中一个额外的方法就是ready，当页面载入时就会用它去执行一个JavaScript函数（同时避免了JavaScript内置onLoad函数的问题）。

举例来说（并为我们最终的函数做准备），在模板中的<head>标签里，包含jQuery的代码后面加上下面的代码：

```
<script type="text/javascript" language="javascript">
$(document).ready(function() {
 alert("Hello world!");
})
</script>
```

JavaScript和Python一样，函数是“first-class”的或普通的对象。在高级JavaScript开发中会大量修改对象以及传递函数。在这里，\$(document).ready接受一个函数并在页面准备好的时候执行它，而且我们当场匿名地定义了这个函数——这个Python里lambda的手法很像，不过这里它能够把函数写成多行。如果一切正常，而且模板能包含jQuery的话，我们的JS代码就会在你刷新页面的时候弹出一个对话框显示“Hello World!”。这只是最简单的例子，后面还有更加精彩的实现效果。

### 在模板里嵌入JavaScript

为了让事情变得有趣一点，我们把以上代码改一下，在载入页面的时候给我们的列表动态地加上一个<div>（来表示一个Update对象，虽然这有点硬编码了）。当然，这和我们稍后即时更新代码做的事情是一样的。

```
<script type="text/javascript" language="javascript">
$(document).ready(function() {
 $("#update-holder").prepend('<div class="update">\
```

Update更新到数据库里的频率——你可能会碰到很多相同时间的数据!()。但是, 使用ID简单(不需要在查询时间戳字符串解析成Python的datetime对象), 也更加高效(根据ID比根据时间戳V.S. ID

这里实际上有两种排序Update对象的方法: 根据ID和根据时间戳字符串。根据ID比较其解析并将其包裹到HTML里去。

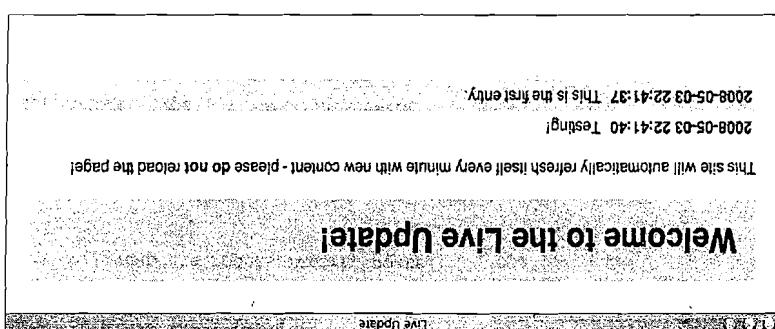
假设我们Ajax视图的URL是 /updates-after/<id>, 这里<id>是请求JavaScript能识别的最新的ID。我们的视图只需要根据这个ID号去做一个简单的查询, 然后返回所有比它新的Update对象就可以了。返回的格式是一个JSON编码版本的模型对象, 这样客户端的JS就可以轻松地对数据进行操作了。

要实现这个需求特性(让我们的JavaScript能请求所有比当前的显示更新的内容)最方便快捷的方法就是设置一个普通的URL。如果我们要在请求里发送多个信息, 那么我们就可能需要基于POST的API, 但是这里我们可以采用简单一点的方案。

假设我们想要实现一个小小的养老服务API, 然后就可以为最后一步进入jQuery的请求机制了。

#### 创建视图函数

图9.4 展示我们即时更新的JavaScript



我们来看看载入页面后的结果, 如图9.4。

选中之后, 我们调用prepend方法把一个对象或HTML字符串加到选中内容之前——所以这里定义的<div id="update-holder">标签, 所以使用了CSS里用来选择对象的#字符。

然后图数返回的结果就是一个代表了所有匹配对象的Query。在这里, 我们查找的是之前模板里定义的<div class="text"><text>Testing!</text></div>。所以使用了CSS的字符串, 它允许你向\$(()查询函数传递一个类似CSS的字符串,

```
<script>
 $(function() {
 var updateHolder = $('#update-holder');
 var updateText = $('#update-text');

 updateText.val('');

 setInterval(function() {
 var update = updateHolder.find('div.text');
 updateText.val(update.text());
 }, 1000);
 });
</script>
```

这个方法假设的是ID永远都是自动增长的，就像Django里的自动ID一样——可是现实世界里并不一定总是如此。

当然，技术上知道怎么做是一回事，而知道怎么正确地把需求翻译成实实在在可行的方案又是另一回事了，通常那都要比看起来棘手的多。

在应用程序的URLconf文件liveupdate/urls.py里加入以下这行代码。

```
liveupdate/urls.py.

url(r'^updates-after/(?P<id>\d+)/$',
 'liveproject.liveupdate.views.updates_after'),
```

以下是liveupdate/views.py里响应的视图函数。

```
from django.http import HttpResponseRedirect
from django.core import serializers

from liveproject.liveupdate.models import Update

def updates_after(request, id):
 response = HttpResponseRedirect()
 response['Content-Type'] = "text/javascript"
 response.write(serializers.serialize("json",
 Update.objects.filter(pk__gt=id)))
 return response
```

利用Django内置的序列化库我们省掉了不少麻烦，它能把模型对象翻译成任意一种文本格式，包括XML、JSON以及YAML。serializers.serialize函数接受一个QuerySet，根据主键（pk）选择对象——在这里，我们只需要ID比传入的参数id更大的那些对象。

然后函数按照我们选择的格式返回字符串结果（这里是JSON），我们再把它写到HttpResponse里去。最后，要让JavaScript能正确解析和使用响应的正文，还需要在响应里设置Content-Type头。

使用可读的序列化文本格式的好处就是我们可以轻松的调试它们。图9.5显示的是在浏览器里手动访问URL <http://localhost:8000/updates-after/0/> 时的结果，这时数据库里只有一个测试用的Update。

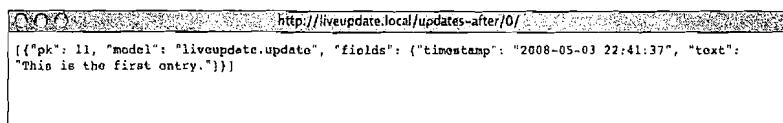


图9.5 在浏览器里测试JSON视图

我们马上就要大功告成了——最后一步是重中之重，即编写和这个API视图交互的JavaScript，并用它的响应来更新页面。

## 通过JavaScript调用视图函数

JavaScript提供了一个内置的计时函数用来在一定的时间间隔里重复执行任意代码，它就是setInterval，它接受一个要执行的函数名称字符串和一个毫秒的时间间隔，比如：

```
setInterval("update()", 60000);
```

这就会每隔60000毫秒或60秒执行一次update()函数。加上另一个好用的jQuery方法getJSON（它会向一个URL发出迷你请求，并将结果解析为JSON），我们就最终完成了Ajax。这里就是最终的代码：

```
<script type="text/javascript" language="javascript">
 function update() {
 update_holder = $("#update-holder");
 most_recent = update_holder.find("div:first");
 $.getJSON('/updates-after/' + most_recent.attr('id') + '/',
 function(data) {
 cycle_class = most_recent.hasClass("odd")
 ? "even" : "odd";
 jQuery.each(data, function() {
 update_holder.prepend('<div id=' + this.pk
 + '" class="update "' + cycle_class
 + '><div class="timestamp">' +
 this.fields.timestamp
 + '</div><div class="text">' +
 this.fields.text
 + '</div><div class="clear"></div></div>');
 });
 cycle_class = (cycle_class == "odd")
 ? "even" : "odd";
 });
 }
};

$(document).ready(function() {
 setInterval("update()", 60000);
})
</script>
```

update内部的逻辑应该相当清晰明了，以下是一些小结：

1. 我们通过jQuery的各种选择方法来获取容器对象（update\_holder）和最近更新的条目（most\_recent）。
2. most\_recent的HTML ID属性（方便起见我们就用服务器端数据库的ID来填充）被用来构建URL，它也是getJSON的第一个参数。
3. 第二个参数通常是匿名函数，它包含了下面的几点。
4. 函数的第一行初始化了“even/odd” CSS类变量。
5. 然后用jQuery的each函数迭代视图里的JSON数据，得到一个序列化的Update对象列表。

6. 把那些要用来构建新的HTML块的Update对象加到

容器里。

7. 最后，在循环结束的地方循环这个CSS类来改变行的颜色。

定义完update之后，在ready里实际执行的代码就是前面提到的setInterval而已。在发布额外的blog条目后，你就可以在保存后一分钟内看到它们被自动加载到网页上。虽然我们没办法实际演示这些代码（动画GIF或视频不太适合以书本形式显示），图9.6是经过若干次更新后的站点的截图。

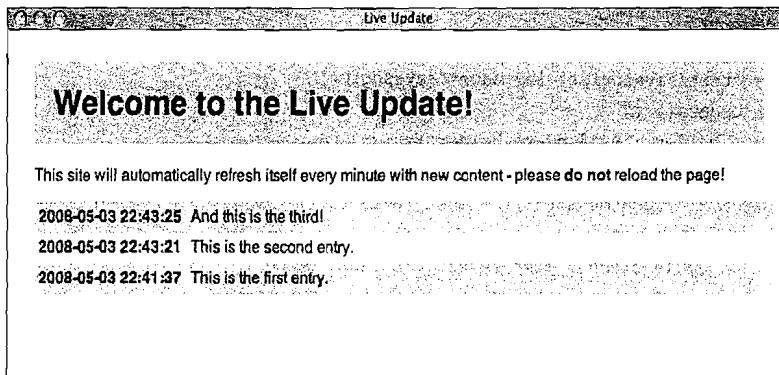


图9.6 包含多条消息的liveblog最终的形态

## 9.5 总结

虽然我们使用了jQuery强大的选择语法和它的getJSON函数来提供客户端的功能，但对在这里设置的服务器端来说，这两项技术都不是必需的。这不仅仅是因为绝大多数其他的Ajax数据库也都有类似的工具，而且实际上不依赖任何特殊的类库而直接使用我们自己的API视图也是相当简单的。

这里关键的地方在于所有的组件都是通过HTTP来交流的（我们的视图接受的常用的GET HTTP请求，换成POST当然也是可以的），而且（经由HTTP发送的）返回值也是一个公开的格式，JSON。就和Django内部的组件尽量保持灵活和模块化一样，配合Ajax使用Django也是依赖于两者都具备开放和良好定义的特性。

这一章更多的关注于Ajax里迷你请求的部分，但是还有很多阅读资料介绍了JavaScript在UI前端能做到的各种惊人效果。我们推荐你可以读一下附录D，“发现，评估，使用Django应用”，记住——虽然在应用程序里加入很多花哨的附属功能很诱人，但是用户还是会希望能够恰到好处。

以下是怎样使用models.py。它定义了一个有五个变量的数据结构，一些

## 10.1 定义模型

Django去做了。

选择是多写一点代码来获得多一些自定义行为还是让Django帮我们处理的时候，我们还是交给Django去做了，不过不是你写的代码就代表了也不用你去调试。所以在该例子中，当遇到人会说这也太懒了，虽然这个例子的精髓在于展示框架能为我们做多少事情。虽然有

这个应用示例的代码和Django社区的pastebin的第一版dpaste.com很相似。虽然今天的dpaste.com不再是一个纯通用视图的应用了，但是它简单和实用的特性依然保留在本章的代码里。

注意

属性的模型文件和要显示它的模板，除此之外的其他部分就都由Django代劳了。

我们几乎不用写太多Python代码就能完成所有这些功能。我们只要创建一个描述数据及其

• 完成清除过期的条目

• 调试着色

• 一个允许我们（站点的所有者）编辑或删除现有条目的管理界面

• 每个条目的详细视图

• 一个显示最近条目的可点击列表

• 一个含有一页必填项（内容）和两项可选项（张贴者的名字和一个标签）的表单

特性

在这一节里，我们将介绍一个依赖于通用视图的简单的pastebin应用。这里是一些  
连接到这个响应，完全取决于你自己的设计。

这都是可以理解的，但事实上并非如此。记住，Django的视图就是一段接受一个HttpRequest请求并返回一个HttpResponse响应的Python代码。怎么把数据对象传递给视图，以及怎么用模板来渲染这个视图，完全取决于你自己的设计。Django新手难免会以为它是那种能显示数据的默认模板，并不仅仅是用完即扔的脚手架而已。Django新手难免会以为它是那种能显示数据的默认模板，

“URL, HTTP机制，和视图”里。)

它提供的通用视图又能让你不用写什么代码就可以创建Web应用。（通用视图的内容在第5章，对很多程序员来说，Django最大的吸引力之一就是可以用Python来开发Web应用。而同时，

Meta选项，和一组通用视图和模板要用到的方法。它还把模型注册给admin并设置了一些admin里和列表相关的选项。

```

import datetime
from django.db import models
from django.db.models import permalink
from django.contrib import admin

class Paste(models.Model):
 """A single pastebin item"""

 SYNTAX_CHOICES = (
 (0, "Plain"),
 (1, "Python"),
 (2, "HTML"),
 (3, "SQL"),
 (4, "Javascript"),
 (5, "CSS"),
)

 content = models.TextField()
 title = models.CharField(blank=True, max_length=30)
 syntax = models.IntegerField(max_length=30, choices=SYNTAX_CHOICES, default=0)
 poster = models.CharField(blank=True, max_length=30)
 timestamp = models.DateTimeField(default=datetime.datetime.now, blank=True)

 class Meta:
 ordering = ('-timestamp',)

 def __unicode__(self):
 return "%s (%s)" % (self.title or "#%s" % self.id,
 self.get_syntax_display())

 @permalink
 def get_absolute_url(self):
 return ('django.views.generic.list_detail.object_detail',
 None, {'object_id': self.id})

class PasteAdmin(admin.ModelAdmin):
 list_display = ('__unicode__', 'title', 'poster', 'syntax', 'timestamp')
 list_filter = ('timestamp', 'syntax')

admin.site.register(Paste, PasteAdmin)

```

这里大多数内容都很面熟了，不同的是它们组成了这个pastebin应用里大部分自定义的Python代码。除了在应用程序urls.py里设置的一些简单规则，大部分的工作都由框架负责完成了。

像这样基于通用视图的应用程序真正展示了Django的DRY（Don't Repeat Yourself，不要重复自己）理念的威力。在我们这个将要构建的例子中，上面核心模型里定义的五个变量

```

 (% endblock %)
 </form>
 <input type="submit" name="submit" value="Paste" id="submit">
 {{ form.content }}

 Syntax: {{ form.syntax }}

 Poster: {{ form.poster }}

 Title: {{ form.title }}

 <form action="" method="POST">
 <h1>Paste something</h1>
 {{ block content %}}
 {% block title %}{% add %}{% endblock %}
 {% extends "base.html" %}


```

有了基础模板之后，我们来创建一个允许用户向我们的应用提交粘贴代码的表单。把下列代码保存为 `pastebin/pastebin/paste_form.html`。（下面会解释为什么这里有一个看似多余的路径名。）

```

</html>
</body>
{% block content %}{% endblock %}
<p>Add one • List all</p>
</body>
</head>
</style>
pre { padding: 20px; background-color: #ddd; }
h1 { background-color: #ccc; padding: 20px; }
body { margin: 30px; font-family: sans-serif; background-color: #fff; }
style type="text/css">
<title>{{ block title }}{{ endblock }}</title>
<head>
<html>

```

现在我们来创建一些基本的模板把内容渲染出来显示给用户。首先我们需要一个基础模板 `base.html`，这种手法在之前的第7章里已经见过。把以下文件保存在 `pastebin/templates` 下。

## 10.2 创建模板

- `create_update` 通用视图，它生成并处理了新添加的 `paste` 目录提交的表单
- `object_detail` 通用视图，它从模型里获取实例，然后把它们发送给模板系统去显示
- `admin` 应用，它为我们自己的数据生成了一个编辑界面
- 创建数据库里的数据表，即 `manage.py syncdb` 命令
- (`content`, `title`, `syntax`, `poster`, `timstamp`) 分别用于:

接着是显示列表的模板。它显示了所有最近张贴的条目，让用户可以从中点击选择。把它保存为pastebin/templates/pastebin/paste\_list.html。

```
{% extends "base.html" %}

{% block title %}Recently Pasted{% endblock %}

{% block content %}

<h1>Recently Pasted</h1>

 {% for object in object_list %}
 {{ object }}
 {% endfor %}

{% endblock %}
```

最后是详细页面的模板。人们绝大多数时间都会停留在该页面上。把它保存为pastebin/templates/pastebin/paste\_detail.html。

```
{% extends "base.html" %}

{% block title %}{{ object }}{% endblock %}

{% block content %}

<h1>{{ object }}</h1>
<p>Syntax: {{ object.get_syntax_display }}

Date: {{ object.timestamp|date:"r" }}</p>
<code><pre>{{ object.content }}</pre></code>
{% endblock %}
```

### 10.3 设计URL

因为我们这个应用程序的结构非常清晰，所以为之创建URL也相当地简单。惟一巧妙的就是利用到通用视图的部分。我们需要设计三个URL模式——一个列出所有的条目，一个显示单独的条目，还有一个负责添加新条目。

```
from django.conf.urls.defaults import *
from django.views.generic.list_detail import object_list, object_detail
from django.views.generic.create_update import create_object
from pastebin.models import Paste

display_info = {'queryset': Paste.objects.all()}
create_info = {'model': Paste}

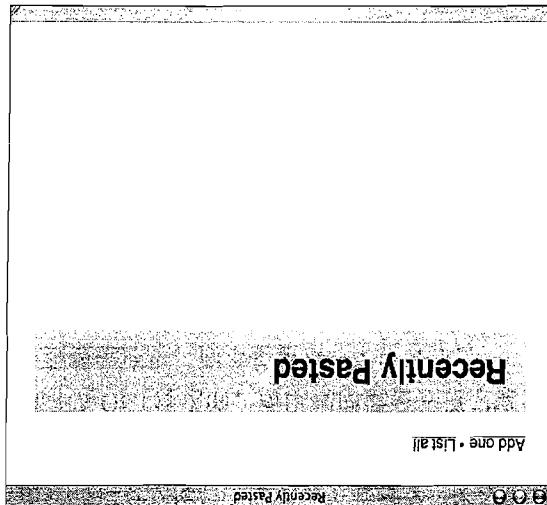
urlpatterns = patterns('',
 url(r'^$', object_list, dict(display_info, allow_empty=True)),
 url(r'^(?P<object_id>\d+)/$', object_detail, display_info),
 url(r'^add/$', create_object, create_info),
)
```

以上就是我们应用程序的核心了。我们从django.views.generic导入了三个函数对象：

- django.views.generic.list\_detail.object\_list

我们来想一下获得这个页面Django帮我们做了哪些事情。首先它解释了我们的URL路

图 10.1 索引的 Pastebin



虽然还没有写下很多代码，不过我们已经有了一个可以工作的应用程序了。现在来跑一下看看。启动应用程序后，我们可以看到如图10.1所示的画面——一个空的pastebin。

10.4 請進行一下

在你大量使用这些\_info字典的时候，有一个值得注意的特殊规则。不是每个请求都会对URLconf里的代码求值。也就是说，如果没什么特别的事情的话，我们在那里设置的对象.allview集有可能会在新对象被用户或网站管理员添加、编辑，或删除后过期。

除了所有Django视图（（不算是不是通用的）都接受的第一个参数HttpRequest外，它们收到了一个额外的字典值，我们在之前的URLconf里定义了两个不同的字典。它们收到了一个名为info的字典，（display\_info和create\_info）在这里虽然还是随便取的（虽然info是这些字典用的后缀），但是它们的内容却是为我们的通用视图专门定制的。list\_detail视图接受一个包含所有合法对象的 queryset。而create\_update视图接受的则是模型类（不是实例）。在object\_list里，我们给字典加上allow\_empty=True来告诉视图即使数据库里没有对象我们也要显示这个页面。

在这个字典里我们还可以包含很多其他可能的值。因为这是自定义通用视图的主要手段，所以它们的数目还是很大的。要了解这些视图完整的可选项，你可以参考Django官方文档。现在我们先尽量保持简单。

- `django.views.generic.list_detail.object_detail`
  - `django.views.generic.list_detail.create_update.create_object`

URL。

这通常是我们提交后最希望看到的结果——即那个提交的视图和一个可以发送给别人的 Pastebin\_detail.html 模板把我们提交的内容渲染出来，如图 10.3 所示。

假设用户正确填写完表单并点击 Paste 按钮，Django 机会（再次通过 create\_update 通用视图）处理表单输入并保存到数据库里。然后，它会重定向到新建对象的 get\_absolute\_url 并且用

假设并把它们用友好的方式反馈给用户。

这里 create\_update 没有显示的一个地方就是验证。要是用户忽略了吧须的变量怎么办？很简单的，这个表单会再次显示。在我们极简化的模板里没有包含查找或显示验证错误信息的代码，但实际上 Django 确实会在 {{ form.errors }} 模版变量里传递了它们。更强大的做法是查找这些错误并把它们用友好的方式反馈给用户。

如果可能的话（把它保存到数据库里，而不是显示一个空表单）。

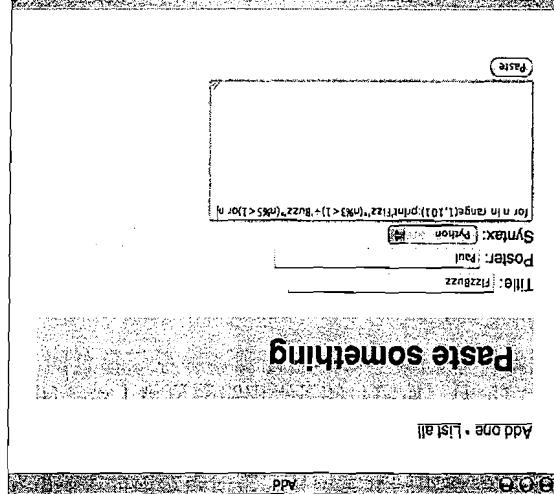
填入内容后点击 Paste 按钮，它就会去再次调用 Django 的 create\_update 通用视图。因为这里数据是经由 HTTP POST 而不是 GET 传递的，所以视图就会知道它需要去处理用户的输入并且

填入内容后点击 Paste 按钮，它就会去再次调用 Django 的 create\_update 通用视图。因为这里

Pastebin 只有方便才会好用，一长串必须的字段那就太麻烦了。

我们友好的 Pastebin 不会逼迫用户做太多的事情。事实上这里只有 Code 字段是必须的。

图 10.2 Add One 表单



这个表单看起来应该像图 10.2 所示。

后将表单显示给用户。注意 <form> 标签和提交按钮都是由我们负责的。

Django 视图和我们的模板通过合作的結果。通用视图 create\_update 重用了我们的一些元数据并生成 HTML 表单元素，并把他们传递给模板里的 {{ form }} 模板变量。我们的模板把这些元数据开放在加上一点内容。点击 Add One 按钮，我们应该能看到一个空的表单。这个表单是

模板文件，用适当的上下文渲染之，最后地结果 HTML 返回给浏览器。

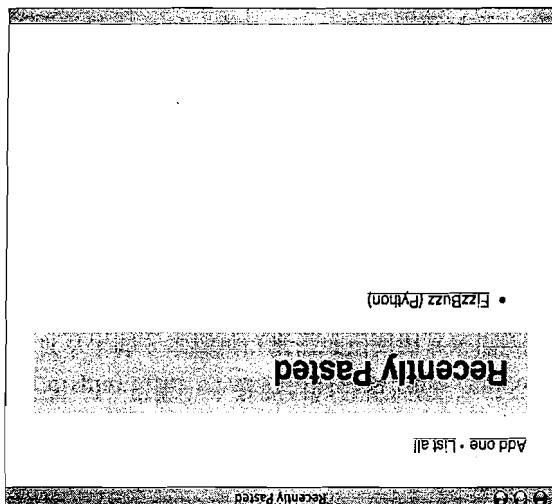
确实要调用哪一個视图，传递一个从我们的模型继承而来的（空的）查询集，找到响应的

## 注意

在实际中，列出所有提交的条目对这类网站来说未必是一个很好的选择。Pastebin用户通常只关心他们自己提交的东西。有一个关于Pastebin的笑话就是，“为什么老人往我的

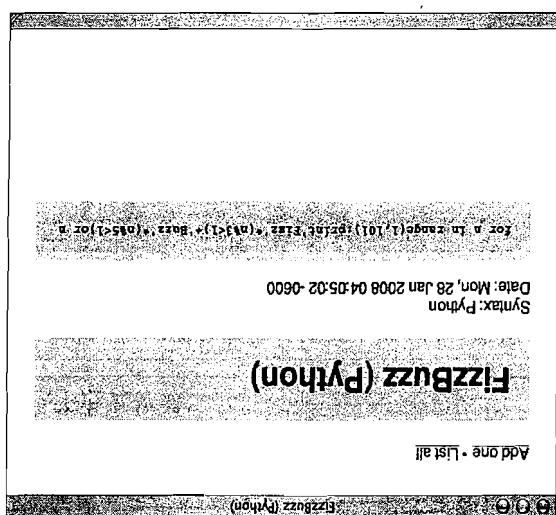
的{% for object in object\_list %}循环生成可点击的列表。有Pastebin的queryset传递给object\_list视图。视图再将对象传给模板，而模板再通过一个简单这个变量漂亮地展示了object-list用视图的能力。URLconf里的display\_info字典把代表所

图10.4 提交条目的列表



object\_list通过视图和pastebin里其他的条目。List All链接会把他导向正确的位置。Django的如果用户要浏览pastebin里的条目。List All链接会把他导向正确的位置。

图10.3 新张贴的条目



除了没有中括号以外，传给slice过滤器的值和我们传给一个普通Python对象的值是完全一样的列表里前十个结果并发送之。

models.py里的ordering = (‘-timestamp’)，它们都按照时间戳排序。然后模板里的for循环取出它的工作原理是：URLconf向模板传递一个代表数据库里所有paste对象的queryset。根据

```
(for object in object_list|slice:"10")
```

只会在paste\_list.html模板的第六行稍微改一下代码，给object\_list切片就行了。

是用来视图，所以管理它的最佳位置就是模板。但是用视图，所以我们这里用的列表很快就会变成庞然大物。要限制这个列表中的条目数量有很多种方法。因为当这个网站变得非常忙碌的时候，

## 10.5 限制最近Paste显示的数量

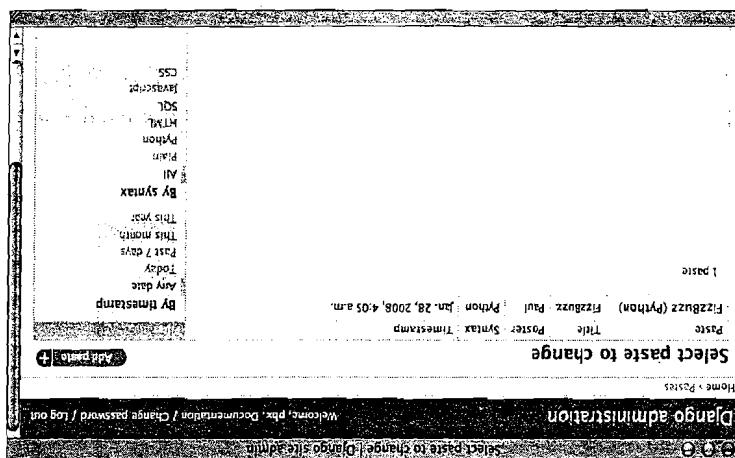
另外还有定期清理旧的数据。

的情况下面出很多改进的。这里就举三个例子：控制最近条目的列表，给代码加上语法高亮，增强这个应用的比较合理的策略应该是开始使用自定义视图，不过你还是可以在保持通用视图到此你已经完成了一个完全由Django通用视图驱动，可用并且好用的应用程序。虽然要再

的模型变量，所以每一项可点击的名字都可以根据可用的消息来调整。

注意因为admin的list\_display里第一个参数用的是模型的“unicode”方法而不是某个特定

图10.5 admin界面



类里设置的选项，它在admin里看起来应该如图10.5所示。

最后别忘了除了我们“设计”的部分之外，还有一个admin应用。按照我们在PasteAdmin

PasteBin里贴一些乱七八糟的东西？”回答是，“制造垃圾信息的人不会‘关心’他们给出的用户发送什么内容。”一个显示最近条目的PasteBin列表就能很方便的完成这个功能，虽然对商业信息来说有点不怎么协调。

```

 {& endblock %}
 </script>
 <script language="javascirpt">highlightAll("code");
dp.SyntaxHighlighter.HighlightAll("code");
 <script language="javascirpt" src="/static/js/shBrushCSS.js"></script>
 <script language="javascirpt" src="/static/js/shBrushSql.js"></script>
 <script language="javascirpt" src="/static/js/shBrushHaskell.js"></script>
 <script language="javascirpt" src="/static/js/shBrushXaml.js"></script>
 <script language="javascirpt" src="/static/js/shBrushPythion.js"></script>
 <script language="javascirpt" src="/static/js/shScore.js"></script>
 <script language="javascirpt" src="/static/css/SyntaxHighlighter.css"></script>
 href="/static/css/SyntaxHighlighter.css"></link>
<link type="text/css" rel="stylesheet"/>
({ object,content })</pre></code>
<code><pre name="code" class="(& object.getSyntaxDisplayLower)">

Date: ({ object.getTimeStamp(date:"F")}</p>
<p>Syntax: ({ object.getSyntaxDisplay })

<h1>{({ object })}</h1>
</block content %}
({ block title %})({ object })({ endblock })
({ extends "base.html" })

```

首先，重新paste\_detail.html模板如下。

们简单地介绍一下怎样给我们的Syntax Highlighter的代码加上语法着色。

在项目的网站上你可以找到完整的安装指导和各种使用Syntax Highlighter的示例，这里我

的名字很简单，就叫Syntax Highlighter，你可以在Google Code上找到它 (<http://code.google.com/p/syntaxhighlighter>)。

有一个叫Alex Gorbatchev的程序员用JavaScript创建了一个出色的国产应用的着色工具。它的名字很简单，就叫Syntax Highlighter，你可以在Google Code上找到它 (<http://code.google.com/p/syntaxhighlighter>)。

但是实现它最简单的途径就是在客户端里用JavaScript完成着色。

要是一个pastebin知道如何正确的将提交上来的代码应用语法高亮的话就会有用的多（也更有吸引力）。要做到这一点方法有很多（包括一个很棒的Python/Pygments，dpaste.com用的

要是一个pastebin知道如何正确的将提交上来的代码应用语法高亮的话就会有用的多（也

## 10.6 语法高亮

更好的地方了。

而且，因为选择在列表里显示多少数目完全是在显示器上的决定，所以再也没有比在模板里嵌套

划算的。所以在URL里指定Paste.objects.all()然后在模板里切片是不会引入任何负面影响的。

如果Django的queryset不是“懒惰”的话（就是说我们要传递整个对象的列表然后除了前

```

[15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1]
>>> print number_list[:10]

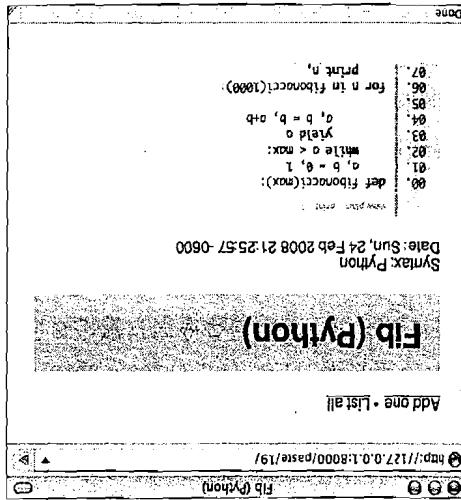
```

样的。所以Python里一个对应的例子就如下所示：

- 张贴到pastebin的代码通常都不会存在很久，所以最好能自动定期清理过期的条目。而最近的方式当然是利用服务器环境之外运行的Django脚本就是Django的“就是Python”Cron job和其他用在Web服务器环境之外运行的Django脚本都是Django的“就是Python”这种方式所展现出来的力量的又一极致体现。编写一个操作你的Django应用里的对象几乎不需要什么Django专用的东西。下面的脚本做出了以下的假设：
  - 环境变量DJANGO\_SETTINGS\_MODULE被设置成包含指向向项目设置文件的Python路径名的字符串（例如，“pastelite.settings”）。
  - 在项目的settings模块里设置EXPIRY\_DAYS。
  - 项目的名字是“pastelite”。如果以上这些是正确的，那么除了测试和部署，你就不再需要做别的事情了。

## 10.7 通过Cron job清除非

图10.6 经过着色后的Python代码



```
我们在<pre>标签上加了name和class属性。这让我们代码块可以在执行时被JavaScript找到。
<pre name="code" class="{{ object.get-'syntax-highlighter'}}">
这样就完成了。当浏览器渲染页面的时候，就会运行语法着色的JavaScript代码，在用户看
到它们之前转换单色的代码。其输出结果如图10.6所示。
```

希望到此你已经承认了Django通用视图的能力。我们的pastebin示例虽然实现的很简单，但是想想它具有的特性：输入验证，根据post重定向，细节和列表视图等。但更重要的是否知道有了Django我们就在于给它打下了坚实的基矗。如果我们要本地化应用，或是通过添加缓存以便承受Digg级别的流量，Django都能为我们做到。

## 10.8 总结

注意 根据数据库引擎的不同，你还可以定期“回收垃圾”，或重新声明被删除记录所留下的空间。你可以在<http://djangosnippets.org>上找到一段类似的操作SQLite数据库的代码。

脚本的最后一行就是关键所在——它使用Django的ORM来选择数据库里所有时间戳比计算出来的 cutoff更老的对象，然后把它们全部删除。



可以说明，admin app是Django皇冠上的明珠。这种说法看有点奇怪，一个在“contrib”目录里得组件能得到这么高的评价，毕竟这种目录的组件通常都是可有可无的插件而绝非核心。不过Django工程师们的规定是正确的——在使用Django时并不一定要使用admin，所以它不应该把所有的东西都塞进admin这个目录中。但是请不要忘记：admin app是一个非常强大和引人注目的应用组件。如果你不应用ModelAdmin子类来进一些定制的行为。比如，在前面的内容里，你已经看到了如何通过ModelAdmin子类来进一些定制的行为。比如，只露设置一下list\_display，list\_filter和search\_fields这几个就能满足你大多数的基本定制需要。另外我们还提到，在以前例子中所使用的admin站点都是“默认的”admin站点——就是说你还可以同时设置多个不同的admin站点来获得更大的灵活性。比如，你可以让不同的用户看到不同的admin界面。

不过最终admin还是不能满足你。经常会有Django的新手在使用admin一段时间后，说出这样的话来，“虽然我已经读过文档了，但就是这件事情我还不知道怎么在admin里完成。要是加上这个功能的话，它就完美了！”

接下来，我们要向你展示一些定制以及扩展admin app的方法来完成人们经常提到的要求。

## 11.1 自定义Admin

在这一章里，我们要讨论一些可以应用到Django上的高级技术，例如生成RSS等聚合方式，自定义admin的行为，以及模板系统的一些高级用途。

自定义admin的行为——你可以在自己的喜好跟着读。不过Django工程师们的决定是正确的——在使用Django时并不一定要使用admin，所以它不应该把所有的东西都塞进admin这个目录中。但是请不要忘记：admin app是一个非常强大和引人注目的应用组件。如果你不应用ModelAdmin子类来进一些定制的行为。比如，在前面的内容里，你已经看到了如何通过ModelAdmin子类来进一些定制的行为。比如，只露设置一下list\_display，list\_filter和search\_fields这几个就能满足你大多数的基本定制需要。另外我们还提到，在以前例子中所使用的admin站点都是“默认的”admin站点——就是说你还可以同时设置多个不同的admin站点来获得更大的灵活性。比如，你可以让不同的用户看到不同的admin界面。

第11章 高级Django编程  
第12章 高级Django部署

# 第四部分 高级Django技术及其特性

根据你打算要深入自定义的程度，有时候与其费尽心力把admin改的面目全非，还不如直接自己开发一个算了。如果现有的技术不适合你的话，完全可以考虑创建你自己的admin。除此之外，下面是在admin里一些比较高级的定制应用。

### 通过Fieldsets改变布局和风格

admin app里的fieldsets设置允许你细致的数据显示的方式。例如希望给某些特定元素加上CSS，按一定的方式组合变量，为选定变量添加JavaScript输入帮助，或是在初启状态下隐藏一些变量时，这个功能就很有用了。

以下是一个简单的模型例子，它通过fieldsets定制了显示的方式。

```
admin.site.register(Person, PersonAdmin)

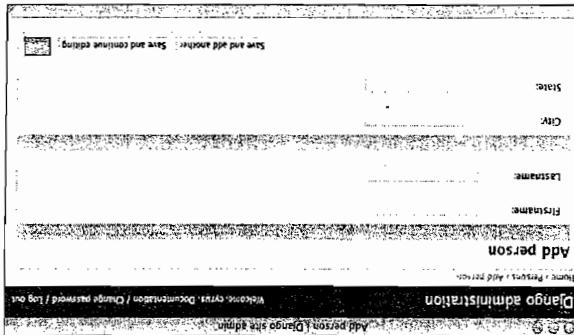
class PersonAdmin(admin.ModelAdmin):
 fieldsets = [
 ('Name', {'fields': ('firstname', 'lastname')}),
 ('Address', {'fields': ('location', 'state')}),
]
 list_display = ['firstname', 'location', 'state']

 def get_queryset(self):
 qs = super().get_queryset()
 qs = qs.order_by('state', 'location')
 return qs

 def save_model(self, request, obj, form, change):
 if obj.state == 'CA':
 obj.location = 'San Francisco'
 else:
 obj.location = 'New York'
 super().save_model(request, obj, form, change)
```

对于Python来说，fieldsets设置就是一组包含两个元组的列表。其中第一个元组是这组变量的标  
题字符串，第二个元组则是对应这个组设置的字典。在字典里的键是选项的名字（稍后列出），对  
应的值则根据特定选项的不同而不同。在上面的例子中，变量选择被设置成一个变量名的元组。  
对于admin来说，fieldsets设置就是两个元组的列表。其中第一个元组是这组变量的标  
题字符串，第二个元组则是对应这个组设置的字典。在字典里的键是选项的名字（稍后列出），对  
应的值则根据特定选项的不同而不同。在上面的例子中，变量选择被设置成一个变量名的元组。  
这个例子的结果就是变量根据fieldsets里的设置组织在一起的显示结果，如图11.1。

图11.1 admin里的新变量



admin.site.register(Person, PersonAdmin)

根据你打算要深入自定义的程度，有时候与其费尽心力把admin改的面目全非，还不如直接自己开发一个算了。如果现有的技术不适合你的话，完全可以考虑创建你自己的admin。除此之外，下面是admin里一些比较高级的定制应用。

根据你对模板系统知识，你一定猜到了开头的extends标签已经在背后为你做了很多工作。没错，真正在所有admin页面靠后的是一个更复杂的模板 base.html。不过其中两个你要定制的东西（页面的<title>和<h1>标题）已经被抽取到 base.html 里了。

只看把 “Django site admin” 和 “Django administration” 改成任何适合你应用需要的值然后保存模板就可以完成定制了。但问题是你要把它保存到哪里去呢？

要覆盖这个模板的内容，你需要创建自己的拷贝，然后帮助Django模板加数据在默认模板之前找到它。你可以把这个文件放在你应用程序里的template目录下（如果你使用的是app\_directories模板加数据），或者放在TEMPLATE\_DIRS里的某个位置下。

```
(% block nav-global %) {# endblock #}
(% endblock %)
<h1 id="site-name">{# trans 'Django administration' #}</h1>
(% block branding %)
(% block title %) {{ title|escape }} {# trans 'Django site admin' #} {# endblock %}
(% load i18n %)
{% extends "admin/base.html" %}
```

（稍作整理以便展示）。

本质上admin就是Django的应用程序，但是从各个方面来说它都是相当复杂的。它错综复杂的模板初学者通常未能驾驭。比较合适的入门模板是base.html。以下是我们代码示例中解释。

当然，实际上你并没有替换它，只不过是你自己的文件重写了原本的模板，我们下面会给出解释。

除了修改简单的设置，你还可以通过替换一系列多个基础模板来大幅调整admin app的外观。

## 扩展基本模板

- `description`: 该选项指定了一个用于描述field组的字符串。你可以把它想象成一个组织级别的help\_text。Django admin会编辑了这部分默认的CSS风格，并将它清楚的显示出来。
- `fields`: 前面说到，该选项指定了要在admin里显示在一起的变量名。如果选项目的这种设计更为类似这样的简单标记语言了不少。
- `help_text`: 该选项指定了一个用于描述field组的字符串。否则该组织就不会显示任何标记。

通过选择添加的CSS类则正好为你自己定义JavaScript代码提供了接口（假若你要给textare添

如果采用的是JavaScript“渐进增强”(progressive enhancement) 的开发方式的话，

### 注意

“monospace”则可用于HTML的textarea显示代码；另外还有“wide”，它比fieldset在admin里获得更大的宽度（虽然这是固定的）。如果你想知道自己默认的“media”目录里装得更多的话，可以去看一下django.contrib.admin里的“media”目录。



你可以通过定义一个包含特殊feed类的单线模块来配置feed，这个模块随后会被直接导入。

## Feed类

以下例子是基于第2章中开发的Blog应用之上。  
为了向将这些对象传递给Syndication app。理解它的工作方式最简单的办法就是直接看代码。  
只有两个必须的步骤：首先是一个生成feed对象的特殊类；其次是在URLconf里的规则，它规定了Syndication app的配置自定义非常高，不过只使用默认值就能轻松地开始使用它了。这里的工作。

Django提供了一个非常方便的工具，Syndication app (`djangoproject.contrib.syndication`) 可以在任何模型对象的集合上生成RSS或是Atom feeds。这极大地简化了为Django项目加入聚合功能的工

## 11.2 使用聚合

通过配合使用Django预定义的权限和应用相关的代码，你可以按照任何你喜欢的方式控制访问。

...  
def secret\_meeting\_create(request):  
 user\_passes\_test(lambda u: u.can\_create\_secretmeeting)  
  
添加新的secret\_meeting的话，相应的装饰器函数就应该如下所示：  
SECRETMeeting模型以及对应在auth系统里的权限，如果你打算创建一个视图来允许特定的用户更新、删除权限，不过在Python代码里也是可以访问这些权限的。比如说你现在有一个app提供了非常方便的权限系统来让你管理用户。你知道每个模型在admin里都有它自己的创建、更新、删除权限。有了lambda，我们可以直接在装饰器里定义“匿名”函数。  
回到我们的例子上来，假设你需要比“是或不是staff”更好的分类。Django的admin和auth

...  
def for\_staff\_eyes\_only(request):  
 user\_passes\_test(lambda u: u.is\_staff)  
  
lambda是什么，请参见第1章）。有了lambda，我们可以直接在装饰器里定义“匿名”函数。

为user\_is\_staff还要专门定义一个函数，是不是太浪费了？这时可以用lambda（如果不知道

print "Next secret staff meeting date: November 14th"  
def for\_staff\_eyes\_only(request):  
 user\_passes\_test(user\_is\_staff)  
  
有了函数定义，就可以这样使用装饰器：

def user\_is\_staff(user):  
 return user.is\_staff  
  
甚至可以很简单：  
一个函数必须接受一个User对象并且返回一个布尔值 (True或False)。这种函数不一定很复杂，

建立一个包含如下代码的feeds.py文件：

```

from django.conf import settings
from django.contrib.syndication.feeds import Feed
from mysite.blog.models import BlogPost
from django.core.urlresolvers import reverse

class RSSFeed(Feed):
 title = "My awesome blog feed"
 description = "The latest from my awesome blog."
 link = "/blog/"
 item_link = item_link

 def items(self):
 return BlogPost.objects.all()[:10]

```

到URLconf里。Django没有规定这个文件的位置（甚至叫什么名字也无所谓），但是比较常见的是放在应用程序目录下，同时把它命名为feeds.py。所以在你的mysite/blog目录里，创建完能处理feed生成的类后，剩下的就是给这个feed安上一个可用的URL了。继续我们blog的例子，编辑更新URLconf文件 (mysite/blog/urls.py) 如下：

```

url(r'^feeds/$', FeedDoesNotExist.as_view())

```

为Feed指定一个URL

Syndication app会通过Python的“duck typing”智能的分辨出你正在用的是哪一个类别。

Syndication app会按照和上面列表的相反顺序查找这三个选项。如果发现了一个名字匹配且接受一个显式参数的方法 (当然引用feed总是有的)

- 接受一个显式参数的方法 (一个独立的feed项)
- 没有显式参数的方法 (当然引用feed的set总是有的)
- 硬编码值 (比如前面的title, description和link)

以及更灵活的指定这些必需字段的方法。理解feed类的很多属性其实都属于下面三个类别之一，对学习定制化很有帮助：

这个类只包含了Syndication app运行所必须具有的组成部分。其实还有很多可选的属性，所以我们可以得到的是最新的十个帖子。

items方法是这个类的核心——它决定了要返回什么对象给Syndication app。这里返回的是blog帖子里的前十篇，因为BlogPost模型在Meta里把ordering属性设置为“-timestamp” (时间递增)，所以我们得到的是最新的十个帖子。

Syndication app会自动去调用它。

调用 (确实是不得不写)，更好的做法是为BlogPost模型添加一个get\_absolute\_url方法，置于/blog/的意思是点击任何单独feed项都会把用户转去blog首页——要是你觉得这点不灵活的话 (确实是很不灵活)，更好的做法是在Meta里把ordering属性设置为“-timestamp” (时间递增)。

item\_link 属性决定了阅读器要看哪个单独feed项的Web页面时载入的页面。item\_link设定了什么页面和feed关联。站点的名字则是由你提供的局部URL决定的。

以上title和description属性是给feed的用户 (例如，一个桌面RSS阅读器) 用来标识feed的。

还有一个包含如下代码的feeds.py文件：

Djanggo凭借的有两点。首先Djanggo的模板语言是基于文本，而不是XML的。如果一个Web的工具而已。但其实Djanggo完全适用于各种类型的内容以及各种传输方式。

Djanggo是一个Web框架，所以自然而然地我们会觉得它就是一个通过HTTP传输生成HTML框架只能处理XML模板的话就会在生成纯文本报告或是email的时候捉襟见肘。

### 11.3 生成下载文件

正要你自己写的机会长平基微。

如果这里介绍的自定义内容还是不能满足你的要求，请进一步查阅Djanggo的官方文档。真很少的代码和配置就能生成功能。

的处理各种类型的（或者说各种糟糕的）发布feed。而通过Djanggo的Syndication app，你只需要自己的Universal Feed Parser (<http://feedparser.org/>) 提供了超过三十个测试用例来确保它能稳定的运行。

表面上看来很简单，但是如果你自己动手写这些feed输出的话还是很麻烦和很容易出错的。整个一个星期（或不定期）更新的站点来说，提供RSS或Atom都是一大进步。虽然这些格式对一个完

再论Feeds

要怎么把请求和正确的Feed类联系起来。

如果回最新开帖子的feed就可以取名为“Latest”。Syndication app对比如没有做任何限制。它只关心虽然在这个字典里的键（“rss”）正好是feed的类型，不过这个名字是可以任意取的——比如其他类型的feed，只需要建立相应的类然后在这个字典里引用它们就可以了。

它是一个字典，里面只包含了两个元素，将字符串“rss”映射到RSSFeed类。如果要添加参数，这个字典用来向视图函数传递更多的参数。这里传递的是一个叫feed\_dict的参数，{“feed\_dict”: {“rss”: RSSFeed}}；在任何urlpatterns元组里，我们都可以说提供像这样的第三个参数：feed；这是从django.contrib.syndication.views里导入的视图函数。

“FOO”传递给视图函数。

也就是在URL /blog/ 下，这意味着它会匹配 /blog/feeds/FOO/ 这样形式的URL，然后把 • rafeeds(?)P<url>.\* /\$: 这是我们的URL正则表达式。因为我们在“blog”应用内部，另外我们还添加了一个看起來很复杂的URL模式。它可以拆成三个部分：

以上只添加了三行代码。其中前两行是import，我们从Syndication app里导入feed view对

```
(

 url(r'^feeds/(?P<url>.*$)', feed, {'feed_dict': {'rss': RSSFeed}}),
 url(r'^$', archive),
 urlpatterns = patterns('',
 from_django.conf.urls.defaults import *
 from django.contrib.syndication.views import feed
 from mysite.blog.views import archive
 from mysite.blog import feeds
 feeds import RSSFeed
 RSSFeed import feed
 feed import views
 views import archive
 archive import feeds
 feeds import mysite.blog
)
)
```

本。它需要的环境是一个System对象，system。

下面的模板可以为每个主机的Nagios生成service-check文件，同样这也是一个简化的版

以下是一个系统-服务体系模型示例的简化（在实际的应用程序里，这些信息被分散到多个模型中）。

用来提供系统和服务的信息。随后这份信息会导出成Nagios可以理解的格式，让用户能跟系统

（除了其他的）生成一部分Nagios配置。它作为每一个Django应用部署在一个中心数据库系统上，

在编写本书的时候，本书的作者之一正在编写一个小型的，高密度定制的内部应用程序来

置文件格式对Django的模板系统来说再好不过了。

和许多类似的项目一样，它也遵循了Unix的惯例，采用纯文本配置文件作为发行格式。这种配

首先和你分享一个现实世界里的例子，即著名的开源监视系统，Nagios (<http://nagios.org/>)。

## Nagios配置文件

确的工具做正确的事。

这里是一些使用Django来生成非Web页面输出的例子。其中有些用到了模板系统，就像刚

一个Content-Disposition头来强制下载。

其次，Django驱动网站的HTTP机制允许你调整HTTP的包头信息。所以，你可以把

我们没有直接返回`HttpServletResponse`，而是先创建了这个对象，设置其内容和`MIME`类型。然后为响应设置`Content-Disposition`头来指定一个附件。在视图最后返回这个响应对象就能自动传输过程，让用户接生成的文件。

## Import von Objekt

它能简单地生成VCARD数据。

VCards是一个基于文本的名片格式。很多流行的地址簿，emaiLi和PIM（Personal Information Manager，个人信息管理器）产品都支持导入这种格式，包括Evolution，The Address Book和Microsoft Outlook。

VCard

上面的模板在使用后会返回给我们一个叫用的Nagios文件，它将system认为Nagios主机，同时输出任何与之相关的服务，并用一个网络检查命令来确认记录的TCP端口正在运行中。

注意不要被这里的单个大括号弄糊涂了，它们是Nagios文件格式的一部分，不会被Django模板系统解析。模板系统只关心两个大括号或是大括号加上百分号。

当用卢靖霖访问到这个视图的时候，还是会返回一个HTTPResponse对象（如果你还记得的话，所有的Django视图都需要求这一点），但是它包含的内容将是text/csv，而非text/html，并且大多数浏览器在遇到Content-Disposition头时，都会触发下载载而不会把它显示在浏览器里。

```

import csv
det_csv_file=requests.get('http://www.webscantech.com/CSVFile/People.csv')
with open('People.csv','w') as f:
 writer=csv.writer(f)
 for row in det_csv_file.json():
 writer.writerow([row['Name'],row['Age'],row['Address'],row['Phone'],row['Email']])
print("CSV File Imported")
#CSV File Imported

#CSV File Read
with open('People.csv','r') as f:
 reader=csv.reader(f)
 for row in reader:
 print(row)
#['Name', 'Age', 'Address', 'Phone', 'Email']
#['John', '25', '123 Main St', '555-1234', 'john@example.com']
#['Jane', '30', '456 Elm St', '555-2345', 'jane@example.com']
#['Bob', '35', '789 Oak St', '555-3456', 'bob@example.com']

#CSV File Write
with open('People.csv','w') as f:
 writer=csv.writer(f)
 for row in people:
 writer.writerow([row['Name'],row['Age'],row['Address'],row['Phone'],row['Email']])
print("CSV File Written")
#CSV File Written

```

看起來還不錯——它自動給包含逗號的頭加上了引號。那麼這跟 Django 捕獲函數有什么關係呢？多亏 Django 的 `HttpResponse` 和 `StringIO` 對象一樣地是一種“类似文件”的對象，所以它們很相像——就是說它們都提供 CSV writer 需要的 write 方法。

```
>>> import csv, StringIO
>>> output = StringIO.StringIO()
>>> output.write("Name,Address,Phone\n")
>>> output.write("Bob,Dobbs,bob@example.edu\n")
>>> output.write("Pat,Patson,pat@example.org\n")
>>> output.write("O'Reilly,O'Reilly,oreilly@example.com\n")
>>> output.getvalue()
'Name,Address,Phone\nBob,Dobbs,bob@example.edu\nPat,Patson,pat@example.org\nO'Reilly,O'Reilly,oreilly@example.com'
```

很多程序员在一开始碰到CSV这类任务的时候都会尝试自己编写解析器和生成器。毕竟CSV这种格式太简单了，对吧？它就是一些逗号加一些数字，可能还有一点字符。要是这些字符里包括逗号的话可能还会用到引号。要是引号里还有引号的话可能还会用到转义符。开始有点复杂了吧。好在已经有了前辈们编写的模块来处理它。

假設現在我們要將人名數據從別人的VCards轉換到一個簡單的數據表里，其中每一列分別是名字、姓氏和email地址。首先，我們先在解釋器里試驗一下確保CSV模塊工作正常。因為CSV模塊主要是設計用來處理文件類的對象，所以這裏我們用Python方便的StringIO模塊來操作輸出。(StringIO為字符串提供了三個操作文件的接口)。

如果你的应用程序要输出表格化的数据，CSV (comma-separated value) 格式绝对是最佳选择。所有的编程语言都可以处理这种格式，而且几乎所有的能处理格式化数据的应用程序（Microsoft Excel, Filemaker Pro）都可以导入导出它。Python也提供了一个CSV模块来帮助你

### Comma-Separated Value (CSV)

Django的ORM查询从数据库里获取它们。根据应用程序的不同，你甚至可以考虑编写一个自定义的模板标签来专门从特定的QuerySet里获取数据，然后把它渲染到PNG图象里去。不过由于当然后，在实际使用视图的时候，我们不会把数据硬编码在视图函数里，而是会通过

图11.2就是生成的图片结果。

器进行下载就可以了。

所以视图函数最后返回的是一张PNG图片，我们只要设置好相应的包头并把它发送给浏览器回给响应。

渲染pycha.pie.PieChart类的函数来把chart画到surface上面，然后用chart的render函数。当绘图完成以后，再用surface的write\_to\_png方法把二进制PNG数据写到一个文件对象里并把它返

这里的关键是获取了一个surface对象（这是Cairo里最重要的绘制元素），并且通过把它传

```
def pie_chart(request):
 import sys, cairo, pycha.pie
 data = {
 ("To on Eyedrops", 61),
 ("Haskell on Hovercrafts", 276),
 ("Lua on Linuxed Oil", 99),
 ("Django", 1000),
 }
 data = [
 (item[0], [0, item[1]]) for item in data
]
 data = {
 "ticks": {"x": {"ticks": ticks}},
 "options": {
 "legend": {"hide": True},
 "axes": {"x": {"ticks": ticks}}
 },
 "chart": addDataset(datasets)
 }
 chart = pycha.pie.Piechart(surface, options)
 chart.render()
 response = HttpResponseRedirect(mimetype="image/png")
 surface.write_to_png(response)
 response["Content-Disposition"] = "attachment; filename=pie.png"
 return response
```

们就可以在响应里设置正确的mimetype并返回它。

在上面CSV的例子里，Django特有的一个重要技巧就是把输出放到一个字符串里，这样我

了两个优点：相对简单的Pythonic语法以及外观美好的默认输出显示。

PyCharm图形界面之上。PyCharm并没有试图容纳所有的输出格式或者配置选项，但是它具备

PyCharm (<http://www.lorenzogill.com/projects/pycha/>) 是Python里一个简单优雅的图表库，

## 用PyCharm图表

处理结构化数据的程序员都能读入CSV。

CSV对解决“能不能从Django获取数据导入到X应用里”这种问题是很方便的。几乎任何

虽然Django的ORM系统没有完全替代SQL的意思，不过对大多数Web应用来说它都绰绰有余了。而且你还可以通过額外的方法直接使用SQL命令来辅助Django ORM查询。不过，自己写Manager给了你提供了一种途径。可能你还不知道其实你已经在下面这样的代码里用到了manager：

```
really_good_posts = Post.objects.filter(gemlike=True)
```

这里Post.objects就是一个manager对象，它是一个models.Manager子类的实例。这个类的方法决定了你能对一个queryset进行的操作——比如这个例子里的过滤。

一个自定义manager也是一个manager对象。它主要有两个用途：修改以返回的数据集（一般是数据表里全部的对象）以及添加操作这个对象集的新方法。

如果不想一遍又一遍的写下前面的查詢代码，而是想要简洁一点的方法来告诉Django你只需要数据库里“gem-like”的帖子该怎么办？很简单，写一个像以下这样的manager就可以了：

```
class GemManager(models.Manager):
 def get_query_set(self):
 return super(GemManager, self).get_query_set().filter(gemlike=True)

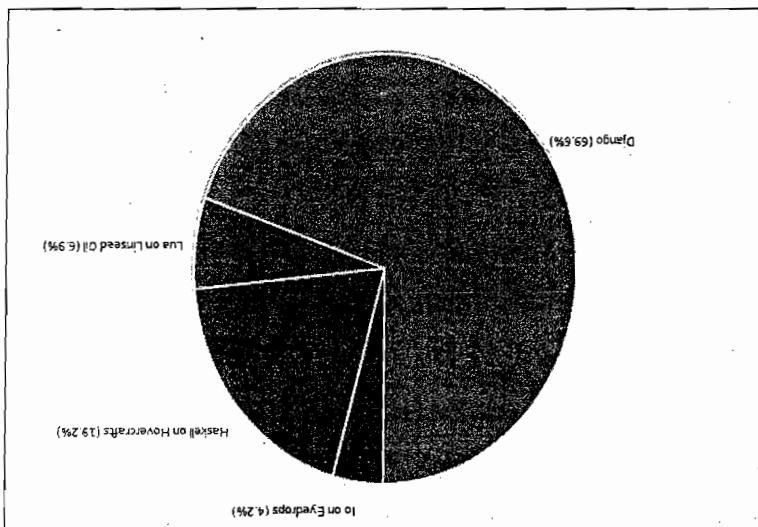
class Post(models.Model):
 title = CharField(max_length=100)
 """ My example Post class """
 class Post(models.Model):
 pass
```

## 修改默认对象集

这个类要怎么用呢？只需要把它赋值为你模型（model）里的一个属性就行了。

## 11.4 用自定义Manager来增强Django ORM

图11.2 Python饼图示例



在这里定义的objects manager和通常Django自动为我们创建的manager是一模一样的。

在这个模型里我们必须显式地定义它的原因在于当Django有额外manager出现时的行为——它看到的第一个manager会当成模型的默认manager，admin会在选择对象时使用它。如果我们

通过gems这个新的manager的代码之后，它的作用大概你就了然于心了。它通过Python的super函数去调用父类 (models.Manager) 的get\_query\_set方法，然后过滤结果，就像之前我们在自定义的manager里做的那样。

那要怎么用它呢？你可能已经猜到了——和默认的objects manager一样。

而且由于这个新的自定义manager返回的也是一个queryset，你还可以进一步对它进行过滤，除非，以及其他queryset方法的操作。

之前定义的自定义manager是一种很有用的语言糖。就是说它可以把你一些本来比较罗嗦的查询：变成更简洁的形式：

```
really_good_posts = Post.objects.filter(gemlike=True)
```

如果要处理很多这类对密集的话，明显能让我们代码更加易读，但是这还不是自定义manager的全部能力。通过为它添加自定义方法允许参数传递，我们能获得更大的灵活性。

继续刚才的例子，假设我们要简洁地指定一个queryset，它只包含了在标题和正文里提到特定单词的帖子。在一般的Django语法里，我们可以这么写：

```
cat_posts = Post.objects.filter(gemlike=True, title__contains='cats', content__contains='cats')
```

这个查询显得稍微长了一点，要是这种查询会重复使用的话，我们需要更简明的方式，比如：

```

img_src = os.path.join(settings.RANDOM_IMG_DIR, img_name)
img_name = random.choice(img_files)
img_files = os.listdir(settings.RANDOM_IMG_DIR)
def home_view(request):

```

假设你要在网站首页上显示一幅随机选择的图片。以你现在的知识水平，要做到这一点非常容易。你只需在视图代码里构建一个图片文件的列表，然后用Python的random.choice函数为你选中一幅，将它的值传给模板就可以了。这个视图函数看起来如下所示：

### 简单的自定义模板标签

首先，我们学习如何创建自定义模板标签过滤器，甚至用第三方的模板系统绕开Django的模板系统。这加快了模板处理的速度，减少了框架整体的复杂度，同时对非程序员来说是一个完整的编程接口。但是有时候你还需要或者想让它更强大一点。在这方面，你会学习到如何创建自定义模板标签过滤器，甚至用第三方的模板系统绕开Django的模板和其他系统的一个不同的地方在于它没有重新发明，或是封装一个完整的编程语言。这加快了模板处理的速度，减少了框架整体的复杂度，同时对非程序员来说是一个完整的编程接口。但是有时候你还需要或希望它更强大一点。在这一节里，你会学习到如何创建自定义模板标签过滤器，甚至用第三方的模板系统绕开Django的模板和其他系统的一个不同的地方在于它没有重新发明，或是封装一个完整的编程语言。这加快了模板处理的速度，减少了框架整体的复杂度，同时对非程序员来说是一个完整的编程接口。但是有时候你还需要或希望它更强大一点。

## 11.5 扩展模板系统

回顾起来，自定义manager并没有提供额外的功能。但是模型的API接口是干嘛的，项目使用它的其他代码就更容易编写和维护。

```

class AllAboutManager(models.Manager):
 def all_about(self, text):
 """Only return posts that are really good and all about X"""
 posts = super(AllAboutManager, self).get_query_set().filter(
 title__contains=text,
 gemlike=True,
 content__contains=text,
)
 return posts

```

以下是一个manager方法的代码：

创建类似的方法和方法时最有意思的就是之一个大概就是给他们起一个好名字。要找到一个能在Post.objects.filter这样的形式里读起来顺口的名字可不容易，毕竟你（或任何会阅读代码的人）都会经常看到它。一般来说，manager应该是一个名词（“objects”，“gems”），而manager方法则应该是动词（“exclude”，“filter”）或是形容词（“all”，“latest”，“blessed”）。

注意

第一个值得注意的就是 `register = template Library()` 这一行。template Library 实例化并为我们提供一些特定的装饰器，它们可以把函数转换成模板函数从而使用 std::function。虽然这个

```

register = template.Library()

def files(path, types=[".jpg", ".jpeg", ".png", ".gif"]):
 fullpath = os.path.join(settings.MEDIA_ROOT, path)
 return [f for f in os.listdir(fullpath) if os.path.splitext(f)[1] in types]

def register.simple_tag(register):
 def random_image(path):
 pick = random.choice(files(path))
 return mark_safe(os.path.join(settings.MEDIA_URL, pick))
 return mark_safe(register.inclusion_tag(random_image))

```

以下是在义标签名的完整代码。和往常一样，我们可以导入任何需要的模块。这里大部分代码都是纯Python，但我们要介绍Django相关的部分。

标签的名字是random\_image，跟在后面加引号的字符串是路劲名。这是相对于MEDIA\_ROOT的相对路劲。Django模板系统会解析参数然后把它传递给你的函数（模板标签

```
 >
```

应该能够做到以下这样的事：这类标签的代码相当简单，所以我们先来看实现，然后再说回来源码细节。这个标签最后已的“专用”版本。

这里的关键是—种叫做自定义模板标签 (custom template tag) 的特性。Django 用来自定义自己的模板标签的机制，作为程序员的你当然也可以一样使用。也就是说你可以为你的设计师创造一个简单的标签。更棒的是，标签还可以接受参数，所以你的设计师还能拥有他/她自己

如果只是这样的话那一切都能够顺利。但假假设你决定要把这个随机图片加到由另一个随机抽取的页面里去。或者你的设计师说说道，“嘿，我有五个要用到这个随机图片功能的地方，其中三个要从不同的目录里读入图片，你有没有办法呀？”当然，答案是没问题！

```

```

最后，在模板里用一个图片标签就能读取传递进来的值了。

settings.py文件里，这样项目里所有的应用都可以访问它，修改起来很方便。)

(这段代码遵循了Django保存配置信息的推荐做法，例如把RANDOM\_IMG\_DIR放在

```
render_to_response("home.html", {"img_src": img_src})
```

# ... other view processing goes here ...

```

```

random\_image标签的使用看起米就如下所示：

在视图里动态指定期取随机图片的目录，然后把它作为image\_dir传递给模板。这样，random\_image标签既可以接受一个字符串，也可以接受一个模板变量作为其参数。所以你可以把标签嵌入模板后，它定义的标签就可以和Django的内置标签一样使用了。你的新标签只要一个模板变量（不用“py”）。

```
(%load randomizers %)
```

要在给定模板里使用这个新标签，你需要在模板顶部使用模板系统编的load标签来加载它。

完成之后，你的新tag就可以用在项目里的任何应用模型里了。

要复习一下为什么Python需要这个文件的话，请参考第1章里的模块和包一节。（要是你需要复习的是你看到这样的错误，别忘了到你的templatetags目录下创建该文件。）

```
from django.template import Library
register = Library()
@register.tag
def randomize(parser, token):
 try:
 name, args = token.split_contents()
 except ValueError:
 raise TemplateSyntaxError("'%s' tag takes one argument" % token.contents[0])
 return RandomizeNode(name, args)
```

这个templatetags目录必须是一个Python的包，也就是说在目录下面要放一个\_init\_.py文件（空的就行）。要是你忘记创建这个\_init\_.py文件的话，你会得到一个对这个问题没什么帮助的错误信息。

所以这个文件需要放在一个叫做templatetags的目录里，并且这个目录要在模板系统里的搜索路径上。那就是在INSTALLED\_APPS里（如果你在settings.TEMPLATE\_LOADERS里有列出TEMPLATE\_DIRS设置里列出的目录之一（如果你在settings.TEMPLATE\_LOADERS里有列出django.template.loaders.filesystem.Loader类的语句），要么是在django.template.loaders.app\_directories.Loader类的语句），并且这个目录的搜索路径上。如果是在这段代码里有什么神奇的地方，那一定就是在random\_image函数上的@register。

虽然这里我们只定义了一个标签，但实际上我们创建的这个文件算是一个能包含很多标签的simple\_tag装饰器了。正是这个装饰器把我们的简单函数转换成模板引擎可以使用的标签。

如果想让其他模板内容标签的话，就可以用它randomizers.py。

虽然这里我们在文件的时候记得给它起个好名字，比如将来我们会在这个库里添加的Django标签。所以在保存文件的时候记得给它起个好名字，比如将来我们会在这个库里添加的Django标签。

为零。在calendar.monthcalendar上用列表推导式就能得到我们想要的结果了。  
现在我们要把列表里的0都变成空白——因为我们不想让月份开始和结束前后的日子显示出来。

```
[0, 0, 1, 2, 3, 4], [5, 6, 7, 8, 9, 10, 11], [12, 13, 14, 15, ...
>>> calendar.monthcalendar(2010, 7)
>>> import calendar
```

一个包含了星期数字列表的列表返回给我们（在月份开始和结束前后的星期数字都用0填充）。  
现在我们来编写这个标签。这个标签是基于Python的calendar模块，这个模块会把月份作为为  
月份编译出一个简单的日历。  
假设你的模板里有一个包含了当前日期的{{date}}变量，然后你想用一个模板标签来为当前的  
Inclusion标签在你要用当前模板context里的一些值来渲染一块内容时是最有用的。例如，  
Inclusion标签。

虽然你可以写一个simple-tag来使用模板引擎，不过Django提供了更方便灵活的方法：  
同样，你也应该尽量避免在模板标签函数里硬编码HTML。  
MVC原则（见第3章）的意思就是将HTML放在模板里，而不应该放在视图函数里，起来挺不错。  
(比如一段动态的HTML)，千万不要用模板标签函数来构建和返回该HTML，即使这个做法看  
在上面的例子中标签返回的是简单的字符串。如果你希望自定义标签返回更复杂的内容  
Inclusion标签。

对于希望给模板的作者（也包括你自己）提供一个简洁可读的方式来自动生成原本需要在多个  
视图函数里通过自写代码才能生成的值的情况，简单的自定义模板标签就是最适合的办法。如  
果你还需要更复杂的办法，请往下读。

```

 <% for month in months %>
 <tr>
 <% for day in week %>
 <td>{{ day }}</td>
 <% endfor %>
 </tr>
 <% endfor %>
</table>
```

我们来为这个月历创建一个小型的模板，这里假设月份是按星期传入的完整列表；在现代CSS布局里，月份还是属于表格外数据，所以我们要用HTML的表格标签来生成它。

```
... , 'January', 'February', 'March', 'April', 'May', 'June', 'July', ...
>>> list(calender.month_name)
['Mon', 'Tue', 'Wed', 'Thu', 'Fri', 'Sat', 'Sun']
>>> list(calender.day_abbr)
['...', '...', '1', '2', '3', '4', '[5, 6, 7, 8, 9, 10, 11]', '[12, 13, 14, ...]
```

calender模块甚至还为我们提供了日期和月份的名字。

数据就不能直接用了。

由于Django的模板引擎并不关心我们传递的是整数还是字符串，所以这里混杂的数据类型不会给我们造成困扰。不过要是我们把这个数据传给Python函数做进一步处理的话，这种混合

如果这一天非零，那么就直接使用它，否则就使用空字符串。

这个巧妙的列表推导式的意思是遍历月份里的每个星期，对每个星期遍历其中的每一天，

```
<... , ... , 1, 2, 3, 4], [5, 6, 7, 8, 9, 10, 11], [12, 13, 14, ...
>>> [[day or '' for day in week] for week in months]
>>> months = calender.monthcalendar(2010, 7)
```

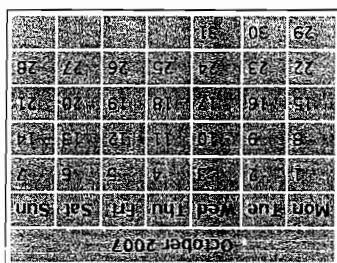
过滤器就是函数——在大多数情况下接受字符串并返回字符串的函数。一个不太复杂又很方便的过滤器语法。

Django的模板系统自带了很多非常有用的过滤器（filter），但是有时候你会需要添加自己的过滤器。过滤器很容易使用，实际上编写起来也不难。和其他你看到的标签一样，它用到的过滤器。

## 自定义过滤器

显示效果看起来还不错。由于商业逻辑和表现层区分的很清楚，你的设计师可以轻易的调整标签的模板文件。同时，你和其他程序员可以对内容作出修改，比如本地化月份和日期的名字，而根本不用去碰模板（或者是它包含的任何其他模板）。

图11.3 我们的日历



这个日历看起来应该如图11.3所示（根据浏览器的不同可能会影响到%calendar-table%）。

当你要使用这个标签的时候，只需要在模板最开始的地方插入{%load inclusion\_tags%}即可。具体调用的时候，在我们希望日历出现的地方插入{%calendar-table%}就行了。

```
<style type="text/css">
 .calendar {
 padding: 10px;
 border: 1px solid black;
 width: 300px;
 margin: auto;
 }
 .calendar td {
 padding: 5px;
 width: 40px;
 height: 40px;
 background-color: #f0f0f0;
 text-align: center;
 font-size: 1.2em;
 }
 .calendar td:hover {
 background-color: #e0e0e0;
 }
 .calendar th {
 padding: 5px;
 width: 40px;
 height: 40px;
 background-color: #f0f0f0;
 text-align: center;
 font-weight: bold;
 font-size: 1.2em;
 }
 .calendar th:hover {
 background-color: #e0e0e0;
 }
 .calendar tr {
 border-top: 1px solid black;
 }
 .calendar tr:first-child {
 border-bottom: 1px solid black;
 }
 .calendar tr td {
 border-right: 1px solid black;
 }
 .calendar tr td:last-child {
 border-right: none;
 }
 .calendar th {
 border-right: 1px solid black;
 }
 .calendar th:last-child {
 border-right: none;
 }
</style>
```

calendar.html文件的顶部。

现在我们给页面加一个简易的样式来稍微装点一下这个日历。将下面的代码放到的地方。

inclusion标签让你既能编写简洁的代码，同时又能把模板的内容保持在模板内部即它本该所在的地方。或者也可以创建一个标签，让它自己去调用模板引擎，不过这样又未免太繁琐。

除了这里说到的内容，inclusion标签并没有其他特别之处。它们用来帮助你方便地将表现

@register.inclusion\_tag来调用。换句话说，字典里所有的键都变成模板里的变量，可以用来显示我们的函数返回的是一个字典，它会在渲染模板的时候作为context字典提供给

```
def my_inclusion_tag(context):
 year = context['year']
 weeks = context['weeks']
 daynames = context['daynames']
 dayabbr = context['day_abbr']
```

这个过滤器的定义是：

```

{{ my_string|grep:"magic" }}

def grep(text, term):
 if text.contains(term):
 return text
 else:
 return text[0:term.length] + "..." + text[term.length:]

def hide_if_shorter_than(text, min_len):
 if len(text) >= int(min_len):
 return text
 else:
 return text[0:min_len] + "..."

register.filter(hide_if_shorter_than="100")

```

本大于一定长度时才打印的话，它应该是如下这样使用的。

过滤器的参数总是用引号括起来并用冒号和过滤器名分开。即使你用的是数字或其他非字符串类型，Django的模板语法也要求使用引号。所以假使你写了一个过滤器让它只存在输入文件中，Django的模板语法也需要使用引号。所以这里用到的是int函数。

过滤器当然可以接受第二个参数了。通常该参数可以用来调整过滤器的行为。例如，假设想让你一个显示字符串的函数只在字符串包含特定的字符串时才显示它（想象一下Unix命令grep）。当然用if/then模板标记也能做到，不过过滤器的方式更加简洁。

过滤器的行为的话要怎么样呢？

上面这个wikify函数接受一个参数，它的值是模板里出现在过滤器前面的 | (管道) 符号左边表达式的结果。不过要是你需要过滤器的用户还可以通过传递一个额外的参数来调整过滤器的行为的话要怎么样呢？

这样的链接。它可以作用在以下这样的模板里（假设变量title和content分别保存了页面的标题和这个标签接受一个字符串，字符串里所有驼峰状的词组都被替换成了 /wiki/CamelCaseWord/

```

@register.filter
def wikify(text):
 wikifilter = re.compile(r'([A-Z]+[a-z]+)([2,])\b')
 text = wikifilter.sub(r'\1', text)
 return wikifilter.sub(r'\1', text)

```

除非 (CamelCase) 的词组转换或适合HTML使用的wikify过滤器。如下所示：

有趣的例子（至少不是太长，要是你不喜欢正则表达式的话可能还看有点晕）就是一个将驼峰状 (CamelCase) 的词组转换或适合HTML使用的wikify过滤器。如下所示：

DjangоМ从一个高质集或的堆叠模型上获得了很多好处。但是它并不是铁板一块——只要你

选择一种替代的模板机制

无需任何第三方模块——这个视图只用到了Python内建的字符串模板语法。

```
return HttpResponse(template.format(request, name))
template = "Hello, your IP address is %s."
def simple_template_view(request, name):
```

以下是一个使用Django模板机制的最简单的例子。

纯文本

非难事。

Django里模板引擎的工作就是准备字符串来作为HttpResponse对象的内容。当你了解了这一点后，要是基于某些原因你决定不使用Django提供的模板语言，而使用另一种模板引擎也并非难事。

## 替换模板

com/documentation/templates-python。

了解更多关于高级自定义模板标签的细节，可以参考Django的在线文档 [www.djangoproject.com/](http://www.djangoproject.com/)。引擎的内部机制。这比编写前面那些简单装饰器的工作量要大多了，而且也不是很常用。想要标签来处理它们所包含的内容。创建这种类型的标签相当复杂，需要直接去操作Django模板引擎来实现。

你还可以定义更复杂的自定义模板标签——例如，用配对的模块风格 (paired block-style)

## 更复杂的自定义模板标签

@stringvalue放在@register.filter前面的话，就只能保证过滤器的输出是字符串了。这里的

@stringvalue效果在@register.filter前面的话，就只能保证过滤器的输出是字符串了。这里的

@stringvalue，最后是函数本身) 让过滤器接受的输入是一个字符串。要是倒过来把

对函数应用装饰器的顺序是有讲究的。我们刚才提到的顺序 (@register.filter，接着是

注意

@stringvalue就行了。

如果你知道过滤器函数接受非字符串的数据，但又希望把它当作字符串来处理的话，你可以给函数加上一个stringvalue装饰器。一个函数可以同时接受多个装饰器，所以如果我们将这个装饰器加到 hide\_if\_shorter\_than函数上的话，我们只需在@register.filter下加上一行

的Web开发是在和字符串打交道，从而让这显得有点不太寻常，不过这的确是可能的。

Python的datetime对象。所以你完全可以在编写一个能处理非字符串对象的过滤器。虽然绝大部分

字符串是字符串。你看一下日期和时间相关的过滤器就知道此言不虚了，它们都是直接操作

字符串有点费解，不过传递给过滤器函数的第一个参数 (实际被“过滤”或修改的值) 不一

基于上述原因，在这里我们显示地将递进来的字符串用min\_len参数转换成整数。

```
return text
```

畢竟，絕大多數情況下你都可以把Django的一部分替換成第三方组件。

特別是替換模版系統非常方便。那為什麼要一個不同的模版系統呢？答案有以下几点：

- 你是從另一個系統轉到Django上來的，那個模版的語法用着比較順手。
- 你還運行了用其他Python Web框架編寫的項目，有一個通用的模版語言會比較好。
- 你要把一個項目从其他Python Web框架上移植過來，而且沒時間轉換現有的模版了。
- 你的表現邏輯需要一些不容易添加到Django的模版語言里的特性。

使用另一種模版引擎：Mako

在這一章里我們要介紹一個非常流行的第三方模版引擎：Mako。它的樣子和設計都和Django的模版語言非常不同，但是它們具有一些相同的优点：速度快，不绑定XML，拥有相似的繼承机制。

Mako 取代了Python的模版框架Myghty，而Myghty則又是基於另一個影响深远的Perl系統HTML::Mason。Mako已經被用在了像 reddit.com 和 python.org這樣的網站上，並且同時還是另一個Python Web框架Python的默認模版引擎。所以要是你想試試另一種模版系統的話，Mako是一個很不錯的起點。它的影響力也波及到了Django的模版系統，所以虽然語法上它們有些許區別，但是概念上兩個系統还是有重合之處的。

因為后者努力地在其模版里限制編程逻辑的數量。這兩種方式各有千秋。Mako的理念是為了尽量方便Python程序员，而且Python清晰的語法也不會阻擋模版設計師的腳步。

在用Mako編寫Django視圖之前，以下提供一個簡單的例子，看看Mako在解釋器中是怎樣工作的。

```
>>> t = Template("My favorite ice cream is ${name} | ${entity}")
>>> t.render(name="Emack & Bollo's")
"My favorite ice cream is Emack & Bollo's"
>>> t = Template("My favorite ice cream is ${name} | ${entity}")
>>> t.render(context={name:"Emack & Bollo's"})
"My favorite ice cream is Emack & Bollo's"
```

在第一個例子裏，我們显式地传递了变量的名字，而在第二個例子中，我們用的則是默認的context。回憶一下，context就是一個传递给render方法的字典而已。所以這個簡單的例子看起來應該也是很明顯的，因為它和Django模版引擎里的工作方式几乎一模一樣。

Mako還有和Django很相似的过滤语法。

現在我們來把Django視圖模版放到Mako模板上去。

```
from mako.template import Template
```

在介紹的時候我們就說到，這一章的內容擴闊了多個主題，我們希望讓Django框架開啟一下眼界，並且讓你們体会到這個框架真正的灵活性和可擴展性。下一章將會通過介紹另一部分高級話題來給本書的第四部分畫上一個完滿的句號。

## 11.6 總結

當然如果你真的打算使用Mako的話，你需要把模板存放在文件系統（或是數據庫）上，讓Django可以像查找到自己的模板那樣方便地找到它們（即無須指定完整路徑），然後創建一個能理解Mako的render\_to\_response方法等。所幸這些工作已經有其他Django/Mako的先驅為你做好了。Django Snippets (<http://www.djangosnippets.org/snippets/97/>) 上有一些很有用的代碼，頁面上還有一些相應的blog帖子能指導你使用它們。

```
def mako_view(request):
 t = Template("Your IP address is ${REMOTE_ADDR}")
 output = t.render(**request.META)
 return HttpResponse(output)
```

這個視圖里完成的功能和前面在交互解釋器里的一樣的——創建一個新的Mako模板，然後用來自請求的META對象里的context來渲染它，這個META對象我們在第5章里曾經見過。

```
def vacuum_db():
```

```
 import os
 import sys
 os.environ['DJANGO_SETTINGS_MODULE'] = "dpaste.settings"
 from django.conf import settings
```

它每天晚上都会由cron控制执行。（如果是Windows系统，你可以把它设置成“服务”的形式。）文件达到4GB大小，但是我们宁可还是不要测试这个底限。下面就是dpaste.com清理脚本的示例。如果不定期清理的话，数据库很快就会变得臃肿无比。虽然SQLite号称能够支持的数据量在有大量操作（删除旧记录和添加记录）的SQLite（和一些PostgreSQL）数据库里，定期清理回收闲置空间是很有用的。拿dpaste.com为例，大多数数据都只在数据库里停留一个月后就被删除了。这就是说每星期大概有25%的淘汰率。

### 负责打扫清理的Cronjobs

在这方面Python是非常有价值的。你只需要在编写Django工具脚本的Python代码时配上一点Django所需要的环境设置就可以了。下面是一些Django工具脚本的例子。我们会逐个解释它们的代码，你可以自行决定哪一个是最适合你的项目。

- 执行清理任务（例如，删除过时的会话）

- 生成报告

- 发送定期的E-mail提醒

- 将数据导入到Django模型里

- 为你的每日构建（或每小时）创建缓存数据或文件

这些常见的Django工具脚本用例包括：

但是却没有为之创建一个完整的Web界面的需要。

用Python而不是其他Web专用的语言（比如ColdFusion或PHP）来编写Django的优点之一就在于它可以命令行环境里使用。比如你要定期或是偶尔对Django应用里的数据进行一些操作，虽然Django是一个Web框架，但是这不代表你就不能在浏览器之外与之交互。事实上，选

## 12.1 编写工具脚本

和第11章一样，这一章也是由几个相互没有关系的小节组成，分别讨论了不同的主题。第11章讨论的主要还是和应用程序代码有关的话题。而在这一章里，我们要讨论的是一些更加正式的话题，包括部署应用，更新它们运行的环境，以及修改Django本身。

## 第12章 高级Django部署

这个脚本假设DjangоДjango本身以及你项目的父目录都在Python路径里。它们可以从site-packages符号链接过来，安装为Python eggs，或者包含在PYTHONPATH环境变量里。要在脚本中手动设置它们的话，只要在最初的两个import语句之后加上以下两行代码：

```
sys.path.append('/your/django/coderebase')
sys.path.append('/your/django/projects')
```

就可以替换你的路径了——第一行指向的是你系统上Django源码的目录（就像我们这种直译从Subversion里检出运行Django的集成分支）。第二行则是把我们的项目目录加到sys.path里去，这样就可以通过import语句来访问它们了。

记住编写Django工具脚本其实就是编写Python脚本。只要告诉Python知道怎么找到Django和你的项目，以及怎么让Django找到你的设置文件就一切OK了。

```

from django import import connection
cursor = connection.cursor()
cursor.execute("VACUUM")
connection.close()

if __name__ == "__main__":
 print "Vacuuming database..."
 before = os.stat(settings.DATABASE_NAME).st_size
 print "Size before: %s bytes" % before
 print "Vacuuming database..."
 cursor = connection.cursor()
 cursor.execute("VACUUM")
 connection.close()

 after = os.stat(settings.DATABASE_NAME).st_size
 print "Size after: %s bytes" % after
 print "Print 'Recalimed: %s bytes' % (before - after)"

```

要么花费掉你大量的精力，要么还错很多（更惨的是两者皆有）。

下面这个简单的模型可以用来存储email消息——实际上它和portal.com用来处理垃圾邮件

件存储的模型非常类似。

```
class Message(models.Model):
 subject = models.CharField(max_length=250)
 body = models.TextField()
 date = models.DateTimeField()

 def __str__(self):
 return self.subject[:50]
```

假设你手上有了这样一个模块，并且在项目的settings.MAILBOX里也设置好了mbox文件

的路径，你就可以像以下这样把附件导入到模型里去：

```
try:
 import os, mailbox, email, datetime
 from email.utils import parsedate # Python >= 2.5
 except ImportError:
 from email.Utils import parsedate # Python < 2.5
 os.environ['DJANGO_SETTINGS_MODULE'] = 'YOURPROJECT.settings'
 from django.conf import settings
 from yourappp.models import Message
 from yourapp.settings import MAILBOX
 mbox = open(SETTINGS.MAILBOX, 'rb')
 for message in mailbox.PortableInbox(mbox, email.message_from_file):
 date = datetime.datetime.strptime(message['date'], '%d %b %Y %H:%M:%S %Y')
 msg = Message()
 msg.date = date
 msg.subject = message['subject']
 msg.body = message.get_payload(decode=False)
 msg.save()
```

Django项目正处在开发活跃的阶段。不过随时和Django的主人或“trunk”版本保持一致能够帮助你进入Django内部“修正”问题的原因是很可能根本不需要费那个劲。我们不修改Django内部代码是最后不得不才会考虑的手段。这不是因为它很难（Python怎么会理解你的需求还有那么干净的codebase，以及大量docstring和注释形式的内部文档支持）。我们不

本上是安全的，因为它还是很稳定的。随着新特性的不断添加和旧有bug的不断修正，你可以通过随时和Django的主人或“trunk”版本保持一致能够帮助你进入Django内部“修正”问题的原因是很可能根本不需要费那个劲。

## 12.2 自定义Django codebase

区间浏览一下Python的“stdlib”(<http://docs.python.org/lib/>)，至少也要知道它大概提供了什么功能。这只是一个展示如何编写Django项目的脚本的小例子。Python的标准库（假设这里没有第三方的类库）已经相当的完备了。要是你真的打算成为一名合格的Django程序员，你一定要花点时间浏览一下Python的“stdlib”(<http://docs.python.org/lib/>)，至少也要知道它大概提供了什么功能。

```
open(MAILBOX, "w").write("")
Depending on your application, you might clear the mbox now:
print "Archive now contains %s messages" % len(message.objects.all().count())
msg.save()
(
```

```
body=message.get_payload(decode=False),
date=date,
subject=message['subject'],
msg = Message(
date = datetime.datetime.strptime(message['date'], '%d %b %Y %H:%M:%S %Y'),
mbox = open(settings.MAILBOX, 'rb')
for message in mailbox.PortableInbox(mbox, email.message_from_file):
 date = datetime.datetime.strptime(message['date'], '%d %b %Y %H:%M:%S %Y')
```

```
from yourappp.models import Message
from django.conf import settings
os.environ['DJANGO_SETTINGS_MODULE'] = 'YOURPROJECT.settings'
```

```
try:
 import os, mailbox, email, datetime
 from email.utils import parsedate # Python >= 2.5
 except ImportError:
 from email.Utils import parsedate # Python < 2.5
```

```
os.environ['DJANGO_SETTINGS_MODULE'] = 'YOURPROJECT.settings'
from django.conf import settings
```

```
date = datetime.datetime.strptime(message['date'], '%d %b %Y %H:%M:%S %Y')
```

```
subject = message['subject'][:50]
```

```
body = message.get_payload(decode=False)
```

```
msg = Message(date=date, subject=subject, body=body)
```

```
msg.save()
```

```
$ ab -n 1000 http://127.0.0.1:8000/blog/
```

怎么也比它要快吧!

不要太在意这个例子里的绝对数字，因为它们是运行在一台三年前的笔记本上——你的服务器里是一个在第2章里的blog应用实例上运行ab的输出。这里说的直白一点，就是“每秒的请求数”。它工作的方式是你提供一个URL和要提交的请求数量，它把最后的性能统计返回给你。这（关于如何使用的消息，请参见其手册 <http://httpd.apache.org/docs/2.2/programs/ab.html>）。

Apache的安装目录下找到，例如，C:\Program Files\Apache Software Foundation\Apache2.2\bin

OS X），它应该可以直接受到（已经包含在Path路径里了）。在Windows系统上，它可以在Linux或者Mac

安装了Apache的话，那么ab也随之一起安装好了。在任何POSIX系的系统上（Linux或者Mac

ab是一个基础服务器测试的常用工具，它的全称是Apache Benchmark tool。如果你已

经不一样，所以要知道网站缓存究竟起了多大的作用，只有亲手去测试一下。

缓存的全部意义就在于提升网站的性能，所以在事前做一点点评估是很必要的。每个站点

画一条底线

里需要关心的就是有多少了。

Django的缓存框架对初学者来说就像是一堆莫名其妙的配置。虽然每个网站的需求（以及每台服务器的能力）都一样，不过要是我们用一个实际的例子来开头的话，你就能更好地掌握这个工具。而且，这个配置还能够满足大部分网站的需要，所以说不定你在Django缓存

每台服务器的能力）都一样，不过要是我们用一个实际的例子来开头的话，你就能更好地掌握这个工具。而且，这个配置还能够满足大部分网站的需要，所以说不定你在Django缓存

基本的缓存方法

，允许你特别指明渲染页面里的哪一部分应该缓存。

对大量网站来说，不管后台使用的是什么技术，缓存都是必不可少的技术。Django可以对每次都会执行相对昂贵的数据库操作或是计算步骤了。

每次都会执行相对昂贵的数据库操作或是计算步骤了。对每次都会执行相对昂贵的数据库操作或是计算步骤了。这样就不用每次都去执行同样的操作就是使用缓存——即保存生成数据的一份拷贝，这样就不

大流量的网站很少因为Web服务器发送数据的性能而受到限制。这里的瓶颈几乎总是和数

## 12.3 缓存

里向Django项目贡献代码的部分。

最后，要是你实在忍不住要hack一下Django的话，不妨考虑一下这些为自己所做的修改是否是你可以接受的其他用户也受益呢？如果你觉得可行的话，请一定要读一读附录F的一点忙（请参见附录C）。

紧跟 [code.djangoproject.com](http://code.djangoproject.com/) 上的报告并在适合的时候进行升级。但要是你自己弄了一个山寨版的话，就算是自己排除在主干的更新之外了。至少你要花费大量的精力才能把更新和你自己的修改合并起来。如果你有非这么做不可的理由的话，分布式版本控制系统或许能帮上自己的修改合并起来。如果是你自己排除在主干的更新之外了。至少你要花费大量的精力才能把更新和你

将会更显著。

settings.py里简单的两行代码就提升了几乎三倍的速度。记住这还是在这个blog应用非常简单的情况下，无论是数据库查询还是商业逻辑都很轻巧，对于更复杂的应用来说，性能的提升会更加明显。

```
Requests per second: 114.29 (#/sec) (mean)
...
Time taken for tests: 8.750 seconds
...
Finished 1000 requests
...
Benchmarking 127.0.0.1 (be patient)
...
$ ab -n 1000 http://127.0.0.1:8000/blog/
```

这就在Django里打开了基本的全场缓存。现在我们来看看带缓存的新网站表现如何。

### 体验一下

(这种独特的伪URL的设置风格在你看过其他类型后台后就会比较有体会了，它们都使用了这种URL格式来封装配置参数。由于这是默认的后台，严格说起来除非我们要用其他的后台否则不需要设置它。不过Python的专家们说过，“不要依赖默认行为，显式定义你要做的事情”，所以在这里放一个设置的话，以后要修改或添加配置参数都会方便一点。)

```
CACHE_MIDDLEWARE = "locmem://"
```

在settings.py里加入下面一行：

Livejournal.com的超高性能缓存系统)。

Django默认的缓存在后台，一个叫做locmem的本地内存缓存。它把缓存数据放在内存里，这样存放的数据就非常快。虽然很多缓存方案都选择把缓存放在磁盘上，不过内存型缓存的性能要好得多。(如果你有疑惑，请参见下面关于Memcached的讨论，这是一个原本设计用来支撑的多。

### 设置缓存类型

Django的缓存特性是通过一个默认关闭的中间件来实现的。打开settings.py文件，在MIDDLEWARE\_CLASSES设置里加上django.middleware.cache.CacheMiddleware。一般你要把它们放在最后一行上，因为有些特定的中间件(特别是SessionMiddleware和GZipMiddleware)可能会需要根据缓存框架来修改HTTP的巨头。

### 加入中间件

差不多是每秒36个请求。现在我们打开缓存看看有什么不同。

```
Requests per second: 36.07 (#/sec) (mean)
...
Time taken for tests: 27.724 seconds
...
Finished 1000 requests
...

```

```

... retrieve stories and return HttpResponse
def top_stories_yesterday(request):
 cache_page(24 * 60 * 60)

from django.views.decorators.cache import cache_page

```

假设这两个列表分别用两个不同的视图生成，那么只要应用一个装饰器就能为它们打开缓存。

假设你站点里的每一个部分都应该被缓存在相同的时间。不过你可能有别的打算。比如，你运行的新站点要跟踪每个故事的流行度，然后把这些统计收集起来生成一个最受欢迎的故事列表。虽然“昨天最受欢迎的故事”列表可以缓存长达24小时。而“今天最受欢迎的故事”则会不停地发生变化。要在保持内容更新和服务器负载合理之间寻找一个平衡点，我们可能希望这个页面只保持五分钟。

如果Django的页面能符合你的要求，那么上面这些信息基本上就是你全部需要了解的内容了。不过它并不适用于所有情况，我们来看看Django提供的粒度更细的缓存方式及其适用的场景。

如果Django的页面不能符合你的要求，那么上面这些信息基本上就是你全部需要了解的页面上。

设置`CACHE_MIDDLEWARE_ANONYMOUS_ONLY`（`settings.py`）能立刻反应在admin站点的[站点](#)里并不是总能发挥作用，因为一个给定URL的内容会随时变化以响应用户的输入。即使你的站点不涉及用户创造的内容（比如使用Django的admin应用），你也会希望把这

个值设置为True，以保证你做的修改（添加，删除，编辑）能立刻反应在admin站点的[站点](#)里不会发生键值冲突。你可以在这些设置里使用任何唯一的字符串——站点要在几个Django站点之间共享的话，不管是内存，文件，还是数据库里，这都能保证缓存内应该被用到。默认值是60秒（十分钟）。

`CACHE_MIDDLEWARE_SECONDS`：一个被缓存的页面在被一份新的请求替换前多少秒内应该被用到。最简单的用法我们已经看过了，以下是一些`settings.py`里的设置它们可以帮助你调试。

前面我们使用的是全站式的缓存特性。Django会缓存所有带有GET或POST参数的请求结果。最简单的用法我们已经看过了，以下是一些`settings.py`里的设置它们可以帮助你调试。

## 缓存策略

Django 还允许你控制HTTP头。通常，正是用URL作为键来缓存的。但是你可以用一些其他参数来影响一个特定URL返回的内容——例如，已登录的用户可以看到和匿名用户不一样的页面，或者可以依据用户的浏览器类型或语言设置来自定义响应。所有这些参数经由一些其他的参数来影响一个特定URL返回的内容——例如，已登录的用户可以看到和匿名用户不同的页面。

比没有好，免得自己动手修改HttpResponse对象的头了。

大多数网站都用不到cache\_control装饰器这么细致的控制功能。但是在需要的时候，有总

```
...
def revvalidate_must_revalidate(request):
 cache_control(must_revalidate=True)

from django.views.decorators.cache import cache_control
...
```

它们缓存的内容还未过期)，你可以如下装饰你的视图函数：

例如，要想迫使客户端缓存必须“重新验证”你的页面（检查它是否被改动过，即使 `s_maxage`）。

`Web客户端缓存`(downstream cache)。你可以设置它的任意六个开关 (`public`, `private`, `no_cache`, `no_transform`, `must_revalidate`, `proxy_revalidate`) 和两个整数值 (`max_age`,

`cache_control`: `djangoviews.decorators.cache.cache_control`。

以上的代码告诉所有的接收者不要缓存该页面。并且只要它们遵守标准 (RFC2616)，页面

```
...
def top_stories_this_second(request):
 never_cache

... we don't want anybody caching this
```

Django 给予你的最基本的额外缓存控制形式就是一个“永不缓存”(never cache) 装饰器。

在实际使用中，缓存是一个在你的服务器和连接到它上面的客户端之间的对话 (包括你可能不能控制的外部缓存服务器)。这个对话由特定的包头控制，即随着HTTP响应传递的“缓存

控制”(cache-control) 头。在论坛中，我们主要关注的是怎么设置服务器来决定缓存内容每次重新生成一次。

缓存控制相关的包头

URL为键以字典形式保存起来。然后用后一个对象给HTTP响应设置和缓存相关的包头。

HttpResponse对象的特点，用前一个对象来确定被请求的是哪一个URL，被缓存的数据则用

per-view装饰器利用了所有的Django视图都接受一个HttpRequest对象并返回一个

别的了。

`cache_page`装饰器只接受一个参数，即这个页面要缓存的秒数。除此之外，你就不再管

```
...
def top_stories_today(request):
 cache_page(5 * 60)

... retrieve stories and return HttpResponseRedirect
```

Django的缓存机制在缓存时要连同包头一起考虑。例如，要根据请求的Accept-Language头，从同一个URL返回不同的内容，你可以告知到返回的内容。

```
...
def localized_view(request):
 vary_on_headers('Accept-Language')
 from django.views.decorators import vary_on_header
 vary_on_header('Content-Language')

 # ...
 def __call__(self, request, *args, **kwargs):
 response = super().__call__(request, *args, **kwargs)
 response['Content-Language'] = self._get_language(request)
 return response
```

都是通过请求里的HTTP头来通知服务器的。“改变”响应头就能够明确定义哪些参数会影响到的内容。

因为“Cookie”头的變化也是很常见的一种情况，所以Django还提供了一个vary\_on\_cookie装饰器方便你使用。

```
...
def cookie_vary_on_headers(headers='Accept-Language'):
 from django.views.decorators import vary_on_header
 vary_on_header(headers)

 # ...
 def __call__(self, request, *args, **kwargs):
 response = super().__call__(request, *args, **kwargs)
 response['Set-Cookie'] = self._get_cookie(request)
 return response
```

Django的缓存机制在缓存时要连同包头一起考虑。

前面的缓存类型都是专注于缓存页面——若是全站式的缓存，那么就是指网站里的每一页；若是per-view的缓存，那么就是指单独的页面（视图）。这些方案都非常容易实现。不过有时候你还可以利用这个缓存架构来直接保存一些单块的数据。

假设你运营着一个繁忙的网站，其中很多页面都包含了某些需要很复杂的操作才能得到的信息——例如，它可以是一个定期更新的巨大文件的处理结果。相比起来页面的生成倒是很快速，既然这个信息要在很多页面上显示，那么在这里使用对缓存来说很有意义了。

Django的对缓存在（实际上就是一组可以设置有效期的键值对）允许你存放任意对象，这样你就可以专心处理那些需要大量资源的对象了。以下的代码来自我们的图片示例，还没有被修改第三行里的列表推导式写得很漂亮，但是对于体积巨大的log文件来说它其实也快不到哪里去。我们要做的是避免为每一个请求都生成stat\_list。解决这个问题的主要手段就是从django.core.cache里把cache.get方法去掉。

```
...
def stats_from_log(request, stat_name):
 log_file = file('/var/log/imginary.log')
 start_list = [line for line in log_file if line.startswith(stat_name)]
 if start_list == None:
 start_list = cache.get(stat_name)
 else:
 start_list = cache.set(stat_name, start_list, 60)
 log_file.close()
 template = template_for_line_in_logfile_if_line_startswith(stat_name)
 cache.display(start_list, which_template=template)
```

虽然第三行里的列表推导式写得很漂亮，但是对于体积巨大的log文件来说它其实也快不到哪里去。我们要做的是避免为每一个请求都生成stat\_list。解决这个问题的主要手段就是从django.core.cache里把cache.get方法去掉。

虽然第三行里的列表推导式写得很漂亮，但是对于体积巨大的log文件来说它其实也快不到哪里去。我们要做的是避免为每一个请求都生成stat\_list。解决这个问题的主要手段就是从django.core.cache里把cache.get方法去掉。

```
...
def stats_from_log(request, stat_name):
 log_file = file('/var/log/imginary.log')
 start_list = [line for line in log_file if line.startswith(stat_name)]
 if start_list == None:
 start_list = cache.get(stat_name)
 else:
 start_list = cache.set(stat_name, start_list, 60)
 log_file.close()
 template = template_for_line_in_logfile_if_line_startswith(stat_name)
 cache.display(start_list, which_template=template)
```

## 注意

上面这个例子假设你向模板传递了RequestContext对象，它根据你的context处理器的设置给模板context添加了额外的变量。默认情况下djangocore.context\_processors.cache会触发并负责设置LANGUAGE\_CODE变量。更多关于context处理器的细节请参见第6章。

所以为了适应list\_of\_stuff在每个语言都不同的这个现实，你的cache标签可以这样定义：list\_of\_stuff) 的名字来创建新的键。  
专门为这种情况准备的话，你会希望缓存选择用的键能反映出这一点。所幸 cache标签还有第三个专  
喜好来自渲染数据的话。假如网站有本地化版本或者要根据当前用户的语言  
在该例子中，光用静态的键是不够的。假如网站有本地化版本或者要根据当前用户的语言  
计算)，以及缓存里要用的键。

上面for循环的输出被整个的缓存起来。cache标签接受两个参数：内容要缓存多久(以秒

```
... other uncached parts ...
 ...
 {# endcache %}
 {# endfor %}
 {{ item|do_expressions_rendering_step }}
 {# for item in really_long_list_of_items %}
 {# cache 300 list-of-stuff %}
 ... Various uncached parts of the page ...
 {# load cache %}

```

缺点。不过有了cache模板标签的话，我们就可以直接在需要的地方应用缓存了。  
示循环加上循环里的昂贵的方法调用，所以无论是在视图还是在模型代码里都没有一个单独的类  
没什么好处，而且这个长列表最多五分钟才会更新一次。由于这个列表输出的“昂贵之处”是显  
时费力。同时除了这个列表以外，页面每次载入时都会变化，所以简单地把整个页面缓存下来  
举个来说，假设有一个模板要显示很长的物品列表信息，而生成这个信息的过程又比较费  
时费力。同时除了这个列表以外，页面每次载入时都会变化，所以简单地把整个页面缓存下来

cache标签  
缓存而无须修改视图代码。虽然有些程序员不喜欢这种优化，因为缓存出现在了表现层上，但  
Django提供的最后一类缓存机制就是：cache模板标签。它让你可以从模板里直接使用对象  
缓存中间件——我们只需要导入djangocore.cache即可，无须修改任何设置或是添加任何中

间的字典。如果你还没有注意到，我们在那里告诉你其实对象缓存没有依赖于Django  
另外还有一个get-many方法，它接受一个键的列表，并返回一个包含那些键的(还在有效的  
CACHE\_BACKEND。细节见下。

cache.set方法则接受一个键(字符串)，一个值(任何Python的pickle模块可以处理的值)，  
和一个可选的有效期(以秒计算)。如果你省略了有效期参数，那么它的值就会设置为

## 缓存后台的类型

- 前面介绍的Django缓存类型都是“locmem”。以下是一份完整的有关缓存类型的列表：
- dummy：只用于开发，实际上根本没有缓存，不过能让你保持其他缓存设置不变，以便它们能在实际网站上（使用下例后台之一）正常工作。
  - locmem：一个可靠的进程安全的内存缓存。也是默认内存缓存。
  - db：一个数据库缓存（需要在你的数据库里创建一张特殊的表）。
  - memcached：一个高性能，分布式内存缓存。也是最强大的内存缓存。
  - CACHES：缓存里保存最多的记录数，默认值是300。记住，通常很小的一部分数据就会占用大部分的服务器资源，所以用不着把所有的内容都放到缓存里才能提升性能。而且由于有放期的工作方式，缓存里应该绝大多数都是最常用的数据。如果你的内存比较大或者缓存里数据过大的话，尝试减小这个值。要是你有很多内存又或者缓存的对象比较小的话，可以尝试增大这个值。
  - max\_entries：缓存里保存最多的数据数，默认值为300。记住，通常很小的一部分数据用来决定哪些数据要从缓存里移除，它还用来创建各种包头来告知Web客户端所发送数据。这个数字不单是用来决定哪些数据要从缓存里移除，它还用来创建各种包头来告知Web客户端所发送数据。
  - CACHES的三个可选参数如下所示：
    - max\_entries：缓存里保存最多的记录数，默认值是300。记住，通常很小的一部分数据用来决定哪些数据要从缓存里移除。它还用来创建各种包头来告知Web客户端所发送数据。这个数字不单是用来决定哪些数据要从缓存里移除，它还用来创建各种包头来告知Web客户端所发送数据。
    - CULL\_PERCENTAGE：这个参数的名字起的很糟糕，它根本就不是代表百分比。它指定了当缓存达到max\_entries的时候要删除多少比例的记录数。默认值为3，即每次缓存满了以后，每老的1/3数据会被删除。
    - CULL\_TIMEOUT：数据在缓存里可以存活的时间，默认值为300（秒，即5分钟）。这个数字不单是用来决定哪些数据要从缓存里移除，它还用来创建各种包头来告知Web客户端所发送数据。

```

$ python manage.py createcacheable cache
最后一个是数据表的名字，虽然我们推荐单地叫它cache就行了，不过你可以选择任何喜欢的名字。设置好数据表以后，CACHE_BACKEND就变成了
CACHE_BACKEND = "db://cache"
这张表结构很简单，它只有三列：cache_key(表的主键)，value(实际缓存的数据) 和
expires(一个datetime字段，Django在这一列上设置了一个索引来加快速度)。
Memcached是Django里最强大的缓存机制。同时，它也是设置起来最复杂的一个。但是如果需要的话，这些努力还是值得的。它原本是由Livejournal.com创建用来缓解他们数据库上两千万PV一天的压力。之后它又被移植到Wikimedia.org, Fotolog.com, Slashdot.org等其他著名网站上去。Memcached的主页是http://danga.com/memcached。
Memcached相比其他选择最大的优势就是可以轻易地被分布到多台服务器上。Memcached就是“一张巨大的虚拟的哈希表”。你可以像访问Python字典的键值对那样使用它，但是它无法地把你给它的数据发散到很多服务器上去。
虽然Memcached比这里其他的缓存机制要重量级得多，不过它仍然是一个内存型的缓存机制，而不是一个对象数据库。经常有Memcached的FAQ问道，“Memcached有什么用？”，“Memcached怎么处理故障转移？”，以及“怎么把数据导入Memcached？”，答案是“没有”，而不是“不需要，不需要！”因为你可以把持久化数据放进Memcached。Python的memcached库在这方面已经很完善了，你可以很容易地找到相应的软件包，对于Mac OS X系统则可以去Darwin Ports 或者通过Django应用的服务端，你得告诉Python怎么连接Memcached。纯Python的客户端需要安装memcached库（http://gjsbert.org/gmemcached）。Python-memcached也可以通过Easy Install来直接简易安装设置。
在运行Django应用的服务端，你需要告诉Python怎么连接Memcached。纯Python的客户端
有Python-memcached (http://lumy.com /Community/software/python-memcached) 或者速度更快的C库 (http://gjsbert.org/gmemcached)。python-memcached还可以通过Easy
安装memcached。Windows下的Memcached可以在那里找到 http://splinedancer.com/memcached-ports 或 Macports查找。在Linux下你应该可以很容易地找到相应的软件包，对于Mac OS X系统则可以去Darwin Ports 它。在运行Django应用的服务端，你得告诉Python怎么连接Memcached。纯Python的客户端
win32。

```

memcached提供了丰富的命令行选项，如果你有兴趣的话，可以去看看它的手册或是其他教程没有配置文件。下面这行告诫memcached启动进程方式启动，使用2GB内存，监听IP地址为10.0.1.1。

```
$ memcached -d -m 2048 -l 10.0.1.1
```

设置服务器让其在系统启动的时候自动运行memcached守护进程(daemon)。这个守护进程没有配置文件。下面这行告诫memcached启动进程方式启动，使用2GB内存，监听IP地址为10.0.1.1。

```
install来直接简易安装设置。
```

在运行Django应用的服务端，你得告诉Python怎么连接Memcached。纯Python的客户端
有Python-memcached (http://lumy.com /Community/software/python-memcached) 或者速度更快的C库 (http://gjsbert.org/gmemcached)。python-memcached也可以通过Easy
安装memcached。Windows下的Memcached可以在那里找到 http://splinedancer.com/memcached-ports 或 Macports查找。在Linux下你应该可以很容易地找到相应的软件包，对于Mac OS X系统则可以去Darwin Ports 它。在运行Django应用的服务端，你得告诉Python怎么连接Memcached。纯Python的客户端
win32。

更快而已 (关于Memcached创建和架构的详细细节，请参考文章 <http://www.linuxjournal.com/article/7451>)。

虽然Memcached比这里其他的缓存机制要重量级得多，不过它仍然是一个内存型的缓存机制，而不是一个对象数据库。经常有Memcached的FAQ问道，“Memcached有什么用？”，“Memcached怎么处理故障转移？”，以及“怎么把数据导入Memcached？”，答案是“没有”，而不是“不需要，不需要！”因为你可以把持久化数据放进Memcached。Python的memcached库在这方面已经很完善了，你可以很容易地找到相应的软件包，对于Mac OS X系统则可以去Darwin Ports 它。在运行Django应用的服务端，你得告诉Python怎么连接Memcached。纯Python的客户端
更快而已 (关于Memcached创建和架构的详细细节，请参考文章 <http://www.linuxjournal.com/article/7451>)。

memcached提供了丰富的命令行选项，如果你有兴趣的话，可以去看看它的手册或是其他

Python 通过两个补充模块 (doctest 和 unittest) 和其他一些流行的独立工具对测试提供了出色的支持。这一章 (以及 Django) 主要关注的是那两个内置的系统，要是你对更广泛的 Python 世界的支撑有点兴趣的话，可以参考上面 URL 的内容。

环境信息是 Web 应用很难测试的。这是由它们固有的不整洁性造成的，比如各种现实世界的应用项目的时候能稍微轻松一点。在讨论 Django 的测试之前，我们先来复习一下 Python 测试的基础。

一个 doctest 是一个包含在模块、类或函数的 docstring 里的 Python 会话的拷贝。用 doctest 模块的 testrunner 可以发现这些测试，并能对其进行执行和验证。你可以参考第 1 章来复习一下关于 doctesting 的内容和它们的用法。

```
def double(x):
 return x * 2
```

例如，我们可以轻松地为这个简单的函数编写测试。

若是在解释器里手动测试它的話，可以这样：

这是一章假設你確實相信測試帶來的益處，且專注于如何測試而不是為什麼要測試。要是你覺得需要更多的理由來說服你測試，請參考 [withdjangocom](http://withdjangocom) 上提供的阅读資料。

## 注意

一个公认的观点是拥有一个自动化的测试集对你的应用程序是有好处的。对动态类型的语言来说特别如此，比如 Python，它不在编译期为你提供安全的类型检查。

### 12.4 测试Django应用

Django 会要求你指定端口。Memcached 默认运行在 11211 端口上，刚才的命令行里我们没有指定其他的端口，所以 Memcached 会去监听这个端口。如果有多个服务器的话，可以用反斜杠把它们隔开。

```
CACHE_BACKEND = "memcached://10.0.1.1:11211;10.0.5.5:11211"
```

最后，虽然 Memcached 比其他后台需要更多的设置，但是它依然是一个后台，就是说当你正骑安装 Memcached 后，它和其他后台的工作方式是一样的。

一个公认的观点是拥有一个自动化的测试集对你的应用程序是有好处的。对动态类型的语言来说特别如此，比如 Python，它不在编译期为你提供安全的类型检查。

Django运行测试的命令是：

## 运行测试

以上例子是一个完整的脚本。在执行的时候它会执行自己的测试集。这是由unittest.main()调用完成的，它会搜索所有unittest.TestCase的子类，并且调用所有以test开头的方法。

```
if __name__ == '__main__':
 import unittest
 class IntegerArithmetTestCase(unittest.TestCase):
 def testAdd(self):
 self.assertEqual(self.add(2, 3), 5)
 def testMult(self):
 self.assertEqual(self.add(4, 4), 8)
 unittest.main()
```

源自Smalltalk中的单元测试框架)。unittest在Python里最典型的用法如下：

unittest模块和doctests是殊途同归。它借鉴了Java的测试框架JUnit的理念(jUnit的灵感则是

## unittest基础

```
def double(x):
 """ Doubles the provided number. We hope,
 this function should double the provided number.
 >>> double(2)
 4
 """
def double(x):
 return x * 2
```

文档文本)，所以我们还能给函数增加一些可读的介绍。  
test number还能智能地跳过非测试的文本（比如不是以>>>开头，或是紧跟在它后面的普通  
那么就一切正常。反之则会报错。

当用doctests模块的test runner测试这个函数的时候，double(2)会被执行。要是输出为“4”，

```
def double(x):
 """ Doubles the provided number.
 >>> double(2)
 4
 """
def double(x):
 return x * 2
```

你葫芦瓢照搬到函数的docstring就可以给它添加doctest测试了。  
这正是我们要的结果，所以可以宣布函数通过了测试。然后，我们只要把这段交互的会话

```
>>> double(2)
```

```

>>> p.age_on_date(date(1950, 1, 1))
26
>>> p.age_on_date(date(2008, 8, 10))
27
... state="NJ", birthdate=date(1982, 7, 15))
...
>>> p = Person(firstname="Jeff", lastname="Forcier", city="Jersey City",
28
>>> from datetime import date
29

```

所示：

若要手动测试这个年龄段方法，可在Python解释器里创建一些样例对象然后调用方法，如下所示：

```

在以下age_on_date方法的代码里，从所用知地容易滋生“fencepost”这类错误，即不当的
边界条件 (boundary conditions，例如测试一个人的生日) 会产生不正确的结果。有了doctest,
我们就能防止这类错误发生。

```

```

def __unicode__(self):
 return "%s %s (%s, %s)" % (self.first, self.last)

class Person(models.Model):
 first = models.CharField(max_length=100)
 last = models.CharField(max_length=100)
 birthdate = models.DateField()

 def age_on_date(self, date):
 if date < self.birthdate:
 return 0
 else:
 return (date - self.birthdate).days / 365

```

from django.db import models

代码如下：

例如，假设有一个包含了生日变量的Person模型，你有一个根据这个日期计算年龄段的模型。一旦你添加了模型方法后，你就可以使用它了。

```

因为这时你的模型只是代表了数据表示而已，而Django内部的逻辑早已经过充分测试了。不过
模型的应用寻找它们。如果你的模型只包含数据变量的话，其实没有太多可以测试的东西。
models.py同级的目录）。所以你可以随意按照自己的喜好把单元测试放在任意一个地方。

```

## 测试模型

此外，test命令还会在应用程序子目录下查找任何名为test.py的单元测试文件（和你的blog.Post。

Django会自动检测到INSTALLED\_APPS设置里列出的所有应用程序里的models.py文件中一个单独的应用，甚至只测试应用里某个模型，例如manage.py test blog或是manage.py test blog.Post。

测试（上面介绍的任意一种）。你还可以给test命令提供额外的数据来缩小测试的范围，指定某

./manage.py test

最后，虽然doctests能满足你大多数时候的需要，但是如果你要测试更复杂的商业逻辑或是摸行到两行句号，加上一两个间或出现的E或F（分别代表了意外错误和测试失败）。

一个小小的测试就产生了这么多输出，但你完整地测试了整个模型层次以后，你会得到一

```
Destroying test database...
OK
```

```
Ran 1 test in 0.003s
```

```
Instaliling index for admin.LogEntry model
Instaliling index for auth.Message model
Instaliling index for auth.Permission model
Creating table myapp_person
Creating table django_admin_log
Creating table django_site
Creating table django_content_type
Creating table auth_message
Creating table auth_user
Creating table auth_group
Creating table auth_permission
Creating test database...
user/opt/code/myproject $./manage.py test myapp
最后，运行前面提到的manage.py命令来运行测试：
```

```
def age_on_date(self, date):
 """
 Returns integer specifying person's age in years on date given.
 """

 >>> p = Person(firstname="Jeff", lastname="Forcier",
... city="Jersey City", state="NJ", birthdate=date(1982, 7, 15))
>>> p.age_on_date(date(2008, 8, 10))
26
>>> p.age_on_date(date(1950, 1, 1))
0
>>> p.age_on_date(p.birthdate)
0
if date < self.birthdate:
 return 0
else:
 return (date - self.birthdate).days / 365
```

Django框架自身包含了大量的测试集。每一个bugfix都会配合有一个回归测试来保证已经修正的bug不会悄悄重现。新的功能也都会配有关联的测试来保证它按照预期正常工作。

測試Django代碼本身

除了上述三个工具（Django调试器，Twill和Selenium）之外，你还可以在 <http://www.awardeck.com/tutorials.html#test> 以及里面的链接上找到更多关于Web应用测试的内容。

使用Selenium最好的起点就是它的IDE (<http://selenium-ide.openqa.org/>)。这是一个FireFox浏览器的集成开发环境（IDE），你可以将Web会话记录下来然后通过回放来测试。它还能将测试输出为各种Selenium RC支持的语言，以便你进一步增强或修改它们。你可以设置断点然后一步一步调试这些测试。因为它是在FireFox上的，所以一个经常会被问到的问题就是有没有一步让Firefox（IE）版。答案是没有。不过IDE的“记录模式”（Record mode）可以让你通

Selenium RC (<http://selenium-rc.openqa.org/>) 让用户可以用不同的编程语言创建完整的自动化测试。你自己编写的测试应用，Selenium Core会运行它们——你可以把它想像成在Core之上的一个脚本层（scripting layer），或者“编码模式”（code mode）。

Selenium Core (<http://selenium-core.openqa.org/>) 是（手动和自动）测试 Web 应用的核心。有人将它称为“机器人模式”（bot mode）的 Selenium。粗话重话都由它干了。除了 Web 应用的系统功能测试外，它还可以进行浏览器兼容性测试。

另一个最近比较火的测试工具是Selenium（参见<http://selenium.openqa.org/>）。和其他两个不同，这是一个基于HTML/Javascript的测试工具，专门从浏览器的角度来测试Web应用。它支持大多数平台上的主流浏览器，而且因为它是基于Javascript的工具，所以还可以测试Ajax功能。它能分成2.5到3种操作模式：Selenium Core, Selenium RC (Remote Control, 远程控制) 和 Selenium IDE (集成开发环境, Integrated Development Environment, 集成开发环境)。

并测试特定条件。

第一个要关注的就是内建到Django里的工具，在本书编写的时候它还是相当得新。它叫做“Django测试客户端”，你可以在Django官方网站上找到它的文档 <http://www.djangoproject.com/documentation/testing/tools>。测试客户端提供了一个简单的方法来模仿访客请求—响应循环，

自上而下地测试Web应用不是一件容易的任务，而且也没有办法用同样的脚本达到100%自动化，因为每个Web应用都是不同的。不过还是有一些工具是很有用的。

测试整个Web应用

类型关系的时候，也不要对设置模型相关的单元测试有顾虑。要是你是测试新手，可能要花一点时间才能分辨出什么时候要用哪个方法——不过千万不要放弃！

1

在运行测试集过程中看到一个失败不一定代表你发现了Django的错误——如果你不确定的话，最好先到Django的用户邮件列表上把你的配置细节和测试失败的输出贴出来给大家讨论。

(这里的“...”代表为节约篇幅，剩下的输出省略了)。

从为的设置失败还是Django确实有问题。

以上E字符代表了一些建议或出现了错误，后面会显示给出具体失败的总结，以便你掌握这里

..... 33 3 .....

在-VI下的输出是这样的：

这是一个相对安静的过程，因为测试只在失败的时候才会生成输出。由于测试集的运行需要一段时间，所以你会在这个过程中不断收到反馈。runtests.py命令会接受一个--verbosity的参数。

```
$ tests/runtests.py -settings=mydummyproject.settings
```

其中就包含 `runner` 其调用命令如下：

(即没有应用的项目)，然后只在settings.py文件里填入DATABASES的设置。

运行Django的测试集相当简单，只需要一点小设置即可：你需要指向一个配置文件，这样它才知道怎么创建测试数据库。这可以是任何项目的配置文件，或者你可以创建一个空项目。

你可以自己运行这些测试。当你在一个比较少见的平台或是在一个不常见的配置下运行 Dango 有问题时非常有用。虽然你总是应该先检查自己的代码，不过也有可能是你发现了一个之前从未见过的不寻常 bug。在内置测试集上的失败会为你生成一份 bug 报告，以便稍后仔

这一章覆盖了好几个高级的主题，算上第11章，我们希望你能充分了解到Django开发的深度。当然，这些主题只是一个展示可能性的例子：Web应用开发和其他大多数计算机分支一样，能够处理各种情况和技术。

到此，你几乎就要读完这本书了——恭喜！剩下的还有附录，和最后两章一样，它们也是本书重要的一部分，涵盖了很不同的主题，从命令行使用到安装部署Django，还包括一系列的外部资源和开发工具等。

最后，你可能还会想要回去重读一下（至少是浏览）本书前面的部分。在了解了我们所有要讨论的东西后，之前的代码示例和解釋或许能让你温故知新。对任何技术书籍来说都是如此，本书也不例外。

## 12.5 总结



和点击鼠标、填写文本、用鼠标驱动的界面不同，类UNIX的服务器操作系统是由命令行操作的。

## A.1 在命令行模式下输入命令

如果你是Mac的用户，那算你运气。Mac OS X是BSD (Berkeley Software Distribution) Unix继承而来的一个分支，所以它具备一台服务器的许多功能。你只需打开终端 (Terminal) 应用程序（可以在 /Applications/Utilities 下找到），就可以输入命令了。从现在开始，我们假设你已经进入UNIX的命令行模式了。

要是在Windows系统上尝试这些命令的话，可以试试一个模仿Linux的环境Cygwin。<http://cygwin.com>获取。但是它可不能就此把你的PC机变成一台服务器。更多关于Cygwin的信息可以从它的官方网站以及一系列UNIX用户常用的命令行工具来说，在这一节里将会介绍其中的一些。

如果你来自Windows的世界，本附录里的知识可能不是那么直观，不过我们还是建议你阅读一下（至少浏览一遍吧），将来总会影响到这些东西的。而且通常来说，一个程序员接触的语言、平台和技术越多，他越是能够更好地利用这些现有的和新的工具。

如果在此之前从来没有接触过在这种环境中常见的命令行界面的话，本附录会为你做一个基础。如果你之前从来没有接触过在这样的机器上运行框架POSIX的操作系统上，比如 Linux、FreeBSD，还有各种类UNIX系统等，Django的Web服务器当然也不例外。几乎所有的Django核心团队以及很大部分的社区都是在这样的机器上运行框架的。如果你之前从来没有接触过在这种环境中常见的命令行界面的话，本附录会为你做一个基础。这能让你更从容易理解本书中的例子。

## 附录A 命令行基础

附录F 参与Django项目

附录E 在Google App Engine上使用Django

附录D 发现，评估，使用Django应用程序

附录C 实用Django开发工具

附录B 安装运行Django

附录A 命令行基础



```

--version output version information and exit
--help display this help and exit
--verbose explain what is being done
-v, -R, --recursive remove contents of directories recursively
--preserve-root fail to operate recursively on '/'
-no-preserve-root do not treat '/' specially (the default)
--interactive prompt before any removal
-e, -force ignore nonexistent files, never prompt
-d, -directory unlink FILE, even if it is a non-empty directory

Remove (unlink) the FILE(s).
Usage: [bin/rm [OPTION]... FILE...
$ rm -h
rm: cannot remove 'temp': Is a directory
$ rm temp
tempfile tempfile
$ ls temp
$ rm temp
$ rm documents/test.py
test.py
$ rm documents
$ ls documents
documents code temp
$ ls
$ user@example$.

```

这里用到的两个命令都是位于当前执行路径中的程序，或者说二进制文件（下一节会详细讨论它们的执行方法。Unix命令一般由三个部分组成：命令的名字，控制命令行为的选项，指定运行时的基准参数）。虽然理论上程序的行为是没有任何限制的，但是我们还是设立了一些指定参数和选项（比如路径）。上面这个例子的最后一条命令来说，rm是程序的名字（rm就是remove，代表删除的意思），\$后面跟着要操作的文件名等参数。

从上面这个例子可以看出，它列出了当前目录的内容，并且从子目录里删除一个文件。下面是一个简单的例子，它列出了当前目录的内容，并且从子目录里删除一个文件：

```

$ rm documents/test.py
test.py
$ rm documents
$ ls documents
documents code temp
$ ls
$ user@example$.

```

这里用区分开哪些是输入的命令，哪些是命令的输出。下面这个简单的例子展示了如何列出我们当前目录的内容，以及从子目录里删除一个文件。注意\$是一个提示符，在下面的例子，\$前面的命令是输入的命令，\$后面的是命令的输出。

除了这个出现在用户名之前的符号之外，很多提示符还可以给出更多的信息，是>>>提示符。除了这些命令之外，还有很多其他的命令，比如>或者%。（Python解释器用的除了\$，每个shell都可能使用不同的字符来作为提示符，比如>或者%。）

解释器或者说shell来驱动的，文字提示符接受命令并且逐个执行。如果你有编程的经验，你一定不难。

熟悉编程语言的表达式：打印一个字符串、向调用的函数传递一些参数等。命令行基本上也差不多。

前面我们说到进阶和参数是程序规范里两个完全不同的部分，其实这么说并不完全正确。在

A.2 选项和参数

很多UNIX程序的名字看上去都非 常莫名其妙，像什么rm、ls、sed等。这些名字有些有些简单易辨别的，比如rm（移除remove）和ls（列出list），而有些就晦涩一点，比如sed代表的是“stream editor”（流编辑器）。其他的是一些，特别是较新的程序员的名字，则是通过各种不同的缩写组合起来的，例如流行的编译器GCC代表的是“GNU C compiler”，这里的GNU指的是GNU项目，它本身也是“GNU's Not UNIX”的缩写。

总的来说，UNIX命令都喜欢用缩写来表示——一旦你知道了这些命令的作用后，你自然而然的就会理解这些简写名字的意义了。

如果内置的帮助也不能满足你，或者你想更深入地了解某个选项的详细信息的话，可以参考man系统（就是用户手册manual的意思），它提供了每一个命令的完整信息，更加详尽地解释参数的作用，有时还会给出例子。当然，man本身也是一个程序，一般它只接受一个参数，即你需要查阅文档man page的程序名。一个经常给UNIX新手举的例子就是man man，即man命令自己查看自己的帮助手册。

注意 你的输出可能和这里有所不同，这是因为UNIX系统的表现实在是太多了，每个平台上对这些程序的实现都会稍有不同。即便是rm和ls这样的核心命令也不例外，不过它们之间至少有一部分是一样的。

在一殷情况下，rm是不能删除目录的，这就是为什么一开始我们失败了。但是插入-t和-f选项后——方便起见，你可以只用一个选项将号把它们组合在一起（详见稍后的A.2节）——就可以让rm强制递归地删除目录了，所以它就可以无障碍地删除temp目录。

§

这里的例子都为了展示概念而尽量的保持简洁，要是你在网上查找这些命令行程序用法的例子（或是查阅各种种 man page 和 --help 输出）时，你会发现这些命令行程序提供了大量的功能和灵活性来改变它们的行为。在下一页里，我们会介绍一种完全不同的UNIX命令行工作方式。

首先。

虽然Linux程序的这种风格非常好用（比如你直到命令打完了才想起来忘了某个选项的时候），我们还是建议你尽量地遵循正统UNIX系统上那种更严谨的格式。否则到时候你会发现在使用FreeBSD或是Mac OS X时，不能收到程序警告你的参数顺序。相信我，我们吃过这样的亏

或者：

```
$ rm -rf temp --verbose
```

字解命令来说，很多Linux的应用更加宽容一点，例如：

```
$ head myfile.txt -n5
```

在默认情况下（即不带-n选项），head会显示头五行内容。另外，就算在多个选项组合到一起的情况下，选项的参数也不一定要和它的选项用空格隔开，即放在一起也是可以的。

```
$ head -n 5 myfile.txt
```

myfile.txt文件的前五行：

文本的开头几行。下面的例子展示了返回的行数是可以通过-n参数来控制的，这里返回了根据选项的不同，它们自身也可以携带参数。例如，head程序的作用是返回给定文件或者

```
$ rm -rf --verbose temp
```

样是可以的：

不过请注意你不能把短选项和长选项组合在一起，因为那样做是没有意义的，但是下面这

```
$ rm -rf temp
```

也可以把这些选项组合起来，少打几个字符，就像我们之前做的那样：

```
$ rm -rf temp
```

个参数（有的程序根本不接受参数）。选项可以是一次一个，用空格分开：

标准是这样写的：工具名/程序名，后面跟着一个或者多个任意类型的选项，再跟着一个或者多

```
Usage: rm [OPTION]... FILE...
```

```
$ rm --help
```

个连字符再加上选项的完整名字——比如前面提到的--help选项。

通常程序所有的参数和选项都是用空格隔开，选项一般出现在参数之前，以连字符-作为前缀开始，而参数则没有前缀。有的程序还接受一种叫“长选项”的格式，通常这种格式使用两个连字符--再加上选项的名字，比如 rm --help。因此，这里的标准非常简单：完全取决于程序的作用者在大多数情况下严格遵守了标准。它都可以。但是，这个内部，选项和参数只不过就是一个传递给程序的长字符串而已，程序不管按什么规则解释

本质上来说，命令行主要是处理文本的输入输出。但是除了从用户那里获得输入输出之外，UNIX程序还可以通过一种叫管道（pipe）的I/O抽象机制来和磁盘上的文件交流。顾名思义，每个UNIX程序都能处理三种潜在的输入输出：输入、常规输出和与错误有关的输出。管道是一种能在终端用户、程序和文件之间定向文本流的机制。

一般情况下，程序交互的是所谓的“标准”管道，即用下面对的文字交流。例如，当你使用cat命令（或是连接concatenate）来导出文件的内容时，cat会打开参数中所列出的文件并把它们的内容放到标准输出流（stdout stream）里。就跟在下面这个例子里，我们cat一下grocery列表的文件到标准输出流（stdout stream）里。输出一个错误消息，在默认情况下，这里直接运行cat的结果是stdout会把文件的结果输出到终端上。要是cat发生错误，比如给定的文件不存在的话，它会向标准错误流（stderr stream）输出一个错误消息，在默认情况下，也是直接返回给用户。

```
cat: foo.txt: No such file or directory
$ cat foo.txt
```

这里直接运行cat的结果是stdout会把文件的结果输出到终端上。要是cat发生错误，比如给定的文件不存在的话，它会向标准错误流（stderr stream）输出一个错误消息，在默认情况下，也是直接返回给用户。

举个例子，我们用grocery列表来重翻看一下head的用法，这次我们用它来查看列表中的第一个项目。在这里我们没有告知head要处理的文件，而是用管道将cat命令的输出重定向给了head命令。其结果和我们直接将文件传递给head是一样的。

```
$ cat groceries.txt | head -n1
```

注意，这里我们没有告知head要处理的文件，而是用管道将cat命令的输出重定向给了head命令。其结果和我们直接将文件传递给head是一样的。

举个例子，我们用grocery列表来重翻看一下head的用法，这次我们用它来查看列表中的第一个项目。把我们grocery列表中不含“can”这个词的行都过滤掉：

```
$ grep -i "can" groceries.txt
```

一个更真实的例子是使用grep工具，它可以用来逐回匹配完全正则表达式（参见第一章）。

### A.3 管道和重定向

列出了 env 的一小部分输出：

和 Python 以及其他语言一样，你可以用赋值运算符来给这些变量赋值。简单起见，这里只讲程序在决定是否使用着色输出或是如何解释击键的时候都要引用它。PWD 则是当前的目录等。

到 EDITOR 变量调用指定的程序来替换时保有用户的输入。TERM 变量则决定了终端的类型，很多环境变量都是给常用的 UNIX shell 工具或是 shell 本身使用的——比如 Subversion 就会用到 EDITOR 变量来替换时保有用户的输入。TERM 则是当前的目录等。TERM 变量则决定了终端的类型，很多环境变量都是给常用的 UNIX shell 工具或是 shell 本身使用的——比如 Subversion 就会用

```
HOME=/home/user
EDITOR=vim
PWD=/home/user
PATH=/usr/local/bin:/usr/bin:/usr/games:/bin
USER=user
SHELL=/bin/bash
TERM=linux
$ env
```

行 env 命令会打印出当前状态下的 environment，如下所示：

命令行命令时可以引用它们，甚至命令自身也可以使用它们（即访问用户的 environment）。执行环境变量里的 namespace 一样，它可以保存很多制定了变量名的不同的字符串，用户可以在执行命令时可以引用它们，甚至命令自身也可以使用它们（即访问用户的 environment）。执行环境变量里的 namespace 一样，它可以保存很多制定了变量名的不同的字符串，用户可以在执行命令时可以引用它们，甚至命令自身也可以使用它们（即访问用户的 environment）。

## A.4 环境变量

和 > 一样，如果新文件不存在的话，它会自动创建。

最后，你还可以用两个重定向符号 (<>) 来将输出添加到现有文件的末尾，而不是覆盖它。

这三条命令创建了一个叫做 filtered.txt 的新文件（小心，它也能覆盖现有文件！），其内容为“Canned veggie”。注意这条路命令不会在终端上产生任何输出——因为我们已经把 stdout 重定向给了文件了，所以当然看不到了。

```
$ grep -i "can" groceries.txt | head -n1 | sed -e "s/corn/veggies/" > filtered.txt
```

（当然，复制粘贴除外）。现在让我们来把它重定向给一个新文件。

我们说过你可以把文字流用 < 和 > 来给重定向给文件。和管道从左到右的作用方式一样，< 用来将 stdout 重定向给文件，而 > 则是将文件重定向给 stdin。例如在上面的例子中我们完成了查找和替换，虽然它不是那么有用，但是一旦结果显示到屏幕上以后，我们的劳动成果就要失掉了（当然，复制粘贴除外）。现在让我们来把它重定向给一个新文件。

```
$ grep -i "can" groceries.txt | head -n1 | sed -e "s/corn/veggies/" > canned_veggies
```

前面说过，我们可以在一个命令里使用多个管道。现在让我们来用 grep 的姐妹命令 sed 把单行面说过，我们可以在一个命令里使用多个管道。现在让我们来用 grep 的姐妹命令 sed 把单

```
$ grep -i "can" groceries.txt | head -n1
```

然后用 head 来把结果限制在第一页上。

```
最后一个要点是：环境变量不是只能保存单个单词的简单字符串而已，它能保存任何字符串。在上面的例子中，shell展开$EDITOR后，将它和行里剩下的内容拼在一起，然后试图将它们当作一个命令来执行。这显然是行不通的（因为没有$! vim=pic0这个名字的程序），但是我们能够利用这一点来简陋地打几个字符，比如下面的例子，我们将一个命令字符串值给变量，然后加上不同的参数重複使用。
$ FINDMILK="grep -ai milk"
$ FINDMILK $FINDMILK groceries.txt
$ FINDMILK todo.txt
$ FINDMILK 1:Milk
$ Powdered milk
$ Powdered milk
1:search grocery list for milk
$ FINDMILK from-reader.txt
$ FINDMILK email@example.com examples?
```

在该例子中，默认的EDITOR值是著名的vim编辑器，我们让它指向了一个不那么强大的编辑器picco。但是和你猜想的一样，环境变量是在每个shell会话开始的时候初始化的。shell启动时都会去读shell的配置文件，除非我们对这个配置文件作出修改来永久的保留设置，否则

<p>Unix执行文件目录</p> <p>Unix系统传统上为不同类别的程序准备了多个目录，比如/usr/bin存放的是普通用户的二进制文件（即可执行文件，就和那些我们已经看到的工具一样），而/usr/local/bin则是存放用户安装的程序（即不是随操作系統安装的程序）。而其他的二进制文件，请如属于系统管理員而非普通用户的程序，则是存放在/sbin和/usr/sbin之下。</p> <p>有些第三方应用程序，比如JAVA实现或旧版的Mozilla套件，会把它们的数据放在/opt/media下，而把执行文件放在/opt/Mozilla/bin下面。路径通常包含了系统执行文件目录的一部分，第三方执行文件目录和其他目录等。</p>
-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

当你在shell里输入一条命令时，shell会逐个检查每一个列出的目录直到它找到你要求的可执行文件，然后，它就会用你提供的参数和选项运行那个文件。比如，当我们键入echo时，shell真正执行的是/bin/echo，当键入man时，它找到并执行的是/usr/bin/man。（下面会解釋这些不同的bin目录。）

```
$ echo $PATH
/usr/local/bin:/usr/bin:/bin:/usr/games:/bin
```

当你键入命令时shell会在这些目录里查找。

可能最重要的环境变量之一就是路径了，通常都是保存在PATH里，它包含了一系列目录，

### A.5 路径

我们的FINDMILK变量每次都被shell展开，变成grep -ni milk groceries.txt这样的命令。不过这个例子有点太矫情了——在大多数这种情况下，你真正需要的是参数化这个静态的程序调用，写一个简单的shell脚本就好。例如你写一个接受两个参数的脚本，用来指定要查找的字符串以及在哪里查找。

shell脚本编程的细节已经超出了本章的范畴，你可以参考有关shell单行的man page，网上也有大量优秀资料可以查阅。

shell提供的路径功能相当的方便和快捷。当然对不在路径中的可执行程序也可以通过指定完整或相对的路径来执行它们。事实上，shell可以轻松地找到它们。

因为你~代表的是用户的home目录，这意味着当用户直接输入他们个人执行文件目录里的名字时，shell可以轻松地找到它们。

当然对不在路径中的可执行程序也可以通过指定期望或相对的路径来执行它们。事实上，

```
$ /tmp/package-name-2.0.1/bin/program
```

shell提供的路径功能相当的方便和快捷。

由于/tmp/package-name-2.0.1/bin/（如果你的程序没有完整安装的话，这可能是可执行文件的位置）不太可能在你的PATH路径里，你必须明确的告诉shell怎么找到它。虽然非常简单，不过路径这个概念为你提供了一条捷径。

最后记住，在PATH里包含的应该是包含执行文件的那个目录，而不是执行文件本身——

要引用的东西。

这一点和Python的模块路径是一样的道理（参见第1章）。把路径想象成一组容器，而不是一组来让这些程序一起和文件工作，以及怎样用环境变量和路径设置来节省大量的时间。到此你已经学习了相当多的内容：如何用不同的选项和参数运行程序，通过管道和重定向但是我们在那里讲授的东西即便是普通的UNIX shell程序也还是有用的——很多shell自身就是一个个完整的编程环境，包含有条件语句、循环等等。当熟练使用文件系统和运行命令后，你会发现花点时间深入了解一下系统的基本shell是非常有好处的，和其他编程工具一样，它能为你节约大量的时间和精力。

## A.6 总结

是直接下載源碼。

UNIX/Linux

把新版本设置为系统默认版本。

Python 在Mac上是默认安装的，但是除非你使用的是一带Python 2.5的OS X 10.5 (“Leopard”) 的话，你Python的版本应该是2.3，我们推荐你最好升级。Python 2.4和2.5都可以直接从python.org获得安装包，或者你也可以通过MacPorts (<http://macports.org>) 这样的软件仓库来安装。如果你决定使用MacPorts的话，你还需要安装python-select包，这样你就可以很方便地安装Python 2.5了。

Mac OS X

Python 2.5.1

着（注意是大写的“V”）。如果Python正确编译安装的话，它会显示版本号然后退出。

如果不确定你机器上安装的Python是什么版本，可以打开命令行终端并且输入python -V 查看

包。下面我們圖要介紹一小間半合上要注意的地方。

在Python的官方网站<http://www.python.org/download/>可以下载到大多数主流平台的安装

問2.6.這樣，同時還有一代的新版本3.0。) 請參考第一章來更多地了解這些版本的差異。

不过我们推荐还是使用你自己的类来重写新数据模型的版本比较好。（在编写本书时，Python 在从 2.3 到

和大多数语言一样，版本号是翻转过来的。Django 从这个文件在任何 Python 2.3 以上的版本里，

B.1 Python

开发Django首当其冲要进行安装，而安装的方式则取决于你选择的操作系统和手头上的工具。运行Django最简单的环境包是Python 2.5，一个轻型的SQL数据库包，以及Django内置的开发Web服务器。

附录B 安装运行Django

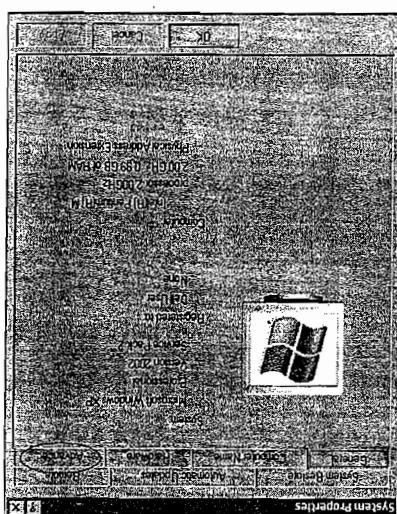
B.4所示。

选择你要修改的变量然后点击“编辑”按钮，按照将其把C:\Python25添加到列表里，如图B.4所示。这还取决于你是否有权限那么做。

修改/更新你自己的路径变量（“USER的用户变量”）PATH，还是为整个系统（“系统变量”）作修改/更新，你会看到有两个面板（如图B.3）可以让你修改环境变量。你可以选择只是添加/修改/更新你自己的路径变量（如图B.3）。跳过这些直接点击窗口下方的“环境变量”按钮。

你会看到三个主要的部分（如图B.2）。通过这些直接点击下方的“环境变量”按钮。

图B.1 系统属性



图B.1所示。

右键点击“我的电脑”，选择“属性”进入系统属性对话框。然后选择“高级”标签，如下：

这一步通常都已经自动完成了。但是Windows用户则需要手动地把它添加到路径里去，步骤如下：

OS X这样的类Unix系统上，如果Python是安装在常见的/usr/bin，/usr/local/bin等目录下的话，这一步通常都已经自动完成了。但是Windows用户则需要手动地把它添加到路径里去，步骤如下：

Python 安装完成后，你可能还需要把可执行文件添加到系统的路径里去。在Linux和Mac OS X这样的类Unix系统上，如果Python是安装在常见的/usr/bin，/usr/local/bin等目录下的话，这一步通常都已经自动完成了。但是Windows用户则需要手动地把它添加到路径里去，步骤如下：

### 重新你的路径

另外，还有一个可选的Windows平台专用的Python库，叫做Python Extensions for Windows (即win32all)。这个库让你能直接开发原生的Windows Python应用。就算Python的新用户还没有这种系统集成库，他也会发现PythonWin这样的IDE需要这个包。

《Core Python》一书的网站http://corepython.com上的Python下载页面来获得各种平台上Python的最新版本。

Windows系统默认不自带Python，所以你需要到官方网站上去下载。或者你可以访问《Core Python》一书的网站http://corepython.com上的Python下载页面来获得各种平台上Python的最新版本。

Windows

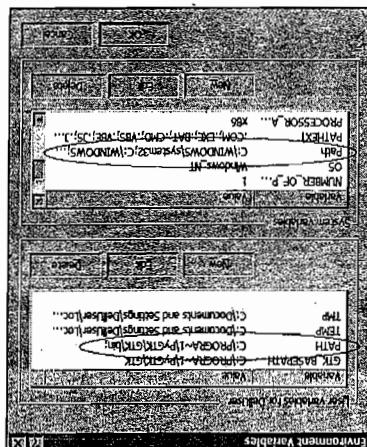
路径C:\Python25\python.exe，而只需要输入python就能启动解释器了（参考下一节）。当你点击OK后，可以打开一个新的DOS命令行窗口，这时你应该已经可以在不需要完整的文件名的情况下运行脚本了。若还没有这个变量，那就按照加粗的就可以了——只要保证所有的文件夹都是用分号隔开的就行了。

如果PATH里已经有文件夹了，你可以把它添加到任何位置，只要保证所有的文件夹都是用分号隔开的就行了。若还没有这个变量，那就按照加粗的就可以了——只要一个文件夹是完全正确的。

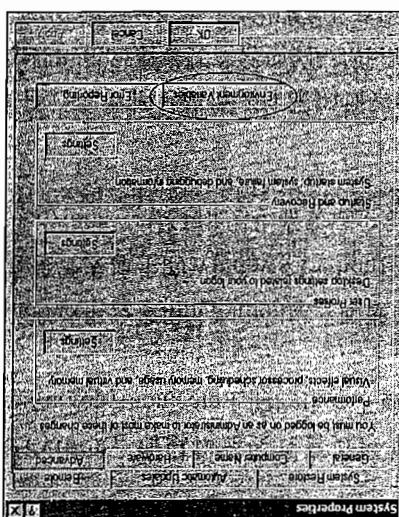
图B.4 修改路径



图B.3 环境变量



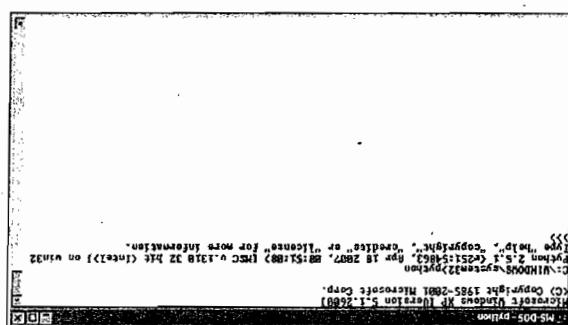
图B.2 环境变量



要确认你的Python安装正确，只需要运行Python主程序启动交互解释器就可以了。在shell或是终端窗口里输入python（可能还需要带上版本号，比如python2.4），如果一切正常的

话，应该可以看到如下的显示结果：

```
$ python
Python 2.5.1 (r251:54863, Mar 7 2008, 04:10:12)
[GCC 4.1.3 20070929 (prerelease) (Ubuntu 4.1.2-1ubuntu2)] on Linux2
Type "help", "copyright", "credits" or "license" for more information.
>>>
```



图B.5 cmd.exe下的Python

Python代码。要退出可以输入Ctrl-D (UNIX shell或是IDLE) 或者Alt (DOS命令窗口)。三个大于号提示符(>>>)表示你正在交互解释器的提示符下，你可以键入任何合法的

如果你得到一个类似“command not found”(找不到命令)或是“python is not recognized as an internal or external command”(python不在一个可识别的内部或外部命令)这样的错误信息的话，说明你没有把文件夹正确地添加到PATH变量里去。请重新检查你的设置，仔细检查你的安装路径，确认python.exe的存在后，再把文件夹的名字加入到PATH里去。

最后为方便起见，你还可以把Scripts文件夹放到路径里去。这样你就可以用和启动Python解释器一样的方法使用 Django的管理工具django-admin.py了。

如果这是你第一次使用Python，而且你还没有读过任何教程或是本书的第一章的话，我们建议你在继续下去之前至少先粗略的看一下它们。在正式开始使用任何工具之前，稍微做点测试总是有好处的。

## 可选的插件

除了Python，这里还有两个值得推荐的（但不是必须的）工具：Easy Install和Python。

## Easy Install

Python 最强大的特性之一就是它默认提供了一个丰富的标准库来帮助你完成任务，相当的“傻瓜”。要是这还不够的话，还有更多的第三方库可供选择。所以在你重新工作之前，先到

Searching for ipython  
C:\>easy\_install ipython

已经安装有最新版IPython for Windows的PC上尝试安装时的结果：  
在具备管理员权限的情况下，Windows系统上的安装过程基本上也差不多。这里是在一台

```
easy_install ipython
 Downloading http://ipython.scipy.org/dist/ipython-0.8.4-py2.4.egg
 Processing ipython-0.8.4-py2.4.egg
 Creating /usr/lib/python2.4/site-packages/ipython-0.8.4-py2.4.egg
 Adding ipython 0.8.4 to easy-install.pth file
 Installing ipyloader script to /usr/bin
 Installing ipyclient script to /usr/bin
 Installing ipylib script to /usr/bin
 Installing ipykernel script to /usr/bin
 Installing ipylib script to /usr/bin
 Installing ipykernel script to /usr/bin
 Installing ipylib script to /usr/bin
 Installing ipykernel script to /usr/bin
 Installing ipylib script to /usr/bin
 Installing ipykernel script to /usr/bin
 Installing ipylib script to /usr/bin
 Installing ipykernel script to /usr/bin
 Installing ipylib script to /usr/bin
 Installing ipykernel script to /usr/bin
 Installing ipylib script to /usr/bin
 Installing ipykernel script to /usr/bin
 Installing ipylib script to /usr/bin
 Installing ipykernel script to /usr/bin
 Installing ipylib script to /usr/bin
 Installing ipykernel script to /usr/bin
 Installing ipylib script to /usr/bin
 Installing ipykernel script to /usr/bin
 Installing ipylib script to /usr/bin
 Installing ipykernel script to /usr/bin
 Installing ipylib script to /usr/bin
 Installing ipykernel script to /usr/bin
 Installing ipylib script to /usr/bin
 Installing ipykernel script to /usr/bin
 Installing ipylib script to /usr/bin
 Installing ipykernel script to /usr/bin
 Installing ipylib script to /usr/bin
 Installing ipykernel script to /usr/bin
 Best match: ipython 0.8.4
 Reading http://ipython.scipy.org/dist
 Reading http://pypi.python.org/simple/ipython/
 Searching for ipython
 Password:
$ sudo easy_install ipython
```

为例讲述一例（用sudo来获得超级用户权限）：

IPython 是一个随 Python 发行的标准交互解释器的第三方实现。它在 Python 原版的基础上添加了许多有用的功能，包括自动生成变量名和属性名、命令行历史、自动对齐、可以轻松的访问 docstring 和参数签名等。基于这些优点，当你用 Django 项目的 manage.py 管理脚本启动 shell 时，它会将其作为默认的解释器。更详细的关于 IPython 的信息可在 <http://ipython.scipy.org> 下找到。

同样，升级和卸载软件包也是一样的简单。

Easy Install 通过 “PyPI” 来获取所需软件的最新（或者你所要求的）版本，自动下载（及其他你的软件包）并且为你安装，这一切只需要一个简单的 shell 命令就可以完成了。

```
easy_install NEW_PROJECT_SOFTWARE
easy_install "PyPI"
 Python Package Index (或者叫 "PyPI", http://pypi.python.org) 上瞧瞧有没有适合你的工具。
 如果知道了你需要的工具，管理你的Python安装就变得麻烦起来。你得关心Python版本的兼容性，软件之间的依赖关系，以及将新的模块包集成进来以便你的应用程序员可以导入它们等。
 好在一个叫 Easy Install 的工具可以帮助你管理这一切，它的位置在 http://peak.telecommunity.com/DevCenter/EasyInstall。你只需下载一个ez-setup.py 文件，运行它就可以了（需要 sudo 或者管理员权限）。它能大大简化新软件包的安装过程，如下：
```

要使Django工作起来有三种方法：  
脚本和信息。

代码——不仅有框架（Python模块本身，django目录），还包括了文档以及测试文件等一系列其他  
前的位置上会有一个新目录叫做Django-1.0或是django\_trunk。目录里包含了Django的完整代  
解开tar.gz文件（对打包发行版来说）或检出Django的源码（对开发版来说）后，在你当

## 安装

```
$ svn co http://code.djangoproject.com/svn/django/trunk django_trunk
```

当有了Subversion之后，只需一行命令即可获得Django的最新版本：

（在OS X上）获得，或者直接从Subversion的网站 (<http://subversion.tigris.org>) 下载也行。

Django的开发版里包含了稳定的版本没有的新特性，它需要使用Subversion版本管理客户端  
能取得。和Python类似，Subversion通常可以通过软件包管理器（在Unix上）或是MacPorts

## 开发版本

(<http://www.cygwin.com>)。

//gnuwin32.sf.net/packages/libarchive.htm)，或者“Windows上的Unix”环境工具如Cygwin  
则需要额外的软件支持，比如7-Zip (<http://7zip.org>)，命令行下的LibArchive (<http://gnuwin32.sf.net/packages/libarchive.htm>)，或者“Windows上的Unix”环境工具如Cygwin  
的Unix格式tar.gz打包——Unix和Mac系统可以无须任何设置就能打开它们，而Windows用  
的com/download/（或者通过你系统上的软件包管理器）获得。网站上下载的官方版本是以常  
见的Unix格式tar.gz打包—— Unix和Mac系统可以无须任何设置就能打开它们，而Windows用  
户则需要额外的软件支持，比如7-Zip (<http://7zip.org>)，命令行下的LibArchive (<http://gnuwin32.sf.net/packages/libarchive.htm>)，或者“Windows上的Unix”环境工具如Cygwin  
的com/download/（或者通过你系统上的软件包管理器）获得。网站上下载的官方版本是以常

## 打包发行版

版总的来说是比较保险的。当然如果你很激进的话，下面的信息会帮助你使用Django的开发版本。  
版正在开发。本书的内容主要是基于1.0版，所以我们推荐你使用该版本，而使用最新的稳定  
安装完Python后，接下来就轮到Django了。在编写本书时，Django 1.0已经发布，同时1.1

## B.2 Django

```
Best match: ipython 0.8.2
Processing ipython-0.8.2-py2.5.egg
ipython 0.8.2 is already the active version in easy-install.pth
Deleting c:\python25\scripts\ipython.py
Installing ipython-script.py script to c:\python25\scripts
Installing ipython-exe script to c:\python25\scripts
Installing ipython-script.py script to c:\python25\scripts
Processing c:\python25\scripts\ipython
Using c:\python25\lib\site-packages\ipython-0.8.2-py2.5.egg
Processing dependencies for ipython
 Finishing processing dependencies for ipython
```

的问题。早发现，早解决。

而且更好的是，你还可以找到一个更加接近真实的环境进行开发——这有助于及早发现部署上面的问题建议。当你的开发走到这一步时，你最好还是花点时间设置一下下面要提到的服务器环境。最后，虽然开发服务器完全可以处理静态文件（参见官方文档或者[whdjangocom](http://whdjangocom)），我们建议的是，你还可以找到一个更加接近真实的环境进行开发——这有助于及早发现部署上面的问题。

但是，要构建一个高级、可靠、安全的Web服务器不是一项简单任务，明智的Django团队决定不涉及这个领域。开发服务器没有经过大量的测试，所以它绝对不应该被用于任何测试以外的情况，例如部署到 Internet上。Django的文档在说到这个问题时开玩笑地说：“如果你真的考虑要把它用在发布环境里的话，我们就需要告诉你的Django机器。”这一点本身是相当严肃的。

这就是开发服务器的区别之处：快速地实现在新特性。用它来调试也是相当不错的，因为它决定了你已经见过它了，在第二章中你已经见过它了，它可以通过manage.py方便的测试一个简单的Web应用。的例子）。在第二章中你已经见过它了，它可以测试一个纯Python标准库的服务，这是一个基于Python内置的BaseHTTPServer的服务器（一个绝佳的利用Python标准库的示例）。

实际上使用的最简单的Web服务器是Django内置的“runserver”服务器，也叫开发（“dev”）内置服务器：不能用于发布

## 内置服务器：不能用于发布

安装完Python和Django之后，要使用Django就算前进了一大步了。下一个重要的步骤是Web服务器，它的作用是把你动态生成的HTML页面返回给浏览器。Web服务器还负责处理图片和CSS等静态的内容，以及各种系统级别的事情（负载平衡，代理等）。

在 Python解释器里简单地输入 `import django` 即可检查 Django是否正确安装到了 PYTHONPATH里去了。如果一切正常的话，安装就完成了！如果你得到一个ImportError: No module named django错误信息的话，重新检查你的步骤，或者试试别的方法。

## 测试

要知道更多关于Python路径机制的信息，请参阅第一章。我们建议如果你打算使用Subversion来跟踪开发最新的进展的话，第一或者第二种方法都可以（第二种方法里只能使用这个命令，所以我们不推荐这种方法）。

• 在新目录下（以管理员身份）执行`python setup.py install`命令，它会自动将Django安装到将django子目录移动，复制，或者链接到Python下的site-packages目录里去。

• 将django子目录添加到PYTHONPATH里去。

• 把这个新目录添加到你的PYTHONPATH里去。例如，如果你把trunk检出到/home/username/.下的话，你应该把/home/username/django\_trunk（而不是django子目录）添加到PYTHONPATH里去。

Apache Web服务器和它的mod\_python模块一直是Django站点推荐的部署方式。这也是开发Django的Lawrence团队在他们自己的网站上所采用的组合，时至今日，这仍然是经受过最完整的测试以及文档最丰富的部署方案。

如果你的情况不允许使用mod\_python的话（比如在一个共享的hosting环境中，或者是一个非Apache的服务器），请参考下面一节。但是如果你有服务器（或是虚拟服务器实例）的控制权，或者是有稳定的mod\_python支持的话，这仍然是最稳妥的做法。

Django安装后的设置要考慮两个主要的问题，当通过mod\_python部署Django时，在哪里连接制版本。

Windows用户分别可以在http://httpd.apache.org和http://www.modpython.org找到编译好的二进制包。如果Mac 10.4之前的用户就需要用MacPorts来更新（只有Leopard才支持Apache 2）。变化，比如Mac 10.4以上的用户就需要用MacPorts来更新（只有Leopard才支持Apache 2）。分支开来）和mod\_python 3.0版本以上。和Python的情况类似，具体情况根据平台的不同会有所不同。

Django 需要Apache 2.0或2.2以上（有的软件包管理器用apache2来和更古老的apache1.3区别开来的）和mod\_python 3.0版本以上。和Python的情况类似，具体情况根据平台的不同会有所不同。

首先要注意的是Django要处理多少域名的URL——是整个站点（如www.example.com），还是其中的一个或几个部分（如 www.example.com/foo/，这里 www.example.com/或 www.example.com/bar/会由其他的如PHP或是静态HTML来处理）。我们还可能用到多个Django项目，每一个都有自己的部分。

要在Apache中挂接Django。

```
<Location "/">
 SetHandler python-program
 PythonHandler django.core.handlers.modpython
 SetEnv DJANGO_SETTINGS_MODULE mysite.settings
 PythonDebug On
</Location>
```

上面的<Location>会让Django的mysite处理器任何时候匹配文件名。如果希望 Django覆盖URL空间中的一部分，只要做出如下更改即可：

```
<Location "/foo/">
 SetHandler python-program
 PythonHandler django.core.handlers.modpython
 SetEnv DJANGO_SETTINGS_MODULE mysite.settings
 PythonDebug On
</Location>
```

片段是在<VirtualHost>里，或者也可以是你自己的apache2.conf（有些系统上是httpd.conf）里。要挂接一个Django项目需要在Apache中作出如下配置，如果不是虚拟主机的话，通常这个项目，每一个都有自己的部分。

首先是其中的一个或几个部分（如 www.example.com/foo/，这里 www.example.com/或 www.example.com/bar/会由其他的如PHP或是静态HTML来处理）。我们还可能用到多个Django项目，每一个都有自己的部分。

要在Apache中挂接Django。

Django以及在哪里处理静态文件。

安装完后的设置要考慮两个主要的问题，当通过mod\_python部署Django时，在哪里连接制版本。

Windows用户分别可以在http://httpd.apache.org和http://www.modpython.org找到编译好的二进制包。如果Mac 10.4之前的用户就需要用MacPorts来更新（只有Leopard才支持Apache 2）。变化，比如Mac 10.4以上的用户就需要用MacPorts来更新（只有Leopard才支持Apache 2）。分支开来）和mod\_python 3.0版本以上。和Python的情况类似，具体情况根据平台的不同会有所不同。

Django 需要Apache 2.0或2.2以上（有的软件包管理器用apache2来和更古老的apache1.3区别开来的）和mod\_python 3.0版本以上。和Python的情况类似，具体情况根据平台的不同会有所不同。

首先要注意的是Django要处理多少域名的URL——是整个站点（如www.example.com），还是其中的一个或几个部分（如 www.example.com/foo/，这里 www.example.com/或 www.example.com/bar/会由其他的如PHP或是静态HTML来处理）。我们还可能用到多个Django项目，每一个都有自己的部分。

要在Apache中挂接Django。

```
<Location "/">
 SetHandler python-program
 PythonHandler django.core.handlers.modpython
 SetEnv DJANGO_SETTINGS_MODULE mysite.settings
 PythonDebug On
</Location>
```

上面的<Location>会让Django的mysite处理器任何时候匹配文件名。如果希望 Django覆盖URL空间中的一部分，只要做出如下更改即可：

```
<Location "/foo/">
 SetHandler python-program
 PythonHandler django.core.handlers.modpython
 SetEnv DJANGO_SETTINGS_MODULE mysite.settings
 PythonDebug On
</Location>
```

```
</LocationMatch>
 SetHandler none
<LocationMatch "/foo/(images|css|js)/*">
 SetHandler none
</LocationMatch>
</Location>
```

个这样的“洞”，就是说你可以以为图片、CSS和JavaScript准备不同的目录，如下：

/foo/media/images/userspic.gif的请求则会转去找Apache的document root（这是由你的

/foo/media/images/userspic.gif代码可以毫不问题地处理 /foo/users/请求，但是对

```
</Location>
 SetHandler none
<Location "/foo/media/">
 SetHandler none
</Location>
```

有了这个以后，Django代码可以毫不问题地处理 /foo/users/请求，但是对

这样的“洞”，就是说你可以以为图片、CSS和JavaScript准备不同的目录，如下：

/foo/media。但現在如果把Python代码放在 /foo/，我们就需要给这些静态内容让条路出来。

/foo/media。但現在如果把Python代码放在 /foo/，我们就需要给这些静态内容让条路出来。

要做到这一点非常简单：你只要在处理Django项目的<Location>下加上另一个<Location>

来告诉Apache如果mod\_python不要处理某个特定的位置就可以了。

现在Apache已经能正确处理Django了，但是还缺了点什么：图片和JavaScript/CSS（还可以

有视频或PDF文件等任何网站可以服务的内容）。

一般来说这些文件都是放在应用程序里的某个特定的URL下，比如说你的app是 /foo/ 的话，你的图片和样式表 (stylesheet) 则可以是

Apache的信息，请参阅http://modpython.org 上的mod\_python的文档。

如你所见，使用mod\_python来管理Django代码是非常灵活的。要知道更多关于Python和

```
</Location>
 PythonDebug On
<Location "/bar/">
 PythonInterpreter barsite
 SetEnv DJANGO_SETTINGS_MODULE barsite.settings
 PythonHandler django.core.handlers.modpython
 SetHandler python-program
</Location>
```

Python解释器可以在同一个块里定义一个唯一的（但不是任意的）Python解释器说明。

当然，你得告诉mod\_python将它们在内存里分开放在，不然的话会发生未定义的行为。这

如果你想要在同一块下面添加多个Python项目，你只要定义多个这样的<Location>就可以了

```
<Location "/foo/">
 PythonInterpreter foosite
 SetEnv DJANGO_SETTINGS_MODULE foosite.settings
 PythonHandler django.core.handlers.modpython
 SetHandler python-program
</Location>
```

你需要在每一个块里定义一个唯一的（但不是任意的）Python解释器说明。

当然，你得告诉mod\_python将它们在内存里分开放在，不然的话会发生未定义的行为。这

如果你想要在同一块下面添加多个Python项目，你只要定义多个这样的<Location>就可以了

```
<Location "/>
 PythonDebug On
<Location "/foo/">
 PythonInterpreter mysite
 SetEnv DJANGO_SETTINGS_MODULE mysite.settings
 PythonHandler django.core.handlers.modpython
 SetHandler python-program
</Location>
```

你需要在每一个块里定义一个唯一的（但不是任意的）Python解释器说明。

当然，你得告诉mod\_python将它们在内存里分开放在，不然的话会发生未定义的行为。这

如果你想要在同一块下面添加多个Python项目，你只要定义多个这样的<Location>就可以了

```
<Location "/>
 PythonInterpreter barsite
 SetEnv DJANGO_SETTINGS_MODULE barsite.settings
 PythonHandler django.core.handlers.modpython
 SetHandler python-program
</Location>
```

你需要在每一个块里定义一个唯一的（但不是任意的）Python解释器说明。

当然，你得告诉mod\_python将它们在内存里分开放在，不然的话会发生未定义的行为。这

WSGI (Web Server Gateway Interface) 和mod\_wsgi是Python Web hosting技术中正在冉冉升起的一颗新星。Django对WSGI的支持相当完善，越来越多的Django程序员（还有Python爱好者）觉得它比mod\_python更好用。WSGI是一种非常灵活的协议，旨在把Python代码到任何兼容的Web服务器上，不仅限是Apache，还有IIS、Nginx（http://nginx.net）、CherryPy（http://cherrypy.org），甚至微软的IIS。

虽然WSGI技术还相当新，但是它已经能够在前面提到的所有Web服务器上工作，并且经受住了大量的Web框架的测试，包括Django，和一些流行的独立Python Web应用如Moin Moin wiki引擎和Trac软件管理系统。

mod\_wsgi的主要卖点（除了Web服务器无关这个特点）在于相比mod\_python它占用的内存更小性能却更好，所有的WSGI应用用统一的接口标准（包括Django之外的应用），以及支持守护进程模式，可以轻松地让一个WSGI进程只属于系统上的某个特定用户。

mod\_wsgi最大的缺点是（到编写本书的时候）它还没有被软件包管理系统广泛接受，所以目前则可以使用mod\_wsgi自己编译的一些非官方的Windows二进制包。

只要在http.conf里添加一行代码即可：

```
LoadModule wsgi_module /usr/lib/apache2/modules/mod_wsgi.so
```

然后在configuration块里加上：

```
Alias /media/ "/var/django/projects/myproject/media"
```

```
<Directory /var/django/projects/myproject/media">
```

```
 Order deny,allow
```

```
 Allow from all
```

```
</Directory>
```

```
WSGIScriptAlias / /var/django/projects/myproject/mod.wsgi
```

最后创建一个mod.wsgi的脚本，引用上面的最后一行：

```
import os,sys
```

```
os.environ[DJANGO_SETTINGS_MODULE] = 'myproject.settings'
```

```
sys.path.append('/var/django/projects')
```

```
import os,sys
```

我们要注意的最后一种可行的Web服务器部署方法是Python模块flup，它不仅是一种使用WSGI的桥梁（flup初级的目标），而且还支持另外一种类似的协议如FastCGI（有时简称FCGI）。和WSGI一样，FastCGI的目标是在用户的应用代码和Web服务器之间架起一座沟通的桥梁，同时它也具有以下优点：由于运行在单独的进程中，所以具有更好的潜在性能；另外以Web服务器公开的用户身份运行也增加了安全性。

就目前的情况来说，FastCGI在其hosting平台上比WSGI支持的更好，部分原因是由于它支持多种语言，另一部分则是因为它存在的时间更长。因此，如果没办法使用Apache而你的环境中又不支持WSGI的话，FastCGI或许是个很不错的选择。

Django的官方文档上（你可以在withdjango.com上找到链接）有一个非常棒的关于如何设置FastCGI之外还支持了SCGI和AJP两种额外的协议，如果你的部署要求不太常见的话，或许这会使用。

## B.4 SQL数据库

Python代码（Python和Django）已经准备就绪，Web服务器（Apache、Lighttpd等）也准备好了。下面的介绍之外，Django的官方文档里还专门给出了一些关于各种平台上常见的问题（可以在withdjango.com上找到链接）。如果你遇到数据库选择上的问题的话，可以先看看这份资料。

SQLite 是一个名副其实的“轻型”SQL数据库实现。和PostgreSQL、MySQL，以及各种商业数据库如Oracle或MS SQL等不同的是，SQLite不是以一个独立的服务器的形式运行的，它只是提供了一个访问磁盘上数据库文件的接口库。和其他复杂的“轻型”实现一样，SQLite也有优点（易用，低耗）和缺点（功能少，数据量变大时性能较差）。

所以，如果你不想在你的小站点上引入完整的数据库服务器的话，SQLite正适合你的需求（就像你在第二章中已经看到的一样）。不过一旦你过了学习阶段进入正式部署阶段后，最好还是升级成更加适合的数据库。

## SQLite

SQLite 是一个名副其实的“轻型”SQL数据库实现。和PostgreSQL、MySQL，以及各种商业数据库如Oracle或MS SQL等不同的是，SQLite不是以一个独立的服务器的形式运行的，它只是提供了一个访问磁盘上数据库文件的接口库。和其他复杂的“轻型”实现一样，SQLite也有优点（易用，低耗）和缺点（功能少，数据量变大时性能较差）。

所以，如果你不想在你的小站点上引入完整的数据库服务器的话，最好还是升级成更加适合的数据库。

MySQL。可以在withdjango.com上找到链接）。如果你遇到数据库选择上的问题的话，可以先看看这份资料。

## SQLite

```
$ createuser -P diango_user
Enter password for new role:
Password:
$ createdb -U diango_user diango_db
$ createdb -U diango_user diango_db
Shall the new role be allowed to create more roles? (y/n) n
Shall the new role be allowed to create databases? (y/n) y
Shall the new role be a superuser? (y/n) n
Enter it again:
Enter password for new role:
$ createuser -P diango_user
Enter password for new role:
$ createdb -U diango_user diango_db
$ createdb -U diango_user diango_db
Shall the new role be allowed to create more roles? (y/n) n
Shall the new role be allowed to create databases? (y/n) y
Shall the new role be a superuser? (y/n) n
Enter it again:
Enter password for new role:
$ createdb -U diango_user diango_db
$ createdb -U diango_user diango_db
$ createdb -U diango_user diango_db
```

建立一个新的数据库和用户：

得到PostgreSQL用户的用户名和密码后，你可以参考下面的例子来为你的Django项目创建一个新数据库和用户（如果安装的话）。

请参阅相关文档来确认这一点——官网上的或是你自己系统里的都行（如果你是通过软件包管理系统安装的）。用户名。已经为你创建了一个超级用户的账号——有时它是一个PostgreSQL的用户，有时则是你自己已经为自己创建了一个超级用户的账号。根据安装系统和安装方式的不同，数据库可能在PostgreSQL里创建数据库和用户十分方便。默认安装包括了独立的命令行工具如createuser和createdb，从字面上就能知道它们的作用了。根据安装系统和安装方式的不同，数据库可能在PostgreSQL里创建数据库和用户十分方便。默认安装包括了独立的命令行工具如createuser和createdb，从字面上就能知道它们的作用了。根据安装系统和安装方式的不同，数据库可能已经为你创建了一个超级用户的账号——有时它是一个PostgreSQL的用户，有时则是你自己已经为自己创建了一个超级用户的账号（如果安装的话）。

Python 要使用PostgreSQL的语言需要psycopg2的支持，最好是用第2版，有时也用psycopg2。psycopg2可以到 <http://midid.org/pub/software/psycopg/> 或是从软件包管理库里获得。这里要注意的一个小地方是Django有两个不同的数据库后台，分别对应不同版本的psycopg。小心确认你用的是正确的版本！

PostgreSQL的官网是<http://www.postgresql.org>，如果你在Windows系统上或是在软件包管理库里找不到的话，可以去<http://www.postgresql.org/fip/binary/>下载，找到最新的版本，然后选择你的平台即可。

PostgreSQL的语言是<http://www.postgresql.org>的文件，最好是用第2版，然后进

得了所有的主流平台，虽然它和MySQL比起来在共享服务器上不是那么流行。PostgreSQL的安装文件是<http://www.postgresql.org>的文件，如果你在Windows系统上或是在软件包管理库里找不到的话，可以去<http://www.postgresql.org/fip/binary/>下载，找到最新的版本（不过Django也能支持第7版）支持了所有的主流平台，虽然它和MySQL比起来在共享服务器上不是那么流行。

PostgreSQL（或者简称为“Postgres”）是一个功能完善的数据库服务器，它提供大量丰富PostgreSQL（或者简称为“Postgres”）是一个功能完善的数据库服务器，它提供大量丰富的特性，并且作为为企业领先的开源数据库之一拥有出色的声誉。Django核心团队推荐使用的就是这款数据库，并对其实现赞赏有加。PostgreSQL的安装文件是<http://www.postgresql.org>的文件，如果你在Windows系统上或是在软件包管理库里找不到的话，可以去<http://www.postgresql.org/fip/binary/>下载，找到最新的版本（不过Django也能支持第7版）支持了所有的主流平台，虽然它和MySQL比起来在共享服务器上不是那么流行。

PostgreSQL（或者简称为“Postgres”）是一个功能完善的数据库服务器，它提供大量丰富的特性，并且作为为企业领先的开源数据库之一拥有出色的声誉。Django核心团队推荐使用的就是这款数据库，并对其实现赞赏有加。PostgreSQL的安装文件是<http://www.postgresql.org>的文件，如果你在Windows系统上或是在软件包管理库里找不到的话，可以去<http://www.postgresql.org/fip/binary/>下载，找到最新的版本（不过Django也能支持第7版）支持了所有的主流平台，虽然它和MySQL比起来在共享服务器上不是那么流行。

和SQLite交互需要一个对应的Python库。如果你使用的是Python 2.5，那么它已经内置支持了SQLite模块，如果是Python 2.4或更高的版本，那么可以访问 <http://www.midid.org>（或者你的

上面的例子创建了一个带密码的PostgreSQL数据库用户django\_user并赋予它创建新数据库的权限，随后利用该用户创建了一个名为django\_db的新数据库。在完成上述步骤并将用户名和数据库名（以及用户密码）添加到settings.py里以后，你就可以使用manage.py来创建和更新数据库了。

注意例子里的那个“password”是Postgres提示我们输入数据库的密码——我们的系统用用户名也是数据库的超級用户。否则我们需要用createuser来指定超级用户的用户名，就像调用createdb时做的那样。

最后，PostgreSQL的SQL命令行程序是psql，它有很多和createdb与createuser一样的选项，比如用-U指定用户名。

MySQL是另一种主流的开源数据库服务器，最新的版本是5（也支持版本4）。虽然MySQL缺乏一些Postgres的高级功能，但是相比起来它更流行一些，这主要是因为它和常用的Web语言PHP的紧密结合。

和其他数据库不同，MySQL在内部有几种不同的数据库类型来支持不同的特性：一种是MyISAM，缺乏事务支持和外键但是可以进行全文搜索；另一种是InnoDB，它比MyISAM更新，特性更丰富，但是目前缺乏全文搜索能力。虽然还有其他的一些类型，但这两个是最常见的。

如果你在Windows上或你的软件包管理系統没有最新版MySQL的话，它的官网http://www.mysql.com 提供了大多数平台的二进制版本。Django推荐的MySQL Python库是http://www.mysql.com 提供了大多數平台的二进制版本。Django推荐的MySQL Python库是MySQLdb，可以在其官网 http://www.sourceforge.net/projects/mysql-python 下下载到，你需要的是1.2.12以上的版本。请注意这里的版本号，有些旧的Linux发行版自带了如1.2.10这样的版本，Django无法兼容这样的旧版本。

MySQL里创建数据库的工作主要是通过全能的admin工具mysqldadmin来完成。和Postgres不一样，你需要先弄清楚安装版的数据库超級用户名和密码才能为Django项目创建新用户。这个超級用户通常叫root，并且没有初始密码（最好尽快修改），所以创建数据库非常地方便：

```
$ mysql -u root
$ mysqladmin -u root create django_db
```

和Postgres不同的是，MySQL的用户完全由数据库自身管理，所以我们要用MySQL的SQL shell来创建数据库用户django\_user。这个shell的名字是mysql。

shell来创建数据库用户django\_user。这个shell的名字是mysql。

```
Welcome to the MySQL monitor. Commands end with ; or \g.
Your MySQL connection id is 6
Server version: 5.0.51a-6 (Debian)

 Read this feature to get a quicker startup with -A
 You can turn off this feature to get a complete list of table and column names
 Redacting table information for completion of table and column names
 A
$ mysql -u root
```

到这里，我们应该已经完整地配置运行了Django及其每个组成部分——Python, Django, Web服务器和数据库。如果你遇到了任何问题，请回忆一下我们在第二章里分章的建议——先备份你的工作，重算所有的东西，确保你没有遗漏文档中的任何一步。

最后，Django自己的安装文档（不仅有Python库，还包含了所有的组件）也是一份非常出色的资料，你可以在withdjangocom上找到链接。

## B.5 总结

Django的数据库支持还在不断的进化中。获得最新信息的最佳来源是官方网站和Django的用户列表。Django目前没有直接支持的两个数据库是微软的SQL Server以及IBM的DB2。Google Code上有一个独立维护的项目为Django提供了对MS SQL的支持：<http://code.google.com/p/django-mssql/>。另外Google Code上还有一个支持DB2的项目Python-DB2：<http://code.google.com/p/django-db2/>。在编写本书的时候，这个项目对Django还不是即插即用的。

### 其他数据库

连接Oracle的Python库是cx-oracle，可以从<http://cx-oracle.sourceforge.net/>获得。常见的，也无法通过Linux软件包管理器练习到。不过，Oracle是有免费版的，Django可以支持！一个数据库驱动开发的新手的话，基本上你不用考虑这个选择，因为Oracle在共享主机上不太常见，也无法通过Linux软件包管理器练习到。不过，Oracle是有免费版的，Django可以支持！

在编写本书时Django目前支持的最后一个数据库是一个商业数据库，Oracle。如果你还是希望通过以下几个步骤，你现在就可以修改settings.py的设置并开始使用manage.py执行数据库相关命令了。

Oracle

```
mysq> GRANT ALL PRIVILEGES ON djangodb.* TO 'djangouser'@'localhost' IDENTIFIED BY 'djangopass';
Query OK, 0 rows affected (0.00 sec)
```

则完全相反，保持主干稳定，把所有的新特性放到分支里去，这样做的好处是可以同时开发好主干上，用分支来表示软件发布（release），从主干分出来后，分支只接受bug修正。另一种方式一种做法是把所有新特性的开发（即会打破向后兼容性或是引入不稳定性的发展）放在主干上，如现在主干和分支上分配任务则取决于项目本身。

版本控制系统往往看起来像一棵树，其中开发的主线被称为主干（trunk），而在之上所进行的复制（于是变成了独立的实体）会按照它们自己的方向发展下去。这些拷贝就是分支（branch），分支在主干和分支上分配任务则取决于项目本身。

## 主干和分支

随着软件开发的不断进步，特别是在开源社区里，大家都意识到这种系统很不适合分布式团队开发，所以就出现了更现代的版本控制系统，例如RCS的改进型和分支型分支，CVS（Concurrent Versions System），以及后来的Subversion项目，Subversion希望成为一个“更好的CVS”并最终取代之。

早期的版本控制系统如SCCS（Source Code Control System）和RCS（Revision Control System）相当简陋，要么是维护一份原始版本的文件加上些许修改（不同版本之间文件的微小改动），要么是反过来——维护一份最新的版本，然后做去新的改动。这类系统的一个主要问题是被控制的文件和修改都在同一个服务器上。

回到一个星期前的那个看似无辜的改动发生前的一个小时。系统为你的项目保留了一个完整的修订历史，允许你将代码回溯到历史上的任何一点（比如，如果你在开发任何形式的软件，而又没有使用版本控制系统的话，就实在太落后了）。版本控制系统相当简陋，要么是维护一份原始版本的文件加上些许修改（不同版本之间文件的微小改动），要么是反过来——维护一份最新的版本，然后做去新的改动。这类系统的一个主要问题是被控制的文件和修改都在同一个服务器上。

## C.1 版本控制

如果你是来自设计或别的领域，这里介绍的内容对你来说可能是全新的体验。好消息是，我们保证只要你花功夫学习掌握这些工具，就可以成倍地提高生产力以及获得极大的灵活性，绝对能让你省不少心。

如果你是从软件开发转入Web开发来的話，那对你要使用现有的软件开发工具和技术都将是电子头上的虱子——明摆着的。这些工具具有版本控制，bug追踪，以及强大的开发环境等。如果是这样的話，你只需要快速浏览一下本书的内容就行了，确认没漏掉任何能帮助你提高工作效率的东西。

再小规模的Web开发也是软件开发。只要你的站点接受某种形式的用户输入并且处理这些输入，它就算是一个Web应用。它和（或者说应该和）其他任何类型软件的开发都是相似的。

## 附录C 实用Django开发工具

几个庞大的新特性。

## 合并不

Djanggo自己用的是一种介于上面两者之间的方法：它同时维护了两条分支，一条是发布，另一条是特性，而主干则保持折衷——既不是完全稳定的也不是完全不稳定的。会产生严重影响稳定的改动会在分支里完成，而小修小补则直接在主干上进行。

## 分散型的版本控制

分散型版本控制(decentralized version control)是未来的趋势所在。一个分散式系统可以完成集中式系统所能做到的一切，而且还具有更强大的特性，即项目的每个“checkout”自身也是一个完整的仓库。在这种分布式版本控制之下，你不再需要连接到中心服务器来记录每一组改动。本地仓库就能帮你记住它们。只有在需要的时候，它们才会通过网络传递出去/进来。

开源的分散式版本控制系统有Git, Mercurial, Darcs, Bazaar, SVK和Monotone。商业系统则包括Biketeer和TeamWave。

Mercurial (<http://www.selenic.com/Mercurial/>) 是今天最流行的分布式版本控制系统之一。Mercurial主要是用Python编写的，所以很多大型的项目都采用了Mercurial，其中包括Mozilla浏览器和Sun的OpenSolaris系统。

Git (<http://git.or.cz>) 算是Mercurial的竞争对手，它们和Bazaar一起瓜分了开源DVC(分布式版本控制系统)的份额。Linux Torvalds和其他Linux内核团队的成员一起用C编写了这个软件。Git是一个相当符合Unix精神的工具，原本是创造出帮助处理Linux内核项目这种极端复杂的源码控制系统。Ruby on Rails等其他很多Rails影响下的项目，以及WINB项目和X.org (Linux的图形系统)，Fedora Linux等都不约而同地使用了Git作为其版本控制系统。

以下为不知道怎么在Django项目里使用版本控制的读者提供了一些基本指导。在这里我们用的是Mercurial (hg命令)。

### 为你的项目添加版本控制

首先我们建立一个Django项目框架。

```
$ hg init
$ hg add .
$ hg commit -m "Initial Django project"
$ hg serve
```

现在把这个工作目录添加到仓库里去。

接着添加一个基本的hgignore文件(你可以像下面那样用echo命令，也可以用文本编辑器)，告诉Mercurial略过那些后缀名是.pyc的字节码文件，这些文件都是由Python的字节码编译器自动生成的，我们没必要跟踪它们的历史。

```
$ echo "\.pyc" > .hgignore
```

如此这般有条不紊地继续进行项目。修改代码，在开发服务器上测试，用简单明了的提交到那一章上去就可以了。

当记录下一个改动集时，它会为项目在当时的“快照”做保存下来。如果你发现在改动集0所作的改动之后完全把事情搞砸了的话，只需要用命令 `hg revert -r 0` 将工作目录回滚

现在我们来做一个小小修改，看看发生了什么，然后提交它。  
里的ID号)。

每次向 Mercurial 仓库提交代码时，Mercurial 都会用两个不同的数字来标记一个改动集（change set）：一个只在本仓库里有效的自增整数和一个十六进制的哈希数（标记了它在主仓库

```
$ hg commit -m "Initial version of my project"
No username found, using pbxexample.org instead
$ hg log
changeset: 0:e991df3d3205
tag: tip
user: pbxexample.org
date: Sun Oct 07 13:49:14 2008 -0400
summary: Initial version of my project
```

然后，用`git commit`命令提交修改，并输入`Mercurial`在`hg log`里有地它们记录下来。

在默认情况下，`hgiignore`文件里的模式都是由正则表达式组成。接着我们来添加文件：

有相同的內容和歷史。

在live站点上也可以用相同的hg revert命令來放棄之前的改动——比如我們發現剛剛加入的改動虽然通過了dev站点的測試，却在live站点上產生了预料之外的問題。项目的每一份克隆都在

```
1 files updated, 0 files merged, 0 files removed, 0 files unresolved
added 1 changesets with 1 changes to 1 files
adding file changes
adding manifests
adding changesets
searching for changes
pulling from /stuff-dev-site
$ hg pull -u
$ cd .. /stuff-live-site
```

些改動加進來即可：

當充分測試這些修改並準備好部署時，我們只要轉到“live”分支下，执行以下命令把那

```
$ hg commit -m "Also, dogs are liked."
$ echo "I LIKE DOGS = True" >> settings.py
$ cd stuff-dev-site
```

現在我們在原始（開發）目錄上作一點修改。

以這些文件並沒有被加入到仓库里去，所以當然也就不會顯示在這裏了。  
等一下，那些py文件都到哪裏去了？因為前面我們告訴過Mercurial讓其忽略掉它们，所

```
$ ls stuff-live-site/
--init__.py manage.py settings.py stuff-app urls.py
```

這就會對原始的仓库生成一份和工作目錄一模一樣的拷貝（clone）。

```
8 files updated, 0 files merged, 0 files removed, 0 files unresolved
$ hg clone stuff-dev-site stuff-live-site
$ cd ..
```

一份部署版本的拷貝，那么相應的命令就是：

原始（開發）目錄的路徑就行了。比方說，我們現在在原始目錄下，希望在同一級目錄下創建一來事情就變得相當簡單，只需要從前工作目錄轉到要創建拷貝的地方，然後輸入hg clone和

為了簡單起見，我們假設部署到網站上的拷貝和開發中的拷貝都在同一台服務器上。這樣

更好的办法是直接从仓库里复制一份拷贝出来，同时让Mercurial记住拷贝的出处，因为这样只要遠端地获取任何在这份拷贝上的更新就可以了。这正是hg clone命令所完成的工作。

你可以把所有的文件打包起来，复制到要部署的服务器上，然后解压文件就行了。如果你是大堆巨難像，将來修正和修復這個項目的可能性非常低，這種重複打包—复制—解包的过程从此不再对这个应用做任何修改的话，这种方式也算不上很糟糕。但事实你开发的是软件，不

信息记录修改。接着就要准备部署你的项目了。

向Python官方网站上的Emacs页面 (<http://www.python.org/emacs/>)。以上都内置了python-mode。对于旧一点的版本，或是Emacs的其他类型如XEmacs等，可以访问Django项目的Django程序员，你手中最重要的武器就是python-mode，它具有语法高亮，自动对齐，以及其他很多提高生产力和方便编写Python代码的功能。Emacs第22版作为第一个高效使用Emacs的Django程序员，你手中最重要的武器就是python-mode。

## Emacs

是一些常见编辑器里的小技巧。编辑Django项目不需要任何特殊的应用。任何程序员的编辑器都可以胜任这一工作。以下

### C.3 文本编辑器

你需要扩展Trac的话，可以先看看 <http://trac-hacks.org/> 上有没有适合的插件和“hack”技巧。另外，虽然Trac默认支持的是Subversion，但是也可以通过插件让Trac支持其他系统（比如前面提到的那些）。你可以在 <http://trac.edgewall.org/> 下载到Trac以及阅读它的完整文档。如果你需要扩展Trac的话，可以先看看 <http://trac-hacks.org/> 上有没有适合的插件和“hack”技巧。

Trac是编写本书过程中的主要工具——更多信息请参考本书扉页。

#### 注意

要是在这些还不够的话，你要知道Django代码仓库本身 <http://code.djangoproject.com/> 也是运行在Trac之上的。（只不过由于那些漂亮的自定义风格，所以你没有认出来吧。）还有一些常用的wiki标记可以用来轻松地链接源码版本，bug记录和文档笔记等内容。它无须太多设置即可完美运行，特别是能和Subversion等版本控制系统很好地集成在一起。它们通常大多数时候即可以完美运行；不过Trac是我们在项目管理软件领域里最喜欢的一款产品。我们承认这里有一点偏激：不由Edgewall Software公司负责维护。Trac是一个开源的wiki，bug跟踪，源码和项目管理系统，它由Edgewall Software公司负责

## Trac

区里用的最多的一个就是Trac，它也是用Python编写完成的。这样的软件市面上有很多，其中大多是以服务的形式出现，例如Google Code，SourceForge，Launchpad和Basecamp。其他的则是有一些可以自行安装的独立应用程序。开源社区还用到了一种经常和版本控制系统“搭档”的Web应用程序，它不但提供了源码仓库及其历史的Web界面，同时还能跟踪问题，TODO条目，以及文档等。

用版本控制系统管理代码很有用，不过有些程序员却对此为止了。其实很多其他程序员甚至在Mercurial，Mercurial等系统的手册，请参见 [withdjangocom](http://withdjangocom)。

### C.2 项目管理软件

这里介绍的只是一些皮毛，不过希望至少能让你有一点概念。当然了，版本控制里，特别是在Mercurial里，要学会的东西比这里展示的还要多很多。更详细的信息，包括免费的Subversion，Mercurial等系统的手册，请参见 [withdjangocom](http://withdjangocom)。

Vim 编辑器来自UNIX下的工具vi。它大大增强和改进了原有的工具。你可以在vim的主页<http://www.vim.org/>上下载到它。和Emacs一样，Vim也有Python语法mode。Djangowiki上可以找到一个专门介绍如何在Vim（以及它的变型）里使用Python的页面。

### Vim

另外你还可以在Django的wiki上找到为Django模板所准备的第三方mode。<http://code.djangoproject.com/wiki/Emacs>。

TextMate是OS X下非常流行的商业编辑器，它对Django的支持非常出色。TextMate把它的某种语言和语法的支持组织成一个个“插件”（bundle），其中的Python插件能大大加速Python代码的编写，并通过代码着色令它更容易阅读。此外，在TextMate公共的插件仓库里还有两个专为Django准备的插件：一个是为Python准备的，另一个则是给Django模板准备的。更多的信息可以在Djangowiki上找到 (<http://code.djangoproject.com/wiki/TextMate>)。

### TextMate

Eclipse IDE提供了一个强大的Python开发模块PyDev。你可以从它在SourceForge上的页面里获取更多的信息和代码 (<http://pydev.sourceforge.net/>)。

- 看看更新的链接和推荐吧。
- 由于Djangosites.org社区越来越大，寻找Django应用的方法也越来越多。经常到withdjango.com 等Django应用程序，并随着Git变得越来越好。
  - Github.com：这是一个使用Git版本控制系统存储源码仓库和社交媒体站点，它也包含了一些开放了源码的项目。
  - Djangostatic.org：这个简单的Django站点有一个“sites with source”类别可以選擇只显示关于Django应用程序的信息。
  - Djangoplugins.com：在编写本书的时候，这还是一个相对较小的站点，它收集了各种插件的地方。在项目查找到Django就能得到一份长长的列表。
  - Google Code：免费的站点和干净的界面，让Google Code迅速成为放置Django项目之展示给全世界，它的连接为 <http://code.djangoproject.com/wiki/DjangoResources>。
  - Django项目wiki上的DjangoResources页面：很多应用程序的作者选择把作品放到这里来询问量，这里我们列出几个流行的网站。

## D.1 到哪里去找应用程序

无论你是简单到用户注册这种功能，还是复杂的论坛系统，或者你经常会发现他们把Django应用放在自己的网站和项目管理系统，因此你经常会发现他们把Django应用放在blog或者是个人Trac系统里。不过，更多的程序员会选择将项目放到集中的列表网站上以期获得更多的访问量，这里我们列出几个流行的网站。

有些程序员拥有自己的网站和项目管理系统，因此你经常会发现他们把Django应用放在博客里，以及是否可以尽量简单地使用（并持续更新）它？

毫无疑问，怎么判断它的质量？以及是否可以尽量简单地使用（并持续更新）它？

毫无疑问，如果一个开源的Django应用可供你使用了。但是要怎么找电子商店解决方案，基本上一定已经有一个开源的Django应用可供你使用了。但是要怎么找MIT风格的协议对商业应用更友好，在你作出选择时必须将协议的问题作为考量之一。

BSD/MIT风格的协议对商业应用更友好。这里没有任何对两种协议品头论足的意思，只不过事实上支持这一点，但却不是必须的。通过在网络上发布（相比只能是通过网络提供服务）里的代码，这个应用程序的代码也必须公开。BSD/MIT风格的协议虽然内部接受。GPL要求如果一个基于GPL协议的产品用在一个会被重新发布的应用程序（相比只是这些应用因此更容易被不方便使用大多数开源协议（GNU Public License或GPL）的组织来说，大多数这些应用都沿袭着Django的脚步采用了限制很少的BSD/MIT风格的版权协议。作为榜样，它树立了一个典范，即有问题的时候，不妨将应用源码开放。

编写Django应用是一件很有意思的事情，但是有时候没必要重新编写。随着框架越来越流行，一系列开源的Django应用如雨后春笋一般出现。这得益于Django自己就是一个开源项目，

## 附录D 发现、评估、使用Django应用程序

## D.2 如何评估应用程序

- 下面这些问题是你找到一个预期的应用和项目时要确认的问题。
- 它是不是能活跃？最近一次的特性更新或bug修正是多少之前？虽然不是每个项目都需要一周更新好几次，但是如果你一个项目最后一次可见的活动是一年前的话基本上可以认为它已经被抛弃了。任何pre-1.0版之前的Django项目都需要及时更新，否则很快它就没什么人关心了。“注意：这一节需要更新！”的警告标志是不是有定期维护？有没有很多有用的回复的话，这就是一个好迹象。
  - 作者是谁？Google一下作者的名字看看他们的经验水平如何，以及他们的工作和社区的贡献。如果你能在项目的邮件列表里得到许多有用的回复的话，这就是一个好迹象。
  - 代码质量如何？如果你是一名有经验的Python程序员，感受项目最好的办法之一就是下载然后读一下它的源码。源文件组织得怎么样？函数和方法里有没有很好的docstring解释它们的意思和用法？有没有测试，而且是不是都通过了？
  - 代码量如何？如果你是一个特定个人或项目的某个性定，你可能会问：“有多少社区支持？”很多Django项目一开始只是为了解决某个人或项目的某些特定需求，然后慢慢培养出自己的用户群和开发社区。应用组件越是复杂，其背后越是可能有一个由资深用户组成的社区提供帮助。当然，不是说每个应用组件背后都一定要有一个活跃的社区才算可用的，但是社会基础也可以在一定程度上反映出应用组件的复杂程度。
  - 第三方的（包括你的）Django应用组件，其实就是Python模块。要在个项目里使用它们，只需要在你项目的settings.py文件中的INSTALLED\_APPS设置里加上一个包含应用组件的路径。第三方的（包括你的）Django应用组件，其实就是Python模块。要在个项目里使用它们，你需要把它们放到任何地方，不过一般来说有三个可选项：
    - 内嵌：如果只是拿它作为一个单独的项目，你可以选择把它加到你项目的目录下和其他应用组件放在一起。可以说这是最简单的办法。缺点是当要在同一个服务器的其他项目里使用它时，你得粘贴复制。
    - 分享一个“共享应用”目录：另一个选择是为这些共享应用单独建立一个目录（比如shared\_apps）然后把它添加到Python路径里。这样所有第三方的代码就部署在同一个地方，在跨项目导入时十分方便。要把应用组件添加到一个给定项目时，只要把它的名字（无需额外前缀）添加到项目的INSTALLED\_APPS设置里即可。
    - 安装到你的site-packages目录：你还可以把Django应用添加到系统范围中的Python库里，这通常是一个名为site-packages的目录。（在Python提示符下输入import sys; sys.path即可以显示你系统里的路径，以及Python路径里的目录。）

如果你要跟踪一个不断演化的项目，你可以选择直接从项目的版本控制系统里获取代码而不是下载打包文件。这样的话，根据你选用的不同方式，只需要把最新的代码检出（check out）到你的项目、共享应用目录，或是系统site-packages下就行了。记住在更新这类型外部应用的时候要格外小心，不要破坏你编译它们的项目。

#### D.4 分享你自己的应用程序

作为一个不断进步的Django应用程序员，总有一天你会发现你的一些成果可以用来帮助别人。我们强烈推荐你在这种情况下愿意用开源协议开放这些应用的源码，让其他Django用户能使用并改进它。Google Code, SourceForge, GitHub以及其他的应用托管服务都可以让你轻松地和其他人分享代码（还不用自己创建一个新的网站来管理）。到时候记得告诉我们你的成果哦！

App Engine应用程序的配置被定义为一个YAML文件(`app.yaml`)，它看起来如下：  
当你的代码准备好发布时，就可以用它把应用部署到App Engine的“云”上去。  
执行开发：这包括SDK的开发服务器(`dev-appserver.py`)以及应用部署上线程序(`appcfg.py`)。纯Google App Engine应用程序一般创建保存在单个目录下，用App Engine提供的工具包进

## E.2 纯Google App Engine应用程序

这和PHP的情况不同（虽然PHP自身也有很多问题，但至少语言本身不是），通常它不用担心这个类的问题。App Engine的出现在Python的Web工程师也能享受到这种好处，即不必再关心编程平台的问题了。而且，它还让我们能利用Google现成的（巨型的）基础设施。

让你能运行Python，我们别无选择。

们和Web服务器设置打交道（部署Apache或nginx，是用mod\_python还是FCGI等），是为了要可能App Engine作出的最重要的事情就是它让部署和服务器维护不会再成为一个问题。我们对它们的兴趣。

的软件开发者社区对Python和Django进行了一次很好的宣传，引起了很多以前从未考虑过这些 Google App Engine对那些正在使用和打算使用Django的人来说绝对是一个利器。它广泛的

## E.1 为什么App Engine如此重要

我们主要关心的是Django——为Google App Engine自身开发应用可以写成另一本书了！  
们不必须介绍。

如果你希望了解更多关于“纯”App Engine应用的内容，请参考最后一给出的相关链接，这里我们在下一篇文章讨论这个。另外两个方式引入了更多的Django特性，这才是本附录主要关心的内容。

第一个方式（最少限度的）使用每个App Engine应用所提供的Django特性和组件——我们  
• 导入为App Engine编写新的Django应用  
• 将一个现有的Django应用移植到App Engine上去  
• 全新的纯App Engine应用  
这里根据包含“Django元素”的多少，可以分出几种为App Engine开发应用的方法：

是要详细了解这一技术的方方面面是不现实的。你可以在本书的最后一章找到更多关于这些信息的  
于Django且包含很多Django特性的可扩展Web应用平台。虽然我们会尽量覆盖所有的基础，但  
这篇文章将为你介绍如何把你的Django应用移植到Google App Engine上去，这是一个部分基  
链接。

## 附录E 在Google App Engine上使用Django

首先，我们要下载必要的软件。你可以在SDK项目的主页 <http://code.google.com/p/>

## 索取SDK和帮助程序

Google App Engine Helper for Django 是让你感觉使用 App Engine 真正”在进行 Django 开发的关键所在。这是一个由 Google 资助的开源项目（Python 2 及 Guido van Rossum 也是作者之一），旨在为那些有经验的 Django 工程师在 App Engine 上提供一个更熟悉的环境。你甚至可以用它把 App Engine 脚带的 Django 替换成最新的版本。

### E.4 Google App Engine Helper for Django

0.96.1 版。

在本书编写的时候，App Engine 提供的 Django 组件是有缺点的不过还算稳定的 Django  
组件

编写同时也能够部署在普通服务器和 App Engine 上的应用，这些组件就完全不够了。

除去 Django 的这些部分，剩下的就是其核心功能了：URLconf，视图和模板。虽然用这些组件也足够构建起任何类型的网站了，但是如果你在 App Engine 上使用现有的 Django 应用或是

在一节里，我们会讨论如何为 App Engine 改写你的应用程序。

比如 admin，认证系统，通用视图等。

以重写应用程序里相关的部分来绕过它，但是还有很多其他的 Django 应用组件是你不想碰的，全新的应用程序，但是这一限制已经基本上把你所有的 Django 应用排除出去了。虽然你也可以用 Django 的其他组件并且把数据模型替换成基于 Google BigTable 的 ORM 来编写

SQL 语句，没有关系，也没有 JOIN。

com/papers/bigtable.html 和 <http://en.wikipedia.org/wiki/BigTable> 找到更详细的说明。这里没有 Google 依赖的是它自己的 BigTable 在键索统而非关系数据库——你可以在 <http://labs.google.com/papers/bigtable.html> 和 <http://en.wikipedia.org/wiki/BigTable> 找到更详细的说明。这里没有

从 Django 工具的角度来说，Google 的实现里缺少的最重要一环就是 Django 的 ORM。

### E.3 App Engine 架构的局限性

完全照搬了 Django 的实现。

你可以看到这里的 handler——和 Django 的 URLconf 作用相似，它将一个类 URL 的字符串和相对应要执行的脚本挂钩。App Engine 开发里的另一个和 Django 类似的特性就是它的模板系统，它

```
application: helloworld
version: 1
runtime: python
api_version: 1
handlers:
 - url: /*
 script: helloworld.py
```

```
application: helloworld
version: 1
runtime: python
api_version: 1
handlers:
 - url: /*
 script: helloworld.py
```

在本书编写的时候，帮助程序（helper）还处于初级的阶段。它的主要目标是消除Django对APP Engine的陌生感，要完成这一点还有很多工作要做。如果想看它的源码的话，你得准备好好学习一点APP Engine和Django内部的原理。

这个帮助程序不是要把它APP Engine变成Django，它的作用是让过渡更加平滑。虽然下面的例子假没使用到帮助程序，不过直接使用APP Engine的功能也是完全可以的。就像我们谈论Django时常说的，“它其实就是Python而已”。不要因为APP Engine复杂的表象就停止你探索的脚步。有人说，APP Engine“想要不扩展都难”。如果你愿意为部署的便捷性及其潜在的无限可能付出一点点代价（就是前面提到的那些限制）的话，请往下读。

帮助程序以一个Django项目框架（skeletal Django project）的形式发布，它包含了一个叫appengine\_django的应用。你可以在这个骨架上编写项目，或是把这个appengine\_django应用就可以上通过URL：`http://localhost:8000/_ah/admin`（假设你的APP Engine运行在8000端口上）来访问它。和普通的admin应用不同，这个工具只能用于数据库里已经有记录的数据（model）。如果你的数据还没有保存在任何数据的话，就不能通过Development Console来创建新数据。

## E.5 建成APP Engine

在这一节里，我们要把第2章里开发的那个简单的blog移植到APP Engine上去。我们打算遵循Google App Engine Helper for Django的README文件里列出的步骤来进行。这个文件可以在<http://code.google.com/p/google-app-engine-django/>找到；<http://code.google.com/p/google-app-engine-django/source/browse/trunk/README>。

在下面的例子中，我们假设项目的名字是mysite，应用程序的名字是blog。

解压下载的帮助程序包后，你会得到一个名为appengine\_helper\_for\_django的目录。把目

## 把APP Engine代码复制到项目里来

在这一节里，我们要把第2章里开发的那个简单的blog移植到APP Engine上去。我们打算遵循Google App Engine Helper for Django的README文件里列出的步骤来进行。这个文件可以在<http://code.google.com/p/google-app-engine-django/>找到；<http://code.google.com/p/google-app-engine-django/source/browse/trunk/README>。

在这一节里，我们想要把那个简单的blog移植到APP Engine上去。我们打算

不兼容的Django功能——我们之前描述的所有功能。方法是执行diffsettings命令。

备份原始的项目，在必要时还可以退回去！为什么呢？因为我们不得不丢掉所有和APP Engine相关的部分。有了这个“新的”manage.py后，我们就可以用它来生成一个新的settings.py文件。建议先来访问帮助程序的插件。不过，就算已经改过的manage.py，也能继续执行很多重要的命令。

这样APP Engine就部分地实现了Django的管理命令，并且可以加入APP Engine相关代码

```
from django.core.management import execute_manager
```

```
from appengine_django import installAppengineHelperForDjango()
installAppengineHelperForDjango()
```

```
#!/usr/bin/env python
```

Django应用一样管理我们的应用程序。方法是在文件开头的地方加上两行代码（指定shell的注释之后，导入execute\_manager的代码之前），所以文件的前四行应该如下所示：

下一步是要把帮助程序集成到负责指挥控制的manage.py里去。这样就可以和管理独立的

## 集成APP Engine Helper

```
$ ln -s $THE_PATH_TO/google_appengine ./google_appengine
链接APP Engine代码。在POSIX系统上（Linux或Mac OS X），可以调用命令：
```

如果你安装APP Engine SDK时用的不是Windows或是Mac OS X的安装程序，那你就得手动

不用填的。

这里把app.yaml或APP Engine里注册的名字实际上不是必须的，只有在你决定要把代码上传到APP Engine的时候才需要。当代码运行在SDK的开发服务器上时，这个信息是

注意

现在的Django文件urls.py和settings.py都在一个目录下。

同样的Django文件urls.py和settings.py都在APP Engine的Admin Console下注册的应用程序名。下一步就是要允许你的项目访问APP Engine自身的代码。

现在编辑app.yaml文件，把应用名字改为你在APP Engine的Admin Console下注册的应用

里的内容复制到Django blog项目的主目录下，即app.yaml，main.py以及appengine\_django目录

```
INFO:root:Server: appengine.google.com
$ manage.py runserver
```

现在是时候来测试一下这个翻新的应用程序了。当运行 manage.py runserver 启动开发服务器时，你不会再看到 Django 的验证码模型，它也不会再告诉你正在使用的是哪个设置文件或是展示友好的启动信息。取而代之的是 App Engine 服务器的输出。

## 测试一下

由于 App Engine 没有使用 Django ORM，所以你必须把视图里的代码改成 BigTable 查询。App Engine 数据库 API 提供了两种不同的查询方式：标准 Query 或是类 SQL 的 GqlQuery。因为我们要的是 Django ORM 并且希望尽量保留原有的数据访问，所以这里采用的是 Query。

如果 blog 正文超过这个长度的话，你可以改用 TextProperty。TextProperty 可以储存超过 500 字节的内容，缺点是它不能被索引，也不能过滤和排序。好在通常你也不会去过滤或是排序 body 属性，所以问题也不是很大。

这里的代码是绝对相似的——这些对象和之前用到的对象基本是等价的。一个显著的区别是我们不再需要为 title 设置 max\_length。App Engine 的 StringProperty 最大支持 500 字节。简单起见，这里 body 也用一样的属性 (property)。

```
class BlogPost(BaseModel):
 title = db.StringProperty()
 body = db.TextProperty()
 timestamp = db.DateTimeProperty()
```

```
from appengine_django.models import BaseModel
from appengine_django.models import db
```

修改 models.py 如下：

现在我们要调整一下应用程来使用 App Engine 对象，首先是 models.py 文件，我们需要用 appengine\_django.models.BaseModel 来替换 django.db.models.Model。备份你的模型文件，然后

## 将应用程移植到 App Engine 里去

在输出中所有 “WARNING” 打头的行经常是指那些被移除的不兼容的内容或是不正确的设置。所有跟在 WARNING 和 INFO 之后的行则组成了新 settings.py 文件的内容。

```
INSTALLED_APPS = ('django.contrib.auth', 'django.contrib.admin', 'mysite.blog',
 'appengine_django')
MIDDLEWARE_CLASSES = ('django.middleware.AuthenticationMiddleware',
 'django.contrib.auth.middleware.AuthenticationMiddleware',
 'common.middleware.CommonMiddleware',
 'appengine_django.middleware.AppEngineMiddleware')
SECRET_KEY = 'w*sb(p($wzxra*a9_64_z0s(qi(9x3(w-ribbaaa4(z~w1
SERIALIZATION_MODULES = {'xml': 'appengine_django.serializer.xml'}, '#'
SETTINGS_MODULE = 'settings', #'
SITE_ID = 1 #'
TEMPLATE_DEBUG = True
TIME_ZONE = 'America/Los_Angeles'
ROOT_URLCONF = 'mysite.urls', #'
```

App Engine的DateTimeProperty能识别datetime.datetime对象，所以我们可以通过直接导入使用

```
In [1]: from blog.models import BlogPost
In [2]: entry = BlogPosts()
In [3]: entry.title = 'last blog entry'
In [4]: entry.body = 'this is my last blog post EVER!!'
In [5]: from datetime import datetime
In [6]: entry.timestamp = datetime.today()

Help --> Python's own help system.
object? --> Details about object. Object also works, ?? prints more.
? --> Introduction and overview of IPython's features.
ipython? --> Quick reference.

IPython 0.8.2 - An enhanced Interactive Python.

Python 2.5.1 (r251:54863, Mar 7 2008, 04:10:12)
Type "copyright", "credits" or "license" for more information.

$ manage.py shell
```

之前说过，App Engine里不能使用Django的admin应用。所以我们自己手动添加第一篇blog帖子（通过Python shell），让App Engine的数据输入机制意识到这个特定模型（model）的存在。在台机器上我们有安装IPython，所以可以看到它的启动信息和提示符。

## 添加数据

图E.1 还没有文章的“新”blog



这里当然还没有显示任何帖子，因为你要用的是App Engine的存储而不是原来Django ORM的数据

打开浏览器窗口输入`http://localhost:8000/blog`。你应该可以看到熟悉的blog页面，如图E.1。

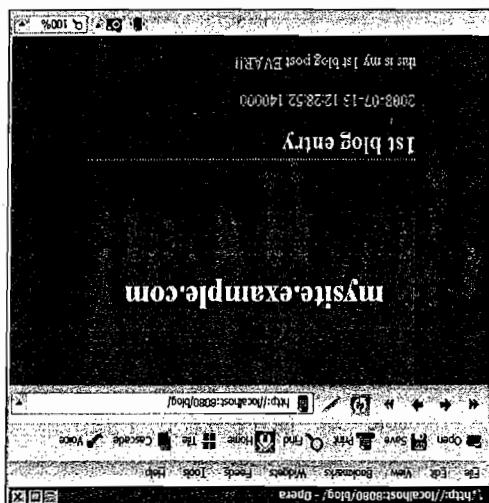
```
INFO:root:Running application mysite on port 8000: http://localhost:8000
```

上一节的经验，下面这些在App Engine上创建全新的Django应用的步骤是很容易的：把现有的Django应用移植到App Engine实际上是一件相对麻烦的工作。比较起来其实大多数时候新建一个应用会简单一点——因为你不用“负担”一个现成的Django应用。如果你有了以上的经验，下面这些在App Engine上创建全新的Django应用的步骤是很容易的。

## E.6 为App Engine创建一个全新的Django应用

希望这段小小 的例子能让你体会到 Django 的威力。当然了，同时你也就看到了无论网站的状态如何，Python shell 在操作数据时都是十分强大的。

图E.2 我们的blog有新文章了！



算帖子了，如图E.2所示。

确认你的服务器还在运行中，刷新页面 `http://localhost:8000/blog`。你现在应该可以看到那新增加的数据可以从此shell里直接访问。剩下的就是将它在应用服务器里显示出来了。`http://code.google.com/appengine/docs/datasource/creatinggettinganddeletingdata.html` 就像你看到的一样，有对应用类型的实体的查询”。（这段话引自数据库文件，你可以在那里找到它：`http://code.google.com/appengine/docs/datasource/CreatingGettingAndDeletingData.html`）就像你看到的一样，`BlogPost`类的`all()`方法返回了一个Query对象，它表示了一个取得所有数据库API文件说到，“Model类的`all()`方法返回了一个Query对象，它表示了一个取得所

```
list blog entry : this is my list blog post EVER!! (2008-07-13 12:28:52,140000)
.....
..... print post.title, ':', post.body, '(s)', & post.timestamp
in [9]: for post in query:
in [9]: print post.title, ':', post.body, '(s)', & post.timestamp
in [8]: query = BlogPost.all()
out[7]: datasotre_types.Key.from_path('BlogPost', 1, _app=u'mysite')
in [7]: entry.put()

In [7]: entry.put()
```

保存对象就行了。

它。然后调用today()函数来获取当前的日期和时间并赋值给timestamp，接着就只需要用put方法

- 最重要的事是你作为一名Django程序员可以在继续编写Django的同时，享受到了Google App Engine带来的好处：简单的部署，出色的可扩展性，以及利用Google成熟的基础设施的能力。
- 随后我们深入展示了如何把一个应用示例移植到App Engine上去。以及通过一些简单的步骤创建一个全新的应用，而无须关心之前的代码。这里我们把如何将logging应用重写为一个100%的App Engine应用留给你作为练习。
- 这一节附录里，我们向你介绍了Google App Engine及其能力，和另一种开发Django应用的环境。虽然为了一些新功能，Django不得不牺牲一些自身的特性，这给了Django程序员直接转向App Engine开发时增加了一些挑战。好在App Engine Helper的出现简化了这个任务。
- 最重要的是你作为一名Django程序员可以在继续编写Django的同时，享受到了Google App Engine带来的好处：简单的部署，出色的可扩展性，以及利用Google成熟的基础设施的能力。
- 下面是一些很有用的在线资源，完整的列表可以在本书的网站 [withdjangocom](http://withdjangocom) 上找到。
- Google App Engine <http://code.google.com/appengine/docs/gettingstarted/>
  - App Engine Tutorial <http://code.google.com/p/googleappengine/>
  - App Engine SDK Project <http://code.google.com/appengine/>
  - Google App Engine <http://code.google.com/p/appengine/>

## E.8 在线资源

更多信息请参考App Engine的文档。

- 用户API (User API)
  - URL获取API (URL Fetch API)
  - 内存缓存API Memcache API
  - 邮件API
  - 图形API
  - 数据存储API
  - Python运行时 (Runtime)
- 在构建应用的时候，记住不要编写“纯Django”的代码，比如要用App Engine Helper的模型（model）而不是Django的。不过有失也有得，你现在可以访问App Engine所有强大的API了。
6. 构建你的应用程序！
  5. 运行 manage.py startapp NEW\_APP\_NAME。
  4. 需要链接Google App Engine的代码。
  3. 用新的应用程序名修改app.yaml文件（就是向App Engine Admin Console注册的名字）。
  2. 把App Engine代码（app.yaml, main.py, appengine\_django）复制到项目目录里。

## VIDEOS

[http://code.google.com/appengine/articles/appengine\\_helper\\_for\\_django.html](http://code.google.com/appengine/articles/appengine_helper_for_django.html)

- Using the Google App Engine Helper for Django (Matt Brown, May 2008)

<http://code.google.com/p/google-app-engine-django/>

- Rapid Development with Python, Django, and Google App Engine (Guido van Rossum, May 2008)  
<http://sites.google.com/site/rapid-development-with-python-django-and-google-app-engine/introducing-gae-at-google-campfire> (various, Apr 2008, 7 videos)

<http://imovationstarups.wordpress.com/2008/04/10/google-app-engine-youtube/>

<http://djangopeople.net>注册，这是一个全球性的Django程序员聚集地（如果你有兴趣组织一支队伍加入邮件列表和IRC频道之外，另一个成为Django社区成员的方法就是在

风格到邮件列表礼仪等。

Django文档（在官网和[withdjango.com](http://withdjango.com)上都有链接）有关于这方面的细节，从编码

## 注意

区作出贡献了。如果你碰到了我们中的任何一个，别忘了打招呼哦！  
列出表或是在irc.freenode.net上的#django IRC频道里花点时间到处看看。很快你就会开始为社区贡献自己的力量！最好的方法就是加入django-users或是django-developers用户列表或是

• 在Django上创建一个高质量的应用并开源发布

• 找到一个被广泛要求但是还未实现的特性，然后实现它

• 为Django中某个还未实现的语言提供本地化支持（如果你能找到的话！）

最后是一些比较艰巨的任务，它们会对整个Django社区产生深远的影响：

• 参与Django上那个正在开发新特性的分支

• 定居在的问题

• 在不太流行的系统或是不常见（但是应该支持）的软硬件下运行Django测试来帮助外界  
• 为一个已知的bug提交patch

如果你不介意动手一下的话，还有一些不需要太大工作量的选择：

• 有效的报告

• 帮助整理[code.djangoproject.com](http://code.djangoproject.com)上的bug报告，关闭无效的报告，并且帮助确认（或推翻）  
在IRC频道和[djangoproject.com](http://djangoproject.com)用附件列表回答新手的问题

• 提交改正文档里的笔误

最简单的方法甚至连编辑都不需要：

• 贡献方面的方面。

Django是开源项目的典范（Python之父Guido van Rossum语），它提供了很多方式让感兴趣的  
人加入进来。和其他很多人一样，作为一个Web程序员你会发现在Django让你的工作变得更加  
简单而有趣，而这些随之而来的乐趣和快乐又能激发你回馈社区的灵感。下面是一些你可以贡  
献的人数过200名贡献者，此外还有很多为项目的运行作出或多或少贡献的人。  
个社区的程序员，测试，翻译，回答问题的人和来自全球的志愿者。Django的AUTHORS文件  
列出了超过200名贡献者，此外还有很多为项目的运行作出或多或少贡献的人。

## 附录F 参与Django项目

团队的话也可以通过这里来寻人）。另一个是Django程序员找工作的网站 <http://djangogigs.com>，它也有一个程序员的列表 (<http://djangogigs.com/developers>)。最后，你还可以从Django官方微博客，从那里可以获得更多关于Django开发的新闻。有关团圆的信息，以及其他社区工具的链接，都可以在以下网址找到：<http://djangoproject.com/community>。

在编写这本书的时候，我们需要把它当作软件开发那样来对待。我们需要使用高质量的开源工具来确保工作顺利完成，特别是那些适合团队合作的工具。虽然这里没有介绍所有的或最理想化的東西，不过在出版界的流程里居然还存在大量你熟悉某些独有的文字处理软件格式和email附件这种現象，我们在希望你能打破这种局面。以下是一份我们在准备、编写和编辑本书的过程中用到的一些重要的开源工具。

用于控制我们的手稿和项目文件版本的软件是Subversion（也有Git和Mercurial）。虽然很难完整地说明清楚到底版本控制对这类工作好在哪里，这里还是列出了一些：完整的项目的分支；允许并行工作而不会影响到他人，甚至是同一份文件的不同部分；任何时候都至少有四人完整的项目拷贝。

Trac是一个Python编写的轻型软件项目管理系統，它让我们可以方便地跟踪任何修改，另外还提供了基于wiki的共享笔记。Trac和Subversion（或是其他任何通过插件支持的版本控制

diff能让你清楚地看到那天完成的工作量，以及在最后一轮修改中到底改变了哪些地方。后台）非常强大，甚至对我们这样每一个（很大程度上）非编程的项目来说也不例外。彩色的文件组织。它们由Markdown文本标记系统写成，可以方便地生成HTML, PDF, 和其他输出格式。选择Markdown文本标记系统是因为它良好的可读性和极少量的标记，这两点对写作都非常重

我们用Mailman（Python编写的邮件列表管理器）设置的邮件列表来进行内部交流。

Mailman除了能提供基本的邮件服务外，还能保存每条消息，这样我们就可以随时回去查看而不用担心要把它们保存到本地的mailto客户端上。

在编写一个复杂的软件系统时，操作系统一定要有一个成熟的软件包管理系統。能够安装SQLite, Memcached, Makro, PostgreSQL及Apache等。从内建在Debian, Ubuntu和FreeBSD系统上出色的软件包管理系統，以及OS X上的MacPorts系统，我们获益良多。

Python语言在这里的重要性更是不言而喻。除了显而易见的优点，Python也支撑了我们在质量控制上作出的努力。我们用简单的Python脚本解析手稿文件，通过doctests模块自动测试代码示例。在手稿里的交互Python示例，并且更新我们从更大的正常工作的作用程序里抓出来的代码示例。我们还用Python代码来扫描所有的手稿文件并更新对应的目录文件，这样就能在任何时候看到我们的进度。Python还被用来执行Markdown-Python编译器从而将所有的文本文件转换成HTML。我们对能在这项工作中如此大量的使用Python感到非常高兴，希望你也一样。

我们制作的Maketile还会包含一些整理生成内容的指导，例如把所有的HTML合并为一个文件，压缩生成ZIP文件，为每个HTML文件打开新的浏览器窗口，以及生成PDF文件等（通过html2ps [不是PHP的那个，是另一个] 和Ghostscriptr的ps2pdf的帮助）。

最后，本书的网站也是用Django开发的哦！

通过html2ps这个伟大的神器，我们可以轻松地将复杂的Python代码、SQL语句，或是“MVC”是什么意思等。

荷兰的Web应用开发者和那些希望在网上发表作品的人联系起来，让他们不用去关心如何编写复杂的服务器代码，而是直接使用Django的模板语言。

荷兰也是Python之父Guido van Rossum的出生地。而Django，也希望能作为一座桥梁，把封面上的大师是位于荷兰鹿特丹的Erasmus大师。

软件	链接
Subversion	<a href="http://subversion.tigris.org">http://subversion.tigris.org</a>
Trac	<a href="http://trac.edgewall.org">http://trac.edgewall.org</a>
Mailman	<a href="http://www.gnu.org/software/mailman">http://www.gnu.org/software/mailman</a>
Maketile	<a href="http://daringfireball.net/projects/maketile">http://daringfireball.net/projects/maketile</a>
Tables	<a href="http://brrian-jarrell.livejournal.com/5978.html">http://brrian-jarrell.livejournal.com/5978.html</a>
TextMate	<a href="http://macromates.com">http://macromates.com</a>
Vim	<a href="http://www.vim.org">http://www.vim.org</a>
Ghostscript	<a href="http://pages.cs.wisc.edu/~ghost/">http://pages.cs.wisc.edu/~ghost/</a>
html2ps	<a href="http://userit.uu.se/~jan/html2ps.html">http://userit.uu.se/~jan/html2ps.html</a>
Firfox	<a href="http://mozilla.com/firfox">http://mozilla.com/firfox</a>
Ubuntu	<a href="http://ubuntu.com">http://ubuntu.com</a>
FreeBSD	<a href="http://freebsd.org">http://freebsd.org</a>
Macports	<a href="http://macports.org">http://macports.org</a>
Python	<a href="http://python.org">http://python.org</a>
Django	<a href="http://djangoproject.com">http://djangoproject.com</a>

我们的Maketile还会包含一些整理生成内容的指导，例如把所有的HTML合并为一个文件，压缩生成ZIP文件，为每个HTML文件打开新的浏览器窗口，以及生成PDF文件等（通过html2ps [不是PHP的那个，是另一个] 和Ghostscriptr的ps2pdf的帮助）。

最后，本书的网站也是用Django开发的哦！

通过html2ps这个伟大的神器，我们可以轻松地将复杂的Python代码、SQL语句，或是“MVC”是什么意思等。

荷兰的Web应用开发者和那些希望在网上发表作品的人联系起来，让他们不用去关心如何编写复杂的服务器代码，而是直接使用Django的模板语言。

荷兰也是Python之父Guido van Rossum的出生地。而Django，也希望能作为一座桥梁，把封面上的大师是位于荷兰鹿特丹的Erasmus大师。