

Learning Perl

第五版
涵盖 Perl 5.10

Perl 语言入门



O'REILLY®

东南大学出版社

Randal L. Schwartz,
Tom Phoenix & brian d foy 著

盛春 蒋永清 王晖 译

Perl 语言入门

第五版

*Randal L. Schwartz,
Tom Phoenix & brian d foy* 著
盛春 蒋永清 王晖译

O'REILLY®

Beijing • Cambridge • Farnham • Köln • Sebastopol • Taipei • Tokyo

O'Reilly Media, Inc. 授权东南大学出版社出版

东南大学出版社

图书在版编目 (CIP) 数据

Perl 语言入门:第 5 版 / (美) 施瓦茨 (Schwartz, R. L.),
(美) 菲尼克斯 (Phoenix, T.), (美) 福瓦 (Foy, B. D.) 著;
盛春, 蒋永清, 王晖译. —南京: 东南大学出版社, 2009.8

书名原文: Learning Perl, Fifth Edition

ISBN 978-7-5641-1763-4

I. P… II. ①施… ②菲… ③福… ④盛… ⑤蒋… ⑥王…

III. PERL 语言—程序设计 IV. TP312

中国版本图书馆 CIP 数据核字 (2009) 第 125922 号

江苏省版权局著作权合同登记

图字: 10-2008-353 号

©2009 by O'Reilly Media, Inc.

Simplified Chinese Edition, jointly published by O'Reilly Media, Inc. and Southeast University Press, 2009. Authorized translation of the English edition, 2008 O'Reilly Media, Inc., the owner of all rights to publish and sell the same.

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

英文原版由 O'Reilly Media, Inc. 出版 2008。

简体中文版由东南大学出版社出版 2009。英文原版的翻译得到 O'Reilly Media, Inc. 的授权。此简体中文版的出版和销售得到出版权和销售权的所有者——O'Reilly Media, Inc. 的许可。

版权所有, 未得书面许可, 本书的任何部分和全部不得以任何形式重制。

Perl 语言入门 (第五版)

出版发行: 东南大学出版社

地 址: 南京四牌楼 2 号 邮编: 210096

出版人: 江 汉

网 址: <http://press.seu.edu.cn>

电子邮件: press@seu.edu.cn

印 刷: 扬中市印刷有限公司

开 本: 787 毫米 × 980 毫米 16 开本

印 张: 21.5 印张

字 数: 361 千字

版 次: 2009 年 8 月第 1 版

印 次: 2009 年 8 月第 1 次印刷

书 号: ISBN 978-7-5641-1763-4

印 数: 1~4000 册

定 价: 48.00 元 (册)

本社图书若有印装质量问题, 请直接与读者服务部联系。电话 (传真): 025-83792328

O'Reilly Media, Inc.介绍

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

从最畅销的《The Whole Internet User's Guide & Catalog》（被纽约公共图书馆评为二十世纪最重要的 50 本书之一）到 GNN（最早的 Internet 门户和商业网站），再到 WebSite（第一个桌面 PC 的 Web 服务器软件），O'Reilly Media, Inc.一直处于 Internet 发展的最前沿。

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

目录

前言	1
第一章 简介	9
问题与答案	9
“Perl” 这个词是什么意思?	12
如何取得 Perl?	17
我该怎么编写 Perl 程序?	20
走马观花	25
习题	26
第二章 标量数据	28
数字	28
字符串	31
Perl 内建警告信息	35
标量变量	37
用 print 输出结果	39
if 控制结构	43

获取用户输入	44
chomp 操作符	45
while 控制结构	46
习题	48
第三章 列表与数组	49
访问数组中的元素	50
特殊的数组索引值	51
列表直接量	52
列表的赋值	54
字符串中的数组内插	57
foreach 控制结构	58
标量上下文与列表上下文	60
列表上下文中的 <STDIN>	63
习题	65
第四章 子程序	66
定义子程序	66
调用子程序	67
返回值	68
参数	69
子程序中的私有变量	71
长度可变的参数列表	72
关于词法 (my) 变量	74
use strict 编译命令	75
return 操作符	77
非标量返回值	79
持久性私有变量	80
习题	81

第五章 输入与输出	83
读取标准输入	83
钻石操作符输入	85
调用参数	87
输出到标准输出	88
使用 printf 格式化输出	91
文件句柄	93
打开文件句柄	95
用 die 处理严重错误	98
使用文件句柄	101
复用标准文件句柄	102
使用 say 来输出	103
习题	104
第六章 哈希	105
什么是哈希?	105
访问哈希元素	108
哈希函数	113
哈希的典型应用	115
%ENV 哈希	117
习题	118
第七章 漫游正则表达式王国	119
什么是正则表达式?	119
使用简易模式	120
字符集	126
习题	128

第八章 以正则表达式进行匹配	129
以 m// 进行匹配	129
可选修饰符	130
锚位	132
绑定操作符 =~	133
模式串中的内插	134
捕获变量	135
通用量词	141
优先级	142
模式测试程序	144
习题	145
 第九章 用正则表达式处理文本	 146
用 s/// 替换	146
可选修饰符	148
split 操作符	149
join 函数	150
列表上下文中的 m//	151
更强大的正则表达式	152
习题	159
 第十章 其他控制结构	 160
unless 控制结构	160
Until 控制结构	161
条件修饰词	162
裸块控制结构	163
elsif 子句	164
自增和自减	165
for 控制结构	167
循环控制	170

三目操作符 ?:	174
逻辑操作符	175
习题	180
第十一章 Perl 模块	181
寻找模块	181
安装模块	182
使用简单模块	183
习题	190
第十二章 文件测试	191
文件测试操作符	191
localtime 函数	200
按位运算操作符	201
习题	202
第十三章 目标操作	203
在目录树中移动	203
文件名通配	204
文件名通配的另一种语法	205
目录句柄	206
递归的目录列表	208
操作文件与目录	208
删除文件	208
重命名文件	210
链接与文件	211
建立及移除目录	216
修改权限	217
更改隶属关系	218

修改时间戳	218
习题	219
第十四章 字符串与排序	221
在字符串内用 index 搜索	221
用 substr 处理子串	222
高级排序	226
习题	232
第十五章 智能匹配与 given-when 结构	233
智能匹配操作符	233
智能匹配操作的优先级	236
given 语句	237
多个项目的 when 匹配	242
习题	243
第十六章 进程管理	245
system 函数	245
exec 函数	249
环境变量	250
用反引号捕获输出结果	250
将进程视为文件句柄	254
用 fork 开展地下工作	256
发送及接收信号	256
习题	259
第十七章 高级 Perl 技巧	260
用 eval 捕获错误	260
用 grep 来筛选列表	263

用 map 对列表进行转换	264
不带引号的哈希键	265
切片	265
习题	270
 附录 A 习题解答	 271
 附录 B 超越小骆驼	 305

前言

欢迎阅读《Perl 语言入门》第五版，此版本顺应 Perl 5.10 及其后续版本的新特性而更新。当然，如果你还在用 Perl 5.6（这个版本已经发布很久了，你还没想过升级？），这本书同样适用于你。

假如你正在寻找用 30 到 45 小时就能掌握 Perl 语言编程的最佳方式，那么你已经找到了！在后面的 300 多页里，我们会提供精心安排的入门指引，介绍这个在互联网中担负重任的程序语言。它也是最受全世界系统管理员、网络黑客（web hacker）及业余程序员青睐的程序语言。

我们不可能只花几小时就把 Perl 的全部知识传授给你，会这么保证的书大概都撒了一点谎。相对地，我们慎选了 Perl 中完整又实用的部分供你学习。这些材料足以编写 128 行以内的小程序，大约 90% 的 Perl 程序都不需要很多篇幅。当你准备继续深入时，建议您阅读《Intermediate Perl》这本书，该书涵盖了许多本书舍去不讲的部分。此外，我们还纳入了许多知识点以便后续的扩展和研习。

每章的内容并不多，可以在一两个小时之内读完。各章后面都有一系列的习题，帮助你巩固刚学到的知识，在附录 A 中还附有习题解答，供你比对思考。因此，本书可以说是相当适合作为“Perl 入门”的课堂教材。我们对此有第一手的经验，因为本书的内容几乎是逐字逐句从“Learning Perl”教学中萃取出来的，而这正是我们教过上千名学生的招牌课程。话虽如此，我们也将本书设计成适合自学的形式。

虽然 Perl 是活生生的“Unix 工具箱”，但你并不需要是 Unix 大师，甚至也不必懂 Unix 就可以使用本书。除非特别注明，否则我们所提到的一切都可以同样应用到 Windows 版本的 ActivePerl（ActiveState 出品），以及许许多多其他新潮的 Perl 版本。

阅读本书之前，虽然不需事先具备任何 Perl 的基础，但我们还是衷心希望你先熟悉一

些写程序的概念，像变量 (variable)、循环 (loop)、子程序 (subroutine) 和数组 (array) 以及最重要的“用你最熟悉的文本编辑器来编辑源代码”。我们不会花时间来尝试说明这些概念。有些人平生所学的第一个程序语言就是 Perl，并因学习本书而获得成功，我们相当高兴有这样的例子，但是我们并不敢保证每个人都能取得一样的成果。

排版约定

本书使用以下的字体惯例：

等宽字 (*Constant width*)

用于方法名称 (method name)、函数名称 (function name)、变量 (variable)、属性 (attribute) 以及程序代码范例。

等宽黑体字 (**Constant width bold**)

用于用户所输入的内容。

等宽斜体字 (*Constant width italic*)

用于程序代码中可被替换的项目 (例如：filename，你可以将它代换成实际的文件名)。

斜体字 (*Italic*)

用于正文所提到的文件名称、URL、主机名称、第一次提及的重要词汇以及命令。

脚注

一般附加在括号之内，初次 (也许是第二次、第三次) 阅读本书时应该略过。有一些不完全正确的用语是为了简化说明，而脚注会说明事实。通常脚注中的资料是高级主题，不会在本书其他部分讨论到。

如何与我们联系

本书的内容都经过测试，尽管我们做了最大的努力，但错误和疏忽仍然在所难免。如果你发现有什么错误或者是对将来的版本有什么建议，请通过下面的地址告诉我们：

美国：

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

中国：

100035 北京市西城区西直门南大街 2 号成铭大厦 C 座 807 室
奥莱利技术咨询（北京）有限公司

你也可以发送电子邮件给我们。若是想要加入我们的邮件列表，或是索取书籍类目手册，可以发邮件到：

info@oreilly.com

询问技术问题或对本书进行评论，请发送邮件到：

bookquestions@oreilly.com

info@mail.oreilly.com.cn

我们还为本书建立了一个专属网页，你可以在此找到关于本书的相关信息，包括范例程序、勘误表，以及将来再版的计划等。该网页还提供一系列可下载的文本文件（以及一些 Perl 程序），在做本书练习的时候可以直接使用。此页面地址为：

<http://www.oreilly.com/catalog/9780596520106>（原版书）

<http://www.oreilly.com.cn/book.php?bn=978-7-5641-1763-4>（中文版）

其他关于本书或其他书籍的信息，请访问 O'Reilly 的网站：

<http://www.oreilly.com>

<http://www.oreilly.com.cn>

使用程序代码范例

本书的目的是让你能够在实践中解决问题。基本上，你不用事先联络我们就可以使用本书所提供的程序代码及文件，除非是大量复制。举例来说，在你的程序中，若用到几段本书中的程序代码，不需要经过我们的同意；但是做成光盘发布、销售 O'Reilly 书中的例子，则必须经过授权。回答别人的问题时，引用本书的文字和程序代码，也不需要经过我们的同意；但在你的产品文件中，若大量加入本书的文字与程序代码，则必须经过授权。虽非必要，但我们会十分感谢你在引用本书的内容和范例时提到出处。完整的信息通常包括书名、作者、出版商及 ISBN 编号。例如：“Learning Perl, Fifth Edition, by Randal L. Schwartz, Tom Phoenix and brian d foy. copyright 2008 O'Reilly Media, Inc., 978-0-596-52010-6。”。如果你的情况有别于上述情形，并心存疑问的话，请给我们来信 permissions@oreilly.com。

本书的历史

为了满足读者的好奇心，Randal 在这里告诉你关于这本书的来历：

1991 年我跟 Larry Wall 写完第一本《Perl 语言编程》之后，硅谷的 Taos Mountain Software 公司跟我联络，要我准备一些培训课程，内容包含 12 次左右的课程，并训练他们的教师继续开课。我就按约写了这个课程给他们【注 1】。

在课程进行了三四次之后（1991 年底），有个人走到面前跟我说：“不瞒你说，我真地很喜欢《Perl 语言编程》这本书，但是这堂课的教材更容易吸收，你真地应该写一本像这个课程的书。”这听起来像是个好机会，所以我开始认真地考虑这个点子。

我写信给 Tim O'Reilly，附上了一份企划书。这是以 Taos Mountain 课程纲要为基础，再根据课堂上的观察调整并修改了一些章节。这可能是有史以来我的企划书最快被接受的记录——我在 15 分钟后收到了 Tim 的回信：“我们一直在等待你的第二本书。《Perl 语言编程》太热销了。”接下来的一年半时间里，我就努力完成了第一版的《Perl 语言入门》。

在那段时间里，我找到硅谷以外教授 Perl 的机会【注 2】，所以我就以正在编写阶段的《Perl 语言入门》为蓝本制作了一套课程。我为许多不同的客户教课（包括我的主要签约人 Intel Oregon），并利用上课所得到的响应进一步微调本书的草稿。

第一版在 1993 年的 11 月 1 日【注 3】问世，销售空前成功，甚至很快就追上了《Perl 语言编程》的销量。

在第一版的封底上这么写着：“由卓越的 Perl 讲师所著”，事后证明这是正确的预言。随后的几个月里，我收到来自美国各地的电子邮件，邀请我到他们那里教 Perl。接下来的 7 年中，我的公司成了全球领先的 Perl 现场培训公司，我个人的飞行里程数也飙升到了百万英里。之后互联网的兴起更是锦上添花，许多站长都采用 Perl 作为内容管理、交互式 CGI 及网站维护的语言。

注 1： 在合约中，我保留了习题的所有权，我希望有一天能以不同的方式来使用它们，比如说我以前曾经写过的杂志专栏。习题是 Taos 公司的课程里唯一还能在本书中出现的東西。

注 2： 我与 Taos 公司的合约有条独特的条款，因此不能在硅谷教授类似的课程，我也遵守了此条款很多年。

注 3： 这个日期我记得很清楚，因为那也是我由于一些跟计算机有关的行为在家被逮捕的日子。关于我与 Intel 公司的合约，我被判有罪。

我跟 Stonehenge 的首席培训师兼内容经理 Tom Phoenix 密切合作了两年。我请他对 Llama 课程做实验，把某些东西移来移去，再打散一些内容。当他带着我们认为最好的修订本出现时，我就联络 O'Reilly，说：“是该有本新书的时候了！”于是第三版就这么诞生了。

在小骆驼书第三版问世的两年后，我和 Tom 决定把一些“高级”的课程移出来成为一本独立的、专门给需要写“100 到 10000 行代码”的人看的书，那就是在 2003 年完成的羊驼书。

不过，在我的同事 brian d foy 从海湾战争回来之后，同样是讲师的他注意到教材必须进一步适应普通学生的需求，因此这两本书都应该适当地改写。于是他对 O'Reilly 推销这个想法，希望在 Perl 6 完成之前进行小骆驼书与羊驼书的最后一次改版（但愿如此）。而此版本的确反映了那些变动的需求。我很少需要给 brian 什么建议，他一向都是顶尖的作者，在写作团队里面他给人的感觉就像是尽责的英国管家。

2007 年 12 月 18 日，“Perl 5 掌门人”（perl 5 porters）发布了 Perl 5.10，一个标志性的版本，融入了众多新特性。之前的 5.8 版本专注于 Perl 的基础架构改良和 Unicode 支持。而最新的版本，以稳固的 5.8 为基础，增加了一系列崭新的特性，特别是那些取自正在开发中（尚未发布）的 Perl 6 的一些理念。其中某些特性，诸如正则表达式里的命名捕捉，比起传统做法来要好很多，对 Perl 初学者来说也更容易掌握。我们未曾想过本书会有第五版，但 Perl 5.10 实在是太有趣了，我们无法故步不前。

读者可能注意到了这一版跟前一版本的某些差异：

- 某些内容为 Perl 5.10 而更新，某些代码仅在此新版本中可用。在讨论 Perl 5.10 的特性时，我们会在行文中加以提示说明。对于那些代码片段，我们也一律使用特殊的 `use` 语句加以区别，提示你使用正确的版本：

```
use 5.010; # 当前脚本需要 Perl 5.10 或更高版本
```

如果在代码示例中没有看到 `use 5.010` 这句话，就说明它也可以在 Perl 5.6 以上版本中工作。要看当前使用何种版本，可在命令行使用 `-v` 参数查看：

```
prompt% perl -v
```

这里列举一些我们将要讨论到的 Perl 5.10 新特性，及其相应章节。介绍这些新特性的同时，我们还会展示如何用老的方法实现相同的功能：

- 有很多 Perl 5.10 的新特性是有关于正则表达式的，包括相对反向引用（第七章），新的字符集简写（第七章）和命名捕捉（第八章）。

- Perl 5.10 包含一个新的条件切换语句，称作 `given-when`。我们会在第十五章中同智能匹配操作符一同讲解。
- 子程序现在可以像 C 一样拥有静态变量，不过在 Perl 里称为 `state` 变量。这类变量能在子程序的多次调用之间保留其中的值，且作用域限于子程序内。我们会在第四章中阐述此特性。

致谢

来自 Randal

我想要感谢 Stonehenge 过去与现在的讲师们（Joseph Hall、Tom Phoenix、Chip Salzenberg、brian d foy 与 Tad McClellan），谢谢他们愿意每周到教室授课并且带回自己的笔记，注明哪部分有用（及哪部分没用），如此我们才能调整这本书的内容。我要特别点名 Tom Phoenix，我的共同作者与事业伙伴，他花了大量时间改进 Stonehenge 的 Llama 课程，也为本书注入了主要的原始内容。还有 brain d foy，在第四版中担任了主要的写作任务，从而完成了我从收件箱转发的无数待办事项。

此外，我还要感谢 O'Reilly 的每一位人员，尤其是我们非常有耐心与眼光的编辑 Allison Randal（不是我的亲戚，但她的姓氏拼法真好）。还有 Tim O'Reilly 本人，是他让我在一开始就有写作大骆驼与小骆驼这两本书的机会。

我由衷感谢过去购买本书的读者，这些钱让我免于流浪街头与夜宿囚牢；感谢我班上的学生，他们把我训练成为一名更好的讲师；还有“财富一千”上大排长龙、在过去选购我们的课程、未来也会继续捧场的客户们。

和以前一样，我得特别感谢 Lyle 与 Jack，他们教会了我几乎所有关于写作的知识，我永远不会忘记你们。

来自 Tom

我必须附和 Randal 对 O'Reilly 的每个人致上谢意。在第三版的时候，我们的编辑是 Linda Mui，她细心地指出书中太过火的玩笑与脚注，当然留下来的那些也不是她的错。她与 Randal 在整个写作过程中不断指导我，我非常感激。在第四版中，Allison Randal 成为了我们的新编辑，我也由衷感谢她。