

Learning Perl

第三版

Perl 语言入门



O'REILLY®

中国电力出版社

Randal L. Schwartz & Tom Phoenix 著

李晓峰 译

Perl 语言入门

第三版

Randal L. Schwartz & Tom Phoenix 著

李晓峰 译

O'REILLY®

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

O'Reilly & Associates, Inc. 授权中国电力出版社出版

中国电力出版社

图书在版编目 (CIP) 数据

Perl 语言入门 (第三版) / (美) 斯库尔茨 (Schwartz, R.) 等著; 李晓峰译.
- 北京: 中国电力出版社, 2002.9

书名原文: Learning Perl, Third Edition

ISBN 7-5083-1113-2

I. P... II. ①斯... ②李... III. Perl - 程序设计 IV. TP312

中国版本图书馆 CIP 数据核字 (2002) 第 053057 号

北京市版权局著作权合同登记

图字: 01-2001-4018 号

©2001 by O'Reilly & Associates, Inc.

Simplified Chinese Edition, jointly published by O'Reilly & Associates, Inc. and China Electric Power Press, 2002. Authorized translation of the English edition, 2001 O'Reilly & Associates, 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 & Associates, Inc. 出版 2001。

简体中文版由中国电力出版社出版 2002。英文原版的翻译得到 O'Reilly & Associates, 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 & Associates, 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。

书 名 / Perl 语言入门 (第三版)

书 号 / ISBN 7-5083-1113-2

责任编辑 / 夏平

封面设计 / Edie Freedman, 张健

出版发行 / 中国电力出版社 (www.infopower.com.cn)

地 址 / 北京三里河路 6 号 (邮政编码 100044)

经 销 / 全国新华书店

印 刷 / 北京市地矿印刷厂

开 本 / 787 毫米 × 1092 毫米 16 开本 24 印张 344 千字

版 次 / 2002 年 8 月第一版 2002 年 8 月第一次印刷

印 数 / 0001-5000 册

定 价 / 48.00 元 (册)

目录

前言	1
第一章 简介	9
问题和答案	9
“Perl” 代表什么意思?	12
如何得到 Perl?	17
如何写一个 Perl 程序?	22
Perl 的旋风式旅行	27
习题	29
第二章 标量数据	30
什么是标量数据?	30
数字	31
字符串	34
Perl 的内置警告	38
标量变量	39
用 print 输出	42

if 控制结构	47
取得用户输入	49
chomp 操作符	50
while 控制结构	51
undef 值	51
defined 函数	53
习题	53
 第三章 列表和数组	55
访问数组的元素	56
特殊的数组索引	57
列表直接量	58
列表赋值	61
在字符串中替换数组	64
foreach 控制结构	65
Perl 最喜欢的缺省变量: \$_	66
标量和列表上下文	68
列表上下文中的 <STDIN>	72
习题	74
 第四章 子例程	75
系统与用户函数	75
定义一个子例程	76
调用子例程	77
返回值	77
参数	80
子例程中的私有变量	81
local 操作符	82
可变长参数列表	84
词法 (my) 变量的说明	87

use strict pragma	88
return 操作符	90
习题	93
第五章 散列	94
什么是散列?	94
散列元素访问	98
散列函数	103
散列的典型使用	106
习题	108
第六章 I/O 基础	109
从标准输入进行输入	109
从钻石操作符进行输入	112
调用参数	114
向标准输出进行输出	115
用 printf 进行格式化的输出	119
习题	122
第七章 正则表达式的概念	124
什么是正则表达式?	124
使用简单的模式	126
一个模式测试程序	129
习题	130
第八章 正则表达式提高	132
字符类	132
通用数量符	135
锚位符	135
记忆的小括号	137

优先级	140
习题	142
第九章 使用正则表达式	143
使用 m// 进行匹配	143
选项修饰符	144
绑定操作符 =~	145
在模式中替换	146
匹配变量	148
用 s/// 进行替换	152
split 操作符	155
join 函数	157
习题	158
第十章 更多的控制结构	160
unless 控制结构	160
until 控制结构	162
表达式修饰符	162
裸块控制结构	164
elsif 子句	165
自递增与自递减	166
for 控制结构	168
循环控制	172
逻辑操作符	177
习题	182
第十一章 文件句柄和文件测试	183
什么是文件句柄?	183
打开一个文件句柄	185

用 die 表明致命错误	188
使用文件句柄	191
重新打开一个标准文件句柄	193
文件测试	194
习题	205
第十二章 目录操作	206
改变目录树	206
glob	207
glob 的另一种语法	209
目录句柄	210
递归的目录列表	212
习题	212
第十三章 处理文件和目录	213
删除文件	213
重命名文件	215
链接和文件	217
建立和删除目录	222
修改权限	225
改变所有者	225
改变时间戳	226
使用简单的模块	227
习题	232
第十四章 进程管理	234
system 函数	234
exec 函数	238
环境变量	240

使用反引号捕获输出	241
文件句柄形式的进程	245
用 fork 进行深入和复杂的工作	247
发送和接收信号	248
习题	251
第十五章 字符串与排序	252
用 index 寻找子字符串	252
用 substr 处理一个子字符串	254
用 sprintf 格式化数据	256
高级排序	258
习题	265
第十六章 简单数据库	267
DBM 文件和 DBM 散列	267
用 pack 和 unpack 处理数据	270
固定长度的随机访问数据库	271
可变长 (文本) 数据库	275
习题	280
第十七章 一些高级 Perl 技术	281
用 eval 捕获错误	281
用 grep 从一个列表中选择条目	284
用 map 转换一个列表的条目	286
不加引号的散列键	287
更强大的正则表达式	288
切片	291
习题	297

附录一 习题答案	299
附录二 尚未谈到的主题	334
词汇表	359

前言

欢迎阅读本书第三版。

如果你刚开始学习 Perl 程序设计语言，希望找到一条途径以最高效地使用最初的 30 到 45 个小时，那你算找对地方了。在后面的章节中，你会发现关于这种 Internet 主力语言的按部就班的介绍。这种语言还是系统管理员、Web 高手以及临时程序员的选择。

我们不可能在几个小时中把 Perl 的全部内容都介绍给你。那些自称如此的书一般都有吹嘘之嫌。我们的做法是为你精心选择了 Perl 的一个完整有用的子集，足够使你能编写出一些像样的程序，这类程序占了常用程序的 90%。如果你还想进一步学习，我们也为你指明继续学习的一些参考资料。

本书的每一章都可以在 1~2 小时内读完。每章的后面是一组习题，供你复习刚刚学到的内容，附录一提供了答案。因此本书非常适合做“Perl 入门”一类课程的教材。我们对此非常清楚，因为本书的内容基本上就是逐字地从我们的旗舰（flagship）课程“Learning Perl”搬过来的，全球有几千人都修过该课。不过本书也是专为自学者而写的。

尽管 Perl 一直被认为是“Unix 的工具箱”，但你不必是个 Unix 高手才能阅读本书，甚至不必是 Unix 用户。除非专门指出，书中所有的内容都可以很好地应用于

ActiveState 的 Windows ActivePerl (译注 1)，还有后来为 Macintosh 推出的 MacPerl，以及几乎所有新近实现的 Perl 版本。

尽管你不必需要任何 Perl 的知识来阅读本书，不过我们仍希望你熟悉基本的程序设计概念，如变量、循环、子例程及数组，还有最重要的“用你最喜欢的编辑器编辑源文件”。我们打算花时间来解释这些概念。我们很高兴地获知一些读者已读完本书，并熟练掌握了作为他们第一门程序设计语言的 Perl。不过我们并不能保证其他人也能如此。

排版约定

本书采用如下的英文字体约定：

等宽字体 (*Constant width*)

用来表示方法名、函数名、变量以及属性。也用来表示代码示例。

等宽黑体 (**Constant width bold**)

用来表示需要用户输入的内容。

等宽斜体 (*Constant width italic*)

用来表示代码中的可替换项（例如，*filename* 将被替换为一个实际的文件名）。

斜体 (*Italic*)

用来表示文件名、URL 链接、主机名、文本中的命令、第一次提到的重要术语以及强调。

脚注

脚注用来附加一些额外的说明，在你第一遍（甚至第二遍、第三遍）阅读本书时可以略去它们。有时为了简化说明，我们可能会使用一些不准确的说

译注 1：MS Windows 下有三个常用的 Perl 版本：ActivePerl、Cygwin Perl 和 MKS Perl，其中 MKS Perl 最稳定。

法；而在脚注中我们会纠正这些说法。通常脚注中的内容较为高级，甚至在本书的其他地方也不会讨论它们。

建议与评论

本书的内容都经过测试，尽管我们做了最大的努力，但错误和疏忽仍然是在所难免的。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 & Associates, Inc.
101 Morris Street
Sebastopol, CA 95472

中国：

100080 北京市海淀区知春路 49 号希格玛公寓 B 座 809 室
奥莱理软件（北京）有限公司

你也可以发送电子邮件。若希望加入我们的邮件列表或得到书目，请发 email 到：

info@oreilly.com

若询问有关技术问题，或对本书有任何评论，请发 email 到：

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

本书有一个网页，在那里我们列出了例子、勘误以及未来版本的计划。那里还有一组可下载的文本文件（和一组 Perl 程序），这些内容对你做练习很有帮助，但并不是必需的。该网址是：

<http://www.oreilly.com/catalog/lperl3/>

欲获得本书及其他书籍的更多信息, 请访问 O'Reilly 的网站:

<http://www.oreilly.com>

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

代码示例

我们建议你拷贝本书中的代码示例, 并按你的实际需要适当调整。你不必手工录入, 最好从 <http://www.oreilly.com/catalog/lperl/> 直接下载这些代码。

本书的故事

为了满足一些读者的好奇心, 下面是 Randal 讲述的关于本书如何产生的故事:

在我和 Larry Wall 完成了第一本《Programming Perl》(译注 2) 之后 (在 1991 年), 硅谷的 Taos Mountain Software 公司找到我, 希望我能给他们提供培训。这包括由我讲授开始的十几次课, 并培训他们的员工以使他们能继续讲授课程。我为他们编写了教材 (注 1) 并按承诺完成了该课程。

在第三次或第四次讲授此课时 (在 1991 年底), 有人对我说: “你知道, 我非常喜欢《Programming Perl》, 不过在此课程中讲述的内容学起来要比这本书容易得多——你应该写一本类似于本课的书。”这给了我一个启发, 因此我开始考虑这个问题。

我给 Tim O'Reilly 写了一封信并附上写一本新书的提议, 提纲类似于我为 Taos 公司讲授的内容——不过我根据在课堂上的观察对一些章节做了调整和修改。我

译注 2: 本书最新版的中文版《Perl 语言编程 (第三版)》已由中国电力出版社引进出版, 详情请访问 www.infopower.com.cn。

注 1: 根据合同, 我保留了使用练习题的权力, 希望某一天能以其他方式重新利用它们, 比如, 用在目前我为某杂志专栏撰写的文章中。练习题是书中唯一取自 Taos 教材的内容。

想那是有生以来我的提议被最快接受的一次——不到15分钟我就收到了Tim的回信，说“我们一直在等你抛出第二本书——《Programming Perl》卖得太火了。”由此我们开始了18个月的努力，从而完成了本书的第一版。

在此期间，我也在寻找机会到硅谷以外的地方讲授Perl课程（注2），因此我以本书的内容为基础设计了一门课，并为许多客户讲授了该课（其中包括我的主要订约人Intel Oregon公司），我根据反馈意见进一步改进了该书的内容。

第一版在1993年11月的第一天（注3）问世了，并立即取得了巨大的成功，经常卖得比《Programming Perl》还要快。

第一版书的封底纸套上写道“由首席Perl培训师撰写”，后来它的确成了自我实现的一个预言。几个月后，我开始不断收到从美国各地发来的电子邮件，希望我能够到他们那里讲课。在随后的七年里，我的公司成了全世界最主要的Perl现场培训公司，我个人积累了百万英里的飞行里程。Web应用也在那个期间逐渐发展起来，而Web站点的管理员们选择了Perl语言来进行内容管理、CGI交互以及日常维护。

在过去的两年里，我一直在和Tom Phoenix合作，他作为Stonehenge公司的培训师负责人及内容管理员，我给他全权去随意试验“小骆驼(Llama)”课程。当我们认为该课已经达到最佳效果时，我便联系O'Reilly，说“是出一本新书的时候了！”而现在你正在阅读这本书。

你可能会发现本书和以前版本的一些区别：

- 内容是全新的，而不是以前版本的简单拷贝和粘贴。本书的内容是从我们为Stonehenge公司编写的“Learning Perl”课件及由我们创作并实地测验的教

注2：我和Taos的合同中有不竞争条款，因此我必须在硅谷以外讲授类似的课程。这个约定我已经遵守了很多年。

注3：关于那一天我记得很清楚，因为就在同一天我由于与Intel的合同有关的一些计算机相关活动而在家里被捕，我受到一系列的重罪指控并在后来被认定有罪。上诉的斗争仍在继续——细节请参考<http://www.lightlink.com/fors/>。

师讲义中抽取出来的。(有些习题与以前的版本类似是因为我们直到最近一直用它们做我们的教材,即使如此我们在重写时也做了修改。)

- 我们把“难以囫圇吞枣”的正则表达式那一节分成了易于理解的三节。
- 我们在编写习题时同时考虑了 Unix 和 Windows 平台。
- 我们取消了别扭的“控制结构”一章,把 while 和 if 语句的内容前移,而把有关 foreach 和 for 循环的介绍后移。这样我们在其他章节,如标量一章,就可以使用更多有用的例子和习题了。
- 我们把子例程的内容前移很多,这样之后的习题就可以在问题和答案中使用子例程。
- 我们现在讲授数组和散列表时先讲元素语法后讲集合语法。这样做使课堂效果出奇的好,因为几乎所有的初学者都不会再犯一个常见的错误,即把切片(slice)与元素混淆。冒天下之大不韪,我们承认希望其他的 Perl 讲师或书籍也能按我们的方式组织。
- 习题更加接近实际并且更加循序渐进。
- 我们包含了关于 use strict、警告和模块的资料,尽管主要是为了给进一步学习的一个引子。
- 我们尽量将内容组织得不针对 Unix 系统管理员,而主要针对一般的程序员。诸如“类似于 C”的句子基本上被完全去掉了。
- 小幽默更加精彩。(我们在课堂上总是即兴讲些小笑话,其中的一些最终成为我们标准 Stonehenge 讲义的内容,而最好的一些则出现在这里。你真应该也看看那些没有包含在这里的笑话!)
- 我们非常遗憾的是这一版没有 Larry Wall 写的精彩诙谐的前言,因为本书付印时他正在忙着定义 Perl 6。Larry 对我们的工作一直大力支持,我们知道在精神上(如果不是在文字上)他仍然是本书的一部分。祝愿你在富饶的 Perl 领地度假时一切顺利。

致谢

Randal的话：我要感谢Stonehenge过去的和现在的培训师们（Joseph Hall, Tom Phoenix, Chip Salzenberg, brian d foy, 还有Tad McClellan）。他们周复一周地站在教室前讲课，然后又带回他们的讲义，告诉我们哪些部分很成功（或不成功），这样我们可以不断地改进本书的内容。我特别要感谢本书的作者之一及我的业务伙伴Tom Phoenix，他花了很多很多的时间来改善Stonehenge的小骆驼课程，并为本书的主要部分提供了精彩的核心部分。

我还要感谢O'Reilly的每个人，特别是给我们鼓励并很有耐心的编辑Linda Mui，还有Tim O'Reilly本人，是他首先给了我写作大骆驼书和小骆驼书的机会。

我当然还要感谢成千上万购买小骆驼书以前版本的人们，是他们使我能够有钱“取保候审”；我还要感谢我的学生们，是他们培训了我，使我成为更好的培训师；更要感谢财富1000（Fortune 1000）中的广大客户，他们在过去购买过我们的课程，并打算在将来继续购买。

像以往一样，我还要特别感谢Lyle和Jack，他们教给我关于写作的所有知识。我永远不会忘记你们。

Tom的话：我也要响应Randal对编辑Linda Mui的致谢，她总是耐心地指出哪个笑话和脚注是多余的，并指出她也不能为剩下的那些承担责任。她和Randal在我写作的全过程中给予了指导，我非常感激。

我还要对Randal及其他Stonehenge的培训师表示感谢。每当我为试验一种新的讲授技巧而突然更改课程内容时，你们从未抱怨过。你们就教学方法提出了很多我从未听说过的不同观点。

我曾在俄勒冈科学工业博物馆（Oregon Museum of Science and Industry, OMSI）工作过多年，我想要感谢那儿的同事们，在我学着在每次活动、展览或学术演讲中引入一两个小笑话时，他们磨练了我的讲授技巧。