

Perl

参考大全

Martin C. Brown 著
顾凯等译

(第二版)

涵盖Perl结构、代码块、语句以及更丰富的构建模块。

包含数据处理、Web和网络编程的实际解决方案。

含有Perl内置函数和Perl标准库的全部细节。

Perl 参考大全

(第二版)

Martin C. Brown 著

顾凯 等 译

人民邮电出版社

图书在版编目 (CIP) 数据

Perl 参考大全 (第二版) / (美) 布鲁因 (Brewin, C. M.) 著; 顾凯译.
—北京: 人民邮电出版社, 2002.2

ISBN 7-115-09943-X

I. P... II. ①布... ②顾... III. PERL 语言—程序设计 IV. TP312

中国版本图书馆 CIP 数据核字 (2001) 第 088055 号

版 权 声 明

Martin C. Brown: The Complete Refrence Perl (Second Editon)

ISBN 0-07-212950-6

Copyright © 2001 by the McGraw-Hill Companies, Inc.

Authorized translation from the English language edition published by McGraw-Hill, Inc.

All rights reserved.

本书中文简体字版由人民邮电出版社和美国麦格劳-希尔国际公司合作出版。未经出版者书面许可, 不得以任何方式复制或抄袭本书的任何部分。

版权所有, 翻印必究。

Perl 参考大全 (第二版)

- ◆ 著 Martin C. Brown
译 顾 凯 等
责任编辑 张立科
 - ◆ 人民邮电出版社出版发行 北京市崇文区夕照寺街 14 号
邮编 100061 电子函件 315@pptph.com.cn
网址 <http://www.pptph.com.cn>
读者热线 010-67180876
北京汉魂图文设计有限公司制作
北京顺义振华印刷厂印刷
新华书店总店北京发行所经销
 - ◆ 开本: 787×1092 1/16
印张: 45.75
字数: 1 508 千字
印数: 1-2 000 册
- 2002 年 2 月第 1 版
2002 年 2 月北京第 1 次印刷

著作权合同登记 图字: 01 - 2000 - 2211 号

ISBN 7-115-09943-X/TP · 2667

定价: 75.00 元

本书如有印装质量问题, 请与本社联系 电话: (010) 67129223

内 容 提 要

本书是一本关于 Perl 语言的参考大全，涵盖了 Perl 语言的各个方面，包括 Perl 编程的基础知识、使用 Perl 进行编程、开发应用程序、Perl 的高级技巧，附录提供了函数参考、Perl 标准库和网上资源特定等内容。书中提供了许多实用的例程，并且给出了一些用 Perl 解决实际问题的建议。

本书并不是为任何特定的 Perl 用户组而写的，作为一本参考书，无论对于初学者，还是专业用户，这本书都是非常有价值的。

前言

写作这本书的目的是向 Perl 程序员提供一个全方位的参考。在本书中，我尽可能多地涵盖 Perl 的各个方面，也就是说，从 Perl 的内核到各种不同的函数以及对 Perl 语言各方面的用途都进行分析阐述。我甚至也讲述了如何调试软件、编写文档以及建立和发布应用程序或者模块的内容。另外我也顺便提及了 Perl 的一些非标准的特性和扩展程序，包括更好的网络、Web 编程和设计界面等内容。

本书提供了许多实用的例程，并且给出了一些用 Perl 解决实际问题的建议——而不是仅仅停留在如何使用 Perl 这个软件的功能的层面上。如果要快速地查找某个主题的相关参考，可以使用书后的附录和网站 www.osborne.com 上的相关内容，其中后者包括了 Perl 的基本组件、函数、标准库、CPAN(Comprehensive Perl Archive Network)、警告和错误的完全参考以及可用资源列表等信息。

与第一版相比有了哪些改进

新版本的最明显的改进在于讲述了 Perl 的最新版本，并且全面地调整了本书的布局。我们收到了关于上一版以及书中的一些漏洞和问题的大量的反馈信息，这些信息表明我们最需要的是关于 Perl 语言的很好的介绍性的参考手册。因此，我们对本书的第一版几乎完全进行了重写，为那些刚开始编程的人和从其他语言转过来的人提供一个丰富的有吸引力的参考。

本书的第 2 部分和第 3 部分做了修正，很多章中提供了面向应用的实际解决方案，这就意味着章节的顺序和内容有了改变。我们在很多方面扩充了内容，包括更多的网络方面的信息、数据库访问（现在包含了扩展的 DBI 和 WIN32::ODBC 模块）和 Web 编程。

本书的最后一部分指明了在把你的 Perl 程序变成一个分布式的应用后应做的一些处理，这包括调试脚本，使用 Perl 编译器来提高程序的效率和在程序中查错。Perl 文档这一章被修改成了教授如何利用已有的文档和信息来写自己的文档。最后，我们讨论了如何开发跨平台的应用和包以及如何分发你的应用或模块等方面的主题。

总之，在整本书中我尽量不以 UNIX 为中心，并且讲述了更多的关于开发跨平台应用的内容，而不是把这些内容压缩在最后的几章中。

本书适合于谁

这本书并不是为任何特定的 Perl 用户组而写的，简单地说，如果你使用 Perl 进行编程，你会发现这本书是有用的。作为一本参考书，大多数人可能会发现这本书非常适合放在桌子或书架上用来对 Perl 的某些函数或特性进行快速的查阅参考。无论对于初学者，还是专业用户，这本书都是“非常有价值”的。

同样，如果想要对于 Perl 的某个专题进行研究的话，这本书是适合的，因为本书中提供

了大量的可以参考的例子。

如何使用这本书

这本书适用于对某些专题进行查询，进而获取相关的信息。本书中的代码可以不需要多大改动就能移植到你的机器上去。但是要注意并不是所有的操作系统都支持 Perl 的全部功能，第 24 章讲述了开发跨平台应用方面的内容。

如果想获得真正的参考信息，就是说，想要查询一个函数的参数和返回的值——那么就可以使用本书后面的附录或者到网站 www.osborne.com 上查询。如果要查询关于某一个专题的详细内容，可以参考本书相关的章节，也可以同时参考书中的章节和书后的附录来获得更详细的内容。

章节内容安排

书中每一章对于 Perl 中的某一个专题进行了讨论。

第 1 章介绍了有关 Perl 的背景知识和 Perl 的基本功能，只有理解了这些内容，才能够更好地使用 Perl。

第 2 章介绍了使用 Perl 编程的基本过程，从在 UNIX、Windows、Mac OS 上安装 Perl 到如何编写和执行你的第一个脚本都有介绍，另外在这一章中还提到了在阅读本书和使用 Perl 时的一些专用术语。

第 3 章介绍了 Perl 语言的基本语法和解析规则，包括 Perl 脚本是如何执行的，以及组成 Perl 脚本的不同组件。

第 4 章介绍了 Perl 的变量类型。Perl 支持 3 种基本的变量类型——标量、数组和哈希结构。

第 5 章介绍了 Perl 语句以及控制 Perl 的流程和程序执行的控制结构的一些细节。

第 6 章详细讨论了 Perl 程序的不同组成部分，包括简单的子程序和比较复杂的模块。

第 7 章介绍了文件的使用等内容，包括基本的文件操作和从多个文件中读取数据。

第 8 章详细讨论了数据处理——这可能是 Perl 最有用的功能。本章详细介绍了字符串和数组的处理方法，然后仔细地讲述了正则表达式方面的内容。

第 9 章详细阐述了确认和捕捉错误的机制，以及如何在干扰用户的情况下解决问题。

第 10 章讨论了比较复杂的数据结构——数组和哈希结构。我们不局限于讨论这些结构的一般的用途，并且从不同方面观察它们如何用于编程。我们也可以通过参考（Reference）来学习如何访问和使用储存在变量中信息、内嵌的数据结构以及面向对象的编程的内容。本章的最后一节将讨论捆绑的数据接口——一个提供了将外部的数据源和内部的数据结构相关联的系统。

第 11 章讨论了获得系统信息的不同方法，例如关于用户、用户组、时间的信息，以及如何操作这些值，并且介绍了通过改变这些环境设置来影响脚本的执行的方法。

第 12 章介绍了在 Perl 中支持在标准网络套接字间通信的问题，如用于在 Internet 间进行通信所需的操作。

第 13 章讨论了如何使用外部数据库中的信息，包括使用 ODBC、Oracle、DBM 这些系统。

第 14 章介绍了通过标准文件句柄或者系统 V IPC 函数进行进程间通信的内容，另外也

讨论了建立、控制和外部进程进行通信的方法。

第 15 章讨论了执行脚本的一些高级的方法，包括在命令行中向脚本提供参数到在 Perl 解释器中动态地运行脚本(使用 eval)和如何使用线程作为运行多进程的替代等内容。

第 16 章讨论了在命令行上通过具有一定格式的接口在脚本和用户之间进行数据传递方法。

第 17 章详细讨论 Perl 的用户界面设计方面的内容，介绍了 Perl/Tk (Perl 的用户界面开发系统接口)。

第 18 章介绍了编写 Web 脚本的环境，这包括脚本所处的硬件环境和在 Web 服务器、用户浏览器和脚本 (CGI) 之间的通信。还介绍了关于 Web 开发过程和 Perl 的 Web 编程的细节，包括 Web 编程中关于安全性方面的检查列表 (CheckList)。

第 19 章介绍了如何通过许多编译标示来控制脚本的运行和执行的方法，这样便于及早发现脚本中可能出现的问题，提供附加的警告和错误信息以及改变脚本被解释的方式。

第 20 章描述通过外部的 C 函数进行 Perl 的扩展编程，以及将 Perl 嵌入到其他语言编写的程序中的方法。

第 21 章介绍了如何对 Perl 脚本进行调试，也就是说，在 Perl 解释器内部进行简单的或者比较深入复杂的调试。另外，我们讨论如何调试正则表达式以及如何使用 Perl Profiler 来确定哪一部分脚本的代码需要优化。

第 22 章讨论了 Perl 的编译器。不同的系统执行一个脚本并且产生许多不同的报文和输出格式，从简单的层次上说，这包括一个代码树并且描述一个脚本执行过程的描述文件，以及标量的结束，编译器提供了产生独立的二进制文件的机制。

第 23 章讨论了 Perl 文档使用和发展的情况，本章中介绍如何阅读文档以及一些编写自己的文档介绍。

第 24 章集中讨论了如何编写一个能够拥有跨平台兼容性的程序。本章中阐述了如何在不同的环境中使用 Perl 的一些基本问题，值得去注意的一些问题以及让脚本能够更好地适应不同操作系统的一些方法。

第 25 章讨论了向 Perl 团体发布 Perl 模块的过程。特别地，本章讲述了 MakeMaker 工具，它用于建立能够被扩展程序编译和自动装载的 Makefile。

附录中包括了一些能用于编程的参考资料，包括 Perl 的所有函数 (附录 A)、标准 Perl 库 (附录 B)，以及相关资源——书籍、站点、邮件列表和新闻组 (附录 C)。

www.osborne.com 提供了关于 Perl 核心架构的参考 (附录 A)，所有扩展程序的参考 (附录 B)，以及解释器所产生的关于告警和错误信息的完整列表 (附录 C)。

本书的约定

本书中所有的关键字都用黑体表示，但所列出的函数并不使用括号。这是因为在 Perl 中所继承的 C 语言的函数可能会按照下列形式来表达 **this** ()。

对函数描述的格式为用等宽度的字体表示。

注意：关于某些特定主题的一些附加信息用这种小字体来表示。同样，某些在编程时的警告信息也用这种方式来表示。

如何联系本书的作者

欢迎对本书提出宝贵意见,我尤其欢迎关于达到不同编程目的的更好的方法指导和建议,特别是利用 Perl 这样的功能强大而且变化非常多的语言。

读者最好通过电子邮件 (E-mail) 和我联系,我的邮件地址是 books@mcwords.com 或者 mc@whoever.com。另外也可以访问我的个人主页 <http://www.mcwords.com>, 其中包含了许多脚本的资源 和 升级信息, 另外还有本书的全部内容, 这本书的主页可以在 <http://www.mcwords.com/projects/books/pcr2e> 中找到。

目 录

第 1 部分 Perl 编程基础

第 1 章 Perl 的背景	1
1.1 版本和命名协定	1
1.1.1 Perl、perl 还是 PeRL	2
1.1.2 Perl 5.6 之前的各种 Perl 版本	2
1.2 Perl 的历史	2
1.3 Perl 的主要特点	2
1.3.1 免费	3
1.3.2 简单易学、简明易读	3
1.3.3 快速高效	3
1.3.4 可扩展性	3
1.3.5 灵活的数据类型	4
1.3.6 面向对象	4
1.3.7 交流协作	4
1.4 编译器或解释器	4
1.5 与 Perl 类似的语言	4
1.5.1 UNIX Shell	5
1.5.2 Tcl	5
1.5.3 Python	5
1.5.4 Java	6
1.5.5 C/C++	6
1.5.6 awk/gawk	6
1.6 关于 Perl 的误区	7
1.7 Perl 的成功业绩	8
第 2 章 Perl 概述	9
2.1 安装及使用 Perl	9
2.1.1 编写一个 Perl 脚本	10
2.1.2 UNIX 下的 Perl	10
2.1.3 Windows 下的 Perl	12
2.1.4 Mac OS 下的 Perl	15
2.2 Perl 中的组件	18
2.2.1 变量	18

2.2.2	操作符	19
2.2.3	语句	19
2.2.4	子程序(函数)	19
2.2.5	模块	20
2.3	接下去如何使用本书	20
第3章	Perl 解析规则	21
3.1	脚本执行过程	21
3.2	语法和解析规则	21
3.2.1	基本语法	22
3.2.2	注释语句	23
3.2.3	组件确认	24
3.2.4	操作符和优先级	25
3.2.5	独立词	37
3.2.6	上下文	38
3.2.7	逻辑值	39
3.3	Perl 的代码体系	39
第4章	变量和数据	41
4.1	基本命名规则	41
4.2	标量变量	42
4.3	字面量	42
4.3.1	数字字面量	42
4.3.2	字符串字面量	43
4.3.3	字符串引用	45
4.3.4	数组的内插	46
4.3.5	本地文档	46
4.3.6	V-字符串	47
4.4	数组	48
4.5	哈希结构	50
4.6	列表	53
4.7	类型标识(Typeglobs)	55
4.8	被定义的函数和未定义值	55
4.9	默认值	56
4.10	其他标记	56
4.11	特殊变量	56
第5章	语句和控制结构	65
5.1	代码块	65
5.2	条件语句	65
5.3	循环语句	68
5.3.1	while 循环	68
5.3.2	until 循环	68

5.3.3	for 循环	69
5.3.4	foreach 循环	69
5.3.5	continue 块	70
5.3.6	标签	70
5.3.7	循环控制	71
5.3.8	非限定块	72
5.3.9	goto 语句	73
第 6 章	子程序、包和模块	74
6.1	函数	74
6.1.1	参数	76
6.1.2	返回值	82
6.1.3	错误通告	83
6.1.4	上下文	83
6.1.5	属性	84
6.1.6	原型	85
6.2	包	86
6.2.1	包符号表	88
6.3	特殊块	88
6.4	模块	89
6.4.1	建立模块	90
6.4.2	Exporter 模块	91
6.4.3	use 和 require 的比较	91
6.4.4	no 语句	93
6.4.5	do 语句	93
6.5	作用域	93
6.5.1	“my” 的效果	94
6.5.2	“local” 的效果	95
6.5.3	“our” 的效果	95
6.5.4	循环中的作用域	96
6.6	自动装载	96
第 7 章	文件操作	98
7.1	文件句柄	98
7.1.1	打开和关闭文件	99
7.1.2	输入/输出句柄	104
7.1.3	文件定位	109
7.1.4	混合控制函数	110
7.2	文件管理	112
7.2.1	文件信息	112
7.2.2	基本的文件管理	116
7.2.3	目录访问	118

7.2.4	管理目录	120
7.2.5	使用 fcntl 函数进行文件的控制	122
7.2.6	通过 ioctl 函数进行 I/O 控制	122
7.2.7	select 函数	123
7.2.8	文件加锁	125
第 8 章	数据操作	127
8.1	数字处理	127
8.1.1	abs —— 绝对值	127
8.1.2	int —— 将浮点数转换为整数	127
8.1.3	exp —— e 为底的乘方	128
8.1.4	sqrt —— 开方运算	128
8.1.5	log —— 对数函数	128
8.1.6	三角函数	128
8.1.7	进制转换	129
8.1.8	字符和数字之间的转换	129
8.1.9	随机数	130
8.1.10	小整数的操作	131
8.2	字符串操作	131
8.2.1	字符串连接	131
8.2.2	字符串长度	132
8.2.3	改变字符大小写	132
8.2.4	行结束字符的消除	132
8.2.5	字符串定位	133
8.2.6	提取子串	133
8.2.7	栈	134
8.2.8	数组切片	135
8.2.9	join 函数	136
8.2.10	split 函数	137
8.2.11	grep 函数	138
8.2.12	map 函数	138
8.2.13	sort 函数	139
8.2.14	reverse	140
8.3	正则表达式	140
8.3.1	模式修饰符	141
8.3.2	匹配操作符	142
8.3.3	替换操作符	143
8.3.4	转换	145
8.3.5	正则表达式的元素	147
8.3.6	正则表达式变量	156
8.3.7	正则表达式扩展名 / 断言	157

8.3.8	预编译表达式	158
8.3.9	正则表达式的支持函数	159
8.4	Unicode	160
8.4.1	Perl 对于 Unicode 的解释	161
8.4.2	Perl 中的 Unicode 字符	161
8.4.3	Unicode 作用于 Perl 的效果	162
第 9 章	错误和捕获错误	163
9.1	在脚本中加入错误确认的代码	163
9.1.1	错误确认方法	163
9.1.2	错误确认方针	164
9.2	什么是要确认的错误	165
9.3	什么是不需要确认的错误	165
9.4	什么时候停止确认	166
9.4.1	尽量避免画蛇添足	166
9.4.2	函数、返回值和内嵌结构	167
9.5	表示错误的信息和数字	168
9.6	在脚本中报告错误	169
9.6.1	Warn 函数	169
9.6.2	Die 函数	169
9.6.3	指示和标记	170
9.7	在模块中进行报错	170
9.7.1	Carp 函数	171
9.7.2	Cluck 函数	171
9.7.3	Croak 函数	171
9.7.4	Confess 函数	172

第 2 部分 使用 Perl 编程

第 10 章	复杂数据结构	173
10.1	访问打包的数据结构	173
10.2	引用	176
10.2.1	建立硬引用	177
10.2.2	反引用	179
10.2.3	判断引用中的数据类型	181
10.2.4	软引用	182
10.3	复杂结构	182
10.3.1	二维数组	183
10.3.2	二维哈希结构	186
10.3.3	哈希结构的数组	189
10.3.4	数组的哈希结构	191
10.3.5	多于二维的结构	193

10.4	对象	195
10.4.1	关于对象的基础知识	195
10.4.2	建立和使用对象	196
10.4.3	对象方法	197
10.4.4	类和类的继承	200
10.4.5	析构函数和垃圾收集	200
10.4.6	Perl 中对象结构同其他语言的比较	201
10.5	使用捆绑	201
10.5.1	建立新的捆绑类	202
第 11 章	系统信息	206
11.1	用户和用户组	206
11.1.1	获取 UNIX 的 Password	206
11.1.2	获得 UNIX 的用户组	208
11.1.3	Password 加密	208
11.2	Time	209
11.2.1	gmtime 和 localtime	209
11.2.2	Time 函数	210
11.2.3	时间值比较	210
11.2.4	将日期和时间转换为起始时间值	210
11.2.5	时间算法	211
11.2.6	times 函数	213
11.2.7	sleep 函数	214
11.2.8	Alarms	214
11.3	环境变量	215
第 12 章	网络	218
12.1	获得网络信息	219
12.1.1	主机	219
12.1.2	协议	220
12.1.3	服务	221
12.1.4	网络	221
12.2	套接字模块	222
12.2.1	地址解析和转换	222
12.2.2	地址常量	222
12.2.3	套接字结构	223
12.2.4	行结束常量	223
12.3	利用套接字进行通信	223
12.3.1	连接远端套接字	223
12.3.2	监听套接字连接信息	226
12.4	使用 IO::Socket	231
12.4.1	客户端	232

12.4.2	服务器端	232
12.4.3	使用 IO::Socket 和 IO::Select	232
12.4.4	获取套接字地址	234
12.4.5	关闭套接字	234
12.4.6	套接字选项	235
12.4.7	数据传递	235
12.5	Graham Barr 的 libnet 包族	236
12.6	Gisle Aas 的 LWP 包族	238
第 13 章	数据库系统	240
13.1	文本数据库	241
13.1.1	限定数据库	242
13.1.2	固定长度的记录	242
13.2	DBM 数据库	245
13.2.1	DBM 的实现	246
13.2.2	DBM 函数	247
13.2.3	使用捆绑 DBM 数据库	248
13.2.4	在 DBM 各种格式之间转换	250
13.2.5	使用 AnyDBM_File	251
13.2.6	GDBM 的特性	251
13.2.7	Berkeley DB 特性	252
13.2.8	在 DBM 数据库中存放复杂的数据	256
13.2.9	DBM 中的关系型数据库	259
13.3	数据库的文件锁定	260
13.4	使用 DBI 和 Win32::ODBC 工具集	261
13.4.1	DBI	262
13.4.2	Win32::ODBC	262
13.4.3	数据库机制和兼容	263
13.4.4	数据库连接	263
13.4.5	执行简单查询	264
13.4.6	执行扩展查询	265
13.4.7	关闭数据库连接	266
13.4.8	确认错误	266
13.5	复习 SQL 语言	267
13.5.1	SQL 语句	268
第 14 章	进程间通信	274
14.1	进程	274
14.1.1	控制 Perl 运行	274
14.1.2	进程信息	275
14.1.3	进程组	275
14.1.4	进程优先级	275

14.2	信号	276
14.2.1	信号句柄	277
14.2.2	_WARN_和_DIE_信号	279
14.2.3	发送信号	281
14.3	管道	281
14.3.1	匿名管道	281
14.3.2	双向通信	282
14.3.3	命名管道	282
14.3.4	Windows 下的命名管道	284
14.3.5	安全管道	286
14.4	执行附加进程	287
14.4.1	运行其他程序	287
14.4.2	替换当前脚本	288
14.4.3	建立子进程	288
14.5	其他的函数调用	293
14.6	system V IPC	294
第 15 章	其他运行方法	295
15.1	以命令行形式执行 Perl	295
15.1.1	特殊操作	300
15.2	Perl 的环境变量	301
15.3	Perl in Perl (eval)	302
15.3.1	使用 eval EXPR	302
15.3.2	使用 eval BLOCK	303
15.3.3	捕捉意外情况	303
15.3.4	返回信息	303
15.3.5	eval 和_DIE_信号句柄	303
15.4	线程	304
15.4.1	多任务如何实现	304
15.4.2	从多任务到多线程	305
15.4.3	线程和多进程的比较	306
15.4.4	线程和 select () 函数的区别	306
15.4.5	线程和 Perl	306
15.4.6	创建一个新线程	306
15.4.7	通过匿名者函数建立线程	307
15.4.8	线程控制	307
15.4.9	变量控制	308
15.4.10	队列	309
15.4.11	标志 (semaphore)	309
15.4.12	信号	309
15.5	安全性	310

15.5.1 使用 Taint 模式	310
15.5.2 safe 模块和 opcode 模块	312

第3部分 开发应用

第16章 用户界面工具	315
16.1 处理命令行参数	315
16.1.1 Getopt::std	316
16.1.2 Getopt::Long	317
16.2 Perl 输出显示机制	319
16.2.1 标题和脚注	322
16.2.2 格式函数	322
16.2.3 格式变量	323
第17章 高级用户界面	324
17.1 终端中的一些处理	324
17.2 使用 TK	325
17.2.1 在 UNIX 下安装 TK	326
17.2.2 在 Windows 下安装 TK	326
17.2.3 接触 TK	326
17.2.4 事件循环	328
17.2.5 部件	330
17.2.6 控制窗口位置大小	344
17.2.7 简化处理	347
第18章 Web 中的 Perl	353
18.1 HTML	353
18.2 统一资源定位符 URL	354
18.3 Web 操作浏览	354
18.4 环境	355
18.5 通用网关接口 (CGI)	358
18.5.1 数据提取	359
18.5.2 返回相关的数据到用户浏览器中	361
18.5.3 文档主体	364
18.6 Web 编程	364
18.7 CGI 模块	365
18.7.1 cookies	368
18.8 解析 HTML	369
18.9 解析 XML	370
18.10 调试程序	373
18.11 安全性	374
第19章 通过标记控制 Perl 的执行	377
19.1 告警	377