

支持
C
和
PERL

CGI

编程指南

[美] Stephen Asbury 等著

胡宝兰 李梅 译
陈志强 谢建军



WAITE
GROUP
PRESS™



浙江科学技术出版社
西蒙与舒斯特国际出版公司

CGI 编程指南

[美] Stephen Asbury 等著
胡宝兰 李 梅 译
陈志强 谢建军 译

浙江科学技术出版社
西蒙与舒斯特国际出版公司

TP31

Stephen Asbury: CGI HOW-TO

Authorized translation from the English language edition published by
WAITE GROUP PRESS

Copyright©1997 by WAITE GROUP PRESS

All rights reserved. For Sale in Mainland China Only

本书中文简体版由浙江科学技术出版社和美国西蒙与舒斯特国际出版公司
合作出版, 未经出版者书面许可, 本书任何部分均不得以任何方式复制或抄袭。
本书封面贴有西蒙与舒斯特防伪标签, 无标签者不得销售。

版权所有, 翻印必究。

浙江省版权局著作权合同登记号 图字: 11-1998-06 号

CGI 编程指南

[美] Stephen Asbury 等著

胡宝兰 李 梅 译

陈志强 谢建军

浙江科学技术出版社 出版
西蒙与舒斯特国际出版公司

杭州富春印务有限公司印刷

浙江省新华书店发行

开本 787 × 1092 1/16 印张 31 字数 727 000

1998 年 4 月第一版

1998 年 4 月第一次印刷

ISBN 7-5341-1117-x/TP-40

定价: 55.00 元

责任编辑: 熊盛新

封面设计: 金 晖

译 者 序

各种各样的公司在万维网(World Wide Web)上开发的各种应用软件都是用 CGI(Common Gateway Interface, 中文含义为通用网关接口)编写的, CGI 程序的主要作用是在 Web 服务器和 Web 应用程序之间建立起通信链接。

本书通过各种由浅入深的问答将读者一步步地引向 CGI 编程的轨道上来。书中各章的结构是先提出问题, 然后, 依次介绍技术要点和解决问题的步骤, 再谈谈在实际开发中的工作步骤, 最后附上一些技术性的注释。本书是一本内容广泛的指导性书籍, 无论对小型应用程序的开发人员还是大型软件包的开发专家, 它都是一本不可多得的工具书。

本书是由 Stephen Asbury、Gary Jason Mathews、Selena Sol 和 Kevin Greer 这四位网络专家共同执笔的, 他们都有多年从事网络编程的经验, 并且, 目前仍然在相当重要的网络机构中工作。其中, Stephen Asbury 是美国 Paradigm Research 研究所的工程师, 主要负责技术培训工作; Gary Jason Mathews 是美国航空航天局 Goddard 太空发射中心的研究员, 主要负责互联网开发工作; Selena Sol 是一位 CGI 编程专家, 曾在美国多家大型网络机构中工作; Kevin Greer 是一位常年工作在美国和加拿大的网络咨询和软件开发专家, 他们的阅历和丰富的经验足以使本书成为一本让读者喜爱的好书。

WAITE GROUP 出版社是西蒙与舒斯特国际出版公司下属的一家子公司, 该出版社为本书和相关的编程书籍专门设立了一个主页, 在这里, 你可以找到本书的所有测试程序, 并能找到该出版社最新出版的各种编程书籍的内容简介和相应的程序代码。该出版社的主页是 <http://www.informant.com/wi/witest.htm>。

当然, 一本书的出版与很多人的努力是分不开的, 参加本书翻译工作的还有张立平、崔如文、程刚、王伍、何珊、洪键等, 他们为本书的顺利出版作出了很大的努力。

最后, 希望读者阅读本书后能有所收益。由于时间仓促及译者水平有限, 书中的错误和不足之处在所难免, 恳请读者批评指正!

译 者

1998 年 3 月于杭州

015115/0

作者序

1. 本书的主要内容

越来越多的公司在万维网(World Wide Web)上开发各种应用软件,它们涵盖的范围极广,从公司的内部信息管理到外部的商业服务,应有尽有。这些软件都是用 CGI(Common Gateway Interface,中文含义为通用网关接口)编写的。CGI 在 Web 服务器和 Web 应用程序之间建立了通信链接。CGI 项目可大可小,小到简单的 Perl 程序,大到复杂的 C 程序。本书从简单的开发问题出发到介绍如何处理 Web 应用程序的高级技巧,各种不同的信息可满足广大读者在实际开发中的不同需求。

CGI Script 是一种由 Web 服务器运行的程序(本书都将其统称为 CGI 程序),它的作用是处理 Web 服务器的请求。这里所指的请求有多种形式,最常见的一种形式是用户通过 HTML 页面上的表格发出的,用户只要点击指向 CGI 程序的链接,而 HTML 页面则包含程序端的所有资源。也就是说,是 CGI 程序为用户提供了这么一个图形式的界面。当 Web 服务器接收到用户信息后,它就会把用户的输入信息传送到相应的 CGI 程序,只要有对应的资源,程序就会把应答返回到服务器,并由服务器将结果反馈给用户。

CGI 程序用于提供 Web 页背后的逻辑,这种逻辑可将用户通过 Web 表格输入的数据转换成报告,或将同样的数据插入到相关的数据库中。编写 CGI 程序涉及到 HTTP 或其他强大的功能,因此,我们常要开发应用程序组,这就要求我们能在程序间共享数据。本书用了几章的篇幅专门介绍从多个 CGI 程序中创建 Web 应用程序。在这一部分里,我们还探讨了其他一些问题,如程序的安全性和安装问题。

总的说来,本书是一本内容广泛的指导性图书,无论是小型应用程序的开发人员还是大型软件包的开发小组,本书都可作为他们的指导工具书。

2. 问答格式

本书通过各种由浅入深的问答将读者一步步地引向 CGI 编程的轨道上来。书中各章的结构是先提出问题,然后,依次介绍技术要点和解决问题的步骤,再谈谈在实际开发中的工作步骤,最后附上一些技术性的注释。Web 服务器将用户提供给它的数据进行解码,而 HTTP 服务器正是程序与客户所使用的 Web 浏览器之间的接口。

CGI 程序也提供数据给客户浏览器。例如, CGI 程序可根据用户安全级别显示不同的 Web 页面。当结合程序的综合反应行为时,程序中的提供数据行为将根据用户的输入发出相应的数据库查询请求,并把查询结果反馈给用户。产生输出是 HTTP 服务器和 CGI 程序的伴随物,本书的中间部分正是介绍 CGI 程序可提供的各种输出。其中,值得一提的是动态产生的 HTML,尤其是 HTML 文档转换和动态数据的插入。动态产生的 HTML 允许开发人员根

据 Web 页的设计来设计 HTML，而不用把 HTML 转换成 C 或 Perl 代码。

3. 本书的主要内容

本书共分 12 章和两个附录，分别从不同的角度来介绍 CGI 编程中所遇到的各种问题，它们的主要内容如下：

第 1 章“通过发出表格请求来读取数据”；第 2 章“使用传送到 CGI 程序中的数据”；第 3 章“通过图像映射请求读取数据”；第 4 章“对请求的标准应答”；第 5 章“在动态输出过程中操纵现有的 HTML 文件”；第 6 章“对请求的高级应答”；第 7 章“访问其他服务器”；第 8 章“将多个脚本文件连接到一个应用程序中”；第 9 章“测试程序”；第 10 章“安装程序”；第 11 章“CGI 安全问题”；第 12 章“充分利用客户机和服务器”。两个附录为：附录 A “CGI 支持的环境变量”；附录 B “HTML 表格元素”。

4. 本书作者简介

Stephen Asbury，美国 Paradigm Research 研究所的工程师，主要负责技术培训工作。

Gary Jason Mathews，美国航空航天局 Goddard 太空发射中心的研究员，主要负责互联网开发工作，他的网址为 <http://coney.gsfc.nasa.gov/Mathews/>。

Selena Sol 是一位 CGI 编程专家，曾在美国多家大型网络机构工作，他的网址为 <http://www.eff.org/~erict/Scripts/>。

Kevin Greer 是一位常年工作在美国和加拿大的网络咨询和软件开发专家，他的网址为 <http://fxfx.com/kgr>。

目 录

第 1 章 通过发出表格请求来读取数据	1
1.1 阅读环境变量	1
1.1.1 问 题	1
1.1.2 技术要点	2
1.1.3 步 骤	2
1.1.4 工作方式	2
1.1.5 注 释	2
1.2 确定所发出的表格请求的类型	4
1.2.1 问 题	4
1.2.2 技术要点	4
1.2.3 步 骤	4
1.2.4 工作方式	4
1.2.5 注 释	5
1.3 从 GET 请求中读取数据	5
1.3.1 问 题	5
1.3.2 技术要点	5
1.3.3 步 骤	6
1.3.4 工作方式	9
1.3.5 注 释	10
1.4 从 POST 表格请求中读取数据	12
1.4.1 问 题	12
1.4.2 技术要点	12
1.4.3 步 骤	12
1.4.4 工作方式	16
1.4.5 注 释	17
1.5 解释来自表格请求的数据	20
1.5.1 问 题	20
1.5.2 技术要点	20
1.5.3 步 骤	20
1.5.4 工作方式	25
1.5.5 注 释	25
1.6 对表格请求中的数据进行解码	35
1.6.1 问 题	35
1.6.2 技术要点	35
1.6.3 步 骤	35
1.6.4 工作方式	39
1.6.5 注 释	40

1.7 保存表格请求中的数据	43
1.7.1 问题	43
1.7.2 技术要点	43
1.7.3 步骤	43
1.7.4 工作方式	47
1.7.5 注释	47
1.8 读取命令行中传送给程序的数据	51
1.8.1 问题	51
1.8.2 技术要点	51
1.8.3 步骤	51
1.8.4 工作方式	54
1.8.5 注释	54
1.9 在同一个程序中同时支持 GET 请求和 POST 请求	55
1.9.1 问题	55
1.9.2 技术要点	55
1.9.3 步骤	55
1.9.4 工作方式	59
1.9.5 注释	60
1.10 解释表格数据中的一键多值现象	64
1.10.1 问题	64
1.10.2 技术要点	64
1.10.3 步骤	64
1.10.4 工作方式	71
1.10.5 注释	71
第2章 使用传送到 CGI 程序中的数据	80
2.1 确定服务器软件的类型	80
2.1.1 问题	80
2.1.2 技术要点	80
2.1.3 步骤	80
2.1.4 工作方式	81
2.1.5 注释	83
2.2 确定自我引用的 URL 的服务器名字	87
2.2.1 问题	87
2.2.2 技术要点	88
2.2.3 步骤	88
2.2.4 工作方式	90
2.2.5 注释	91
2.3 确定服务器所使用的协议	92
2.3.1 问题	92
2.3.2 技术要点	92
2.3.3 步骤	92
2.3.4 工作方式	93

2.3.5 注 释	94
2.4 确定所使用的 CGI 的版本	97
2.4.1 问 题	97
2.4.2 技术要点	97
2.4.3 步 骤	97
2.4.4 工作方式	99
2.4.5 注 释	99
2.5 确定程序的路径	101
2.5.1 问 题	101
2.5.2 技术要点	101
2.5.3 步 骤	101
2.5.4 工作方式	103
2.5.5 注 释	104
2.6 确定客户机的 IP 地址	106
2.6.1 问 题	106
2.6.2 技术要点	106
2.6.3 步 骤	106
2.6.4 工作方式	107
2.6.5 注 释	108
2.7 确定客户机的名字	109
2.7.1 问 题	109
2.7.2 技术要点	109
2.7.3 步 骤	109
2.7.4 工作方式	110
2.7.5 注 释	111
2.8 确定客户所使用的浏览器	113
2.8.1 问 题	113
2.8.2 技术要点	113
2.8.3 步 骤	113
2.8.4 工作方式	115
2.8.5 注 释	116
2.9 确定客户所能理解的数据类型	118
2.9.1 问 题	118
2.9.2 技术要点	118
2.9.3 步 骤	118
2.9.4 工作方式	121
2.9.5 注 释	122
第 3 章 通过图像映射请求读取数据	125
3.1 操纵来自静态图像的输入	125
3.1.1 问 题	125
3.1.2 技术要点	125
3.1.3 步 骤	125

3.1.4 工作方式	136
3.1.5 注 释	138
3.2 操纵来自动态图像映射的输入	140
3.2.1 问 题	140
3.2.2 技术要点	140
3.2.3 步 骤	141
3.2.4 工作方式	145
3.2.5 注 释	146
第 4 章 对请求的标准应答	152
4.1 选择输出类型	152
4.1.1 问 题	152
4.1.2 技术要点	153
4.1.3 步 骤	153
4.1.4 工作方式	154
4.1.5 注 释	154
4.2 对输出进行初始化	154
4.2.1 问 题	154
4.2.2 技术要点	155
4.2.3 步 骤	155
4.2.4 工作方式	158
4.2.5 注 释	158
4.3 如何输出对本地文档的引用	160
4.3.1 问 题	160
4.3.2 技术要点	160
4.3.3 步 骤	160
4.3.4 工作方式	162
4.3.5 注 释	162
4.4 输出完整的 URL	163
4.4.1 问 题	163
4.4.2 技术要点	163
4.4.3 步 骤	163
4.4.4 工作方式	166
4.4.5 注 释	166
4.5 输出一个本地文档	168
4.5.1 问 题	168
4.5.2 技术要点	168
4.5.3 步 骤	168
4.5.4 工作方式	171
4.5.5 注 释	171
4.6 输出动态的 HTML 页面	174
4.6.1 问 题	174
4.6.2 技术要点	174

4.6.3 步 骤	174
4.6.4 工作方式	178
4.6.5 注 释	178
第 5 章 在动态输出过程中操纵现有的 HTML 文件	183
5.1 将 HTML 转换成标记符和主体文本	183
5.1.1 问 题	183
5.1.2 技术要点	183
5.1.3 步 骤	184
5.1.4 工作方式	195
5.1.5 注 释	196
5.2 为表格设置行为或请求方法	207
5.2.1 问 题	207
5.2.2 技术要点	207
5.2.3 步 骤	207
5.2.4 工作方式	211
5.2.5 注 释	211
5.3 找到输入项目并确定其类型	213
5.3.1 问 题	213
5.3.2 技术要点	213
5.3.3 步 骤	213
5.3.4 工作方式	217
5.3.5 注 释	217
5.4 修改文本/隐含文件/口令输入项目的设置值或大小	221
5.4.1 问 题	221
5.4.2 技术要点	221
5.4.3 步 骤	221
5.4.4 工作方式	225
5.4.5 注 释	225
5.5 管理复选框的状态	229
5.5.1 问 题	229
5.5.2 技术要点	229
5.5.3 步 骤	229
5.5.4 工作方式	233
5.5.5 注 释	234
5.6 管理一系列圆按钮的状态	236
5.6.1 问 题	236
5.6.2 技术要点	237
5.6.3 步 骤	237
5.6.4 工作方式	241
5.6.5 注 释	241
5.7 使 Submit 按钮有固定的名称	244
5.7.1 问 题	244

5.7.2 技术要点	244
5.7.3 步 骤	244
5.7.4 工作方式	248
5.7.5 注 释	248
5.8 改变文本区域的设置值和大小	250
5.8.1 问 题	250
5.8.2 技术要点	250
5.8.3 步 骤	250
5.8.4 工作方式	254
5.8.5 注 释	254
5.9 管理选择列表中的选项	256
5.9.1 问 题	256
5.9.2 技术要点	256
5.9.3 步 骤	257
5.9.4 工作方式	262
5.9.5 注 释	262
第 6 章 对请求的高级应答	268
6.1 支持自定义的 HTML 命令	268
6.1.1 问 题	268
6.1.2 技术要点	268
6.1.3 步 骤	268
6.1.4 工作方式	271
6.1.5 注 释	272
6.2 将使用当前格式的数据插入到现有的 HTML 文件中	274
6.2.1 问 题	274
6.2.2 技术要点	274
6.2.3 步 骤	274
6.2.4 工作方式	277
6.2.5 注 释	277
6.3 将 HTML 链接插入到现有的 HTML 文件中	278
6.3.1 问 题	278
6.3.2 技术要点	278
6.3.2 步 骤	279
6.3.4 工作方式	281
6.3.5 注 释	282
6.4 将选择列表插入到现有的 HTML 文件中	283
6.4.1 问 题	283
6.4.2 技术要点	283
6.4.3 步 骤	283
6.4.4 工作方式	286
6.4.5 注 释	286
6.5 将 Submit 按钮的列表插入到现有的 HTML 文件中	287

6.5.1 问 题	287
6.5.2 技术要点	287
6.5.3 步 骤	288
6.5.4 工作方式	290
6.5.5 注 释	290
6.6 直接将 HTTP 传送给客户机	292
6.6.1 问 题	292
6.6.2 技术要点	292
6.6.3 步 骤	292
6.6.4 工作方式	296
6.6.5 注 释	298
6.7 使用同一输入执行另一个 CGI 程序	298
6.7.1 问 题	298
6.7.2 技术要点	298
6.7.3 步 骤	299
6.7.4 工作方式	301
6.7.5 注 释	302
6.8 使用新的输入执行另一个 CGI 文件	302
6.8.1 问 题	302
6.8.2 技术要点	302
6.8.3 步 骤	302
6.8.4 工作方式	307
6.8.5 注 释	307
第 7 章 访问其他服务器	316
7.1 通过 CGI 程序发送电子函件	316
7.1.1 问 题	316
7.1.2 技术要点	316
7.1.3 步 骤	316
7.1.4 工作方式	321
7.1.5 注 释	322
7.2 动态创建图像	327
7.2.1 问 题	327
7.2.2 技术要点	327
7.2.3 步 骤	328
7.2.4 工作方式	333
7.2.5 注 释	334
7.3 存取 dbm 文件	340
7.3.1 问 题	340
7.3.2 技术要点	340
7.3.3 步 骤	340
7.3.4 工作方式	344
7.3.5 注 释	346

第 8 章 将多个脚本文件连接到一个应用程序中	354
8.1 将隐含变量与格式数据一起传送	354
8.1.1 问 题	354
8.1.2 技术要点	354
8.1.3 步 骤	355
8.1.4 工作方式	358
8.1.5 注 释	359
8.2 用一个程序为多种表格服务	359
8.2.1 问 题	359
8.2.2 技术要点	359
8.2.3 步 骤	360
8.2.4 工作方式	365
8.2.5 注 释	367
8.3 创建多分支查询系统	368
8.3.1 问 题	368
8.3.2 技术要点	368
8.3.3 步 骤	368
8.3.4 工作方式	371
8.3.5 注 释	373
8.4 使用 URL 传送数据	373
8.4.1 问 题	373
8.4.2 技术要点	373
8.4.3 步 骤	374
8.4.4 工作方式	376
8.4.5 注 释	376
8.5 创建一个可以在 HTML 页面搜索关键词的搜索引擎	377
8.5.1 问 题	377
8.5.1 技术要点	377
8.5.3 步 骤	377
8.5.4 工作方式	382
8.5.5 注 释	382
8.6 在程序之间传递变量	383
8.6.1 问 题	383
8.6.2 技术要点	383
8.6.3 工作方式	388
8.6.4 注 释	388
8.7 创建一个基于 Web 的公告系统	388
8.7.1 问 题	388
8.7.2 技术要点	388
8.7.3 步 骤	389
8.7.4 工作方式	406
8.7.5 注 释	406

8.8 创建一个完善的商业系统	407
8.8.1 问 题	407
8.8.2 技术要点	407
8.8.3 步 骤	409
8.8.4 工作方式	425
8.8.5 注 释	426
第9章 测试程序	427
9.1 建立程序测试方案	427
9.1.1 问 题	427
9.1.2 方 法	427
9.1.3 步 骤	427
9.1.4 工作方式	433
9.2 在使用浏览器和服务器的情况下测试 CGI 程序	435
9.2.1 问 题	435
9.2.2 方 法	435
9.2.3 步 骤	436
9.2.4 工作方式	438
9.2.5 注 释	438
9.3 在只使用 Web 服务器, 而不使用浏览器的情况下测试程序	438
9.3.1 问 题	438
9.3.2 步 骤	438
9.3.3 工作方式	439
9.3.4 注 释	439
9.4 用 Web 服务器和浏览器测试程序	439
9.4.1 问 题	439
9.4.2 技术要点	439
9.4.3 步 骤	439
9.4.4 工作方式	440
9.4.5 注 释	440
9.5 测试过程自动化	440
9.5.1 问 题	440
9.5.2 方 法	440
9.5.3 步 骤	440
9.5.4 工作方式	442
9.5.5 注 释	442
9.6 避免常见的编程错误	442
9.6.1 问 题	442
9.6.2 方 法	442
9.6.3 步 骤	442
9.6.4 工作方式	444
9.6.5 注 释	444

第 10 章 安装程序	445
10.1 在 UNIX 上安装 CGI 程序	445
10.1.1 问 题	445
10.1.2 方 法	445
10.1.3 步 骤	445
10.1.4 工作方式	449
10.1.5 注 释	450
10.2 在 Windows NT 上安装程序	450
10.2.1 问 题	450
10.2.2 技术要点	450
10.2.3 步 骤	450
10.2.4 工作方式	451
10.2.5 注 释	452
10.3 在 Windows 3.1 上安装程序	452
10.3.1 问 题	452
10.3.2 技术要点	452
10.3.3 步 骤	453
10.3.4 注 释	453
10.4 在 Macintosh 中安装程序	453
10.4.1 问 题	453
10.4.2 技术要点	453
10.4.3 步 骤	454
10.4.4 工作方式	454
10.4.5 注 释	454
第 11 章 CGI 安全问题	455
11.1 如何避免常见的 CGI 安全问题	455
11.1.1 问 题	455
11.1.2 方 法	455
11.1.3 步 骤	456
11.1.4 工作方式	458
11.1.5 注 释	459
11.2 使用 NCSA 服务器请求用户身份验证	459
11.2.1 问 题	459
11.2.2 方 法	459
11.2.3 步 骤	459
11.2.4 工作方式	460
11.2.5 注 释	461
11.3 使用 CERN 服务器请求用户身份验证	461
11.3.1 问 题	461
11.3.2 方 法	461
11.3.3 步 骤	461
11.3.4 工作方式	462

11.3.5 注 释	463
11.4 使用 Netscape 服务器请求用户身份验证	463
11.4.1 问 题	463
11.4.2 方 法	463
11.4.3 步 骤	464
11.4.4 工作方式	464
11.4.5 注 释	464
第 12 章 充分利用客户机和服务器	465
12.1 使用服务器端的包含文件	465
12.1.1 问 题	465
12.1.2 技术要点	465
12.1.3 步 骤	465
12.1.4 工作方式	468
12.1.5 注 释	469
12.2 迫使客户机不断更新	469
12.2.1 问 题	469
12.2.2 技术要点	469
12.2.3 步 骤	470
12.2.4 工作方式	471
12.2.5 注 释	471
12.3 运用现有条件提供永久的数据	471
12.3.1 问 题	471
12.3.2 技术要点	471
12.3.3 步 骤	472
12.3.4 工作方式	472
12.3.5 注 释	473
12.4 给 Netscape 服务器添加特性	473
12.4.1 问 题	473
12.4.2 技术要点	473
12.4.3 步 骤	473
12.4.4 工作方式	474
12.4.5 注 释	475
附录 A CGI 支持的环境变量	476
附录 B HTML 表格元素	477