

<<正则表达式Cookbook>>

图书基本信息

书名：<<正则表达式Cookbook>>

13位ISBN编号：9787564119317

10位ISBN编号：7564119314

出版时间：2010-1

出版时间：东南大学出版社

作者：（美）高瓦特斯 著，（美）莱文森 著

页数：493

版权说明：本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问：<http://www.tushu007.com>

前言

Over the past decade, regular expressions have experienced a remarkable rise in popularity. Today, all the popular programming languages include a powerful regular expression library, or even have regular expression support built right into the language. Many developers have taken advantage of these regular expression features to provide the users of their applications the ability to search or filter through their data using a regular expression. Regular expressions are everywhere. Many books have been published to ride the wave of regular expression adoption. Most do a good job of explaining the regular expression syntax along with some examples and a reference. But there aren't any books that present solutions based on regular expressions to a wide range of real-world practical problems dealing with text on a computer and in a range of Internet applications. We, Steve and Jan, decided to fill that need with this book. We particularly wanted to show how you can use regular expressions in situations where people with limited regular expression experience would say it can't be done, or where software purists would say a regular expression isn't the right tool for the job. Because regular expressions are everywhere these days, they are often a readily available tool that can be used by end users, without the need to involve a team of programmers. Even programmers can often save time by using a few regular expressions for information retrieval and alteration tasks that would take hours or days to code in procedural code, or that would otherwise require a third-party library that needs prior review and management approval. Caught in the Snarls of Different Versions As with anything that becomes popular in the IT industry, regular expressions come in many different implementations, with varying degrees of compatibility. This has resulted in many different regular expression flavors that don't always act the same way, or work at all, on a particular regular expression.

<<正则表达式Cookbook>>

内容概要

本书提供了超过100条的锦囊妙计，帮助你利用正则表达式处理数据、操纵文本。每位程序员都能找到正则表达式的用武之地，但想要充分发挥它的威力却未必容易。纵使经验丰富的用户也常会遇到性能不佳、误判、漏判或者令人费解的错误。

《正则表达式Cookbook》对涉及此工具的最常见任务做了逐步讲解，此外还包括在C#、JaVa、JavaScript、Perl、PHP、Python、Ruby和VB.NET语言中使用正则表达式的诀窍。

阅读本书，你将：

- 通过简洁的教程理解正则表达式的基本原理
- 在多种编程和脚本语言中高效地应用正则表达式
- 学习如何验证和格式化输入
- 操纵单词、行、特殊字符和数值
- 找到在URL、路径、标记和数据交换中使用正则表达式的方法
- 学习更高级的正则表达式特性
- 理解在不同语言中正则表达式的应用程序接口、语法和行为的不同
- 针对特定需要，编写更加优化的正则表达式

无论你是初学者还是经验丰富的用户，《正则表达式Cookbook》都将有助于你对这一独特而不可替代的工具的理解。

你将学到功能强大的新技巧，避免和语言相关的陷阱，利用这一经过实践检验的方法解决现实世界中的难题，从而节省宝贵的时间。

<<正则表达式Cookbook>>

作者简介

Jan Goyvaerts , 经营Just Great Software 软件公司 , 在这家公司他负责设计和开发一些最流行的正则表达式软件。

Steven Levithan , 是一位JavaScript正则表达式权威专家 , 同时他还负责一个以正则表达式内容为中心的流行博客。

书籍目录

Preface.1. Introduction to Regular Expressions Regular Expressions Defined Searching and Replacing with Regular Expressions Tools for Working with Regular Expressions2. Basic Regular Expression Skills 2.1 Match Literal Text 2.2 Match Nonprintable Characters 2.3 Match One of Many Characters 2.4 Match Any Character 2.5 Match Something at the Start and/or the End of a Line 2.6 Match Whole Words 2.7 Unicode Code Points, Properties, Blocks, and Scripts 2.8 Match One of Several Alternatives 2.9 Group and Capture Parts of the Match 2.10 Match Previously Matched Text Again 2.11 Capture and Name Parts of the Match 2.12 Repeat Part of the Regex a Certain Number of Times 2.13 Choose Minimal or Maximal Repetition 2.14 Eliminate Needless Backtracking 2.15 Prevent Runaway Repetition 2.16 Test for a Match Without Adding It to the Overall Match 2.17 Match One of Two Alternatives Based on a Condition 2.18 Add Comments to a Regular Expression 2.19 Insert Literal Text into the Replacement Text 2.20 Insert the Regex Match into the Replacement Text 2.21 Insert Part of the Regex Match into the Replacement Text 2.22 Insert Match Context into the Replacement Text3. Programming with Regular Expressions Programming Languages and Regex Flavors 3.1 Literal Regular Expressions in Source Code 3.2 Import the Regular Expression Library 3.3 Creating Regular Expression Objects 3.4 Setting Regular Expression Options 3.5 Test Whether a Match Can Be Found Within a Subject String 3.6 Test Whether a Regex Matches the Subject String Entirely 3.7 Retrieve the Matched Text 3.8 Determine the Position and Length of the Match 3.9 Retrieve Part of the Matched Text 3.10 Retrieve a List of All Matches 3.11 Iterate over All Matches 3.12 Validate Matches in Procedural Code 3.13 Find a Match Within Another Match 3.14 Replace All Matches 3.15 Replace Matches Reusing Parts of the Match 3.16 Replace Matches with Replacements Generated in Code 3.17 Replace All Matches Within the Matches of Another Regex 3.18 Replace All Matches Between the Matches of Another Regex 3.19 Split a String 3.20 Split a String, Keeping the Regex Matches 3.21 Search Line by Line4. Validation and Formatting 4.1 Validate Email Addresses 4.2 Validate and Format North American Phone Numbers 4.3 Validate International Phone Numbers 4.4 Validate Traditional Date Formats 4.5 Accurately Validate Traditional Date Formats 4.6 Validate Traditional Time Formats 4.7 Validate ISO 8601 Dates and Times 4.8 Limit Input to Alphanumeric Characters 4.9 Limit the Length of Text 4.10 Limit the Number of Lines in Text 4.11 Validate Affirmative Responses 4.12 Validate Social Security Numbers 4.13 Validate ISBNs 4.14 Validate ZIP Codes 4.15 Validate Canadian Postal Codes 4.16 Validate U.K. Postcodes 4.17 Find Addresses with Post Office Boxes 4.18 Reformat Names From "FirstName LastName" to "LastName, FirstName" 4.19 Validate Credit Card Numbers 4.20 European VAT Numbers5. Words, Lines, and Special Characters 5.1 Find a Specific Word 5.2 Find Any of Multiple Words 5.3 Find Similar Words 5.4 Find All Except a Specific Word 5.5 Find Any Word Not Followed by a Specific Word 5.6 Find Any Word Not Preceded by a Specific Word 5.7 Find Words Near Each Other 5.8 Find Repeated Words 5.9 Remove Duplicate Lines 5.10 Match Complete Lines That Contain a Word 5.11 Match Complete Lines That Do Not Contain a Word 5.12 Trim Leading and Trailing Whitespace 5.13 Replace Repeated Whitespace with a Single Space 5.14 Escape Regular Expression Metacharacters6. Numbers 6.1 Integer Numbers 6.2 Hexadecimal Numbers 6.3 Binary Numbers 6.4 Strip Leading Zeros 6.5 Numbers Within a Certain Range 6.6 Hexadecimal Numbers Within a Certain Range 6.7 Floating Point Numbers 6.8 Numbers with Thousand Separators 6.9 Roman Numerals7. URLs, Paths, and Internet Addresses 7.1 Validating URLs 7.2 Finding URLs Within Full Text 7.3 Finding Quoted URLs in Full Text 7.4 Finding URLs with Parentheses in Full Text 7.5 Turn URLs into Links 7.6 Validating URNs 7.7 Validating Generic URLs 7.8 Extracting the Scheme from a URL 7.9 Extracting the User from a URL 7.10 Extracting the Host from a URL 7.11 Extracting the Port from a URL 7.12 Extracting the Path from a URL 7.13 Extracting the Query from a URL 7.14 Extracting the Fragment from a URL 7.15 Validating Domain Names 7.16 Matching IPv4 Addresses 7.17 Matching IPv6 Addresses 7.18 Validate Windows Paths 7.19 Split Windows Paths into Their Parts 7.20 Extract the Drive Letter from a Windows Path 7.21 Extract the Server and Share from a UNC Path 7.22 Extract the Folder from a Windows Path 7.23 Extract the Filename

<<正则表达式Cookbook>>

from a Windows Path 7.24 Extract the File Extension from a Windows Path 7.25 Strip Invalid Characters from Filenames8. Markup and Data Interchange 8.1 Find XML-Style Tags 8.2 Replace Tags with 8.3 Remove All XML-Style Tags Except and 8.4 Match XML Names 8.5 Convert Plain Text to HTML by Adding

章节摘录

插图：At the time of this writing, the Win32 version of Delphi does not have any built-in support for regular expressions. There are many VCL components available that provide regular expression support. I recommend that you choose one based on PCRE. Delphi has the ability to link C object files into your applications, and many VCL wrappers for PCRE use such object files. This allows you to keep your application as a single .exe file. You can download my own TPerlRegEx component at <http://www.regexp.info/delphi.html>. This is a VCL component that installs itself onto the component palette, so you can easily drop it into a form. Another popular PCRE wrapper for Delphi is the TJclRegEx class part of the 3CL library at <http://www.delphi-jedi.org>. TJclRegEx descends from TObject, so you can't drop it into a form. Both libraries are open source under the Mozilla Public License.

<<正则表达式Cookbook>>

媒体关注与评论

“这是一本精心打造的著作，内容翔实。

我仅仅读了介绍的章节就学会了一些新技巧。

”——Nikolaj Lindberg，计算机语言学家，STTS Speech Technology Services “《正则表达式Cookbook》对当前面临的问题做出了优雅的答案。

总之，我对这些技巧之详尽感到震惊。

”——Zak Greant，开放技术倡导者和战略家

<<正则表达式Cookbook>>

编辑推荐

《正则表达式Cookbook(影印版)》由东南大学出版社出版。

<<正则表达式Cookbook>>

版权说明

本站所提供下载的PDF图书仅提供预览和简介, 请支持正版图书。

更多资源请访问:<http://www.tushu007.com>