第一图书网, tushu007.com

<<正则指引>>

图书基本信息

书名：<<正则指引>>

13位ISBN编号：9787121165511

10位ISBN编号：7121165511

出版时间：2012-5

出版时间：电子工业出版社

作者：余晟

页数：336

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

第一图书网, tushu007.com

前言

前 言　　提到正则表达式，许多人很有点不屑一顾：这东西，不登大雅之堂，再说也不是总要用到，何必专门花时间学习？

没错，正则表达式并不“总要用到”，但到了需要的场合用不上，往往产生“一分钱难倒英雄汉”的尴尬。

经常需要处理文本的程序员自然会知道正则表达式的价值，其他的程序员如果不会正则表达式，即便开发的领域与文本处理没什么关系，也难免“躺着中枪”的命运--前几天我遇到一个问题，将一行长长的地址拆分成多行，负责这部分的程序员的日常工作只是制作PDF而已，拆分地址是很“边缘”的功能，但不会正则表达式就无法准确折行（一般需要在标点符号出现的地方折行，而不能只在空白字符处折行，但是不同语言中的标点符号各有不同），结果一筹莫展；相反，如果了解正则表达式，就可以很容易地处理各种语言中的标点字符。

以我的开发经验来看，专门花点时间掌握正则表达式，确实是非常有必要的。

目前可以见到的关于正则表达式的书籍和资料有不少，但又各有不足。

在互联网上，流传着一些编程语言的正则文档和《30分钟教会你正则表达式》之类的帖子。这类资料的好处是简单直接，查到了，如果有现成的例子，而且适用于自己的语言，则可以直接拿来用；然而，其坏处也是简单直接，因为缺乏背后原理的讲解，如果找不到现成的例子，或者找不到能在自己所使用语言中行得通的例子（需知道，同样的正则表达式并不能直接套用到不同的语言中），则束手无策。

在正式的出版领域，已经有《精通正则表达式》、《正则表达式必知必会》之类的书籍出版，尤其是前者，堪称关于正则表达式的经典著作，如果想认真学习正则表达式，这类书籍是必须阅读的。但这类书籍也有一个弱点，即都是由英文版本翻译而来的，更多地侧重英文文本的处理，身为中文世界的开发人员，我们经常需要处理中文文本--对于处理英文之外的字符，正则表达式已经提供了足够丰富的功能，但如何用对、用好这些功能，资料却很匮乏。

我经常需要给人讲解正则表达式的相关知识，时常惋惜的是，开发人员为这些问题所困扰；正因为如此，本书的写作动机便是着力弥补现有资料的缺陷。

相对于正则文档和速成教学帖子，本书深入讲解了匹配背后的原理，往往会举一反三，告诉读者，这里为何这样写，如果改成其他形式，会造成什么结构；并且，集中讲解和比较了多种语言中正则表达式用法的异同，方便读者把现成的正则表达式“移植”到自己的工作环境中。

相对于《精通正则表达式》等正式的书籍，本书辟出专门的内容讲解语言和编码，告诉读者如何设定编码，如何正确处理中文等字符。

另外，本书还涵盖了.NET、Java、JavaScript、PHP、Python、Ruby六种常用语言，对每种语言给出专门章节，不但详细介绍了语言中正则表达式的用法，更点明了版本之间的细微差异，不但可以作为专门学习的教材，还可以成为有用的参考手册。

本书结构　　本书可以分为三大部分。

第一部分主要讲解正则表达式的基础知识，覆盖常见正则表达式中的各种功能和结构。看完前面3章，就可以基本弄明白现在流行的各种正则表达式；尤其是如果你之前有一些经验，会觉得阅读起来并不困难。

但是我也希望读者不要忽略其他的内容，断言和匹配模式现在已经是正则表达式的“标准配备”了，而且确实可以派上大用场，所以第4章和第5章的内容，即便不是很熟悉，阅读起来可能有一些麻烦，也不应该忽略。

最后的第6章，则厘清了正则表达式在使用中的若干疑惑，了解它们，你就可以相对自由地在正则表达式的世界里行走了。

第二部分主要讲解关于正则表达式的更深入的知识，这一部分用3章的内容，详细探讨了编码问题、匹配原理、解题思路。

这部分内容更抽象，需要多花一点时间来阅读和理解，但是它们确实可以帮你在正则表达式的世界里登堂入室，脱离“术”的层面，掌握万变不离其宗的“道”。

第一图书网, tushu007.com

第三部分的作用是接地气，将之前介绍的各种知识落实到六种常用语言.NET、Java、JavaScript、PHP、Python、Ruby中来。
每一章的开头有正则功能列表，其中的功能都对应着前面部分的讲解，这些功能的具体应用实例，以及不同版本之间的差异，则在章节中详细讲解，每一章的最后还给出了常见任务的示例代码，方便日后查询。
在最后，第16章简要介绍了正则表达式在Linux下常用工具vi、grep、awk、sed中的使用，并通过一个实际的例子将这几种工具串起来，对比说明了它们适合解决的问题。
在本书的最后提供了用作参考的3个附录。
附录A是正则表达式的常用功能在不同语言中的比对，希望能给需要在多种语言中使用正则表达式或者移植正则表达式的读者提供一份有用的参考；附录B给出了若干常见的正则表达式，比如匹配邮政编码、身份证号、手机号、QQ号、电子邮件地址等，希望能成为常见问题的“速查手册”；附录C列出了常用正则表达式的工具和资源，方便大家调试自己的正则表达式，以及继续深入学习。
本书读者　　本书适合以下几类读者。
经常需要进行文本处理（比如日志分析或网络运维）的技术人员。
这些读者或许已经熟悉了正则表达式的基本用法，但面对日益复杂化和海量化的数据，阅读本书可以帮助你更准确、更高效地处理文本，提升自己工作的价值。
熟悉常用开发语言的程序员。
虽然这些读者不需要专职进行文本处理，但源代码和许多数据其实也是文本，如果不会正则表达式，在偶然遇到处理源代码或文本数据的任务时，往往会产生躺着中枪的无力感。
本书第三部分可以帮你迅速找到有关的例子，并落实在自己的编程语言中。
当然前两部分也非常有必要，因为它们可以帮你夯实基础。
对正则表达式已经有一定了解的读者。
这些读者虽然能用正则表达式解决常见的任务，但未必了解正则表达式的编码问题、匹配原理、解题思路，仔细阅读本书的第二部分，可以深化完善对正则表达式的理解；而第三部分详细比较了使用正则表达式时各种语言，以及同一种语言中各种版本的差异。
所有这一切，应该可以让你对正则表达式的掌握更上一层楼。
致谢　　一本书的完成，必然离不开众多人的帮忙。
首先需要感谢的是周筠老师和徐定翔、卢鹄翔两位编辑，他们在我写作的最初阶段做了大量细心、耐心的工作，完全可以说，没有他们的这些工作，我就不会有写作这本书的念头，或者坚持写完的动力。
然后要感谢的是电子工业出版社的杨福平副总编和张月萍编辑，没有他们的关照和辛劳工作，这本书的出版定然会遇到更多的困难。
感谢我的朋友霍炬和韩磊，虽然我之前阅读过《精通正则表达式》，但与翻译和写作结缘，他们给了我莫大的帮助，有了这个契机，才有了现在的《正则指引》。
尤其值得一提的是霍炬的夫人西乔，精心手绘了这本书的封面，在这里表示诚挚的谢意。
感谢我曾工作过的盛大创新院以及创新院的各位同事（李骏、郝培强、庄表伟、丁宇、许式伟、莫华枫、李道兵、赵劼、樊一鹏、张一宁等），创新院给了大家宽松自由的工作环境，与各位同事的讨论加深了我对正则表达式的理解，也为我提供了许多例子。
感谢张东亮、陆亦斌、孙勇、叶劲峰等各位朋友，愿意拨冗阅读本书的草稿，并提出了大量专业的意见。
感谢何源、陈钢、贺钧、陈驰等读者，试读本书之后提出了大量的宝贵意见，在最后关头打消了我心中的许多忐忑。
在更早之前，我的父母从小就鼓励我研究和了解各种科学原理（“玩也要动脑筋”），我之所以有兴趣探究正则表达式背后的世界，而不满足于“够用/凑合”，归源都是受益于这种思维行为习惯。
此外，在中小学阶段，我的语文老师罗碧玉、郭志鸿、易玺铭培养了我对于文字的兴趣，在大学阶段

第一图书网, tushu007.com

，东北师范大学文学院的王确老师给了我这个理科生非常多的帮助和指引。
对各位师长，在此一并表示感谢，能遇到你们是我的幸运。
　　最后还需要感谢许多为这本书做出过贡献的人，你们的名字我可能暂时无法记起，或者无法一一罗列，但我会在心中存留对你们的谢意。

第一图书网, tushu007.com

<<正则指引>>

内容概要

《正则指引》针对作者在开发中遇到的实际问题，以及其他开发人员咨询的问题，总结出一套使用正则表达式解题的办法，并通过具体的例子指导读者拆解、分析问题。
全书分为三大部分：第一部分主要讲解正则表达式的基础知识，涵盖了常见正则表达式中的各种功能和结构；第二部分主要讲解关于正则表达式的更深入的知识，详细探讨了编码问题、匹配原理、解题思路；第三部分将之前介绍的各种知识落实到6种常用语言.NET、Java、JavaScript、PHP、Python、Ruby中，不但详细介绍了语言中正则表达式的用法，更点明了版本之间的细微差异，既可以作为专门学习的教材，也可以作为有用的参考手册。

本书适合经常需要进行文本处理（比如日志分析或网络运维）的技术人员、熟悉常用开发语言的程序员，以及已经对正则表达式有一定了解的读者阅读。

第一图书网, tushu007.com

作者简介

余晟，毕业于计算机系，副修中文，非正统型技术爱好者。
曾任抓虾网、银杏泰克主力程序员，盛大创新院高级研究员，现任华南某电商公司技术部总监。
坚信计算机可以无限延伸人的能力，前提是人必须理解计算机的逻辑，所以对任何技术都不应该浅尝辄止，仅仅满足于“ 会用 ”。

已经翻译出版《精通正则表达式（第3版）》和《技术领导之路》，审阅《软件架构师应该知道的97件事》和《REST in
Practice中文版》。

第一图书网, tushu007.com

<<正则指引>>

书籍目录

第一图书网, tushu007.com

第一图书网, tushu007.com

第一图书网, tushu007.com

第一图书网, tushu007.com

第一图书网, tushu007.com

第一图书网, tushu007.com

第一图书网, tushu007.com

章节摘录

第一图书网, tushu007.com

<<正则指引>>

编辑推荐

《正则指引》适合经常需要进行文本处理（比如日志分析或网络运维）的技术人员、熟悉常用开发语言的程序员，以及已经对正则表达式有一定了解的读者阅读。

第一图书网, tushu007.com

名人推荐

掌握正则表达式应该是IT工程师的一项标准技能，遗憾的是，过去，不少人多多少少忽视了这一点，所以在工作中总要应对正则表达式带来的【麻烦】。
我相信只有掌握并熟练运用它才有可能成为一个高效率的工程师。
期待每个人手边都有一本正则表达式的参考书，当然，最好就是你现在看到的这本。
就在写这句话的几分钟前，我又从这本书中学到了一个有用的技巧。
——冯大辉 配合恰当的案例，大量的反问，使读者自问、思考，扣人心扉，比较有代入感。
加上配图，很容易让读者全面认识正则表达式。
在原理讲解的章节，对比两种理论模型的区别，顺其自然地引入NFA引擎的关键要素:回溯。
使读者从匹配原理上了解回溯，写出高效、严谨的正则表达式。
——陈驰 本书由浅入深地讲述了正则表达式，在正则的应用和调优方面有非常详细的介绍，特别在正则表达式处理中文方面有独到的阐述。
对于需要经常处理中文的国内技术人员来说，无疑是非常值得拥有的一本手册。
——贺钧 正则表达式是程序员的必备知识。
如果您还没有使用过这个强大的工具，或者学习正则表达式总不得要领。
确实可以读读《正则指引》。
——何源 这是一本通俗版的“精通正则表达式”。
高手很难挑出毛病，一般程序员会受益匪浅，普通用户一步步读下去也能登堂入室。
——张东亮 余晟在之前翻译业内名著的基础上，结合中文环境和自己的丰富经验，再接再厉推出自己的原创著作，实在是我等码农的一大幸事。
——陈钢

第一图书网, tushu007.com

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