

第二章：字符串和文本

几乎所有有用的程序都会涉及到某些文本处理，不管是解析数据还是产生输出。这一章将重点关注文本的操作处理，比如提取字符串，搜索，替换以及解析等。大部分的问题都能简单的调用字符串的内建方法完成。但是，一些更为复杂的操作可能需要正则表达式或者强大的解析器，所有这些主题我们都会详细讲解。并且在操作Unicode时候碰到的一些棘手的问题在这里也会被提及到。

Contents:

- 2.1 使用多个界定符分割字符串
- 2.2 字符串开头或结尾匹配
- 2.3 用Shell通配符匹配字符串
- 2.4 字符串匹配和搜索
- 2.5 字符串搜索和替换
- 2.6 字符串忽略大小写的搜索替换
- 2.7 最短匹配模式
- 2.8 多行匹配模式
- 2.9 将Unicode文本标准化
- 2.10 在正则式中使用Unicode
- 2.11 删除字符串中不需要的字符
- 2.12 审查清理文本字符串
- 2.13 字符串对齐
- 2.14 合并拼接字符串
- 2.15 字符串中插入变量
- 2.16 以指定列宽格式化字符串
- 2.17 在字符串中处理html和xml
- 2.18 字符串令牌解析
- 2.19 实现一个简单的递归下降分析器
- 2.20 字节字符串上的字符串操作