第10章 模式匹配

10.1 模式匹配的相关函数

字符串标准库提供了基于模式的4个函数，包括find,gsub,match和gmatch。

10.1.1 函数string.find

函数string.find找到一个模式后，会返回两个值：匹配到模式开始位置和结束位置的索引。如果没有任何匹配，返回nil。

string.find有两个可选参数，第3个参数是一个索引，用于说明从哪个位置开始索引，第4个参数是一个布尔值，用于说明管是否进行简单搜索。

10.1.2 函数string.match

string.match和string.find非常相似。不过string.match返回的是目标字符串与模式相匹配的那部分子串。

string.match(date,"%d+/%d+/%d+")

10.1.3 函数string.gsub

string.gsub有3个必选参数：目标字符串，模式和替换字符串。

该函数还有一个可选的第4个参数，用于限制替换的次数。

10.1.4 函数string.gmatch

string.gmatch返回一个函数，通过返回的函数可以遍历一个字符串中所有出现的指定模式。

编程：找出指定字符串s中出现的所有单词

模式'%a+'会匹配一个或多个字母组成的序列，也就是单词。

10.2 模式

Lua语言中的模式使用百分号作为转义符。

预置的字符分类及其对应的含义：

任意字符：. 字母:%a 控制字符：%c 数字：%d 除空格外的可打印字符：%g

小写字母：%l 标点符号:%p 空白字符:%s 大写字母：%u 字母和数字：%w 十六进制数字：%x

这些类的大写形式表示类的补集，例如，'%A'代表任意非字母的字符。

当在模式中使用时，还有一些被称为魔法字符的字符具有特殊含义。

魔法字符包括：() . % + - \* ? [ ] ^ $

还可以使用字符集来创建自定义的字符分类，只需要再方括号内将单个字符和字符分类组合起来即可。

[%w\_]匹配所有以下划线结尾的字母和数字 [01]匹配二进制数字

还可以在字符集中包含一段字符范围，例如[0-9]

在字符集前加一个补字符^就可以得到这个字符集对应的补集。

还可以通过描述模式中重复和可选部分的修饰符来让模式更加有用。Lua语言提供了4种修饰符：

重复一次或多次:+ 重复零次或多次：\* 重复零次或多次（最小匹配）：- 可选（出现零次或一次）：?

以补字符^开头的模式表示从目标字符串的开头开始匹配。

以$结尾的模式表示匹配到目标字符串的结尾。

模式'%b'匹配成对的字符串。

10.3 捕获