

# 正则表达式

正则表达式是一个特殊的字符序列，计算机科学的一个概念。通常被用来检索、替换那些符合某个模式(规则)的文本。

许多程序设计语言都支持利用正则表达式进行字符串操作。在Python中需要通过正则表达式对字符串进行匹配的时候，可以使用`re`模块。`re` 模块使 Python 语言拥有全部的正则表达式功能。

## 特点：

1. 灵活性、逻辑性和功能性非常强；
2. 可以迅速地用极简单的方式达到字符串的复杂控制。
3. 对于刚接触的人来说，比较晦涩难懂。

## Python中的正则表达式

与大多数编程语言相同，正则表达式里也使用 `\` 作为转义字符，这就可能造成反斜杠困扰。假如你需要匹配文本中的字符 `\`，那么使用编程语言表示的正则表达式里将需要4个反斜杠 `\\`：前两个和后两个分别用于在编程语言里转义成反斜杠，转换成两个反斜杠后再在正则表达式里转义成一个反斜杠。

```
print(re.match('\\\\\\', '\\hello')) # 需要使用四个反斜杠来匹配一个 \
```

Python里的原生字符串很好地解决了这个问题，有了原生字符串，你再也不用担心是不是漏写了反斜杠，写出来的表达式也更直观。在Python 字符串前面添加 `r` 即可将字符串转换成为原生字符串。

```
print(re.match(r'\\', '\\hello')) # 使用两个反斜杠即可匹配一个 \
```