**正则表达式**

正则表达式是一种特殊的文本模式，包括普通字符(字母数字下划线等)和元字符。正则表达式使用单个字符串来描述、匹配一系列匹配某个句法规则的字符串，通常被用来检索、替换那些符合某个模式的文本

**头和尾**

^The表示以"The"开始的字符串

it$表示以it为结尾的字符串

notice表示包含notice的字符串

**[]**

表示一个字符,该字符可以是[]中指定的内容

[abc]：这个字符可以是a或b或c  
[a-z]：表示任意一个小写字母  
[a-zA-Z]：表示任意一个字母  
[a-zA-Z0-9\_]：表示任意一个数字字母下划线  
[^abc]：该字符只要不是a或b或c

**预定义字符**

. :表示任意一个字符,没有范围限制  
\d :表示任意一个数字,等同于[0-9]  
\w :表示任意一个单词字符,等同于[a-zA-Z0-9\_]–单词指字母/数字/\_  
\s :表示任意一个空白字符  
\D :表示不是数字  
\W :不是单词字符  
\S:不是空白字符

**量词**

|  |
| --- |
| ? :表示前面的内容出现0-1次 例如: [abc]? 可以匹配:a 或 b 或 c 或什么也不写 |
| + :表示前面的内容最少出现1次 例如: [abc]+ 可以匹配:b或aaaaaaaaaa…或abcabcbabcbabcba… 但是不能匹配:什么都不写 或 abcfdfsbbaqbb34bbwer…(整个串都必须完全匹配，而不是其中一部分) |
| \*: 表示前面的内容出现任意次(0-多次)—匹配内容与+一致，只是可以一次都不写 例如: [abc]\* 可以匹配:b或aaaaaaaaaa…或abcabcba…或什么都不写，但是不能匹配:abcfdfsbbaqbb34bbwer… |
| {n} :表示前面的内容出现n次 例如: [abc]{3} 可以匹配:aaa 或 bbb 或 aab 或abc 或bbc 但是不能匹配: aaaa 或 aad |
| {n,m} :表示前面的内容出现最少n次最多m次 例如: [abc]{3,5} 可以匹配:aaa 或 abcab 或者 abcc 但是不能匹配:aaaaaa 或 aabbd |
| {n,} :表示前面的内容出现n次以上(含n次) 例如: [abc]{3,} 可以匹配:aaa 或 aaaaa… 或 abcbabbcbabcba… 但是不能匹配:aa 或 abbdaw… |

**()**

()表示组，即将括号内的内容看做是一个整体

例如：

(abc){3}表示abc整体出现3次. 可以匹配abcabcabc

但是不能匹配aaa 或abcabc  
 (abc|def){3}表示abc或def整体出现3次

可以匹配: abcabcabc 或 defdefdef 或 abcdefabc

但是不能匹配abcdef 或abcdfbdef