

匹配字符

[abc]	匹配a,b,c中的任意一个字符
[^abc]	不匹配a,b,c中的所有字符,“^”只有作为第一个字符出现才有效
[a-g]	匹配a-g范围内的任意一个字符
[^a-g]	不匹配a-g中的所有字符
[H-N]	匹配H-N范围内的任意一个字符
[0-9]	匹配0-9范围内的任意一个字符
[a-gH-N]	匹配a-g,H-N范围内的任意一个字符
通配符:	
.	[^\n\r],匹配非换行符
\s	space,空白字符,即换行,制表,空格等字符
\S	[^\s]
\d	digit: [0-9]
\D	[^\d]
\w	word: [0-9A-Za-z_]
\W	[^\w]

匹配组

(a b)	匹配a或b并捕获
(...)	捕获组
(?:...)	非捕获组
(?<name>...)或(?:'name'...)	命名捕获组,name可以自定义
(?(condition>true_regex false_regex)	条件匹配

group(0)用于获取整个匹配串的匹配结果,group(>0)则可以从整个匹配串的结果中获取子捕获组。

注意:java1.7之后才支持命名捕获组(group方法传入字符串name作为参数),js和python暂时还不支持命名捕获组。

频次范围

{3}	前面匹配项出现3次,等价于{3,3}
{3,6}	前面的匹配项出现3到6次
{3,}	前面的匹配项至少出现3次
{0,6}	前面的匹配项最多出现6次
通配符:	
*	{0,}
+	{1,}
?	{0,1}
\w*	贪婪模式
\w*?	非贪婪模式

定点符

^	字符串开始符
\$	字符串结尾符
\b	boundary,单词边界
\B	非单词边界

定点断言

(?=exp)	向后肯定断言(exp后缀)
(?<=exp)	向前肯定断言(exp前缀)
(?!exp)	向后否定断言(没有exp后缀)
(?<!exp)	向前否定断言(没有exp前缀)

\b(w+)\b等价于(?<=W?)(w+)(?<=W?)

