

正则表达式字符表

- 元字符 #反斜杠后的元字符变为普通字符
 - . 匹配任意 (除"\n") a.c abc
 - \ 转义 元-》普通字符 a|.c;a||c a.c;a|c
 - * 匹配字符0或多次 abc* ab;abccc
 - + 匹配字符1次或无限次 abc+ abc;abccc
 - ? 匹配个字符0次或1次 abc? ab;abc
 - ^ 匹配字符串开头 ^abc abc
 - \$ 匹配字符串末尾 abc\$ abc
 - | 匹配|左右表达式任意一个, 从左到右, 没有(), 范围整个正则表达式 abc|def
abc def
 - {} {m}匹配前一个字符m次, {m,n}匹配前一个字符m至n次, 若省略n, 则匹配m至无限次 ab{1,2}c abc abbc
 - [] 字符集。可以是任意字符。可以逐个列出, 也可给出范围, 如[abc]或[a-c]。
[^abc]表示取反, 即非abc。特殊字符在字符集中失去特殊含义。 \反斜杠恢复特殊含义。 a[bcd]e abe ace ade
 - () 分组, 有编号 (abc){2}; a(123|456)c abcab; a456c
- 预定义字符集 #斜杠后字母变为特殊字符
 - \d 数字:[0-9] a\bc a1c
 - \D 非数字:[^\d] a\Dc abc
 - \s 匹配任何空白字符:[<空格>\t\r\n\f\v] a\s c a c
 - \S 非空白字符:[^\s] a\S abc
 - \w 匹配包括下划线在内的任何字符:[A-Za-z0-9_] a\wc abc
 - \W 匹配非字母字符, 即匹配特殊字符 a\Wc a c
 - \A 仅匹配字符串开头,同^ \Aabc abc
 - \Z 仅匹配字符串结尾, 同\$ abc\Z abc
 - \b 匹配\w和\W之间, 匹配一个单词边界, \babc\ba\b!bc 空格abc空格a!bc
 - \B [^\b] a\Bbc abc
- 特殊分组用法
 - (?P<name>) 分组, 除了编号外再指定额外的别名 (?P<id>abc){2} abcab
 - (?P=name) 引用别名为<name>的分组 (?P<id>\d)abc(?P=id) 1abc15abc5
 - \<number> 引用编号为<number>的分组 (\d)abc\1 1abc15abc5

