

Java程序员：一刻钟精通正则表达式 PDF转换可能丢失图片或格式，建议阅读原文

https://www.100test.com/kao_ti2020/290/2021_2022_Java_E7_A8_8B_E5_BA_8F_c104_290407.htm 想必很多人都对正则表达式都头疼。今天，我以我的认识，加上网上一些文章，希望用常人都可以理解的表达方式来和大家分享学习经验。开篇，还是得说说 ^ 和 \$ 他们是分别用来匹配字符串的开始和结束，以下分别举例说明： "^The"：开头一定要有"The"字符串； "of despair\$"：结尾一定要有"of despair" 的字符串；那么， "^abc\$"：就是要求以abc开头和以abc结尾的字符串，实际上是只有abc匹配。 "notice"：匹配包含notice的字符串。你可以看见如果你没有用我们提到的两个字符（最后一个例子），就是说模式（正则表达式）可以出现在被检验字符串的任何地方，你没有把他锁定到两边。接着，说说 *， ， 和 ?， 他们用来表示一个字符可以出现的次数或者顺序。他们分别表示： "zero or more"相当于{0, }， "one or more"相当于{1, }， "zero or one."相当于{0, 1}， 这里是一些例子： "ab*"：和ab{0, }同义，匹配以a开头，后面可以接0个或者N个b组成的字符串（"a", "ab", "abbb", 等）； "ab"：和ab{1, }同义，同上条一样，但最少要有一个b存在（"ab", "abbb", 等。）； "ab?"：和ab{0, 1}同义，可以没有或者只有一个b； "a? b \$"：匹配以一个或者0个a再加上一个以上的b结尾的字符串。要点， *， ， 和 ? 只管它前面那个字符。你也可以在大括号里面限制字符出现的个数，比如 "ab{2}"：要求a后面一定要跟两个b（一个也不能少）（"abb"）； "ab{2, }"：要求a后面一定要有两个或者两个以上b（如"abb", "abbbb", 等。）；

"ab{3, 5}" : 要求a后面可以有2-5个b ("abbb", "abbbb", or "abbbbb")。现在我们把一定几个字符放到小括号里, 比如:
"a(bc)*" : 匹配a后面跟0个或者一个"bc"; "a(bc){1, 5}" : 一个到5个"bc." 还有一个字符 , 相当于OR操作:
"hi|hello" : 匹配含有"hi" 或者 "hello" 的字符串; "(b|cd)ef" : 匹配含有 "bef" 或者 "cdef"的字符串; "(a|b)*c" : 匹配含有这样多个 (包括0个) a或b, 后面跟一个c的字符串;
一个点(.)可以代表所有的单一字符, 不包括"\n" 如果, 要匹配包括"\n"在内的所有单个字符, 怎么办? 对了, 用[\n.] 这种模式。
"a.[0-9]" : 一个a加一个字符再加一个0到9的数字
"^. {3}\$" : 三个任意字符结尾. 中括号括住的内容只匹配一个单一的字符 "[ab]" : 匹配单个的 a 或者 b (和 "a|b" 一样)
; 100Test 下载频道开通, 各类考试题目直接下载。详细请访问 www.100test.com