常用正则表达式符号和特殊字符

|  |  |  |
| --- | --- | --- |
| 表示法 | 描述 | 正则表达式示例 |
| 符号 |  |  |
| literal | 匹配文本字符串的字面值 literal | Foo |
| re1|re2 | 匹配正则表达式 re1 或者 re2 | foo|bar |
| . | 匹配任何字符（除了\n 之外） | b.b |
| ^ | 匹配字符串起始部分 | ^Dear |
| $ | 匹配字符串终止部分 | /bin/\*sh$ |
| \* | 匹配 0 次或者多次前面出现的正则表达式 | [A-Za-z0-9]\* |
| + | 匹配 1 次或者多次前面出现的正则表达式 | [a-z]+\.com |
| ? | 匹配 0 次或者 1 次前面出现的正则表达式 | goo? |
| {N} | 匹配 N 次前面出现的正则表达式 | [0-9]{3} |
| {M,N} | 匹配M~N次前面出现的正则表达式 | [0-9]{5,9} |
| […] | 匹配来自字符集中的任意单一字符 | [aeiou] |
| [..x−y..] | 匹配 x ～ y 范围中的任意单一字符 | [0-9], [A-Za-z] |
| [^…] | 不匹配此字符集中出现的任何一个字符，包括某一范围的字符（如果在此字符集中出现） | [^aeiou], [^A-Za-z0-9] |
| (\*|+|?|{})? | 用于匹配上面频繁出现/重复出现符号的非贪婪版本（\*、+、?、{}） | .\*?[a-z] |
| (…) | 匹配封闭的正则表达式，然后另存为子组 | ([0-9]{3})?,f(oo|u)bar |
| 特殊字符 |  |  |
| \d | 匹配任何十进制数字，与[0-9]一致（\D 与\d 相反，不匹配任何非数值型的数字） | data\d+.txt |
| \w | 匹配包括下划线的任何单词字符，与[A-Za-z0-9\_]相同（\W 与之相反） | [A-Za-z\_]\w+ |
| \s | 匹配任何空白字符，与[\n\t\r\v\f]相同（\S 与之相反） | of\sthe |
| \b | 匹配任何单词边界,也就是指单词和空格间的位置（\B 与之相反） | \bThe\b |
| \t | 匹配一个制表符 |  |
| \n | 匹配一个换行符 |  |

Python

术语“匹配”（matching）指的是模式匹配，主要用两种方法完成模式匹配

“搜索”即在字符串任意部分中搜索匹配的模式；

“匹配”是指判断一个字符串中能否从起始处全部或部分地匹配某个模式；