

拓展：使用正则表达式查询

正则表达式通常被用来检索或替换那些符合某个模式的文本内容，根据指定的匹配模式匹配文本中符合要求的特殊字符串。例如，从一个文本文件中提取电话号码，查找一篇文章中重复的单词或者替换用户输入的某些敏感词语等，这些地方都可以使用正则表达式。正则表达式强大而且灵活，可以应用于非常复杂的查询。

MySQL中使用REGEXP关键字指定正则表达式的字符匹配模式。下表列出了REGEXP操作符中常用字符匹配列表。

选项	说明	例子	匹配值示例
^	匹配文本的开始字符	'^b'匹配以字母 b 开头的字符串	book, big, banana, bike
\$	匹配文本的结束字符	'st\$'匹配以 st 结尾的字符串	test, resist, persist
.	匹配任何单个字符	'b.t'匹配任何 b 和 t 之间有一个字符的字符串	bit, bat, but, bite
*	匹配零个或多个在它前面的字符	'f*n'匹配字符 n 前面有任意个字符 f 的字符串	fn, fan, faan, fabcn
+	匹配前面的字符 1 次或多次	'ba+'匹配以 b 开头后面紧跟至少有一个 a 的字符串	ba, bay, bare, battle
<字符串>	匹配包含指定的字符串的文本	'fa'匹配包含 fa 的字符串	fan, afa, faad
[字符集合]	匹配字符集合中的任何一个字符	'[xz]' 匹配包含 x 或者 z 的字符串	dizzy, zebra, x-ray, extra
[^]	匹配不在括号中的任何字符	'[^abc]'匹配任何不包含 a、b 或 c 的字符串	desk, fox, f8ke
字符串 {n,}	匹配前面的字符串至少 n 次	b{2,}匹配 2 个或更多的 b	bbb, bbbb, bbbbbbb
字符串 {n,m}	匹配前面的字符串至少 n 次,至多 m 次。如果 n 为 0, 此参数为可选参数	b{2,4}匹配含最少 2 个、最多 4 个 b 的字符串	bb, bbb, bbbb

1. 查询以特定字符或字符串开头的记录 字符'^'匹配以特定字符或者字符串开头的文本。

在fruits表中，查询f_name字段以字母'b'开头的记录，SQL语句如下：

```
mysql> SELECT * FROM fruits WHERE f_name REGEXP '^b';
```

2. 查询以特定字符或字符串结尾的记录 字符'\$'匹配以特定字符或者字符串结尾的文本。

在fruits表中，查询f_name字段以字母'y'结尾的记录，SQL语句如下：

```
mysql> SELECT * FROM fruits WHERE f_name REGEXP 'y$';
```

3. 用符号"."来替代字符串中的任意一个字符 字符'.'匹配任意一个字符。在fruits表中，查询f_name字段值包含字母'a'与'g'且两个字母之间只有一个字母的记录，SQL语句如下：

```
mysql> SELECT * FROM fruits WHERE f_name REGEXP 'a.g';
```

4. 使用"*"和"+"来匹配多个字符 星号'*'匹配前面的字符任意多次，包括0次。加号'+'匹配前面的字符至少一次。

在fruits表中，查询f_name字段值以字母'b'开头且'b'后面出现字母'a'的记录，SQL语句如下：

```
mysql> SELECT * FROM fruits WHERE f_name REGEXP '^ba*';
```

在fruits表中，查询f_name字段值以字母‘b’开头且‘b’后面出现字母‘a’至少一次的记录，SQL语句如下：

```
mysql> SELECT * FROM fruits WHERE f_name REGEXP '^ba+';
```

5. 匹配指定字符串 正则表达式可以匹配指定字符串，只要这个字符串在查询文本中即可，如要匹配多个字符串，多个字符串之间使用分隔符‘|’隔开。

在fruits表中，查询f_name字段值包含字符串“on”的记录，SQL语句如下：

```
mysql> SELECT * FROM fruits WHERE f_name REGEXP 'on';
```

在fruits表中，查询f_name字段值包含字符串“on”或者“ap”的记录，SQL语句如下：

```
mysql> SELECT * FROM fruits WHERE f_name REGEXP 'on|ap';
```

之前介绍过，LIKE运算符也可以匹配指定的字符串，但与REGEXP不同，LIKE匹配的字符串如果在文本中间出现，则找不到它，相应的行也不会返回。REGEXP在文本内进行匹配，如果被匹配的字符串在文本中出现，REGEXP将会找到它，相应的行也会被返回。对比结果如下所示。

在fruits表中，使用LIKE运算符查询f_name字段值为“on”的记录，SQL语句如下：

```
mysql> SELECT * FROM fruits WHERE f_name like 'on';  
Empty set (0.00 sec)
```

6. 匹配指定字符中的任意一个 方括号“[]”指定一个字符集合，只匹配其中任何一个字符，即为所查找的文本。

在fruits表中，查找f_name字段中包含字母‘o’或者‘t’的记录，SQL语句如下：

```
mysql> SELECT * FROM fruits WHERE f_name REGEXP '[ot]';
```

在fruits表中，查询s_id字段中包含4、5或者6的记录，SQL语句如下：

```
mysql> SELECT * FROM fruits WHERE s_id REGEXP '[456]';
```

7. 匹配指定字符以外的字符 “[^字符集合]” 匹配不在指定集合中的任何字符。

在fruits表中，查询f_id字段中包含字母a~e和数字1~2以外字符的记录，SQL语句如下：

```
mysql> SELECT * FROM fruits WHERE f_id REGEXP '[^a-e1-2]';
```

8. 使用{n,}或者{n,m}来指定字符串连续出现的次数 “字符串{n,}”表示至少匹配n次前面的字符；“字符串{n,m}”表示匹配前面的字符串不少于n次，不多于m次。例如，a{2,}表示字母a连续出现至少2次，也可以大于2次；a{2,4}表示字母a连续出现最少2次，最多不能超过4次。

在fruits表中，查询f_name字段值出现字母‘x’至少2次的记录，SQL语句如下：

```
mysql> SELECT * FROM fruits WHERE f_name REGEXP 'x{2,}';
```

在fruits表中，查询f_name字段值出现字符串“ba”最少1次、最多3次的记录，SQL语句如下：

```
mysql> SELECT * FROM fruits WHERE f_name REGEXP 'ba{1,3}';
```