正则表达式

-Author: bavdu

-Email: bavduer@163.com

-Gtihub: https://github.com/bavdu

- 正则表达式基本元字符
- 正则表达式拓展元字符

正则表达式（regular expression, RE）是一种字符模式, 用于在查找过程中匹配指定的字符. 在大多数程序里, 正则表达式都被置于两个正斜杠之间

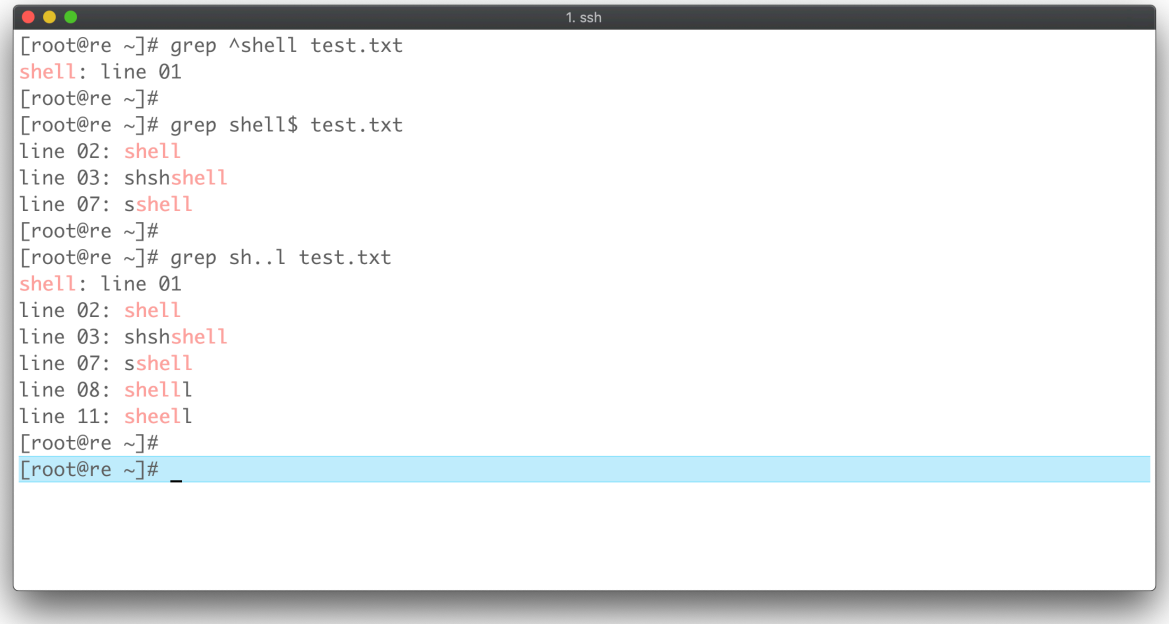
```
grep password /var/log/mysqld.log -精确查找
grep passw* /var/log.mysqld.log   -模糊查找

love lOve l ve lve
```

例如`/l[oO ]ve/`就是由正斜杠界定的正则表达式, 它将匹配被查找的行中任何位置出现的相同模式. 在正则表达式中,元字符是最重要的概念

No.1 正则表达式基本元字符

元字符	功能描述	示例
`^`	行首定位符	`^shell`
`$`	行尾定位符	`shell$`
`.`	匹配单个字符	`sh..l`
`*`	匹配前导符0次到多次	`sh*ell`
`.*`	任意多个字符	`None`
`[]`	匹配指定字符中的一个字符	`[Ss]hell`
`[ - ]`	匹配指定范围内的一个字符	`[a-z]hell`
`[^ ]`	匹配不在指定组内的一个字符	`[^a-z]hell`
`\`	转义字符	`\$SHELL`
`\(..\)`	匹配后面使用的字符标签	`:% s/\(172.\)\(16.\)\(130.\)1/\1\2\35/`




1. ssh
[root@re ~]# grep sh*ell test.txt
shell: line 01
line 02: shell
line 03: shshshell
line 04: shhhhhell
line 07: sshell
line 08: shelll
line 15: sell
line 17: sellhhh
[root@re ~]#


1. ssh
[root@re ~]# grep s.* test.txt
shell: line 01
line 02: shell
line 03: shshshell
line 04: shhhhhell
line 07: sshell
line 08: shelll
line 09: sshheell
line 11: sheell
line 12: sheeell
line 13: sheeeell
line 14: sheeeeell
line 15: sell
line 16: shelelelellll
line 17: selllhhh
line 18: shl
[root@re ~]#


1. ssh
[root@re ~]# grep [she]l test.txt
shell: line 01
line 02: shell
line 03: shshshell
line 04: shhhhhell
line 05: 3hell
line 07: sshell
line 08: shelll
line 09: sshheell
line 11: sheell
line 12: sheeell
line 13: sheeeell
line 14: sheeeeell
line 15: sell
line 16: shelelelelllll
line 17: sellllhhh
line 18: shl
[root@re ~]#


1. ssh
[root@re ~]# grep s[^e]l test.txt
line 18: shl
[root@re ~]#
[root@re ~]# grep [a-z]he test.txt
shell: line 01
line 02: shell
line 03: shshshell
line 04: shhhhhell
line 07: sshell
line 08: shelll
line 09: sshheell
line 11: sheell
line 12: sheeell
line 13: sheeeell
line 14: sheeeeell
line 16: shelelelelllll
[root@re ~]#

No.2 正则表达式拓展元字符

元字符	功能描述	示例
`+`	匹配1个或多个前导字符	`[a-z]+hell`
`?`	匹配0个或 1 个前导字符	`sh?ll`
`var01|var02`	匹配var01或var02	`shell|python`
`()`	匹配括号中的词组	`(bavduer|python)`
`x\{m\}`	规定字符x重复出现m次	`she\{2\}ll`
`x\{m,\}`	规定字符x重复出现至少m次	`she\{2,\}ll`
`x\{m,n\}`	规定字符x重复出现m到n次	`she\{2,5\}ll`

练习题:

- 匹配文件中所有的固定电话号码、所有的手机号码、所有的固话和手机号码

```
^1[3578][0-9]\{9\}$                          --手机号
^0[1-9][0-9]\{1,2\}-[1-9][0-9]\{6,7\}        --固话

(^1[3578][0-9]\{9\}$)|(^0[1-9][0-9]\{1,2\}-[1-9][0-9]\{6,7\})

[root@re ~]# egrep "(^1[3578][0-9]{9}$|^0[1-9][0-9]{1,2}-[1-9][0-9]{6,7})"
regex.txt
010-88889999
0352-68861234
13811012345
13187651101
15897550681
17789564321
18911012345
```

- 匹配文件中所有的北京市身份证号

```
^110[0-9]\{14\}x|X
^110[0-9]\{15\}

[root@re ~]# egrep "(^110[0-9]{15}|^110[0-9]{14}[xX])" regex.txt
11011118897987123x
110111188909071512
110111188789071258
```

- 匹配文件中所有的IP地址

```
256.789.567.134: 点分十进制中所有的十进制数,都不能大于255
0.1.1.1:

[1-255]: ([1-9]|[1-9][0-9]|1[0-9][0-9]|2[0-4][0-9]|25[0-5])
[0-255]: ([0-9]|[1-9][0-9]|1[0-9][0-9]|2[0-4][0-9]|25[0-5])
[1-255]: ([1-9]|[1-9][0-9]|1[0-9][0-9]|2[0-4][0-9]|25[0-4])

[root@re ~]# egrep "^([1-9]|[1-9][0-9]|1[0-9][0-9]|2[0-4][0-9]|25[0-5])\.
(([0-9]|[1-9][0-9]|1[0-9][0-9]|2[0-4][0-9]|25[0-5])\.){2}([1-9]|[1-9][0-
9]|1[0-9][0-9]|2[0-4][0-9]|25[0-4])$" regex.txt
```

- 匹配文件中所有的日期

```
1900-2099: (19|20)[0-9]{2}
1-12: 01~09/1~12: 10/11/12: (0?[1-9]|1[0-2])
1~31: 01~09/1-9: 10~30: 31: (0?[1-9]|[12][0-9]|30|31)

egrep "(19|20)[0-9]{2}-(0?[1-9]|1[0-2])-(0?[1-9]|[12][0-9]|30|31)"
regex.txt
```

- 查找文件中所有的邮箱

附录: 基础知识练习文件

```
[root@re ~]# cat test.txt
shell: line 01
line 02: shell
line 03: shshshell
line 04: shhhhhell
line 05: 3hell
line 06: 192.168.161.3
1.1.1.100
line 07: sshell
line 08: shelll
line 09: sshheell
line 10: printf and println
line 11: sheell
line 12: sheeell
line 13: sheeeell
line 14: sheeeeell
line 15: sell
line 16: shelelelellll
line 17: selllhhh
line 18: shl
```

```
[root@re ~]# cat regex.txt
[phone]
010-88889999
0352-68861234
123-12345678
13811012345
13187651101
15897550681
17789564321
18911012345
12345678911
28911012346

[userid]
120111188565671231
130111188343474561
150111188212171809
170111188909071231
11011118897987123x
110111188909071512
150111188200071430
```

```
130111188211171734
110111188789071258

[ipaddress]
1.1.1.1
12.12.12.12
123.123.123.123
255.256.225.101
127.0.0.1
0.255.246.322

[date]
1988-12-12
1977-10-28
1994-12-24
2008-09-25
0123-13-24
3221-12-09

[email]
bavduer@163.com
qfcc@sina.com
bavduer@126.org.cn
opera@139ftpcom
853942672@qq.com
aaa@1000phone.com
13843838438@aliyun.com
```