网站工具：

[测试网站1](https://regex101.com/)

[测试网站2](https://regexr-cn.com/)

[学习网站](https://codejiaonang.com/#/course/regex_chapter1/0/0)

[博客](https://www.cnblogs.com/meowv/p/12895081.html)

[可视化工具](https://jex.im/regulex)

# 修饰符和优先级

## （一）修饰符

一般正则表达式放在//中，也可以 /表达式/修饰符 对正则表达式功能扩充。

|  |  |  |
| --- | --- | --- |
| i | ignore 忽略大小写 | /表达式/i |
| g | global 全局匹配,查找所有的匹配项。 | /表达式/g |
| m | multi line多行匹配。通常和^和$匹配中。 | /表达式/m |
| s | .可以匹配换行符\n。 | /表达式/s |

默认.只能匹配换行符以外的字符

## （二）优先级

|  |  |
| --- | --- |
| **运算符** | **描述** |
| \ | 转义符 |
| (), (?:), (?=), [] | 圆括号和方括号 |
| \*, +, ?, {n}, {n,}, {n,m} | 限定符 |
| ^, $, \任何元字符、任何字符 | 定位点和序列（即：位置和顺序） |
| | | 替换，"或" 操作 字符具有高于替换运算符的优先级，使得"m|food"匹配"m"或"food"。若要匹配"mood"或"food"，请使用括号创建子表达式，从而产生"(m|f)ood"。 |

# 二、元字符

## （一）边界符

### 1.^

在区间内代表取反，区间外代表多行匹配的**行开始。**

[^a-z]：不含小写字母的数据

^java：java开头的

### 2.$

匹配开头， 多行匹配中匹配**行**结尾

Java&：匹配以java结尾的字符串

## （二）量词

表示某个模式重复次数

### 1.\*？+

\* ：{0, } 0个或多个

？ ：{0,1} 0个或一个

+ ：{1, } 1个或多个

### 2.{n}

前一个位置的字符重复了N次

\d{3}-\d{5}：010-1111

### 3.{n,}

\d{1,} :开启区间。匹配1个数字或者无数个数组。

### 4.{m,n}

前一个位置区间

\d{3,4}:可以同时匹配**123**和**1234**，**1234**5

\d{3,4}？:只能匹配**123**和**123**4，**123**45，**非贪婪模式**，能匹配3个数字，不会匹配4个数字

## （三）字符类

### 1.[]

匹配[]中**任何一个**即匹配成功。

[Jj]ava ：匹配大写java和小写java

**-简化：**

[0-9]：有数字

[a-z]：所有小写

[A-Z]：所有大写

[0-9a-zA-Z]：以上所有

### 2.预定义

指的是常见模式的简化

|  |  |
| --- | --- |
| 预定类 | 说明 |
| \d | 所有数字，相当于[0-9] |
| \D | 所有非数字，相当于[^0-9] |
| \w | 所有字符。字母数组下划线，相当于[A-Za-z0-9\_] |
| \W | 所有非字符，可以匹配特殊字符，相当于[^A-Za-z0-9\_] |
| \s | 匹配空格（换行符，制表符，空格符等），相当于[\t\r\n\v\f] |
| \S | 非空白，相当于[^\t\r\n\v\f] |
| \b | 单词边界 **单词和非单词间的位置（\w\W），**包括单词的**首尾，**单词和**空格**  s t1: 位置s位置 位置t位置1 |
| 点. | 除了\r 和\n之外的任意字符 |

# 三、正则特性

## （一）位置特性

正则表达式要目匹配字符，要么匹配位置

字符串stt： 位置1 s 位置2 t 位置3 t

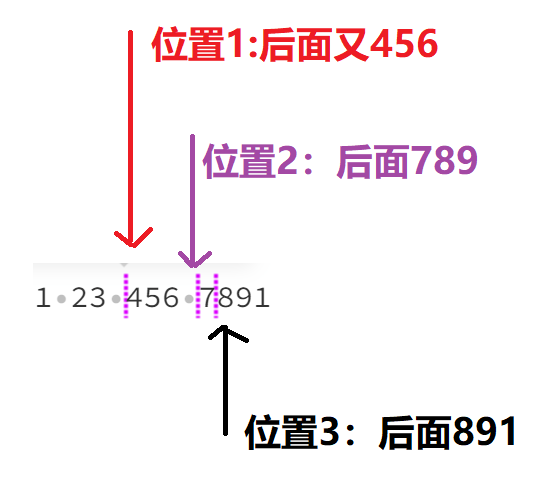
匹配特定位置3组正则

**^、$**  ：开头和结尾 ^stt$ 位置stt 位置

**\b、\B** ：单词边界 1s2 1位置s位置2

**(?=p)、(?!p)** ：断言 stt(?=t) s位置t位置t

(?=\d{3}):匹配一个位置，后面是连续的3个数字



## （二）贪婪和非贪婪

贪婪模式在整个表达式匹配成功的前提下，尽可能多的匹配，而非贪婪模式在整个表达式匹配成功的前提下，尽可能少的匹配。非贪婪模式只被部分NFA引擎所支持。

|  |  |
| --- | --- |
| 贪婪 | 非贪婪 |
| {m,n} | {m,n}? |
| {m,} | {m,}? |
| ? | ?? （用途少） |
| + | +? |
| \* | \*? |

贪婪模式

## （三）分组

### 1.分组（）

[分组](https://www.cnblogs.com/meowv/p/12895081.html)，可以进行数据的提取。也可以称为隔离。

分组后，每个分组所捕获的值，都有一个分组编号。

<p>(\w.\*)</p>：

匹配<p>hello</p>中的hello

(\d{4}).?(\d{1,2}).?(\d{1,2})：

同时匹配

2020-1-2

2020-2-2

2020-01-02

2020/01/20

### 2.|

或一般配合分组，（p1|p2）起到**分支选择**的作用

(.avi|.mp4|.wmv|.rmvb)：

同时匹配

1.avi

abc.mp4

chapter1.wmv 

chapter2.rmvb

### 3.（:?）

非捕获分组。使用括号的原始特性。不会被引用。

(?:\d{2})-(\d{5})：01-3412 其中3412会被提取，01不会

### 4.分组回溯

\数字：可以引用之前定义分组的东西

(\w)(\w)\2\1：匹配**abba**,**acca**, sf**adda**s

## （四）环视

先行断言和后行断言也称环视，预搜索。共分为：

（?<=）(?<!) 文本（?=）(?!)

### 1.正向先行断言

文本（?=条件）

（?=表达式）：表示所在位置右侧必须匹配表达式，**该位置**会被提取

1995(?=stt):1995的右侧必须是stt，1995会被提取

**1995**stt

1996stt

1995tts

(?=.\*?[a-z])(?=.\*?[A-Z])(?=.\*?[0-9]).{8,}：至少一个大小，至少一个小写，至少一个数字，至少8个字符

### 2.反向先行断言

文本（?!条件）

1995(?!stt): 1995右侧不能是stt

### 3.正向后行断言

（?<=表达式）文本

（?<=s）tt：tt左侧必须是s

### 4.反向后行断言

（?<!表达式）文本

（?<!l）tt：tt的左侧不能是l

# 四、常用正则

### 1.数字校验

|  |  |  |
| --- | --- | --- |
| **描述** | **正则表达式** | **备注** |
| 数字 | ^[0-9]\*$ |  |
| n位数字 | ^\d{n}$ |  |
| 至少n位数字 | ^\d{n,}$ |  |
| m~n位数字 | ^\d{m,n}$ |  |
| 整数 | ^(-?[1-9]\d\*)$ | 非0开头，包括正整数和负整数 |
| 正整数 | ^[1-9]\d\*$ |  |
| 负整数 | ^-[1-9]\d\*$ |  |
| 非负整数 | ^(([1-9]\d\*)|0)$ |  |
| 非正整数 | ^((-[1-9]\d\*)|0)$ |  |
| 浮点数 | ^-?(?:[1-9]\d\*\.\d\*|0\.\d\*[1-9]\d\*|0\.0+|0)$ | 包括正浮点数和负浮点数 |
| 正浮点数 | ^(?:[1-9]\d\*\.\d\*|0\.\d\*[1-9]\d\*)$ |  |
| 负浮点数 | ^-(?:[1-9]\d\*\.\d\*|0\.\d\*[1-9]\d\*)$ |  |
| 非正浮点数 | ^(?:-(?:[1-9]\d\*\.\d+|0\.\d\*[1-9]\d\*)|0\.0+|0)$ | 包含0 |
| 非负浮点数 | ^(?:[1-9]\d\*\.\d+|0\.\d+|0\.0+|0)$ | 包含0 |
| 仅一位小数 | ^-?(?:0|[1-9][0-9]\*)\.[0-9]{1}$ |  |
| 最少一位小数 | ^-?(?:0|[1-9][0-9]\*)\.[0-9]{1,}$ |  |
| 最多两位小数 | ^-?(?:0|[1-9][0-9]\*)\.[0-9]{1,2}$ |  |
| 连续重复的数字 | ^(\d)\1+$ | 例如：111，222 |

### 2.字符校验

|  |  |  |
| --- | --- | --- |
| **描述** | **正则表达式** | **备注** |
| 中文 | ^[\u4E00-\u9FA5]+$ |  |
| 全角字符 | ^[\uFF00-\uFFFF]+$ |  |
| 半角字符 | ^[\u0000-\u00FF]+$ |  |
| 英文字符串（大写） | ^[A-Z]+$ |  |
| 英文字符串（小写） | ^[a-z]+$ |  |
| 英文字符串（不区分大小写） | ^[A-Za-z]+$ |  |
| 中文和数字 | ^(?:[\u4E00-\u9FA5]{0,}|\d)+$ |  |
| 英文和数字 | ^[A-Za-z0-9]+$ |  |
| 数字、英文字母或者下划线组成的字符串 | ^\w+$ |  |
| 中文、英文、数字包括下划线 | ^[\u4E00-\u9FA5\w]+$ |  |
| 不含字母的字符串 | ^[^A-Za-z]\*$ |  |
| 连续重复的字符串 | ^(.)\1+$ | 例如：aa，bb |
| 长度为n的字符串 | ^.{n}$ |  |
|  |  |  |
|  |  |  |
|  |  |  |
| ASCII | ^[ -~]$ |  |

### 3.日期和时间校验

|  |  |  |
| --- | --- | --- |
| **描述** | **正则表达式** | **备注** |
| 日期 | ^\d{1,4}-(?:1[0-2]|0?[1-9])-(?:0?[1-9]|[1-2]\d|30|31)$ | 弱校验，例如：2022-06-12 |
| 日期 | ^(?:(?!0000)[0-9]{4}-(?:(?:0[1-9]|1[0-2])-(?:0[1-9]|1[0-9]|2[0-8])|(?:0[13-9]|1[0-2])-(?:29|30)|(?:0[13578]|1[02])-31)|(?:[0-9]{2}(?:0[48]|[2468][048]|[13579][26])|(?:0[48]|[2468][048]|[13579][26])00)-02-29)$ | 严格校验，考虑平闰年 |
| 时间 | ^(?:1[0-2]|0?[1-9]):[0-5]\d:[0-5]\d$ | 12小时制，例如：11:21:31 |
| 时间 | ^(?:[01]\d|2[0-3]):[0-5]\d:[0-5]\d$ | 24小时制，例如：23:21:31 |
| 日期+时间 | ^(\d{1,4}-(?:1[0-2]|0?[1-9])-(?:0?[1-9]|[1-2]\d|30|31)) ((?:[01]\d|2[0-3]):[0-5]\d:[0-5]\d)$ | 例如：2000-11-11 23:20:21 |

### 4.日常生活相关

|  |  |  |
| --- | --- | --- |
| **描述** | **正则表达式** | **备注** |
| 中文名 | ^[\u4E00-\u9FA5·]{2,16}$ |  |
| 英文名 | ^[a-zA-Z][a-zA-Z\s]{0,20}[a-zA-Z]$ |  |
| 车牌号 | ^[京津沪渝冀豫云辽黑湘皖鲁新苏浙赣鄂桂甘晋蒙陕吉闽贵粤青藏川宁琼使领][A-HJ-NP-Z][A-HJ-NP-Z0-9]{4}[A-HJ-NP-Z0-9挂学警港澳]$ | 不含新能源 |
| 车牌号 | ^[京津沪渝冀豫云辽黑湘皖鲁新苏浙赣鄂桂甘晋蒙陕吉闽贵粤青藏川宁琼使领][A-HJ-NP-Z](?:(?:[A-HJ-NP-Z0-9]{4}[A-HJ-NP-Z0-9挂学警港澳])|(?:(?:\d{5}[A-HJK])|(?:[A-HJK][A-HJ-NP-Z0-9][0-9]{4})))$ | 包含新能源 |
| 火车车次 | ^[GCDZTSPKXLY1-9]\d{1,4}$ | 例如：G1234 |
| 手机号 | ^(?:(?:\+|00)86)?1[3-9]\d{9}$ | 弱匹配 |
| 手机号 | ^(?:(?:\+|00)86)?1(?:(?:3[\d])|(?:4[5-79])|(?:5[0-35-9])|(?:6[5-7])|(?:7[0-8])|(?:8[\d])|(?:9[189]))\d{8}$ | 严格匹配 |
| 固话号码 | ^(?:(?:\d{3}-)?\d{8}|^(?:\d{4}-)?\d{7,8})(?:-\d+)?$ |  |
| 手机IMEI码 | ^\d{15,17}$ | 一般是15位 |
| 邮编 | ^(?:0[1-7]|1[0-356]|2[0-7]|3[0-6]|4[0-7]|5[1-7]|6[1-7]|7[0-5]|8[013-6])\d{4}$ | 例如：211100 |
| 统一社会信用代码 | ^[0-9A-HJ-NPQRTUWXY]{2}\d{6}[0-9A-HJ-NPQRTUWXY]{10}$ |  |
| 身份证号码(1代) | ^[1-9]\d{7}(?:0\d|10|11|12)(?:0[1-9]|[1-2][\d]|30|31)\d{3}$ | 15位数字 |
| 身份证号码(2代) | ^[1-9]\d{5}(?:18|19|20)\d{2}(?:0[1-9]|10|11|12)(?:0[1-9]|[1-2]\d|30|31)\d{3}[0-9Xx]$ | 18位数字 |
| QQ号 | ^[1-9][0-9]{4,}$ | 一般是5到10位 |
| 微信号 | ^[a-zA-Z][-\_a-zA-Z0-9]{5,19}$ | 一般6~20位，字母开头，可包含字母、数字、-、\_，不含特殊字符 |
| 股票代码 | ^(s[hz]|S[HZ])(000[\d]{3}|002[\d]{3}|300[\d]{3}|600[\d]{3}|60[\d]{4})$ | A股，例如：600519 |
| 银行卡卡号 | ^[1-9]{1}(?:\d{15}|\d{18})$ | 一般为19位 |

### 5.互联网相关

|  |  |  |
| --- | --- | --- |
| 描述 | 正则表达式 | 备注 |
| 域名 | ^a-zA-Z0-9{0,62}(?:\.a-zA-Z0-9{0,62})+$ | 例如：r2coding.com |
| 网址 | ^(?:https?:\/\/)?a-zA-Z0-9{0,62}(?:\.a-zA-Z0-9{0,62})+$ | 例如：<https://www.r2coding.com/> |
| 带端口号的网址(或IP) | ^(?:https?:\/\/)?[\w-]+(?:\.[\w-]+)+:\d{1,5}\/?$ | 例如：<http://127.0.0.1:8888/> |
| URL | ^https?:\/\/(?:www\.)?[-a-zA-Z0-9@:%.\*+*~#*=]{1,256}*\*.[a-zA-Z0-9()]{1,6}*\*b(?:[-a-zA-Z0-9()!@:%*\+.~#?&\/\/=]\*)$ | 例如：<https://www.r2coding.com/>#/README?id=1 |
| 邮箱email | ^[A-Za-z0-9\u4e00-\u9fa5]+@a-zA-Z0-9{0,62}(?:\.a-zA-Z0-9{0,62})+$ | 支持中文，例如：[codesheep@cs.com](mailto:codesheep@cs.com) |
| 用户名 | ^[a-zA-Z0-9\_-]{4,20}$ | 4到20位 |
| 弱密码 | ^[\w]{6,16}$ | 6~16位，包含大小写字母和数字的组合 |
| 强密码 | ^.\*(?=.{6,})(?=.\*\d)(?=.\*[A-Z])(?=.\*[a-z])(?=.\*[!@\.#$%^&\*? ]).\*$ | 至少6位，包括至少1个大写字母，1个小写字母，1个数字，1个特殊字符 |
| 端口号 | ^(?:[0-9]|1-9{1,3}|1-5{4}|60-4{3}|650-4{2}|6550-2|6553[0-5])$ | 例如：65535 |
| IPv4地址 | ^(?:(?:\d|[1-9]\d|1\d\d|2[0-4]\d|25[0-5])\.){3}(?:\d|[1-9]\d|1\d\d|2[0-4]\d|25[0-5])$ | 例如：192.168.31.1 |
| IPv4地址+端口 | ^(?:(?:\d|[1-9]\d|1\d\d|2[0-4]\d|25[0-5])\.){3}(?:\d|[1-9]\d|1\d\d|2[0-4]\d|25[0-5])(?::(?:[0-9]|1-9{1,3}|1-5{4}|60-4{3}|650-4{2}|6550-2|6553[0-5]))?$ | 例如：192.168.31.1:8080 |
| IPv6地址 | ^(([0-9a-fA-F]{1,4}:){7,7}[0-9a-fA-F]{1,4}|([0-9a-fA-F]{1,4}:){1,7}:|([0-9a-fA-F]{1,4}:){1,6}:[0-9a-fA-F]{1,4}|([0-9a-fA-F]{1,4}:){1,5}(:[0-9a-fA-F]{1,4}){1,2}|([0-9a-fA-F]{1,4}:){1,4}(:[0-9a-fA-F]{1,4}){1,3}|([0-9a-fA-F]{1,4}:){1,3}(:[0-9a-fA-F]{1,4}){1,4}|([0-9a-fA-F]{1,4}:){1,2}(:[0-9a-fA-F]{1,4}){1,5}|[0-9a-fA-F]{1,4}:((:[0-9a-fA-F]{1,4}){1,6})|:((:[0-9a-fA-F]{1,4}){1,7}|:)|fe80:(:[0-9a-fA-F]{0,4}){0,4}%[0-9a-zA-Z]{1,}|::(ffff(:0{1,4}){0,1}:){0,1}((25[0-5]|(2[0-4]|1{0,1}[0-9]){0,1}[0-9])\.){3,3}(25[0-5]|(2[0-4]|1{0,1}[0-9]){0,1}[0-9])|([0-9a-fA-F]{1,4}:){1,4}:((25[0-5]|(2[0-4]|1{0,1}[0-9]){0,1}[0-9])\.){3,3}(25[0-5]|(2[0-4]|1{0,1}[0-9]){0,1}[0-9]))$ | 例如：CDCD:910A:2222:5498:8475:1111:3900:2020 |
| IPv6地址+端口 | ^\[(([0-9a-fA-F]{1,4}:){7,7}[0-9a-fA-F]{1,4}|([0-9a-fA-F]{1,4}:){1,7}:|([0-9a-fA-F]{1,4}:){1,6}:[0-9a-fA-F]{1,4}|([0-9a-fA-F]{1,4}:){1,5}(:[0-9a-fA-F]{1,4}){1,2}|([0-9a-fA-F]{1,4}:){1,4}(:[0-9a-fA-F]{1,4}){1,3}|([0-9a-fA-F]{1,4}:){1,3}(:[0-9a-fA-F]{1,4}){1,4}|([0-9a-fA-F]{1,4}:){1,2}(:[0-9a-fA-F]{1,4}){1,5}|[0-9a-fA-F]{1,4}:((:[0-9a-fA-F]{1,4}){1,6})|:((:[0-9a-fA-F]{1,4}){1,7}|:)|fe80:(:[0-9a-fA-F]{0,4}){0,4}%[0-9a-zA-Z]{1,}|::(ffff(:0{1,4}){0,1}:){0,1}((25[0-5]|(2[0-4]|1{0,1}[0-9]){0,1}[0-9])\.){3,3}(25[0-5]|(2[0-4]|1{0,1}[0-9]){0,1}[0-9])|([0-9a-fA-F]{1,4}:){1,4}:((25[0-5]|(2[0-4]|1{0,1}[0-9]){0,1}[0-9])\.){3,3}(25[0-5]|(2[0-4]|1{0,1}[0-9]){0,1}[0-9]))\](?::(?:%5B0-9%5D%5C%7C%5B1-9%5D%5B0-9%5D%7B1,3%7D%5C%7C%5B1-5%5D%5B0-9%5D%7B4%7D%5C%7C6%5B0-4%5D%5B0-9%5D%7B3%7D%5C%7C65%5B0-4%5D%5B0-9%5D%7B2%7D%5C%7C655%5B0-2%5D%5B0-9%5D%5C%7C6553%5B0-5%5D))?$ | 例如：[CDCD:910A:2222:5498:8475:1111:3900:2020]:9800 |
| 子网掩码 | ^(?:254|252|248|240|224|192|128)\.0\.0\.0| 255\.(?:254|252|248|240|224|192|128|0)\.0\.0| 255\.255\.(?:254|252|248|240|224|192|128|0)\.0| 255\.255\.255\.(?:255|254|252|248|240|224|192|128|0)$ | 例如：255.255.255.0 |
| MAC地址 | ^(?:(?:[a-f0-9A-F]{2}:){5}|(?:[a-f0-9A-F]{2}-){5})[a-f0-9A-F]{2}$ |  |
| Version版本号 | ^\d+(?:\.\d+){2}$ | 例如：12.1.1 |
| 图片后缀 | \.(gif|png|jpg|jpeg|webp|svg|psd|bmp|tif)+ | 可按需增删扩展名集合 |
| 视频后缀 | \.(swf|avi|flv|mpg|rm|mov|wav|asf|3gp|mkv|rmvb|mp4)+ | 可按需增删扩展名集合 |
| 图片链接 | (?:https?:\/\/)?a-zA-Z0-9{0,62}(?:\.a-zA-Z0-9{0,62})+.+\.(gif|png|jpg|jpeg|webp|svg|psd|bmp|tif) | 可按需增删扩展名集合 |
| 视频链接 | (?:https?:\/\/)?a-zA-Z0-9{0,62}(?:\.a-zA-Z0-9{0,62})+.+\.(swf|avi|flv|mpg|rm|mov|wav|asf|3gp|mkv|rmvb|mp4) | 可按需增删扩展名集合 |
| 迅雷链接 | thunderx?:\/\/[a-zA-Z\d]+= |  |
| ed2k链接 | ed2k:\/\/\|file\|.+\|\/ |  |
| 磁力链接 | magnet:\?xt=urn:btih:[0-9a-fA-F]{40,}.\* |  |
|  |  |  |

### 6.日常生活

|  |  |  |
| --- | --- | --- |
| **描述** | **正则表达式** | **备注** |
| MD5格式 | ^(?:[a-f\d]{32}|[A-F\d]{32})$ | 32位MD5，例如：7552E7071B118CBFFEC8C930455B4297 |
| BASE64格式 | ^\s\*data:(?:[a-z]+\/[a-z0-9-+.]+(?:;[a-z-]+=[a-z0-9-]+)?)?(?:;base64)?,([a-z0-9!$&',()\*+;=\-.\_~:@/?%\s]\*?)\s\*$ | 例如：data:image/jpeg;base64,xxxx== |
| UUID | ^[a-f\d]{4}(?:[a-f\d]{4}-){4}[a-f\d]{12}$ | 例如：94f9d45a-71b0-4b3c-b69d-20c4bc9c8fdd |
| 16进制 | ^[A-Fa-f0-9]+$ | 例如：FFFFFF |
| 16进制颜色 | ^#?([0-9a-fA-F]{3}|[0-9a-fA-F]{6})$ | 例如：#FFFFFF |
| SQL语句 | ^(?:select|drop|delete|create|update|insert).\*$ |  |
| Java包名 | ^(?:[a-zA-Z\_]\w\*)+(?:[.][a-zA-Z\_]\w\*)+$ | 例如：com.r2coding.controller |
| 文件扩展名 | \.(?:doc|pdf|txt) | 可按需增删扩展名集合 |
| Windows文件路径 | ^[a-zA-Z]:(?:\\[\w\u4E00-\u9FA5\s]+)+[.\w\u4E00-\u9FA5\s]+$ | 例如：C:\Users\Administrator\Desktop\a.txt |
| Windows文件夹路径 | ^[a-zA-Z]:(?:\\[\w\u4E00-\u9FA5\s]+)+$ | 例如：C:\Users\Administrator\Desktop |
| Linux文件路径 | ^\/(?:[^/]+\/)\*[^/]+$ | 例如：/root/library/a.txt |
| Linux文件夹路径 | ^\/(?:[^/]+\/)\*$ | 例如：/root/library/ |