

Regular Expression introduction

正则指引

小龙

正则是什么鬼？

```
</?[a-z][-a-z0-9_:.]*(?  
=[\s>])('^[']*'|"^["]*"|  
^[^'"]*>)*>
```

模式匹配

- ❖ ``
- ❖ ``
- ❖ `</h1>`
- ❖ `
`

正则基础

- ❖ 字符组: `[abc]` `[012345]` `[a-z]` `[0-9]`
- ❖ 排除型字符组: `[^abc]` `[^012345]` `[^a-z]` `[^0-9]`
- ❖ 元字符: `^` `$` `.` `\w` `\W` `\s` `\S` `\d` `\D` `\b`
- ❖ 量词: `*` `+` `?` `{n}` `{n, m}`
- ❖ 忽略优先量词: `*?` `+?` `??` `{n, m}?`
- ❖ 字符转义: `\.` `*`

正则基础

- ❖ 分组: $(ab)^+$
- ❖ 多选结构: $(ab|cd)$
- ❖ 断言: $(?=...)$ $(?!...)$ $(?<=...)$ $(?<!...)$
- ❖ 反向引用 非捕获分组 匹配模式 NFA DFA ...

举个栗子

❖ `url(r'^api/(?P<model>(sku| hiring |
news | enterpriseneeds)+)/v2/?
$', api.CommentCreateNewApi.as_view()
)`,

再举一栗

- ❖ var reg = /[^].+\ / (activitydocument\ / .+)
\$/; /static/upload/
activitydocument/
2016/05/23/3883507937.pdf
- ❖ var reg = /[^].+?\ / (activitydocument\ / .
+)\$/;

再看什么鬼

```
</?[a-z][-a-z0-9_:.]*(?  
=[\s>])('^[']*'|"^["]*"|  
^[^'"]*>)*>
```

一些注意点

- ❖ .号并不能匹配所有字符
- ❖ 注意贪婪和非贪婪的使用
- ❖ 正则不是万能的

深入正则

- ❖ 精通正则表达式
- ❖ 正则指引