

# 正则表达式快速参考手册

胡志飞 <WisdomFusion#gmail.com>

2016 年 2 月 26 日

# 目 录

<b>1</b>	<b>简介</b>	<b>1</b>
<b>2</b>	<b>基本语法</b>	<b>1</b>
<b>3</b>	<b>高级语法</b>	<b>6</b>
<b>4</b>	<b>举些栗子</b>	<b>8</b>
<b>5</b>	<b>正则表达式“流派”</b>	<b>8</b>
<b>6</b>	<b>应用场景</b>	<b>9</b>
6.1	工具箱 . . . . .	9
6.2	应用案例 . . . . .	9

# 1 简介 INTRODUCTION

中文文字  
中文文字

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

# 2 基本语法 BASIC SYNTAX

特性	语法	描述	举个栗子
字符	除 <code>[\\^\$. ?*\+O</code> 以外的任意字符	除了 <code>[\\^\$. ?*\+O</code> 以外的任意字符，和也是文字文本，除了下面说到的成对出现的量词语法，如 <code>n</code> 和 <code>m,n</code> 等。	<code>a</code> 匹配 <code>about</code> 中的 <code>a</code>
	转义	<code>\\t, \\?, \\*, \\+, \\., \\ , \\{, \\}, \\[, \\], \\(, \\)</code>	

To be continued...

(续表)

特性	语法	描述	举个栗子
	<div>\n, \r</div> <div>\cA 到 \cZ,</div> <div>\ca 到 \cz</div> <div>\a</div> <div>\f</div> <div>\v</div>		
基本特性	<div>.</div> <div> </div>	<div>Matches any single character except line break characters. Most regex flavors have an option to make the dot match line break characters too.</div>	<div>. 匹配 about 中的任意一个字符</div>
字符类	<div>[abc], [^\[]</div> <div>[^abc]</div> <div>\d, \D</div> <div>\w, \W</div>		

To be continued...

(续表)

特性	语法	描述	举个栗子
	\s, \S		
POSIX	[:alnum:]		
	[:alpha:]		
	[:ascii:]		
	[:blank:]		
	[:cntrl:]		
	[:digit:]		
	[:graph:]		
	[:lower:]		
	[:print:]		
	[:punct:]		
	[:space:]		
	[:upper:]		

To be continued...

(续表)

特性	语法	描述	举个栗子
	<code>[:word:]</code>		
	<code>[:xdigit:]</code>		
锚点	<code>^</code>		
	<code>\$</code>		
	<code>\A</code>		
	<code>\G</code>		
	<code>\Z</code>		
	<code>\b</code>		
	<code>\B</code>		
	<code> &lt;</code>		
	<code> &gt;</code>		
	<code> `</code>		
	<code> &amp;</code>		

To be continued...

(续表)

特性	语法	描述	举个栗子
	\'		
量词	?		
	??		
	?+		
	*		
	*?		
	*+		
	+		
	+?		
	++		
	{n}		
	{n,m}		

To be continued...





(续表)

特性	语法	描述	举个栗子
分组与反向引用	(regex) (?:regex) \1 到 \9 \10 到 \99 \g{1} 到 \g{99} \g{-1}, \g{-2}, etc. (?<name>regex) \k<name>, \g{name}		
高级分组	(?#comment)		

To be continued...

(续表)

特性	语法	描述	举个栗子
	(?=Regex)		
	(?!Regex)		
	(?<=regex)		
	(?<!regex)		

4 举些栗子 REGEX EXAMPLES

一些栗子

5 正则表达式“流派” REGEX FLAVORS

流派

## 6 应用场景 APPLICATION SCENARIOS

### 6.1 工具箱 TOOLBOX

总有一款适合你，Windows 下的记事本太鸡肋，Word 处理方式主要是“通配符”而不是正则表达式。

#### RegexBuddy

JGsoft 开发的一个强大的正则表达式测试工具，

#### grep

#### PowerGREP

RegexBuddy 的兄弟软件，同是 JGsoft 开发，是 grep 在 Windows

平台的实现和增强。

#### UltraEdit, Notepad++

#### Vim

#### GNU Emacs

#### sed & awk

### 6.2 应用案例 APPLICATION CASES

#### Dreamweaver 表格处理

#### VBA 中使用正则表达式

```

1 Sub IndentParaWithRegEx()
2 ' PowerPoint VBA 批量给指定字符开头段落加动画
3 Dim oSld As Slide
4 Dim oShp As Shape
5 Dim i As Integer
6 ' 正则相关变量
7 Dim regx As Object, oMatch As Object
8
9 ' 这里写查找的正则，参考 http://msdn.microsoft.com/en-us/library/ms974570.aspx
10 strPattern = "^开头字符串"
11
12 Set regx = CreateObject("vbscript.regexp")
13 With regx

```

```
14 .Global = True
15 .IgnoreCase = True
16 .Pattern = strPattern
17 End With
18
19 For Each oSld In ActivePresentation.Slides
20     For Each oShp In oSld.Shapes
21         If oShp.HasTextFrame Then
22             If oShp.TextFrame2.HasText Then
23                 With oShp.TextFrame2.TextRange
24                     For i = 1 To .Paragraphs.Count
25                         With .Paragraphs(i)
26                             '可能会出现多个匹配项的
27                             If (regx.Test(.Text) = True) Then
28                                 .ParagraphFormat.FirstLineIndent = 0
29                             End If
30                         End With
31                     End With
32                     Next i 'para
33                 End With
34             End If 'has text
35         End If 'has textframe
36     Next oShp
37 Next oSld
End Sub
```

## InDesign GREP

使用 **Perl** 正则表达式处理文件

神的编辑器之正则