**术语**

• 通配符指的是您可以用来代表一个或多个字符的键盘字符。例如，星号 (\*) 通常代表一个或多个字符，问号 (?) 通常代表单个字符。

• 对我们来说，正则表达式指的是您可以用来查找和替换文本模式的文本字符和通配符组合。文本字符指的是必须存在于目标文本字符串中的文本。通配符指的是目标字符串中可能各不相同的文本。

1. 为什么要在Word中使用正则表达式？

在Word不支持正则表达式前，查找替换都是用很硬的方式，就是找什么或是把什么替换成什么，没有太多的规则或是说模糊性，但是现实世界往往又不是如此的清晰，有些规则带有这样那样的条件，在这种情况下，正则表达式就可以发挥出巨大的功效了。

举个简单的例子，一篇文章中，为了分割，有的地方用一个空格，有的地方用两个，有的地方用n个，现在要把所有这样的地方都用一个空格来替换。

2. 如何在Word中使用正则表达式

和普通的查找替换稍有不同，当用户选择“use wildcards”（中文应该是支持通配符）后，用户的查找和替换就支持正则表达式了。如下图，可以查找6位及以上的数字。

对上面举的例子，只要在Find What中输入“ {1,}”，在替换为中输入“ ”，就可完成此项任务，是不是超简单吧。

本节中的步骤介绍了如何使用正则表达式转置姓名。请记住，始终使用“查找和替换”对话框来运行您的正则表达式。同时请记住，如果表达式没有按预期工作，你始终可以按下 CTRL + Z 来撤销您的更改，然后尝试其他表达式。

**转置姓名**

1. 启动 Word，然后打开一个新的空白文档。

2. 复制此表格，将它粘贴到该文档中。

Josh Barnhill

Doris Hartwig

Tamara Johnston

Daniel Shimshoni

3. 在“开始”选项卡上的“编辑”组中，单击“替换”以打开“查找和替换”对话框。

4. 如果您没有看到“使用通配符”复选框，请单击“更多”，然后选中该复选框。如果您没有选中该复选框，Word 会将通配符视作文本。

5. 在“查找内容”框中键入以下字符。请确保您在两组括号之间包含了空格：

(<\*>) (<\*>)

6. 在“替换为”框中，键入以下字符。请确保您在逗号和第二个斜杠之间包含了空格：

\2, \1

7. 选择该表格，然后单击“全部替换”。Word 会转置这些姓名并使用逗号分隔它们，如下所示：

Barnhill, Josh

Hartwig, Doris

Johnston, Tamara

Shimshoni, Daniel

此时，您可能会想知道：如果您的姓名中有一部分或全部包含中间名首写字母，该怎么做？ 请参阅使用正则表达式中的第一个示例以了解更多信息。

**两个汉字中间有一个空格，要把这个空格替换掉，但两个英文之间的空格不替换；**

查找内容：([一-﨩])( )([一-﨩])

替换为：\1\3

搜索选项：使用通配符

With Selection.Find

.Text = "([一-﨩])( )([一-﨩])"

.Replacement.Text = "\1\3"

.Forward = True

.Wrap = wdFindContinue

.Format = False

.MatchCase = False

.MatchWholeWord = False

.MatchByte = False

.MatchAllWordForms = False

.MatchSoundsLike = False

.MatchWildcards = True

End With

Selection.Find.Execute Replace:=wdReplaceAll

**正则表达式的工作原理**

从此处开始，请记住这条原则：文档内容决定了您绝大多数（并非全部）正则表达式的设计。例如，在您之前使用的示例表格中，每个单元格都包含了两个单词。如果单元格包含两个单词和一个中间名首写字母，您将使用不同的表达式。

让我们由内而外检查每个表达式：

在第一个表达式 (<\*>) (<\*>) 中：

• 星号 (\*) 会返回单词中的全部文本。

• 小于和大于符号 (< >) 分别标记了每个单词的开始和结束。它们确保搜索会返回单个单词。

• 它们之间的括号和空格将单词拆分成不同的组：（第一个单词） （第二个单词）。括号还指示了您希望搜索以何种顺序处理每个表达式。

换而言之，该表达式说：“查找两个单词。”

在第二个表达式 \2, \1 中：

• 斜杠 (\) 与数字放在一起充当占位符。（您也可以使用斜杠来查找其他通配符。请参阅下一节以了解更多信息。）

• 第一个占位符后的逗号在两个被转置的姓名之间插入正确的标点符号。

换而言之，该表达式说：“写入第二个单词，添加逗号，写入第一个单词。”

**要查找和替换的项目的通配符**

可以使用通配符搜索文本。例如，可使用星号 (\*) 通配符搜索字符串（例如，使用“s\*d”将找到“sad”和“started”）。

使用通配符查找和替换文本

1. 在“开始”选项卡上的“编辑”组中，单击“查找”旁边的箭头，然后单击“高级查找”。

2. 选中“使用通配符”复选框。

如果看不到“使用通配符”复选框，请单击“更多”。

3. 请执行下列操作之一：

• 若要从列表中选择通配符，请单击“特殊格式”，再单击通配符，然后在“查找内容”框中键入任何其他文本。要了解更多信息，请参阅可用的通配符表。

• 在“查找内容”框中直接键入通配符。要了解更多信息，请参阅“可用的通配符”表格。

5. 如果要替换该项目，请单击“替换”选项卡，然后在“替换为”框中键入要用作替换的内容。

6. 单击“查找下一处”、“查找全部”、“替换”或“全部替换”。

若要取消正在执行的搜索，请按 Esc。

**可用的通配符**

注释

• 选中“使用通配符”复选框后，Word 将只查找与指定的内容精确匹配的文本。请注意，“区分大小写”和“全字匹配”复选框将不可用（灰显），以表示这些选项已自动开启。您无法关闭这些选项。

• 若要搜索已被定义为通配符的字符，请在该字符前面键入一个反斜杠 (\)。例如，键入 \? 来查找问号。

• 可使用括号对通配符和文本进行分组，以指明处理顺序。例如，键入 <(pre)\*(ed)> 可以找到“预排序的”和“已阻止的”。

• 可使用 \n 通配符搜索表达式，然后将其替换为经过重新排列的表达式。例如，在“查找内容”框中键入 (Ashton) (Chris)，并在“替换为”框中键入 \2 \1。Word 将找到 Ashton Chris 并将其替换为 Chris Ashton。

|  |  |  |  |
| --- | --- | --- | --- |
| 用于查找 | 类型 |  | 示例 |
| 任一字符 | ? | s?t | 可找到“sat”和“set”。 |
| 任何字符串 | \* | s\*d | 可找到“sad”和“started”。 |
| 单词开头 | < | <(inter) | 可找到“interesting”和“intercept”，但找不到“splintered”。 |
| 单词结尾 | > | (in)> | 可找到“in”和“within”，但找不到“interesting”。 |
| 指定字符之一 | [ ] | w[io]n | 可找到“win”和“won”。 |
| 此范围内的任一字符 | [-] | [r-t]ight | 可找到“right”和“sight”。范围必须是升序。 |
| 除了括号内范围中的字符之外的任一字符 | [!x-z] | t[!a-m]ck | 可找到“tock”和“tuck”，但找不到“tack”或“tick”。 |
| 前一个字符或表达式的 n 个匹配项 | {n} | fe{2}d | 可找到“feed”，但找不到“fed”。 |
| 前一个字符或表达式的至少 n 个匹配项 | {n,} | fe{1,}d | 可找到“fed”和“feed”。 |
| 前一个字符或表达式的 n 到 m 个匹配项 | {n,m} | 10{1,3} | 可找到“10”、“100”和“1000”。 |
| 前一个字符或表达式的一个或多个匹配项 | @ | lo@t | 可找到“lot”和“loot”。 |
| 表达式查找，尤其用于多个关键词的组合查找 | （） |  | 在 “查找内容”中键入 “（America）（China）”，在 “替换为”中键 入 “\2\1”，Word 找到 “America China”并替换为 “China America”。 |

**使用正则表达式**

这些示例会向您展示一些在 Microsoft Word 中使用正则表达式的方法

示例 1：转置带中间名首写字母的姓名

此示例使用通配符和字符代码组合来转置包含中间名首写字母的姓名。如果您不熟悉字符代码，请参阅“Word 帮助”主题查找和替换 Word 2010 文件中的文本和其他数据。

在操作过程中，请记住下列事实：

• 在对表格中的姓名使用此表达式时，您必须首先将该表格转换成文本。

• 如果表格包含多个列，请将包含姓名的列复制到空白文档中，然后在那里将它转换成文本。

• 转置姓名后，请将文本重新转换成表格。然后您可以删除原始列并使用您已更改的数据替换它。

准备示例数据

1. 如果您还没完成这一步，请启动 Word 并创建一个新的空白文档。

2. 将空白表格插入到该文档中。将表格设为宽一列高四行。

3. 逐个复制这些姓名，并将每个姓名粘贴到空白表格单元格中：

Joshua Quentin Barnhill

Doris X. Hartwig

Tamara Y. Johnston

Daniel Shimshoni

您的表格应如下所示：

Joshua Quentin Barnhill

Doris X. Hartwig

Tamara Y. Johnston

Daniel Shimshoni

4. 选择该表格，然后在“表格工具布局”选项卡上的“数据”组中，单击“转换成文本”。

5. 选择“段落标记”作为文本分隔符，然后单击“确定”。

转置带首写字母的姓名

1. 在“开始”选项卡上的“编辑”组中，单击“替换”以打开“查找和替换”对话框。

2. 选中“使用通配符”复选框（您可能需要单击“更多”以查看该复选框），然后在“查找内容”框中键入以下表达式：

(\*) ([! ]@)^13

请确保在两组括号之间和感叹号之后输入空格。如果您之前没有看到 ^13 字符，我们会在下一节介绍它的作用。

3. 在“替换为”框中，键入以下表达式：

\2, \1^p

4. 选择姓名列表，然后单击“全部替换”。Word 会转置这些姓名以及中间名首写字母或中间名，如下所示：

Barnhill，Joshua Quentin

Hartwig，Doris X.

Johnston，Tamara Y.

Shimshoni，Daniel

将已更改的文本重新转换为表格

1. 请选择已转置的姓名列表。

2. 在“插入”选项卡上的“表格”组中，单击“表格”，然后单击“将文本转换成表格”。

会打开“将文本转换成表格”对话框。

3. 在“文字分隔位置”下，单击“段落”，然后单击“确定”。

逐个查看表达式

让我们从“查找内容”框中的表达式开始，逐个查看这些表达式，从而了解它们是如何工作的。

整个表达式会查找两组模式：一个带有中间名首写字母（或中间名）的名和一个姓。(\*) 会查找全部名。请注意，它后面有一个空格。

表达式的此部分与姓相匹配：

([! ]@)^13

感叹号会排除括号中指定的任何字符。在此案例中，[! ] 代表“查找除空格外的全部信息”。它的作用是从姓前面剪裁掉空格。

@ 字符会查找前一个字符的一个或多个匹配项，因此，它的作用是确保删除姓前面的全部空格。

我们需要知道姓的结尾处，因此我们还使用 ^13 字符来搜索每一行结尾处的段落标记。但是，由于我们不准备重新使用该段落标记，所以我们会用括号将所有其他的内容括起来。

您可以通过以下步骤来尝试这一操作：将姓名再次复制到您的测试文档中（请确保使用段落标记分隔它们），然后使用 ([! ]@)^13 在“查找内容”框中进行搜索。搜索会与每个姓相匹配。

因为搜索会在下一行的起始处重新开始，所以我们使用星号通配符 (\*) 来匹配此处到下一个姓的起始处之间的全部内容。

由于我们不准备在姓前面重新使用空格，我们使用括号来将它从两组中排除：

(\*) ([! ]@)^13

要点 请谨慎使用 ^13 字符代码。通常，您可以 ^p 字符代码来搜索段落标记。但是，无法在通配符搜索中使用该代码。您需要替换为使用代码 ^13。尽管无法在通配符搜索中使用 ^p 字符代码，您应当在通配符替换操作中使用它。这是为什么？ ^p 字符包含了格式信息，^13 字符则没有。此外，您完全无法将样式信息分配给 ^13 字符。在替换操作中错误使用 ^13 代码会在实质上将您的文档转换成无法设置格式的文件。

“替换”表达式 (\2 \1) 会完成实际转置。在“替换为”框中，\2 字符让搜索首先写入第二个模式，然后在该模式后添加逗号。\1^p 字符指定了写入第一个模式的位置并让搜索在该模式后写入段落标记。

示例 2：转置日期

您可以使用此处显示的正则表达式来将欧洲格式的日期转换为美国格式的日期。

转置日期

1. 复制以下日期并粘贴到您的文档中：28th May 2003

2. 打开“查找和替换”对话框，然后在“查找内容”框中键入以下表达式：

([0-9]{1,2})([dhnrst]{2}) (<[ADFJMNOS]\*>) ([0-9]{4})

请确保在以下左括号和右括号之间插入空格： 2}) (<[ and \*>) ([0.

3. 在“替换为”框中输入以下表达式：

\3 \1, \4

请确保在每个字符集之间插入空格。

4. 单击“全部替换”。

搜索会将 28th May, 2003 替换为 May 28, 2003。

逐个查看表达式

让我们从“查找内容”框中的表达式开始。该表达式的作用是将日期分解为四个部分，由多组括号来表示。每个模式都包含您在全部日期（以您在示例中使用的样式书写）中查找的组件。按从左到右的顺序：

• 数字范围 [0-9] 匹配第一个模式中的一位数数字。因为日期可以由两个数字组成，我们让搜索返回一位数或两位数的日期：{1,2}。结果为第一个模式：([0-9]{1,2})。

序数词组成了第二个模式。序数词包含“th”、“nd”、“st”和“rd”，因此我们将这些字母添加到范围 [dhnrst] 中。因为序数词总是由两个字母组成，我们将字母计数限制为 2：([dhnrst]{2})。

• 接下来是空格，之后是查找月份名的文本字符和通配符。所有月份名都以这些大写字母开头：ADFJMNOS。我们不知道每个大写字母后有多少个字符，因此我们在它们之后使用星号 (\*)。我们只对月份名本身感兴趣，因此我们使用大于和小于字符来将结果限制为单个单词。结果为第四个模式：(<[ADFJMNOS]\*>)。

• 最后，我们搜索年。我们使用相同的数字范围，但这次我们将计数限制为四个字符 ([0-9]{4})。

请注意，我们在“替换为”框中只写入四个地址模式中的三个。我们从日期中省略了序数词 ("th")，因为美国格式的日期不使用序数词。如果您想在日期中保留序数词，请在“替换为”框中输入 \3 \1\2, \4。在本例中，您在 3 和逗号之后都输入了空格，其他地方则没有。

此时，您可能要问：如何处理没有拼写月份名的日期？例如， 28/05/03。使用此表达式搜索：

([0-9]{1,2})/([0-9]{1,2})/([0-9]{2})

使用此表达式替换：

\3/\1/\2

如果日期使用的格式为 28/05/2003，则在最后一个模式中使用 {4}，而非 {2}。

关于在正则表达式中使用列表分隔符

上面的例子使用以下参数来查找一位数或两位数日期：{1,2}。在本例中，逗号将两个值分隔开。但是请记住，您在 Windows 中的区域设置会控制您使用的列表分隔符。如果您的区域设置将分号的使用指定为列表分隔符，则您必须使用分号，而非逗号。

要找出您的操作系统指定的列表分隔符，请执行以下操作：

Windows 7

1. 单击“开始”按钮 ，然后单击“控制面板”。

2. 单击“时钟、语言和区域”。

3. 单击“更改日期、时间或数字格式”，然后单击“其他设置”。

4. 单击“数字”选项卡，然后找到“列表分隔符”条目。

Windows Vista

1. 单击“开始”按钮 ，然后单击“控制面板”。

2. 单击“时钟、语言和区域”。

3. 单击“区域和语言选项”。

4. 在“格式”选项卡上的“当前格式”下，单击“自定义此格式”。

5. 单击“数字”选项卡，然后找到“列表分隔符”条目。

Windows XP

1. 单击“开始”，然后单击“控制面板”。

2. 双击“区域和语言选项”。

3. 在“区域”选项卡上，单击“自定义”。

4. 单击“数字”选项卡，然后找到“列表分隔符”条目。

示例 3：向称呼添加句号，或从称呼删除句号

在某些国家，尊称（Mr.、Mrs. 等）不包括句号。此示例为您展示了如何向尊称添加句号，或从尊称删除句号。从现在开始，我们假定您知道如何使用“查找和替换”对话框。

此表达式会查找没有句号的 Mr、Ms、Mrs 和 Dr：

<([DM][rs]{1,2})( )

请注意，表达式会使用另一个包含空格的模型。该空格将紧随尊称之后（如果那里没有句号）。此表达式会添加句号：

\1.\2

要反向操作，请使用此表达式进行搜索：

<([DM][rs]{1,2})。

使用此表达式进行替换：

\1

示例 4：查找重复段落或行

在使用此表达式时，您可能希望首先对列表排序，以便将重复行相邻放置。同样的，您需要删除全部空白段落标记。换言之，如果您使用空白段落来分隔文本块，如下所示：

Joshua Quentin Barnhill¶

¶

Joshua Quentin Barnhill¶

¶

Doris X. Hartwig¶

您需要删除那些段落，如下所示：

Joshua Quentin Barnhill¶

Joshua Quentin Barnhill¶

Doris X. Hartwig¶

您可以用自己喜欢的方法来删除空白段落，但是，我们正在谈论正则表达式，因此这里就有一个能够查找两个连续段落字符的方法。使用此表达式进行搜索（@ 字符会重复查找和替换操作并删除全部多个空行）：

(^13)\1@

使用此表达式替换结果：

^p

现在，我们来看看替换文本的方法。此表达式会查找两个连续相同段落的任意序列：

(\*^13)\1

此表达式还会匹配在段落中结束的较长文本重复项。例如，对以下列表运行该表达式：

Joshua Quentin Barnhill¶

Doris X. Hartwig¶

Joshua Quentin Barnhill¶

Doris X. Hartwig¶

Tamara Y. Johnston¶

搜索会查找前四行，只有当总体模式更改时才会停止。与之相反，如果您对此列表运行该表达式：

Joshua Quentin Barnhill¶

Joshua Quentin Barnhill

Doris X. Hartwig¶

Doris X. Hartwig¶

表达式会只查找前两个段落。

要搜索更多的相同项目，请添加更多占位符。例如，此表达式会查找三个连续相同段落：

(\*^13)\1\1

您还可以使用大括号来达到同样的目的。以下示例会分别查找两个和三个相同段落：

(\*^13){2}

(\*^13){3}

或者，您可以查找两个或三个相同段落：

(\*^13){2,3}

您还可以查找两个或更多相同段落：

(\*^13){2,}

可以使用以下字符串替换这些表达式中的任意一个：

\1

此外，您可以在必要时重复查找和替换操作，以便替换您文档中的全部重复段落，或者您可以添加 @ 通配符，让表达式为您重复该操作：

(\*^13)\1@

您还可以使用此方法来替换表格中的重复行。要做到这一点，请首先删除合并的单元格，然后为表格排序以相邻放置重复单元格。接着，请将您的表格转换成文本。（在“表格”菜单上，指向“转换”，然后单击“表格到文本”；出现提示时，使用制表符分隔符。）完成替换后，请将文本重新转换成表格。

更多示例

要了解更多如何在 Word 中使用正则表达式的示例，请参阅 MVP FAQ 网站上的使用通配符查找和替换字符。

转自：<http://office.microsoft.com/zh-cn/word-help/HA102350661.aspx#_Available_wildcards>

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **GNU grep** | **GNU Emacs** | **Perl** | **MS Word** |
| 任何单个字符 | . | . | . | ? |
| 符合其中的任一字符 | [...] | [...] | [...] | [...] |
| 符合除此之外的任一字符 | [^...] | [^...] | [^...] | [!...] |
| 子表达式 | /(*re*/) | /(*re*/) | (*re*) | (*re*) |
| 表达式重复多次 | *re*/{.../} | *re*/{.../} | *re*{...} | *re*{...} |
| 表达式出现0或1次 | *re*/? | *re*? | *re*? | 无 |
| 表达式出现0或多次 | re\* | re\* | re\* | 无 |
| 表达式出现1或多次 | *re*/+ | *re*+ | *re*+ | *re*@ |
| 表达式出现0或1次(lazy mode) | 无 | *re*?? | *re*?? | 无 |
| 表达式出现0或多次(lazy mode) | 无 | re\*? | re\*? | 无 |
| 表达式出现1或多次(lazy mode) | 无 | *re*+? | *re*+? | 无 |
| 单词起始 | /< | /< | 无 | < |
| 单词结尾 | /> | /> | 无 | > |
| 单词边界(起始或结尾) | /b | /b | /b | 无 |
| 非单词边界 | /B | /B | /B | 无 |
| 任意字符串 | .\* | .\* | .\* | \* |
| 行首 | ^ | ^ | ^ | 无 |
| 行末 | $ | $ | $ | 无 |
| 换行符 | /n | C-q C-j | /n | ^13 |
| ”或“ | re/|re | re/|re | re|re | 无 |
| 引用上次找到的内容 | 不知道 | /& | 不知道 | ^& |
| 引用被括号捕获的内容 | /digit | /digit | /digit | /digit |
| 含义 | GNU grep | GNU Emacs | Perl | MS Word |

注：

re 代表一个正则表达式 (RegExp)

MS Word里多次匹配，最少是一次，没有0次。

/digit 指 /后面加数字 如 /1 /2

正则表达式里括号子表达式可以嵌套使用，Word里好像不行

正则表达式30分钟入门教程

http://manual.phpv.net/regular\_expression.html#mission