Regex replace to change the order of words in a sentence

Ask Question



If I have a string like this:

4

"word1 'word2' word3"



is it possible to use a regex replace to change the string to this:



3

"word1 word3 'word2""

I know what word1 and word3 will be, but do not know what word2 will be, but it will always be in single quotes.

regex

edited Sep 14 '10 at 18:12



itsmatt

27.1k 9 93 155

asked Sep 14 '10 at 18:00



Jeremy

24k 59 169 288

Will there always be three words like this and always need to swap the second and third? – R. Martinho Fernandes Sep 14 '10 at 18:04 ✓

4 Answers



You can replace "word1 ('\w+') word3" With "word1 word3 \1". The replace syntax might be different in other regex engines; I'm using .NET's which is based on Perl's.



 \w+ matches a series of word characters, aka a word. You can change this if it does not fit your definition of word;



- The parenthesis are used to delimit groups within the expression. The first is group one, the second is group two, etc. Usually regex engines count group zero as the entire match;
- The \1 in the replace string means to use group one from the match, \2 means group two, etc.

edited Sep 14 '10 at 18:16

answered Sep 14 '10 at 18:05



R. Martinho Fernandes **165k** 59 383 470

The question states "I know what word1 and word3 will be", hence no need for ($\w+$). – z5h Sep 14 '10 at 18:13

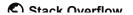
@z5h: yeah, you're right, I didn't notice. I'll make it more specific. – R. Martinho Fernandes Sep 14 '10 at 18:15 ✓

It's ok, I derived what I needed from the answer. Thanks! – Jeremy Sep 14 '10 at 23:10

If you are doing a replace inside of Visual Studio's

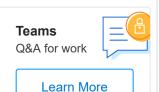
Home

PUBLIC



Users

Jobs





I would say:

4

s/"(word1)\s+('.+?')\s+(word3)"/"\$1 \$3 \$2"/



answered Sep 14 '10 at 18:20



Toto

67.4k 17 58 100



Search for: \"word1 (\'[a-zA-Z0-9]+\') word3\" And replace with "word1 word3 \1"



You could also use (.+) for word2 (in capture group #1), if you want to match more than just alphanumeric characters. I think \w could also be used, but for some reason I don't use that very often. Maybe I don't always agree with the regex engine about what should be a "word" character, so I prefer to be more explicit.

edited Sep 14 '10 at 18:19

answered Sep 14 '10 at 18:05



21.1k 27 118 176



Following the other suggestions...

this expression will make sure that word3 is the final word



allows for any characters for your 'words' except single and double quotes.

edited Sep 14 '10 at 18:17

FrustratedWithFormsDe signer

21.1k 27 118 176

answered Sep 14 '10 at 18:08



Brad

2.2k 6 30 50