
Subject: Re: Removing (or replacing) substrings in a string array
Posted by [Michael Galloy](#) on Thu, 23 Jan 2014 02:50:05 GMT
[View Forum Message](#) <> [Reply to Message](#)

On 1/22/14, 3:57 pm, Michael Galloy wrote:

> On 1/22/14, 7:13 AM, Mats Löfdahl wrote:

>> Say I have a string array where each string may or may not begin with
>> a dot (".") and may or may not end with a dot. What is an efficient
>> (non-loop?) way of removing those dots in IDL 7?

>>

>> The reason I specify IDL 7 is that I realize this could be done with
>> strsplit in IDL 8 but in IDL 7 strsplit does not support string
>> arrays.

>>

>>

>> Alternatively, if I could make it so those dots are not there in the
>> first place, that's even better. What I'm trying to do is to parse an
>> array of strings, where the fields are separated by dots but I cannot
>> identify the substrings by their position. What I can do is to
>> identify them with regular expressions because they all have
>> different forms. And stregex supports string arrays also in IDL 7.

>>

>> So I can do something like

>> strmid(stregex(strlist, '\.[0-9]{5}\.', /extr), 1, 5) if I know I have a
>> five-digit number surrounded by dots. I remove the dots with the
>> strmid command.

>>

>> But any field can also be first or last, so to find it I need to do
>> stregex(strlist, '(\.|^)[0-9]{5}(\.|\$)', /extr). But then one of the
>> dots will be missing in those cases when the field really is first or
>> last, so the simple strmid operation does not work.

>>

>> And some fields can have variable lengths but be identified by an
>> initial character, so I could find it with something like
>> stregex(strlist, '(\.|^)[+-][0-9]+(\.|\$)', /extr). In this case strmid
>> does not work even when there is a dot at the end, because I don't
>> know its position.

>>

>> Would be great if I could tell stregex to return everything but the
>> dots (if present) but I don't know if that is possible. Can't you do
>> something like that with regular expressions in the shell? Or was it
>> in elisp?

>>

>>

>> I wanted to educate myself and went looking for "Making Regular
>> Expressions Your Friends" by Mike Galloy but the links seem to be
>> dead in <http://michaelgalloy.com/2006/06/11/regular-expressions.html>
>> and the link in

>> http://www.exelisvis.com/docs/Learning_About_Regular_E.html leads to
>> a page telling me that "The mgunit project site has moved to Github"
>> and when I follow the provided link to github I can't find the
>> document.
>>
>>
>> Hmm... I guess if I substitute '.'+strlist+'.' for strlist in all
>> stregex calls I don't have to worry about the case where there are no
>> dots. But that seems not so elegant and it still does not solve the
>> problem with unknown lengths.
>>
>
> Check out MG_STREPLACE:
>
>
> https://github.com/mgalloy/mglib/blob/master/src/strings/mg_streplace.pro
>
> I have that "Making Regular Expressions Your Friends" article here
> somewhere too. I will update the link on my website and post here when I
> find it. Have to run now...
>
> Mike

I updated the article at:

<http://michaelgalloy.com/2006/06/11/regular-expressions.html>

with working links to the article and the example code. I would use
MG_STREPLACE from my Github repo over the STR_REPLACE linked to in the
article.

Mike

--

Michael Galloy

www.michaelgalloy.com

Modern IDL: A Guide to IDL Programming (<http://modernidl.idldev.com>)

Research Mathematician

Tech-X Corporation
