Subject: Re: Efficient way to split an array of strings.
Posted by David Fanning on Thu, 15 Nov 2007 01:39:32 GMT

bberkey@gmail.com writes:

> have a lot of data that comes to me in tabular files with a nice mix
> of strings and numerical data. Typically I will read the data into a
> string array. And then loop over each line in the file calling
> STRSPLIT repeatably.
>
> Is there a better way to accomplish this?

Why don't you just read the data from the file
directly:

```
OpenR, lun, 'test.dat', /Get_LUN
rows = File_Lines('test.dat')
struct = {f:0.0D, I:0, s:''}
data = Replicate(struct, rows)
ReadF, lun, data
Free_Lun, lun

Print, data.f, Format='(F10.2)'
Print, data.I
Print, data.s
```

Cheers,

David

--
David Fanning, Ph.D.
Fanning Software Consulting, Inc.
Coyote's Guide to IDL Programming: http://www.dfanning.com/
Sepore ma de ni thui. ("Perhaps thou speakest truth.")

Subject: Re: Efficient way to split an array of strings.
Posted by bberkey on Thu, 15 Nov 2007 09:16:10 GMT

On Nov 14, 3:39 pm, David Fanning <n...@dfanning.com> wrote:
> bber...@gmail.com writes:
>> have a lot of data that comes to me in tabular files with a nice mix
>> of strings and numerical data. Typically I will read the data into a
>> string array. And then loop over each line in the file calling
>> STRSPLIT repeatably.

>
>> Is there a better way to accomplish this?
>
> Why don't you just read the data from the file
> directly:
>
>   OpenR, lun, 'test.dat', /Get_LUN
>   rows = File_Lines('test.dat')
>   struct = {f:0.0D, I:0, s:''}
>   data = Replicate(struct, rows)
>   ReadF, lun, data
>   Free_Lun, lun
>
>   Print, data.f, Format='(F10.2)'
>   Print, data.I
>   Print, data.s
>
> Cheers,
>
> David
>
> --
> David Fanning, Ph.D.
> Fanning Software Consulting, Inc.
> Coyote's Guide to IDL Programming:http://www.dfanning.com/
> Sepore ma de ni thui. ("Perhaps thou speakest truth.")

David,

Thanks, I think I will start doing that from now on. But since my
last post I though of a much better way to split an array of strings.
Something like the following:
   reform(strsplit(strjoin(data,delimiter),delimiter,/
extract),numstrings,rows)

I quick performance check shows that both of these methods are about
5x faster then my old method.

Thanks
-Ben

Subject: Re: Efficient way to split an array of strings.
Posted by Brian Larsen on Thu, 15 Nov 2007 15:13:28 GMT
View Forum Message <> Reply to Message

Wow, I had forgotten (or maybe never knew) that you could feed readf
an array and it would read more than one line...

I just tested it on one read routine I have where the file has a bunch of header then 8192 integers one per line, then more header, then another 8192 integers and so on.

Going to the "array readf" was a x140 improvement in speed!!

This is why I read this group, thanks a million.

Cheers,

Brian

 -- --------------
Brian Larsen
Boston University
Center for Space Physics