Subject: Re: file_lines and expand path
Posted by Paul Van Delst[1] on Fri, 04 Aug 2006 19:35:15 GMT
View Forum Message <> Reply to Message

Jean H. wrote:
> Paul van Delst wrote:
>> FL wrote:
>>
>>> Hi,
>>>
>>> I am implementing file_lines in FL and was wondering what to return
>>> when the input path expands to multiple files. The IDL reference
>>> guide says nothing. I have run IDL and got curious result:
>>>
>>> IDL> help, file_lines('*.pro')
>>> <Expression>   LONG64   =   15
>>>
>>> I have 448 .pro files in the current directory, with a total of 10k
>>> lines.
>>> IDL seems to be randomly picking a single file and reporting its
>>> length (probably the first match in readdir()).
>>>
>>> What do you expect to get?
>>>
>>> Possibilities:
>>>
>>> 1. never expand (but why do we have NOEXPAND then?)
>>> 2. give an error message if the expansion results in more than one file
>>> 3. return the total length of files
>>> 4. return the total length of files if a new keyword TOTAL was set
>>> 5. do as IDL, pick a single file silently
>>
>>
>> Well, if the routine allows for wildcard inputs, then I wouldn't
>> expect *any* of the above.
>>
>> What I would expect would be an array (in this case 448 elements in
>> size) containing the lengths of *each* file that matched the wildcard
>> expression, e.g.
>
> what about a structure containing the array you said, and another one
> containing the complete file name, so then it becomes fairly easy to
> open the files (and the order in wich the files are read in the
> directory is not a problem). This would allow you to read only files
> that have more than X lines in a very convenient way!

I think that would be marvelous (I've never used file_lines so maybe it does something like that already?)

However, to keep it more compatible with IDL, maybe rather than a structure it can have an optional output keyword argument, i.e.

```
IDL> fl=file_lines('*.pro',files=filenames)
IDL> help, fl, files
FL              LONG64    = Array[448]
FILENAMES       STRING    = Array[448]
```

?

Hang on a minute.... in the IDL file_lines documentation it states:

Return Value

Returns the number of lines of text contained within the specified file or files. If an
array of file names is specified via the Path parameter, the return value is an array with
the same number of elements as Path, with each element containing the number of lines in
the corresponding file.

Arguments

  Path
A scalar string or string array containing the names of the text files for which the
number of lines is desired.

So, the IDL version of file_lines *should* return an array when multiple files are given. Does this mean IDL has a bug in file_lines() with regards to how it handles wildcard input? The words associated with the /NOEXPAND_PATH suggest that wildcard inputs are allowed (although it doesn't specifically say it).

paulv

--
Paul van Delst   Ride lots.
CIMSS @ NOAA/NCEP/EMC   Eddy Merckx
Ph: (301)763-8000 x7748
Fax:(301)763-8545