
Subject: Re: Read Total lines in an ASCII file
Posted by [David Burrridge](#) on Fri, 13 Dec 2002 18:08:39 GMT
[View Forum Message](#) <> [Reply to Message](#)

Hi Maria,

Actually, I beleive there is a new feature in IDL 5.6 that does this. But assuming you're still on IDL 5.5, something like:

```
noLines = 0L
line = ""
OpenR, lun, 'mytextfile.txt', /Get_lun
While Not EOF (lun) Do Begin
    ReadF, lun, line
    noLines = noLines + 1
EndWhile
```

Will do the trick. If (as I suspect) you then want to read the file, don't forget to:

```
Point_Lun, lun, 0
```

before you do. Otherwise, you'll need a:

```
Free_Lun, lun
```

to close the file and make the allocated file unit available to IDL again.

On the other point, will a simple:

```
mychar EQ 'P'
```

in a statement do the trick? Examples

```
If mychar EQ 'P' Then Print, 'Hello'
```

```
mybool = mychar EQ 'P'
```

Hope this helps.

Best regards,

David

David Burrridge

Burridge Computing

18 The Green South

Warborough

Oxon

OX10 7DN

Tel: 01865 858279

Mobile: 0780 244 1748

Email: davidb@burridgecomputing.co.uk

"Maria" <msmimb@hotmail.com> wrote in message
news:d19b702e.0212130942.1d6d6dbf@posting.google.com...
> I bet my question seems simple to all of you but...does anybody know
> how to read the total number of lines in an ASCII file?
>
> Also, is there any command in IDL such as if (variable_char is char)
> then give me a boolean (true or false)?
>
> Thanks a lot!
> Maria.

Outgoing mail is certified Virus Free.

Checked by AVG anti-virus system (<http://www.grisoft.com>).

Version: 6.0.404 / Virus Database: 228 - Release Date: 15/10/2002

Subject: Re: Read Total lines in an ASCII file
Posted by [Med Bennett](#) on Fri, 13 Dec 2002 18:39:43 GMT
[View Forum Message](#) <> [Reply to Message](#)

David Burridge wrote:

> Hi Maria,
>
> Actually, I beleive there is a new feature in IDL 5.6 that does this. But
> assuming you're still on IDL 5.5, something like:

>
> noLines = 0L
> line = "
> OpenR, lun, 'mytextfile.txt', /Get_lun
> While Not EOF (lun) Do Begin
> ReadF, lun, line
> noLines = noLines + 1
> EndWhile
>
> Will do the trick. If (as I suspect) you then want to read the file, don't
> forget to:
>
> Point_Lun, lun, 0
>
> before you do. Otherwise, you'll need a:
>
> Free_Lun, lun
>
> to close the file and make the allocated file unit available to IDL again.
>
> On the other point, will a simple:
>
> mychar EQ 'P'
>
> in a statement do the trick? Examples
>
> If mychar EQ 'P' Then Print, 'Hello'
>
> mybool = mychar EQ 'P'
>
> Hope this helps.
>
> Best regards,
>
> David
>
> David Burrridge
>
> Burrridge Computing
>
> 18 The Green South
>
> Warborough
>
> Oxon
>
> OX10 7DN
>

> Tel: 01865 858279
>
> Mobile: 0780 244 1748
>
> Email: davidb@burridgecomputing.co.uk
>
> "Maria" <msmimb@hotmail.com> wrote in message
> news:d19b702e.0212130942.1d6d6dbf@posting.google.com...
>> I bet my question seems simple to all of you but...does anybody know
>> how to read the total number of lines in an ASCII file?
>>
>> Also, is there any command in IDL such as if (variable_char is char)
>> then give me a boolean (true or false)?
>>
>> Thanks a lot!
>> Maria.

That should work fine, but I would think that it would be slow for large files because of the loop. I do this instead - also somewhat crude but maybe faster:

```
junk=strarr(1000000L)
on_ioerror,done
openr,lun,'mytextfile.txt'
readf,lun,junk
stop,'need to increase array size'
done:
close,lun
junk = junk[where(strlen(junk) gt 0)]
end
```

The readf will generate an error unless your file is bigger than the initial array; if that is the case, you have to make it bigger.

Subject: Re: Read Total lines in an ASCII file
Posted by [David Fanning](#) on Fri, 13 Dec 2002 18:53:59 GMT
[View Forum Message](#) <> [Reply to Message](#)

Med Bennett (no.spam@this.address.please) writes:

> That should work fine, but I would think that it would be slow for large files
> because of the loop. I do this instead - also somewhat crude but maybe
> faster:

I once thought of writing an article about all the goofy ways people have devised to count the lines in their files.
(Have you ever wondered why we are so attached to columns

of data? I wonder if it has something to do with the hard-wiring in our brains. Something about bananas hanging from the top of a tree, maybe.)

Anyway, IDL 5.6 has FILE_LINES, which will give us all a consistent way to count lines from now on. Now, all we have to do it convince everyone to upgrade... :-)

Cheers,

David

--

David W. Fanning, Ph.D.
Fanning Software Consulting, Inc.
Phone: 970-221-0438, E-mail: david@dfanning.com
Coyote's Guide to IDL Programming: <http://www.dfanning.com/>
Toll-Free IDL Book Orders: 1-888-461-0155

Subject: Re: Read Total lines in an ASCII file
Posted by [wmconnolley](#) on Fri, 13 Dec 2002 22:42:53 GMT
[View Forum Message](#) <> [Reply to Message](#)

David Fanning <david@dfanning.com> wrote:

> Anyway, IDL 5.6 has FILE_LINES, which will give us
> all a consistent way to count lines from now on. Now,
> all we have to do it convince everyone to upgrade... :-)

If you are lucky enough to be running under unix, then

```
spawn,'cat filename|wc -l',number_of_lines
```

should work.

-W.

--

William M Connolley | wmc@bas.ac.uk | <http://www.nerc-bas.ac.uk/icd/wmc/>
Climate Modeller, British Antarctic Survey | Disclaimer: I speak for myself
I'm a .signature virus! copy me into your .signature file & help me spread!

Subject: Re: Read Total lines in an ASCII file
Posted by [Mark Hadfield](#) on Sun, 15 Dec 2002 21:09:52 GMT
[View Forum Message](#) <> [Reply to Message](#)

"Med Bennett" <no.spam@this.address.please> wrote in message news:3DFA296F.65A8E1A0@this.address.please...

[IDL code for counting lines in code omitted]

```
> That should work fine, but I would think that it would be slow for
> large files because of the loop. I do this instead - also somewhat
> crude but maybe faster:
>
> junk=strarr(1000000L)
> on_ioerror,done
> openr,lun,'mytextfile.txt'
> readf,lun,junk
> stop,'need to increase array size'
> done:
> close,lun
> junk = junk[where(strlen(junk) gt 0)]
> end
```

Hmmm. This doesn't actually determine the number of lines in the file. It reads & returns all the non-empty lines.

Of course, when people ask for a way to count the lines in a file, usually what they want to do next is to read the file contents. Addressing the problem, "how do I read all the data from an ASCII file with an unknown number of lines?". The first thing to do is to go to David's site and see what he says. I found this

http://www.dfanning.com/tips/unknown_rows.html

It points to some useful routines, but doesn't really discuss the general approaches. I can think of three:

1 - Pre-allocate an array big enough to hold the maximum expected amount of data, read the data into it, then trim the array.

2 - Read the file once to count lines, allocate a data array of exactly the right size, then read the file again to store the data.

3 - Read the file once, storing the data in an extensible data structure, then (optionally) copy the data out of the extensible structure into an array.

No. 3 is the most flexible and arguably the most aesthetically pleasing, but unfortunately you will have to write the "extensible data structure" yourself (or use someone else's), since IDL doesn't have anything suitable. IDL arrays *look* like they can be extended, but in fact every time you extend an IDL array you create a new one.

No. 2 seems very inefficient, but with disk caching it often turns out that reading a file twice doesn't take much longer than reading it once.

No. 1 (as Med proposes) is probably the fastest, but has the disadvantage of a built-in hard limit that will bite you when you are reading **really** big files. To some extent the choice between them will depend on what it is that you want to read out of each line. Will you want to skip any lines?

I have done some comparisons in the past and will dig them out if I can.

--

Mark Hadfield "Ka puwaha te tai nei, Hoesa tatou"
m.hadfield@niwa.co.nz
National Institute for Water and Atmospheric Research (NIWA)
