Subject: Re: Determining length of a text file?
Posted by Craig Markwardt on Tue, 21 Sep 2004 15:18:11 GMT
View Forum Message <> Reply to Message

lloyd@evilprofessor.co.uk (Lloyd Watkin) writes:

> [...]
>> It is usually more time efficient to allocate a big text array, read,
>> ignore the error, and truncate the empty text strings thus avoiding a
>> FOR loop. Since all null strings point to the same (null) memory (I
>> think) you only use a small temporary memory. For large table this can
>> improve speed a lot.
>
> I have a file of a few million lines to read in. Obviously it takes
> awhile to read in (using while EOF), and would like to speed it up.
>
> How do I get the code to ignore the error generated?

You can trap errors using ON_IOERROR.

Another way you might approach this is to read the data in as
bytes. For example, a 256 kB array like this, and read it
unformatted,

```
  bb = bytarr(256L*1024L)

  ; ...

  bb(*) = 0
  readu, unit, bb
```

You will still need to trap the I/O error which will probably come on
the last read, since your file will probably not be a multiple of
256k. Initializing BB to zero is important since if you only get a
partial read, then the previous contents will remain.

Then just search for the line-ending character of choice, usually
either ascii 13 or 10,

```
  wh = where(bb EQ 13, ct)
  ; or
  wh = where(bb EQ 10, ct)

  total = total + ct
```

Roll that into a loop, and you probably can't get faster than that
from within IDL. Since you are reading a large amount of data at
once, 256k, the overhead for looping is very low.

Craig

--

```
 --------------------------------------------------------- -------------
```
Craig B. Markwardt, Ph.D.   EMAIL: craigmnet@REMOVEcow.physics.wisc.edu
Astrophysics, IDL, Finance, Derivatives | Remove "net" for better response
```
 --------------------------------------------------------- -------------
```