Subject: Re: Continuing trouble with ASCII file. Trouble with line ends
Posted by Peter Brooker on Wed, 05 Jan 2000 08:00:00 GMT
View Forum Message <> Reply to Message

I have determined the following information:

DOS/Windows: End of line is CR/LF
Unix: End of line is LF
Mac: End of line is CR

The ASCII file I am trying to read was produced on a MAC
so most likely the EOL character is a single CR. I will start a
new post in which I will ask for help in this specific issue.
The subject will be "Help Reading MAC ascii files in Windows"

peter brooker wrote:

> I have been having allot of trouble reading in a multiline
> ascii file to IDL. I am reading each line into a string by a
> call to readf for each line. Much thanks to Liam for
> showing me how to do that! The problem I am having has
> something to do with how the file represents a line
> end/carriage return. The line end is not being interpreted
> correctly by IDL. An example of what I am talking about is
> below.
>
> DEVICE MEAS
> J27H   0.3806
>
> In my editor (NT Notepad) there is a black vertical
> rectangle between "meas" and "J27H". This actually is the
> end of the first line. The way I want IDL to read this would
> be to read "DEVICE MEAS" on the first call to readf and then
> "J27H   0.3806" on the second call to readf. Instead it
> just reads in "DEVICE MEAS" on the first call to readf and
> then returns eof on the next call.
>
> Other editors such as MSWord recoginze the line end
> character. If I copy the above text and paste it into a word
> document it pastes in the text as
>
> DEVICE MEAS
> J27H 0.3806
>
> I can load the file also into excel. In recognizes the file
> as "Windows ANSI" and "tab delimited".
>
> I have discovered a bizarre work around. I enclosed the file

> in an email, send it, open it, and then drag it back to the
> desktop. When I open the file it appears correctly as two
> lines in the NT notepad editor. IDL can then read it in
> correctly also.
>
> Is there someway that IDL can read the file correctly
> without having to resort to the work around?
>
> thanks
>
> Peter Brooker
>
> * Sent from AltaVista http://www.altavista.com Where you can also find related Web Pages,
Images, Audios, Videos, News, and Shopping. Smart is Beautiful

---

Subject: Re: Continuing trouble with ASCII file. Trouble with line ends
Posted by m218003 on Thu, 06 Jan 2000 08:00:00 GMT
View Forum Message <> Reply to Message

In article <2107d358.8afa2864@usw-ex0110-076.remarq.com>,
 peter brooker <ra5589NOraSPAM@email.sps.mot.com.invalid> writes:
> The problem I am having has
> something to do with how the file represents a line
> end/carriage return. The line end is not being interpreted
> correctly by IDL. An example of what I am talking about is
> below.
>
> DEVICE MEAS
J27H 0.3806
>
> In my editor (NT Notepad) there is a black vertical
> rectangle between "meas" and "J27H". This actually is the
> end of the first line. The way I want IDL to read this would
> be to read "DEVICE MEAS" on the first call to readf and then
> "J27H   0.3806" on the second call to readf. Instead it
> just reads in "DEVICE MEAS" on the first call to readf and
> then returns eof on the next call.
>
> Other editors such as MSWord recoginze the line end
> character. If I copy the above text and paste it into a word
> document it pastes in the text as
>
> DEVICE MEAS
> J27H   0.3806
>
Have you tried to save the file again using Word? If you select TEXT
formatted it should put it out ok (i.e. replace CR by CR/LF). On a Unix

system you would probably prefer a shell script using perl, sed or awk
(this allows batch processing in case you have many of those files).
I have so far only converted Windows to Unix (remove the CR from CR/LF).
It would involve a little toying around to do it for the Mac since the
Mac's line ends are not recognized as such.

Regards,
Martin

--
[[[[[[[[[[
[[Dr. Martin Schultz   Max-Planck-Institut fuer Meteorologie   [[
[[Bundesstr. 55, 20146 Hamburg   [[
[[phone: +49 40 41173-308   [[
[[fax: +49 40 41173-298   [[
[[martin.schultz@dkrz.de   [[
[[[[[[[[[[

Subject: Re: Continuing trouble with ASCII file. Trouble with line ends
Posted by Peter Mason on Thu, 06 Jan 2000 08:00:00 GMT
View Forum Message <> Reply to Message

peter brooker <ra5589NOraSPAM@email.sps.mot.com.invalid> wrote:
> ...The problem I am having has something to do with how the file
> represents a line end/carriage return. The line end is not being
> interpreted correctly by IDL.
<...>
> Other editors such as MSWord recoginze the line end character.
<...>

Well, those ASCII files of yours sure are putting up a good fight!
I once came across an "ASCII" format that used character 13 ("carriage
return") as a line terminator and experienced similar problems to
yours.   I wasn't inclined to get too fancy with a solution and so
resorted to steam power as well, but with *automated* coal shoveling.
I modified my reader to do the following:
1] Read the entire file into a byte array BDAT.
2] Replace all bytes in BDAT that have value=13b with value 10b
(a.k.a. "line-feed" - something that IDL understands to mean "end of
line")
3] Write the modified BDAT to a temporary file.
4] Business as usual with the temporary file, which is now really ASCII.
5] Delete the temporary file.

Parts 1-3 look like this:
  Openr,u,fname,/get_lun,/noautomode

```
N=fstat(u) &n=n.size
Buf=bytarr(n)
Readu,u,buf
Free_lun,u
;
J=where(buf eq 13b,tc)
If tc gt 0L then buf[temporary(j)]=10b
;
Openw,u,'just.pvisiting',/noautomode,/get_lun
Writeu,u,temporary(buf)
Free_lun,u
```

Part 4 starts like this:
```
  Openr,u,'just.visiting',/noautomode,/get_lun,/delete
```
..and then it's over to you.

Part 5 takes care of itself when you FREE_LUN,U at the end.

I'd recommend that you dump the first, say 120 bytes of one of your
files and see what your rogue character is.   (e.g., Use the section of
code above except specify BUF=BYTARR(120<n) and then just have an extra
line that does PRINTF,BUF after the read.)   It might not be character
13 as it was in my case.

Good luck,
Cheers
Peter Mason

Subject: Re: Continuing trouble with ASCII file. Trouble with line ends
Posted by Struan Gray on Tue, 11 Jan 2000 08:00:00 GMT
View Forum Message <> Reply to Message

Peter Brooker, ra5589@email.sps.mot.com writes:

> DOS/Windows: End of line is CR/LF
> Unix: End of line is LF
> Mac: End of line is CR

 Forgive me if I'm teaching you how to suck eggs, but if you're
programming IDL (or anything else) on a Mac it's well worth getting
hold of the text editor BBedit or the free version, BBedit Lite. Among

other things it allows you to choose which line ending convention to use. Bare Bones Software have more info on their website: www.barebones.com

Struan