
Subject: Re: A very simple FORMAT question

Posted by [David Fanning](#) on Wed, 21 Apr 2004 19:31:26 GMT

[View Forum Message](#) <> [Reply to Message](#)

"Andry William (Please remove \".spam\")" <andry@ya.com.spam> writes:

> Sorry for this very simple question but I would appreciate if
> someone can help in the use of FORMAT in READF command.

>

> Here is the simple data:

>

> ----- test.datos -----

> |020922 10:00:00 | 339| 348| 63| 356| 400| 88|

> |020922 10:00:00 | 339| 348| 63| 356| 400| 88|

> |020922 11:00:00 | 345| 349| 111| 422| 482| 140|

> |020922 12:00:00 | 339| 340| 113| 431| 533| 137|

> |020922 13:00:00 | 335| 333| 99| 410| 492| 127|

> ----- End data -----

>

> Here is the simple program to read it.

>

> ----- Simple IDL program -----

> flnm= 'test.datos'

> SPAWN, "wc -l "+ flnm+ " | awk '{print \$1}'", lineas

> ; This can be done in IDL without any problem

>

> lineas= LONG(lineas[0])

> data= INTARR(11, lineas)

>

> OPENR, 1, flnm

> READF, 1, data, format= '(1x,i6,3(1x,i2),4x,7(1x,i6))'

> CLOSE,1

> END

> ----- End IDL program -----

>

> For some reason (I don't understand why) the READF command

> reaches the end of the file and can not close it.

It looks to me that you have 10 things to read on each line,
and you are trying to read 11 in your format statement. Could
that be the problem?

Cheers,

David

--

David Fanning, Ph.D.

Subject: Re: A very simple FORMAT question
Posted by [wmconnolley](#) on Wed, 21 Apr 2004 19:38:44 GMT
[View Forum Message](#) <> [Reply to Message](#)

"Andry William (Please remove \".spam\)") <andry> wrote:
> ----- test.datos -----
> |020922 10:00:00 | 339| 348| 63| 356| 400| 88|
> ----- End data -----

> lineas= LONG(lineas[0])
> data= INTARR(11, lineas)

I count 10 not 11?

> OPENR, 1, flnm
> READF, 1, data, format= '(1x,i6,3(1x,i2),4x,7(1x,i6))'

should be 6 not 7(ix,i6)?

BTW, you could start of with

spawn,"perl -i -pe 's/|:/ /g' test.datos"

which would change your : and |'s into spaces, whereupon your
reading could be simplified.

-W.

--

William M Connolley | wmc@bas.ac.uk | <http://www.antarctica.ac.uk/met/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: A very simple FORMAT question
Posted by [Andry William \(Please](#) on Thu, 22 Apr 2004 00:29:15 GMT
[View Forum Message](#) <> [Reply to Message](#)

Here is the simple data:

----- test.datos -----
|020922 10:00:00 | 339| 348| 63| 356| 400| 88|
|020922 10:00:00 | 339| 348| 63| 356| 400| 88|

```
|020922 11:00:00 | 345| 349| 111| 422| 482| 140|
|020922 12:00:00 | 339| 340| 113| 431| 533| 137|
|020922 13:00:00 | 335| 333| 99| 410| 492| 127|
----- End data -----
```

Here is the simple program to read it.

```
----- Simple IDL program -----
flnm= 'test.datos'
SPAWN, "wc -l "+ flnm+ " | awk '{print $1}'", lineas
; This can be done in IDL without any problem

lineas= LONG(lineas[0])
data= INTARR(10, lineas)

OPENR, 1, flnm
READF, 1, data, format= '(1x,i6,3(1x,i2),4x,7(1x,i6))'
CLOSE,1
END
```

> It looks to me that you have 10 things to read on each line,
> and you are trying to read 11 in your format statement. Could
> that be the problem?

That was just a mistyping. I just reproduced part of my data
in order to have it fill the e-mail editor. It should read 10
instead of 11. Here the result when reading the file:

```
IDL> print,data
20922 10 0 0 339 348 63 356 400 88
20922 11 0 345 349 111 20922 12 0 339
340 113 20922 13 0 335 333 99 0 0
0 0 0 0 0 0 0 0 0 0
```

Where does go the second "0" of the second line? The problem
keeps occuring even for more columns of data.

Andry

Subject: Re: A very simple FORMAT question
Posted by [Andry William \(Please\)](#) on Thu, 22 Apr 2004 00:35:41 GMT
[View Forum Message](#) <> [Reply to Message](#)

wmc@bas.ac.uk wrote:
>> ----- test.datos -----
>> |020922 10:00:00 | 339| 348| 63| 356| 400| 88|

```
>> ----- End data -----
>
>> data= INTARR(11, lineas)
>>READF, 1, data, format= '(1x,i6,3(1x,i2),4x,7(1x,i6))'
>
> I count 10 not 11?
> should be 6 not 7(ix,i6)?
```

You are right. It was just a mistyping on the e-mail editor. The original code has the above variables.

```
>
> BTW, you could start of with
>
> spawn,"perl -i -pe 's/|:/ /g' test.datos"
>
> which would change your : and |'s into spaces, whereupon your
> reading could be simplified.
```

Does it make some difference if I have "|" or ":" instead of space?

On the other side, the data I have come from some other database software and generated by some other people with a lot more data. I am also not supposed to alter the data since after my checking I should pass it to another program (with the original format), pretty picky (I agree).

Thanks again for any input.

Andry

Subject: Re: A very simple FORMAT question
 Posted by [wmconnolley](#) on Thu, 22 Apr 2004 08:20:59 GMT
[View Forum Message](#) <> [Reply to Message](#)

"Andry William (Please remove \"spam\")" <andry> wrote:

```
> wmc@bas.ac.uk wrote:
>>> ----- test.datos -----
>>> |020922 10:00:00 | 339| 348| 63| 356| 400| 88|
>>> ----- End data -----
>>
>>> data= INTARR(11, lineas)
>>> READF, 1, data, format= '(1x,i6,3(1x,i2),4x,7(1x,i6))'
>>
>> I count 10 not 11?
>> should be 6 not 7(ix,i6)?
```

> You are right. It was just a mistyping on the e-mail editor. The
> original code has the above variables.
>>
>> BTW, you could start of with
>>
>> spawn,"perl -i -pe 's|:/ /g' test.datos"
>>
>> which would change your : and |'s into spaces, whereupon your
>> reading could be simplified.

> Does it make some difference if I have "|" or ":" instead of
> space?

Not as you've written it. But if you read in as

```
data= INTARR(11, lineas)
readf,1,data
```

and skip the format (as you can if you've removed the |'s etc in favour
of spaces) you have a simpler life.

> On the other side, the data I have come from some other database
> software and generated by some other people with a lot more data.
> I am also not supposed to alter the data since after my checking
> I should pass it to another program (with the original format),
> pretty picky (I agree).

-W.

--

William M Connolley | wmc@bas.ac.uk | <http://www.antarctica.ac.uk/met/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: A very simple FORMAT question
Posted by [Chris Lee](#) on Thu, 22 Apr 2004 09:06:16 GMT
[View Forum Message](#) <> [Reply to Message](#)

In article <4086EE70.9070707@ya.com.spam>, "Unknown" <andry@ya.com.spam>
wrote:

> Hi all,
> ----- test.datos -----
> |020922 10:00:00 | 339| 348| 63| 356| 400| 88|
> |020922 10:00:00 | 339| 348| 63| 356| 400| 88|
> -----
> SPAWN, "wc -l "+ flnm+ " | awk '{print \$1}'", lineas

```
> ...
> READF, 1, data, format= '(1x,i6,3(1x,i2),4x,7(1x,i6))'
```

Hi,

The line counting line confused me, I use a function called FILE_LINES (I'm not sure if it's local or not, it appears to be undocumented if it's a library function).

As for the formatting, it works fine for single line reads. So you can either read one line at a time or use STRSPLIT on each line (which can then be read in all at once)

e.g.

```
flnm= 'test.datos'
n_data=10
n_lines=file_lines(flnm)
d=strARR(n_lines)
data=intarr(n_data,n_lines)
```

```
OPENR, 1, flnm
readf,1,d
CLOSE,1
```

```
for i=0, n_lines-1 do $
  data[* ,i]= fix(strsplit(d[i],[' :'],'/regex,/ex))
```

I don't know why the FORMAT is broken, only that READF seems to lose its position in the file and applies the format to the wrong string.

The 'missing second zero' is actually the first zero (the minutes in the time string), which is being used as the the third dummy argument in the format string. After that, everything blows up.

Chris.

Subject: Re: A very simple FORMAT question
Posted by [Chris Lee](#) on Fri, 23 Apr 2004 11:18:34 GMT
[View Forum Message](#) <> [Reply to Message](#)

```
> In article <4086EE70.9070707@ya.com.spam>, "Unknown" <andry@ya.com.spam>
> wrote:
>
>> Hi all,
>> ----- test.datos -----
>> |020922 10:00:00 | 339| 348| 63| 356| 400| 88| |020922
>> 10:00:00 | 339| 348| 63| 356| 400| 88|
>> -----
```

```
>> SPAWN, "wc -l "+ flnm+ " | awk '{print $1}'", linesas
>> ...
>> READF, 1, data, format= '(1x,i6,3(1x,i2),4x,7(1x,i6))'
> Hi,
>
```

Aha, L'inspiration, or something.

```
READF, 1, data, format= '(1x,i6,3(1x,i2),4x,6(1x,i6))'
```

does NOT work

```
READF, 1, data, format= '((1x,i6,3(1x,i2),4x,6(1x,i6)))'
```

does work fine (note the extra brackets).

It's not an IDL bug either, Fortran (at least the Intel and f77 compiler) both have the same behaviour. I do think it's bizzare that the everything after the repeat statement (i.e 3(1x,i2)) is repeated in the READ, but not the bit before it.

Chris.

Subject: Re: A very simple FORMAT question

Posted by [Andry William \(Please\)](#) on Fri, 23 Apr 2004 18:35:57 GMT

[View Forum Message](#) <> [Reply to Message](#)

```
>>> ----- test.datos -----
>>> |020922 10:00:00 | 339| 348| 63| 356| 400| 88| |020922
>>> 10:00:00 | 339| 348| 63| 356| 400| 88|
>>> -----
>>> SPAWN, "wc -l "+ flnm+ " | awk '{print $1}'", linesas
>>> ...
>>> READF, 1, data, format= '(1x,i6,3(1x,i2),4x,7(1x,i6))'
>
> Aha, L'inspiration, or something.
Bravo,
>
> READF, 1, data, format= '(1x,i6,3(1x,i2),4x,6(1x,i6))'
>
> does NOT work
>
> READF, 1, data, format= '((1x,i6,3(1x,i2),4x,6(1x,i6)))'
>
> does work fine (note the extra brackets).
```

This is just the command I am missing. I have never thought I

need to write 2 brackets in order to be able to read the whole data with the specified format. It works (as it should be)....!!! :)
I have been looking for this solution for very long time.

Thank you very much for all of the suggestion.

Any explanation for these extra brackets? (for IDL, ifc, ...)

Cheers,

A.

Subject: Re: A very simple FORMAT question
Posted by [Andry William \(Please\)](#) on Fri, 23 Apr 2004 18:38:30 GMT
[View Forum Message](#) <> [Reply to Message](#)

> The line counting line confused me, I use a function called
> FILE_LINES (I'm not sure if it's local or not, it appears to
> be undocumented if it's a library function).

FILE_LINES is documented in the IDL help (at least with the V 6.0).

A.

Subject: Re: A very simple FORMAT question
Posted by [Christopher Lee](#) on Sat, 24 Apr 2004 11:26:52 GMT
[View Forum Message](#) <> [Reply to Message](#)

> Any explanation for these extra brackets? (for IDL, ifc, ...)
>
> Cheers,
>
> A.

The answer to this lies with our old friend Google....

<http://www.hef.kun.nl/~pflk/teaching/fortran/ftn-AB.html>

"The rule for rescanning formats is complicated by the presence of brackets surrounding groups of format codes. The full rule for rescan is as follows. If there are no nested brackets, rescan returns to the beginning of the format. Otherwise, the rescan returns to the left bracket which matches the right bracket nearest to the closing bracket of the format. If this left bracket is preceded by a repeat count, this count takes effect during the rescan."

so...

```
print, findgen(6),format='(i," A",i," B",i," C")'
```

```
  0 A      1 B      2 C  
  3 A      4 B      5 C
```

```
print, findgen(6),format='(i," A",(i," B"),i," C")'
```

```
  0 A      1 B      2 C  
  3 B      4 C  
  5 B
```

If you want the read to include the terms before the first `_real_` brackets in a repeat, you need an extra set of opening brackets.

Chris.
