Subject: Re: READF data containing numbers and strings
Posted by Doug Rowland on Thu, 30 May 2002 14:51:05 GMT

View Forum Message <> Reply to Message

I would start with reading the entire line in as a single string, then parse
it as follows:

First of all, do you expect the text to be a single word, or to have spaces
in it?

If the text is a single word, then you can use the STRTRIM and STRCOMPRESS
functions in conjunction with STR_SEP.
STRTRIM cuts off leading and trailing blanks from a string
STRCOMPRESS can either compress multiple whitespaces to a single space or
remove them entirely.
Then you could use STR_SEP as usual. Or you could use READS if you know the
format and number of characters in each variable.

But if you are using a recent version of IDL (5.5, though it may be present
in earlier versions) you should use the STRSPLIT routine instead of STR_SEP.
It reproduces the functionality of STR_SEP, which is now obsolete, but lets
you use regular expressions. Those regular expressions might make all this
even easier.

If your text could have multiple words in it, this might all work if you
know exactly what format the other variables have (ie if they are all
four-digit integers or what have you). Then you could still do the
strcompress and use some strmid routines to extract the string, knowing the
start and end position of it from your knowledge of the other column widths.

If you can't get this to work, why don't you explicitly post a few lines of
your data file and tell us exactly what you can count on in the file in
terms of known formats etc.

Doug Rowland
School of Physics and Astronomy
University of Minnesota
rowland@fields.space.umn.edu

On 05/29/02 8:39 PM, in article 3cf5821f$1@nntp.kla-tencor.com, "Eric Vella"
<eric.vella.remove@kla-tencor.com> wrote:

> I need to read data from a file which contain lines with mixed variables
> including strings, for example "1 2 text 3". I tried "READF, int1, int2,
> str1, int3", but str1 gobbles up the rest of the line, leaving nothing for

> int3. This happens even if I supply an explicit format like A4 for the
> string. I can read the entire line into a single string, but then how do I
> separate the variables? STR_SEP does not work, since the separation
> character has to be a blank, and there can be multiple blanks between
> variables. I am hoping there is something simple I am missing (as is
> usually the case with IDL, simple only after you know the answer) ....
>
>

Subject: Re: READF data containing numbers and strings
Posted by Doug Rowland on Thu, 30 May 2002 15:24:36 GMT
View Forum Message <> Reply to Message

You might also try using ASCII_TEMPLATE and READ_ASCII to see if you can
interactively develop a template to read your data files.

On 05/30/02 9:51 AM, in article B91BA689.2C20%rowland@fields.space.umn.edu,
"Doug Rowland" <rowland@fields.space.umn.edu> wrote:

> I would start with reading the entire line in as a single string, then parse
> it as follows:
>
> First of all, do you expect the text to be a single word, or to have spaces
> in it?
>
> If the text is a single word, then you can use the STRTRIM and STRCOMPRESS
> functions in conjunction with STR_SEP.
> STRTRIM cuts off leading and trailing blanks from a string
> STRCOMPRESS can either compress multiple whitespaces to a single space or
> remove them entirely.
> Then you could use STR_SEP as usual. Or you could use READS if you know the
> format and number of characters in each variable.
>
> But if you are using a recent version of IDL (5.5, though it may be present
> in earlier versions) you should use the STRSPLIT routine instead of STR_SEP.
> It reproduces the functionality of STR_SEP, which is now obsolete, but lets
> you use regular expressions. Those regular expressions might make all this
> even easier.
>
> If your text could have multiple words in it, this might all work if you
> know exactly what format the other variables have (ie if they are all
> four-digit integers or what have you). Then you could still do the
> strcompress and use some strmid routines to extract the string, knowing the
> start and end position of it from your knowledge of the other column widths.

>
> If you can't get this to work, why don't you explicitly post a few lines of
> your data file and tell us exactly what you can count on in the file in
> terms of known formats etc.
>
> Doug Rowland
> School of Physics and Astronomy
> University of Minnesota
> rowland@fields.space.umn.edu
>
>
>
>
> On 05/29/02 8:39 PM, in article 3cf5821f$1@nntp.kla-tencor.com, "Eric Vella"
> <eric.vella.remove@kla-tencor.com> wrote:
>
>> I need to read data from a file which contain lines with mixed variables
>> including strings, for example "1 2 text 3". I tried "READF, int1, int2,
>> str1, int3", but str1 gobbles up the rest of the line, leaving nothing for
>> int3. This happens even if I supply an explicit format like A4 for the
>> string. I can read the entire line into a single string, but then how do I
>> separate the variables? STR_SEP does not work, since the separation
>> character has to be a blank, and there can be multiple blanks between
>> variables. I am hoping there is something simple I am missing (as is
>> usually the case with IDL, simple only after you know the answer) ....
>>
>>
>