Subject: Re: byte/unicode mismatch
Posted by Heinz Stege on Thu, 20 Nov 2008 19:08:17 GMT
View Forum Message <> Reply to Message

On Thu, 20 Nov 2008 09:23:52 -0800 (PST), mgalloy@gmail.com wrote:

> On Nov 20, 3:19ï¿½am, Reimar Bauer <R.Ba...@fz-juelich.de> wrote:
>> Hi
>>
>> the ascii table is gone.
>>
>> IDL> print,byte('ï¿½')
>> ï¿½195 188
>>
>> A char has now two bytes
>>
>> IDL> help, byte('ï¿½')
>> <Expression> ï¿½ ï¿½BYTE ï¿½ ï¿½ ï¿½= Array[2]
>>
>> This means all of the fast string replacing routines which are related
>> to iso encoded ascii one byte characters are broken in 7.0
>>
>> What is the name of the function to convert byte('ï¿½') into 252b ?
>
> I guess it is how you type/enter the ï¿½:
>
> IDL> u = string(252B)
> IDL> print, u
> ï¿½
> IDL> help, u
> U   STRING   = 'ï¿½'
> IDL> print, byte(u)
> 252
>
> Mike

Probably most readers here don't have an ï¿½-key on their keybord. So
here is another example:

IDL> print,!version
{ x86 Win32 Windows Microsoft Windows 7.0 Oct 25 2007   32   64}
IDL> mu='ï¿½'   ; (the greek letter)
IDL> help,mu
MU   STRING   = 'ï¿½'
IDL> help,byte(mu)
<Expression>   BYTE   = Array[2]
IDL> print,byte(mu)
 194 181

The string entered in the workbench command line is encoded in UTF8.
Using this string as a title in direct graphics results in a mu
preceded by an "A" with a hat. Direct graphics don't like UTF8. It
would need string(181b) for a mu.

If I don't miss something, Reimar is asking for a function to convert
the UTF8 string to ISO8859(?).

Heinz