
Subject: Re: byte/unicode mismatch

Posted by [R.Bauer](#) on Tue, 25 Nov 2008 13:03:38 GMT

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

me has forwarded a feature request to creaso for an en/de- coding parameter for open and had 5 minutes ago a phonecall about that. Lets see.

Reimar

Allan Whiteford schrieb:

> Reimar Bauer wrote:

>> Allan Whiteford schrieb:

>>> Reimar Bauer wrote:

>>>> That is all orthogonal.

>>>>

>>>> How can I decode and how can I encode?

>>>>

>>>> cheers

>>>> Reimar

>>>>

>>> Reimar,

>>>

>>> The question (and answer) isn't all that straightforward, byte values

>>> over 127 aren't well defined without an encoding system or a codepage.

>>>

>>> However, the answer you're probably looking for is:

>>>

>>> b=byte('iġ½') ; assumption 2

>>> print,b[1]+(b[0] eq 195)*64 ; assumption 1

>>>

>>> which is assuming:

>>>

>>> 1) you want byte values from (two byte) UTF-8 to ISO-8859-1

>>>

>>> and

>>>

>>> 2) that the u-umlaut character has entered the interpreter from a UTF-8

>>> environment.

>>>

>>> Please don't just cut and paste the above assuming all will be well.

>>>

>>> Thanks,

>>>

>>> Allan

>>>

>>

