
Subject: Re: Choosing the best way to go full UNICODE

Posted by [mirek](#) on Tue, 30 May 2017 09:03:01 GMT

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

cbpporter wrote on Tue, 30 May 2017 10:31

Get return the code unit.

What is 'code unit' in your terminology?

If I go by definition here: https://en.wikipedia.org/wiki/Character_encoding, it does not make sense with your explanation.

Now to explain what I am trying to do now is to provide some means to add full support to U++ editors in future. To that end, I need to define "cursor position" within the UNICODE text.

So far I believe that a "index of UNICODE codepoint" is a good cursor position. Alternatively, grapheme could be a good cursor position too, but then sometimes you would like to edit grapheme internals, like remove some combining mark, which could be difficult if grapheme is indivisible.

Anyway, it is definitely something to think about.

Mirek
