Subject: Re: 16 bits wchar
Posted by mirek on Sun, 21 Oct 2007 21:57:16 GMT

View Forum Message <> Reply to Message

cbpporter wrote on Sun, 21 October 2007 17:46luzr wrote on Sun, 21 October 2007 20:19
Well, I now might sound stupid, but I got a little bit lost in regard what problem we are really trying to solve.

In fact, I have already asked in some of previous posts...

My suggestion back then was that perhaps, if we are about rigid Unicode processing, we should not error-escape at all.

Well, what might help me: Do you have any real world scenario that can be solved using your routines? Maybe considering it will tell us something about what we are trying to do.

Mirek

Well the my routines are meant to be used this way:

```
// obtain a possibly invalid Utf-8 in s
if (!CheckUtf8(s))
    s = ToUtf8EE(s);
// pass s to other methods handling only valid Utf-8
```

The routines are done and tested, I'll post them on Monday (I don't have them on my home computer, which brings up the problem of forum submitting. Can I zip you my whole file or something?). I'm not sure if this is what you wanted to know.

Ah, I see.

Anyway, what are "other methods" supposed to do?

(I just want to see the bigger picture - IME, the only reasonable way of working with codepoints is to convert it to WString...).

Mirek

P.S.: Consider other aspect too - I have to be a little bit hesitant when adding things to Core - everything in chrset.cpp will bloat the Linux binaries...