
Subject: Re: Choosing the best way to go full UNICODE
Posted by [cbpporter](#) on Mon, 19 Jun 2017 08:51:35 GMT
[View Forum Message](#) <> [Reply to Message](#)

Yeah, that's why those decompositions are special and should not be used in day to day use.

Canonical decomposition in Unicode is not something to convert to when needed to solve some task. It can be used as the sole encoding format for all your string because it doesn't change the meaning of the string. Example: LoadFile("c:\\utf8.txt"). You could return the string as is, or you could return one of the two canonical Unicode forms, with the original string never stored.

On the other hand, , <box> and other non standard decompositions change the meaning of the text. They are computed when needed and you can't store your string as such 100% of the time.
