
Subject: Re: Choosing the best way to go full UNICODE
Posted by [cbpporter](#) on Tue, 06 Jun 2017 07:28:38 GMT
[View Forum Message](#) <> [Reply to Message](#)

You mean Unicode General Category?

I have that encoded in my two table scheme.

If that is all you can want to include, it should be extremely small, something like (number of characters / chunk size) * 1.2.

If you mean scripts, that is a bit more complex, mostly because I want the data small.

I'm also working on Unicode names. It turns out that there is under 200 unique words in the names. I want to use a huge string buffer where each byte in the string is a word, not a character.
