
Subject: Re: Choosing the best way to go full UNICODE
Posted by [cbpporter](#) on Tue, 06 Jun 2017 09:18:44 GMT
[View Forum Message](#) <> [Reply to Message](#)

There is no easy solution here I'm afraid.

You probably know the file:
<http://www.unicode.org/Public/UCD/latest/ucd/UnicodeData.txt>

What I do is read that file, compile all the information in RAM and write out C++ tables.

The file has a lot but not all of the needed information. The question is how much of it you need and how are you going to store it.

I don't understand the final point you were making about ambiguities? Characters are uniquely defined, so are the canonical composition and decomposition rules, together with compatibility substitutions.
