
Subject: Re: Will UPP support full UNICODE (21bits long codepoint)?

Posted by [Oblivion](#) on Sat, 15 Aug 2020 13:11:14 GMT

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

You are correct.

Quote:But hangul graphemes is what worries me... Not that I have searched too much, but so far I have failed to find simple algo how to combine hangul characters...

I'm no expert on Eastern-Asian languages either, but I've been studying unicode database for some time and related existing codes that deal with UC The one source code that can give an idea is in perl):

Maybe this could help (as a starter)?: <https://metacpan.org/pod/Lingua::KO::Hangul::Util>

Ps. The code does not support Hangul letters (jamo) assigned after Unicode 5.2, but it can serve as a starting point. (See also the references below the page).

Also: Hangul syllable types database for composition/decomposition:

<http://www.unicode.org/Public/UNIDATA/HangulSyllableType.txt>

Best regards,
Oblivion